# Two genes, *dig-1* and *mig-10*, involved in nervous system development in *C.elegans*

By

## Christopher T. Burket

A Dissertation

submitted to the Faculty of

## WORCESTER POLYTECHNIC INSTITUTE

In partial fulfillment

of the requirements for the

**Degree of Doctor of Philosophy**

**in Biomedical Sciences**

**Christopher T. Burket**

**Approved:**

**Dr. Elizabeth F. Ryder**                    **Dr. Samuel M. Politz**

**Dr. Joseph Bagshaw**                    **Dr. Stephen Lambert**

# Abstract

We are using genetic and molecular techniques to study a simple model organism, *C. elegans*, to determine the cues involved in the formation of the nervous system. Two molecules currently being studied in the laboratory play roles in the formation of the IL2 neurons, a class of sensory neurons in *C. elegans*. The first gene, *dig-1*, influences the sensory process or dendrite and is involved in adhesion as well as potentially providing directional information during development. The second gene, *mig-10*, influences the axon and may be involved in a cell signal cascade.

Genetic screens of *C. elegans* using Ethyl methyl sulfonate (EMS) as a mutagen resulted in the isolation of mutants with defects in the IL2 sensory map; sensory processes followed aberrant paths, appearing to be defasciculated. Complementation tests showed that the mutations failed to complement *n1321,* a known allele of *dig-1*; thus, these new mutations were alleles of *dig-1* (Ryder unpub. results). Several of these new alleles of *dig-1*, including *nu336* and *n1480,* have been further studied to elucidate the role of this gene in sensory map formation.

A *dig-1* candidate gene was identified that encodes a protein that is a member of the immunoglobulin super-family (IgSF). The candidate gene is predicted to be a large gene, with a transcript of approximately 45Kb. The encoded protein contains three distinct regions and is similar to the hyalectan family of proteoglycans. N terminal region 1 contains immunoglobulin and fibronectin-like domains. Central region 2 is an

area that is highly repeated with a potential to have GAGs attached. C-terminal region 3 contains domains associated with adhesion.

Polymerase chain reaction (PCR) products from alleles *nu336* and *n1480* were amplified and sequenced from the candidate gene. The DNA lesion present in the candidate gene from both alleles fit the method for how that mutation was generated. The point mutation in allele *nu336* removes a potential glycosylation site. The large re-arrangement in allele *n1480* truncates the transcript, suggesting that the protein is also truncated. The sequencing results along with rescuing data (R. Proenca, personal communication) showed that the candidate gene for *dig-1* was the gene of interest. Each of the alleles was further studied to determine how severe that allele was by looking at the neuronal process aspect and the brood size as well as displacement of the gonad. In general, alleles with severe defects in the nervous system also had severe gonad displacement, suggesting the gene functions similarly in the two tissues.

To determine if the gene was expressed at the RNA level, reverse transcriptase polymerase chain reaction (RT-PCR) was used. Most of the RT-PCRs amplified a cDNA of the appropriate size that showed *dig-1* was expressed at the RNA level. RT-PCR further suggested that all three regions were in one transcript as well as confirming part of the predicted exon structure to be correct. In addition, northern analysis showed the presence of a large transcript in wildtype worms as well as a smaller truncated transcript from allele *n1480*. To investigate developmental differences mixed stage of RNA and embryonic RNA from wildtype animals were compared using gene specific primers. The

initial RT-PCR showed potential alternative splicing occurring at the 5' end of the gene during development.

To examine expression at the protein level, two recombinant proteins from *dig-1* were successfully made by cloning cDNA products from the 5' and 3' end of *dig-1*. The constructs were sequenced and shown to be in frame. The recombinant proteins (Ant1Con1 and Ant3Con3) were mass produced and sent to a commercial source for injection into pre-screened rabbits. Western analysis showed the presence of an antibody in the serum from two of the rabbits. These antibodies should prove useful in future determination of correctness of our models of DIG-1 function.

IgSF members have been shown to have many roles in nervous system development. DIG-1 could act in either an attractive or a repellent role to position sensory processes during development. DIG-1 might also change its function over time; early in development DIG-1 could be adhesive and later become repellant as more sugars are added.

The gene *mig-10* is involved in sensory map formation. To localize MIG-10 expression, several transgenic animals were generated by injection of two constructs that should recombine in the worm to create a MIG-10::GFP fusion protein. Ten transgenic lines were generated and screened by PCR for the presence of the correct recombinant construct. If this construct makes functional, rescuing protein, the GFP expression should reflect the expression pattern of the MIG-10 protein.

# Acknowledgements

-"If it is not one thing, it's about five others but we have them all covered"

I would like to thank my thesis advisor for all of the advice and laughs about *dig-1* and life in general. Without this advice, *dig-1* might still be some candidate gene that no one knows what to do with.

Thanks to my thesis committee for providing direction and helpful insights into science.

Special thanks to Sam, Jill and JoAnn, for their help with proteins, membranes, anitbodies and a bunch of advice in general.

A big thanks to all the individuals that provided plasmids, procedures and worm strains.

Thanks to the GAANN fellowship, Dr. Ryder's NSF grant and WPI for providing financial support.

Thanks to all of my friends here, who have listened to me complain about the same "stuff" over and over again but still listened anyway.

To Rick and Rachel Stock thanks for being there when I really needed someone to be there.

Thanks to my sister, brother-in-law and my godson Alec for reminding me there is a life outside of the lab

Very special thanks to my parents, Dean and Freya Burket, who have provided financial, mental, and emotional support through the last five years.

> -In dedication to my father: Who stayed as long as he could to see his little boy reach one of his biggest dreams in life, which seemed so far away. That dream came true, thanks dad.

> Mom what can I say, but there is not anything that I can say to express my appreciation for all the help and support, even in the last few disheartening years.

# Table of contents

# List of Tables

# List of Figures

# 1.0 Literature review

## 1.1 Introduction

The development of the nervous system leads to the formation of an organ system that allows information to be acquired from the environment and transferred to the brain. Information is transferred to the brain by intricate circuits made by neurons forming synapses with other neurons in an organized way. In mammals, this means that trillions of neuronal cells extend their neurites in a highly ordered manner to make connections with their appropriate targets during embryogenesis. During embryogenesis, many cues give positional information to the neurites and direct them to the correct synapses. The cues that are involved in the formation of the nervous system have diverse functions that range from adhesion to signal transduction. We are using genetic and molecular techniques to study a simple model organism, *C. elegans*, to determine the cues involved in the formation of the nervous system.

Two molecules currently being studied in the laboratory play roles in the formation of a class of sensory neurons in *C. elegans*. The first gene, *dig-1*, influences the sensory process or dendrite and is involved in adhesion as well as potentially providing directional information during development. The second gene, *mig-10*, influences the axon and may be involved in a cell signal cascade.

**1.2 Axon guidance**

The process of an axon finding its correct synapse during development is an essential milestone in the formation of the nervous system. An axon is a process that extends from a neuronal cell body and terminates at a synapse. Axons migrate in a highly stereospecific manner from the cell body and invade target areas. Axon guidance, the process by which an axon projects onto the correct target, has been studied extensively. Experimental evidence suggests that axon migrations occur through the presence of molecular cues that guide the axon to the correct target area in the developing organism. The process by which dendrites make connections has not been as extensively studied, but some of the cues are conserved between these two processes (Agarwala et al., 2001; Keith and Wilson, 2001; McAllister, 2002). An example of a gene that is involved in both axon and dendrite guidance is *roundabout* (Robo 1,2, and 3), which was shown to have roles in both dendritic guidance and synaptic connectivity (Godenschwege et al., 2002). Since axon guidance has been more extensively studied, the underlying mechanisms and molecules involved in this process will be reviewed here.

Axons extend to the correct target by four general mechanisms, which act either at short range through cell to cell contact, or at a long range, via diffusable gradients (Keynes and Cook, 1995; Tessier-Lavigne and Goodman, 1996). The mechanisms are either permissive ones, which stabilize axonal interactions, or informative ones, which determine the pathway choice (Van Vactor, 1998) (**Figure 1**). The mechanisms of axon pathfinding are executed by molecules (proteins) that are present on adjacent cells, in the

2

extracellular matrix, or as secreted molecules. The molecules are detected by the axon growth cone.
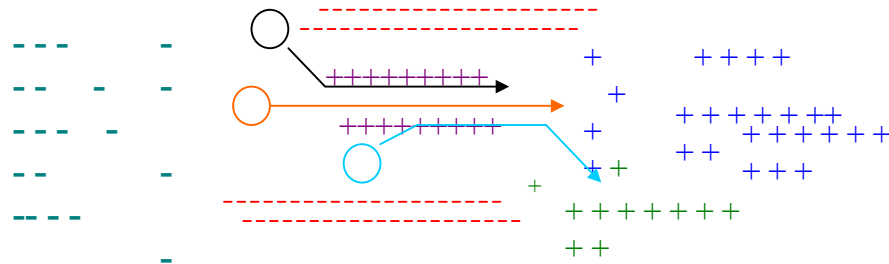


**Figure 1. Schematic depiction of the mechanisms of axon guidance.** The neuron in orange extends a pioneering axon away from the diffusible repellent (-). The contact repellent molecule (-) repels all axons. The other neurons fasciculate or adhere to the pioneer axon by contact attraction, ( + ) molecules. The light blue axon shows selective defasciculation from the pioneering axon and travels toward a diffusible attractant (+). The other two axons travel on to their targets following the second diffusible attractant (+). Figure adapted from Tessier-Lavigne and Goodman, 1996.

Growth cones are organelles that can expand and contract from the distal tips of dendrites and axons. The growth cone directs the dendrites and axons to the correct synaptic target (Lankford et al., 1990; Bentley and O'Connor, 1994; Tanaka and Sabry, 1995; Crino and Eberwine, 1996, **Figure 2**). Three distinct cytoplasmic domains are present in growth cones: the peripheral domain (lamellipodia and filopodia), the central domain (organelles), and the transition zone, which has actin based ruffling activity (Suter and Forscher, 2000). The lamellipodia and filopodia are used to probe the microenvironment by forming transient contacts with surrounding cells and the extracellular matrix (Gordon-Weeks, 1987; Kater and Rehder, 1995; Korey and Van Vactor, 2000; Skaper et al., 2001). This allows the receptors present on the growth cone

to interact with the guidance molecules that are present on surrounding cells or in the extracellular matrix as well as diffusible molecules.
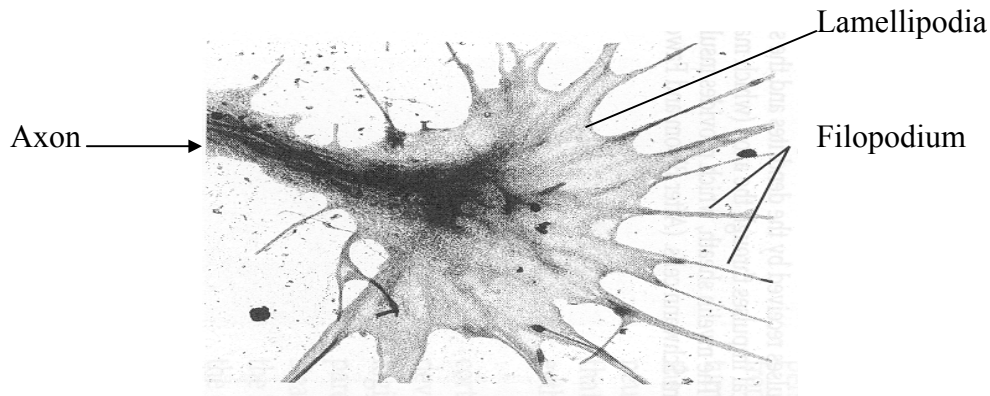


**Figure 2. Axon growth cone.** This is a picture of an axon growth cone, showing the lamellipodia and filopodia. Guidance and adhesion molecules would be found in the micro-environment around the growth cone. Picture adapted from Gilbert (1997).

The major structure that allows the growth cone to respond to guidance cues is the cytoskeleton (Suter and Forscher, 2000; Korey and Van Vactor, 2000; Skaper et al., 2001). The cytoskeleton has two primary components, F-actin and microtubules. F-actin is found in the peripheral domain as a meshwork in lamellipodia and as bundles extending into the filopodia. Microtubules project from the axon shaft into the central domain with their dynamic ends entering the peripheral domain (Bentley and O'Connor, 1994; Korey and Van Vactor, 2000). The actin structures of the growth cone are influenced by the cues present in the environment that cause them to collapse or expand. The collapse of these structures is associated with actin being degraded as quickly as it is polymerized. Zhou and Cohan (2002) have shown that actin bundle loss is associated with growth cone turning in response to a repulsive signal. The expansion of these

structures is associated with a slowing of the depolymerization of actin, which then allows for forward motion (Korey and Van Vactor, 2000).

Many different signaling pathways play a role in regulating actin polymerization, including tyrosine phosphatases and kinases, the Ena/VASP family of proteins, and the small GTPases of the Rho family (Korey and Van Vactor 2000; Dickson, 2001). TRIO, which is a guanine nucleotide exchange factor that activates Rho-GTPases, has been shown to be involved in neurite outgrowth in PC12 cells, suggesting that Rho-GTPases are necessary for regulation of the cytoskeleton of the growth cone (Estrach et al., 2002). Attractive cues may activate Rac, Cdc24 and Rho G, which promote actin stabilization, whereas repulsive cues activate Rho A, which may inhibit actin stabilization (Dickson 2001; Estrach et al., 2002). On the other hand, the GTPases may only be needed to polymerize actin but not be necessary for pathfinding; they may act permissively by making the actin backbone for the lamellipodia and filopodia (Dickson, 2001). Kim et al., (2002) have shown that the activities of Cdc24 on filopodia extension and axon pathfinding are regulated independently of each other. Signaling events in the growth cone cytoskeleton seem to be of two distinct types: changes that allow growth cone motility (turning), and those that regulate the polymerization of actin.

The growth cone does not advance continuously to the target location; the advancement (pathfinding) is broken into short periods of advancement to intermediate targets where the growth cone decides either to continue in the same direction or turn based on the molecules present. Guidepost cells that provide the growth cone with

information on where to go typically produce these intermediate targets. This decision making process can be seen in changes in growth cone morphology and changes in direction of the growth cone as it contacts or passes an intermediate target (Kaprielian et al., 2001). The growth cone moves from these intermediate choice points by changing how it responds to the molecules provided by the guidepost cells.

An example of this process occurs at the floor plate, ventral midline tissue in vertebrates. The commisural axons migrate to the midline in response to an attractant secreted by the floor plate, but upon reaching the floor plate, the axons lose the ability to respond to that chemo-attractant (Shirasaki et al., 1998). More recent studies of commissural axon migration to the midline further support the idea that neurons change their response to hierarchical cues present. Initially, an axon expressing the netrin receptor, deleted in colorectal cancer (DCC) is attracted by netrin to extend to the midline (Stein and Tessier-Lavigne, 2001). After crossing the midline, axons increase their expression of Robo, the receptor for the chemo-repellent molecule Slit, which is also expressed by the floor plate (Stein and Tessier-Lavigne, 2001; Giger and Kolodkin, 2001). When the growth cone is only expressing DCC, the axon is attracted to the midline; after the up-regulation of Robo expression, the receptors for Netrin (DCC) and Slit (Robo) interact to silence the attractive activity of netrin. Slit then repels the growth cone by interacting with Robo. Another example of change in responsiveness involves the second receptor for netrin in *C. elegans*, UNC-5. When UNC-5 and DCC form a complex, the binding of netrin results in a repulsive signal instead of the usual attractive signal (reviewed by Kaprielian et al., 2001). Thus, the interaction of the molecules

determines whether a particular cue is repulsive or attractive.  Since the context

determines the activity of guidance cues, assigning axon guidance molecule to a specific

class (attractive or repulsive) is impossible.

Most growth cones do not have to make new tracts; instead, the axon adheres to a

pioneering axon, forming a bundle or fascicle.  The process of fasciculation is highly

selective, with some growth cones displaying a specific affinity for one fascicle over

another (Goodman et al., 1984; Fambrough and Goodman, 1996).  The growth cone can

move from one fascicle to another, exposing itself to a different environment. The axon

can then selectively de-fasciculate to reach the target area (**Figure 2**). After reaching the

target area, the axon then locates its appropriate synapse using the local cues present. For

example, the secreted protein product of the *beaten path* gene controls selective

defasciculation at motor axon choice points in *Drosophila* (Fambrough and Goodman,

1996).  In *beaten path* mutants, axons fail to defasciculate and miss their muscle target

regions.

In summary, axon guidance is a dynamic process that is dependent on the

regulation of the cytoskeleton of the growth cone.  Changes in the cytoskeleton provide

force for the growth cone to find its way to the target.  The changes in the cytoskeleton

are controlled or influenced by the guidance cues present.

**1.3 Molecules involved in axon guidance**

The many families of axon guidance molecules act in concert to guide the axon to its target area by promoting adhesion, outgrowth, steering, and motility of the growth cone. The molecules range from membrane bound receptors to molecules in the ECM. Many of these molecules have multiple functions, such as promoting adhesion as well as outgrowth or motility, during axon guidance. The activities of several of these families of molecules will be briefly reviewed.

The immunoglobulin superfamily (IgSF) is one of the largest families of molecules, with several members that are involved in axon guidance. This family has a diverse range of molecules, all containing at least one immunogobulin (Ig) domain. The Ig domain is made up of between 70-110 amino acids with two cysteine residues that are 55-75 residues apart. A conserved tryptophan is 10-15 residues downstream from the first cysteine. IgSF molecules have been shown to mediate adhesion between cells, axon fasciculation, neurite pathfinding, and neurite outgrowth, by homophilic and heterophilic binding independent of calcium. There are many sub-families of the IgSF, represented by the receptor protein tyrosine kinases (RPTK), receptor protein tyrosine phosphatases (RPTP), netrin/slit-receptors (DCC and Robo), and immunogobulin cell adhesion molecules (IgCAMs, Brummendorf and Lemmon, 2001). These molecules have diverse functions, ranging from acting as a receptor for secreted molecules to adhesion during axon guidance (**Figure 3**). In most IgSF members, the Ig domain is followed by several fibronectin type III domains (FNIII), which are usually associated with adhesion. FNIII domains are 90 amino acids in length and related to domains found in fibronectin

(reviewed by Walsh and Doherty, 1997). The cytoplasmic domains present in Ig SF

members in general activate signal cascades.
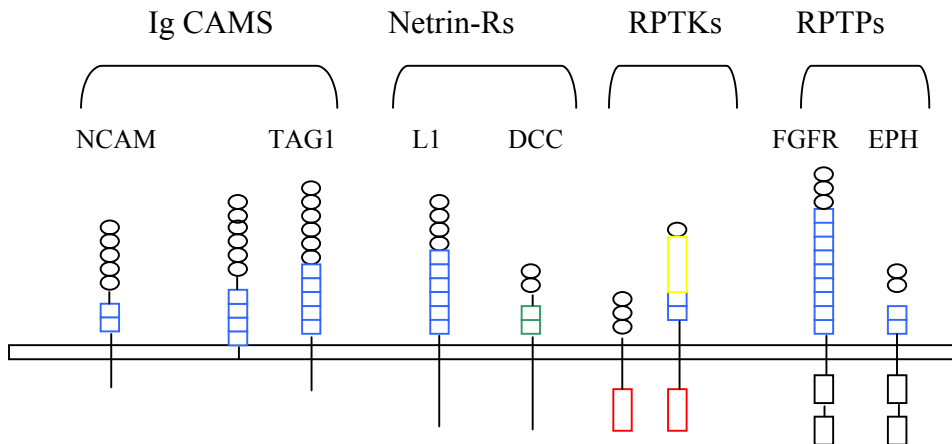


**Figure 3. Representation of some of the Ig-Superfamily molecules.** The immunoglobulin domains are represented as circles. The fibronectin domains are represented as blue squares. The green squares are thrombospondin type 1 domains. The red boxes are protein tyrosine kinase domains. The yellow box is a cysteine rich region. The black boxes are protein tyrosine phosphatase domains. Figure adapted from Tessier-Lavigne and Goodman, 1996.

A second family of molecules involved in axon guidance is the semaphorin

family, which has 20 members placed into eight groups based on their sequence and

structure (Goshima et al ., 2002). All semaphorins contain a conserved 500 amino acid

sequence, the sema domain, and are associated with repulsion of axons, but there are also

attractive molecules such as Sema3C (Goshima et al., 2002). Molecules of this family

are mainly membrane bound or cell-surface molecules, but there are several members

that are diffusible. Several of the diffusible members of this family have one Ig domain.

Neuropilins and the plexins, two known semaphorin receptors, are from two distinct

transmembrane receptor families (Goshima et al., 2000). In some cases, members of

these two families of receptors form a complex that has increased binding activity for semaphorins. The plexin family of receptors, whose members also contain a sema domain, has been implicated in regulating Rho GTPases, which mediate the growth cone collapsing activity of the semaphorins (Goshima et al., 2000; Patel and Van Vactor, 2002).

A third family is the ephrins, which are cell surface molecules. This family contains two subfamilies, ephrins A and B. Ephrin A family members are cell surface molecules that are anchored to the cell by glycosylphosphatidyinositol (GPI) and bind EphA receptors. Ephrin B family members, on the other hand, are transmembrane proteins that bind to EphB and EphA4 receptors. These ligand/receptor pairs signal through the Rho family of GTPases, as do the semaphorins; both families usually mediate repulsive signals. A proposed molecule that links the ephrin/Eph complex to the Rho GTPase pathway is ephexin (Knoll and Drescher 2002; Patel and Van Vactor, 2002).

A fourth family of molecules is the netrins. This family of molecules contains several members, e.g., netrin 1-4 in mice, which are typically small diffusable molecules that are about 600 amino acids in length. One netrin family member has a GPI anchor. The netrin ortholog in *C. elegans* is the gene product of *unc-6* that shares about 50% homology with other family members (Keynes and Cook, 1995; Kaprielian et al., 2001). The N-terminus of netrin is related to the laminin-b subunit (Keynes and Cook, 1995). The ECM can influence diffusion of these molecules. Netrins have been shown to be necessary for attracting commissural axons to the midline but have also been shown to be

chemo-repellents for the trochlear motor axons (Kennedy et al., 1994; Colamarino and

Tessier-Lavigne, 1995; Serafini et al., 1996).  The primary netrin receptor is DCC, which

can complex with other molecules to change the activity of the molecule in axon

guidance (Livesy, 1999).

ECM molecules are also involved in axon guidance.  These include molecules

that guide axons in the ECM such as laminin, tenascin, and the proteoglycans (PGs).  Of

the diverse groups of molecules, the group that is of the most relevance to this work is the

PGs.

PGs are made up of a core protein that has glycosaminoglycans (GAGs) attached

to it by either N- or O- linkages.  GAGs are long polysaccharide chains made from

repeating disaccharide units with a highly negative charge. The disaccharide units have at

least one amino sugar, either *N*-acetylglucosamine or *N*-acetylgalactosamine.  PGs are

classified based partly on the molecular make up of the GAGs attached to the core

protein.  The different groups of GAGs are heparan/heparin sulfate (HS), chondroitin

sulfate (CS), dermatan sulfate and keratan sulfate.  There are at least three families of

core proteins involved in nervous system development. These families are the syndecans,

the glypicans, and the hyalectans.  There are four syndecan members that are

transmembrane proteins with a majority of HS GAGs.  The glypicans have 6 members

that are all GPI anchored with a majority of HS GAGs. The hyalectan family of core

proteins has 8 members, one being GPI anchored, that have mainly CS side chains that

interact with hyaluronate and lectin (Schwartz, 2000). PGs with CS side chains are the

most abundant PG in the organisms nervous system and are often found in growing axon tracts (reviewed Bovolenta and Fernaud-Espinosa, 2000).

PGs are involved in stimulating or inhibiting axon and dendrite outgrowth, mediating cell connections, and mediating presentation of signals to cells (Bovolenta and Fernaud-Espinosa, 2000; Bandtlow and Zimmermann, 2000; Huaiya, 2001). The structural diversity of proteoglycans is contingent on differential expression of the genes encoding the core protein, alternative splicing, transcription termination, variations of the GAG side chains and proteolytic cleavage of the core protein (Bandtlow and Zimmermann, 2000). These molecules bind several types of ligands such as growth factors, cell adhesion molecules, matrix components, enzymes and enzyme inhibitors. PGs can affect the activity of molecules involved in guidance; for example, heparan sulfate has been shown to enhance the repellent activity of SLIT2 by modulating its binding to ROBO (Huaiyu, 2001). PGs bind growth factors such as FGF-2 and in doing so potentiate their activity (Milev et al., 1998). PGs can also form networks of molecules. For example, aggrecan and versican can bind fibulin-2, which can then cross link hyaluronan-lectican complexes (Olin et al., 2001). Mutations of genes encoding PGs can affect their activity by changing the core protein structure, whereas mutations in the enzymes necessary for addition and extension of the sugar side chains can affect these modifications of the core (Lander and Selleck, 2000).

Finally, the integrins are molecules that span the cell membrane and allow the cell cytoskeleton to be linked to the ECM. On the extracellular side, the integrins bind the

amino acid sequence RGD of adhesive molecules in the ECM. Integrins are coupled to the cytoskeleton by talin and α-actinin, two proteins that bind actin. Integrins are made up of two subunits, α and β. The β1 subunit is the main form of this subunit in neuron development. Integrins have been proposed to mediate axon growth and modulate adhesive signals (McKerraracher et al., 1996; Stevens and Jacobs, 2002). The data from *C. elegans* and *Drosophilia* do not support integrins mediating growth but do support the idea of modulating adhesion signals, since in many mutants the axons are defasciculated or miss the synaptic target (Stevens and Jacobs, 2002). Stevens and Jacobs (2002) propose that the strength of adhesion signaling influences the threshold for response to the chemo-repellent Slit.

**1.3.1  Molecules involved in adhesion during axon guidance**

One of the genes presented later in this work appears to play a role in adhesion. Why is adhesion important during development? In mature organisms adhesion molecules maintain cell-cell connections such as synapses. The temporal and spatial regulation of adhesion molecules is required for proper cell migration, axon guidance and synapse formation during nervous system development (Lee and Benveniste, 1999). Adhesion molecules allow cells to adhere to one another and form aggregates for organ formation and stable structures for cells to migrate across. Recently, adhesion molecules have also been implicated in cell signal cascades.

There are several classes of molecules involved in adhesion. The first class, which are a subclass of the IgSF, are the immunogobulin cell adhesion molecules

(IgCAMs) that act to bind cells to one another. The cadherins are a second class of molecules, which for the most part are involved in early development and are dependent on calcium for their activity (Ranscht, 2000). A third class are cell junction molecules such as connexin, which provide channels between cells. A fourth class of adhesion molecules, are substrate adhesion molecules (SAMs) that bind cells to molecules in the extra cellular matrix. The IgCAMs and proteoglycans are the most pertinent to the current work, and the following review focuses on these molecules.

**1.3.2 IgCAMs**

IgCAM members are grouped based on their domain organization, amino acid sequence similarity and type of membrane anchorage (Brummendorf and Rathjen, 1996; Suter and Forscher, 2000). The classical mode of binding for IgCAMs is considered to be by homophilic or heterophilic binding between two cells (in trans). More recently it has been shown that they can also bind in cis to molecules on the same cell (Kamiguchi and Lemmon, 2000; Rutishauser, 2000). IgCAMs also play a role in maintaining the connections of the nervous system after it is formed; an example of a gene family that mediates this function is the *zig* gene family in *C. elegans* (Aurelio et al., 2002).

IgCAMs' adhesiveness is regulated by four mechanisms. The first mechanism is grouping the molecules or lateral oligomerization. An example of this is one molecule of L1, an IgCAM, that binds to one ankyrin molecule, which is next to a second ankyrin molecule that is bound to a second L1 molecule. Both ankyrin molecules are bound to the same substrate (spectrin) that is attached to the cytoskeleton. This would then

effectively "attach" two molecules, by indirect attachment, which are involved in adhesion, thus increasing the adhesiveness. The second mechanism is regulating the number of molecules involved in adhesion by removing the molecules from the membrane. For axons, this means that the leading edge of the growth cone would have a higher percentage of the adhesion molecules while the trailing end would have fewer. The third mechanism is proteolytic cleavage of the adhesion molecules. L1 is again a good example in that it can be cleaved by plasmin and disintegrin-metalloproteinase (ADAM) family proteases at two different sites, which removes the Ig domains (reviewed by Kamiguchi and Lemmon, 2000). The fourth mechanism is spatial and temporal by transcriptional regulation (mechanisms reviewed, Kamiguchi and Lemmon, 2000; Brummendorf and Lemmon, 2001). One example of transcriptional regulation is alternative splicing. An extreme example of structural diversity generated by alternative splicing is the *Drosophila* DSCAM gene, which has a potential for 38,000 splice isoforms (Schmucker et al., 2000; Brummendorf and Lemmon, 2001).

One example of a well-studied IgCAM is neural cell adhesion molecule (N-CAM). N-CAM has three isoforms as well as different levels of adhesion and outgrowth stimulation based on the presence of an alternative 10 amino acid sequence encoded by the VASE exon and the presence or absence of polysialic acid (PSA, Cunningham et al., 1987; reviewed by Walsh and Doherty, 1996; Skaper et al., 2001). The presence of the VASE domain, which is added into the fourth Ig domain, reduces the growth properties of N-CAM (reviewed by Walsh and Doherty, 1996). Early in development, N-CAM has PSA chains present that promote plasticity; later in development PSA is removed and N-

CAM becomes more adhesive (reviewed by Walsh and Doherty, 1996). The modifications change the function of N-CAM at different times in development, allowing the molecule in one instance to promote plasticity and in another, adhesion (reviewed by Skaper et al., 2001).

**1.3.3 Hyalectan family of PGs as mediators of axon guidance**

The hyalectan family of PGs has a unique structure that is similar to the structure of the predicted DIG-1 protein presented later in this work. In hyalectans, a single Ig domain at the N-terminus is followed by a hyaluronan binding tandem repeat. The middle of the protein is poorly conserved and is the area for GAG attachment. The C-terminus has epidermal growth factor like repeats as well as a C-type lectin domain. Most GAGs associated with hyalectan family members are CS. GAGs are proposed to keep the molecule in an extended conformation.

Neurocan, a member of the hyalectan family, binds to both NCAM and L1, which are both IgCAM members (Retzler et al., 1996; Oleszewski et al., 1999). Neurocan also regulates the function of cadherins and integrins (Li et al., 2000). The proposed function of neurocan is to inhibit neurite outgrowth and modulate the adhesion of the neurons to the ECM. How might neurocan modulate adhesion during development? A potential model is that neurocan prevents inappropriate adhesion of a particular neuron to a certain tract. The inhibitory function prevents neurons from leaving that tract or adhering inappropriately. In support of this idea, aggrecan, a second member of the hyalectan family, which is also inhibitory, has been shown to up-regulate integrin expression in

16

embryonic neurons (Condic et al., 1999).  This suggests that it prevents the inappropriate growth but stimulates expression of integrins, which adhere the neuron to the correct tract.  In contrast to the inhibition, in cat development neurocan seems to be necessary for distinguishing between efferent and afferent pathways (reviewed Bandtlow and Zimmermann, 2000).  Again the different environments and forms of the molecules can affect how the molecule influences the axon.

**1.3.4 Other activities of molecules that mediate adhesion**

It should also be mentioned that molecules that were classically thought to be involved only in adhesion could also be involved in neurite outgrowth. The neurite outgrowth seen with CAMs (N-CAM, L1 and N-Cadherin) is dependent on the tyrosine kinase activity of the fibroblast growth factor receptor (FGFR) (reviewed Crossin and Krushel, 2000).  NCAM (NCAM140) has been shown to signal through lipid rafts that are required for signaling by the FGFR in cultured neurons (Mizoguchi et al., 2002). The signaling process is not well understood, but it is thought that it involves the influx of calcium to a localized area of the growth cone, via the activation of the FGFR (Doherty and Walsh, 1996; Doherty et al., 2000).  The influx of calcium has been shown to be sufficient for the neurite outgrowth in response to stimulation by N-CAM, N-cadherin and L1 (reviewed Skaper et al., 2001).  NCAM has also been shown to stimulate the Ras-MAPK pathway as well as CREB phosphorylation in neuronal cells (Schmid et al., 1999).

A second activity mediated by CAMs is motility (Prag et al., 2001). One proposed model is substrate-cytoskeletal coupling. The principle of this model is that a growth cone can move forward if it is capable of coupling intracellular actomyosin movement to a fixed substrate in the ECM via cell surface receptors (Suter and Forscher, 2000). The receptors link the ECM to the actin cytoskeleton, allowing the actinomyosin contractions to pull the growth cone forward.

In summary, the many families of guidance molecules have diverse functions that are necessary for guiding growth cones to their target areas. Many of these molecules can be placed into several different families based on domain structure. A simple model system to study the roles of these molecules is represented by the nervous system of *C. elegans*.

## 1.4 *C. elegans* as a model system to study nervous system development

### 1.4.1 General aspects of *C. elegans*

In 1974, Sydney Brenner identified the free-living soil nematode, *C. elegans,* as an ideal organism for studying development. *C. elegans* is an advantageous model organism for studying development for several reasons, including a rapid life cycle, simple nervous system, sequenced genome, and ease of manipulation in the laboratory (Riddle et al., 1997). Also, the mechanisms of axonal guidance, as well as DNA sequences of the molecules involved, have been evolutionarily conserved from nematode to higher vertebrate (Tessier-Lavigne, 1994). An example of conservation is that there

are at least one SLIT protein (SLT-1) and one robo receptor (SAX-3) in *C. elegans* that

mediate midline guidance (Hao et al., 2001).

The generation time of *C. elegans* is about 3 days at 20°C; development proceeds

through embryogenesis (14hrs), and four larval stages (L1-L4), finally producing a

sexually mature adult.  Adult worms are either hermaphrodites or males (males constitute

0.1% of the population).  The lineage of all 959 somatic cells has been determined and is

invariant (Riddle et al., 1997).  The *C. elegans* nervous system, which starts to form

connections at about 400 minutes into embryogenesis, has also been well defined, and

contains 302 neurons that make about 10,000 connections (Riddle et al., 1997).

**1.4.2  A model circuit map in *C. elegans* is comparable to lower vertebrates**

In vertebrates, sensory systems, which are part of the nervous system, are

topographically mapped.  These systems allow information to be mapped from the

environment to the brain in a coordinate manner (Kaplan, 1996).   A widely studied

topographic sensory map is the visual field map of lower vertebrates (**Figure 4**).  This

figure illustrates the definition of a topographic map, in that neurons that are close

together and have neighboring sensory fields are projected onto neurons that are close

together in a target field in the brain.  In *C. elegans*, several classes of head sensory

neurons such as IL1 and IL2 neurons are topographically mapped (White et al., 1986;

Kaplan, 1996).  One difference between the sensory map in vertebrates and that of *C.

elegans* is a specialized receptor cell, which receives information from the environment

in vertebrates.  The acquired information is then relayed through several neurons before it

reaches the brain. In *C. elegans,* on the other hand, the receptor cell is also the same cell

that synapses on the nerve ring or brain in the worm (**Figure 4**).    Thus, in the simple

topographic map in *C. elegans,* the function of the map is dependent on both the sensory

process and the axon finding the correct area during development.

Panel A



Panel B

Cell body

Brain

Sensory processes          Axons
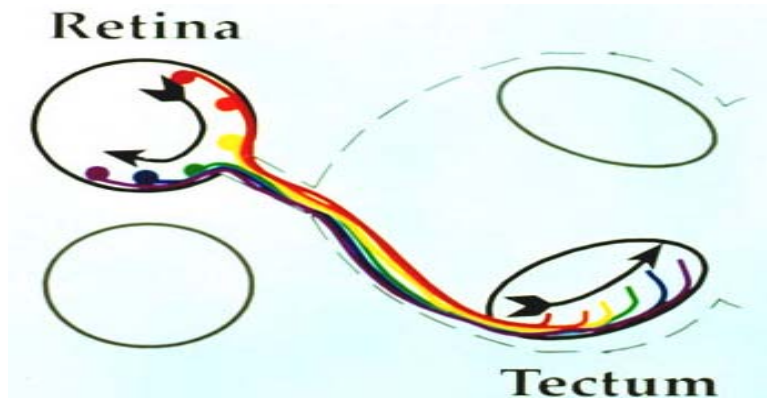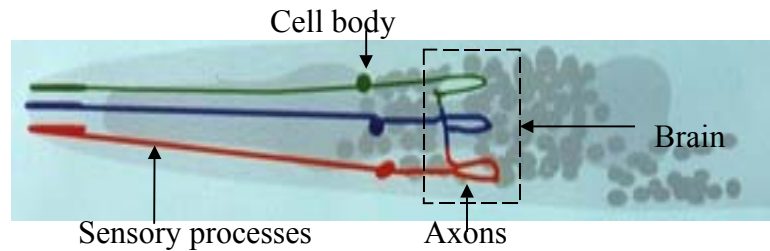
**Figure 4.  Comparison of sensory maps between lower vertebrates and *C. elegans*.** Panel A, Schematic representation of the visual field map of lower vertebrates. Dotted line is the boundary of the brain.  Neurons projecting their axons from the retina (large circle) to the tectum (large ovals) are shown in red, orange, yellow, green, blue and purple.  The arrow represents information from the environment that is coordinately mapped to the brain. Panel B, IL2 sensory map in *C. elegans* showing the sensory process, cell body and axons.  The nerve ring is the "brain" of the worm, outlined in a dotted box. Dark gray structure is the pharynx; dark ovals are other cell bodies of neurons in the worm's head .

In both cases of topographic map formation, vertebrate and worm, the formation

of the sensory map is dependent on the neurons making the correct connections.  As

previously stated, neurons make the correct connections via the molecular cues present

during development. We are interested in finding molecules that are involved in the

formation and maintenance of the nervous system.  Such molecules might include those

involved in outgrowth, adhesion, and motility such as IgCAMs and the proteoglycans, as

well as those molecules that are involved specifically in topographic sensory map

formation, similar to the ephrins in vertebrates.


## 1.5 Genes involved in the formation of the IL2 sensory map

### 1.5.1   Identification of *dig-1*, a *C. elegans* gene encoding a putative adhesion molecule

Genetic screens of *C. elegans* using Ethyl methyl sulfonate (EMS) as a mutagen

and DiO staining of the IL2 neurons resulted in the isolation of mutants that were

aberrant in the IL2 sensory map (**Figure 5**).  Electron micrograph (EM) data showed that

many sensory ending were missing or misplaced.  In addition, the cuticle was detached

from epithelial cells (Kaplan, and Horvitz, unpub. results).  These mutations were then

mapped on the genetic map to a specific location on chromosome III (Ryder, map

results).  Complementation tests showed that the mutations failed to complement *n1321,*

a known allele of *dig-1*; thus, these new mutations were alleles of *dig-1* (Ryder unpub.

results). Thomas et al. (1990) originally isolated *dig-1* due to a **di**splaced **g**onad

phenotype, and *dig-1* was mapped close to the cloned gene *sma-3* on chromosome III (M.

Basson personal communication).  The many different aspects of the phenotype seen in

the mutants suggested that these aspects were all caused by loss of adhesion (Kaplan and

Horvitz, unpub results).  Several of these new alleles of *dig-1*, including *nu336* and

*n1480,* have been further studied in the present work to elucidate the role of this gene in sensory map formation.

Panel A

Panel B
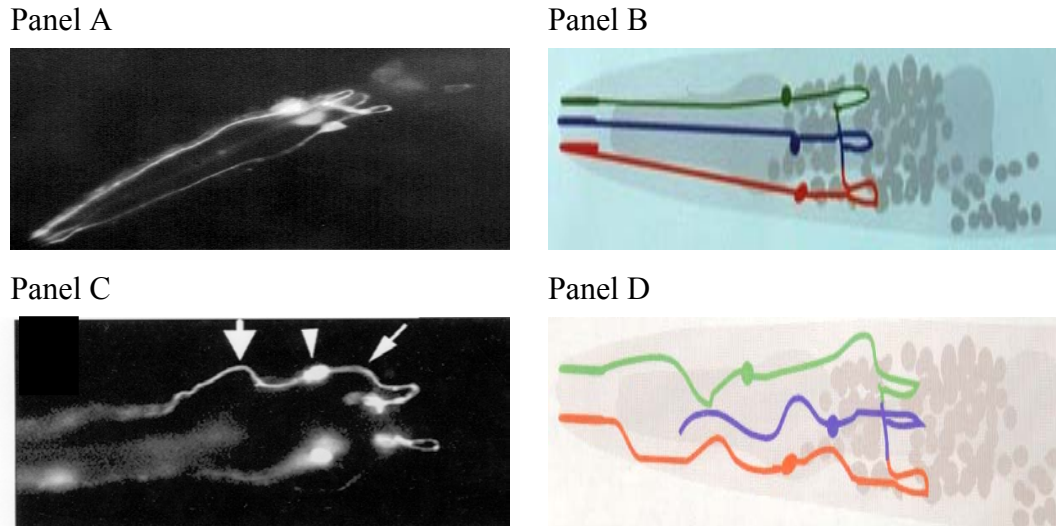
Panel C

Panel D



**Figure 5. Wild type and mutant IL2 sensory map.** Panel A, DiO stain of wildtype worms showing the normal phenotype of the IL2 neurons. Panel B, Schematic of the wild type IL2 neurons. Panel C DiO stain of mutant worms showing the aberrant IL2 sensory phenotype. Large white arrow shows the sensory process. Arrowhead shows the cell body. The thin arrow shows the axon. Panel D, Schematic of the IL2 mutant phenotype.

**1.5.2 *mig-10,* a second gene involved in formation of the IL2 sensory map**

Mutations in *mig-10* cause incomplete long-range migration of canal-associated neurons (CANs), hermaphrodite-specific neurons (HSNs), and anterior lateral microtubule cells. The defects seen in migration may be due to either the migration machinery being affected or the cellular development prior to migration being affected (Manser et al., 1997). Mutant *mig-10* animals cause aberrant axon outgrowth or guidance in the IL2 sensory map (Ryder, unpub. results).

Allele *ct41* is a null or nearly null mutation that results in most of the gene function being abolished, with 60% of the worms possessing the mutant phenotype (Manser et al., 1997). The axon defect has two forms, either the "stop" or the "branch" (Mason, 1997) (**Figure 6**). In the stop form the axon stops and does not reach the nerve ring. In the "branch" defect, on the other hand, the axon fails to make a posterior loop but does appear to reach the nerve ring.

Panel A

Panel B
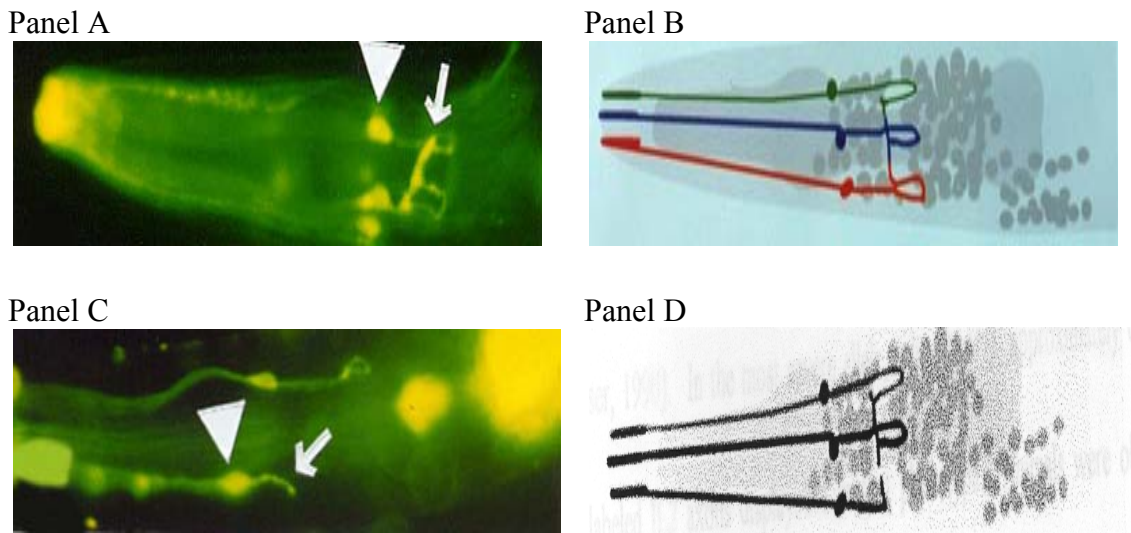
Panel C

Panel D



**Figure 6. DiO staining of the IL2 sensory map in wildtype worms and *mig-10* mutant worms.** Panel A, DiO staining of a wildtype worm showing the axons making the normal loops back to the nerve ring. The large arrowhead shows the cell body. The arrow shows the axon. Panel B, Schematic representation of the IL2 neurons. Panel C, DiO stain of *mig-10* (*ct41*) showing the stop defect. The arrowhead shows the cell body. The arrow shows the "stopped" axon. Panel D, Schematic representation of the stop defect observed IL2 neurons in *mig-10* mutant animals. (Mason, 1997)

Manser et al. (1997) cloned and sequenced *mig-10* through transformation rescue experiments using overlapping clones F10E9 and F49A3. *mig-10* is located on chromosome III between *unc-86* and *unc-116* (Manser et al., 1997). There are two

potential MIG-10 proteins depending on whether exon 1a or 1b is used. Transgenic

rescue of a *mig-10* mutant can be obtained without 1a, which is 6.2 Kb upstream from the

rest of the gene (**Figure 7**). MIG-10 shares a large region of homology with mammalian

SH2 domain proteins, Grb-7 and Grb-10 (Manser et al., 1997). The Grb proteins contain

an SH2 domain, pleckstrin homology (PH) domain and proline-rich regions. MIG-10

shares all of these regions, except the SH2 domain. The PH domains have been indicated

in mediating several types of signal transduction events (Lemmon et al., 1996). Thus, it

is likely that MIG-10 is involved in signal transduction and is not a secreted molecule or

cell surface molecule. Recently, *mig-10* was shown to contain RA-associating domains,

suggesting that it maybe an effector of G proteins of the ras family (Wojcik et al., 1999).

Manser *et al*, 1997.
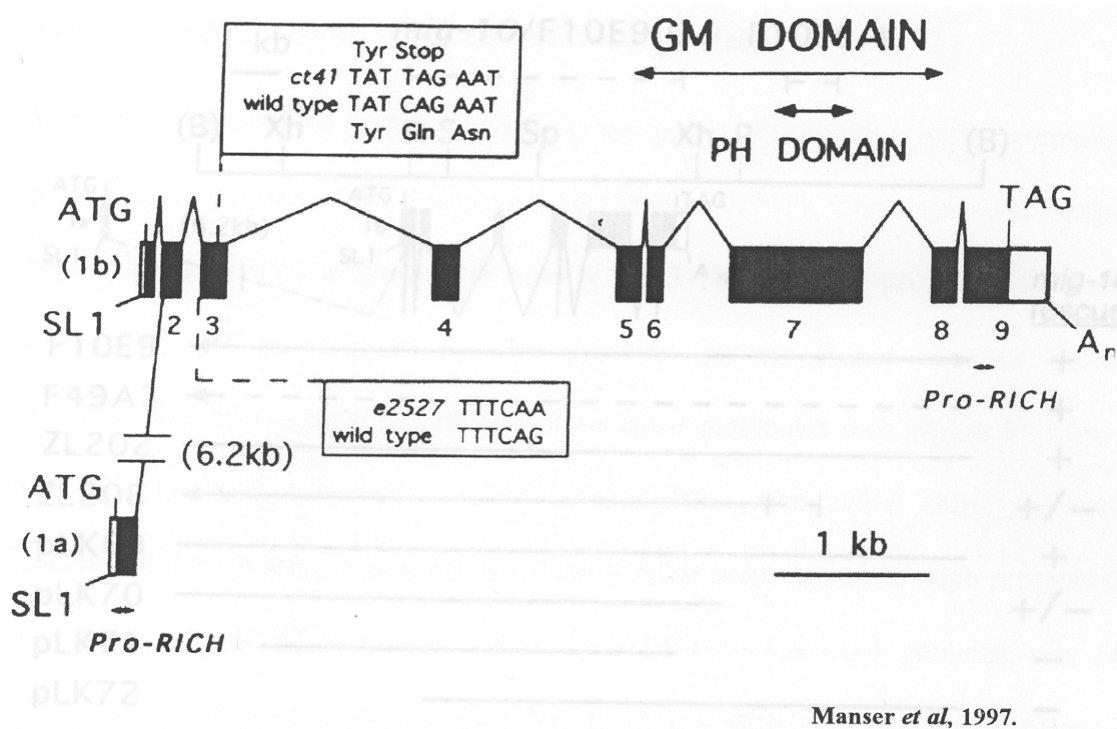
**Figure 7. Genomic organization of the *mig-10* gene.** Exons and introns of *mig-10* coding sequences are shown in detail. Solid boxes indicate coding sequence, while open boxes indicate untranslated regions. Exons are numbered, including the alternative first exons 1a and 1b. Dashed lines indicate the positions and nature of the mutations *ct41* and *e2527*. Arrowheads denote regions of sequence homology: GM (Grb and Mig), PH (Pleckstrin homology domain), and Pro-Rich (proline rich regions). Other abbreviations are as follows: SL1 (trans-splice leader), A_n (site of polyadenylation), ATG (start codon), TAG (stop codon). Figure is taken from Manser *et al*. (1997).

## 1.6 Research goals

### 1.6.1 *dig-1*

Why would it be important to study *dig-1*? By studying *dig-1* we can begin to understand how this gene functions during development of nervous system in *C. elegans*. The first goal was to confirm that a gene encoding a candidate adhesion molecule located near where *dig-1* mapped was the gene for *dig-1*. This result was achieved using PCR and sequence analysis of two *dig-1* alleles, *nu336* and *n1480*. A second goal was to study the structure and function of DIG-1, by analysis of the severity of several phenotypic aspects.

A third goal was to show that the predicted gene structure is substantially correct by reverse transcriptase polymerase chain reaction (RT-PCR), sequence analysis, and northern analysis. A fourth goal was to show potential differences in expression at the RNA level between mutant and wild type animals due to RNA nonsense mediated decay, by using RT-PCR and northern analysis. A fifth goal was to determine where DIG-1 was localized in the worm by raising an antibody to recombinant proteins, generated from cDNA from the 5' and 3' ends of the *dig-1* gene.

### 1.6.2  *mig-10*

Why would it be important to localize the activity of MIG-10? By studying MIG-10 localization, we can begin to elucidate the function that *mig-10* has during axon guidance of the axons projected by the IL2 neurons. To localize the expression of MIG-10, we generated transgenic animals by injection of two constructs to generate a GFP::MIG-10 (Green fluorescent protein) fusion protein by recombination.

## 2.0 Materials and Methods

## 2.1 Nomenclature and genetic procedures for worm strains

Genetic nomenclature is as described (Horvitz et al., 1979). Some worm strains were generated by mating, recombination events, and injection of DNA (**Appendix B**, Table 1, **section 2.20**). Other strains were obtained from the *Caenorhabditis* Genetic Center or were thawed from a stock in the Ryder laboratory (**Apendix B**, Table 2).

## 2.2 Maintaining worm strains

Strains were maintained and grown as described by Brenner (1974). Nematode Growth Medium (NGM) agar plates with a spot of *E. coli* OP50 bacteria were used to maintain worms (Brenner, 1974). Worms were picked to new plates every 2-4 days depending on the growth temperature. Most worms, except for injected (**section 2.20**) animals and their offspring (25°C) or mating plates (15°C), were grown at 20°C. Worms were also placed at 15°C when they were not being maintained for an extended period. For long term storage, worms were frozen at -80°C. Briefly, worms were washed off freshly starved plates with 2 ml of room temperature S basal (0.1M NaCl, 50mM $KH_2PO_4$, 5 µg/ml Cholesterol) and chilled on ice for 15 min. Excess S basal was removed until 2 ml were left. Two ml freezing solution plus agar (0.1M NaCl, 0.04mM $KH_2PO_4$, 300g glycerol, 5.6mM NaOH, 0.3mM $MgSO_4$ and 0.4 g agar per liter at 55°C), was added and mixed. Worms were then placed into 1.5ml screw cap tubes and frozen. Frozen worms were recovered by removing 200µl from the frozen stock and placing it onto a plate. Plates were then checked the next day for recovered worms.

If bacterial or fungal contamination was encountered on a maintenance plate, 10 gravid worms were picked to a clean NGM agar plate containing a drop of bleach solution: 80% M9 buffer (Brenner, 1974) and 20% bleach (5% sodium hypochlorite). This killed the adult worms and the contaminant, but spared the eggs. The offspring (L1-L4) were picked to a clean plate and maintained.

## 2.3 Labeling neurons

In order to visualize IL2 neurons, worms were stained using a lipophilic long-chain fluorescent dye, DIO (dialkylcarbocyanine 3,3'-dioctadecyloxacarbocyanine perchlorate, Molecular Probes Cat. Num. D-275, C. Bargmann, personal communication). One ml volume of 50mM Calcium Acetate (CaAc) was used to wash well-fed worms off of a plate. The worms were transferred with a Pasteur pipette to a 1.5ml tube. Several plates of worms were washed, combined, and the worms were allowed to settle to the bottom for 5 minutes. After removing the supernatant, the worms were rinsed with 1 ml of 50mM CaAc and allowed to settle for an additional 5 min. The supernatant was removed and replaced with 1ml of 50mM CaAc plus DiO, at a final DiO concentration of 10 µg/ml (DiO stock was 1 mg/ml in dimethyl formamide stored at -80°C wrapped in foil). The tubes were then wrapped in aluminum foil and gently rotated for 2hrs. The worms were then allowed to settle for 5 min and the supernatant removed. 1 ml of $dH_2O$ was then added to wash the worms and removed after the worms settled. The rinsing step was repeated one time. Worms were then moved to a clean plate with a pasteur pipette and allowed to feed for approximately a half-hour. This step removed excess background fluorescence in the gut. Three

microliters of M9 buffer was put on an agarose pad (2% agarose, 10mM sodium azide) on a microscope slide.  Worms were picked to the M9 and covered with a cover slip.  IL2 neurons were then examined by fluorescence microscopy.


## 2.4  Counting adult progeny from several worm strains

To determine the total number of adult progeny per worm from a single strain of worms, the adult progeny were counted in the following manner.  Ten single L4 hermaphrodites were picked from either N2 or an allele of *dig-1* to separate 60mm NGM agar plates and allowed to lay eggs overnight at 20°C.  The next day the worm was moved to another plate and allowed to lay eggs at 20°C; this was repeated until day 10.  The progeny were allowed to grow at 20°C.  The adult progeny were then counted and removed from the plate.  The progeny were then totaled per day.  The total number of worms from each parent were summed to get the total number of adult progeny for that single worm.  Differences between brood sizes of worm strains were compared by ANOVA using the computer program SPSS.


## 2.5  Washing and freezing worms from plates for either DNA or RNA extractions

Worms were washed (2X) off 100 mm, non-starved (RNA) or starved (DNA), half enriched agarose plates with 2 ml ice cold EN (10mM EDTA pH 8.0, 100 mM NaCl) per plate into a 15 ml conical tube.  The worms were then centrifuged for 2 min at 500 X g (1400 rpm) in a GH-3.7 horizontal rotor at 4°C using a Beckman GPR centrifuge and the supernatant was removed.  Two ml of ice cold EN was added to the pellet, mixed and the mixture was left on ice for 10 min.  The worms were then centrifuged as before and

2.5 ml sucrose solution  (68% sucrose, 3 mM EDTA pH 8.0) was added and the sample was gently vortexed (Lewis and Fleming, 1995).  A layer of EN (3 ml) was then carefully added over the sucrose and worm mix.  This was then centrifuged at 1,100 X g (2200 rpm) at 4°C in the Beckman centrifuge for 5 min.  A sterile Pasteur pipet was then used to remove the upper band of worms to a fresh tube; 10 ml of EN was added to the worms to dilute any sucrose remaining.  The worms were then centrifuged for 2 min at 4°C at 500 X g (1400 rpm) and the supernatant discarded.  The pelleted worms were then transferred to either a 1.5 ml Eppendorf tube (DNA ) or a 1.5 ml screw cap tube (RNA) and spun in a microfuge for 30 sec.  For DNA preparations the supernatant was removed and the tube placed at -80°C until extraction.   For RNA preparations the supernatant was removed and the tube placed into liquid nitrogen for 1 minute, then stored at -80°C until extraction.

**2.6 DNA isolation from worms**

DNA was isolated for polymerase chain reaction (PCR, **section 2.11**) and Southern analysis in the following manner.  The frozen worms  (**see section 2.5**) were allowed to thaw and all the excess liquid was carefully removed with a sterile Pasteur pipet.  Then 500 µl worm lysis buffer  (130mM Tris pH 8.5, 1% SDS, 50mM EDTA, 0.1 M NaCl, 1% betamercaptoethanol and 100µg/ml proteinase K) was added and incubated at 65°C overnight.  The tube was mixed every 15min for the first 90min.  After incubation, tubes were allowed to cool and 1ml of phenol was added and mixed by vigorous shaking.  The tube was then centrifuged (Eppendorf Micro centrifuge 5415 c) at top speed for 5 min and the supernatant (aqueous phase) was removed to a clean tube.

One ml of phenol /chloroform (50/50 v/v) was added, and the sample was then mixed and centrifuged at top speed for 4 min.  The supernatant was then removed to a clean tube. This step was repeated and then 1 ml of chloroform was added.   The sample was then mixed and then centrifuged at top speed for 3 min. The supernatant was removed to a clean tube and 1 ml of 100% ethanol was added.  DNA was precipitated at room temperature for 15 min with mixing on a nutator (Clay Adams, brand).   The DNA was then pelleted after a 5min centrifugation and washed with 70% ethanol.  After drying, the DNA was re-suspended in an appropriate amount of TE (10mM Tris·Cl, 1mM EDTA pH 7.6).


A second DNA extraction method was performed in cases where it was difficult to grow worms until starved or unnecessary to have a large preparation of DNA from a particular worm strain.  Ten, or more, L4 or younger worms were picked and placed in 2.5 µl of digest buffer  (50mM KCl, 10mM Tris pH 8.2, 2.5mM $MgCl_2$, 0.45% NP-40, 0.45% Tween 20, 0.01 % gelatin, 0.12mg/ml proteinase K) in the cap of a PCR tube.  The tube was placed over the cap, centrifuged for 10 sec to pellet the worms, a mineral oil overlay was added and the tube was placed at -80°C.  After a minimum of 30 min, tubes were removed and placed into a thermal cycler (Progene, Techne, Inc).  The worms were then digested at 65°C for 1 hr.  After digestion, proteinase K was inactivated by incubation at 95°C for 15min.  The sample was then ready for PCR (**section 2.11**).  This procedure can also be performed on a single worm.

**2.7 RNA isolation from worms**

RNA was isolated for both reverse transcriptase polymerase chain reaction (RT-PCR, **section 2.12**) and northern analysis in the following manner.  The frozen worms (**section 2.5**) were thawed with 1 ml of Trizol$^{tm}$ (GibcoBRL Cat# 15596-026) and vortexed for 5 min.  The tube was then incubated at room temperature for 10 min followed by another 5 min vortex.  Chloroform  (200 µl per ml of Trizol$^{tm}$) was added and the tube was shaken vigorously for about 20 sec.  The sample was then incubated at room temperature for 5 min and centrifuged at 12,000 x g for 15 min at 4°C.  The upper, aqueous phase containing RNA was then recovered and the red organic phase containing DNA and protein was left behind.  RNA was then precipitated with Isopropyl alcohol (500 µl per ml of Trizol$^{tm}$).  The RNA was allowed to precipitate at room temperature for 15 min after the tube was inverted ten times.  RNA was then pelleted at 12,000 X g for 10 min at 4°C.  After decanting the supernatant, the pellet was washed with 70% ETOH by pipetting until the pellet was dislodged from the tube.   The pellet was re-centrifuged at 7,500 X g for 5 min at 4°C and the supernatant removed.  The RNA pellet that was for RT-PCR (**section 2.12**) was air dried and re-suspended in 100 µl of nuclease free water (Promega RNA kit). The RNA pellet that was for northern analysis (**Appendix A**) was air dried and re-suspended in a maximum of 50µl of nuclease free water with formamide (final concentration of formamide is the same as northern loading buffer, Ausubel et al., 1998). To help re-suspend the RNA, samples were heated for 5 min at 60°C.  RNA was stored at -80°C, until use  (Protocol provided by the Hekimi laboratory, McGill University).

## 2.8 Plasmid isolation and preparation

From a frozen stock, bacteria containing the plasmid of interest were streaked onto LB plates containing 50 μg/ml of either Ampicillin or Kanamycin and grown overnight at 37°C.  Single colonies were then inoculated into test tubes containing 5, 10 or 100 ml of LB with 50 μg/ml of Ampicillin or Kanamycin.  Samples were grown overnight at 37°C at 250 rpm, in a New Brunswick Scientific controlled environment shaker incubator.

Plasmids were isolated using a Qiaprep® Spin Miniprep Kit (Qiagen Cat. No. 27104).  Larger amounts of plasmids were isolated using a Plasmid Maxi kit  (Qiagen Cat. No.12162).  Isolated plasmids were eluted from spin columns in 50μl of (TE, pH 7.6) for general applications or 50μl of buffer EB (10mM Tris-Cl, pH 8.5) for sequencing or enzymatic digestion.  For injections, plasmids were precipitated by addition of 1M KAc to a final concentration of 0.1M KAc and 2 volumes of ethanol. Precipitated DNA was pelleted by centrifugation and re-suspended in sterile injection buffer (2% polyethylene glycol, 20 mM potassium phosphate, 3.0 mM potassium citrate, pH 7.5). To determine the concentration and purity of plasmid samples, absorbance of diluted samples readings at 260nm and 280nm was measured on a Beckman DU 650 spectrophotometer.

## 2.9 Plasmid clean-up after digestion or to increase concentration
   for sequencing

In some cases, it was necessary to clean plasmids before use after digestion, for injections or to concentrate the plasmid for sequencing.  To concentrate or clean plasmids

a Qiaquick™ Gel extraction kit (Qiagen Cat. No. 28704) was used.  This procedure was modified from the Qiaquick™ Spin handbook (gel extraction kit protocol page 24).  First 5X volume of QG buffer was added to each sample.  One volume (equal to starting volume) of isopropanol was added and mixed.  The sample was then added to a QIAquick™ column.  The column was placed in a 2ml collection tube and centrifuged at 16,000 X g for 1 min.  The flow-through was discarded and 0.5 ml of buffer QG was added to the column (the kit protocol indicates that the second wash in QG is necessary for injections).  The column was placed into the same collection tube.  After centrifuging at 16,000X g for one minute the collection tube was emptied and 0.75ml of buffer PE was added to the column.  The column was centrifuged again for one minute and the flow through was discarded. The column was centrifuged for an additional minute to dry.  DNA was either eluted in 50µl of heated TE (pH 7.6) and placed on the column for 1 minute, or for sequencing in 50µl of buffer EB placed on the column and left for 1 min.  The samples were then eluted by centrifugation at 16,000 X g into a 1.5ml eppendorf tube.  Samples that were used for injections were then placed into injection buffer as described (**section 2.8**).

## 2.10 Primer design

Primer names describe, in order, the primer source, base pair position (Kb), orientation (forward t, reverse b) and special notes for use (**Appendix B,** Table 4,5).  In addition, some primers were designed with restriction sites for cloning or to include either the T7 or T3 promoter to produce RNA. RT-PCR primers spanned an intron to allow differentiation between DNA and RNA amplified products.

In general, several rules were followed for primer design.  First, primers were designed to contain 50% or higher C and G content. Second, primers were designed so that at the 3' end three out of the last five bases were either C or G, but no runs of either of those were allowed (Dieffenbach and Dveksler, 1995).  Third, primers were subjected to a BLASTN analysis to determine how many times a primer matched with the sequence of interest (Sanger Center).  If a potential primer matched perfectly at more then one place in the *dig-1* gene it was in general not used; several exceptions to this rule were made in areas of the gene that were highly repetitive. For sequencing, potential primers were not used if there was a second match site that differed by 3 or fewer base pairs from the target site.

## 2.11 Polymerase chain reaction (PCR)

## 2.11.1 PCR using purified genomic DNA as template

Oligonucleotide PCR primers were purchased as a lyophilized powder from Fisher/Genosys.  The powder was dissolved in an appropriate amount of TE (pH 7.6) to make 100 µM stocks that were stored at -80°C.  A 3µM working primer stock diluted into reagent grade water was kept at -20°C, until use in PCR.  Deoxynucleoside tryphosphates (dNTPs) were purchased as a set from Perkin Elmer.  A dNTP mix (2.5mM) stocks was made by mixing equal amounts of the four separate nucleotide stocks (10mM).

For short PCR (2Kb or less), a master mix was made on ice, to insure that samples were homogenous and to reduce contamination.   A 50µl reaction contained 26.75µl of water, 0.5µl of 5 u/µl *Taq* DNA polymerase (Fisher Scientific), 5µl of 10X PCR buffer B

(Fisher Scientific), 5µl of 25 mM MgCl$_2$, 1.75µl of 2.5 mM dNTP mix (Perkin Elmer), 1µl of template DNA, and 5µl of each primer (3µM stock), in a 0.2 ml thin walled tube. Reactions were then placed into a Progene thermal cycler (Techne, Inc, Princeton, NJ). Thermal cycler conditions were: 94°C for 10min, 30 cycles of 94°C for 30 sec (denaturing), 64°C for 1min (annealing), and 72°C for 2 min (primer extension) followed by an incubation at 72°C for 10 min, as a final extension. The annealing temperature was sometimes altered if no product was amplified after annealing at 64°C. After amplifications, reactions were left at 4°C, until removal from the thermal cycler.

For long range PCR (used for PCR product of greater than 2Kb), two separate master mixes were made on ice. The first master mix consisted (per 50µl reaction) of 7µl of water, 7µl of 2.5mM dNTP mix (Perkin Elmer), 1µl of template DNA and 5µl of each primer of interest (3µM working stock). The second master mix was made with 19.25µl of water, 5µl of buffer 3, and 0.75µl of enzyme mix (Roche Expand Long Template PCR System, Cat # 1 681 834). Twenty-five µl of master mix one was added to a 0.2 ml thin wall tube. An oil overlay was added before master mix 2 was added. Twenty-five µl of master mix 2 was added and centrifuged through the mineral oil overlay. Samples were immediately placed into a thermal cycler. The cycling conditions were: 94°C for 2 min; 20 cycles of 94°C for 10 sec, 64°C for 30 sec, 68°C (time dependent on length of product, 1min per 1kb +20sec per cycle), followed by an incubation at 68°C for 7min. The annealing temperature was 68°C for all reactions. After amplification, reactions were then left at 4°C, until removal from the thermal cycler.

**2.11.2 Ten or one worm PCR**

Short range PCR of 10 or 1 worm lysates was done in the following manner. A master mix was made, consisting (per 25µl reaction) of 11µl of water, 2.5µl of 2.5mM dNTPmix, 2.5µl of 10X PCR buffer 3 (from long PCR kit), 0.75µl enzyme, and 2.5µl of both primers (3µM working stock). 22.5µl of the master mix was added to each tube containing 2.5µl of lysate (**section 2.11.2**) for a final reaction volume of 25µl in 0.2ml thin wall tubes. Lysates were then cycled as described for short range PCR (**section 2.11.1**). Several different concentrations of enzyme were used, but 1.5µl/reaction gave the best results consistently. Magnesium concentration was also altered; in some instances it optimized PCR yield and in others it did not. Other parameters, such as primer concentration and dNTP concentration, were kept constant. To increase the yield of PCR products for digestion, PCR products were sometimes cleaned with the Promega PCR clean kit and re-amplified (**section 2.11.3**).

Long range PCR of 10 or 1 worm lysates was done in the following manner. A master mix was made, consisting of (per 25µl reaction) 11.75µl of water, 2.5µl of 2.5mM dNTP mix, 2.5 10X PCR buffer 3, 1.5µl of enzyme (Roche Expand Long Template PCR System). 2.5µl of each primer (3µM stock) were added per reaction; if a single primer set was used for all reactions, this primer set could be added to the master mix. Lysates were then cycled as described for long range PCR (**section 2.11.1**). The two master mix method was tried but never worked for 10 worm lysates.

### 2.11.3  PCR product purification and re-amplification

Ten worm PCR plus restriction digestion were used to screen recombinant worms from KP951 *dig-1*(*nu336*) to detect an additional EcoRI restriction site that was present in this mutant.  The majority of these amplifications from recombinant worms did not have a discernable PCR product on an agarose gel.  Since it was necessary to see the product after digestion, the products were re-amplified. To re-amplify PCR products it was determined that the product should be cleaned prior to re-amplification.  PCR product cleaning was done using the Promega Wizard® PCR Preps DNA Purification System.  Samples were eluted in TE (pH 7.6) and 5µl was used in a second PCR.


### 2.12 Reverse Transcriptase Polymerase Chain Reaction (RT-PCR)

RT-PCR was performed on wildtype (N2) RNA using the Promega Access RT-PCR system  (catalog# 1280).  Several different primer sets were used to generate cDNA from different sites in *dig-1*.  To optimize the RT-PCR, a titration of $MgSO_4$ was done on most primer sets (for SL1 experiments and when using *dig-1* (*n1480*) RNA, a standard 2mM $MgSO_4$ was used).  A range of 1.0mM to 3.0mM $MgSO_4$, in 0.5mM increments, was used with each set of primers. The optimal concentration of $MgSO_4$ for a primer set was used to amplify a large quantity of cDNA, which was later used for cloning. Standard components for Mg optimization per reaction were, 13.4µl water, 10µl 5 X buffer, 8.3µl of each primer of interest, 2µl of $MgSO_4$, 1µl Amv reverse transcriptase (except for PCR control), 1µl Tfl DNA polymerase and 1µl RNA (amount depended on concentration of RNA; the target range of the kit is 10 pg to1µg).  A master mix was made and 45µl aliquot was removed from the master mix and added to the PCR (no RT)

control tube and 5µl of water was added to make a 50µl reaction. Seven µl of Amv reverse transcriptase (5U/µl) was added to the master mix. Forty-six µl of the master mix was then added to each RT-PCR tube. To vary the $MgSO_4$ concentration, it was necessary to add appropriate amounts of either water or $MgSO_4$ per reaction, ranging from 1µl to 4µl. A concentration of 2mM $MgSO_4$ was determined to be optimal. After the samples were mixed, an oil overlay was added. Cycling conditions for RT-PCR were an initial 48°C for 45min for first strand synthesis, followed by 40 cycles of: 94°C for 30sec, 64°C for 1min, and 68°C for 2min. After cycling, a final extension was done at 68°C for 7min followed by 4°C until samples were removed from the thermal cycler.

Subsequent RT-PCR reactions used the optimal concentration of Magnesium. For RT-PCR using the SL1 primer as the only standard per reaction primer a second primer was added directly to the reaction after the master mix was allocated. For RT-PCR reactions that used a different template (N2, N2 embryonic, mutant) a master mix was made and divided in half after the addition of the DNA polymerase. Template was then added to the master mixes. The master mix was then added to the PCR tubes and the AMV reverse transcriptase was added to the master mix. The RT-PCR tubes then had an appropriate amount of master mix added. In most cases in which two different templates were amplified, several different primer sets were used. Thus, these tubes had less of the master mix added, since separate primer sets were added to each tube.

## 2.13  Agarose Gels

Standard 0.8% agarose gels were used for most electrophoresis separations of DNA. Seventy-five ml of 1X TBE were added to 0.6 g of agarose and microwaved until all the agarose was liquified.  The agarose was poured into a gel caster (Owl Scientific), with comb, and allowed to cool for several hours. After the gel had cooled sufficiently, the comb was removed and the gel placed into the electrophoresis unit (Owl Scientific). Running buffer (1 X TBE 0.089 M Tris·base, 0.089 M boric acid, 0.002 M EDTA) was then poured into the unit until it just covered the agarose gel. Samples were mixed with 0.25 volume of gel loading buffer (0.25% Bromophenol blue, 0.25% xylene cyanol, 30% glycerol) and loaded into the gel. Wells were then loaded with sample and the gel was run at 75 V for three hours.  Along with the samples, molecular weight markers (MWM) were also run.  Two standard MWM used were Lambda DNA-Hind III digest: 23,130bp-2,027bp (New England Biolabs, Inc.) and 1 Kb DNA ladder: 10kb-500bp (New England Biolabs, Inc.).

To visualize DNA or dsRNA, Ethidium Bromide (EtBr) was added to the gel at a final concentration of $6X10^{-3}$ mg/µl.  After the gel was run, it was then removed from the electrophoresis unit and placed on an Ultra-Violet trans-illuminator for visualization. The DNA was then visualized and photographed with a DS34 Polaroid Direction Screen Instant camera (film was Polaroid 667) orange #15 filter, with shutter speed set at ½ sec and F stop set at 15 or 16.

## 2.14  Agarose gel purification of DNA and cDNA

Several PCR and RT-PCR amplified products required purification on a TAE agarose gel before cloning.  DNA products were electrophoresed on a 0.8% or 1% agarose gel was run (**section 2.13**) until all bands were separated.  The band of interest was then identified and removed from the gel, using a hand held UV light source at low setting and a spatula that had been flamed prior to use. The gel slice was then placed into a pre-weighed 1.5ml eppendorf tube and then re-weighed.  After determining the weight of the gel slice an appropriate amount of buffer QG from the Qiagen QIAquick gel extraction kit was added.  The protocol for the kit was then followed.


## 2.15  RNA production for RNA interference injections (RNAi)

RNA was produced using a Riboprobe® in vitro Transcription system kit (Promega).  Prior to in vitro transcription, the plasmid of interest was cut with two different enzymes that flanked the target DNA of interest.  Two reactions were set up using the linearized DNA template to produce RNA complementary to each DNA strand.


After transcription, the DNA template was removed by treatment with DNase (RQ1 RNase-Free Dnase, Promega). Both RNA samples were re-suspended in TE.  Equal amounts of RNA samples were heated to 95°C for 1min, mixed and incubated at 37°C for 30min.  A small amount of this mixture was electophoresed in a lane next to the PCR product in an agarose gel to test if the RNA was present and in a double stranded form.  Double stranded RNA was then injected as described (**section 2.20**).

**2.16 Restriction enzyme digestion**

Plasmids, PCR products and RT-PCR products were subjected to restriction enzyme digestion (**Appendix B**). Digestions were normally done in 20μl volumes. In some instances it was necessary to digest more then the normal amount of vector or insert for cloning, thus the volumes for the digest were adjusted. In general for a standard digestion (no BSA needed) the reaction mix contained 12 μl or 7 μl of water, 2 μl of appropriate 10X buffer, 1 μl of enzyme and either 5 μl or 10 μl of template for either plasmid or RT-PCR product. For enzymes that require BSA, 2 μl of BSA were added and the amount of water was reduced.

**2.17 Competent cells**

Cells were made competent in the following manner. From a LB plate streaked with bacteria (*DH5α*, *BL21* or *BL21*(DE3)) a single colony was picked to 10 ml of LB media and grown over night at 37°C in a shaker bath (250 rpm). From this 10 ml culture, 1 ml was used to inoculate 100 ml of LB media. The sample was grown to an $OD_{590}$ of 0.375. Cells were allocated (15 ml) and pelleted at 950 X g (2,000 rpm) for 10 min at 4°C on the Beckman GPR centrifuge. The supernatant was decanted and the cells re-suspended in 7.5 ml of ice cold 1M $CaCl_2$. Cells were pelleted as before and the supernatant removed. The pelleted cells were re-suspended in 1 ml of ice cold 1M $CaCl_2$, allocated into 100 μl samples, and frozen at -80°C until transformations were performed (**section 2.19**).

**2.18 Ligations**

Inserts and vectors were prepared in the following manner for ligations.  Inserts were generated by either PCR (**section 2.11**) or RT-PCR (**section 2.12**), with unique restriction sites engineered (**section 2.10**) at the ends of the products. Alternatively, the adenosine overhang (A-tail) produced by Taq was used in cloning.  Restriction sites were designed to match unique restriction sites in the multiple cloning site of a vector and, if needed, to put the insert and vector in frame for expression.  Vectors were prepared by a mini-preparation (**section 2.8**), followed by digestion with appropriate enzymes (**section 2.16**).  Both insert and vector, after digestion, were purified by gel electrophoresis (**section 2.13**) and Qiagen cleaning (**section 2.9**).

Ligation reactions were done in 0.2ml thin wall tubes, in the following manner. Ligation reactions contained 1μl of 10X ligation buffer, 1μl of vector, 7μl of insert or water and 1μl (400u/μl) of T4 DNA ligase (New England Biolabs, Catalog number 202S).  In some cases, depending on the concentration of either the vector or the insert, up to 10 μl of the vector and 7 μl of the insert were used in a 20μl reaction.  The ratio of vector to template was approximately 1:8 to 1:10 molar ratio.  The amount of enzyme and buffer were then also changed to reflect the volume changes.

For cloning cDNA products, for sequencing a TOPO TA cloning kit or T-tailed vector was used (Invitrogen TOPO TA Cloning Kit, pCR 2.1 Topo vector, TOP10 cells). The reaction mix contained: 4 μl fresh PCR product, 1 μl of vector (10ng) and 1μl of salt solution. The ratio of vector to insert was approximately 1:8 molar ratio.  Incubation

times were 30 min for ligation and 1hr for transformation. Controls for ligations were those provided by the kit and water blank, where only vector was present. A second T-tailed system pGEM-T cloning kit (Promega) was also used to clone constructs for the *mig-10* injections. Of the two kits, the TOPO TA kit was determined to be the most efficient at cloning by A-T overhang method.

### 2.18.1    Adding an A-tail to long range PCR products for T-tailed vector cloning

In some instances where long PCR products were cloned without specific restriction sites, it was necessary to A-tail the insert. The long Taq polymerase mix has exonuclease activity that cleaves off the A-tail over-hang that is typically generated by Taq polymerases. PCR products were A-tailed in the following manner. 7 µl of PCR product (Gel separated and Qiagen column purified) were mixed with 1µl 10X PCR Buffer, 1µl dATP (0.2 mM), and 1 µl of taq polymerase (5U). Reaction mix was then incubated at 70°C for 30 min (Promega, pGEM®-T and pGEM®-T easy vector systems, Technical manual).

### 2.19    Transformations

Cells were transformed in the following manner. To a 100µl sample of competent cells, thawed on ice, 5µl of a ligation mix or plasmid sample were added and incubated on ice for 30 min. The cells were re-suspended and heat shocked at 37°C for 2 min. Pre-warmed (37°C) LB media (0.5 ml) was added and the sample gently mixed. The transformation reaction was incubated at 37°C for an hour and then the cells were re-suspended. Cells were plated by spreading 10µl, 100µl, and concentrated cell suspension

(cells pelleted and resuspended in a drop of LB) on LB plates containing 55µg/ml of ampicillin (or appropriate selection marker) and allowed to grow overnight at 37°C. Colonies were screened for the correct construct by plasmid isolation (**section 2.8**) and digestion (**section 2.16**). A glycerol stock of each colony screened was frozen and stored at -80°C.

## 2.20   Microinjections and transgenics

Worms were microinjected with DNA to produce transgenic lines (Mello et al., 1991; Mello and Fire, 1995). Worms were injected under Nomarski optics using a Zeiss Axiovert S100 inverted microscope with rotating glide stage, using 5X and 40X Zeiss Achromat lenses. Needles were mounted on a Narashige microtool collar and positioned using a Narashige micromanipulator (type M0-102L, Tokyo, Japan) with joystick control. Tygon® tubing (1/16 inch) (Norton Performance Plastics, Akron, OH) attached to a nitrogen tank was used to pressurize the needle. Pressure control was obtained with a Western Medica compressed gas regulator (Tritech Research, Los Angeles, CA).

A single Kwik-Fil™ Borosilicate glass capillary tube (Item number 1B120F-4, World Precision Instruments, Inc. Sarasota, FL) was used to make injection needles. A single capillary tube was placed into a PUL-1 horizontal pipet puller (World Precision Instruments, Inc.) and the automatic button was pushed, resulting in two needles being pulled. The heat and delay settings were adjusted to produce needles of desired length and diameter. Settings were adjusted periodically as the heating cell aged.

Needles were etched with hydrofluoric acid (HF). This was done to sharpen the needle as well as to insure that the injection materials were able to flow out of the needle. To etch needles, 50 µl of HF and 50 µl of water were placed on a plastic petri dish in the fume hood. Needles were then attached to the Tygon® tubing and placed into the HF, while nitrogen was being blown through the needle at 60 PSI or 90 PSI. After a few seconds or if bubbles could be seen, the needle was moved into the water drop, with pressure, for a few seconds. Needles were then stored in a box to prevent dust accumulation. Some needles, even after etching, did not flow and were deliberately broken on the worm or a salt crystal on the pad to improve flow.

DNA solutions for injections, after a 5min centrifugation at maximum speed, were loaded into the injection needle via a drawn-out 1ml syringe (Becton Dickinson, Cat# 309602). Drawn-outs were made by flaming the 1ml syringe until it melted. The end was then pulled causing the melted area to be "drawn out." This was pulled until the plastic was a desired diameter. The drawn-out was then filled with 1 to 5µls of injection materials and the drawn out tip was placed into the needle. The needle tip was then filled, by applying a small amount of pressure to the syringe plunger. The very tip of the needle was filled by capillary action. In some instances, an air bubble was present in the tip of the injection needle. Air bubbles were removed in several ways. First, the air bubble could be removed with the drawn-out (only large bubbles). Second, the needle could be placed into a hydration chamber for about half an hour. Third, the needle could be "flicked" until the air bubble was dislodged. After air bubbles were removed, the

needle was placed into the Tygon® tubing, secured with parafilm and fastened to the microtool collar (Narashige).

Young adult hermaphrodites with a row of eggs in utero were picked from a plate using Series 700 halocarbon oil (Halocarbon Products, River Edge, NJ) and immobilized on a 2 % agarose injection pad. To immobilize the worms on the pad the worms were allowed to settle on to the pad or were manipulated with either the pick or the injection needle. Injection pads were made the day of injection. 2 % agarose in either water or TE was melted in a test tube in the microwave. A drop of agarose was then placed on a 24 X 60mm Corning cover glass and a second cover glass was placed over this. After a minute, the cover glass sandwich was separated. The agarose pad was then placed into the drying oven at approximately 60°C for an hour to two hours. The pads were then used immediately. After the worm was immobilized, the cover slip was placed onto the stage of the microscope with the vulva oriented to the left, which placed the distal gonad to the right, making it easier to inject. The worm was then positioned in the middle of the optical field under the 40X objective and then the objective was changed back to the 5X for gross positioning of the needle. With the needle near the desired injection point, the objective was changed back to the 40X and the needle was more precisely positioned near the gonad. The pad was pushed toward the needle until the needle either punctured the worm or made an indentation in the worm. If the needle did not immediately puncture the worm, the side of the scope was tapped until the needle punctured the body. Once the needle was in the gonad, the DNA mix was injected under $N_2$ pressure until the material filled the bend in the gonad. After injection, the needle was removed. Usually

only one side of the gonad was injected. Injected worms were recovered under the

dissection scope. Worms were gently washed off the injection pad with M9 buffer and

picked to an NGM plate containing OP50 bacteria.


Two co-injection markers were used, *dpy-20* (Clark et al., 1995) (pMH86,

100ng/µl) or *elt-2*::GFP (pJM68, 75ng/µl ). When pMH86 was used, injections were

made into *dpy-20* (*e1282ts*) hermaphrodites. The worms were then placed at 25°C for

recovery. This temperature is the non-permissive temperature for this temperature

sensitive allele. The phenotype is first observed at the L3 stage of larval development

and results in a pronounced difference between rescued and non-rescued animals (more

so at 25°C then 20°C) visible in the dissecting microscope on plates. Rescued animals

with the non-Dpy phenotype were selected. Those animals that had non-Dpy self

progeny (indicating inheritance of the transgenic *dpy-20* marker) in the F2 generation

were selected and screened for expression of GFP encoded by the injected construct of

interest (Chalfie et al., 1994). When *elt-2*::GFP was used, injections were made into

either *mig-10* (*ct41*) or *mig-10* (*ct41*)/*dpy-17*(*e164*) *unc-32*(*e189*) animals. These worms

were screened for the expression of GFP in the intestine with the dissecting fluorescence

scope (Zeis M²BIO Stemi Su II).


## 2.21 Recombinant protein production and isolation

An initial small-scale protein production run was done to determine which of the

constructs produced protein (**Appendix B**, Table 3). Three ml of LB containing 55µg/ml

ampicillin was inoculated with *BL21*(DE3) containing one of the expression constructs

and allowed to grow to an $OD_{600}$ of 0.5. The 3ml of *BL21*(DE3) cells were added to 100ml of LB plus ampicillin (55µg/ml) and incubated at 37°C with shaking. This was grown to an $OD_{600}$ between 0.5 and 1. Cells were then split into two 50ml samples. An ampicillin booster was added to both samples to a final concentration of 55µg/ml, prior to being induced. One sample was induced with 500µl of 0.1M IPTG and incubated at 37°C for 2hrs. The second un-induced sample was placed at 37°C until harvest. After induction, a 1.5ml sample was taken to isolate plasmid. Plasmids were isolated as described before, in order to determine if there was plasmid in the sample after induction. After the plasmid sample was taken, samples were harvested by centrifugation at 6,500 X g for 10min at 4°C in a Beckman centrifuge (Model j-21b). The decanted media and cell pellet were frozen at -80°C. This procedure was performed on each of the expression constructs.

The frozen cell pellet was thawed on ice for both induced and un-induced cells. Two mls of BugBuster™ Protein Extraction Reagent (Novagen) were added and the cells mixed with gentle vortexing. The thawed cell mix was removed to clean 1.5 ml Eppendorf tubes and incubated with rotation on a nutator (Clay Adams, Brand) for 10 min at RT. After incubation, the cell debris was removed by centrifugation at 16,000 X g for 20 min at 4°C. The supernatant was then removed and stored at -20°C for further analysis. The cell pellet was also frozen at -20°C to extract inclusion bodies if necessary. To determine the amount of protein present in samples, a Bio-Rad protein assay was done. This was used to determine the amount of protein to load on a SDS-Polyacrylamide gel electrophoresis (SDS-PAGE, **section 2.22**). It was determined by

SDS-PAGE that the recombinant protein from construct 1 and 3 was present in inclusion bodies.

**2.21.1 Large scale recombinant protein preparation from *BL21*(*DE3*)**

To produce the recombinant protein in sufficient amounts for antibody production, a 5L bio-reactor (WPI, Bioprocess Lab) was used to grow the bacteria. After induction, the bacteria were harvested by centrifugation and the cell pellets collected. The pellets were scraped out of the centrifuge bottle with a clean, sterile spatula and placed into a zip lock bag. The sample was split into two, representing 2.5L of the original 5L culture. The bags of bacteria were frozen at -80°C. Protein was then isolated using the Bug Buster reagent as before, except that the bags were rinsed with Bug Buster to remove as much of the bacteria as possible. The inclusion bodies were then purified according to the Bugbuster Protein Extraction Manual. The protein was then resuspended in 10ml of 1X PBS with 8M urea.

To further purify the recombinant proteins, the samples were separated over a column containing Ni-NTA His•bind resin (Novagen, Cat# 70666-3). The recombinant proteins were engineered to have a His•tag at what would be the N-terminus. His•tags are 6 to 10 consecutive histidine residues that bind Nickel cations ($Ni^{+2}$). The $Ni^{+2}$ cations are immobilzed by chelating activity of reactive groups that are covalently attached to the solid support in the column (Novagen). NonHis·tag proteins do not have as many histidines in a row, preventing tight binding to the column. Thus, nonHis·tag proteins are washed off the column.

2.5 ml of the Ni·NTA His·bind slurry was mixed with 10 ml of the inclusion body mix. The mixture of His·bind slurry and protein was then mixed gently on a nutator (Clay Adams brand) at 4°C for 60 min. The mixture was then washed with wash buffer (8M urea, 0.1 M $NaH_2PO_4$, 0.01 Tris·Cl, pH 6.3) and the His·bind slurry with bound protein was pelleted at 1400 X g. The mixture was then washed and pelleted again. The protein was then eluted with elution buffer (8M urea, 0.1 M $NaH_2PO_4$, Tris·Cl, pH 4.5) and the His·bind slurry was pelleted by centrifugation at 1400 X g. The supernatant containing the recombinant protein was then removed. The elution was repeated 3X. The protein was eluted from the His·bind slurry as described. The protein was then dialyzed (Spectra/Por molecularporous membrane m.w. cutoff 3,500; Spectrum Medical Industries, Inc) to remove the Urea and then freeze dried.

## 2.22    Polyacrylamide gel electrophoresis (protein)

Polyacrylamide gels were done in the following manner. Before the monomer mix of acrylamide and bis-acryalmide were polymerized, two glass plates were washed with detergent and water. After the glass plates were washed, they were rinsed with 95% ethanol, dried with clean paper towel and then allowed to air dry. While these were drying the gel spacers were also rinsed with 95% ethanol and dried. The spacers were then placed on top of one of the glass plates (short side) and lined with the edges of the plate. The second glass plate was placed over the first glass plate on the spacers. This was clamped with the gel clamps. After the clamps were tightened enough to hold the plates, the plates and clamps were stood on end and aligned flush with the bench top.

The clamp screws were tightened (just to the point of being tight) and the edge where the glass will fit against the rubber gasket in the gel caster was tested.  If there was not a smooth edge the clamps were loosened and re-aligned.  After the plates were aligned, the sandwich was placed into the gel caster with the screws facing away from the lower buffer chamber.  Cams were then placed into position with the cam handle down and the cam ridge at the top or facing up.  The cams were then twisted 180° sealing the glass plates against the gasket in the bottom of the gel former.  At this point melted 1% agarose could be placed along the edge between the gasket and the glass plate to prevent leaking of the gel.

The initial separation of the crude cell lysate and cell pellet material was done on a 12% polyacrylamide separating gel and a 4% resolving gel.  To pour the separating gel, 12ml of the monomer mix (30% acrylamide (BioRad)), 0.8% bis-acrylamide (BioRad)) was mixed with 7.5ml of separating gel buffer (1.5M Tris-Cl pH 8.8), 0.3ml SDS (10%), 10ml of water, 150μl of Ammonium per-sulfate (10%), and 10μl of N,N,N,N'-Tetra-methyl-ethylenediamine (TEMED).  This solution was then drawn into a 60cc syringe fitted with an 18gauge needle (Precisionglide, Becton Dickinson & Co. Rutherford N.J.) and expelled into the chamber between the glass slides to just above the 15cm mark.  After the mix was added, resolving gel overlay buffer (0.375M Tris-Cl pH 8.8, 0.1% SDS) was added to the top of the gel to prevent inhibition of free radical formation by $O_2$.  The monomer mix was allowed to polymerize for an hour. To test if polymerization had occurred, a small amount of gel mix was drawn into a pasteur and left beside the gel.  The pasteur bubble was then pushed on lightly to see if the material was solid or not.  After

polymerization, the gel overlay buffer was removed by decanting and water was added to rinse the top of the gel. The stacking gel was then poured over the separating gel. The stacking gel consisted of 2.66ml of the monomer mix, 5ml stacking gel buffer (0.5M Tris-Cl pH 6.8), 0.2ml SDS (10%), 12.2ml water, 100μl Ammonium persulfate (10%), and 5μl of TEMED. This was dispensed into the chamber as before. Mix was added until it filled the chamber to almost overflowing. To place the comb into the stacking gel, it was placed at an angle on one side of the chamber and slid into the stacking gel. This was done to prevent bubbles from getting into the spaces where the gel wells were forming. The gel was then allowed to polymerize for an hour, with the same pasteur test set up as before. After the gel was polymerized, the comb was removed and the wells washed with Tank buffer (0.025 Tris pH 8.3, 0.192M Glycine, 0.1% SDS). The wells were then filled with tank buffer, for sample loading.

Samples were loaded and run in the following manner. Equal amounts of protein sample and 2X Treatment buffer (0.125M Tris-Cl (pH6.8), 4% SDS, 20% glycerol, 10% 2-mercaptoethanol) were mixed and placed into boiling water for 90 sec. Samples were then placed on ice until loading. Before loading the samples, 5μl of 0.025% Bromophenol Blue were added as a tracking dye. Samples were then loaded into the gel. After loading samples, the upper buffer chamber was attached and tightened into position with cams (cams go in with the cam handle up or edge facing down). In addition, a silicone rubber gasket with a slot cut in it was placed on the underside of the upper chamber. The upper and lower buffer chambers were filled with Tank buffer (250-350ml). To fill the upper chamber, a syringe was used to gently add buffer to the

chamber to avoid disturbing the samples in the wells. After loading the buffer chambers, the safety lid was put into place and the samples were electrophoresed at 60V for 2hr in the resolving gel. When the samples and tracking dye had all stacked at the resolving gel/separating gel interface, the voltage was increased to 120 V for about 5hr. When the tracking dye reached the bottom of the gel, the voltage was turned off and the gel assembly was disassembled. The glass plates were separated from the gel with a plastic wedge.

The gel was stained and de-stained in the following manner. After separating the glass plates, the gel was slid from the glass plate into stain (0.125% Coomassie blue R-250 (BioRad), 50% Methanol, 10% Acetic acid) and stained for 4-12 hours. De-staining was then done with several 1hr washes of de-stain buffer 1 (50% Methanol, 10% acetic acid). If de-staining was done overnight, de-stain buffer 2 (5% Methanol, 7% Acetic acid) was used to control the level of de-staining. After de-staining, a photo was taken using a DS34 Polaroid Direction Screen Instant camera, Polaroid 667 film and a yellow #8 filter (Coomasie Blue), with shutter speed set at 1/12.5 sec and F-stop set at 32. Following the picture, the gel was dried for storage.

The gel was dried in the following manner. After the picture was taken, the gel was placed into several consecutive washes of water. The gel was then left in water overnight and allowed to equilibrate. Before removing the gel to dry, a glass plate was washed as before and water was added to the plate. A piece of wetted cellophane (Bio Rad) larger than the gel dimensions was placed over the glass plate and more water

added. The gel was placed over the cellophane in the middle of the glass plate. Water was then added and a second wetted cellophane piece was applied. If bubbles were present, rolling the surface with a pipette was used to remove them. Gel spacers were then placed around the gel on the glass plate. The extra cellophane was removed and binder clamps were placed on the gel spacers. The gel was allowed to dry horizontally for two days. After the gel was dry, it was removed and taped into a notebook.

## 2.23 Staining of adult worms to localize DIG-1

### 2.23.1 Fixation and collagenase treatment

Worms were washed off a standard plate with M9 into a 1.5ml Eppendorf tube and centrifuged briefly to pellet the worms. The supernatant was then removed. The pelleted worms were placed on ice for a minimum of 15min (which slows pumping and causes the worms to stop moving). 0.5ml of ice cold fixative (4% paraformaldehyde) was added and the worms were incubated overnight at 4°C. The worms were then washed 3 times with 1X PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mM $Na_2HPO_4 \cdot 7H2O$, 1.4 mM $KH_2PO_4$). The fixed worms were incubated with 1.5mls of a solution containing 5% β-mercaptoethanol, 1% Triton X-100, and 0.1 M Tris-HCL (pH6.9) overnight with gentle mixing at 37°C. After incubation, the worms were washed 3 times in 1XPBS. The worms were then digested with 0.4 mls of collagenase (2,000 U/ml, Sigma, C-5138) for 30min to an hour with stirring at 37°C. After digestion the worms were rinsed 3 times with ice cold 1XPBS and stored at 4°C in storage buffer (1XPBS, 0.1% Triton X-100).

**2.23.2 Antibody staining of adult worms**

Worms were removed from 4°C storage and the supernatant was removed. Worms were incubated in 0.4ml of Antibody solution (1XPBS, 1% BSA, 0.5% Triton X-100, 0.05% NaN$_3$) for one hour at room temperature. After incubation, the supernatant was removed and 20µl of appropriate antibody (1:2000 anti-*lin-26* serum, 1:400 anti-Ant1Con1 or anti-Ant3Con3) was applied to the samples, which were then incubated overnight at 20°C. The control or secondary only sample was left in the antibody solution at 20°C. The samples were then washed three times with antibody wash (1XPBS, 0.1% BSA, 0.5% Trition X-100, 0.05% NaN$_3$) for 15 min each wash. Samples were then incubated in antibody solution for an hour at room temperature. The antibody solution was removed and 20µl of Goat anti-rabbit TRITC conjugate secondary antibody (1:400, Sigma, T-6778) was added. These samples were incubated at 20°C overnight. After incubation, the worms were washed 3X with antibody wash and stored at 4°C in 0.4ml of antibody solution. Worms were then mixed with an equal amount of mounting solution (20mg p-phenylenediamine dissolved in 2ml 1X PBS pH 8.0 and mixed with 18ml glycerol) and placed on a slide. The sample was then covered with a cover slip and clear nail polish was added to seal the sample.

**2.24    Antibody staining of embryos**

**2.24.1  Preparation of slides for embryos staining**

Corning microscope slides (Cat # 2947) were prepared for embryo staining in the following manner. The slides were washed for 5min in acid alcohol (1% HCl, 70%

ethanol) and rinsed three times in reagent grade water. The slides were then incubated in 1/10 diluted poly-L-lysine (diluted in reagent grade water, Sigma, Cat #P-8920) at room temperature for 5min. The coated slides were then dried in an open slide container box at 60°C for 1 hr.

### 2.24.2  Immobilizing and staining embryos

Three micro-liters of water were placed onto a, poly-L-lysine coated slide. Embryos were then picked with bacteria from a 60mm plate and pressed onto the slide where the water was placed. The water was then removed with a chem-wipe and the water allowed to air dry briefly. Prior to the eggs drying completely, but after the water was removed, 10µl of 1% paraformaldehyde was placed over the embryos. A coverslip was placed over the embryos and the excess paraformaldehyde was removed by placing a Kim-wipe along the junction between the coverslip and the coated slide. The slide was then tipped to one side to see if the coverslip would stay in place. If the coverslip moved, more paraformaldehyde was removed with a Kim-wipe. Once the coverslip stayed in one place, the slide was incubated for 30min in a humidified chamber at room temperature. Next, the slides were removed and placed in liquid Nitrogen for 1min or more. The coverslip was then removed with a razor blade, which cracks the embryos, and the slides were incubated in Methanol (pre-equilibrated at -20°C) for 5min. Embryos were re-hydrated in re-hydration buffer (1XPBS, 0.5% Triton X-100) for 30min. The slide was then blocked with blocking buffer (1XPBS, 0.5% Triton X-100, 0.1% BSA) for an hour. After blocking, the excess buffer was removed and 25µl of primary antibody (1:2000 anti-*lin-26* serum, 1:400 anti-Ant1Con1 or anti-Ant3Con3) was added to the embryos and

incubated overnight in a humidified chamber at room temperature. The secondary only slide had blocking buffer added and was incubated the same way. After incubation, slides were washed 3X for 5min each with re-hydration buffer. 20µl Goat anti-rabbit TRITC conjugate secondary antibody (1:400, Sigma, T-6778) was added and incubated as before. After incubation, slides were washed with re-hydration buffer as before. Mounting solution as described before was added and a coverslip placed over the embryos. The slides were sealed as before with nail polish and viewed under the microscope using FITC filters at 40X (procedure from Michel Labouesse).

## 2.25    Western analysis of worm protein

Total worm protein was isolated from N2 and *dig-1* (*n1480*) strains. Worms from 20 non-starved half enriched 100mm plates were washed off the plates in M9 and pelleted. Pelleted worms were then boiled for 5 min in 0.6ml protein extraction buffer (0.125M Tris-Cl (pH 6.8), 4% SDS, 10% 2-mercaptoethanol, 0.004µl/worm, approximately 7,500 worms per plate). The homogenate was then centrifuged at top speed at 4°C for 15 min. The concentration of the protein homogenate (supernatant) was determined by the Bio-Rad Protein Assay. 20µg of total worm homogenate protein as well as 0.6µg of recombinant protein were separated by SDS-PAGE as described (**section 2.22**) and blotted using a Mini Trans-Blot apparatus (Bio-Rad) to Immobilon-P Transfer Membrane (Millipore, transfer buffer (0.02 Tris base, 0.15 M Glycine, 20% Methanol, 0.025% SDS), transfer done overnight at 30V). To determine whether transfer occurred, the gels were stained after the transfer as described (**section 2.22**). The membrane was then blocked overnight at 4°C in blocking buffer (0.025 Tris-Cl, pH 7.5, 0.15 M NaCl,

1% BSA fraction V, 2.0mM EDTA, and 0.02% NaN$_3$ ). Blots were washed 2X with wash buffer (1X PBS, 0.1% Tween-20) for 5 min at room temperature. Primary antibody (1:400 final in 1X PBS, 1% BSA, 0.1% Tween-20) was then added to the blocking buffer and incubated at 4°C overnight. The blots were then washed as before. Blots were incubated with anti-rabbit secondary antibody (1:1000, Bio Rad Opti-4CN Detection Kit, Goat anti-Rabbit, Cat #170-8235) conjugated to horseradish peroxidase (HRP) and detected with 4CN substrate (Bio Rad).

Dot blots were done in the following manner. The membrane was pre-wetted in transfer buffer, same as before. The membrane was then placed onto the Bio Rad dot blot apparatus. One to 1.5μg of a recombinant protein (~20μl) was added to a well on the Bio Rad dot blot apparatus and was drawn onto the membrane with vacuum. Recombinant protein dots were detected according to the Bio Rad Opti-4CN Detection Kit with the secondary antibody dilution at 1:5000.

## 2.26 Computer analysis of the predicted protein

To determine which domains are present in the predicted DIG-1 protein the entire length of the predicted Genefinder protein and regional amino acid sequences were subjected to BLASTP (Sanger Center, NCBI and SMART). The Sanger Center web site also has a database of **p**rotein **fam**ilies (Pfam) alignments and hidden Markov Models (HMMs 7.3) that contain multiple sequence alignments and protein domains (Bateman et al., 2002). Under the Swissprot accession number Q09165 a detailed report on the hypothetical protein for *dig-1* is displayed.

## 3.0  Results

The overall goal of the project was to identify a candidate gene for *dig-1*, determine where mutations were within the gene, and analyze expression and function of the DIG-1 protein.  In order to accomplish this we first examined the phenotype.  Many aspects of the *dig-1* mutant phenotype are consistent with a defect in adhesion. Several aspects of the *dig-1* phenotype were measured, including IL2 sensory process defects and reduction in brood size.  A *dig-1* candidate gene with sequence similarity to known CAM genes was located by genetic mapping.  By Southern analysis, this candidate, K07E12, was shown to have changes in two alleles of *dig-1*, *nu336* and *n1480* (Ryder, unpub. results).  In order to confirm that the candidate CAM gene was *dig-1*, several techniques were used, including sequencing of short PCR products from alleles, rescue of mutant defects in transgenic animals, RNA interference, and RT-PCR from mutants to look for nonsense mediated decay.   To determine if *dig-1* was expressed at the RNA level, RT-PCR, Northern analysis (Higgins, **Appendix A**), and GFP expression were used. To study DIG-1 localization was in worms, recombinant proteins were produced and injected into rabbits to generate antibodies.

## 3.1 Severity of the *dig-1* phenotype

## 3.1.1 IL2 neuronal process severity

To determine the severity of the phenotype for the various alleles, three aspects of the *dig-1* phenotype, IL2 process morphology, number of adult progeny, and gonad displacement (Ryder unpub. results), were examined.   IL2 neurons of wildtype animals (N2) and several *dig-1* alleles were stained with DiO and the sensory processes were

scored (**Table 1**).  A wildtype process followed a smooth path to the nose, while mutants

had various levels of divergence from that path.  N2 had 100% of the processes scored as

wildtype.   Allele *nu336* had the highest percentage, 89%, of processes that were severely

aberrant.  Reference allele *n1321* was the next most severe, with 59% of its processes

being severely aberrant.   Allele *n1480* was the least affected, with only 3% of its

processes being severely aberrant.  The other mutants had 18% to 37% of their processes

severely affected and 34% to 58% moderately affected.

| *dig-1* alleles | SEVERE | MODERATE | SLIGHT | WILDTYPE | N |
|---|---|---|---|---|---|
| N2 at 20°C | 0 | 0 | 0 | 143 (100%) | 143 |
| N2 at 25°C | 0 | 0 | 0 | 50 (100%) | 50 |
| *n1480* | 4 (3%) | 10 (26%) | 22 (56%) | 3 (8%) | 39 |
| *nu319ts** | 15 (18%) | 28 (35%) | 16 (20%) | 21 (26%) | 80 |
| *n2467* | 15 (27%) | 26 (47%) | 11 (20%) | 3 (5%) | 55 |
| *nu345* | 13 (36%) | 21 (58%) | 2 (5%) | 0 | 36 |
| *nu52* | 20 (37%) | 20 (37%) | 9 (17%) | 5 (9%) | 54 |
| *n1321* | 17 (59%) | 10 (34%) | 2 (7%) | 0 | 29 |
| *nu336* | 31 (88%) | 4 (11%) | 0 | 0 | 35 |

**Table 1. DiO stain of N2 and several *dig-1* alleles the severity of the neuronal
processes defect.**  A wild type process ran initially straight anterior from the cell body,
curved around the pharynx, and then continued straight to the nose with no part of the
process going out of the plane of focus.  Mutants were scored by the degree to which
their processes diverged from the wild type.  A slight score was a process that had a slight
"wiggle" to the process, but was not out of the plane of focus, with the process still
following a straight path to the nose.   A moderate score was given to processes with
more "wiggle" with some of the process out of the plane of focus.  A severe score was
given to a process that was not straight and was out of the plane of focus for an extended
length. N, total number of processes scored. * *nu319ts* is a temperature sensitive allele

### 3.1.2   Total number of adult progeny

A second aspect of the *dig-1* phenotype studied was the number of adult progeny

per hermaphrodite parent (**Figure 8**).  Both allele *nu336* and allele *n1321* had

significantly fewer progeny per parent then wild type (p<0.001).  Allele *n1480* was not

significantly different than wild type (ANOVA, Bonferroni post-hoc test).

**Average adult worms**
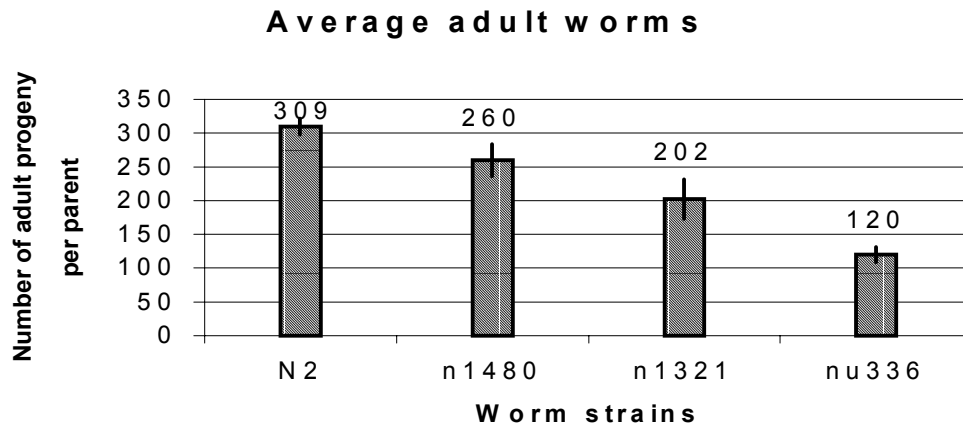


**Figure 8. Average number of adult progeny from several *dig-1* alleles.**  The total progeny from ten single worms over a ten-day period from different strains were counted (Methods). Average numbers per strain are shown at the top of each bar.  The standard deviation for each is N2 (37.4), *n1480* (67.0), *n1321* (71.8), and *nu336* (34.9).  The standard error of the mean for each is N2 (11.8), *n1480* (23.6), *n1321* (25.4) and *nu336* (11.0)

In summary, the neuronal process defect follows the same trend as the number of

adult progeny, indicating that *nu336* is the most severe allele of *dig-1* for these two

aspects of the phenotype.  By contrast, analysis of the gonad displacement showed that

gonads of *nu336* animals were not displaced anteriorly like the other alleles, but had

placement similar to wild type (Ryder unpub. results, **Appendix A**).

## 3.2 Confirming that the candidate CAM gene is *dig-1*

## 3.2.1 Characterization of *dig-1(nu33)*

The candidate CAM gene was suspected to be *dig-1 a priori* because of DNA restriction fragment length polymorphisms (RFLP) detected in alleles *nu336* and *n1480* by Southern analysis. Allele *nu336* had been isolated by EMS mutagenesis (Burket et al., submitted). The mutagenized strain contained a transgenic, nonintegrated GFP construct designed to be used as a marker for IL1 sensory neuron morphology. At the time of mutagenesis an integration of the GFP array occurred near the candidate CAM gene, which initially was thought to have caused the mutation. The EcoRI polymorphism in the candidate CAM gene was originally attributed to this GFP insertion site. Long range PCR (greater than 2 KB) was used to determine if the inserted GFP was in the candidate CAM gene. Additionally, long range PCR and PCR (2Kb and below) products were subjected to digestion to define the placement of the EcoRI polymorphism. There were no large changes in long range PCR products seen between wildtype and *nu336* (data not shown), indicating that the insertion was not in the candidate CAM gene. In addition, long range PCR (primer sets 12T, 16B; 14T, 16B; 21T, 25B; 23T, 27B; 25T, 32B; 30T, 40B) and digestion analysis showed that the restriction fragment length polymorphism (RFLP) was between primer 21T and 27B (data not shown). Further PCR and digestion showed the RFLP to be positioned between primer 25T and 27B (**Appendix B**, Table 4), within the candidate CAM gene (data not shown). When PCR products from primer set 25T and 27b were subjected to digestion, the additional EcoRI digestion site produced a 1.5 Kb fragment and a 0.5 Kb fragment.

To eliminate the GFP insertion site as a candidate *dig-1* mutation, recombinant strains were isolated from *dig-1*(*nu336*)/*dpy-17*(*e164*)*unc-32*(*e189*) worms to separate the RFLP from the GFP insertion (**Figure 9**). Unc non-Dpy or Dpy non-Unc recombinant strains were isolated and screened for the presence or absence of the RFLP by PCR (25T, 27B) and digestion with EcoRI (**Table 2**, **Figure 10**). Strains containing the RFLP and no GFP expression as well as strains without the RFLP that expressed GFP were allowed to homozygose. Worm strains that were determined to be homozygous for the RFLP were screened a second time for the presence of the RFLP with PCR (25T, 27B) and digestion to confirm the presence of the RFLP. These strains were then stained with DiO to observe IL2 neuron sensory processes (data not shown, Materials and Methods). The homozygous RFLP worms had the mutant phenotype for the IL2 sensory processes. The worm strains that were homozygous for the GFP insert did not show the mutant sensory process defect after staining with DiO (data not shown, Materials and Methods). It was therefore concluded that the GFP insert could not be the cause of the *dig-1* phenotype, and conversely the RFLP was closely linked to *dig-1*.
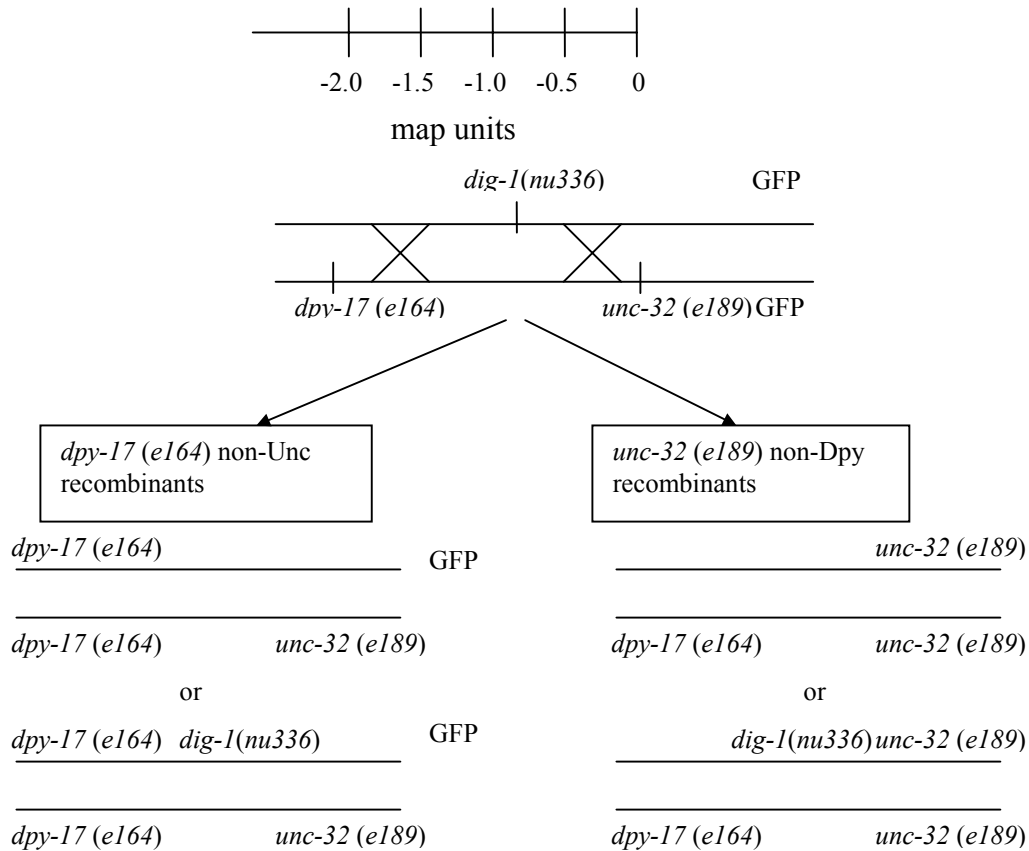
**Figure 9. Schematic of genotypes expected for recombinant worms.** Top part of the panel shows potential crossover events occurring between flanking markers *dpy-17* and *unc -32*. Below the map are shown recombinant genotypes isolated that are Dpy non-Unc or Unc non-Dpy worms, with and without the RFLP in allele *nu336* present. All Dpy non-Unc recombinants isolated had GFP expression, while no Unc non-Dpy recombinants expressed GFP, suggesting that GFP is to the right of *unc-32*. Chromosome 3 genetic map positions are -2.21 *dpy-17*, -0.93 *dig-1* and 0.0 for *unc-32*.

Dpy non-Unc recombinants

Unc non-Dpy recombinants

*dpy-17*  *dig-1*    GFP   *dpy-17*        GFP          *dig-1*   *unc-32*                *unc-32*
_____   _____              _____   _____
*dpy-17*    *unc-32*   *dpy-17*      *unc-32*           *dpy-17*      *unc-32*   *dpy-17*     *unc-32*
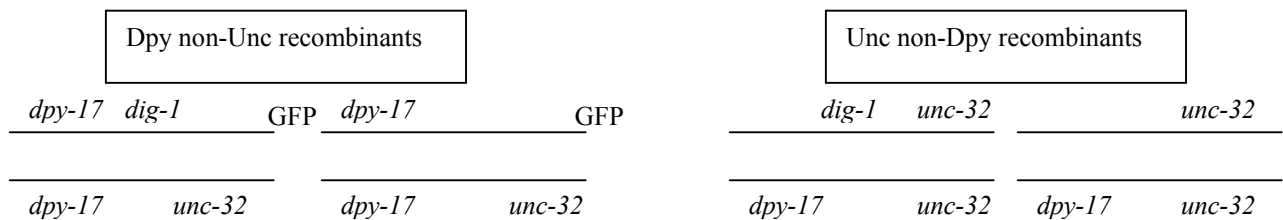
| Recombinant Phenotype | Positive for GFP expression/total | Positive for the presence of the RFLP/total |
|---|---|---|
| Dpy non-Unc | 9/9 | 4/9 |
| Unc non-Dpy | 0/16 | 3/16 |

**Table 2. Heterozygous worm lines that were screened for the presence of the RFLP in *dig-1*(*nu336*).** The upper section shows the possible genotypes. The bottom section is a summary of the recombinants found.
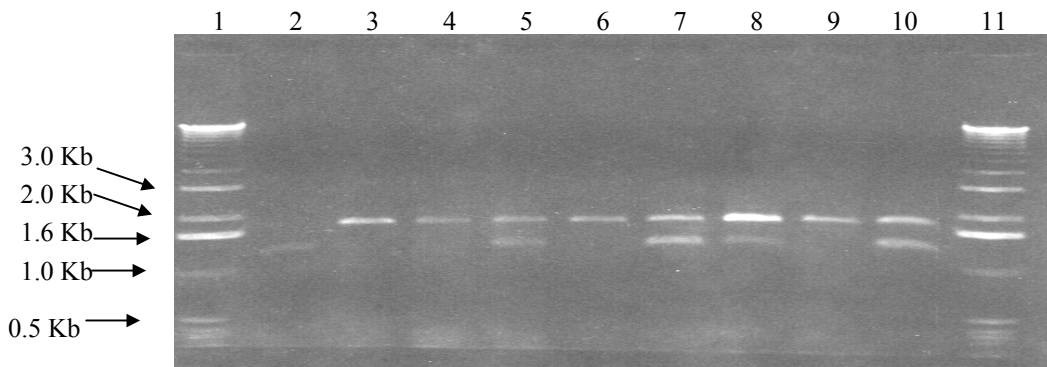


**Figure 10. Digestion of PCR product from heterozygous worms to show the presence or absence of allele *nu336*.** Lane 1 and 11, molecular weight marker. Lanes 2-5, digest of PCR products from *unc-32* (*e189*) recombinants, lanes 2 and 5 showing RFLP present (*dig-1* (*nu336*) *unc-32* (*e189*)/*dpy-17* (*e164*) *unc-32* (*e189*)). Lanes 3 and 4, no RFLP present after digestion (*unc-32* (*e189*)/*dpy-17* (*e164*) *unc-32* (*e189*)). Lanes 6-10 digest of PCR products from *dpy-17* (*e164*) recombinants, lanes 7, 8 and 10 showing RFLP present (*dpy-17* (*e164*) *dig-1* (*nu336*)/*dpy-17* (*e164*) *unc-32* (*e189*)). Lanes 6 and 9, no RFLP present after digestion (*dpy-17* (*e164*)/*dpy-17* (*e164*) *unc-32* (*e189*)).

66

### 3.2.2 Sequencing alleles of *dig-1*

Due to its size, the entire candidate CAM gene was not sequenced for each *dig-1* allele. Instead, short PCR products were amplified and sequenced from the two alleles with known RFLPs, *nu336* and *n1480*. The mutation found in allele *nu336* was a point mutation in exon 19, a transition from G*C to A*T near the 5' end of the gene (**Figure 11**). The type of mutation is consistent with EMS mutagenesis (Anderson, 1995). This mutation results in a Ser to Phe change in amino acid sequence at position 3943 of the predicted protein (**Figure 14**). Control DNA amplified from *unc-32* (*e189*) worms was wildtype at this position. This was the appropriate control, since the mutation was in an *unc-32* (*e189*) background. Also, the *unc-32* (*e189*) mutation by itself does not have the neuronal process defect that is present in *dig-1* animals. Allele *nu336* was subsequently crossed away from *unc-32*(*e189*); after 8-backcrosses to wild type, *nu336* still showed aberrant sensory processes (Long, 2001) with the presence of the EcoRI RFLP.

dig-1( *nu336*)

5'                                                                    3'

25T

exon 18          exon 19

27B

Control sequence: *unc-32*                    mutant sequence: *nu336*

5'-ACT GAC *GAA TCC* GGC ACT-3'   5'-ACT GAC *GAA TTC* GGC ACT-3'
5'-TGA CTG *CTT AGG* CCG TGA-3'   5'-TGA CTG *CTT AAG* CCG TGA-3'
    Thr  Asp  Glu  **Ser**  Gly  Thr            Thr  Asp  Glu  **Phe**  Gly  Thr
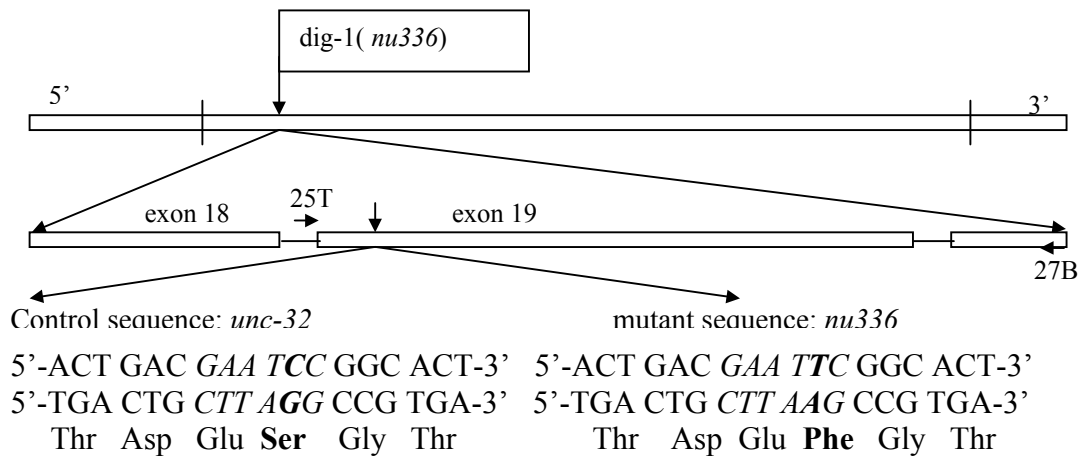
**Figure 11. Schematic representation of *nu336* mutation.** The top section depicts the gene for *dig-1*, showing the approximate position of the DNA lesion in allele *nu336* and the three regions of the gene (**section 3.3**) separated by thin vertical lines. The middle section shows an expanded view of the predicted exons (18, 19 labeled) where PCR products were generated for sequence analysis and the point mutation (downward arrow). The bottom section shows the sequence around the mutation from the control (*unc-32*) PCR amplified product and sequence from the mutant (*nu336*) PCR amplified product. The bold bp is the one that is changed in *nu336*. The italic bps represent the new EcoRI site in *nu336*. Amino acid sequence is also presented with the changed aa in bold (Ser to Phe) below the nucleotide sequence.

The second allele, *n1480*, was shown to have a large rearrangement nearer the 3' end of the gene. The points where foreign DNA interrupts the wild type sequence were determined (Hubbard, 2000; Higgins unpub. results) as well as a partial sequence extending from the insertion breakpoints into the inserted DNA. In this mutant DNA, 930 bp of DNA from predicted exon 34 were replaced with a smaller insert from Chromosome 5 (cloned cosmid K11C4, **Figure 12**). The 5' most insertion point was also confirmed by sequencing cDNA amplified product, which was generated between a primer inside the mutation and one outside the mutation on exon 33 (primer 46.9t to K11c4.25b, **Figure 12**). The cDNA product showed a stop codon in frame with the ORF, 12bp into the insertion (**Figure 12**). The splicing between exon 33 and 34 was as

predicted. Other alleles of *dig-1* were screened by Southern analysis but no additional

DNA lesions were found (Higgins, unpub results).
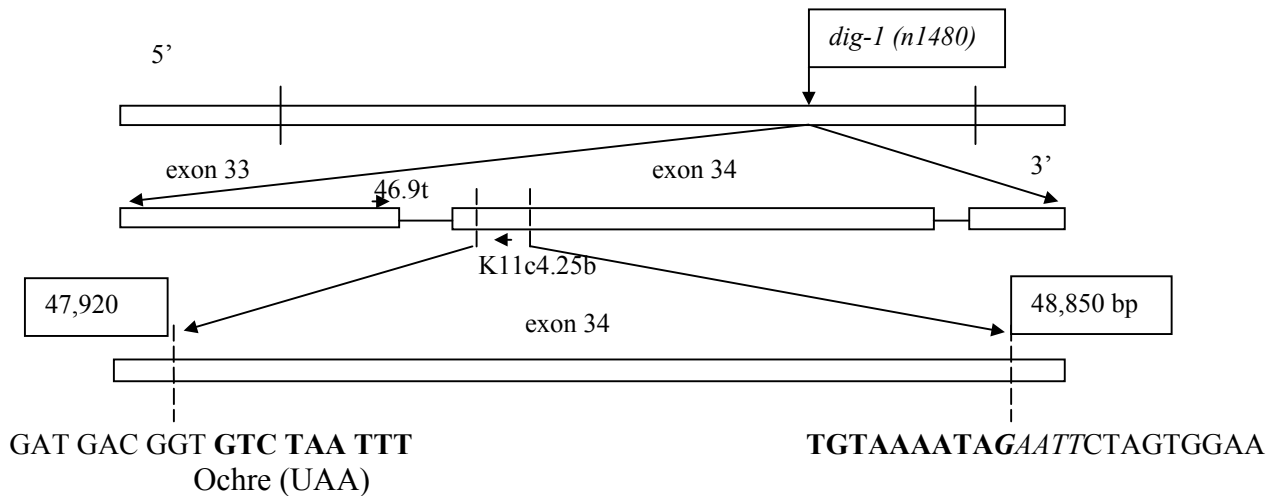


**Figure 12. A schematic representation of allele *n1480*.** The top section of this figure is the gene for *dig-1*, showing the approximate position of the DNA lesion in allele *n1480* and the three regions (**section 3.3,** thin vertical lines). The middle section represents the exons (33, 34 labeled) in the area of mutation (dashed vertical lines). RT-PCR primers are labeled (horizontal arrows). The bottom section is a closer look at the insert points (dashed vertical lines) in exon 34 with bp position indicated. Also shown is the sequence at the insert points; non-bold is the same as wildtype, bold is foreign DNA from chromosome 5. Italic bp show where an additional EcoRI site was generated from the insertion.

### 3.2.3 Rescue of the *dig-1* mutant phenotype

To determine if a wild type form of the candidate CAM could rescue (restore to wildtype) the defects in *dig-1* mutant animals, cosmids from the *C. elegans* sequencing project (*C. elegans* Sequencing Consortium, 1998) containing the DNA encoding the candidate CAM were injected into *dig-1(n1321)* worms (R. Proenca, John Hopkins University, personal communication). The sequence encoding the candidate CAM starts in cosmid K07E12 and ends in cosmid R05H11 (**Figure 13**). Both of these cosmids were co-injected into *n1321* worms, using a dominant allele of *rol-6,* which confers an obvious

Roller phenotype, as a co-injection marker. Progeny of injected animals were rescued for a *dig-1* defect involving the pharynx (R. Proenca, John Hopkins University,  personal communication).


To determine if rescued lines of worms were also rescued for the neuron process defect, worms were stained with DiO and the processes were determined to be either mutant or wildtype (**Table 3**). For N2, of the processes that were visible, 100% of the labeled processes were scored as wildtype. For the rescued line, 94% of the labeled processes were scored as wildtype.  The remaining mutant processes were most likely due to mosaicism of the rescuing array within animals; that is, individual cells may have lost the array.  Some worms from the rescued lines lost the rescuing array completely, and thus were no longer rescued for the *dig-1* phenotype these worms were identified by their loss of the Roller phenotype (**Table 2**, *n1321* no roller).  These animals had numbers of mutant processes similar to those exhibited by *n1321*, showing that the mutation is still present in the rescued strain (**Table 2**).

| Strains | worms scored | total processes | visible processes | wildtype | Mutant |
|---|---|---|---|---|---|
| N2 | 49 | 294 | 86 | 86 (100%) | 0 |
| *n1321* | 32 | 192 | 57 | 3 (5%) | 54 (95%) |
| *n1321*::K07E12;R05H11 | 64 | 304 | 80 | 75 (94%) | 5 (6%) |
| *n1321* (no roller) | 46 | 276 | 53 | 6 (11%) | 47 (89%) |

**Table 3. Rescue of the *dig-1* nervous system phenotype.**  Neuronal processes were scored as either wildtype or mutant after staining with DiO as described in Materials and Methods.   The total processes were the number of processes that were present in the worm (6 IL2 neurons total per worm).  Neurons were scored as described (**Table 1**) with all mutant processes being "lumped" into one score as mutant, regardless of severity. Visible processes were those that took up DiO.

In summary, the DNA sequence changes in *dig-1* alleles and the rescue of the

mutant phenotype by cosmids containing the candidate CAM, taken together, strongly

suggest that the candidate CAM gene is indeed the gene for *dig-1*.


### 3.3 *dig-1* DNA and protein structure

The gene for *dig-1* is predicted to be a large gene (Genefinder) with two splice

forms, K07E12.1a and K07E12.1b (**Figure 13**). K07E12.1a is predicted to have 44

exons, while K07E12.1b has only 36 exons. K07E12.1b leaves out 1a exons 36 to 43,

and uses a different open reading frame for exon 44 at the 3' end.
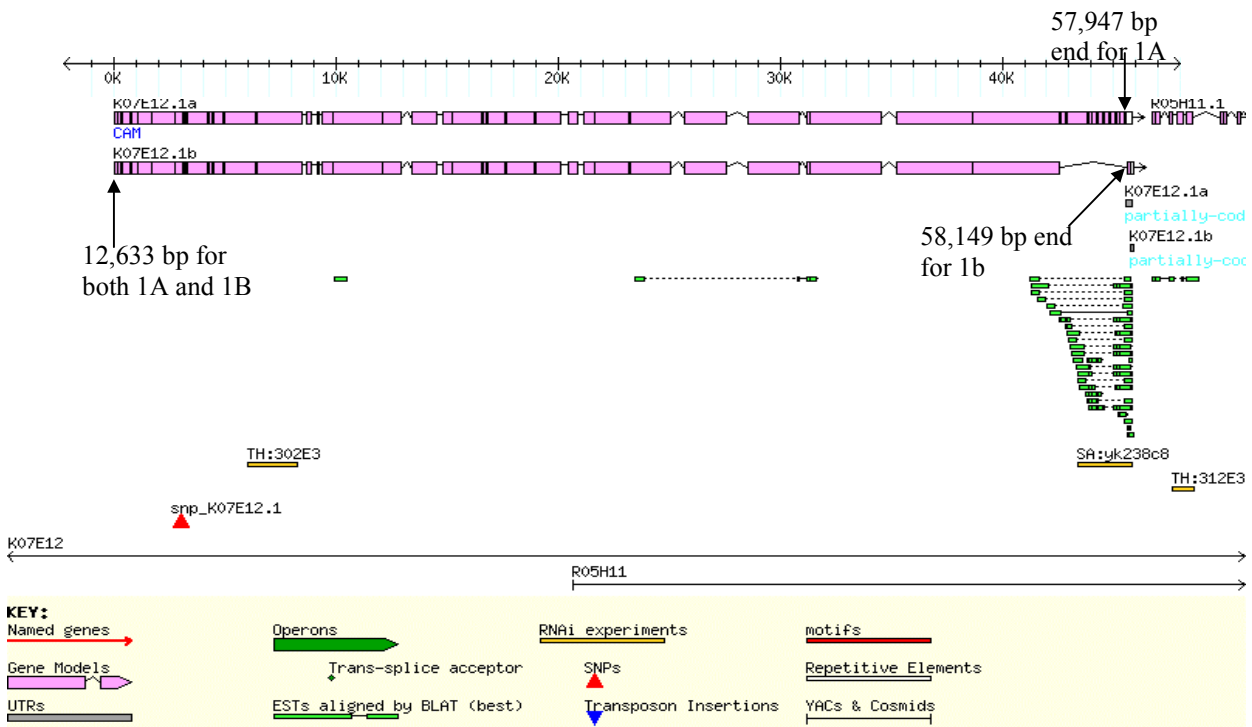


**Figure 13. The exon and intron structure of the gene for *dig-1*.** The structure of the two potential splice forms is shown with exons in purple boxes and introns as lines. The numbers for the position of the gene in the cosmid are shown. Primer numbers used in PCR and RT-PCR refer to the position on the cosmid. Expressed sequence tags (ESTs) are shown as green boxes with dots. The overlap of the cosmids is also shown. The actual K07E12 cosmid ends more 5' than shown (Stein et al., 2001).

The predicted protein encoded by the *dig-1* transcript 1a is 13,055aa in length (**Figure 14**). The N-terminus and C-terminus have several domains that are associated with adhesion. The middle of the protein is highly repeated. Estimated molecular weight is 1436.05 Kda. Regions 1 and 3 were defined by presence of globular domains that have been shown to mediate adhesion in other organisms. Region 2 was defined by the absence of such globular domains, and its highly repetitive nature.
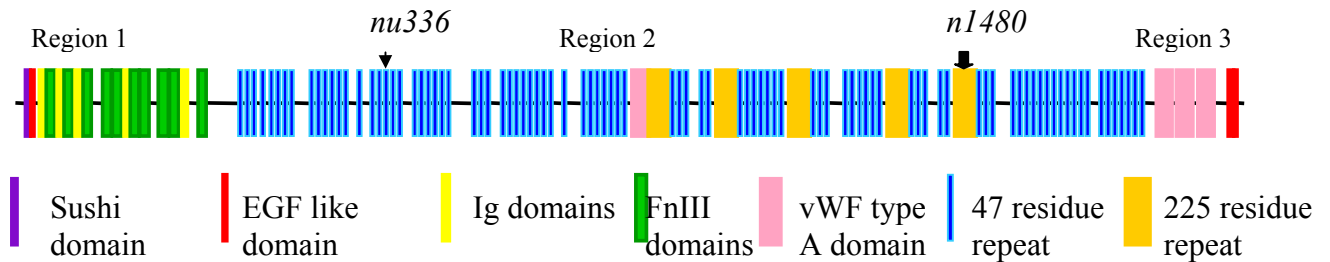


**Figure 14. Schematic of the predicted protein for *dig-1*.** The N-terminus has a complement regulatory-type domain (sushi) and several Fibronectin type III repeats (FNIII) and Immunoglobulin (Ig) domains (All domains involved in adhesion). The middle of the molecule is highly repeated, with a 47 amino acid residue repeat that is repeated and a 225 repeat that is repeated 5 times. There are also two RGD (Arg-Gly-Asp, 1676-1678aa and 4789-4791aa) sequences, which could be used in integrin binding. The middle also has several potential N-glycosylation sites (Prosite, Pole Bio-Informatique Lyonnias). The C-terminal end has Von Willdebrand type A domains and Epidermal Growth Factor domains. Different regions of the protein are also indicated (Region 1,2,3). Appendix A (**A Protein Structure**) contains a more detailed analysis of the protein.

## 3.4 Reverse Transcriptase Polymerase Chain Reaction results (RT-PCR)

### 3.4.1 Results of RT-PCR on RNA isolated from N2 worms

To determine if this gene was expressed at the RNA level, RT-PCR was used to amplify several cDNA products along the entire length of the gene (**Figure 15**).

Amplification of cDNA products was done with N2 mixed population (MP) RNA and

primers that spanned one or more introns.
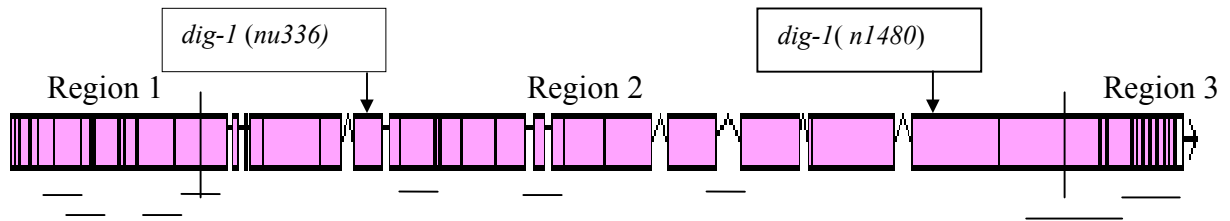


**Figure 15. Schematic representation of overall RT-PCR performed on MP RNA from N2.** The intron and exon (purple boxes) structure for *dig-1* is presented. Both sequenced mutations are labeled by arrows. The vertical lines mark the regions of *dig-1*. Horizontal lines under the gene represent some of the cDNA products amplified.

For most reactions, a band of the predicted size was amplified and was usually the

predominant product, suggesting that the prediction is correct. In addition, smaller bands

were also amplified suggesting that alternative splicing may occur. However, in several

reactions where only one intron was spanned and there were multiple bands, these bands

were likely to be non-specific amplification products from that reaction. Reactions with

multiple bands amplified, that crossed several exons, were considered more likely to be

regions that could have alternative splicing (**Figure 15**). One of these reactions (14t, 16b)

was subjected to Southern analysis using a probe that was flanked by the primer pair

(Higgins unpub. results) and several of the cDNA bands were detected, suggesting that

alternative splicing was occurring. One example of the RT-PCR on N2 MP RNA

(**Figure 16**) is primer set 56t58b that amplified a cDNA of the correct size. This primer

set overlaps with the ESTs (**Figure 13**).
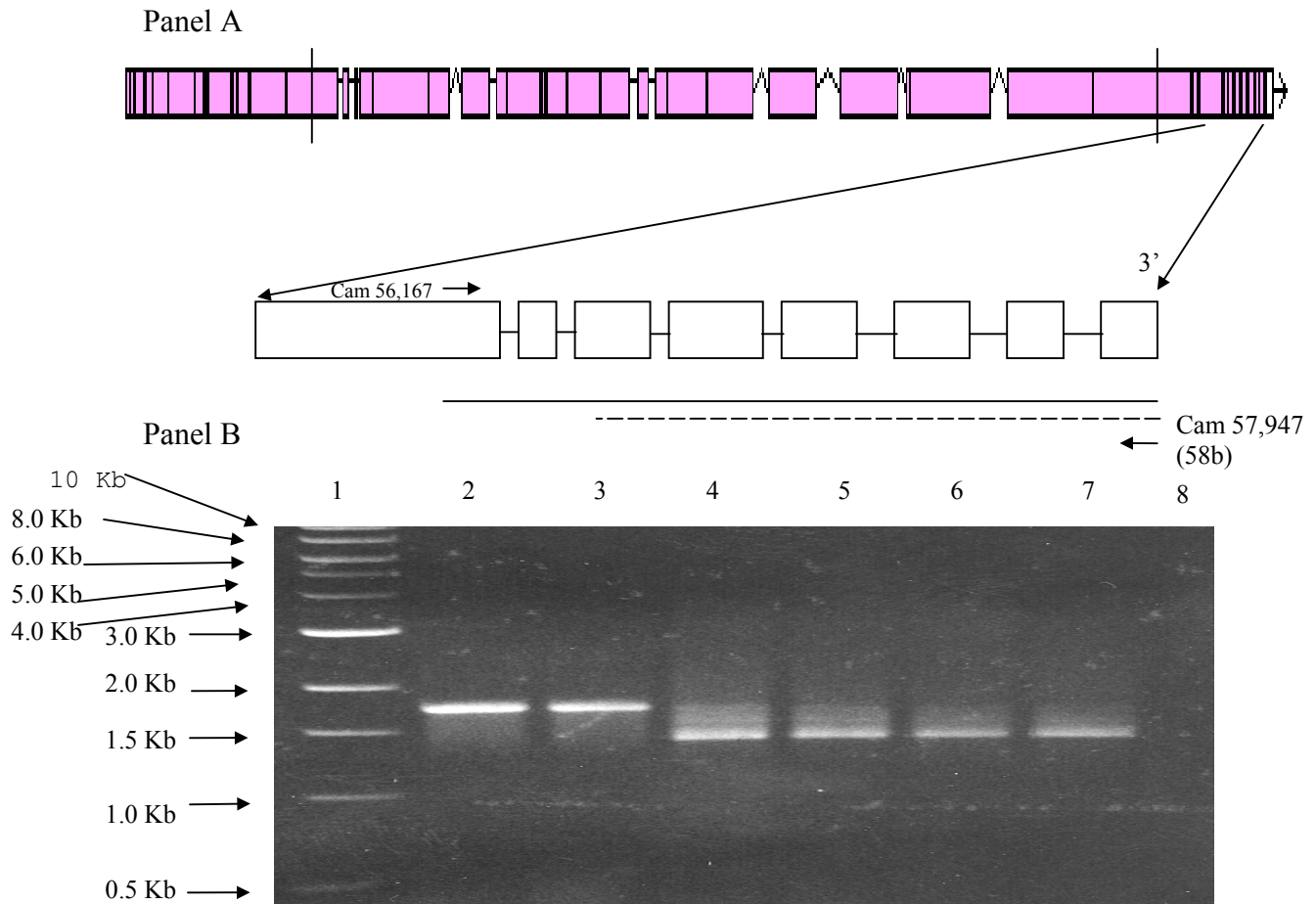
**Figure 16. Gel of RT-PCR with primer set 56t, 58b with schematic showing predicted intron/exon structure.** Panel A, the predicted intron and exon structure of *dig-1*, with a blow-up from the 3'end showing the RT-PCR from primer set 56t to 58b. Primers used to amplify cDNA are represented as arrows and labeled with base pair position. The long black line represents DNA and the dotted black line represents the length of the predicted cDNA product. Panel B, RT-PCR using N2 RNA as template with primers 56t, 58b. Lane 1, Molecular weight markers. Lane 2, PCR control (no RT). Lanes 3-8 RT-PCR with various MgSO$_4$ concentrations (ranges from 0.5 to 3.0mM in 0.5 increments).

Many other primer sets were used to amplify cDNA from all three using N2 MP RNA as template. Most produced a cDNA of appropriate size that fit with the predicted exon/intron structure of *dig-1* (**Table 4**).

| PRIMER SETS | DNA (Kb) | RNA (Kb) | FITS PREDICTED SIZE cDNA |
|---|---|---|---|
| 12T 14.06b | 1.4 | 1.17 | Yes |
| 13t 14b | 1.51 | 1.34 | NO |
| 14.04t 15b | 1.60 | 1.58 | NO |
| 14t 16b | 2.01 | 1.86 | Yes |
| *18t 20b | 1.74 | 1.69 | Yes |
| 25t 26b | 1.67 | 1.28 | Yes |
| 30tabif 32bab | 1.57 | 1.47 | Yes |
| 37t 40b | 2.67 | 2.09 | Yes |
| 37t 43b | 6.19 | 4.71 | NO |
| 37 46b | | | NO |
| #46.9t 47.7b | 0.81 | .023 | Yes |
| 46.9t | | | |
| 50t 51b | 1.13 | 1.03 | Yes |
| *54t 57b | 2.49 | 2.14 | Yes |
| 56t 58b | 1.7 | 1.4 | Yes |

**Table 4.  RT-PCR products amplified from N2 mixed population RNA** *Primer sets that span the change from adhesion to repeated area or repeated area to adhesion. #These primer sets have a second site that has 100% homology to the target site.  This 100% hit could also produce RNA of the appropriate size (Appendix B Table4).

To show that the three unique regions of *dig-1* are all one gene, RT-PCR was performed that spanned regions 1-2, and regions 2-3 (**Figure 17**).  Primer sets flanked at least one adhesive domain in region 1 or 3, and the second primer flanked the repeated region at the respective end of the gene. The cDNA products were then cloned and partially sequenced (**Appendix A**).  The sequences corresponded with the predicted exon/intron structure boundaries for both products, and were consistent with the presence of a single ORF along the entire length of the gene.
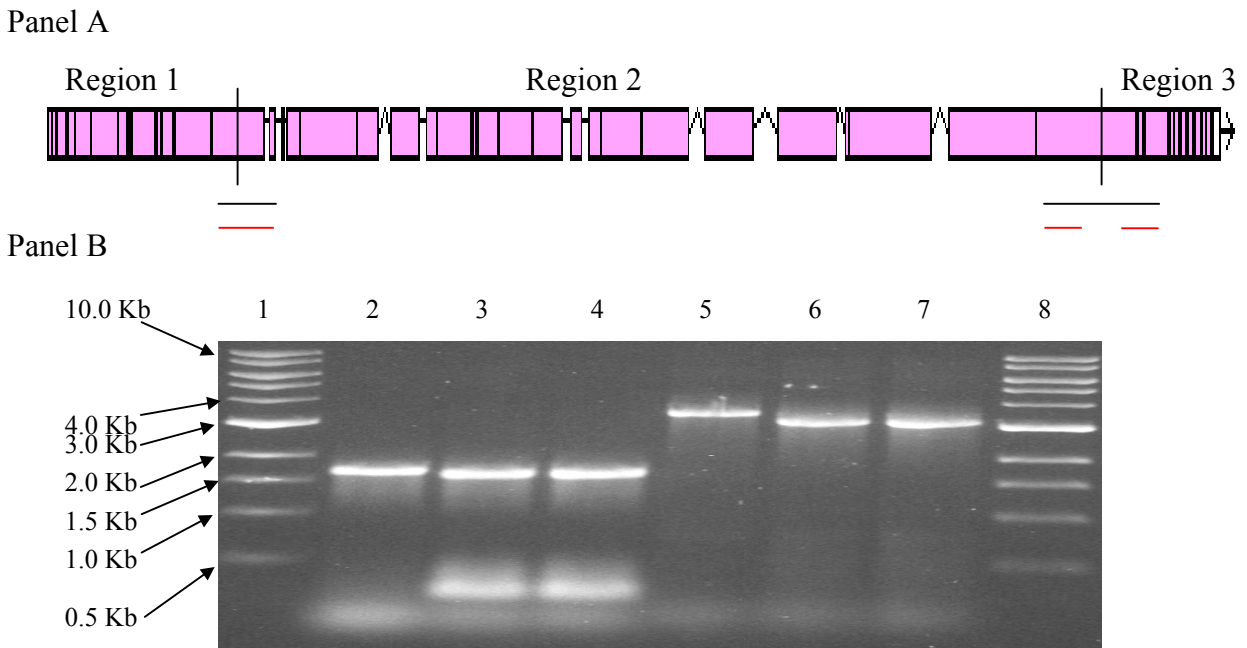
Panel A



Panel B



**Figure 17. RT-PCR amplification from the three regions of *dig-1*.** Panel A, schematic of the intron and exon structure of *dig-1*. The three regions are shown and labeled. The horizontal lines are the areas spanned by the primer sets. Panel B, lane 1,8 Molecular weight marker. Lane 2, PCR of 18t 20b (DNA 1,741bp). Lanes 3,4, RT-PCR of 18t 20b (cDNA 1,698bp); primers spanned region1 to region 2. Lane 5, PCR of 53.9t 57b (DNA 3,462bp). Lanes 6,7, RT-PCR of 53.9t 57b(cDNA 3,108bp); primers spanned region 2 to region 3. The red line under the schematic of the cDNA products shows the area sequenced for each product. Primer set 18t 20b spanned region 1 to region 2 and primer set 53.9t 57b spanned region 2 to region 3. The cDNA products of the expected size were amplified from both areas. The amplification from primer set 18t 20b was also done with Dnase treated RNA (data not shown) since the PCR product and the RT-PCR products are close in size, and an RT-PCR product was still obtained.

### 3.4.2 RT-PCR to determine the 5' end of *dig-1*

In order to confirm that the predicted 5' end was also expressed as predicted, several RT-PCR strategies were used to identify the 5' end of *dig-1* mRNA; none were successful. Genes in *C. elegans* are often trans-spliced with the SL1 leader sequence (Krause and Hirsh, 1987; Huang and Hirsh, 1989; Spieth et al., 1993). One strategy was the use of a primer made from the SL1 leader sequence to amplify the 5' end of *dig-1* (**Appendix B,** Table 4). Three different primers in exons downstream of the predicted 5'

end of *dig-1* were used in first strand synthesis. Included, as a control was the Actin genes (1-4), which has been shown to have a trans-spliced leader sequence (Krause and Hirsh, 1987). The amplification of cDNA from actin resulted in a cDNA of the appropriate size, but the PCR (no RT) lane also had a product of similar size (data not shown). The RT-PCR reactions on *dig-1* produced multiple bands and not a band of the predicted size (data not shown).

A second strategy was to use primers at various distances from the predicted 5' end to amplify cDNA product. The first strand synthesis primer used was primer 14.06b. The PCR control (no RT) resulted in inappropriate sized PCR bands (data not shown). RT-PCR resulted in multiple bands for all primer sets used (data not shown).

A third and final strategy was to determine if an open reading frame (ORF) approximately 10Kb upstream from the predicted 5' start was also included in the gene. This open reading frame, denoted K07E12.2, the second ORF on cosmid K07E12, where as the sequence for *dig-1* is the first open reading frame K07E12.1. RT-PCR was done using primers (**Appendix B**, Table 4) that are in known exons of *dig-1* and a primer specific for K07E12.2. Several of the primer sets amplified cDNA products of an appropriate size. The cDNA products were then cloned and sequenced (data not shown). The sequencing results showed the cDNA product to be non-specific for either *dig-1* (K07E12.1) or K07E12.2 (**Figure 18**).

**Figure 18. Schematic of the open reading frames from cosmid K07E12.** RT-PCR was attempted from KO7E12.1 to K07E12.2.

### 3.4.3 RT-PCR of embryonic isolated RNA

To determine if *dig-1* was expressed during embryogenesis, RT-PCR was used to amplify cDNA products from several areas of *dig-1* from aged embryo RNA (embryos aged approximately 2-3hrs post-fertilization, when the nervous system forms). cDNA products amplified from aged embryos, for the most part, were the same as those amplified from N2 MP RNA (**Figure 19**). Only the 5' end of the gene showed a difference in the banding pattern. In **figure 19**, lanes 11, 13 and 15, 17 show cDNA products from MP RNA and aged embryo RNA respectively. The primer sets were at the 3' end of the gene and the predicted cDNA was produced. Lanes 3 and 5 show cDNA amplified from the 5' end of *dig-1* from MP RNA and aged embryo RNA respectively, with Lane 5 showing multiple bands that were not present in Lane 3
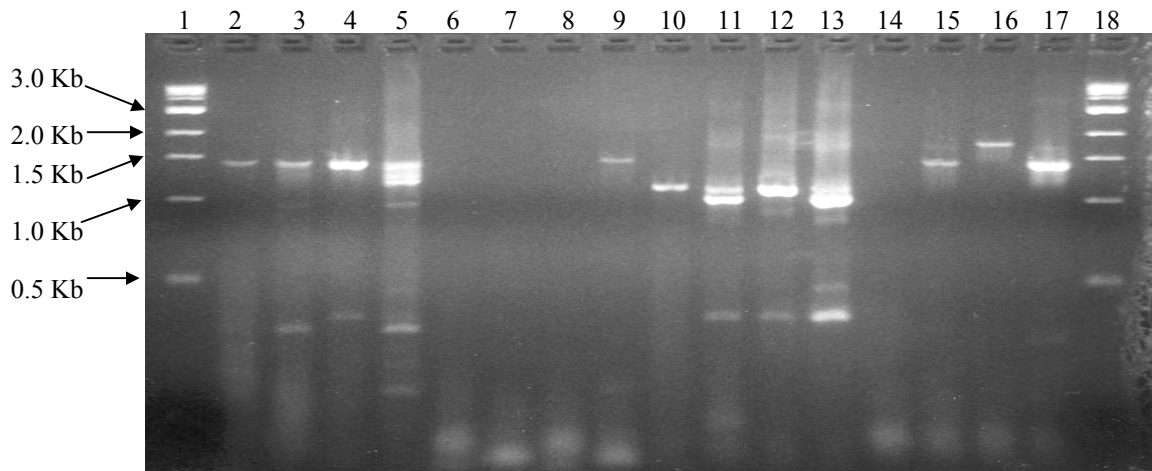
**Figure 19. Gel of RT-PCR amplified cDNA products from N2 MP RNA and aged embryo RNA.** Lanes 1 and 18 are molecular weight marker (1 kb DNA ladder). Lanes 2-5 primer set 12t, 14.06b; Lane 2, PCR N2 MP; Lane 3, RT-PCR N2 MP; Lane 4, PCR aged embryo RNA; Lane 5 RT-PCR aged embryo RNA. Lanes 6-9 primer set 30tabif, 32bab; Lane 6, PCR N2 MP; Lane 7, RT-PCR N2 MP; Lane 8, PCR aged embryo RNA; Lane 9, RT-PCR aged embryo RNA. Lanes 10-13 primer set 50t, 51b; Lane 10 PCR N2 MP; Lane 11, RT-PCR N2 RNA; Lane 12, PCR aged embryo RNA; Lane 13, RT- PCR aged embryo RNA. Lanes 14-17 primer set 56t, 58b; Lane 14, PCR N2 MP; Lane 15, RT-PCR N2 MP; Lane 16, PCR aged embryo RNA; Lane 17, RT-PCR aged embryo RNA. Aged embryo RNA was prepared by Christina Higgins

### 3.4.4  RT-PCR of RNA isolated from *dig-1* alleles

RT-PCR was also used to determine if there were any differences, such as nonsense mediated decay, between the wildtype (N2) RNA and RNA from the alleles of *dig-1*. RNA isolated from each of the alleles and N2 was subjected to RT-PCR using primers at the 5' end, middle, and 3' end of the gene. In cases where the alleles were sequenced, RT-PCR was done across the mutation and compared to N2. An example of one RT-PCR from RNA isolated from N2 and an allele of *dig-1* (allele *nu345*) is shown (**Figure 20**).
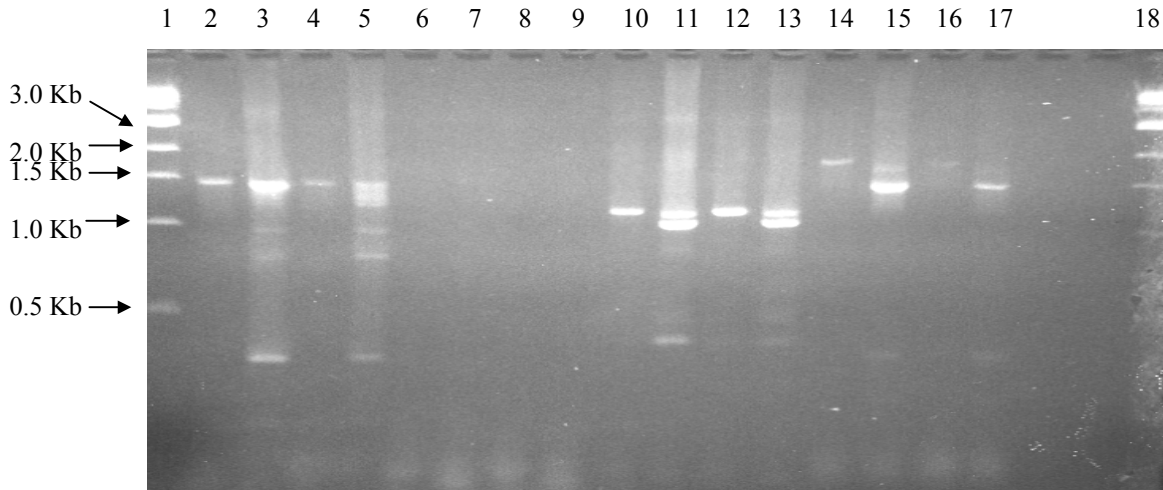
**Figure 20. Gel of cDNA products generated from RNA isolated from N2 MP and allele *nu345*.** Lane 1, 18 molecular weight marker (1 kb DNA ladder). Lanes 2-5, primer set 12t,14.06b; Lane 2, PCR N2 MP; Lane 3, RT-PCR N2 MP, Lane 4, PCR *nu345* MP; Lane 5, RT-PCR *nu345* MP. Lanes 6-9, primers 30tabif, 32bab; Lane 6, PCR N2 MP; Lane 7, RT-PCR N2 MP; Lane 8, PCR *nu345* MP; Lane 9, RT-PCR *nu345* MP. Lanes 10-13, primer sets 50t,51b; Lane 10, PCR N2 MP; Lane 11, RT-PCR N2 MP; Lane 12, PCR *nu345* MP; Lane 13, RT-PCR *nu345* MP. Lanes 14-17, primer set 56t,58b; Lane 14, PCR N2 MP; Lane 15 RT-PCR N2 MP; Lane 16 PCR *nu345* MP; Lane 17 RT-PCR *nu345* MP.

Most of the alleles screened had cDNA products that were the same as cDNA products amplified from N2 MP RNA (**Table 5**). The one exception was allele *n1480* which had no cDNA products 3' of the area of mutation (**Table 5,** primer sets 50t51b, 56t58b and **Figure 21**). The predicted size cDNA products were amplified from *n1480* in the area that is 5' of the mutation (**Figure 21** lanes 2-13). Lanes 3, 7, and 11 are cDNA products from N2 MP RNA that are comparable to lanes 5, 9, and 13, which are cDNA products from *n1480*. cDNA product was not amplified in *n1480* RNA from the area 3' of the mutation (**Figure 21** lanes 14-21). Lanes 15 and 19 show the correct size cDNA product using different primer sets from N2 MP RNA but no comparable cDNA product is amplified from the RT-PCR of *n1480* RNA (lanes 17 and 21). The other sequenced

allele (*nu336*) did not show any changes in the area of mutation or in cDNA products that

were 5' or 3' to the mutation (data not shown).

| STRAINS OF WORMS | 12T 14.06B | *25T26 B | 30TABIF32BA B | 50T51B | 56T58B |
|---|---|---|---|---|---|
| N2 | + | + | + | + | + |
| *n2467* | + | ND | $ | + | + |
| *nu52* | + | ND | $ | + | + |
| *n1480* | + | + | + | - | - |
| *nu345* | + | ND | $ | + | + |
| *nu319ts* | + | ND | + | + | + |
| *nu336* | + | + | + | ND | + |
| *n1321* | | ND | $ | + | + |

**Table 5. Summary of RT-PCR amplification from N2 MP and alleles of *dig-1*.** The alleles screened with RT-PCR and the primer sets used are shown. A + means that cDNA of the predicted size was amplified. A − means that a cDNA was not amplified. A $ means that neither the N2 nor the mutant amplified a cDNA for that particular experiment. ND means that primer set was not used on those alleles. * This primer set spans the point mutation in allele *nu336*.
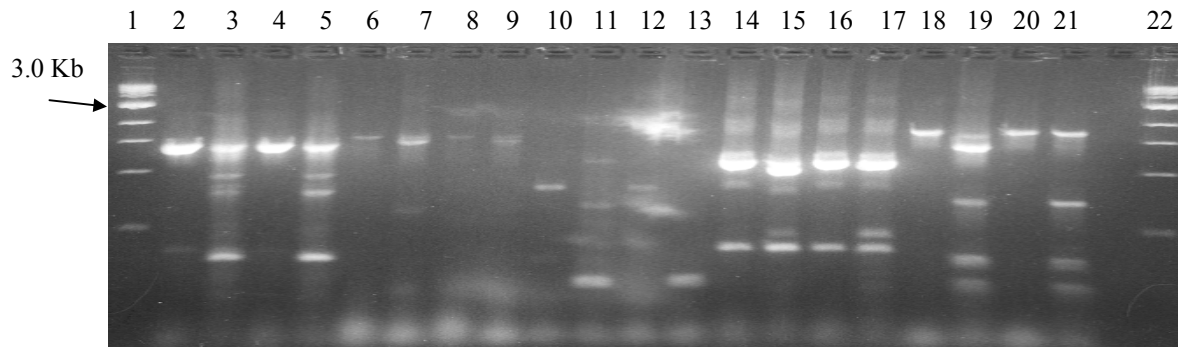


**Figure 21. RT-PCR of N2 and *n1480* with various primer sets.** Panel A, is RT-PCR across the entire gene for *dig-1*. Lanes 1, 22 molecular weight markers, as in Figure 16. Lanes 2-5 primer set 12t,14b; Lane 2, PCR N2 MP; Lane 3, RT-PCR N2 MP; Lane 4 PCR *n1480* MP; Lane 5, RT-PCR *n1480* MP. Lanes 6-9 primer set 30tabif32bab Lanes 10-13 primer set 46.9t47.7B Lanes 14-17 primer set 50t51b; Lane 14, PCR N2 MP; Lane 15, RT-PCR N2 MP; Lane 16, PCR *n1480* MP; Lane 17, RT-PCR *n1480* MP. Lanes 18-21 primer set 56t,58b; Lane 18 PCR N2 MP; Lane 19, RT-PCR N2 MP; Lane 20 PCR *n1480* MP; Lane 21 RT-PCR *n1480* MP.

Allele *n1480* was further studied with RT-PCR, to determine where and how

much of the RNA was affected (**Figures 22 and 23**). Figure **22** is a representation of the

exon/intron structure of allele *n1480* showing the location of the RT-PCR primer sets

used to determine the extent of the changes in the transcript. Figure **23** shows the results

of the RT-PCR across the region of mutation in allele *n1480*.  One RT-PCR using a

primer outside the mutation (46.9t➤ ) and one inside the mutation (K11C4.25b ◄)

produced a cDNA that was then sequenced (**Figure 23**, lane 5).  A second primer set

using a primer that crossed from the normal sequence into the mutation (for first strand

synthesis, *n1480*M ◄) and primer 46.9t failed to produce a cDNA product (data not

shown).  This suggests that the transcript stopped at a point within the mutated region

after the inframe stop codon.
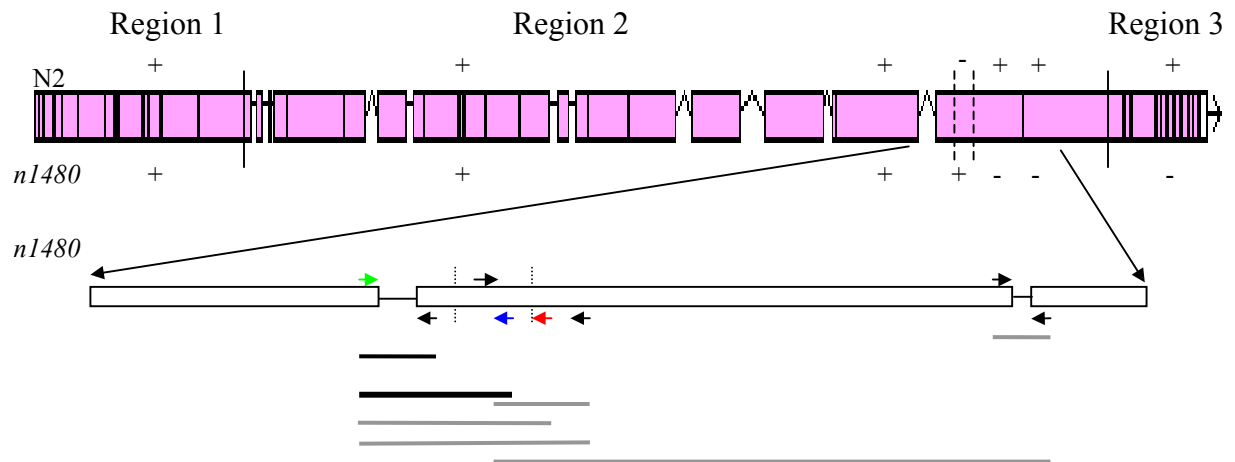


**Figure 22. Schematic representation of RT-PCR across the area of mutation
in allele *n1480*.**  The top section is a representation of where the cDNA products
differ between N2 and *n1480*, also showing the three regions.  A (+) means that an
RT-PCR product of anticipated size was produced.  A (−) means that no cDNA
product was generated. The dotted vertical lines mark the area of mutation in allele
*n1480*. The expanded area shows the RT-PCR done across the mutation in allele
*n1480*.  The black horizontal lines are RT-PCR products. The thick black line is
the sequenced cDNA amplified from *n1480* (see sequence).  The thin black line is
RT-PCR just before the mutation, which produced amplified product of correct
size in both wild type and *n1480*.  The grey bands are all RT-PCRs that failed
using *n1480* RNA; no product was amplified.  The primers are black, green, red
and blue horizontal arrows.
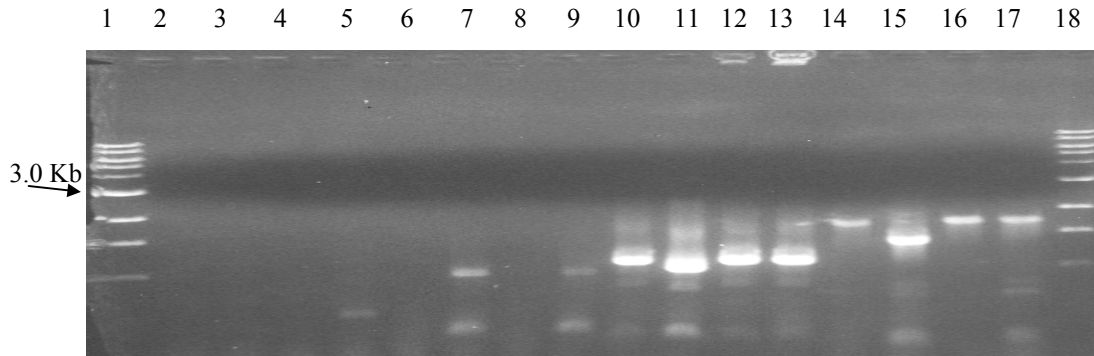
**Figure 23. RT-PCR across the area of mutation in *n1480*.** Lanes 1, 18 Molecular weight markers, as in Figure 16. Lanes 2-5 primer set 46.9t, K11C4.25b; Lane 2, PCR N2 MP; Lane 3, RT-PCR N2 MP; Lane 4, PCR *n1480* MP; Lane 5, RT-PCR *n1480* MP. Lanes 6-9 primer set K11C4.25b, 48b; Lane 6, PCR N2 MP; Lane 7, RT-PCR N2 MP; Lane 8 PCR *n1480* MP; Lane 9, RT-PCR *n1480* MP. Lanes 10-13 primer set 50t,51b; Lane 10 PCR N2 MP; Lane 11 RT-PCR N2 MP; Lane 12 PCR *n1480* MP; Lane 13 RT-PCR *n1480* MP. Lanes 14-17 primer set 556t, 58b; Lane 14 PCR N2 MP; Lane 15 RT-PCR N2 MP; Lane 16 PCR *n1480* MP; Lane 17 RT-PCR *n1480* MP.

## 3.5 Localization of DIG-1

### 3.5.1 Production of recombinant proteins as immunogens

In order to generate antibodies to DIG-1 protein, recombinant proteins were designed. cDNA products from the 5', middle and 3' end of the gene were amplified and cloned into the expression vector pET-15b (Novagen, **Figure 24**, **Appendix B**). All three DNA constructs were sequenced and were shown to be in frame (data not shown). cDNA constructs at either the 5' or 3' end (pET-15bexon3 (Construct 1) or pET-15b56,58 (Construct 3)) were chosen as immunogens, since these are in areas that are predicted to have globular domains, which should generate an immune response. Recombinant protein expression was induced in large scale bacterial cultures and protein was extracted. Recombinant proteins of appropriate size were eluted from His·tag columns (Novagen) and were named Ant1Con1 and Ant3Con3 (Ant-antigen 1 or 3 and Con-construct 1 or 3

which are from the 5' and 3' ends, respectively) (**Figure 25**).  Ant1Con1 and Ant3Con3

were then sent to a commercial source for immunization into rabbits.
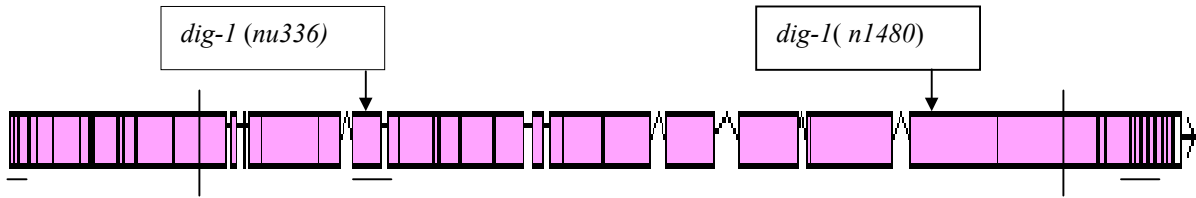


**Figure 24.  Location of cDNA products used to produce recombinant proteins.**  The intron and exon (purple boxes) structure for *dig-1* is presented.  Both sequenced mutations are labeled by arrows.  The vertical lines mark the regions of *dig-1*.   Horizontal lines under the gene represent the cDNA products amplified and cloned. cDNA products were cloned in frame into the expression vector pET-15b.
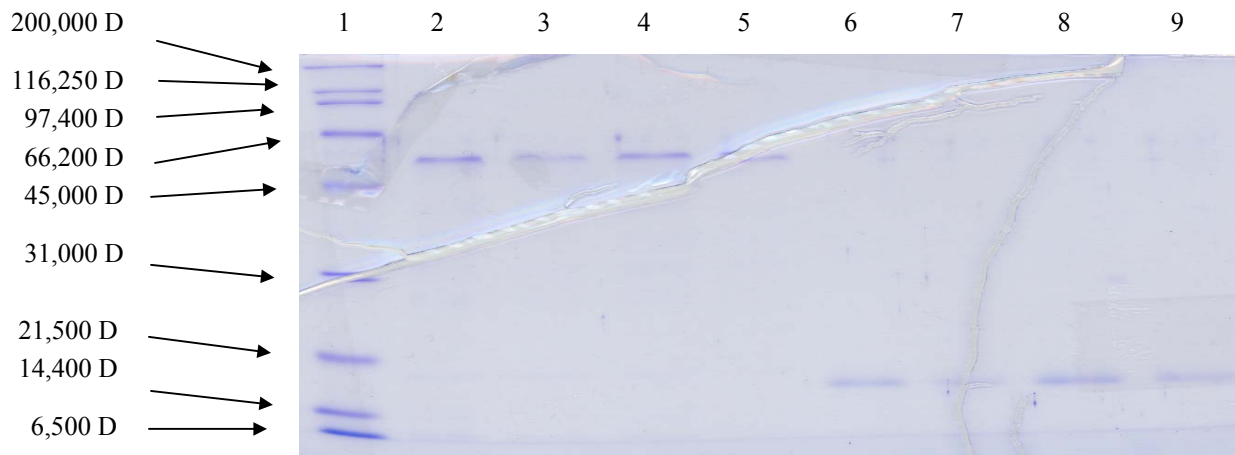


**Figure 25. SDS-PAGE analysis of His·tag purified recombinant proteins.**  Lane 1, SDS- PAGE broad range molecular weight standards (BioRad). Lanes 2-5, Purified recombinant protein from Ant3Con3 (55,759 D), from four different elutions from a His·tag column. Lanes 6-9, Purified recombinant protein from Ant1Con1 (15,755 D), from four different elutions from a His·tag column.

Several pre-immune sera were screened by immunofluorescent staining of adult

worms as well as embryos, as described in Materials and Methods.  Four rabbits (out of

8) were picked that did not have a noticeable staining of either adult worms or embryos.

These selected rabbits were injected with the recombinant protein (2 rabbits for each recombinant).  Rabbits were boosted at 3 week intervals, and production bleeds collected (Covance).

**3.5.2 Western analysis of recombinant protein and N2 homogenate**

Western analysis was used to determine if antibodies specific to DIG-1 protein were generated (**Figure 26**).  Western analysis (lanes 2-5) showed a band of correct size for Ant3Con3 (C-terminal antigen) when probing with the immune (lane 3) but not pre-immune serum (lane 5).  The analysis (lanes 6-9) also showed a band of correct size for Ant1Con1 (N-terminal antigen) when probing with the immune (lane 7) but not the pre-immune serum (lane 9).  Multiple bands in the N2 homogenate lanes (2,4,6 and 8) suggest background, but some bands are not present in the pre-immune lanes.   A dot blot of the recombinant proteins confirmed that antibodies were produced against both proteins and that the antibodies were specific to their respective immunogen (**Figure 27**).
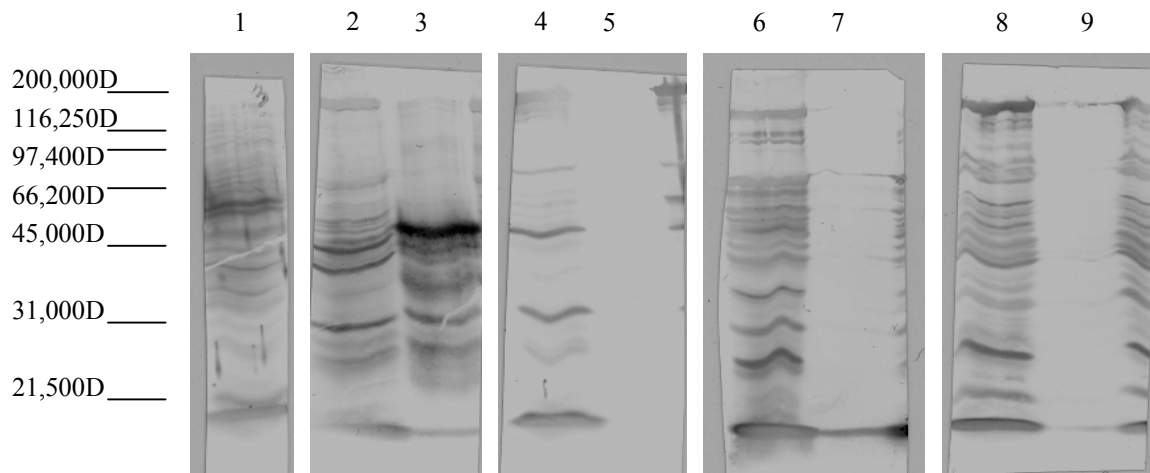
**Figure 26. Western analysis of worm homogenate and recombinant protein.**
Lane 1, Worm homogenate probed with anti-LIN-26 antibody. Lane 2,3 were
probed with anti-Ant3Con3 serum. Lanes 4,5 were probed with pre-immune serum
MA015. Lane 2, N2 protein homogenate. Lane 3, Recombinant protein
Ant3Con3. Lane 4, N2 protein homogenate. Lane 5, Recombinant protein
Ant3Con3. Lanes 6, 7 were probed with anti-Ant1Con1 serum. Lane 8,9 were
probed with pre-immune serum MA018. Lane 6, N2 protein homogenate. Lane 7,
Recombinant protein Ant1Con1. Lane 8, N2 protein homogenate. Lane 9,
Recombinant protein Ant1Con1. The extra bands seen in lane 3 are most likely
break down products of the recombinant protein.

Panel A    Panel B

Panel C    Panel D
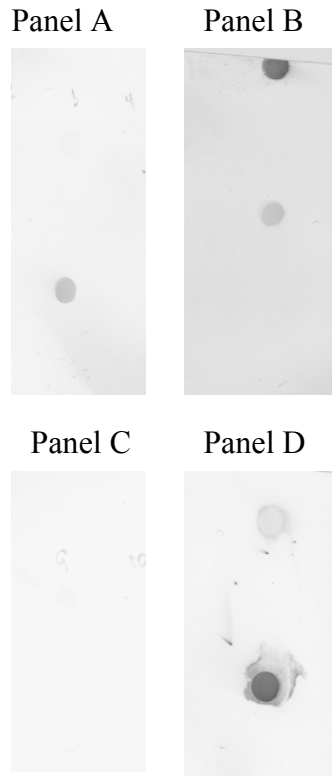
**Figure 27. Specificity of DIG-1 antibodies.** On all blots the upper dot is Ant1Con1 (1.28 μg) and the lower dot is Ant3Con3 (1.14 μg). Panel A. Dots probed with pre-immune serum for Ant1Con1. Panel B. Dots probed with anti-Ant1Con1 serum. Panel C. Dots probed with pre-immune serum for Ant3Con3. Panel D. Dots probed with anti-Ant3Con3 serum. Each antiserum detected its corresponding immunogen. There is some cross reactivity that is either non-specific or recognition of the His·tag.

If anti-Ant3Con3 is reacting specifically to DIG-1 protein, we should expect the reactivity to be gone in *n1480*, which should be missing the C-terminus end of the protein. To test this hypotheis a dot blot was done (**Figure 28**). The blot analysis showed the presence of the antibody that recognized Ant3Con3. Figure (**28**) shows the Ant3Con3 is detected by immune (Panel A) but not pre-immune serum (Panel B), nor lin-26 serum (Panel C). There was no obvious difference in binding to *n1480* and N2, indicating, as did the western analysis that there was a high level of non-specific binding.
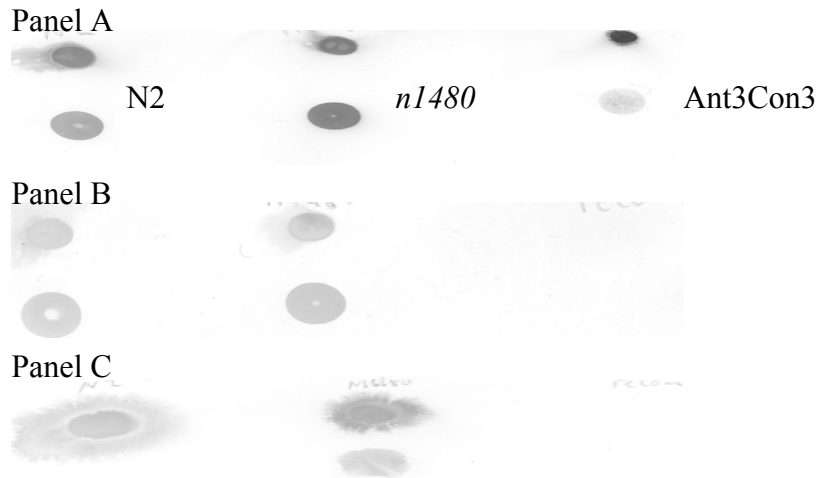
**Figure 28. Dot blot of N2, n1480 homogenates and recombinant protein Ant3Con3.** Panel A, N2 concentrate (top dot) and 1/10 diluted N2 (bottom dot). Concentrate n1480 (top dot) and 1/10 diluted (bottom dot). Concentrated Ant3Con3 (top dot) and 1/10 diluted (bottom dot). Protein detected with anti-Ant3Con3 serum. Panel B, same as in panel A, except that this blot was probed with antiLIN-26 antibody. Panel C, same as in panel A, except that this blot was probed with pre-immune serum for Ant3Con3.

To determine whether antibodies were produced that could be used to localize DIG-1, embryos were stained by indirect immunofluorescence using pre-immune and production bleed sera against each antigen. The embryos that were stained with serum from pre-immune rabbits had staining patterns that were the same as those from production bleed sera. This indicates that the rabbits that were used to generate an antibody had already been exposed to worm proteins or had cross-reacting antibodies present or that high levels of non-specific binding are occurring. In *n1480* embryos had a staining pattern similar to wild type, when stained with immune serum and pre-immune serum. Adult worms were also stained by indirect immunofluorescence for the presence of DIG-1 but did not have a noticeable staining pattern with the production bleeds.

**3.6 Worm lines isolated from micro-injection**

**3.6.1  *dig-1* promoter::GFP constructs were not expressed embryonically**

In order to localize which cells were producing *dig-1*, a reporter construct was
made with 12Kb of the 5' upstream sequence that should represent the *dig-1* promoter
cloned in frame into a GFP promoterless vector with or without a nuclear localization
sequence (pPD95.77.12 and pPD95.96.12, Methods).   In *dpy-20* worms injected with
*dpy-20 +*; pPD95.77.12 or *dpy-20+*;pPD95.96.12 produced six strains exhibiting rescue
of *dpy-20,* were all positive by 10 worm PCR for the presence of GFP.  Embryos from all
six strains were studied for expression of GFP (**Table 6**).  None of the *dig-1*
promoter::GFP containing lines expressed GFP during embryogenesis, which is when the
nervous system is formed.  The gene *dig-1* is known to be required during embryogenesis
(Ryder unpub. results, **Appendix A**)

| EMBRYO STAGE | RY0031 | RY0029 | RY0030 | RY0028 | RY0026 | RY0027 |
|---|---|---|---|---|---|---|
| Late Gastrulation | 4 | 3 | 10 | 11 | 6 | 3 |
| Comma stage | 3 | 4 | 2 | 6 | 2 | 5 |
| 3 fold | 8 | 7 | 10 | 10 | 10 | 14 |
| L1 | 1 | - | 4 | 3 | 10 | 4 |
| Rescued percentage at 25˚C | 29% | 65% | 37% | 50% | 32% | - |

**Table 6. Lack of embryo staining in transgenic *dig-1*::GFP strains.** Numbers shown
are embryos examined at each stage; all were negative for GFP. The percent of rescue by
*dpy-20*, which tells how many worms or embryos should have the *dig-1* promoter::GFP
construct, is also shown.

Adult worms from the rescued lines were also studied and determined to have
expression of GFP.  The expression was not strong in any one line and the pattern of

expression was difficult to determine.  This experiment was terminated when it was determined that the 5' end used in cloning the *dig-1* promoter was incorrect.

**3.6.2 Attempts to phenocopy *dig-1* using RNAi were unsuccessful**

To provide further evidence that the candidate CAM was the gene for *dig-1*, RNA mediated interference (RNAi) was used to attempt to phenocopy the *dig-1* phenotype. RNA was produced from a control vector pJP603 containing the *pie-1* gene (**Appendix B**, provided by the Mello Lab, Worcester MA) and injected into N2 worms.  This produced the embryonic lethal phenotype expected for that injection control (C. Mello, University of Massachusetts Medical School, personal communication).   The first attempt at producing RNA for interference of *dig-1* was unsuccessful using PCR products to produce RNA.  It was determined that the PCR products should be cloned into a suitable vector.  Before the PCR products were cloned, the lab was informed that RNAi does not work effectively in producing the expected phenotype in neuronal cells.  This experiment was terminated.

**3.6.3   Generation of strains expressing *mig-10*::GFP fusion products**

To localize the expression of MIG-10, worms were injected with constructs encoding a fusion protein between GFP and MIG-10. Because *mig-10* is relatively large, two constructs were used, *mig-10*pro::GFP and GFP::*mig-10*cod were used to generate the construct (**Figure 29**, Papenfuss 2000; Rivard 2001; Rusiecki, 1999).  If rescue occurs, with such a construct, it suggests that the GFP expression pattern reflects the location of a functional protein.  Initially, homozygous *mig-10(ct41)* worms were

injected with the constructs, but these produced very few progeny. A healthier

heterozygous line (RY0040) was constructed with the genotype *mig-10(ct41)/dpy-*

*17(e164)unc-32(e189)* (RY0040). RY0040 worms were injected and screened for the

presence of GFP expression in the intestine, where the *elt-2*::GFP co-transformation

marker is expressed (Materials and Methods). Ten lines of worms were isolated from the

injection, with two being *dpy-17unc-32* homozygotes, the other eight being heterozygous.

To isolate *mig-10* homozygous worms for further analysis, worms that had no DpyUnc

progeny were selected.


Preliminary screening using a fluorescence dissecting scope showed that GFP

expression in the head of embryos was difficult to observe, due to the brightness of the

GFP expression in the intestine. Adult worms, on the other hand, did have some easily

visible GFP expression in the head, since they are not folded over themselves, as are

embryos. The expression seen in the head of some lines could have been from the correct

recombination of the injected materials (*mig-10*pro::GFP and GFP::*mig-10*cod) or from

expression of *mig-10*pro::GFP by itself. These lines were further screened with PCR to

determine if the *mig-10*pro::GFP and GFP:: *mig-10*cod had recombined to give a

potential rescuing construct. Seven of the lines were determined to be positive for the

presence of the correct recombined construct by PCR (**Figures 29 and 30**).

Panel A

*mig-10*pro::GFP

Part of exon 1b

*mig-10* promoter

GFP

BsmI

SalI

GFP

*mig-10* coding region

Panel B

GFP::*mig-10*cod

*mig-10*promt

*mig-10* promoter

GFP

*mig-10* coding region

*mig-10*transb

1.6Kb PCR product
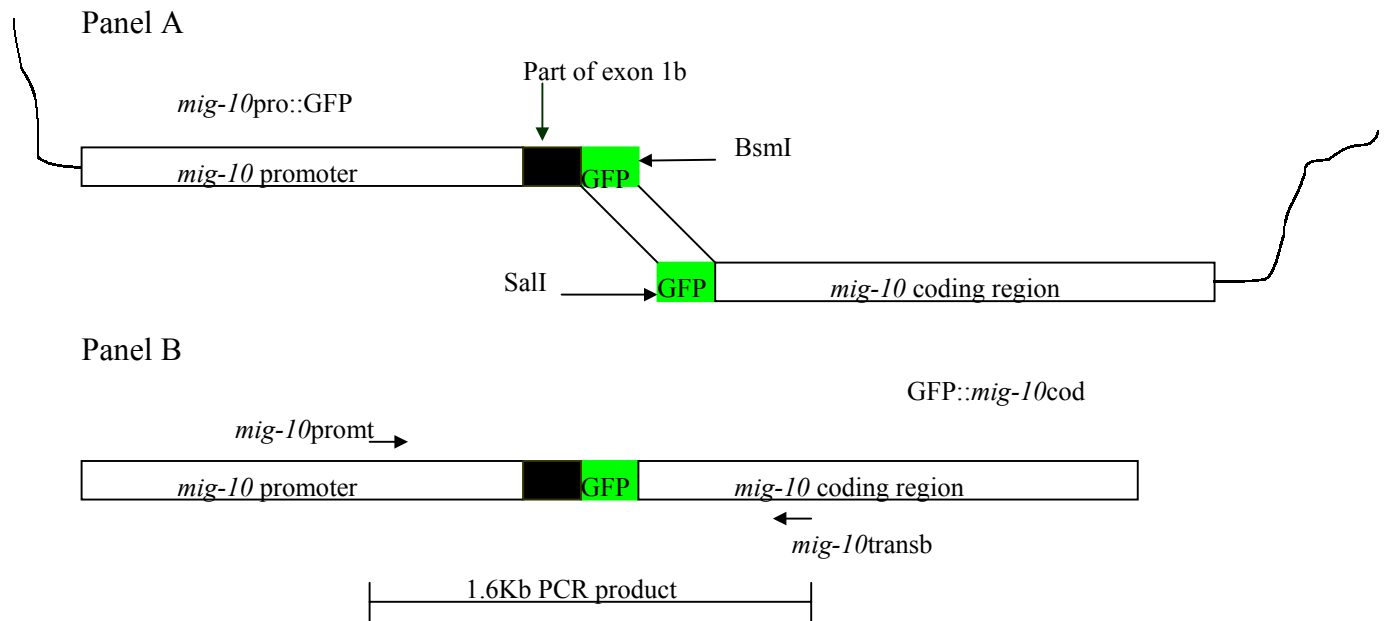
**Figure 29. Schematic representation of recombination of *mig-10*::GFP and GFP::*mig-10*cod, producing the rescuing array to localize MIG-10 activity. Panel A**. The two constructs were cut with appropriate restriction enzymes to place GFP at the end as a target for recombination after injection into RY0040. The recombination is represented by the lines connecting the GFPs on the separate constructs. **Panel B.** PCR primers were designed to flank the area of recombination. The PCR product produced for the recombination array is 1.6Kb, while the non-array PCR product would be 692 bp.
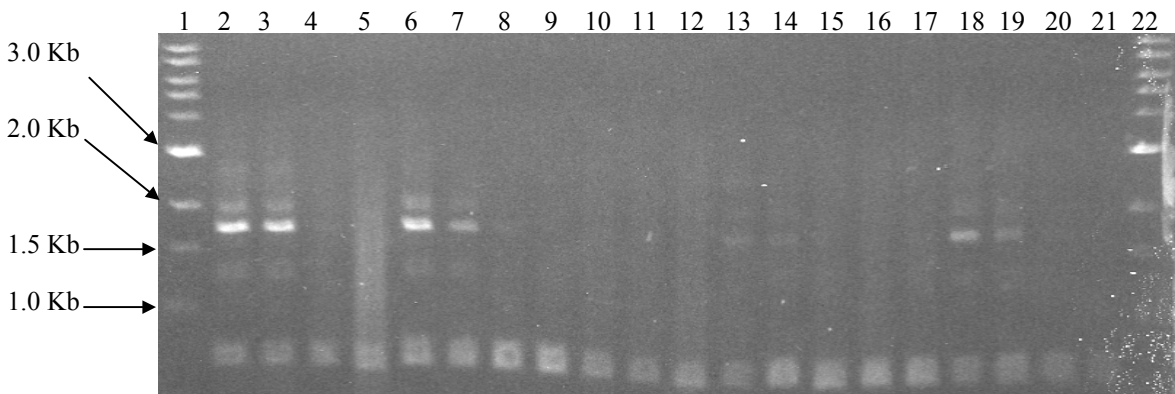


**Figure 30. Screening of homozygous *mig-10* strains from injection for the recombination of *mig-10*::GFP and GFP::*mig-10* by PCR.** All PCR products generated from primers *mi-10*promt and *mig-10*transb. Lanes 1, 22 molecular weight marker. Lanes 2,3 PCR from strain RY0056. Lanes 4,5 PCR from strain RY0065. Lanes 6,7 PCR from strain RY0057. Lanes 8,9 PCR from strain RY0058. Lanes 10,11 PCR from strain RY0059. Lanes 12,13 PCR from strain RY0060. Lanes 14,15 PCR from strain RY0063. Lanes 16,17 PCR from strain RY0064. Lanes 18,19 PCR from strain RY0061. Lanes 20,21 PCR from strain RY0062.

## 4.0 Discussion

### 4.1 *dig-1* has an adhesive function

The nervous system and gonad defects that were observed in *dig-1* mutant animals are consistent with a defect in adhesion. All of the major fascicles of the head contain sensory processes that are affected by mutations in *dig-1* (White et al., 1986). Adhesive defects are also seen in which the gonad migrates incorrectly and fails in some cases to adhere to the ventral body wall (Thomas et al., 1990). Thus, the molecule encoded by this gene appears to provide an adhesive function. Temperature shift assays with a temperature-sensitive *dig-1* mutant have shown that DIG-1 is necessary in the nervous system during embryogenesis, when the head sensory processes are formed, suggesting a role for DIG-1 in process guidance or fasciculation (Ryder unpub. results, **Appendix A**). The temperature shift experiments also showed that *dig-1* is necessary during larval life, most likely playing a maintenance role.

### 4.1.1 Structure and expression of *dig-1*

The open reading frame (ORF) for *dig-1* predicted by Genefinder, K07E12.1a, is approximately 45 Kb in length and encodes a protein molecule containing many motifs typically associated with adhesive molecules. Two lines of evidence support this ORF as being the gene for *dig-1*. The first line of evidence is the rescue of the mutant phenotype by co-injection of two cosmids, K07E12 and R05H11, which when recombined contain the entire ORF encoding DIG-1 (R. Proenca, personal communication). The second line of evidence is the sequence analysis of the two *dig-1* alleles *nu336* and *n1480*. The

sequence analysis showed changes of the DNA in both alleles in this open reading frame. Thus, *dig-1* is contained on the sequenced cosmids K07E12 and R05H11.

The gene *dig-1* encodes a large polypeptide apparently involved in adhesion that is a member of the immunoglobulin superfamily (IgSF). BLAST and the annotated databases SMART and Pfam were used to analyze the gene sequence (Bateman et al., 2002). This analysis revealed three different regions of the gene. Region 1 (N-terminus) contains several Ig domains and several FNIII repeat domains. Region 2 (middle) contains a large highly repeated region as well as two RGD sequences (which are bound by integrins) and several potential N-glycosylation sites (Prosite, Pole Bio-informatique Lyonnais, Bairoch et al., 1997). This region also contains many prolines, which are usually involved in preventing secondary structure, consistent with the idea that the middle region has an extended conformation. Region 3 (C-terminus) contains von Willebrand type A domains (typically involved in adhesion) as well as an EGF domain (typically promotes outgrowth) that might also bind calcium. The importance of the EGF domain potentially binding calcium is unclear, but many molecules in the extracellular matrix need calcium to bind for activity. DIG-1 is most likely not found in the cytoplasm, due to the number of domains typically found extracellularly. DIG-1 is also not likely to be membrane bound since there are no anchor sites such as a GPI attachment site or transmembrane domains predicted on this molecule. The 5' end of the gene predicted by Genefinder did not encode a secretion signal. This predicted 5' end is probably incorrect, with a more likely predicted 5' structure encoding a secretion signal

(R. Proenca, personal communication).  Thus, this molecule is most likely a secreted

ECM molecule.


The open reading frame has at least two splice forms, K07E12.1a and b (Stein et

al., 2001).  Both of the splice variants for the *dig-1* gene are predicted to encode a large

molecule involved in adhesion.  The first splice variant, K07E12.1a, contains 44 exons

and corresponds to the protein structure discussed above.  The second splice variant

contains 36 exons, 35 of which have the same reading frame as K07E12.1a.  The last

exon or exon 36 for K07E12.1b is part of the exon 44 for K07E12.1a, with a different

open reading frame predicted for that exon.  The overall difference in the size of the

transcript is less than 2Kb. The predicted protein for K07E12.1b would be missing the

three von Willebrand type A domains.  Both of these splice variants are supported by

sequence analysis of expressed sequence tags (ESTs) generated in the Kohara lab

(personal communication).  The ten ESTs representing K07E12.1a produce cDNAs of

appropriate sizes; the 5' and 3' ends of the ESTs contained sequence corresponding to the

predicted transcript.  The one EST representing K07E12.1b was completely sequenced

and showed an unusual splicing site between exons 35 and 36.  Since there are many

ESTs corresponding to K07E12.1a and only one EST that is homologous to the

K07E12.1b form, it is likely that 1b is a rare transcript.


**4.1.2 *dig-1* is expressed as at least one very large transcript**

Expression of the *dig-1* gene was detected at the RNA level by RT-PCR and

northern analysis.  A single band of greater than 20Kb was detected in northern blots by

probes from all three regions of the gene, suggesting that the gene encodes at least one very large transcript (Higgins, **Appendix A**). In conjunction with this experiment, RT-PCR was done, which produced products spanning the different regions of *dig-1* predicted from transcript K07E12.1a. cDNA products connecting region 1 to 2 and region 2 to 3 were cloned and sequenced. The sequence data from these two cDNA clones showed the predicted splicing pattern for the K07E12.1a splice form. These results showed the predicted exon/intron structure is correct in this area and suggest that this gene produces one large transcript that could be the predicted K07E12.1a transcript.

Neither of these results rules out the possibility of alternative splicing events occurring, since splicing may still produce large RNA molecules. In particular, the splice variant, K07E12.1b, may not even be discernable as a separate band by northern analysis, since it is less then 2Kb smaller than K07E12.1a. Why would alternative splicing be important to DIG-1 function? In other genes, such as other members of the IgSF molecules and proteoglycans, alternative splicing is one way of producing proteins of altered adhesiveness at different times and locations. DSCAM, a member of the IgSF, is a large gene that is predicted to have 38,000 splice forms, all of which have a similar sized transcript (Schmucker et al., 2000). This is accomplished by the gene being organized with sets of exons of similar size, one arrangement of which is chosen for each splice variant. Thus, many splice variations can be obtained that are all approximately the same size. In *dig-1*, unlike DSCAM, there are not obvious sets of similar sized exons; in particular, the 47 residue repeats of region 2 do not correspond to exon boundaries. Since the size of each exon is minimal when compared to the rest of the gene, alternative

splicing by leaving out a single exon or even groups of exons would not significantly change the size of the transcript.

RT-PCR analysis showed bands of sizes smaller then predicted by Genefinder in several regions of the gene, suggesting that alternative splicing could be occurring. Southern analysis (Higgins) detected both the predicted product and a second cDNA product. This product was smaller than the predicted form, suggesting different exon usage. On the other hand, the size of this cDNA product was not consistent with simply splicing into any of the predicted exons. This result suggests that there could be alternative splicing occurring that uses unusual splice sites that would not be predicted by Genefinder. The use of unusual splice sites is seen in the K07E12.1b splice form, where the splice junction is 10 bp 5' of the predicted splice junction in K07E12.1a. To further investigate the presence of alternative splicing, RT-PCR was done near the 3' end of the gene with primers specific to both splice forms. Only cDNAs from K07E12.1a splice form were generated. This result suggests that K07E12.1b is a rare transcript that is not represented in the RNA preparations that were used for RT-PCR. Additional evidence for alternative splicing comes from screening of several cDNA libraries constructed using various primers homologous to different regions of *dig-1* (R. Proenca, personal communication). These data suggest there is alternative splicing of the K07E12.1a occurring, that would not be detectable by northern analysis of total RNA.

The *dig-1* gene was determined to be necessary during the formation of the nervous system in embryogenesis (**Appendix A**, Temp. shift exp. Ryder, unpub results).

To determine if there were alternative splice forms present in the embryo, cDNA products from N2 mixed population RNA and embryonic RNA were compared. This preliminary result suggested that alternative splicing may occur at the 5' end of the gene during embryogenesis. This work suggests that alternate forms of the gene product could be involved in the formation of the nervous system.

### 4.1.3 Analysis of alleles of *dig-1*

Analysis of the severity of several *dig-1* alleles showed that the alleles formed a series. The allelic series indicated two functions of *dig-1,* one in adhesion, the other in migration. In most *dig-1* alleles, the hermaphrodite gonad is displaced anteriorly from its normal position. In all *dig-1* alleles, a percentage of gonads are displaced dorsally, suggesting that the gonad does not adhere to the ventral body wall. In general, alleles with the most severe neuronal phenotype also had the most severe gonad phenotype, suggesting DIG-1 functions in the same way in both tissues. The exception was hermaphrodites homozygous for *nu336*; although this allele had the most severe nervous system phenotype seen as well as perturbing the gonad attachment to the ventral body wall, the gonad position along the anterior-posterior axis was similar to wild type placement of the gonad. Thus, *nu336* seems to affect gonad positioning differently then adhesion per se.

To understand what is occurring, it is important to look at relevant events in embryogenesis. The gonad primordium is composed of four cells, two germ line cells (Z2 and Z3) arising from the $P_4$ cell, and two somatic cells (Z1 and Z4) derived from

either MSp or Msa, respectively (**Appendix A**). Both of the somatic cells migrate

posteriorly and attach to the germ line cells between 250 and 400 min after fertilization

(Sulston et al., 1983). In *dig-1*(*n1321*) mutants, any dorsal displacement of the gonad

occurring was observed after the posterior migrations were complete (M. Basson, pers.

comm). Thus, anterior and dorsal displacement of the gonad may occur through different

mechanisms.

In the nervous system at about 400 min, the anterior neurons such as IL1 and Il2

move anteriorly towards the tip of the head, and the sensory organs are formed.

Subsequently, the sensory processes are laid down as the neurons move posteriorly

(**Appendix A**, Sulston et al., 1983). Since there are two migration events occurring, the

defects in the sensory process morphology may be due to a perturbation in migration.

Since some of the processes fail to reach the nose, the anterior migration could be

perturbed by DIG-1 not guiding the cell to the tip of the nose, which would then prevent

the sensory process being positioned correctly at the tip of the nose. On the other hand,

the anterior migration might not be affected, but the sensory process may fail to adhere to

the correct position at the tip of the nose. The posterior cell migration could also be

affected, causing an aberrant process placement, or the sensory process might not adhere

properly.

### 4.1.4 Models for function of *dig-1*

The overall organization of this gene is very similar to but not the same as the

hyalectan family of proteoglycans (**Figure 31**). Like DIG-1, three distinct regions make

up these proteoglycans; an Ig domain is found at the N-terminus with an area of low complexity (in some members large) region followed by motifs involved in adhesion at the C-terminus (**Figure 31**).  DIG-1 differs in that it does not have the hyaluronan-binding tandem repeat, which binds hyaluronic acid that always follows the Ig domain in hyalectans as well as having more Ig domains.  DIG-1 further differs by the presence of FNIII domains at the N-terminus and the absence at the C-terminus of a C-type lectin domain (Bandtlow and Zimmermann, 2000).
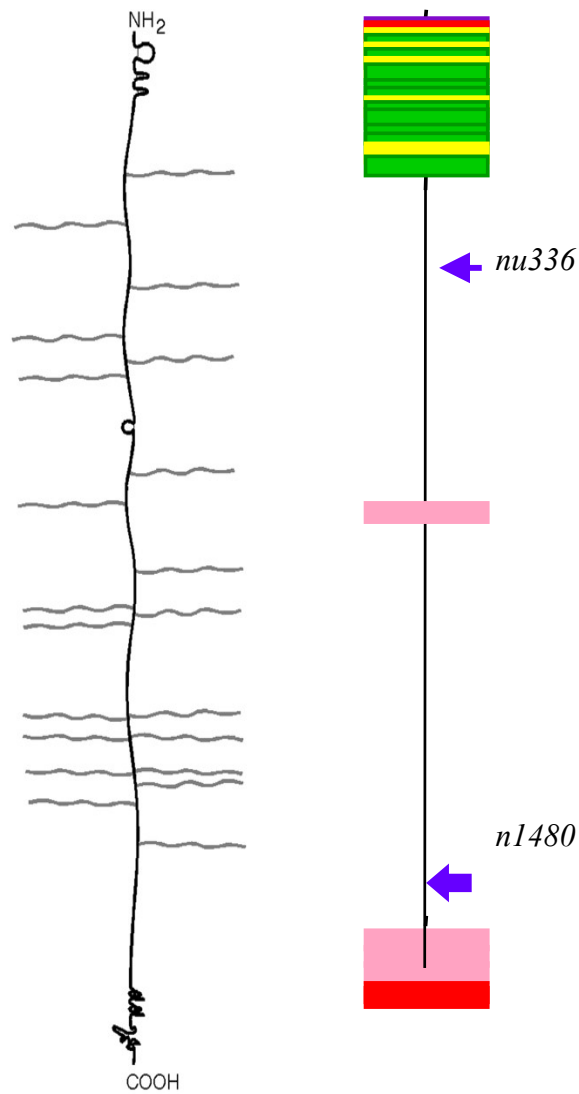
**Figure 31. General comparison of the predicted DIG-1 structure and Versican, a member of the hyalectan proteoglycan family.** The N-terminus has Ig domain(s). The middle of both proteins has many acidic amino acids and prolines, and few cysteines. The C-terminus has domains involved in adhesion. The small blue arrow denotes the position of the lesion in allele *nu336*. The large blue arrow denotes the position of the lesion in allele *n1480* (Bandtlow and Zimmermann, 2000).

How might *dig-1* function as a proteoglycan? The PGs of the hyalectan family

have been shown to interact with other adhesion molecules such as NCAM, NgCAM,

TAG-1 and axonin (Retzler et al., 1996). The majority of the members of the hyalectans

are involved in repelling growth cones (Bandtlow and Zimmermann, 2000). In a model

where DIG-1 is repulsive, it would be expressed around the area that the cell is migrating across and would prevent the cell or process from escaping from the tract (**Figure 32**). The most likely candidates for expression of DIG-1 would be hypodermal or embryonic body muscle cells. Mutations in *dig-1* could prevent the repulsion, allowing the cell to wander around since it is no longer repelled into place.
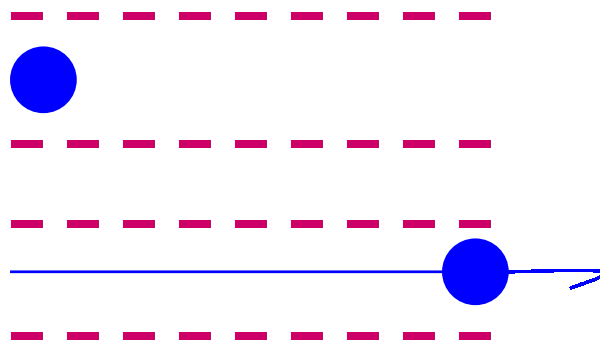


**Figure 32. Schematic representation of DIG-1 as a repulsive molecule.** DIG-1 would be found surrounding the path of cell migrations (represented as large red minus sign). The cell body (blue circle) would then be repelled and "trapped" in this area. The sensory process (blue line) and axon (curved blue line) would then be able to make the correct connections

Alternatively, *dig-1* might function as an attractive adhesion molecule (**Figure 33**). If DIG-1 functions in adhesion, the cells with processes in the same fascicles with those of IL2 neurons are most likely to be expressing DIG-1. In addition to the IL2 neurons, these are the neurons IL1, OLQ, CEP, URY, URA, and a glial like cell, GLR. The domains found at the N- and C-termini of this gene are classically thought to be involved in adhesion. Many of these domains are found in genes that have been shown to promote axon fasciculation and pathfinding via adhesion. RGD sequences are also

involved in adhesion of integrins, which also mediate axon guidance, to molecules in the

ECM (Arnaout et al., 2002). A potential model is that the Ig and FNIII domains bind to

similar domains on the neuron surface, providing directional information to the migrating

cell. Binding of the Ig and FNIII domains might then potentiate the activation (by

conformation change) of the Von Willdebrand factor A domains for tight binding to their

ligand(s), in a manner similar to the integrins.  The ligands of the factor A domain or I-

domain in integrins are usually molecules that contain immunoglobulin or

immunoglobulin-like domains (Colombatti et al., 1993; Lee et al., 1995; Qu and Leahy,

1995; Dickeson and Santoro, 1998; Shimaoka et al., 2002).  The presence of the RGD

sequence on DIG-1 could then be used to couple the extracellular matrix to the

cytoskeleton in the neuronal sensory process. A second possibility is that the

adhesiveness of *dig-1* is modulated by glycosylation similar to N-CAM (Walsh and

Doherty, 1996).  Thus, early in development *dig-1* could provide a permissive substrate,

while later in development the adhesiveness is increased to hold cells in place.  Another

potential mechanism is the differential use of a particular exon. An example of this is the

VASE exon in NCAM, which when utilized, down regulates the neurite growth

promoting activity of NCAM (Walsh and Doherty, 1996).

## Adhesive Model
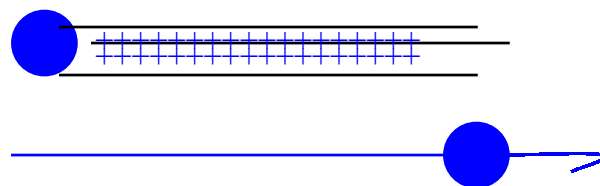


**Figure 33.  DIG-1 could function as an adhesive cue in the path of migration.**
DIG-1 would be located in the path of migration (blue pluses).  The cell body (blue
circle) would be guided along the path by DIG-1 providing a sticky path.
Alternatively, DIG-1 may adhere the sensory process as the cell body is migrating to
its posterior position.

How could the migration function of *dig-1* be mediated?  One model is that different regions of the DIG-1 molecule mediate adhesion and migration.  According to this hypothesis, the N-terminal end would mediate tight adhesion, since this region is most likely perturbed in *nu336*, the mutant with the most severe adhesion phenotype.  The C-terminus could primarily mediate migration, since the *n1480* mutant has a relatively mild adhesive defect but does have an anteriorly displaced gonad.

### 4.1.5 Potential reasons for the severity of the phenotypes of the two sequenced alleles *nu336* and *n1480*

The molecular analysis of *dig-1* provides insight into how this molecule might be involved in the development of the nervous system.  The most severe allele, *nu336*, is not a molecular null, but rather a misssense mutation changing a serine to a phenylalanine. This amino acid change occurs near the beginning of the repeated region of the protein. Why would this change in the repeated region of this gene cause such a severe neuronal process phenotype?  If *dig-1* is a proteoglycan, the altered serine may be a GAG attachment site.  In the hyalectans, one of the GAG's functions is to maintain the extended conformation of the PG core protein (Bandtlow and Zimmermann, 2000, **Figure 34**).  The absence of a GAG could allow the protein to fold differently, causing the N-terminus adhesion area to be unavailable to its adhesive partner.   Along the same lines, if the mutated serine is not a GAG site, the substitution of phenylalanine, an amino acid with a large side chain, could cause a bend in the protein that would decrease the availability of the N-terminus adhesive motifs.  On the other hand, improperly folded proteins are shunted out of the Golgi and degraded (Parodi, 2000).  In this case, no adhesive event could occur, leading to the mutant phenotype.  Although the mutation in

*nu336* is contained in a repeated region, it could be that this particular repeat has some specific role in adhesion. Consistent with this idea, the sequence of this repeat differed somewhat from the consensus (**Appendix A**). For example, a proline that is conserved in 95% of the repeats is replaced by a serine in the *nu336*-containing repeat. It is also possible that a second mutation is present in *nu336* that contributes to the phenotype, since we did not sequence the entire gene. These data suggest that the N-terminal region and the repeated region are both necessary for adhesion.

Unlike *nu336,* the *n1480* mutation causes early transcription termination as well as a stop codon, which would result in truncation of the last third of the protein, removing the C-terminal adhesive motifs as well as part of the C-terminal repetitive region (**Figure 34**). This suggests that the C-terminus of the protein is important for the complete adhesion activity of this molecule. The absence of the C-terminus could perturb tight adhesion; thus, resulting in a little "wobble" in the placement of the sensory process. The fact that this allele is not very severe supports the idea that the N-terminus is more important for mediating tight adhesion. In the *n1480* encoded molecule, the C-terminus would be absent, but the N-terminal adhesive motifs as well as the potential RGD binding sites would be able to interact with their target. This would stabilize the connection between the cells and allow the sensory processes to adhere for the most part at the correct place.

Panel A

Wild type

Panel B

*nu336*

Panel C

*n1480*
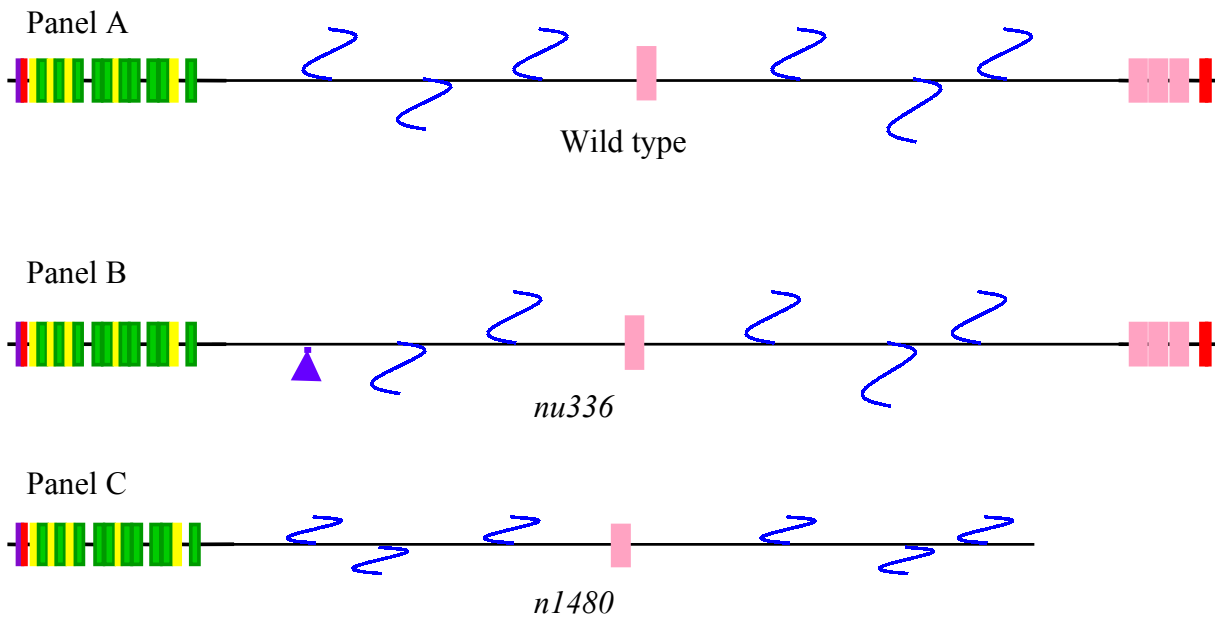
**Figure 34. Potential proteins present in wild type and mutant worms.**
Panel A. Representation of DIG-1 as a proteoglycan. GAGs are represented as blue
squiggles. Panel B. Mutant DIG-1 encoded by allele *nu336,* which is missing the GAG
attachment site. The blue arrowhead indicates the point mutation in *nu336,* which
prevents GAG attachment. Panel C. Mutant DIG-1 encoded by allele *n1480,* with the
predicted truncation shown. Domains are the same as in **Results, Figure 14**.

**4.1.6 Localization of DIG-1**

To obtain a reagent for localization of the DIG-1 protein, in situ we generated

antibodies to recombinant proteins from cDNAs from region 1 and region 3. Western

analysis showed that an antibody was present in serum from two different rabbits for

recombinant proteins from both regions of DIG-1. Western and dot blot analysis of the

antibodies indicate that the antibodies need to be purified for use in localizing DIG-1 in

embryo staining, and that non-specific binding of antibodies in pre-immune serum needs

to be controlled.

In summary, we have identified the gene for *dig-1*, as well as showed that *dig-1* is expressed largely as predicted by Genefinder. We have started to understand the function of the protein through the analysis of several mutations that have indicated at least two roles for the protein during development. The main role that DIG-1 seems to be involved in is adhesion. A second function is providing information for migration. There could also be synergy between the models. For example, initially the adhesion motifs might guide the cell migration, and then later, when sugars are added, the sensory process might be repelled so that it then stays in its target area.

Now that we have produced antibodies, we can begin to determine where this protein is localized. The localization of the protein could then provide information on how the molecule is working. If DIG-1 provides an adhesive pathway, we would expect to localize DIG-1 along the tracts of the neuronal process. The cells expressing DIG-1 would be those found in the fascicle with the IL2 cells. If DIG-1 is repelling the cell body, the protein should localize alongside of the paths, which would prevent the neuronal process from leaving that tract. The cells expressing DIG-1 for the repulsive model could be the hypodermal or embryonic body muscle cells.

## 4.2 *mig-10*

Several transgenic lines of animals containing GFP::*mig-10* constructs were generated by injection. Several of these lines tested positive by PCR for the presence of the injection materials. The PCR results support the idea that the injected materials were recombined in a way that would potentially lead to production of a wild type GFP::MIG-

10 fusion protein.  If the PCR positive lines do turn out to be rescued, these lines could then be used to localize MIG-10, to begin to understand where it functions to promote axon pathfinding.

**4.3 Future directions**

An exciting future direction is to look at the potential alternative splicing of *dig-1* RNA during the time when the nervous system is being formed.  The initial RT-PCR analysis indicated potential splicing events could be occurring at the time the nervous system is being formed.  Cloning and sequencing of those cDNA products could identify which exons are present in the different splice forms.  A second exciting area is in work with the antibodies.  The initial western results indicate the presence of background staining, which will have to be removed.  Since the pre-immune serum did not recognize either of the recombinant proteins, these could be used to purify antibodies that are specific for the recombinant proteins.  After purification, it would be necessary to determine if the antibodies are specific for DIG-1.  In *n1480* animals, DIG-1 should be missing the portion of the protein that anti-Ant3Con3 should detect.  Thus, purified antibodies could be used to screen *n1480* and N2 homogenates. If anti-Ant3Con3 fails to detect DIG-1 in a *n1480* homogenate but does detect something in N2, this should be DIG-1, thus indicating that anti-Ant3Con3 is specific for DIG-1.  After determining the specificity of the antibodies, they could be used in localizing DIG-1 in adults and embryos. If DIG-1 is a proteoglycan, we could test whether it is glycosylated by removal of the GAGs and determine if there was a shift in molecular size.   If DIG-1 changes its structure during development, we could then look for differences in glycosylation states

during development; it could get more or less glycosylated depending on whether it is adhesive or repulsive.  All of these experiments would provide evidence that could be used to refine our model of how DIG-1 functions in the development of the nervous system.

# 5.0 References

Agarwala K.L., S. Ganesh, T. Suzuki, T. Akagi, K. Kaneko, K. Amano, Y. Tsutsumi, K. Yamaguchi, T. Hashikawa, K. Yamakawa. (2001). Dscam is associated with axonal and dendritic features of neuronal cells. J. Neurosci. Res. 66(3): 337-46.

Anderson P. (1995). Mutagenesis. In *Methods in Cell Biology Vol. 48. Caenorhabditis elegans*: *Modern Biological Analysis of an Organism*. (Edited by H.F. Epstein and D.C. Shakes). pp. 31-54. New York: Academic Press, Inc.

Arnaout M., S. Goodman, J. Xiong. (2002). Coming to grips with integrin binding to ligands. Curr. Opin. Cell Biol. 14(5):641

Aurelio O., D.H. Hall, and O. Hobert. (2002). Immunoglobulin-domain proteins required for maintenance of ventral nerve cord organization. Science. 295(5555): 686-702

Ausubel, F., R. Brent, R. Kingston, D. Moore, J. Seidman, J. Smith, K. Struhl, Eds. (1998). Current Protocols in Molecular Biology. John Wiley and Sons, Inc.

Bairoch A., P. Bucher, and K. Hofmann. (1997). The PROSITE database, its status in 1997. Nucleic Acids Res. 25(1):217-221.

Bandtlow C.E., and D. R. Zimmermann. (2000). Proteoglycans in the developing brain: new conceptual insights for old proteins. Physolog. Rev. 80(4): 1267-1290.

Bateman A., E. Birney, L. Cerruti, R. Durbin, L. Etwiller, S.R. Eddy, S. Griffiths-Jones, K.L. Howe, M. Marshall and E.L.L. Sonnhammer. (2002). The Pfam protein families database. Nucl. Acid Res. 30(1): 276-280.

Bentley D., and T. P. O'Connor. (1994). Cytoskeleton events in growth cone steering. Curr. Opin. Neurobio. 4:43-48.

Bovolenta P., and I. Fernaud-Espinosa. (2000). Nervous system proteoglycans as modulators of neurite outgrowth. Prog. Neurobiol. 61: 113-132.

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

Brummendorf T. and V. Lemmon. (2001). Immunoglobulin superfamily receptors: cis-interactions, intracellular adapters and alternative splicing regulate adhesion. Curr. Opin. Cell Biol. 13(5): 611-8.

Brummendorf T., and F.G., Rathjen. (1996). Structure/function relationships of axon-associated adhesion receptors of the immunoglobulin superfamily. Curr. Opin. Neurobiol. 6:584-593.

Burket C.T., C. Higgins, S. Hubbard, E. Roach, and E. F. Ryder. The *C. elegans* gene *dig-1* encodes a novel adhesion molecule that functions during sensory process development in the nervous system. In preparation.

The C. elegans Sequencing Consortium. (1998). Genome sequence of the nematode C elegans: a platform for investigating biology. Science. 282:2012-8.

Chalfie M., Y. Tu, G. Euskirchen, W. Ward, and D. Prasher. (1994). Green fluorescent protein as a marker for gene expression. Science. 263: 802-805.

Clark D.V., D.S. Suleman, K.A. Bleckenbach, E.J. Gilchrist, and D.L. Baillie. (1995). Molecular cloning and characterization of the *dpy-20* gene of *Caenorhabditis elegans*. Mol. Gen. Genet. 247:367-378.

Colamarino S. A., and M. Tessier-Lavigne. (1995). The axonal chemoattractant netrin-1 is also a chemorepellent for trochlear motor axons. Cell. 81:621-629.

Colombatti A., P. Bonaldo, and R. Doliana. (1993). Type A modules: interacting domains found in several non-fibrillar collagens and in other extracellular matrix proteins. Matrix. 13(4):297-306.

Condic M., D.M. Snow, and P. Letourneau. (1999). Embryonic neurons adapt to the inhibitory proteoglycan aggrecan by increasing integrin expression. J. Neurosci. 19(22): 10,036-10,043.

Crino P.B., and J. Eberwine. (1996). Molecular characterization of the dendritic growth cone: regulated mRNA transport and local protein synthesis. Neuron. 17(6):1173-87.

Crossin K.L., L.A. Krushel. (2000). Cellular signaling by neural cell adhesion molecules of the immunoglobulin superfamily. Dev. Dyn. 218(2): 260-79.

Cunningham B.A., J. J. Hemperly, B. A. Murray, E. A. Prediger, R. Brackenbury, and G. M. Edelman. (1987). Neural cell adhesion molecule: structure, immunoglobulin-like domains, cell surface modulation, and alternative RNA splicing. Science. 236: 799-805.

Dickeson S.K., and S.A. Santoro. (1998). Ligand recognition by the I domain-containing integrins. Cell Mol Life Sci. 54(6):556-566.

Dickson L.J. (2001). Rho GTPases in growth cone guidance. Curr. Opi. Neurobio. 11: 103-110.

Dieffenbach C.W., and G.S. Dveksler Eds. (1995). PCR primer: A laboratory Manual. Cold Spring Harbor Press: New York.

Doherty P., and F.S. Walsh. (1996). CAM-FGF receptor interactions: a model for axonal growth. Mol. Cell. Neurosci. 8(2-3): 99-111.

Doherty P., G. Williams, E.J. Williams. (2000). CAMs and axonal growth: a critical evaluation of the role of calcium and the MAPK cascade. Mol. Cell Neurosci. 16(4): 283-95.

Estrach S., S. Schmidt, S. Diriong, A. Penns, A. Blangy, P. Fort, and A. Debant. (2002). The human Rho-GEF Trio and its target GTPase RhoG are involved in the NGF pathway, leading to neurite outgrowth. Curr. Bio. 12: 307-312.

Fambrough D., and C. S. Goodman. (1996). The Drosophila *beaten path* gene encodes a novel secreted protein that regulates defasciculation at motor axon choice points. Cell. 87: 1049-1058.

Gilbert S.F. (1997). Early vertebrate development: neurulation and the ectoderm. Developmental Biology 5$^{th}$ Ed. Sinauer Associates, Inc. Sunderland, MA. pp 253-306.

Giger R. J., and A. L. Kolodkin. (2001). Silencing the siren: guidance cue hierarchies at the CNS midline. Cell. 105: 1-4

Godenschwege T.A., J. H. Simpson, X. Shan, G.J. Bashaw, C.S. Goodman, and R.K. Murphey. (2002). Ectopic expression in the giant fiber system of *Drosophila* reveals distinct roles for roundabout (Robo), Robo2, and Robo 3 in dendritic guidance and synaptic connectivity. J. Neurosci. 22(8):3117-3129.

Gordon-Weeks P.R., (1987). The cytoskeletons of isolated, neuronal growth cones. Neuroscience. 21:977-989.

Goshima Y, T. Ito, Y. Sasaki, F. Nakamura. (2002). Semaphorins as signals for cell repulsion and invasion. J. Clin. Invest. 109(8): 993-8.

Goodman C.S., M.J. Bastini, C.Q. Doe, S. du Lac, S.L. Helfand, J.Y. Kuwada, and J.B. Thomson. (1984). Cell recognition during neuronal development, Science. 225:1271-1279.

Hao J.C., T. W. Yu, K. Fujisawa, J. C. Culotti, K. Gengyo-Ando, S. Maitani, G. Moulder, R. Barstead, and M. Tessier-Lavigne. (2001). *C. elegans* slit acts in midline, dorsal-ventral, and anterior-posterior guidance via the SAX-3/Robo receptor. Neuron. 32: 25-38.

Horvitz H.R., R. Brenner, S. Hodgkin, R. K. Herman. (1979). A uniform genetic nomenclature for the nematode *Caenorhabditis elegans*. Mol. Gen. Genet. 175: 129-133.

Huaiyu Hu. (2001). Cell-surface heparan sulfate is involved in the repulsive guidance activities of Slit2 protein. Nat. Neurosci. 4(7): 695-701.

Huang X.Y., and D. Hirsh. (1989). A second trans-spliced RNA leader sequence in the nematode *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. August 4, 8640-8644.

Hubbard S. (2000). The molecular characterization of an adhesion molecule involved in sensory map formation in *C. elegans*. A Masters Degree Project, Worcester Polytechnic Institute, Worcester, MA.

Kamiguchi H and V. Lemmon. (2000). IgCAMs: bidirectional signals underlying neurite growth. Curr. Opin. Cell Biol.12(5): 598-605.

Kaplan J.M. (1996). Sensory signaling in *Caenorhabditis elegans*. Curr. Opin. Neurobiol. 6:494-499.

Kaprielian Z., E. Runko, and R. Imondi. (2001). Axon guidance at the midline choice point. Develop. Dyn. 221: 154-181.

Kater SB., and V. Rehder (1995). The sensory-motor role of growth cone filopodia. Curr. Opin. Neurobiol. 5(1):68-74.

Keith C.H, and M.T. Wilson. (2001). Factors controlling axonal and dendritic arbors. Int. Rev. Cytol. 205: 77-147.

Kennedy T.E., T. Serafini, J.R. de la Torre, and M. Tessier-Lavigne. (1994). Netrins are diffusible chemotropic factors for commissural axons in the embryonic spinal cord. Cell. 78: 425-435.

Keynes R., and G.M.W Cook.(1995). Axon guidance molecules. Cell. 83: 161-169.

Kim D.M., P.Kolodziej, A.Chiba. (2002). Growth Cone Pathfinding and Filopodial Dynamics Are Mediated Separately by Cdc42 Activation. J. Neurosci. *22*(5): 1794–1806.


Knoll B. and U. Drescher. (2002).  Ephrin-As as receptors in topographic projections. TRENDS in Neurosci. 25(3): 145-149.

Koery C.A., and D. Van Vactor. (2000). From the growth cone surface to the cytoskeleton: one journey, many paths. J. Neurobiol. 44: 184-193.

Krause M., and D. Hirsh. (1987). A trans-splice leader sequence on actin mRNA in *C. elegans*. Cell. 49: 753-761.

Lander A. D., and S. B. Selleck. (2000). The Elusive Functions of Proteoglycans: In Vivo Veritas. J. Cell Bio. 148(2): 227–232.

Lankford K., C. Cypher, and P. Letourneau. (1990). Nerve growth cone motility. Curr Opin Cell Biol. 2(1):80-5.

Lee J.O., P. Rieu, M.A. Arnaut, R. Liddington. (1995). Crystal structure of the A domain from the alpha subunit of integrin CR3 (CD11b/CD18). Cell. 80(4):631-638.

Lee S.J., and E.N. Benveniste. (1999). Adhesion molecule expression and regulation on cells of the central nervous system. J. Neuroimmunol. 98(2): 77-88.

Lemmon M.A., K.M. Ferguson, and J. Schlessinger. (1996). PH domains: Deverse sequences with a common fold recruit signaling molecules to the cell surface. Cell. 85: 621-624.

Lewis J.A., and J.T. Fleming. (1995). Basic culture methods. Meth. Cell Bio. 48: 3-29.

Li H., T.C. Leung, S. Hoffman, J. Balsamo, and J. Lilien.(2000). Coordinate regulation of cadherin and integrin function by the chondroitin sulfate proteoglycan neurocan. J. Cell. Bio. 149(6): 1275-1288.

Livesey F.J. (1999). Netrins and netrin receptors. Cell Mol. Life Sci. 56(1-2): 62-8

Long D. (2001). Screen of *C. elegans* mutations for nerve defects, a Major Qaulifying Project Report, Worcester Polytechnic Institute, Worcester, MA

Mason G. (1997).  Analysis of a sensory map in mutant nematodes, a Major Qaulifying Project Report, Worcester Polytechnic Institute, Worcester, MA

Manser J., C. Roonprapunt, and B. Margolis. (1997).  C. elegans cell migration gene mig-10 shares similarities with a family of SH2 domain proteins and acts cell nonautonomously in excretory canal development. Dev. Bio. 184:150-164.

McAllister A.K.(2002). Conserved cues for axon and dendrite growth in the developing cortex. Neuron. 33(1): 2-4.

McKerracher L., M. Chamoux, and C.O. Arregui. (1996). Role of laminin and integrin interactions in growth cone guidance. Mole. Neurobiol. 12:95-115.

Mello C., and A. Fire. (1995). DNA Transformation. Methods in Cell Biology.  48:451-482.

Mello C.,  J. M. Kramer, D. Stinchcomb, and V. Ambros. (1991). Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and intergration of transforming sequences. The EMBO J. 10(12): 3959-3970.

Milev P., H. Monnerie, S. Popp, R. K. Margolis, and R.U. Margolis. (1998). The core protein of the chondroitin sulfate proteoglycan phosphacan is a high-affinity ligand of fibroblast growth factor-2 and potentiates its mitogenic activity. J. Biol. Chem. 273(34): 21439-21442.

Mizoguchi A, H. Nakanishi, K. Kimura, K. Matsubara, K. Ozaki-Kuroda, T. Katata, T. Niethammer P., M. Delling, V. Sytnyk, A. Dityatev, K. Fukami. (2002). Cosignaling of NCAM via lipid rafts and the FGF receptor is required for neuritogenesis. J. Cell. Bio. 157(3): 521-532.

Oleszewski M., S. Beer, S. Katich, C. Geiger, Y. Zeller, U. Rauch, and P. Altevogt (1999). Integrin and neurocan binding to L1 involves distinct Ig domains. J. Biol. Chem. 274(35): 24,602-24,610.

Olin A.I., M. Morgelin, T. Sasaki, R. Timpl, D. Heinegard, and A. Aspberg. (2001). The proteoglycans aggrecan and versican form networks with fibulin-2 through their lectin domain binding. J. Biol. Chem. 276(2):1253-1261.

Papenfuss K.A. (2000)The role of the MIG-10 gene in topographic sensory map formation in the soil nematode *C. elegans*, A Major Qualifying Project, Worcester Polytechnic Institute, Worcester MA.

Parodi A. (2000). Role of N-oligosaccharide endoplasmic reticulum processing reactions in glycoprotein folding and degradation. Biochem. J. 328:1-13.

Patel B.N., and D. Van Vactor. (2002). Axon guidance: the cytoplasmic tail. Curr. Opin. Cell Bio. 14: 221-229.

Prag S., E.A. Lepekhin, K. Kolkova, R. Hartmann-Petersen, A. Kawa, P.S. Walmond, V. Belman, H.C. Gallagher, V. Berezin, E. Bock, and N. Pedersen. (2001). NCAM regulates cell motility. J. Cell Sci. 115: 283-292.

Ranscht B. (2000). Cadherins: molecular codes for axon guidance and synapse formation. Int. J. Dev. Neurosci. 18(7): 643-51.

Retzler C., W. Gohring, and U. Rauch. (1996). Analysis of neurocan structure interacting with the neuronal cell adhesion molecule N-CAM. J. Bio. Chem. 271(4): 27,304-27,310.

Riddle D.L., T. Blumenthal, B. J. Meyer, and J. R. Priess. (1997). Introduction to *C. elegans*. In *C. elegans* II. Edited by D.L. Riddle, T. Blumenthal, and B.J. Meyers. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY pp 1-22.

Rivard M.V. (2001). An expression pattern study of *mig-10,* a gene required for nervous system development in *C. elegans*. a Major Qualifying Project Report. Worcester Polytechnic Institute, Worcester MA.

Rusiecki D.N. (1999). The role of the *mig-10* gene in *C. elegans* sensory map formation. A Maters Degree Project, Worcester Polytechnic Institute, Worcester, MA.

Rutishauser U. (2000). Defining a role and mechanism for IgCAM function in vertebrate axon guidance. J. Cell Bio. 149(4): 757-759.

Schmid Ralf-Steffen, R.D. Graff, M.D. Schaller, S. Chen, M. Schachner, J.J. Hemperly, and P. F. Maness (1999). NCAM stimulates the Ras-MAPK pathway and CREB phosphorylation in neuronal cells. J. Neurobiol. 38: 542-558.

Schmucker D., J. C. Clemens, H. Shu, C. A. Worby, J. Xiao, M. Mudo, J. E, Dickson, and S. L. Zipursky. (2000). *Drosophila* Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. Cell. 101: 671-684.

Schwartz N.B. (2000). Biosynthesis and regulation of expression of proteoglycans. Frontiers in Biosci. 5: 649-655.

Serafini T., S.A. Colamarino, E.D. Leonardo, H. Wang, R. Beddington, W. C Skarnes, and M. Tessier-Lavigne. (1996). Netrin-1 is required for commissural axon guidance in the developing vertebrate nervous system. Cell. 87: 1001-1014.

Shimaoka M., J. Takagi, and T.A. Springer. (2002). Conformational regulation of integrin structure and function. Annu. Rev. Biophys. Biomol. Struct. 31:483-516.

Shirasaki R., R. Katsumata, F. Murakami. (1998). Change in Chemoattractant Responsiveness of Developing Axons at an Intermediate target. Science. 279(5347): 105-107.

Skaper S.D, S.E. Moore, F.S. Walsh. (2001). Cell signalling cascades regulating neuronal growth-promoting and inhibitory cues. Prog. Neurobiol. 65(6): 593-608.

Spieth J., G. Brooks, S. Kuersten, K. Lea, and T. Blumenthal. (1993). Operons in *C. elegans*: polycistronic mRNA precursors are processed by trans-splicing of SL2 to downstream coding regions. Cell. 73: 521-532.

Stein E., and M. Tessier-Lavigne. (2001). Hierarcical organization of guidance receptors: Silencing of netrin attraction by slit through a Robo/DCC receptor complex. Science. 291(5510): 1928-1938.

Stein E., P. Sternberg, R. Durbin, J. Thierry-Mieg, and J. Speith. (2001). Wormbase: network access to the genome and biology of *Caenorhabditis elegans*. Nucliec Acids Research. 29:82-86.

Stevens A., and J.R. Jacobs. (2002). Integrins Regulate Responsiveness to Slit Repellent Signals. J. Neurosci. *22*(11): 4448–4455.

Sulston J.E., and H.R. Horvitz. (1977). Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. Dev. Biol. 56:110-156.

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

Suter D.M., and P. Forscher. (2000). Substrate-Cytoskeletal Coupling as a Mechanism for the Regulation of Growth Cone Motility and Guidance. J. Neurobiol. 44(2): 97-113.

Tanaka E., and J. Sabry (1995). Making the connection: cytoskeletal rearrangements during growth cone guidance. Cell. 83(2):171-6.

Tessier-Lavigne M. (1994). Axon guidance by diffusible repellants and attractants. Curr. Opi. Gen Dev., 4: 596-601.

Tessier-Lavigne M., and C.S. Goodman. (1996). The Molecular Biology of Axon Guidance. Science. 274:1123-1132.

Thomas J.H., M.J. Stern, and H.R. Horvitz. (1990). Cell interactions coordinate the development of the *C. elegans* egg-laying system. Cell. 62: 1041-1052

Qu A., and D.J. Leahy. (1995). Crystal structure of the I-domain from the CD11a/DC18 (LFA-1, alpha L beta 2) integrin. Proc. Natl. Acad. Sci. USA. 92(22):10277-81.

Van Vactor D. (1998). Adhesion and signaling in axonal fasciculation. Curr. Opin. Neurobiol. 8: 80-86.

Walsh F.S. and P. Doherty. (1997). Neural cell adhesion molecules of the immunoglobulin superfamily: role in axon growth and guidance. Annu. Rev. Cell Dev. Biol. 13: 425-56.

White J.G., E. Southgate, J.N. Thomson and S. Brenner. (1986). The structure of the nervous system of the nematode *Caenorhabditis elegans*. Phil. Trans. Roy. Soc. (Lond.) 314B:1-340.

Wojcik J., J. Girault, G. Labesse, J. Chomilier, J. Mornon, and I. Callebaut. (1999). Sequence analysis identifies a Ras-associating (RA)-like domain in the n-termini of band 4.1/JEF domains and in the Grb7/10/14 adapter family. Biochem. Biophy. Res. Comm. 259: 112-120.

Zhou F.Q., C. M. Waterman-Storer, and C.S. Cohan. (2002). Focal loss of actin bundles causes microtubule redistribution and growth cone turning. J. Cell Bio. 157(5): 839–849.

**Appendix A: Northern analysis of N2 and *n1480* RNA, Migrations of a cell that forms the IL2 sensory map, Migration and adhesion of cells that form the gonad, The severity of the gonad defect defined an allelic series for the seven *dig-1* alleles, Temperature shift assay summary, and Protein repeats.**

**Northern analysis of N2 and *n1480* MP RNA.**
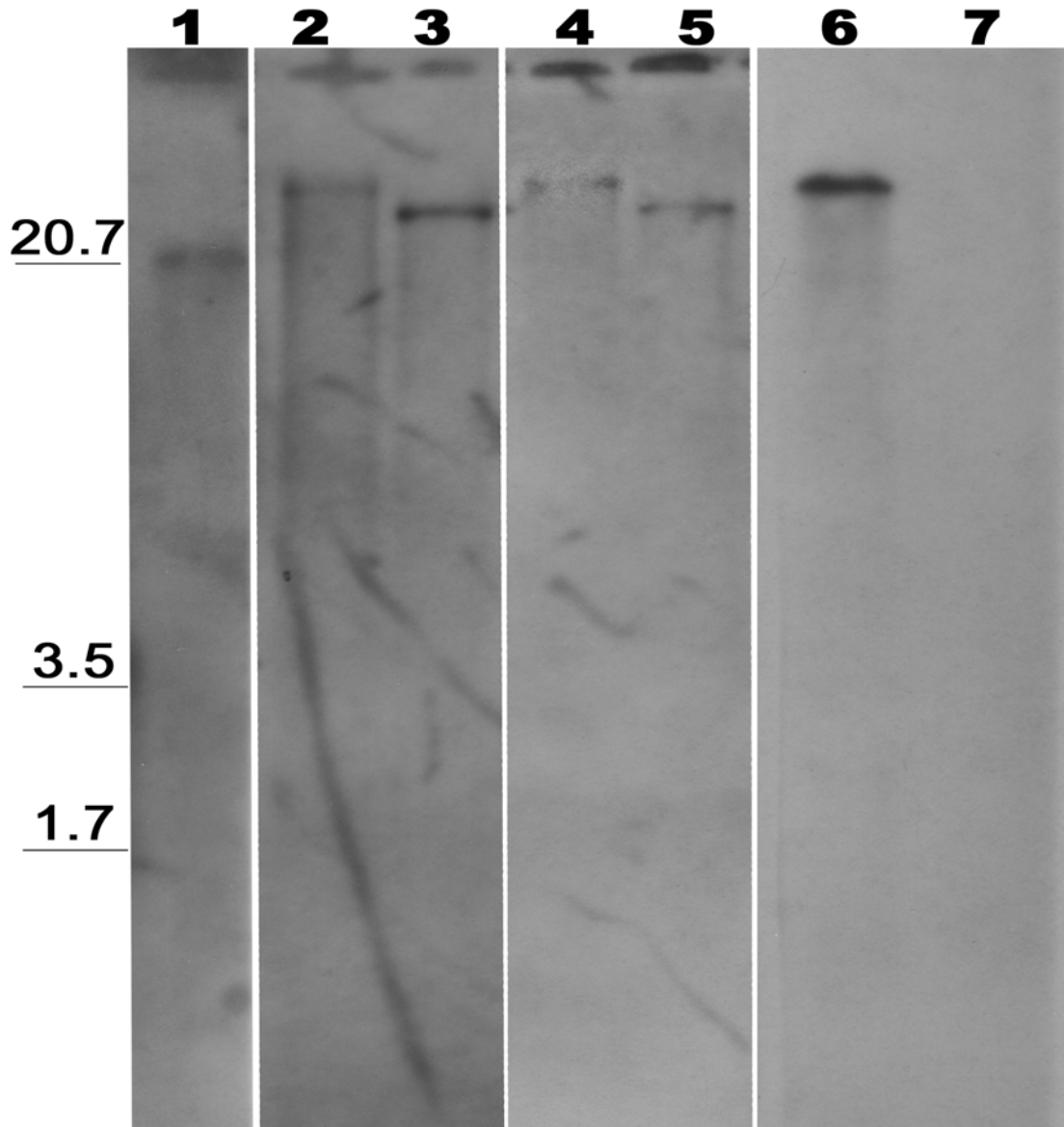


**Figure 1A. Northern analysis of N2 and *n1480* RNA.** Lane 1, *unc-89* control showing the 20.7Kb transcript. Lanes 2-3, N2 and n1480 probed with a probe to the 5' end of *dig-1*, showing size difference of the transcripts. Lanes 4-5, N2 and *n1480* probed with a probe to the middle region of *dig-1*, showing the size difference in the transcripts. Lanes 6-7, N2 and *n1480* probed with a probe to the 3' end of *dig-1*, showing the absence of the transcript in allele *n1480*.
Done by: Christina Higgins, Lynn Hull, and Stacy Hubbard

**IL2 neuron migrations**



Nose

Anterior migration to the nose

Posterior migration to target area near the nerve ring

Wild type

*dig-1*
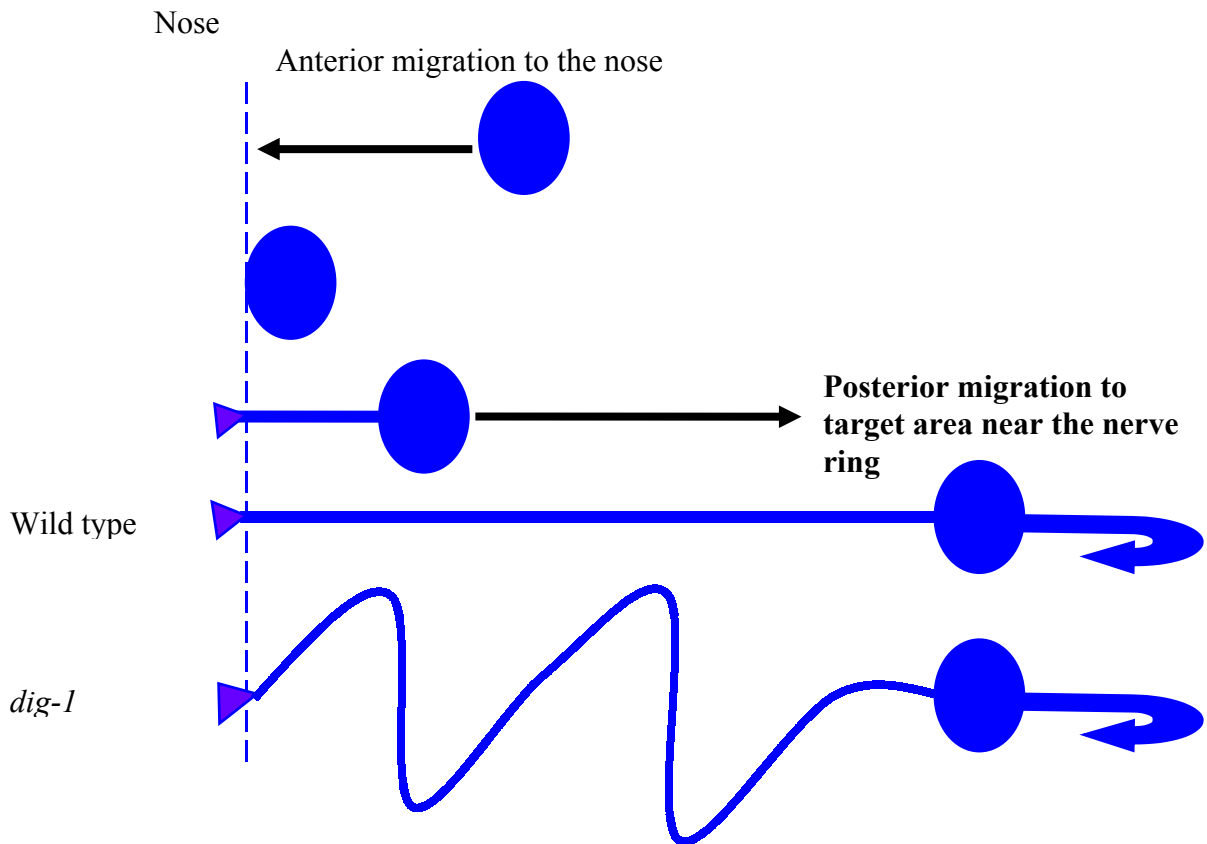
**Figure 2A. Migrations of a cell that forms part of the IL2 sensory map.** The cell first makes an anterior migration to the nose. At the nose the ciliary sensillum is formed and the cell then makes a posterior migration. As the cell makes its posterior migration it lays the down the sensory process (wild type). A severe defect seen in some *dig-1* animals is also shown. The vertical dotted line is the nose. The large circle is the cell body. The solid blue line and wavy line are the sensory processes in wild type and mutants. The curved arrow is the axon. The solid black lines represent the direction of migration.

**Cells that form the gonad migration**
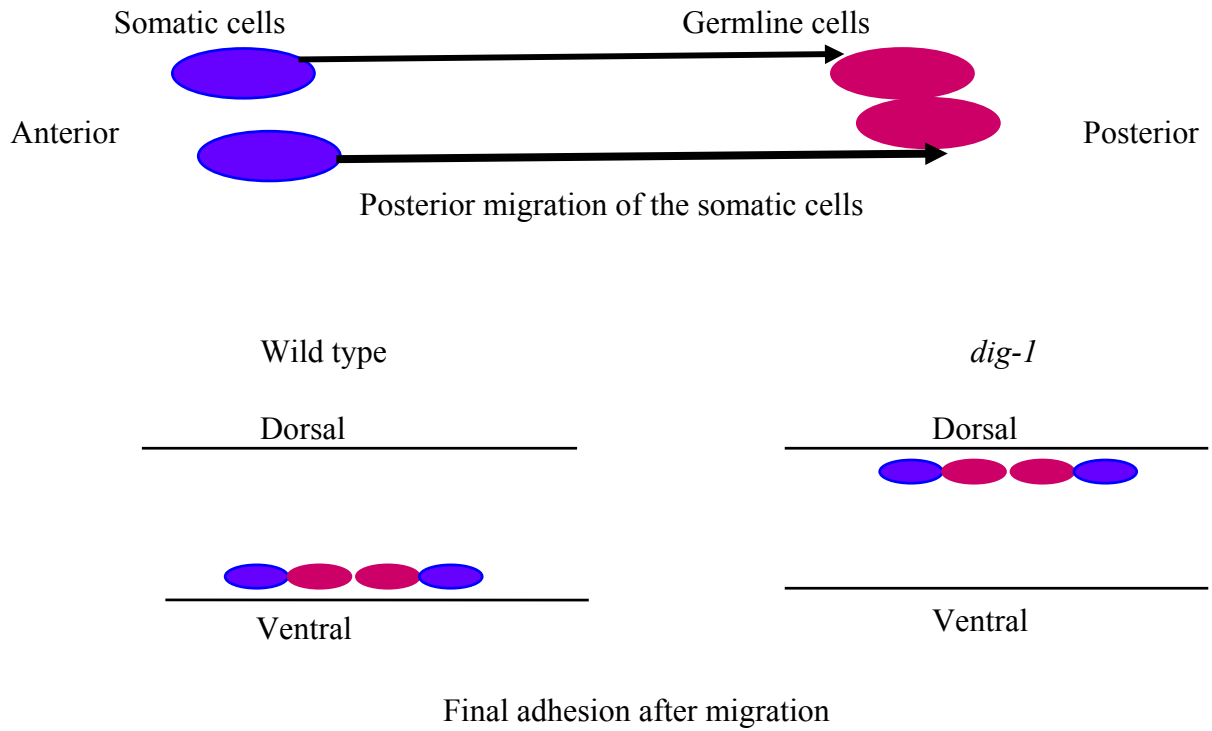


Figure 3A. **Migration and adhesion of cells that form the gonad.** Panel A. The somatic cells (blue circles) migrate posteriorly to contact the germline cells (red circles). The migration occurs first followed by adhesion at the site of gonad formation. Panel B. The wild type adheres to the ventral side. In some *dig-1* animals, the gonad is displaced dorsally.

**Gonad defect summary**

dig-1 allele

| Position of gonad \ Position of P5/6 cell | nu345 | n1321 | n1480 | nu52 | n2467 | nu319ts | nu336 | + |
|---|---|---|---|---|---|---|---|---|
| (Z1 Z2 Z3 Z4) | 2 | 1 | 3 | 14 | 15 | 10 | 19 | 23 |
| | 6 | 7 | 11 | 8 | 10 | 12 | 9 | 7 |
| | 13 | 19 | 4 | 7 | 5 | 1 | 2 | 0 |
| | 9 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| | 5 | 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |

**Figure 4A. The severity of the gonad defect defined an allelic series for the seven dig-1 alleles.** The degree of anterior displacement of the gonad primordium in L1 larvae was measured with respect to either the left or right P5/6 cell (before its migration to the ventral cord; Sulston and Horvitz, 1977). The P5/6 cells are nongonadal cells not affected by the *dig-1* mutation (Thomas et al., 1990). The number of animals displaying each degree of displacement is shown for all the *dig-1* alleles. The assay was performed by viewing L1 animals using Nomarski optics (Thomas et al., 1990). Alleles are ordered by severity of the phenotype. *dig-1(nu319ts)* animals were raised at the non-permissive temperature of 25°C.

Done by: Dr. Ryder

**Temperature-shift experiments**



Graph: Y-axis "percent animals with aberrant IL2 sensory processes" marked at 100, 50. X-axis "Stage of temperature shift" with labels: early embryo, late embryo, L1, L2/L3, never. Curve labeled 25°C→15°C (circles, increasing) and 15°C→25°C (squares, decreasing).
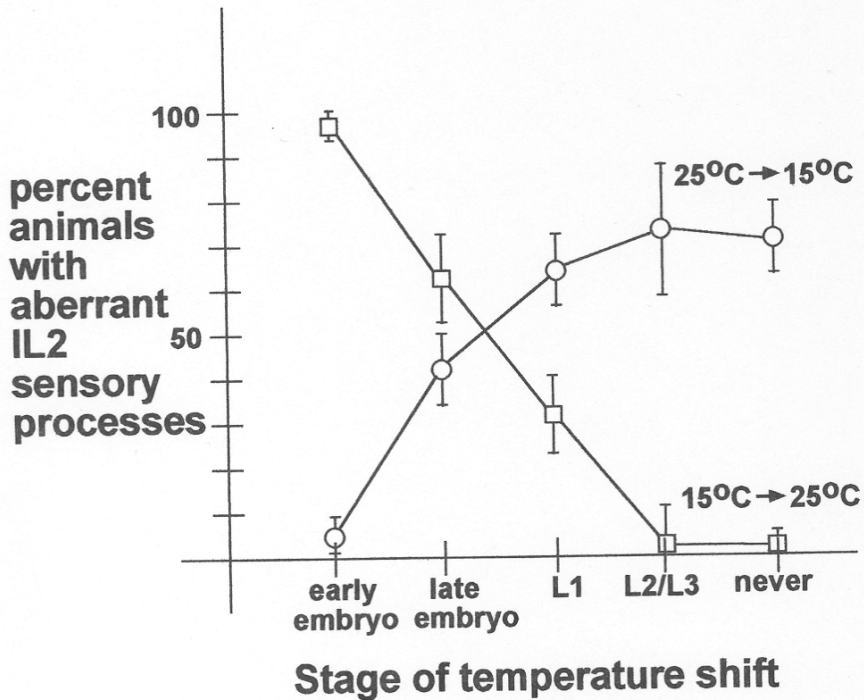
**Figure 5A. Temperature-shift experiments suggest *dig-1* is required during development and early first larvae stage.** *dig-1* (*nu319ts*) animals were grown for one generation at either 15˚C (the permissive temperature) or 25˚C (the restrictive temperature). Progeny were transferred from the initial to final temperatures at the stages shown; animals were grown to adults at the final temperature and stained with DiO (Materials and Methods). Percentages of animals having aberrant IL2 sensory processes are shown; mean +\- SEM. Early embryo, pre-comma stages; late embryo, past 3-fold stage (Sulston et al., 1983); L1, within 8 hours of hatching; l2/3, either of these larval stages. Circles, animals transferred from 25˚C to 15˚C; squares, animals transferred from 15˚C to 25˚C.

Done by: Dr. Ryder

**Protein structure showing the small and large repeat sequence**
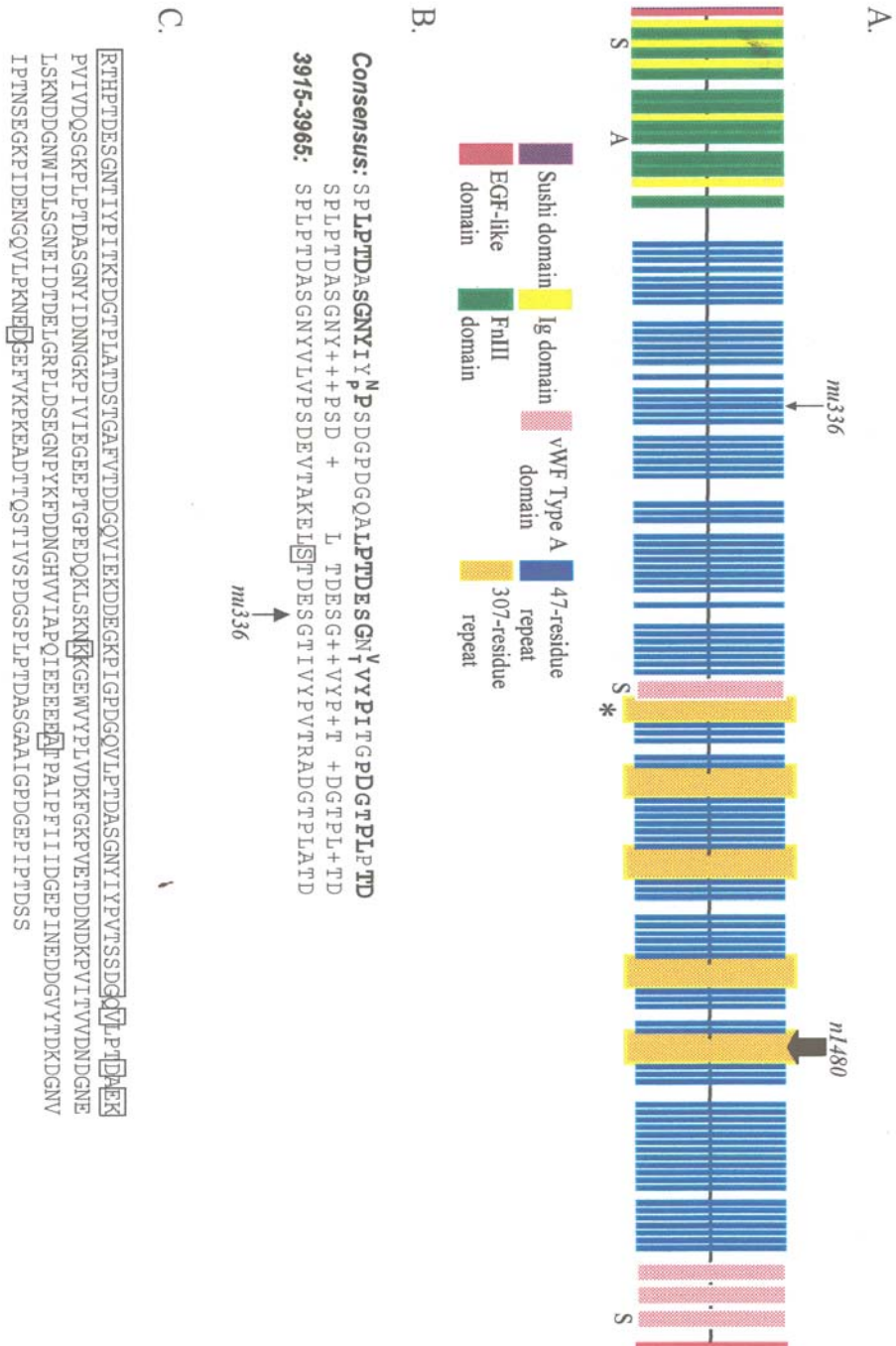


**Figure 6A. Structure of the predicted DIG-1 protein**. A. Schematic of the entire protein. Residues 1-2100 contain numerous Ig and FnIII domains. Residues 2300-12000 are highly repetitive. A 47-residue sequence (blue box) that contains the location of the

nu336 point mutation (indicated by small arrow) is repeated 93 times in this region with at least 41% identity and 55% similarity. A 307-residue sequence (gold box) that encompasses the *n1480* lesion (indicated by large arrow) is repeated with at least 98% identity 4 times and a smaller segment of the repeat occurs with 96% homology (indicated by a*). Residues 12000-13055 contain several von Willebrand Type A domains as well as two EGF domains. Schematic was constructed using information from the Pfam Protein Families Database (Bateman et al., 2002). "S" or "A" under a domain indicates that it was found only in the S.M.A.R.T. database or in the Pfam A data base, respectively.

B. The consensus sequence of the 47-residue repeat that occurs in the repetitive region of the protein is compared to the repeat containing the *nu336* point mutant (amino acid 3915-3965). The wild type version of this repeat is shown; in *nu336*, the arrowed serine is mutated to phenylalanine. Residues with>= 80% similarity among the 93 repeats are shown in bold; those with >=80% identity are shown larger than other residues. A small minority of repeats had two extra inserted residues not shown in the consensus sequence. The boxed serine differs in the 3915-3961 repeat segment but is a conserved proline in 95% of the repeats. + indicates the two amino acids are similar. The $^N_P$ symbol indicates that the residue is a non-polar residue (V,L,I, or P) in 55% of the repeats; in 45% of the repeats this residue is excluded. The $^V_T$ symbol indicates that valine and tyrosine each occur in 37% of the repeats.

C. The 307 residues that are deleted by the *n1480* lesion. The segment indicted by * contains a segment that is 96% homologous to the last 229 residues in the repeated sequence. Those residues not perfectly conserved in all five repeats are boxed.

Done by Christina Higgins and Dr. Ryder.

## Appendix B: Worm lines, Constructs, Primers and Restriction enzymes

## Worm lines

| Strain name | Genotype | Constructed from | Comments/strains made by CB unless noted |
|---|---|---|---|
| RY001 | *dpy-20(e1282es)mig-10(ct41)/dpy-20(e1282es)mig-10(ct41)* | | Double mutant for injections of mig-10. Diana Rusiecki |
| RY007 | *dig-1(nu336)*map1 | EMS mutagenesis | Non-suppressed line from suppressor screen. Contains a lethal, it appears suppressed |
| RY008 | *dig-1(nu336)*map2 | EMS mutagenesis | see RY007 |
| RY009 | *dig-1(nu336)*map3 | EMS mutagenesis | see RY007 |
| RY0010 | *lon-1(e185)unc-32(e189)nuIn10* | recombinant (rec) | Rec#2 from KP0951 w/o RFLP |
| RY0011 | *lon-1*(e185)*dig-1(nu336)*+/*lon-1(e189)*+*unc-32(e189)* | recombinant | Rec#9 from KP0951 w RFLP |
| RY0012 | *dig-1(nu336)/dpy-17(e164)unc-32(e189)* | KP0951 and SP471 | Second strain for isolation of recombinants |
| RY0013 | Never Frozen | | unc rec#5, died off |
| RY0014 | *dig-1(nu336) unc-32(e189)/ dpy-17(e164) unc-32(e189)* | recombinant | unc rec#12, het w respect to *dig-1(nu336)* |
| RY0015 | *dig-1(nu336) unc-32(e189)/ dpy-17(e164) unc-32(e189)* | recombinant | unc rec#18, het w respect to *dig-1(nu336)* |
| RY0016 | *dig-1(nu336) unc-32(e189)/dig-1(nu336) unc-32(e189)* | recombinant | unc rec#12, homoz w respect to *dig-1(nu336)* |
| RY0017 | *dig-1(nu336) unc-32(e189)/dig-1(nu336) unc-32(e189)* | recombinant | unc rec#18, homoz w respect to *dig-1(nu336)* |
| RY0019 | *dig-1(nu336)/dig-1(nu336)* | recombinant | Constructed by Dan Long. Also, stained for severity of neuronal phenotype. |
| RY0026 | *dpy-20 (e1282ts)*pMH86pPD95.69.12/*dpy-20 (e1282ts)* | Injection | Injected with pMH86 and pPD95.69.12 |
| RY0027 | *dpy-20 (e1282ts)*pMH86pPD95.69.12/*dpy-20 (e1282ts)* | Injection | See RY0026 |
| RY0028 | *dpy-20 (e1282ts)*pMH86pPD95.77.12/*dpy-20 (e1282ts)* | Injection | Injected with pMH86 and |

| | | | pPD95.77.12 |
|---|---|---|---|
| RY0029 | *dpy-20* (*e1282ts*)pMH86pPD95.69.12/*dpy-20* (*e1282ts*) | Injection | See RY0026 |
| RY0030 | *dpy-20* (*e1282ts*)pMH86pPD95.77.12/*dpy-20* (*e1282ts*) | Injection | See RY0028 |
| RY0031 | *dpy-20* (*e1282ts*)pMH86pPD95.77.12/*dpy-20* (*e1282ts*) | Injection | See RY0028 |
| RY0040 | *mig-10* (*ct41*)/*dpy-17*(*e164*) *unc-32*(*e189*) | Crossing | injected with *elt-2*::GFP, pDR98 and pGEM-Tconstruct4 |
| RY0054 | *mig-10* (*ct41*)/ *mig-10* (*ct41*)*elt-2*::GFP | injection | injected with *elt-2*::GFP. Formerly 1C |
| RY0055 | *mig-10* (*ct41*)/ *mig-10* (*ct41*)*elt-2*::GFP | injection | injected with *elt-2*::GFP. Formerly 2A |
| RY0056 | *mig-10* (*ct41*)/ *mig-10* (*ct41*)*elt-2*::GFPpDR98pGEM-tconstruct4 | injection | see RY0040 Formerly 7-3-15 |
| RY0057 | *mig-10* (*ct41*)/ *mig-10* (*ct41*)*elt-2*::GFPpDR98pGEM-tconstruct4 | injection | see RY0040 Formerly 7-3-24 |
| RY0058 | *mig-10* (*ct41*)/ *mig-10* (*ct41*)*elt-2*::GFPpDR98pGEM-tconstruct4 | injection | see RY0040 Formerly 7-3-4 |
| RY0059 | *mig-10* (*ct41*)/ *mig-10* (*ct41*)*elt-2*::GFPpDR98pGEM-tconstruct4 | injection | see RY0040 Formerly 7-3-31 |
| RY0060 | *mig-10* (*ct41*)/ *mig-10* (*ct41*)*elt-2*::GFPpDR98pGEM-tconstruct4 | injection | see RY0040 Formerly 7-3-50 |
| RY0061 | *mig-10* (*ct41*)/ *mig-10* (*ct41*)*elt-2*::GFPpDR98pGEM-tconstruct4 | injection | see RY0040 Formerly 7-3-48 |
| RY0062 | *mig-10* (*ct41*)/ *mig-10* (*ct41*)*elt-2*::GFPpDR98pGEM-tconstruct4 | injection | see RY0040 Formerly 7-3-49 |
| RY0063 | *mig-10* (*ct41*)/ *mig-10* (*ct41*)*elt-2*::GFPpDR98pGEM-tconstruct4 | injection | see RY0040 Formerly dpyunc-1 |
| RY004 | *mig-10* (*ct41*)/ *mig-10* (*ct41*)*elt-2*::GFPpDR98pGEM-tconstruct4 | injection | see RY0040 Formerly dpyunc-2 |
| RY0065 | *mig-10* (*ct41*)/ *mig-10* (*ct41*)*elt-2*::GFPpDR98pGEM-tconstruct4 | injection | see RY0040 Formerly 212A |

**Table 1B: Worm lines generated in the Ryder lab (RY) during research.**

| Strain name | Genotype | Used for | Comments |
|---|---|---|---|
| N2 | +/+ | Reference strain | Is wildtype for all genes |
| KP0951 | *dig-1(nu336)*nuIn 10/*lon-1(e185)unc-32(e189)* | Mutant line | Has point mutation as well as insertion of GFP. |
| KP0981 | *dig-1(n2467)* | Mutant line | Stained for severity of |

| | | | phenotype |
|---|---|---|---|
| KP1084 | *dig-1*(*nu345*) | Mutant line | Stained for severity of phenotype |
| KP1102 | *dig-1*(*n1321*) | Mutant line | Stained for severity of phenotype |
| KP1083 | *dig-1*(*n1480*) | Mutant line | Stained for severity of phenotype |
| SP0471 | *dpy-17*(*e164*) *unc-32*(*e189*) | To construct line RY0012 for isolation of recombinants | Crossed into KP0951 |
| CB1282 | *dpy-20*(*e1282ts*) | Injected with a wildtype copy of the gene for marker rescue to determine which worms contained the injected materials. | This is a temperature sensitive mutant. The non-permissive temperature is at 20°C or 25°C. |
| BW315 | *mig-10*(*ct41*) | crosses and injections | injected with elt-2::GFP |

**Table 2B: Worm strains maintained in the Ryder lab for research.**


**Plasmids**

| Plasmids isolated | Description | Provided by | Enzymes used to construct |
|---|---|---|---|
| pjP603 | Contains the *pie-1* gene, which is an embryonic lethal gene. Used to produce RNA, as a control for anti-sense injections. Ampicillin (amp) as a selection marker. | Mello lab (Worcester, MA) | N/A |
| pRF-4 | Contains dominant reporter gene *rol-6* (*su1006*), used as an injection control (amp) | Mello lab (Worcester, MA) | N/A |
| pSKII | Contains both T7 and T3 promoter for RNA production. (amp) | Stratagene | N/A |
| pPD95.77 | Contains the gene for green fluorescent protein without a native promoter. (amp) | Fire lab | N/A |
| pPD95.69 | Contains the gene for GFP and a nuclear localization sequence (NLS) without a native promoter. (amp) | Fire lab | N/A |
| pMH86 | Contains a wildtype copy of the *dpy-20* gene, which is used to rescue the mutant phenotype. Used as an injection marker. (amp) | M. Han | N/A |
| pET-15b | Expression vector, for the production of peptides for antibody production (amp) | Novagen | N/A |

| TOPO PCRII Vector | general T-tailed cloning vector | Invitrogen | N/A |
|---|---|---|---|
| pGEM-T | general T-Tailed cloning vector | Promega | N/A |
| pPD95.77.12 kb | Contains a 12 kb insert, that should represent the *dig-1* promoter. Also contains the gene for GFP. (amp) | Constructed by CB | BamHI, SalI |
| pPD95.69.12 kb | Contains a 12 kb insert, that should represent the *dig-1* promoter. Also contains the gene for GFP and a NLS. (amp) | Constructed by CB | BamHI, SalI |
| pET-15bexon#3 (Ant1Con1) | PCR product from a predicted exon, that is expressed. | Constructed by CB | NdeI, BamHI |
| pET-15b2526 | cDNA insert for protein production, cam25tabif,26bab | Constructed by CB | NdeI, BamHI |
| pET-15b5658 (Ant3Con3) | cDNA insert for protein production, cam56tab,58b | Constructed by CB | NdeI, BamHI |
| pSKII5657 | RNA probe construct containing a PCR product from cDNA of 56tab,58b; for detection of *dig-1*. | Constructed by CB | EcoRI, BamHI |
| pDR98 (*mig-10*::GFP) | Contains *mig-10* promoter + GFP | Constructed by DR | PstI, BamHI |
| pGEM®-t construct4 (GFP::*mig-10*) | Contains SOE construct 4 (made by Kevin Papenfus). Construct 4 has GFP plus the translated and un-translated region of *mig-10* | Constructed by CB | T-tailed vector no unique sites for cloning |
| pRT539 | Contains RT-PCR product from region 2 to region 3 | Constructed by CB | TOPO T-tailed vector. see molecular book 5 pages 1 and 3 |
| pRT185 | Contains RT-PCR product from region 1 to region 2 | Constructed by CB | TOPO T-tailed vector see pRT539 |
| pP48L1 | Contains PCR that spans the mutation in allele *n1480* | Constructed by CB | TOPO T-tailed vector see molecular book page 8 |
| pGT5c1 | Contains short cDNA product from K07E12.1 to K07E12.2 | Constructed by CB | TOPO T-tailed vector. see page 122 Molecular book 4 |
| pGT5c2 | Contains the long cDNA product from K07E12.1 to K07E12.2 | Constructed by CB | TOPO T-tailed vector. see pGT5c1 |
| pT14c1 | Contains the cDNA product from primer set 46.9t and K11c4.25b | Constructed by CB | TOPO T-tailed vector. see page 137 and 149 molecular book 4 |
| pT14c2 | Contains the short cDNA product | Constructed by CB | TOPO T-tailed |

| | | | |
|---|---|---|---|
| | from primer set K11c4.25t and 51b | | vector see page 137 and 149 molecular book 4 |
| pT14c3 | Contains the long cDNA product from primer set K11c4.25t and 51b | Constructed by CB | TOPO T-tailed vector see page 137 and 149 molecular book 4 |

**Table 3B: Plasmids obtained and constructed during research.**

**Primers**

| Primer name | Match site for *dig-1* | Primer sequence 5' TO 3' | Primer use |
|---|---|---|---|
| Cam4t | 454bp | <u>ACGCGTCGAC</u>TAGGGCGATCAAGAGACTAGAGG | To clone 5' regulatory sequence (promoter). LP,S,CB |
| Cam6t | 6,206bp | <u>ACGCGTCGAC</u>CTTTCATTCAGGAGCCATCATCC | See Cam4t |
| Cam8t | 8,617bp | <u>ACGCGTCGAC</u>CCAAACTTATCTGGCTAGAGAGG | See Cam4t |
| Cam11tko | 11,726bp | ACAATCTGATGTATCCGGAGCCG | To find a large deletion that was requested from the *C.elegans* gene knockout (ko) consortium. SP,LP,CB |
| Cam12tko | 12,101bp | GTTACAGCCCTAGCCATTGTTGG | See Cam11tko |
| Cam12.5tseq | 12,479bp | GTCGTCGTTGTGAGTATCGTTCC | To sequence the GFP constructs. CB |
| Cam12t | 12,631bp | GCATGTGGCTTGGTGAGGTAGTG | SP,LP,R,CB |
| Cam12tab | 12,631bp | <u>GGAATTCCATATG</u>GCATGTGGCTTGGTGAGGTAGTG | To clone cDNA for protein expression. R,N,CB |
| Cam12b | 12,653bp | <u>CGCGGATCC</u>CACTACCTCACCAAGCCACATGC | To clone the 5' regulatory sequence (promoter). LP,B,CB |
| Cam12tabif | 12,669bp | <u>GGAATTCCATATG</u>TGTATGCAAGTGGATCGTCGCCG | See Cam12tab (note: this one in frame for protein production) |
| Cam13022t | 13,022bp | <u>GGGAATTCCATATG</u>GTCCGAGCAATGGCATTTGTACC | To clone exon3 for protein production. SP,N,CB |
| Cam13t | 13,226bp | CCTCCAATGTGTCGTGGACAAGG | R,CB |
| Cam13346b | 13,346bp | <u>CGCGGATCC</u>CATCTAACGTCATTCCTCGTGGGC | To clone exon3 for protein production. SP,B,CB |

| Cam13bSL1 | 13,588bp | TCCTCCTGGAGCAACCTCATACG | Used with the SL1A primer to determine the 5' end of *dig-1*. R,CB |
|---|---|---|---|
| Cam14.04t | 14,044bp | CCAGATCTTCGCCCAGACACTGC | R,CB |
| Cam14.06b | 14,066bp | GCAGTGTCTGGGCGAAGATCTGG | R,CB |
| Cam14bab | 14,066bp | GCAGTGTCTGGGCGAAGATCTGG | To clone cDNA into pET-15b for protein production. R |
| Cam14t | 14,720bp | CCGGAATTCCAAGTGCTCCATTCACTGC | SP,LP,E,ER |
| Cam14b | 14,738bp | GTGAATGGAGCACTTGGAAGTCC | R,CB |
| Cam14bSL1 | 14,785bp | GTTCCAACTGGACAACAGGTGGC | See Cam13bSL1 |
| Cam15bko | 15,460bp | GATGGATCCACAGAAGGATCGGC | See Cam11tko |
| Cam15bko2 | 15,505bp | CGGAAGTGAATCGGTTTGAGTCGG | See Cam11tko |
| Cam15b | 15,724bp | ACGTGACGATCATCAACAGTTCC | R,CB |
| Cam16tant | 16,214bp | **ATTAACCCTCACTAAAGGA**CTCCAACTCAACCACTTGGACC | To make RNA for anti-sense injections. T3 promoter. SP, CB/ER |
| Cam16bant | 16,735bp | **TAATACGACTCACTATAG**TCTCATTTGGTGCTGTTCCAGG | To make RNA for anti-sense injections. T7 promoter. SP, CB/ER |
| Cam16b | 16,735bp | CCGGAATTCTCATTTGGTGCTGTTCCAGG | SP,LP,E,ER |
| Cam18t | 18,391bp | CAAGTGTCTTCTCCGAGCCATGC | R,CB |
| Cam19b | 19,346bp | TTGACTGTGACTGGCTCACTGGC | SP,P,CB |
| Cam20b | 20,132bp | GGATCTCCGTCTTCATTGACAGG | R,CB |
| Cam21t | 21,124bp | CCATTACCAACTGATGCCTCCGG | LP,CB |
| Cam22b1 | 22,190bp | TTGTCTGTAGGAAGTGGTGATCC | SP,LP,CB |
| Cam22b | 21,909bp | AAGGTGGAGTAGAGCCAGTGGGG | M,ER(not in use) |
| Cam23t | 23,502bp | CGTGGATCCCCTCTTCCCACTG | SP,LP,CB |
| Cam25t | 25,731bp | AGCTTTGCAACCGCCTTGAGACC | SP,LP,CB |
| Cam25tab | 25,261bp | GGAATTCCATATGGATGATTCCGGCGCAGTAATCGG | To clone cDNA into pET-15b for protein production. R,N,CB |
| Cam25b | 25,462bp | CCGGAATTCAGTTCCAAGTGGAGTACC | SP,LP,CB |
| Camsp26t | 26,075bp | TCCACTTCCAACGGATGCTTCGG | To sequence *dig-1(nu336)*. CB |
| Camsp26b | 26,509bp | GTGCCATCTGGACCGCGAACAGG | See Camsp26t |
| Cam26bab | 26,935bp | GTAGGATGGGATCCACGTACTGG | To clone cDNA into pET-15b. R,CB |
| Cam27b | 27,819bp | CCACTTATCTGGCTTTGTGAAC | SP,LP,CB |
| Cam28230b | 28,230bp | TCTCCACCTTCAGTTACGAAGGAC | R, CB |
| Cam30tabif | 30,510bp | GGAATTCCATATGCCTCTCGCAACTGATTCCACCG | To clone cDNA into pET-15b. R,N,CB |
| Cam30tab | 30,511bp | GGAATTCCATATGTCCTCTCGCAACTGATTCCACC | see Cam30tabif (note: this one is not in-frame) |
| Cam30t | 30,511bp | CGCGGATCCTCTCGCAACTGATTCCACC | SP,LP,B,ER |
| Cam31b | 31,482bp | AATCGTACCTGTCCATCTGGTCC | SP,P,CB |

| | | | |
|---|---|---|---|
| Cam32tant | 31,965bp | **ATTAACCCTCACTAAAGGA**TGTCGATAGAGACGACGAGC | See Cam16tant |
| Cam32bab | 32,088bp | <u>CGCGGATCC</u>ATCTGTCGGCAATGGCTTAGTCG | To clone cDNA into pET-15b. R,B,CB |
| Cam32bant | 32,411bp | **TAATACGACTCACTATA**GTCAGTTGGGATCGGCTTGCC | See Cam16bant |
| Cam32b | 32,411bp | <u>CCGGAATTC</u>AGTTGGGATCGGCTTGCC | SP,LP,E,ER |
| Cam37t | 37,235bp | <u>CGCGGATCC</u>ACTAGGTACCGATTCGAG | SP,LP,B,M,ER |
| Cam38b | 38,368bp | CCACTTGCATCAGTAGGAATTGG | SP,P,M,CB |
| NORPRO-BAMHI+38772t | 38,722bp | <u>CGCGGATCC</u>AATTGTAATTGAAGGAGAAAGAACC | SP,B,P,M,CH |
| Cam38.7t | 38,787bp | GGAGAAGAACCAACTGGACCAG | SP,LP,M,SH |
| NORPRO-EcorI+39435b | 39,435bp | <u>CCGGAATTC</u>CAGAAGCGTCAGTTGGCAATGGC | SP,E,P,M,CH |
| Cam39b | 39,840bp | <u>CCGGAATTC</u>AGTCGGAAGTGGAGATCC | SP,LP,E,M,ER (replaced by Cam40b) |
| Cam40b | 39,904bp | GCTGCAGATTCGTCAGATGGAACG | SP,LP,M,CB |
| Cam40t | 40,632bp | TCCCTACCGGCCTTATTGCCGTC | SP,LP,CB |
| Cam43b | 43,434bp | CTGAACCTTATCCTTCGGGTTGTG | SP,LP,CB |
| Cam44t | 44,132bp | <u>CGCGGATCC</u>AGTCTATCCAGTTCGTGG | SP,LP,B,ER |
| Cam46b | 46,236bp | <u>CCGGAATTC</u>CTGACTCATCTGTCGTCTC | SP,LP,E,M,ER |
| Cam46.5t | 46,524bp | GGACCTAATGGAGAACCGATCCC | To sequence *dig-1(n1480)*. SH |
| Cam46.9t | 46,980bp | CCTCTTTCAACCGATTCTACCGG | See Cam46.5t |
| Cam47.5tSEQ | 47,436bp | GACTAATAGCGGAAGTACGGCAG | See Cam46.5t. CB |
| Cam47.6t | 47,697bp | GGACTAGATGGACAAGCCCTGCC | See Cam46.5t |
| cam47.7b | | CCAGAGGCATCAGTTGGAAGCGG | |
| Cam47.9t | 47,901bp | TTAGTTCCTTCAGATGACGGTG | See Cam46.5t |
| Cam48t | 47,994bp | CTTGCCACAGACTCAACTGGCGC | See Cam46.5t |
| Cam48.5t | 48,528bp | GGACATGTTGTCATTGCTCCAC | See Cam46.5t |
| Cam48.2t | 48,258bp | GAAGGAGAAGAACCAACTGGAC | See Cam46.5t |
| Cam48.7b | 48,760bp | TGAGTTGTGTCAGCTTCCTTCG | See Cam46.5t |
| Cam48.905b | 48,905bp | AGAAGCGTCAGTTGGCAATGGCG | See Cam46.5t |
| Cam48.9b | 48,945bp | CCGTAACTCCTTCGCCTGATGG | See Cam46.5t |
| Cam48955b | 48,955bp | GGAAGTGAATCCGTAACTCC | |
| Cam49022b | 49,022bp | GTCGGTGGCAAGTAATGTTCC | |
| Cam49.1b | 49,185bp | CAGATTCATCAGTTGGCAAGGC | See Cam46.5t |
| Cam49.436b | 49,445bp | CGGTTCCAAGCGGTGTTCCG | See Cam46.5t |
| Cam50t | | CGACCAGACGGAACACCTCTCGG | |
| Can51b | 51,430bp | CCTAAAGCCTCGTCAGAAGGAAC | SP,LP,SH |
| Cam53t | 52,998bp | <u>CGCGGATCC</u>TGATGGTTCACTTCTTGG | SP,LP,B,ER |
| Cam55tant | 54,476bp | **ATTAACCCTCACTAAAGGA**AGATATCAACGGGAAGCCAGC | See Cam16tant |

| | | | |
|---|---|---|---|
| Cam54t | 54,845bp | CGGAGATGGAGCTGAAATCCCTG | SP,LP,R,CB |
| Cam55bant | 54,975bp | **TAATACGACTCACTATA**GTCAAGCGCTTCACGAAGACC | See Cam16bant |
| Cam55b | 54,976bp | <u>CCGGAATTC</u>AAGCGCTTCACGAAGACC | SP,LP,E,ER |
| Cam56t | 56,167bp | AGAGCTTCAGCGGATGAGTTCCG | R,CB |
| Cam56tab | 56,167bp | <u>GGAATTCCATATG</u>AGAGCTTCAGCGGATGAGTTCCG | To clone cDNA into pET-15b. R,N,CB |
| Cam56tBamHI | 56,167bp | <u>CGCGGATCC</u>AGAGCTTCAGCGGATGAGTTCCG | To clone cDNA into pSKII for RNA probes. SP,B,P,CB |
| Cam56752t | 56,752bp | TTCATCTGGACCAATGGTCTGCC | |
| Cam57084t | 57,084bp | TCGCAAGGAAGTCACATGCCGTG | |
| Cam57304t | 57,304bp | GAAGACTACGTTAAGCCGTCTCC | |
| Cam57b | 57,342bp | ACAGTTGCAGCTCGATGGAGA | SP,P,CB |
| Cam57bEcoRI | 57,177bp | <u>CCGGAATTC</u>CTCGTTAGTCATGACACAGACAG | To clone cDNA into pSKII for RNA probes. SP,E,P,CB |
| Cam58b | 57,940bp | AGCCATTTGAAGGCTCTTCGCAC | R,CB |
| Cam58.3t | | CATGCCTCTGGATATGCTGTCGC | R,CB |
| Camn1480M | | CTTCCACTAGAATTCTATTTTAC | R (mutant RNA), CB |

**Table 4B: Primers from *dig-1* used in PCR, RT-PCR, and sequencing.**
Amplifications: SP-short range PCR, LP-long range PCR, R-RT-PCR. Engineered digestion sites: B-BamHI, E-EcoRI, N-NdeI, S-SalI (underlined in primer sequences). Other: T7 and T3 promoter sequences are in bold. Primers with seq are sequencing primers. Primers with the M, mean that they have more then one hit site that is 100% and should be looked up in the nucleic acid folder (in the Ryder lab) for those additional sites. Primers with P were used to generate probe for either Southern analysis or northern analysis. In addition, the designers of the primers are included: CB-Chris Burket, ER-Dr. Ryder, SH-Stacy Hubbard, CH-Christina Higgins. If no initials, then that primer was not designed in the Ryder lab.

| Primer name | Primer sequence 5' to 3' | Primer use |
|---|---|---|
| NORPROBamHI+ACT2t | <u>CGCGGATCC</u>TAGAAGCACTTGCGGTGAACG | For making, Actin probe. SP,B,P,CH |
| NORPROEcorI+ACTb1117 | <u>CCGGAATTC</u>ACGTCATCAAGGAGTCATGG | For making, Actin probe. SP,E,P,CH |
| NORPRO-BamHI24048t | <u>CGCGGATCC</u>GGAACTTGTATGGCTTGTCG | For making *unc-89* probe. SP,B,P,CH |
| NORPRO-EcoRi+24800b | <u>CCGGAATTC</u>CTGACATCACAGAGAGCTTCC | For making *unc-89* probe. SP,E,P,CH |
| T7promoter primer | TAATACGACTCACTATAGGG | For screening antibody clones. |
| T7terminator | GCTAGTTATTGCTCAGCGG | See T7promoter primer |

| primer | | |
|---|---|---|
| GFPtop | TTAGATGGTGATGTTAATGGGCA | Screen potential expressing worm lines for the GFP construct. KP |
| GFPend | CTATTTGTATAGTTCATCCATGC | See GFPtop. |
| Cmo24 | TTGTAAAACGACGGCCAG | Control for anti-sense injections. |
| Cmo25 | GACCATGATTACGCCAAGC | See Cmo24 |
| K11C4.21t | TTGTGACACTTTCGATCTGGCTC | LP,CB |
| K11C4.27b | TGATGAGGGAACCGTTGCAGAGG | LP,CB |
| K11C4-25.4bseq | TCTTCGTGAATAGCTGAAGCCGC | To sequence the mutation in *dig-1*(*n1480*) that contains an insertion of cosmid K11C.4. |
| K11C4-25.4tseq | CTGCGGCTTCAGCTATTCACG | See K11C4-25.4bseq |
| SL1A | GGTTTAATTACCCAAGTTTGAG | SL1 Trans-spliced leader sequence. |
| SL1 | GTTTAATTACCCAAGTTTGA | SL1 Trans-spliced leader sequence. |
| pET-15bseq | ATGCCGGCCACGATGCGTCCGGC | To sequence expression constructs. CB |
| pGEM-Tseqb | ACAGCTATGACCATGATTACGCC | CB |
| mig-10promt | CGTTAATGATATTAGTGCCTTGCG | screen mig-10 injected animals CB |
| mig-10tranb | GGCTAGAAGAAGCTGGAATCCCG | see mig-10promt |
| K07E12.2t | TCGAGGAAAATGGTTCAAGGGTG | R, CB |
| R05H11.1b | GTCCCAATCCAACAACATCAGCC | R, CB |

**Table 5B: Primers used in PCR, RT-PCR, and sequencing; that were not from *dig-1*.** See table 4 for description, with the addition of KP-Kevin Papenfuss.


**Restriction enzymes**

| Enzyme name | Use | Obtained from |
|---|---|---|
| BamHI (BSA) | For cloning and diagnostic | New England Biolabs, Inc. (Beverly, MA) |
| EcoRI | For cloning and diagnostic | see BamHI |
| RsRII | Diagnostic | see BamHI |
| XhoI | Diagnostic | see BamHI |
| NdeI | Diagnostic | see BamHI |
| BglI | Diagnostic | see BamHI |
| SalI (BSA) | For cloning and diagnostic | see BamHI |
| PstI | Diagnostic | see BamHI |
| NcoI | Diagnostic | see BamHI |
| HindIII | Diagnostic | see BamHI |
| PpumI | Diagnostic | see BamHI |
| BsmI | Diagnostic | see BamHI |
| AvaII | Diagnostic | see BamHI |
| BclI | Diagnostic | see BamHI |

**Table 6B: Restriction enzymes used in research**

**Appendix C Web resources**


**http://wormbase.org/**
**http://www.sanger.ac.uk/**
**http://ncbi.nlm.nih.gov/Blast**
**http://elegans.swmed.edu/**