University of Massachusetts Medical School

eScholarship@UMMS

GSBS Dissertations and Theses   Graduate School of Biomedical Sciences

2014-04-29

POS-1 Regulation of Endo-mesoderm Identity in C. elegans: A Dissertation

Ahmed M. Elewa
University of Massachusetts Medical School

Let us know how access to this document benefits you.

Follow this and additional works at: https://escholarship.umassmed.edu/gsbs_diss

Part of the Cell Biology Commons, Cellular and Molecular Physiology Commons, and the Developmental Biology Commons

POS-1 REGULATION OF ENDO-MESODERM IDENTITY IN C. ELEGANS

A Dissertation Presented
By

Ahmed Elewa

Submitted to the Faculty of the
University of Massachusetts Graduate School of Biomedical Sciences, Worcester
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

April, 29th 2014

INTERDISCIPLINARY GRADUATE PROGRAM

POS-1 REGULATION OF ENDO-MESODERM IDENTITY IN C. ELEGANS

A Dissertation Presented
By

Ahmed Elewa

The signatures of the Dissertation Defense Committee signify completion and approval
as to style and content of the Dissertation

_____
Craig C. Mello, Ph.D., Thesis Advisor

_____
Sean P. Ryder, Ph.D., Member of Committee

_____
Albertha J.M. Walhout, Ph.D., Member of Committee

_____
(T. Keith Blackwell), Member of Committee

The signature of the Chair of the Committee signifies that
the written dissertation meets the requirements of the Dissertation Committee

_____
Joel D. Richter, Ph.D., Chair of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences signifies that
the student has met all graduation requirements of the school.

_____
Anthony Carruthers, Ph.D.,
Dean of the Graduate School of Biomedical Sciences

Interdisciplinary Graduate Program

April 29th, 2014

DEDICATION

for Kareem and Omar

ACKNOWLEDGEMENTS

Craig Mello, without your patience this would not have been possible.

Thank you for believing in me.

Masaki Shirayama, I could not have done this on my own. Thank you for making many experiments a reality.

Acknowledgments are due to the entire Mello lab for providing a vital environment for my doctoral research. In particular, I would to thank Don Gammon for his scientific camaraderie. Jean O'Connor provided administrative and moral support throughout my tenure. In genetic parlance, she has been a "Lethal Jean". Sandra Vergara joined our lab, and with that came an end to a long lab winter. She has been a continuous source of fun moments and generous celebrations. Rita Sharma, thank you for expressing a care that is reminiscent of that expressed by dear relatives. Takao Ishidate has been an inspiring example of poise; *Arigato*. Daniel Chaves helped me expand my social circle by introducing me to a wonderful group of UMass friends. Obrigado!

Elaine Youngman deserves special recognition. This dissertation ends with a chapter on *neg-1*. Elaine had called me out during lab meeting for my tongue-in-cheek indifference towards this gene. "But there is nothing known about it, nothing, and no homologues!" That was my excuse. "We spend our careers looking for something new instead of more of the same. You have a shot at something novel and you're throwing it away." That was her response. It was Elaine's interjection that helped me realize that I was before true novelty, for it was even beyond my own preconceived impressions of what qualifies as new. This incident has become part of my scientific being, along with an incident with Craig. When I was about to quit for the second time he told me,

"Ahmed, you have yet to taste the thrill of discovery. Before you leave I want you to taste it. I'll help you get there." A second round of acknowledgement is therefore in order. Thank you Craig.

Joel Richter has been both critic of my work and supporter of my career. Thank you for the delicate balance. Marian Walhout, without your recommendation I would not have been accepted to UMass. Thank you for your advice, encouragement and continued support. Sean Ryder interviewed for a faculty position when I first joined the graduate program and within a few years his graduate students made major contributions to the field of research I share. Thank you for illuminating the targets of POS-1, MEX-3 and MEX-5 and for helping me interrogate their binding to the *neg-1* 3'UTR. In this context, acknowledgment is due to Ebru Kaymak for teaching me how to do *in vitro* binding assays and for doing the second round of assays when I had to dedicate more time to writing my thesis.

UMass Medical School would not be what it is to me without the presence of Allan Jacobson and Victor Ambros. They represent to me the person I would like to become.

My parents have always been there for me and I am grateful for their unconditional support.

ABSTRACT

How do embryos develop with such poise from a single zygote to multiple cells with different identities, and yet survive? At the four-cell stage of the *C. elegans* embryo, only the blastomere EMS adopts the endo-mesoderm identity. This fate requires SKN-1, the master regulator of endoderm and mesoderm differentiation. However, in the absence of the RNA binding protein POS-1, EMS fails to fulfill its fate despite the presence of SKN-1. *pos-1(-)* embryos die gutless. Conversely, the RNA binding protein MEX-5 prevents ectoderm blastomeres from adopting the endo-mesoderm identity by repressing SKN-1. *mex-5(-)* embryos die with excess muscle at the expense of skin and neurons.

Through forward and reverse genetics, I found that genes *gld-3/Bicaudal C*, cytoplasmic adenylase *gld-2, cye-1*/Cyclin E, *glp-1/Notch* and the novel gene *neg-1* are suppressors that restore gut development despite the absence of *pos-1*. Both POS-1 and MEX-5 bind the 3'UTR of *neg-1* mRNA and its poly(A) tail requires GLD-3/2 for elongation. Moreover, *neg-1* requires MEX-5 for its expression in anterior ectoderm blastomeres and is repressed in EMS by POS-1. Most *neg-1(-)* embryos die with defects in anterior ectoderm development where the mesoderm transcription factor *pha-4* becomes ectopically expressed. This lethality is reduced by the concomitant loss of *med-1*, a key mesoderm-promoting transcription factor.

Thus the endo-mesoderm identity of EMS is determined by the presence of SKN-1 and the POS-1 repression of *neg-1*, whose expression is promoted by MEX-5. Together they promote the anterior ectoderm identity by repressing mesoderm differentiation. Such checks and balances ensure the vital plurality of cellular identity without the lethal tyranny of a single fate.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

Tae-Ho Shin uncovered the *gld-3* suppression of the *pos-1* and *gld-1* gutless phenotype. GLD-3 had been identified by Tae-Ho due to its interaction with PIE-1 via Yeast-2-Hybrid. He later isolated an allele of *gld-3* (*alp-1(ne157)*) as a gene required for <u>A</u>symmetric <u>L</u>ocalization of <u>P</u>-granules. However, this strain was lost. In the context of studying the role of GLD-3 in blastomere specification, the genetic interactions with *pos-1* and *gld-1* were discovered.

Sandra Vergara conducted a mutagenesis screen that isolated two temperature sensitive dominant maternal effect sterile hermaphrodites. Since this phenotype, sans temperature sensitivity, was reminiscent of *gld-3/alp-1(ne157)*, I tested if the two mutants were *pos-1* suppressors by feeding them *pos-1(RNAi)*. One of the two mutants laid dead eggs with restored endoderm, evident by gut granules. Craig Mello predicted the mutation was in the *glp-1* locus since a fraction of embryos died with Glp pharynxes. Takao Ishidate confirmed the presence of a lesion by sequencing the *glp-1* coding region.

Meetu Seth performed the antibody staining of Cyclin E as part of her rotation, and Darryl Conte brought my attention to the RPL15 homology with NEG-1.

C13F10.7 was identified as an interactor of NEG-1 in collaboration with Alex Tamburino (Laboratory of Marian Walhout, University of Massachusetts Medical School, MA).

Cell-lineaging of *neg-1(tm6077)* was done in collaboration with Zhou Du (Laboratory of Zhirong Bao, Memorial Sloan Kettering Cancer Center, NY).

Behavioral assays were done in collaboration with Chris Chute (Laboratory of Jagan Srinivasan, Worcester Polytechnic Institute, MA).

In vitro binding assays between *neg-1* 3'UTR and MEX-3, MEX-5 and POS-1 were done in collaboration with Ebru Kaymak (Laboratory of Sean Ryder, University of Massachusetts Medical School, MA).

Extension polyA tail assays were done in collaboration with Traude Beilharz (Monash University, Australia)

GFP tagged *neg-1* worm strains and *neg-1* 3'UTR GFP reporter strains were engineered in collaboration with Masaki Shirayama.

Portions of this work have been presented publically as:

Elewa, A., Shirayama, M., Miliaras, N., Shin, T., & Mello, C. (2010). **gld-3 functions in blastomere specification.** *C. elegans: Development and Gene Expression, EMBL, Heidelberg, Germany.*

Elewa, A., Ishidate, T., Vergara, S., Shin, T., Shirayama, M., & Mello, C. (2012). **Germline expressed GLP-1 regulates embryonic endoderm specification.** *C. elegans: Development, Cell Biology and Gene Expression, University of Wisconsin – Madison, USA.*

Elewa, A., Shirayama, M., Vergara, S., Ishidate, T., & Mello, C. (2013). **Suppressors of *pos-1* identify a novel function for GLP-1 and new players, gld-3, cyclin E and *spos-1*, involved in endoderm specification**. *International Worm Meeting, University of California – Los Angeles, USA*

Background information on Cyclin E has appeared in:

Ishidate T, Elewa A, Kim S, Mello CC, Shirayama M. **Divide and differentiate: CDK/Cyclins and the art of development.** *Cell Cycle* 2014; 13:0 - -1;

Chapter I

Introduction to *C. elegans*

early embryogenesis

The question of blastomere identity determination

In the perennial quest to discover the logic of life the nematode has been a tool and the embryo an ancient tomb. The nematode embryo provides the matter needed to approach a central question of biology; **how a single cell becomes an adult**. Early inquiries into "nemic embryogenesis" noted the regularity of nematode embryo development (Chitwood and Chitwood, 1937). However, it was the contribution of Theodor Boveri in the late 19th century and the publication of the embryogenesis of *Parascaris equoram* (*Ascaris megalochephala*) that laid the foundation of modern nematode embryology by establishing the fact that nematode embryogenesis is invariant and occurs through determinate cleavages (*Ibid*. p. 216). This means that cells are predestined, or programmed, to adopt an identity and then divide and branch into a lineage that forms particular organs or body parts.

The motif that Boveri identified in early nematode embryogenesis is the asymmetric cleavage of a parental germinal cell (P) into an anterior somatic stem cell (S) and posterior parental germinal cell (P). The somatic stem cells generated by this motif produce the blastomeres required for the developing body. S1 cleaves into blastomeres A and B, which form the greater part of the ectoderm and contribute to the formation of the epithelium, pharynx, nervous system and excretory system. S2 cleaves into blastomeres E, which is the sole source of endoderm, and MSt, which in turn divides into blastomeres M for mesoderm and St for stomodeum in the nematodes that have one. S3 is the blastomere responsible for ectodermal epithelium of the posterior part of the body, whereas S4 is a blastomere that provides the embryo with the rectum and rectal glands as well as some muscular tissue. In some nematodes, P4 divides asymmetrically one more

time to generate the final somatic stem cell S5 that forms the epithelium of the gonoducts, and P5. However, in most nematodes P4 gives rise to the germ cells designated with perpetuating the motif in the following generation by maturing into a gamete, sperm or oocyte, which upon fertilization is anointed P0 (*Ibid*. pp. 216-217).

Such an invariable course of development raises the question of **how blastomeres are assigned a particular "identity" at the beginning embryogenesis**. The question in this form is a reiteration of the central question above, how does a single cell divide and differentiate to become an adult?

C. elegans as a model to study blastomere fate determination

In order to dissect the genetic requirements for development and behavior, *Caenorhabditis elegans* was elected by Sydney Brenner as a model organism (Brenner, 1974, 2003). *C. elegans* early embryogenesis follows the main course outlined by Boveri in *Parascaris equoram* (**Figure 1.1**) (Schierenberg, 2006; Sulston et al., 1983). Few important exceptions are worth noting. Anatomically, *C. elegans* does not have a stomodeum, however the designation EMSt was retained as EMS despite the absence of the "S" warranting body part. Additionally, the A and B blastomeres are designated ABa and ABp respectively. The designations S1 through S4 are substituted with AB, EMS, C and D in that order. Deviations not withstanding, *C. elegans* remains faithful to the Boverian account of nematode embryogenesis according to *P. equaroam*, including which body parts are derived from which blastomeres.

C. elegans early embryogenesis, blastomeres and their fates

zygote

2 cell

AB

4 cell

EMS

C

Anterior
Pharynx

Posterior
Pharynx

Hypodermis & Muscle

Intestine

Germline

Adult

Neurons

Hypodermis & Neurons
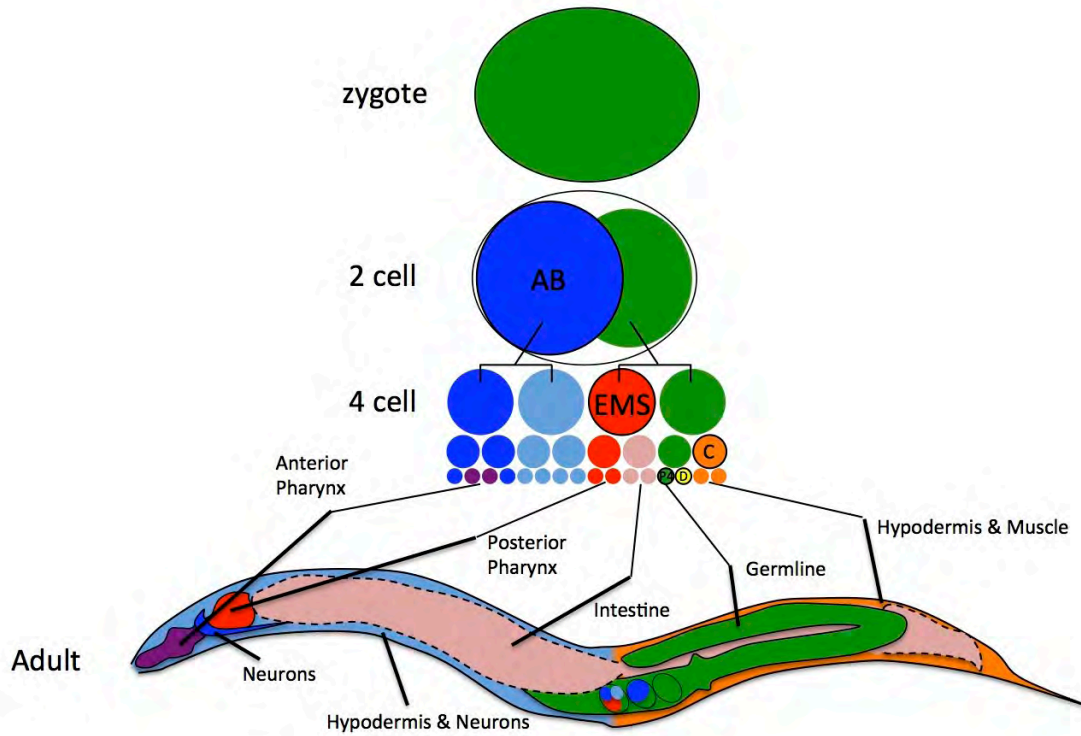
Figure 1.1

Figure 1.1. Blastomere identity and fate is determined in *C. elegans*. The following schematic will be used to represent the early embryo in C. elegans. Circles represent blastomeres and their sizes are reduced to reflect cell cleavage. Colors represent developmental fates.

Two approaches have been used to study the basis of blastomere identity specification in *C. elegans*. The first approach involves separating blastomeres to determine if their identity is arrived at independently in isolation or is informed by a neighboring cell. The second approach is the pursuit of mutants in which blastomeres adopt the wrong identity and identifying the mutated gene responsible for the monstrosity.

Cell interactions are required for blastomere fate specification

The *C. elegans* zygote is able to develop to completion without an accommodating uterus. However, the fates of AB and EMS do not proceed as expected when cultured in isolation or when particular neighboring cells are ablated (Priess and Thomson, 1987). This indicates that cell interactions are important for part of early embryogenesis. Following are the three main cell-cell induction events that pattern the early embryo.

P2 to ABp signaling specifies a blastomere that generates hypodermis and neurons

The AB blastomere is the stem cell for the primary ectoderm. The two daughter cells born after AB division have different developmental fates. The anterior sister ABa produces the anterior pharynx made of muscle, gland, structural cells and neurons (Figure 1.2). On the other hand, the posterior sister ABp gives rise to hypodermal cells (skin) and neurons. This difference in developmental fate is contingent upon ABp receiving an inductive signal from its neighbor P2 (Priess, 2005; Priess and Thomson, 1987). When cultured in isolation, AB divides into two identical sisters that both unroll ABa lineages. The developmental consequence of this ABp to ABa transformation is the doubling of anterior pharynx at the expense of hypodermal cells and neurons. Mutations in a gene

encoding the Delta ligand of Notch signaling, *apx-1*, cause a similar Anterior Phayrnx in eXcess phenotype (Apx). Obstructing the function of *glp-1*/Notch during the 4-cell stage using temperature sensitive alleles interrupts the P2 to ABp signaling resulting in an Apx phenotype as well **(Figure 1.2)** (Mello et al., 1994; Mickey et al., 1996). Notch signaling is therefore required for P2 to ABp cell-cell induction.

MS to AB great granddaughter signaling specifies anterior phayrnx

The Notch signaling pathway is required for a second inductive event that takes place between the MS blastomere and two granddaughters of ABa, ABalp and ABara **(Figure 1.3)** (Hutter and Schnabel, 1994; Priess and Thomson, 1987). If this inductive signal is obstructed, either by culturing ABa in isolation or as a consequence of *glp-1* loss of function, ABa produces neurons and hypodermis but no pharyngeal cells (Priess and Thomson, 1987; Shelton and Bowerman, 1996). The developmental consequence of losing this second Notch inductive signal is a posterior pharynx derived from MS but no anterior pharynx due to the misspecification of ABa. Such a partial pharynx is referred to as Glp since it is arrived at by temperature sensitive *glp-1* alleles originally isolated due to their role in Germ Line Proliferation **(Figure 1.3)** (Austin and Kimble, 1987; Priess et al., 1987).

P2 to EMS signaling specifies endoderm blastomere

When cultured in isolation, EMS divides into two identical sister cells that adopt an MS fate; mesoderm and no endoderm **(Figure 1.4)**. Endoderm differentiation originating from the E blastomere is restored if EMS is brought into contact with P2 (Goldstein, 1992). Mutagenesis screens pursuing More Of Mesoderm (Mom) mutants reveal the role of *mom-2*/Wnt and *src-1*/Src signaling in P2 to EMS signaling (Bei et al.,

2002; Rocheleau et al., 1997; Rocheleau et al., 1999; Thorpe et al., 1997). Briefly, MOM-2/Wnt from the P2 cell signals EMS through the MOM-5/Frizzled receptor. This signaling event acts redundantly with SRC-1 to orient the EMS division axis and to promote cortical-release and nuclear translocation of WRM-1/β-catenin into the posterior nucleus of the dividing EMS that will become the nucleus of the E blastomere (Nakamura et al., 2005). High WRM-1 levels in the E nucleus modulate POP-1 function to promote endoderm differentiation. POP-1 is an HMG-box containing protein belonging to the TCF/LEF family of transcription factors. Loss of *pop-1* function results in both EMS daughters adopting an E fate (Lin et al., 1995).

P2 to ABp signaling and the Apx phenotype
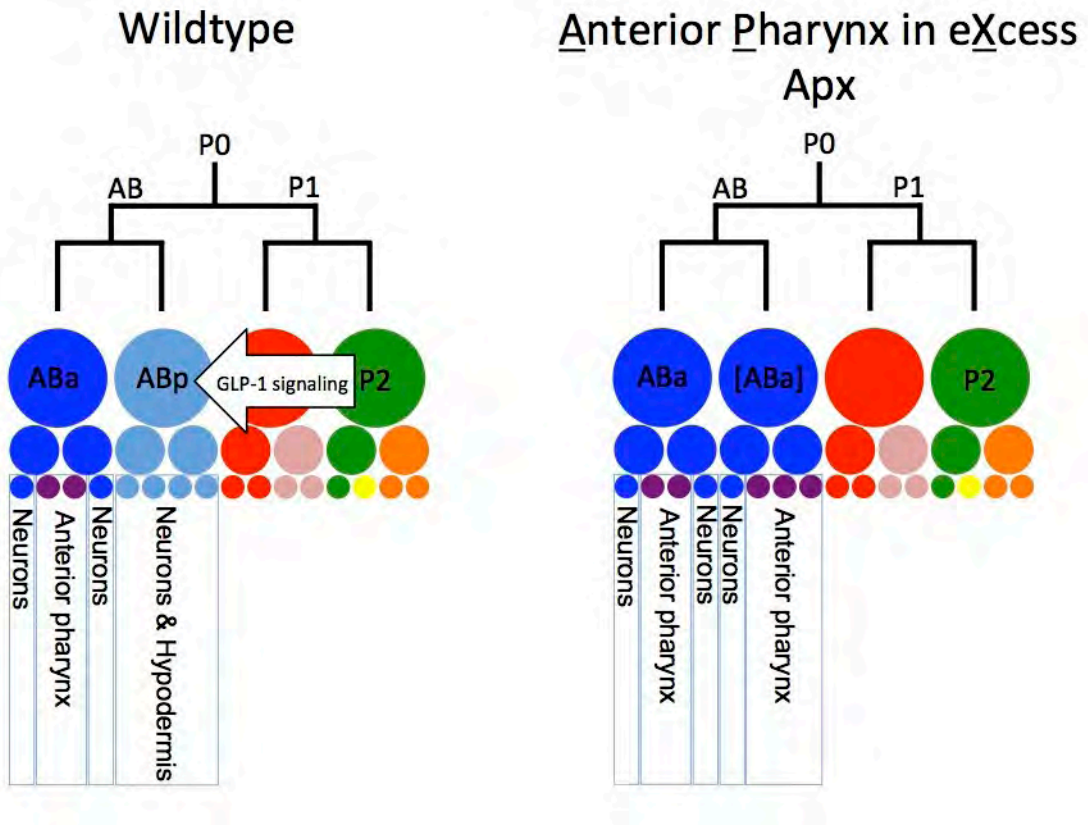
Figure 1.2

Figure 1.2. P2 to ABp signaling is necessary for the later to differentiate into hypodermal and neuronal cells (ectoderm). In the absence of this signaling, ABp remains identical to its sister ABa and response to inductive signals activating its mesodermal differentiation into anterior pharynx.

MS to AB great-granddaughter (AB8) signaling and the Glp phenotype
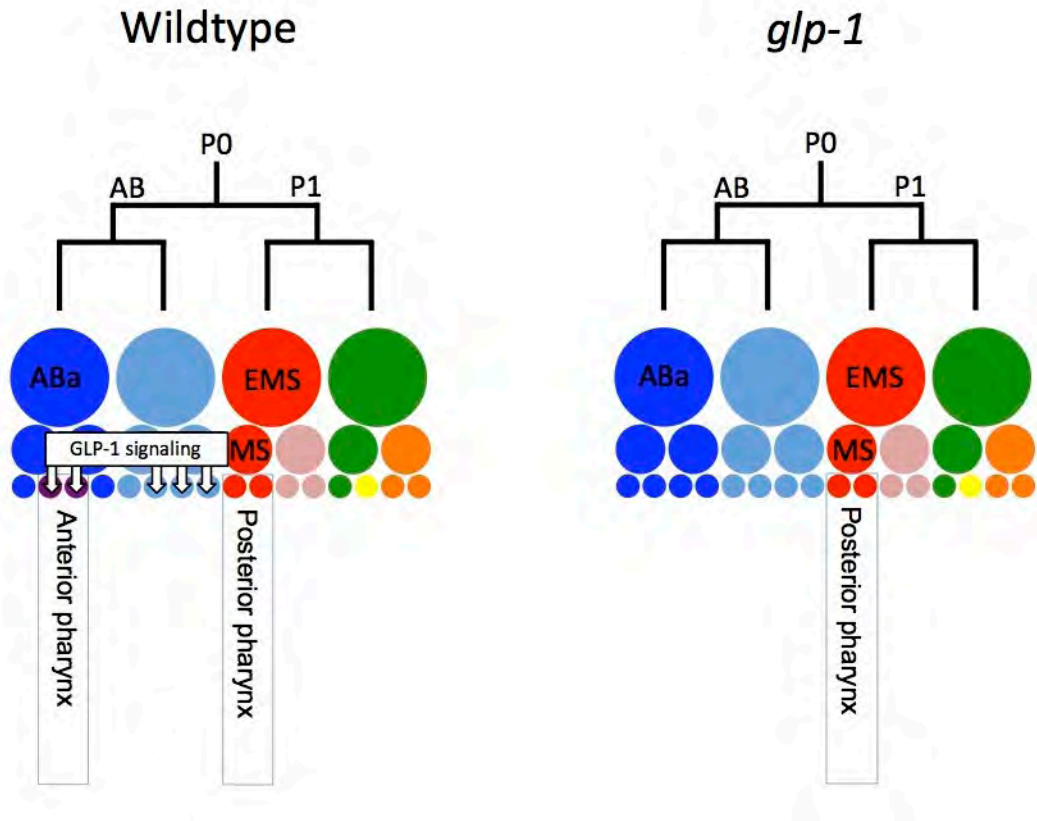
Figure 1.3

Figure 1.3. MS signals several of the 8 AB great granddaughters. However, only the descendants of ABa are responsive and in turn differentiate to build the anterior pharynx. In the absence of the signaling no anterior pharynx develops (Glp phenotype).

P2 to EMS signaling and the Mom phenotype
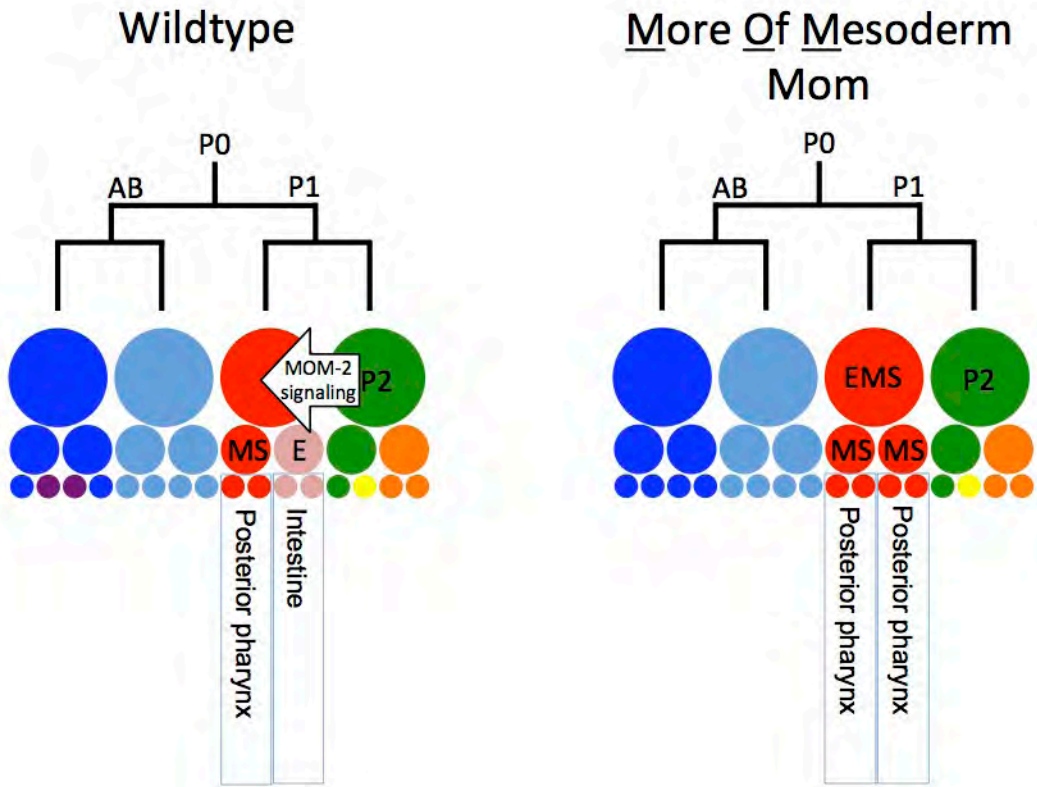
Figure 1.4

Figure 1.4. EMS divides asymmetrically into Endoderm (E) and Mesoderm (MS) blastomeres due to P2 signaling. In the absence of the Wnt/Src signaling input from P2, EMS divides into two mesoderm blastomeres (MS). This is the More Of Mesoderm phenotype Mom.

Polarity axis formation is required for blastomere fate specification

The first cell cleavage is asymmetric, with a larger anterior blastomere (AB) differentiating into a somatic fate and a smaller posterior blastomere (P1) retaining germline potential (Sulston et al., 1983). Thus the asymmetry is of size and identity. The asymmetry in identity is an outcome of selective segregation of differentiation determinants to the appropriate daughter cell. Key loci for such determinants are P-granules, which segregate to the posterior blastomere and are maintained in subsequent P blastomeres (Strome, 2005). Mutations in a series of genes named *par-1* through *par-6* disrupt the initial asymmetric division with subsequent disruption of cell size, division axis, or PARtitioning of developmental determinants (Gonczy and Rose, 2005; Kemphues et al., 1988).

par-4 is of particular interest in the context of blastomere specification. Embryos produced by homozygous *par-4* mothers fail to produce intestinal cells. This means that the E blastomere fails to specify its endoderm fate. Using a temperature sensitive allele suggests that the critical period during which *par-4* is required for the E blastomere specification begins during oogenesis, about 90 minutes before fertilization, and ends before the four-cell stage. In this sense, *par-4* helps set up the requirements for the endoderm generating blastomere well before it is even born (Morton et al., 1992).

Master Transcription Factors are required for blastomere fate specification

As reviewed above, mutagenesis screens illuminated the genetic basis for the necessary cell-cell signaling interactions that dictate the fates of ABalp, ABara, ABp and E. Moreover, it became clear that the initial asymmetrical division is necessary for subsequent blastomeres to adopt the correct fate. Additional studies further illuminated

15

the processes that determinate blastomere fate both positively (induction of fate) and negatively (preventing alternative fates). An important characteristic of these mutants is that they do not appear to function within signaling pathways and do not disrupt cell polarity and division asymmetry.

Skn

SKiNhead mutants *skn-1* and *skn-2* (later to be named *pos-1*) are recessive maternal effect embryonic lethal (Bowerman et al., 1992; Tabara et al., 1999). Heterozygous worms develop and grow to fertile adults. Homozygous progeny also develop and grow to fertile adults. However, these homozygous worms fail to provide their progeny with maternal doses of *skn-1* and *skn-2* mRNA leading to embryonic lethality. Embryos of *skn-1 -/-* or *skn-2 -/-* worms fail to specify the EMS blastomere leading to an embryo without EMS driven endoderm and mesoderm **(Figure 1.5)**. In the case of *skn-1*, EMS transforms and adopts the fate of the C blastomere (Bowerman et al., 1992). The name *SKiNhead* refers to the excess C derived hypodermis observed in this mutant. The focus of this thesis is *skn-2/ pos-1* and will be addressed in detail below.

skn-1 encodes a protein that is asymmetrically distributed in the early embryo **(Figure 1.5)** (Bowerman et al., 1993). At the two-cell stage, SKN-1 is detected in both nuclei but accumulates to much higher levels in P1. The asymmetry is emphasized in the four cells stage where SKN-1 accumulates in EMS and P2 compared to ABa and ABp.

Sequence analysis reveals that SKN-1 is an atypical bZIP transcription factor that lacks a leucine zipper dimerization segment and has an N-terminal arm related to those of homeodomain proteins. Further highlighting the hybrid nature of SKN-1, in vitro studies

demonstrated that SKN-1 recognizes a bZIP half-sites adjacent to 5'AT-rich sequence in the minor groove recognized by homeodomain proteins (Blackwell et al., 1994).

pal-1

pal-1 is a second transcription factor that functions as a fate determinant in early embryogenesis. Originally identified as a homeodomain transcription factor required for male tail development (Waring and Kenyon, 1991). *pal-1* mRNA was later demonstrated to be asymmetrically distributed in the 4-cell stage embryo in a manner similar to *skn-1* transcripts **(Figure 1.5)** (Hunter and Kenyon, 1996). Loss of both maternal and zygotic *pal-1* function in early embryos results in C blastomere misspecification into unfamiliar cells **(Figure 1.5)** (Hunter and Kenyon, 1996). Loss of zygotic *pal-1* function alone results in non-viable L1 larvae with gross posterior defects (Edgar et al., 2001).

How are SKN-1 and PAL-1 asymmetrically distributed and how are their activities manifest in EMS and C, respectively but barred from P2? Answers to these questions became evident upon the characterization of two other categories of blastomere specification mutants.

The transcription factors *skn-1* and *pal-1* specify blastomere fate
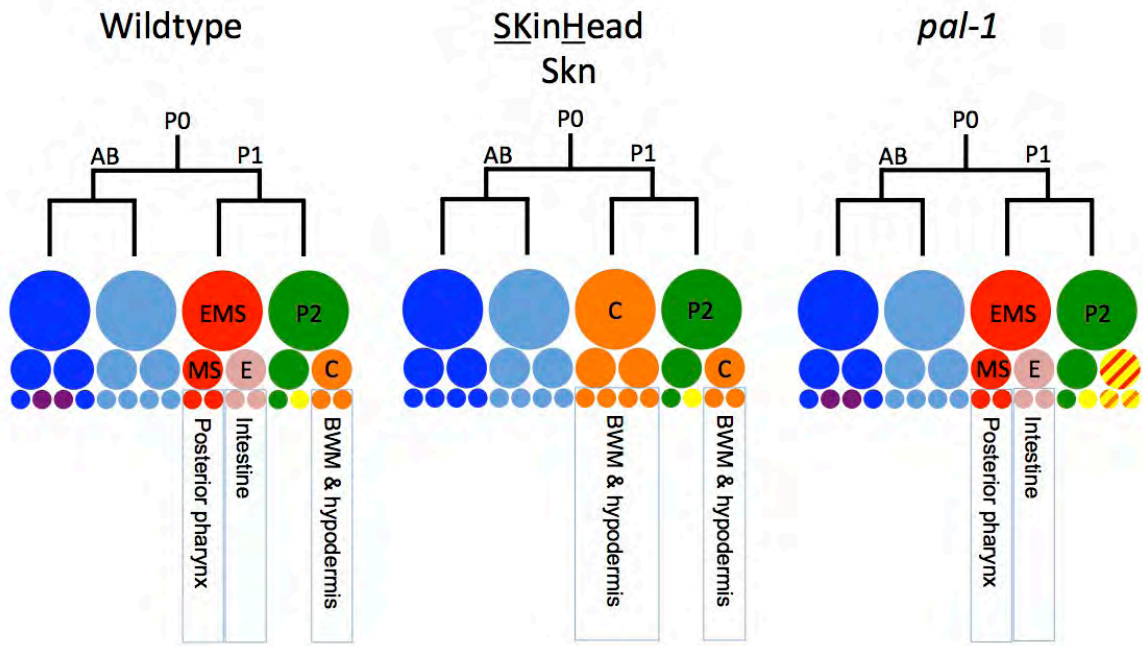
Figure 1.5

Figure 1.5. In *skn-1* mutants, EMS adopts the C fate and develops into hypodermal cells. In *pal-1* mutants, the C fate fails to materialize.

RNA binding proteins are required for Blastomere Specification in *C. elegans*

Pie

Pharynx and Intestine Excess (Pie) mutant embryos arrest with a P2 to EMS transformation leading to double the portion of endo-mesoderm at the expense of C derived hypodermis and body wall muscle, D derived muscle and P4 derived germline **(Figure 1.6)** (Mello et al., 1992). The effect of *pie-1* is recessive maternal effect embryonic lethal. PIE-1 protein is restricted to the P lineage of the early embryo. The P2 to EMS transformation of *pie-1* mutants is suppressed by *skn-1* loss of function (Mello et al., 1992). This genetic interaction uncovered the repressive nature of PIE-1 function in the P lineage in order to maintain germline identity at the expense of somatic differentiation. Although SKN-1 protein is present in the P2 blastomere (Bowerman et al., 1993), it is unable to initiate endo-mesoderm differentiation in the presence of PIE-1. Inquiry into the function of PIE-1 revealed its ability to inhibit zygotic transcription by regulating the activity of RNA Polymerase II (Batchelder et al., 1999; Mello et al., 1996; Seydoux and Dunn, 1997; Seydoux et al., 1996; Zhang et al., 2003).

PIE-1 belongs to a family of proteins with two tandem zinc fingers (TZFs) first discovered due to their early response to mammalian growth factor and insulin (DuBois et al., 1990; Gomperts et al., 1990; Lai et al., 1990). The hallmark of this family is the presence of two tandem zinc fingers of the C(X8)C(X5)C(X3)H format, where X is any amino acid (Mello et al., 1992). The discovery of PIE-1 and its requirement in embryogenesis came hand in hand with that of another TZF required for blastomere specification, *mex-1*.

Mex

In wildtype embryos, MS primarily produces mesodermal cells that build the posterior pharynx and contribute to body wall muscle. In *mex-1* loss of function embryos, the four AB granddaughters adopt the MS fate as well **(Figure 1.6)** (Mello et al., 1992; Schnabel et al., 1996). The consequence of this blastomere fate misspecification is the Muscle EXcess Mex phenotype. As is the case with *pie-1*, reducing *skn-1* function suppresses the ectopic MS differentiation in *mex-1* embryos (Mello et al., 1992; Schnabel et al., 1996).

mex-5 and *mex-6* encode TZFs. Like *mex-1*, *mex-5* also bars SKN-1 activity in the AB lineage. *mex-5(-)* embryos arrest with terminal phenotype similar to that of *mex-1* with AB granddaughters transforming into MS blastomeres leading to multiple pharyngeal clusters and excess muscle **(Figure 1.6)**. *skn-1* reduction tames the monstrosity. On the other hand, loss of *mex-6*, which shares 75% identity with *mex-5* has no consequences on embryonic development. Loss of both *mex-5* and *mex-6* functions, however, result in a severe loss of differentiation as germ cells, anterior muscles and intestine are all absent **(Figure 1.6)**. This suggests that *mex-6* does indeed contribute to blastomere fate specification along side with *mex-5* (Schubert et al., 2000).

Two transcription factors belong to this category of Mex mutants. *mex-2* and *mex-4* were later renamed *efl-1* and *dpl-1*, respectively (Page et al., 2001). Although both transcription factors are expressed and act in the maternal germline, their loss of function results in embryos exhibiting the Mex phenotype. This connection between activity in the maternal germline and effect in the subsequent embryos reflects how embryogenesis is a process that is prepared for well before fertilization. In the case of *efl-1* and *dpl-1*, their

activity induces the transcription of maternal genes including *mex-5*. Sufficient levels of *mex-5* transcripts are thus deposited in the embryo and function to bar posterior determinants such as *skn-1* from the AB lineage (Page et al., 2001).

Ectopic *skn-1* expression in the AB lineage is one way to transform anterior blastomeres to posterior muscle producing stem cells. Misexpression of *pal-1* is a second way. Loss of *mex-3* function results in ectopic *pal-1* expression in the AB lineage and a consequent transformation to C fate **(Figure 1.6)**. This defect is reversible by the loss or reduction of *pal-1* function (Draper et al., 1996; Hunter and Kenyon, 1996). MEX-3 is a KH domain containing RNA binding protein (Draper et al., 1996).

pos-1 and *gld-1*

pos-1 was formerly referred to as *skn-2* due to its loss of muscle development in a manner reminiscent to *skn-1* (Craig Mello personal communication). However, *pos-1* exhibits additional phenotypes that reflect an overall defect in the P lineage (Tabara et al., 1999). In *pos-1* embryos, the P1 blastomere divides into defective P2 and EMS daughter cells **(Figure 1.6)**. P2 fails to signal ABp leading to an Apx phenotype in two out of three embryos. EMS divides to produce two daughters, of which E fails to produce endoderm and MS exhibits signs of mesoderm development but overall fails to display a proper MS lineage. The consequence of this mediocre MS specification is the occasional loss of the second Notch signaling to ABa granddaughters ABalp and ABara. In one out of three embryos, this induction is absent and the embryo develops with no pharyngeal tissue, for MS does not produce the posterior pharynx and does not signal the ABa lineage to produce the anterior part. On the other hand, in cases where MS does succeed in

signaling the ABa lineage an Apx embryo results. In addition to the defects mentioned, *pos-1* embryos also lack germ cells due to a misspecification of the P4 blastomere. Overall, cells of the P lineage exhibit abnormal division rates and it seems that only C and D specify properly **(Figure 1.6)**. However this has not been rigorously tested.

gld-1 was originally isolated as a gene required for germline development (Francis et al., 1995a; Francis et al., 1995b; Jones et al., 1996). *gld-1* homozygous worms are sterile since germ cells fail to enter and properly proceed through meiosis. Knockdown of *gld-1* mRNA has the same effect. However, in the interim between injecting *gld-1* dsRNA and the onset of sterility, three consecutive types of embryos are produced. The first escapes *gld-1* RNAi entirely. The second hatches and grows to become sterile adults. The third type arrests as embryos exhibiting the *pos-1* phenotype **(Figure 1.6)** (Jones et al., 1996)(Craig Mello personal communication).

POS-1 belongs to the TZF family as is the case with PIE-1, MEX-1, MEX-5 and MEX-6 (Tabara et al., 1999). GLD-1 on the other hand affiliates with KH domain proteins such as MEX-3 (Jones and Schedl, 1995).

TZF and KHD proteins required for blastomere specification

Wildtype

Pharynx and Intestine Excess
pie-1

P2 to EMS
misspecification

Muscle EXcess
mex-1/5

AB4 to MS
misspecification

Muscle EXcess
mex-3

AB8 to C
misspecification

Pos
pos-1 or *gld-1*

No anterior
pharynx

Gutless,
No Germline

Anterior
Pharynx
in eXcess

Gutless,
No Germline
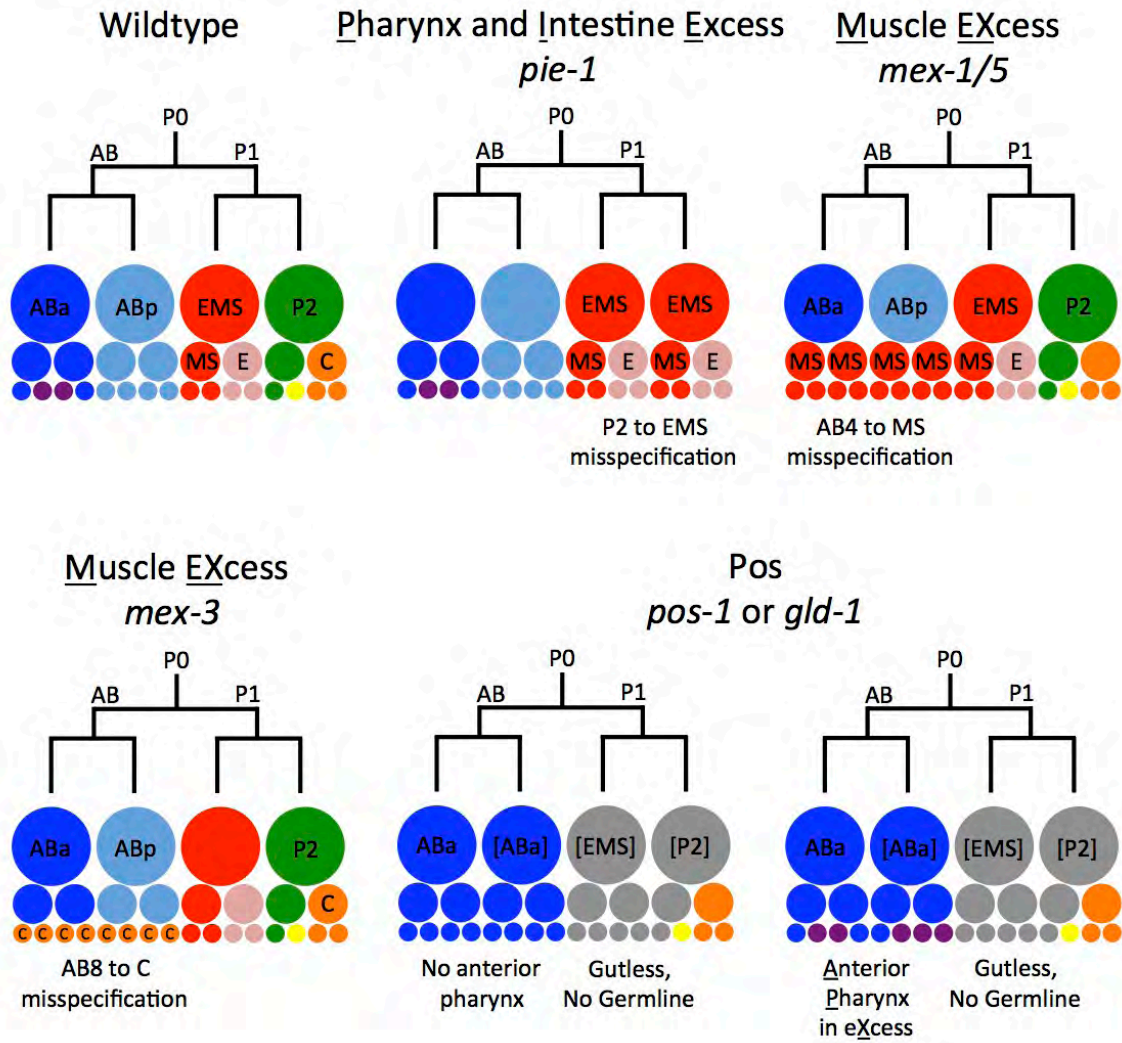
Figure 1.6

Figure 1.6. A number of Tandem Zinc Finger (TZF) and K-Homology Domain (KHD) proteins are required for blastomere specification. Mutations in the genes encoding these proteins or knockdown of their expression, lead to characteristic embryonic lethal phenotypes.

Table 1.1 Summary of embryonic phenotypes.

Gene	Phenotype	Reference	Notes
glp-1	ABp ➔ ABa	Mello *et al* 1994, Mickey *et al*. 1996	Second defect includes loss of anterior pharyngeal induction by MS
apx-1	ABp ➔ ABa	Mello *et al* 1994, Mickey *et al*. 1996.	
mom-2	E ➔ MS	Bei *et al*., 2002; Rocheleau *et al*., 1997; Rocheleau *et al*., 1999; Thorpe *et al*., 1997	
pop-1	MS ➔ E	Lin *et al*., 1995	
pie-1	P2 ➔EMS	Mello *et al*., 1992	
mex-1	AB ➔ 2EMS	Mello *et al*., 1992	
skn-1	EMS ➔ C	Bowerman *et al*., 1992	
pal-1	C ➔ X	Hunter and Kenyon, 1996	
mex-5	AB ➔ 2EMS	Schubert *et al*., 2000	
mex-6	None	Schubert *et al*., 2000	Additional defects when combined with *mex-5(-)*
mex-3	AB ➔8C	Draper *et al*., 1996; Hunter and Kenyon, 1996	
pos-1	ABp ➔ABa, EMS ➔X, P4 ➔X	Tabara *et al*., 1999	
gld-1	ABp ➔ABa, EMS ➔X, P4 ➔X	Jones *et al*., 1996, Craig Mello unpublished results	

TZF and KH proteins that function in blastomere specification are asymmetrically expressed

The embryonic determinants surveyed above are asymmetrically expressed in the early embryo. MEX-5 and MEX-6 are segregated to the anterior of the zygote and in the nascent AB daughter cell (Schubert et al., 2000). This localization is dependent on upstream PAR proteins that function to polarize the zygote after fertilization. MEX-5/6 also exclude the germplasm TZFs PIE-1, POS-1 and MEX-1 from the anterior thus enriching them in the posterior half of the zygote and subsequent P blastomeres (DeRenzo et al., 2003; Schubert et al., 2000). This negative interaction between MEX-5/6 on the one hand and PIE-1, POS-1, and MEX-1 on the other requires the protein ZIF-1, which interacts with the zinc fingers of the germplasm TZFs to mediate their destruction (Cuenca et al., 2003).

MEX-3 is also asymmetrically expressed in the anterior, whereas GLD-1 is restricted to the P lineage and blastomeres EMS, C and D. No GLD-1 is detected in the AB lineage by immuno-staining (Jones et al., 1996).

TZF and KH proteins that function in blastomere specification bind 3'UTRs and regulate expression

Our knowledge of the function of TZF proteins begins with the discovery of the mammalian protein Tristetraprolin (TTP) and its ability to bind AU rich regions in 3'UTRs (Brooks and Blackshear, 2013). Delineating the binding specifies of RNA binding proteins and RNA sequences involves two different approaches. The RNA-centered approach is a reiterative process in which the equilibrium dissociated constant

(K_d) of the interaction between purified recombinant protein and RNA is measured. This measurement is repeated after permutating the RNA to determine the contribution of each nucleotide in each position to the binding. A second protein-centered approach infers the binding site from aligning sequences of bound RNA captured with the protein of interest.

A combination of both approaches derived a conservative consensus sequence for GLD-1, UACU(C/A)A and a relaxed consensus (U>A/C/G)A(C>A)U(C/A>U)A (Ryder et al., 2004). A more complicated quantitative code was inferred in which heptamers were categorized as 'strong', 'medium' and 'weak' GLD-1 binding motifs **(Table 1.2)** (Wright et al., 2011).

MEX-5 was deduced to have a relaxed RNA-binding specificity requiring at least six uridines within a 9 to 13 nucleotide window (Pagano et al., 2007). Starting with a 12 nucleotide fragment recognized by POS-1, the consensus sequence of this protein was deduced to be $UA(U_{2-3})(A/G)(A/G/U)(N_{1-3})G$, where 'N' is any nucleotide **(Table 1.2)** (Farley et al., 2008). The MEX-3 consensus sequence was derived from non-biased *in vitro* selection experiment (SELEX) followed by the identification of 12-nucleotide sequence sufficient for MEX-3 binding. This sequence was then mutated to deduce that the MEX-3 consensus sequence was (A/G/U)(G/U)AGU(U/A/C)UA. Further analysis using 'spacing' mutants revealed that MEX-3 binds to a bipartite recognition element, which can be spaced from each other by up to eight nucleotides. Thus, the consensus sequence of MEX-3 has been derived as $(A/G/U)(G/U)AGN_{0-8}U(U/A/C)UA$ **(Table 1.2)** (Pagano et al., 2009).

Table 1.2: Consensus sequences of RBPs

Protein	Consensus binding sequence	Reference	Notes
GLD-1	Conservative: `UACU(C/A)A` Relaxed:`(U>A/C/G)A(C>A)U(C/A>U)A`	Ryder *et al*., 2004 Wright *et al*., 2011	See Wright *et al*., for alternative approach
MEX-3	`(A/G/U)(G/U)AGN`$_{0-8}$`U(U/A/C)UA`	Pagano *et al*., 2009	SELEX based
MEX-5	`≥75% U in N`$_{9-13}$	Pagano *et al*., 2007	ARE based
POS-1	`UA(U`$_{2-3}$`)(A/G)(A/G/U)(N`$_{1-3}$`)G`	Farely *et al*., 2008	*glp-1* 3'UTR based

POS-1 as negative regulator of expression (repressor)

The POS-1 recognition element is overrepresented in the 3'UTRs of early embryonic transcripts (Farley et al., 2008). Experimental evidence establishes that indeed POS-1 regulates the expression of a number of maternal transcripts in the embryo. Interestingly, POS-1 functions as both repressor and an activator, depending on the transcript in question.

Repression of glp-1

Embryonic expression of *glp-1* is restricted to anterior blastomeres (Evans et al., 1994). GLD-1 and POS-1 are required for the repression of *glp-1* in posterior blastomeres via their interaction with the spatial control region (SCR) of the 3'UTR (Marin and Evans, 2003; Ogura et al., 2003). The SCR consists of two POS-1 binding elements flanking a GLD-1 binding site (Farley and Ryder, 2012).

SPN-4 is required for *glp-1* expression in anterior blastomeres (despite the fact that SPN-4 is localized to posterior blastomeres). SPN-4 binds to a subregion of the temporal control region (TCR) in the *glp-1* 3'UTR (Ogura et al., 2003). SPN-4 and POS-1 physically interact suggesting that the switch from repression to expression may hinge upon the interaction of these proteins.

Interestingly, POS-1 binding to its recognition elements displaces GLD-1 interaction with the SCR. However, both repress *glp-1* translation (Farley and Ryder, 2012). A possible utility for this is that POS-1 repression substitutes GLD-1 repression in the early embryo. The difference between the two repressions is that POS-1 could be responsive to SPN-4 derepression (via physical interaction), but perhaps GLD-1 is not. In

this sense, *glp-1* expression goes from "off" (GLD-1 repression) to "locked" (POS-1 repression) to "unlocked" (SPN-4 derepression).

An alternative mode of countering POS-1 repression could be through MEX-5 and MEX-6, which are required for *glp-1* expression in the anterior blastomeres (Schubert et al., 2000). *glp-1* 3'UTR includes predicted MEX-5 binding regions some of which have been experimentally verified in the TCR (Pagano et al., 2007). It is possible that MEX-5/6 promote the expression of *glp-1* in the anterior by countering the waning repressive grip of GLD-1 and POS-1. However, this remains to be demonstrated.

Repression of zif-1

zif-1 is another gene with an embryonic expression pattern reminiscent of *glp-1* (Oldenbroek et al., 2012). ZIF-1 is a SOCS-box protein that interacts with the E3 ubiquitin ligase subunit elongin C and is required for CCCH finger protein degradation. In the early embryo, ZIF-1 is activated in somatic daughter cells where it binds to germ-plasm TZFs (PIE-1, MEX-1 and POS-1) to mediate their destruction (DeRenzo et al., 2003). This contributes to the proper segregation of these blastomere fate determinants. On the other hand, these targeted proteins repress the translation of *zif-1* in the posterior (Oldenbroek et al., 2012). Such mutual negative regulation contributes to the robust and rapid asymmetry observed.

Using a fragmenting approach the contribution of separate 3'UTR regions to *zif-1* expression was determined and the recognition of these regions by RBPs was tested (Oldenbroek et al., 2012). Removal of Region II lead to loss of *zif-1* expression in the anterior and ectopic expression in the posterior, which is a pattern opposite to that of wildtype embryos. This is interpreted to mean that region II is required for both the

repression in the posterior and the expression in the anterior. POS-1 and MEX-5 bind to region II. Loss of *pos-1* function results in ectopic posterior expression of *zif-1* whereas loss of *mex-5* results in loss of anterior expression. Epistatic interaction suggests that MEX-5 promotes expression by countering POS-1 repression (Oldenbroek et al., 2012). This is reminiscent of the mechanism probably regulating *glp-1* anterior expression.

POS-1 as positive regulator of expression (derepressor/activator)

Derepression of mom-2

mom-2 offers an opposite example to *glp-1* and *zif-1* since its expression is restricted to the posterior (Oldenbroek et al., 2013). Using the same fragmenting approach used to study the *zif-1* 3'UTR, POS-1 was found to bind to region 4. Loss of *pos-1* function results in loss of *mom-2* expression in posterior blastomeres (Oldenbroek et al., 2013). PIE-1 and MEX-1 are also required for *mom-2* expression and bind to region 3. Importantly, SPN-4 is required for *mom-2* repression and binds to region 3 as well. However, PIE-1 and MEX-1 effectively compete with SPN-4 binding of region 3 proposing a mechanism for posterior expression. Although POS-1 cannot compete with SPN-4 binding, the loss of *mom-2* expression upon *pos-1* loss of function suggest that POS-1 derepresses *mom-2* expression (Oldenbroek et al., 2013).

Derepression of nos-2

Embryonic *nos-2* expression is repressed until the P3 blastomere is born. There, and then in P4, *nos-2* is expressed and is required for germ line specification (Subramaniam and Seydoux, 1999). POS-1 binds the 3'UTR of *nos-2* and counters the repressive effect of MEX-3 and SPN-4 (D'Agostino et al., 2006; Jadhav et al., 2008).

Introduction to Research Question

The study of early embryogenesis in *C. elegans* has matured towards an investigation into the regulated expression of maternally provided mRNA to orchestrate the establishment of blastomere identity in the embryo. These blastomeres then activate appropriate genetic programs and unroll a series of cell divisions and differentiate into lineages that interweave and complement one another. The genes surveyed above encode RNA binding proteins that regulate the expression of maternal mRNAs, transcription factors that initiate genetic programs, or signaling molecules that allow blastomeres to confirm their identity *vis-a-vis* neighboring cells.

At this stage in our study of the embryo using *C. elegans* we lack a mechanistic understanding of how expression is modulated by the binding of TZF or KH domain proteins to the mRNA 3'UTRs. Complicating the answer is the fact that the same RNA binding protein could potentially act as a repressor or derepressor (activator) of expression, depending on the transcript and the cellular milieu in which the interaction occurs. The regulation is context dependent. Further challenging the quest is the fact that factors like PIE-1 can act at both the levels of transcription and translation.

As a contribution to this research tradition I have studied *pos-1*. Of the defects manifest in *pos-1* embryos, the loss of gut is paramount. This reflects a loss of endo-mesoderm differentiation in the absence of *pos-1*. Thus the question I set out to answer is the following: **Which maternal mRNA transcripts are regulated by *pos-1* in order to ensure proper endo-mesoderm development and how does *pos-1* regulate these transcripts?**

Chapter II

Identifying suppressors

of

pos-1

PREFACE

Tae-Ho Shin uncovered the *gld-3* suppression of the *pos-1* and *gld-1* gutless phenotype. GLD-3 had been identified by Tae-Ho due to its interaction with PIE-1 via Yeast-2-Hybrid. He later isolated an allele of *gld-3* (*alp-1(ne157)*) as a gene required for Asymmetric Localization of P-granules. However, this strain was lost. In the context of studying the role of GLD-3 in blastomere specification, the genetic interactions with *pos-1* and *gld-1* were discovered.

Sandra Vergara conducted a mutagenesis screen that isolated two temperature sensitive dominant maternal effect sterile hermaphrodites. Since this phenotype, sans temperature sensitivity, was reminiscent of *gld-3/alp-1(ne157)*, I tested if the two mutants were *pos-1* suppressors by feeding them *pos-1(RNAi)*. One of the two mutants laid dead eggs with restored endoderm, evident by gut granules. **Craig Mello** predicted the mutation was in the *glp-1* locus since a fraction of embryos died with Glp pharynxes. **Takao Ishidate** confirmed the presence of a lesion by sequencing the *glp-1* coding region.

Meetu Seth performed the antibody staining of Cyclin E as part of her rotation, and **Darryl Conte** brought my attention to the RPL15 homology with NEG-1.

Portions of this work have been presented publically as:

Elewa, A., Shirayama, M., Miliaras, N., Shin, T., & Mello, C. (2010). **gld-3 functions in blastomere specification.** *C. elegans: Development and Gene Expression, EMBL, Heidelberg, Germany.*

Elewa, A., Ishidate, T., Vergara, S., Shin, T., Shirayama, M., & Mello, C. (2012).

Germline expressed GLP-1 regulates embryonic endoderm specification. *C. elegans: Development, Cell Biology and Gene Expression, University of Wisconsin – Madison, USA.*

Elewa, A., Shirayama, M., Vergara, S., Ishidate, T., & Mello, C. (2013). **Suppressors of *pos-1* identify a novel function for GLP-1 and new players, *gld-3*, cyclin E and *spos-1*, involved in endoderm specification**. *International Worm Meeting, University of California – Los Angeles, USA*

Background information on Cyclin E has appeared in:

Ishidate T, Elewa A, Kim S, Mello CC, Shirayama M. **Divide and differentiate: CDK/Cyclins and the art of development.** *Cell Cycle* 2014; 13:0 - -1;

ABSTRACT

Mutations in the TZF encoding gene *pos-1* result in maternal-effect embryonic lethal phenotypes due to the misspecification of several embryonic cell fates, including the failure to specify endoderm. In order to address the mechanism underlying this loss of endoderm, we conducted RNAi and genetic screens for factors that restore endoderm specification in *pos-1* mutants. RNAi screening identified *gld-3* (Bicaudal C), *gld-2* (a cytoplasmic polyadenylase), *cye-1* (Cyclin E) and the novel gene F32D1.6 (which we have named *neg-1* "Negative Effect on Gut"). Forward mutagenesis screens identified an allele of *glp-1* (Notch). In each case, suppression resulted in restored endoderm and pharyngeal-mesoderm differentiation.

POS-1 has been shown to repress the translation of the *glp-1* mRNA in posterior sister cells at the 4-cell stage of embryogenesis. However, GLP-1, a transmembrane receptor related to Notch, is not required for endoderm specification, and its mis-expression in early *pos-1* embryos has not been linked to the loss of endoderm fate. Interestingly, the *glp-1(ne4298ts)* allele we have identified has a lesion that alters a conserved amino acid in the 4th swi6 motif resulting in a strong temperature sensitive *glp-1* loss of function phenotype. Surprisingly, the temperature-sensitive period (tsp) for *glp-1* suppression of *pos-1* occurs prior to fertilization, indicating that *glp-1(ne4298ts)* alters a maternal function that influences endoderm specification several hours later during early embryogenesis.

GLD-3 has been proposed to act as an RNA-binding cofactor that recruits GLD-2 to stimulate the cytoplasmic polyadenylation of targets important for germline development. We therefore investigated if they regulate the expression of *cye-1* or *neg-1*

37

to affect endoderm specification during embryogenesis. Whereas embryonic levels of *cye-1* were not affected by in absence of *gld-3*, *neg-1* expression was lost. We therefore proceeded to analyze the regulation of *neg-1* and its effect on embryogenesis.

The *pos-1* phenotype

All progeny of *pos-1* homozygous hermaphrodites (hereafter *pos-1* embryos) die with a compound phenotype **(Figure 1.6)**.

Loss of Germ cells

The *C. elegans* zygote divides asymmetrically into a somatic daughter and a daughter cell retaining germline identity called P1. The same asymmetric division occurs in P1, P2 and P3. In wildtype worms, P3 divides into a somatic daughter D and the germ blastomere P4. P4 divides into two germ cells Z3 and Z4 that give rise to the germline. In *pos-1* embryos, P4 identity is lost and the two germ cells Z3 and Z4 are absent (Tabara et al., 1999).

Excess or loss of anterior pharynx

In the absence of maternally provided *pos-1*, P2 fails to signal ABp via the Notch/GLP-1 signaling pathway. In accordance with this defect, *glp-1* and its ligand *apx-1* are misexpressed. In *pos-1(-)* 4-cell stage embryos, GLP-1 is detected in all cells as opposed to the anterior ABa and ABp alone. On the other hand, APX-1 is undetected (Ogura et al., 2003). Failure of P2 to ABp signaling renders the ABp granddaughter cells responsive to mesoderm-inducing signals from the MS blastomere leading to an excess of anterior pharynx (Mello et al., 1994; Mickey et al., 1996). This is the case in the majority of *pos-1(-)* embryos. However, in one out of three *pos-1(-)* embryos no anterior pharynx develops due to failure of MS induction (Tabara et al., 1999).

EMS hypo-specification

The MS failure to induce anterior pharyngeal development echoes a profound

prior defect in EMS specification **(Figure 2.1)**. In *pos-1* embryos, the daughters of EMS fail to differentiate as expected, with 99% of embryos completely defective in endoderm development despite presence of SKN-1 and the fulfillment of P2 to EMS signaling hallmarks (Tabara et al., 1999). While EMS fails to differentiate to mesoderm and/or endoderm, it does not adopt a C-like lineage and fate. Instead, the defect in EMS specification resembles one of hypo-specification rather than misspecfication to an alternative fate. What is the proximate cause for EMS hypo-specification?

Whereas SKN-1 levels are normal in *pos-1* mutant embryos, we found that the downstream target *med-1* fails to maintain robust expression. *med-1* transcription is dependent on SKN-1 and is sustained by its own autoregulation (Maduro et al., 2007). To address the possibility that weak *med-1* expression is the beginning of EMS failure in *pos-1* mutants, we examined the effect of upregulated *med-1* expression using a multicopy *med-1::gfp* strain (*med-1(+++)*). 83% of embryos produced by *pos-1(-); med-1(+++)* developed endoderm tissue reflecting restored EMS identity.

pos-1 is necessary for EMS specification and gut development
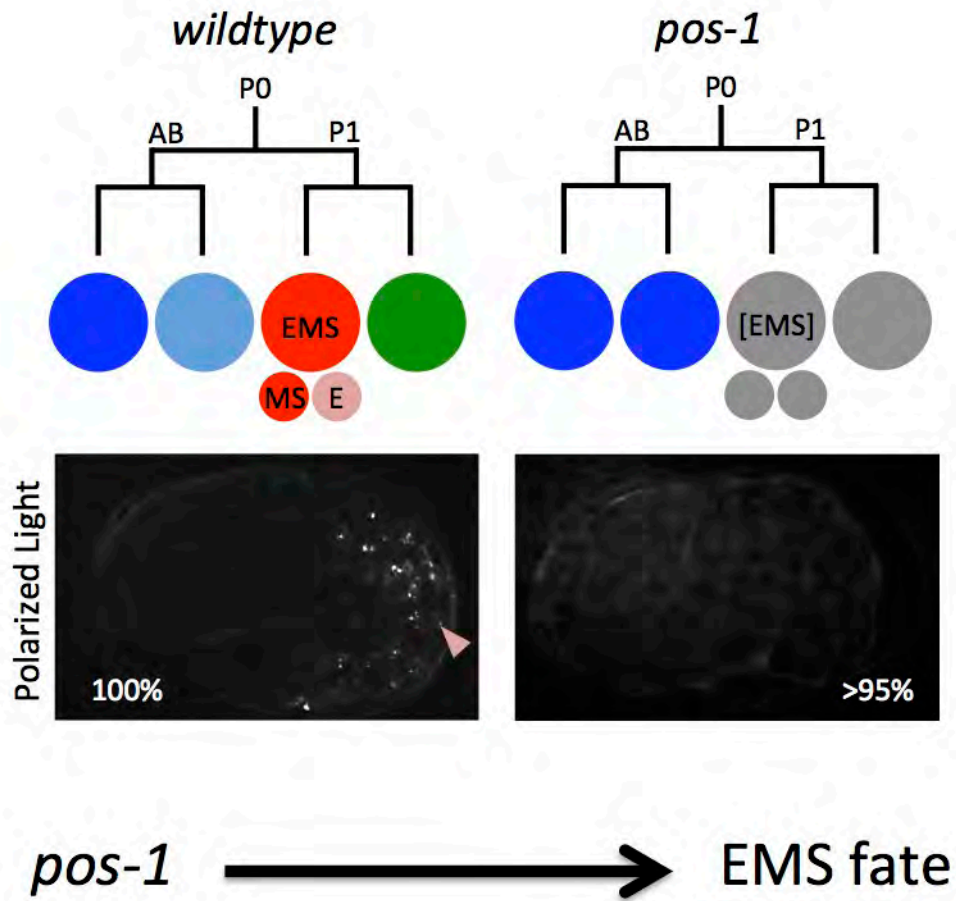
Figure 2.1

Figure 2.1. Embryos of *pos-1(-)* worms die gutless. The EMS blastomere fails to divide into the MS and E blastomeres (red in wildtype and grey in *pos-1*). *pos-1(-)* embryos therefore do not have birefringent gut granules (pink triangle). Another defect in *pos-1* embryos is the failure of ABp differentiation (light blue circle in wildtype, blue in *pos-1(-)*). This is due to impaired cell-cell interaction between P2 and ABp. The developmental consequence of this defect is an unenclosed embryo, which is difficult to discern under polarized light.

gld-1 phenocopies *pos-1*

The involvement of several TZF and KH proteins in blastomere fate specification suggested that other members of these families might be involved as well. Prior to the onset of *gld-1(RNAi)* sterility, a number of embryos are produced, the terminal batch of which die. The terminal phenotype of these *gld-1* embryos is identical to *pos-1* (Craig Mello unpublished observations).

gld-3 and *gld-2* suppress *pos-1*

In a systematic survey of genetic interactions between TZF and KH genes conducted by Tae-Ho Shin (see Preface), our lab found that concomitant loss of *gld-3* function using RNAi restored endoderm specification in the progeny of *pos-1(-)* worms (>95% **Figure 2.2**). *gld-3 (RNAi)* also restored gut development in *gld-1(RNAi)* embryos (Craig Mello personal communication). *gld-3* encodes a KH domain protein homologous to Bicaudal C and required for proper germline development (Eckmann et al., 2002). Since GLD-3 recruits the cytoplasmic poly-A polymerase GLD-2 to promote the polyadenylation and translation of target transcripts required for spermatogenesis (Eckmann et al., 2004; Wang et al., 2002), we tested if *gld-2* also suppresses *pos-1*. We fed *pos-1(zu148)* homozygous worms *gld-2* RNAi food and found that 46% of *pos-1; gld-2(RNAi)* dead embryos had developed endoderm.

Pursuing the proximate cause for loss of EMS specification in *pos-1*

The penetrant loss of EMS fate in *pos-1* and *gld-1* mutants and the dramatic restoration upon *gld-3* suppression allude to an elegant regulation of EMS fate specification. However, the aforementioned genes encode RNA binding proteins with likely roles in translation regulation. Therefore, my analysis focused on identifying the

transcripts that are evidently crucial for EMS fate specification and that rely on *pos-1* and

gld-3 for their proper expression.

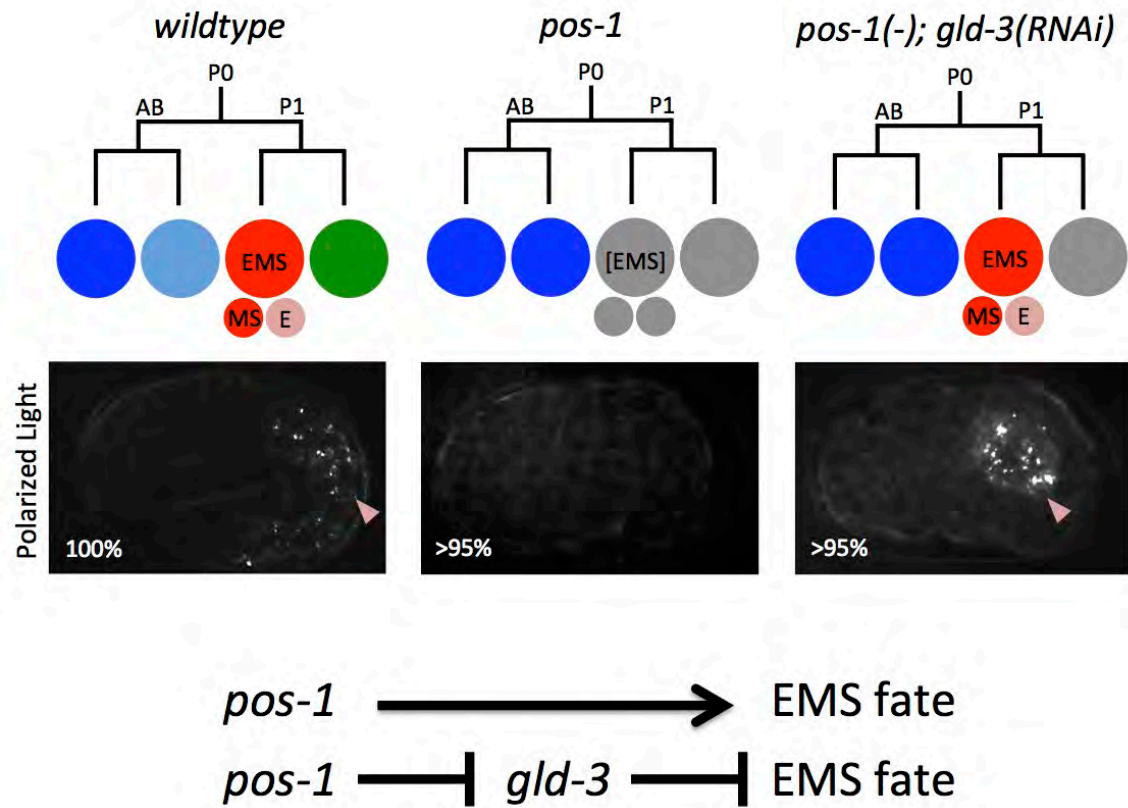**gld-3 loss of function restores gut specification in the *pos-1(-)* background**

Figure 2.2

Figure 2.2. Birefringent gut granules (pink triangle) are restored in *pos-1; gld-3(RNAi)* embryos. The origin of the developing gut is EMS as opposed to alterative blastomeres (data not shown). We represent the *gld-3* suppression of *pos-1* as *pos-1* negatively regulating *gld-3*, which in turn negatively regulates EMS fate.

Limitations and caveats

The limitations of my inquiry include:

1) I assumed *pos-1* and *gld-3* were regulating the translation of the same gene, which was crucial for EMS fate determination. It is conceivable that they were acting on two or more independent genes that function in parallel to one another. It is also conceivable that they regulate a process other than translation.

2) I assumed that *pos-1* was acting as a repressor while *gld-3* was a promoter of this single gene. The Drosophila homolog of GLD-3, Bicaudal C, has been found to function in both poly(A) tail adenylation and deadenylation, depending on the complex it engages with. Such a dual role for GLD-3 is conceivable. Moreover, although *pos-1* is a repressor of *glp-1* expression, it is required for the expression of *apx-1*. Therefore, considering *pos-1* a repressor and *gld-3* an activator is only one way to think of their roles in EMS specification.

3) It follows that I assumed that the single gene in question was an antagonist of EMS fate and that it could therefore be uncovered in a *pos-1* suppressor screen.

RNAi screen of essential genes in pursuit of *pos-1* suppressors.

Despite these limitations I proceeded with an RNAi screen due to its feasibility. I chose to screen approximately 1000 genes that are essential for embryogenesis (Kamath and Ahringer, 2003). This constitutes ~5% of *C. elegans* coding genes but is biased towards those with a role in connection with embryonic development. RNAi knockdown of any of these genes alone would result in a percentage of dead embryos.

The *pos-1* suppressor screen was performed by placing 5 to 10 L2 – L4 *pos-1(zu148) unc-42(e270)* worms on RNAi food and allowing them to reach adulthood and

lay eggs. Embryos were then scored for gut granules. Positive hits were then retested by scoring the embryos of 4 to 8 individual worms fed the RNAi food in order to assess overall gut development and obtain average *pos-1* suppression and standard deviation.

Results of *pos-1* suppressor RNAi screen

The top suppressors of *pos-1* are summarized in **Table 2.1**. Notably, *gld-2* was identified as a suppressor of *pos-1*, which confirmed our earlier finding that it can restore gut specification in the absence of *pos-1*.

rev-1 encodes a translesion DNA polymerase capable of replicating DNA despite the presence of hindering lesions and is a target of the mesoderm determining transcription factor HLH-8/Twist (Holway et al., 2006; Wang et al., 2006b). *cye-1*/Cyclin E initiates the S-phase of the cell cycle and regulates the decision between cell proliferation and differentiation (Ishidate et al., 2014). *lrr-1* encodes a leucine rich repeat-containing protein and is required for regulating the extent of DNA replication during S phase (Merlet et al., 2010). *rba-1* is a chromatin remodeling factor that has been implicated in DNA replication coupled chromatin assembly pertinent to neuronal bilateral asymmetry (Nakano et al., 2011). *rpn-8* is a proteasome regulatory particle that functions in SKN-1 turnover after EMS division (Du et al., 2014). It is conceivable that *rpn-8* knockdown causes an accumulation of SKN-1 sufficient to activate gut differentiation.

A gene that suppressed *pos-1* in a reproducible and robust manner was the uncharacterized gene F32D1.6. The quality of suppression was on par with that of *gld-3* and it was predicted to be a target of *pos-1* by virtue of a POS-1 consensus binding element in its 3'UTR. I pursued this gene for future study.

Table 2.1. Results of *pos-1* suppressor RNAi screen

Sequence Name	Gene Name	RNAi phenotype*	*pos-1* suppression (Gut+)	Function
F32D1.6		**20% Emb, Lvl, Bmd**	78.8% ±10.8 (n=7)	**unknown**
ZK675.2	*rev-1*	100% Emb	76.7% ±14.4 (n=4)	Translesion DNA polymerase
K07A1.11	*rba-1*	100% Emb, Pvl, Unc.Prz	46.7% ±29.2 (n=8)	RBAp48 related
R12E2.3	*rpn-8*	100% Emb, Ste (no ooc)	23.8% ±5.8 (n=4)	Proteasome regulatory particle
ZC308.1	*gld-2*	100% Emb, Ste (no ooc)	41.2% ±12.3 (n=4)	Cytoplasmic poly(A) polyermase
C37A2.4	*cye-1*	10% Emb, Clr	19.1% ±17.7 (n=5)	Cyclin E
F33G12.4	*lrr-1*	100% Emb	29.9% ±33.1 (n=6)	Leucine-Rich Repeat protein

Phenotypes from Ahringer Database.

Emb = Embryonic lethal   Pvl = Protruding vulva

Clr = Clear   Ste = Sterile

Lvl = Larval lethal   Bmd = Body morphology defects

Unc = Uncoordinated   Prz = Paralyzed

Cyclin E regulates cell cycle and differentiation

Cyclins owe their name to their oscillating patterns of expression throughout the cell cycle. This cyclic property enables cyclins to provide an active kinase complex with substrate specificity appropriate for each of the cell cycle phases (Morgan, 1997). Cyclin E binds and functions with its dependent kinase CDK2 to mediate the entry of cells into the DNA synthesis (S) phase and the expression of S-phase specific genes (Moroy and Geisen, 2004). An increasing number of studies have connected Cyclin E with differentiation processes (Ishidate et al., 2014). Four examples are provided below from research on *Drosophila melanogaster* and *C. elegans* somatic and germline development.

Cyclin E in *Drosophila*

Individual neuroblasts produced from corresponding positions in thoracic and abdominal segments usually acquire similar fates in fruit fly. However, the neuroblast NB6-4 lineage is a case where variations exist (Akiyama-Oda et al., 1999). In the thoracic segments of the embryonic central nervous system (CNS), NB6-4t divides asymmetrically to generate one neuronal and one glial precursor cell. The neuronal precursor cell then divides several times in stem-cell mode to generate a neuronal sub lineage. The glial precursor cell, however, divides twice to generate three glia **(Figure 2.3)**. In the abdominal segment of the embryonic CNS, the corresponding neuroblast (NB6-4a) merely divides symmetrically into two glial cells **(Figure 2.3)**. Mutations in homeotic genes *abd-A* and *abd-B* cause abdominal NB6-4 to transform into the more elaborate thoracic lineage. Conversely, overexpression of *abd-A* results in the opposite NB6-4t to NB6-4a transformation (Berger et al., 2005a, b). Genetic and biochemical evidence demonstrate that these homeotic transformations are dependent on Cyclin E.

Loss of function mutations in Cyclin E result in thoracic NB6-4 adopting the fate of its abdominal counterpart and generating only two glial cells. Conversely, ectopic expression of Cyclin E in abdominal NB6-4 transforms its mode of division to one that resembles thoracic NB6-4. *In situ* hybridization reveals that Cyclin E is expressed in thoracic NB6-4 but not in the abdominal counterpart **(Figure 2.3)**. After the first asymmetric division, Cyclin E mRNA is detected in the neuronal precursor but not in the glial precursor daughter cell. In *abd-A* mutant embryos, which exhibit an abdominal to thoracic NB6-4 transformation, Cyclin E mRNA is detected in the transformed NB6-4a lineages. Importantly, overexpression of *abd-A*, which causes the opposite transformation, reduces Cyclin E mRNA levels in the thoracic segment. Perhaps most critical, the homeotic transformation observed in *abd-A* mutants is suppressed by mutations in Cyclin E. This epistatic relationship, along with potential Abd-A-binding sites in the Cyclin E promoter, suggest that homeotic genes in the Bithorax complex regulate the differential expression of Cyclin E, which maintains the neuroblast stem cell fate (Berger et al., 2005a).

An alternative mode by which Cyclin E regulates differentiation has been described in *Drosophila* female germline stem cells (Ables and Drummond-Barbosa, 2013). Germline stem cells deficient in Cyclin E or its canonical partner cyclin-dependent kinase 2 (Cdk2) display normal rates of proliferation as long as they are within the stem cell niche, emphasizing that the cell cycle machinery is still intact. However, these Cdk-2/Cyclin E deficient germ cells fail to maintain their proliferative fate due to a weakened response to niche bone morphogenetic protein (BMP) signals, which are required for germ stem cell maintenance. The outcome of this failure is that germ cells are lost from

the stem cell niche, arrest at G1 stage and undergo excessive growth without committing to a clear differentiated fate. This impaired response to BMP induction suggests that Cyclin E and Cdk2 maintain germline stem cells by modulating its response to niche derived signals (Ables and Drummond-Barbosa, 2013).

Cyclin E is elevated in Drosophila neuroblasts during asymmetric division of NB6-4t
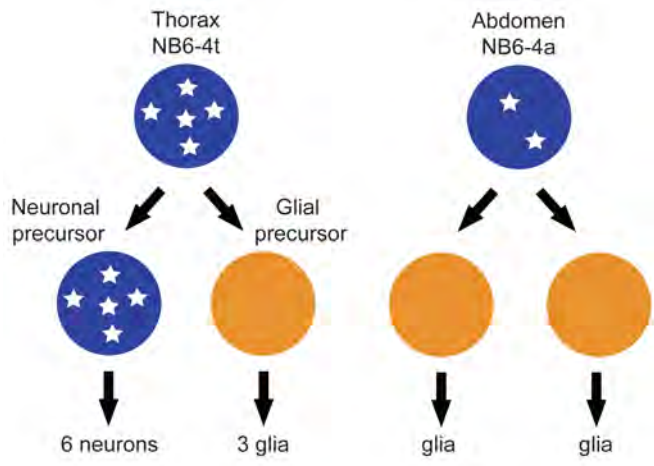

Figure 2.3

Figure 2.3. Asymmetric cell division of thoracic neuroblast 6-4 (NB6-4). NB6-4t divides asymmetrically to produce a neuroblast (blue) and gliablast (orange) whereas NB6-4a divides symmetrically giving rise to 2 glial cells. The neuroblast identity is maintained by high levels of Cyclin E (white stars), which is significantly reduced in NB6-4a.

Cyclin E in *C. elegans*

The *C. elegans* germline consists of two gonad arms connected at the uterus (Kimble and Crittenden, 2005). The distal end of the each gonad arm (relative to the uterus) harbors a stem cell niche of proliferating germ cells known as the 'mitotic zone' **(Figure 2.4)**. However, unlike other stem cell systems in which a small number of stem cells divide asymmetrically, the *C. elegans* mitotic zone harbors >200 syncytial cells that divide into equivalent daughters. The fates of these daughters are not determined by asymmetric division but by their position relative to the niche signal.

The proliferative mode of the mitotic zone is promoted by Notch signaling from the Distal Tip Cell (DTC) (Crittenden et al., 2003; Kimble and Simpson, 1997). The DTC presents the LAG-2/Delta ligands, which activates GLP-1/Notch receptor cleavage in germ cells thereby unleashing the Notch intracellular domain (ICD). Once released from the plasma membrane, Notch ICD enters the germ-cell nucleus where it interacts with CSL family transcription factors to regulate the expression of genes necessary for the proliferative fate (Christensen et al., 1996).

Beyond the range of Notch signaling, mitotic germ-cells enter the transition zone and enter meiosis, in part through the function of GLD-1. GLD-1 is a KH domain RNA binding protein that accumulates at low levels in the distal end of the germline but gradually rises until reaching the maximum level in the transition zone. In the absence of GLP-1/Notch signaling, GLD-1 is ectopically expressed in the distal region of the germline (Hansen et al., 2004). Conversely, overexpression of GLD-1 in the distal end abrogates the proliferative fate and causes the germ cells to enter meiosis prematurely. Consistent with these observations, GLD-1 binds the 3'UTR of *glp-1* mRNA and

represses its translation in the embryo and meiotic region of the germline (Marin and Evans, 2003). Collectively, these genetic and biochemical interactions suggest a mutual repressive feedback between Notch backed proliferation and GLD-1 based differentiation.

The proliferative fate has also been shown to depend on CDK-2/CYE-1, since premature meiotic differentiation is enhanced in a weak *glp-1* mutant background upon *cye-1* or *cdk-2* RNAi (Fox et al., 2011). CDK-2/CYE-1 directly phosphorylates GLD-1 in vitro and *cye-1* RNAi leads to an increase in GLD-1 levels in the distal end of the gonad. Since mutating the phosphorylation site in GLD-1 phenocopies this increase in distal GLD-1 levels it is likely that that CDK-2/CYE-1 could block the switch to meiotic differentiation by targeting GLD-1 and thereby maintain the proliferative mode of mitotic germ cells (Jeong et al., 2011). However, since reduction of *cdk-2* or *cye-1* function leads to premature meiotic entry even in the absence of *gld-1*, CDK-2/CYE-1 must also be promoting proliferation through a GLD-1 independent pathway.

The role of Cyclin E in *C. elegans* vulval development may offer a clue to this additional mode of action. During vulva development in *C. elegans*, the Anchor Cell (AC) signals the vulval precursor cell (VPC) P6.p via the EGFR signaling pathway. Upon this induction, P6.p then signals its neighboring VPCs P5.p and P7.p via the Notch signaling pathway to adopt a different fate. By utilizing a NICD::GFP reporter, it was demonstrated that CYE-1 exerts a stabilizing effect on the intercellular pool of NICD in VPCs (Nusser-Stein et al., 2012). It is enticing to surmise that CYE-1 may exert a similar stabilizing effect on NICD molecules in the germline mitotic zone **(Figure 2.4)**. If this indeed is the case, then Cyclin E would maintain the proliferative mode of mitotic germ

cells by both positively stabilizing the Notch proliferative cue and negatively targeting GLD-1 to prevent differentiation. Interestingly, GLD-1 binds to the 3'UTR of *cye-1* and represses its translation beyond the proliferative zone, emphasizing the mutual negative feedbacks that govern proliferation versus differentiation decisions in the germline (Biedermann et al., 2009) **(Figure 2.4)**.

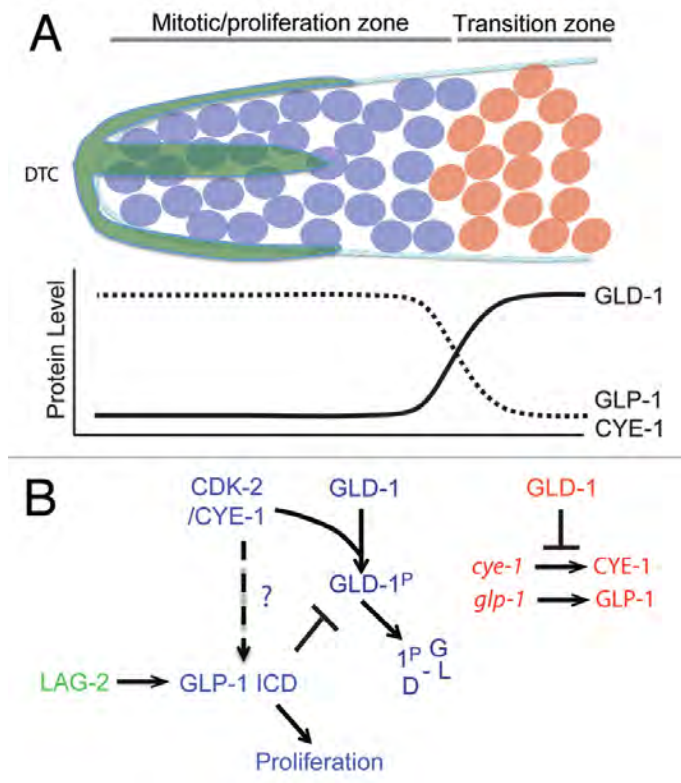**Cyclin E maintains proliferative fate in C.** *elegans* **germline**
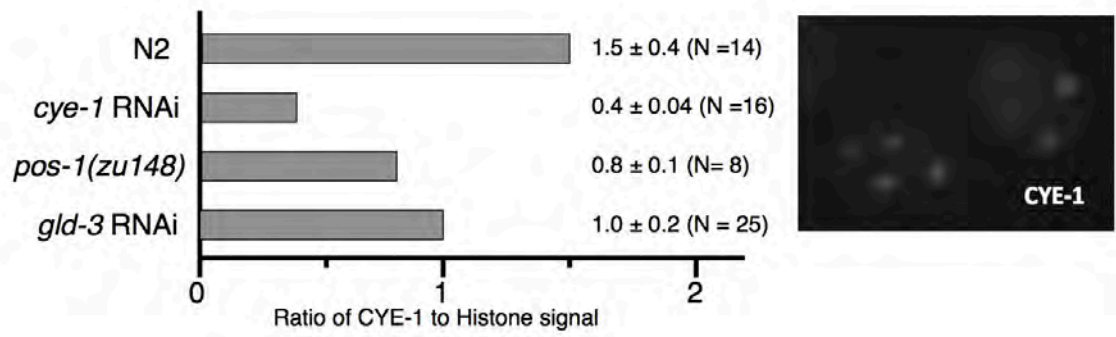

Figure 2.4

Figure 2.4. (**A**) Transition from mitosis into meiosis in *C.elegans* germline. The Distal Tip Cell (DTC in green) caps the distal end of the germline and induces mitotic proliferation (blue). Beyond the reach of the DTC Notch signaling, germ cells transition into the differentiation zone in a GLD-1 dependent manner, where they undergo meiosis (red). (**B**) Model for Cyclin E in maintenance of germ stem-cell fate in *C. elegans*. The boundary between the proliferative and transition zones is maintained by 2 mutual negative regulations: CDK-2/ CYE-1 and GLP-1 inhibit accumulation of GLD-1 in the mitotic proliferation zone, while GLD-1 represses the translation of *cye-1* and *glp-1* transcripts in the transition zone. It remains to be seen if CYE-1 can maintain proliferation zone by enhancing GLP-1 ICD induction in the similar way it does in vulval development.

Embryonic Cyclin E levels are not affected by *pos-1* or *gld-3* loss of function

Given the role of Cyclin E in maintaining the proliferative fate of germ cells, I hypothesized that gut development may be lost in *pos-1(-)* embryos due to an overexpression of Cyclin E preventing endo-mesoderm differentiation. This hypothesized overexpression would require *gld-3*. In this model, *cye-1* is a gut antagonist due to its support of the germline proliferative fate, *pos-1* is a repressor of *cye-1* translation and *gld-3* is an activator of *cye-1* expression. This would explain why *cye-1(RNAi)* restores gut specification in the *pos-1(-)* background.

In order to test this hypothesis we immunostained early embryos from wildtype, *pos-1(zu148)* and *gld-3(RNAi)* to detect Cyclin E CYE-1/Cyclin E levels were quantified and compared to co-stained histone. Contrary to our hypothesis, CYE-1 levels were not upregulated in the absence of *pos-1* function. Moreover, despite a reduction in CYE-1 levels in *gld-3(RNAi)* embryos, the reduction was less than that observed in the absence of *pos-1* function **(Figure 2.5)**. These two results made it improbable that *pos-1(-)* embryos are gutless due to a *gld-3* dependent increase in CYE-1 levels.

CYE-1/Cyclin E is not elevated in *pos-1* mutant embryos

Figure 2.5

Figure 2.5. The intensity of antibody signal recognizing CYE-1 was compared to that from an antibody recognizing histones. The average ratio of CYE-1 : Histone signal in wildtype embryos was 1.5. This ratio was reduced to less than 0.5 after *cye-1(RNAi)*. The ratio did not increase after *pos-1* RNAi. *gld-3* RNAi reduced the CYE-1 : Histone signal to 1, suggesting that GLD-3 may be required for *cye-1* expression. Immunostaining was performed by Meetu Seth.

glp-1ts is a suppressor of *pos-1*

Concurrent with the RNAi suppressor screen, a mutagenesis screen in our lab identified a mutant, *ne4298ts,* as temperature sensitive haplo-insufficient. This means that heterozygous hermaphrodites are sterile at 25°C. Since *gld-3* was initially identified as a haplo-insufficient maternal-effect sterile mutant, I tested to see if *ne4298ts* was also a *pos-1* suppressor and indeed it was. More than 50% of *ne4298ts; pos-1(RNAi)* worms developed gut at 25°C **(Figure 2.6)**. This suppression was not observed at 15°C.
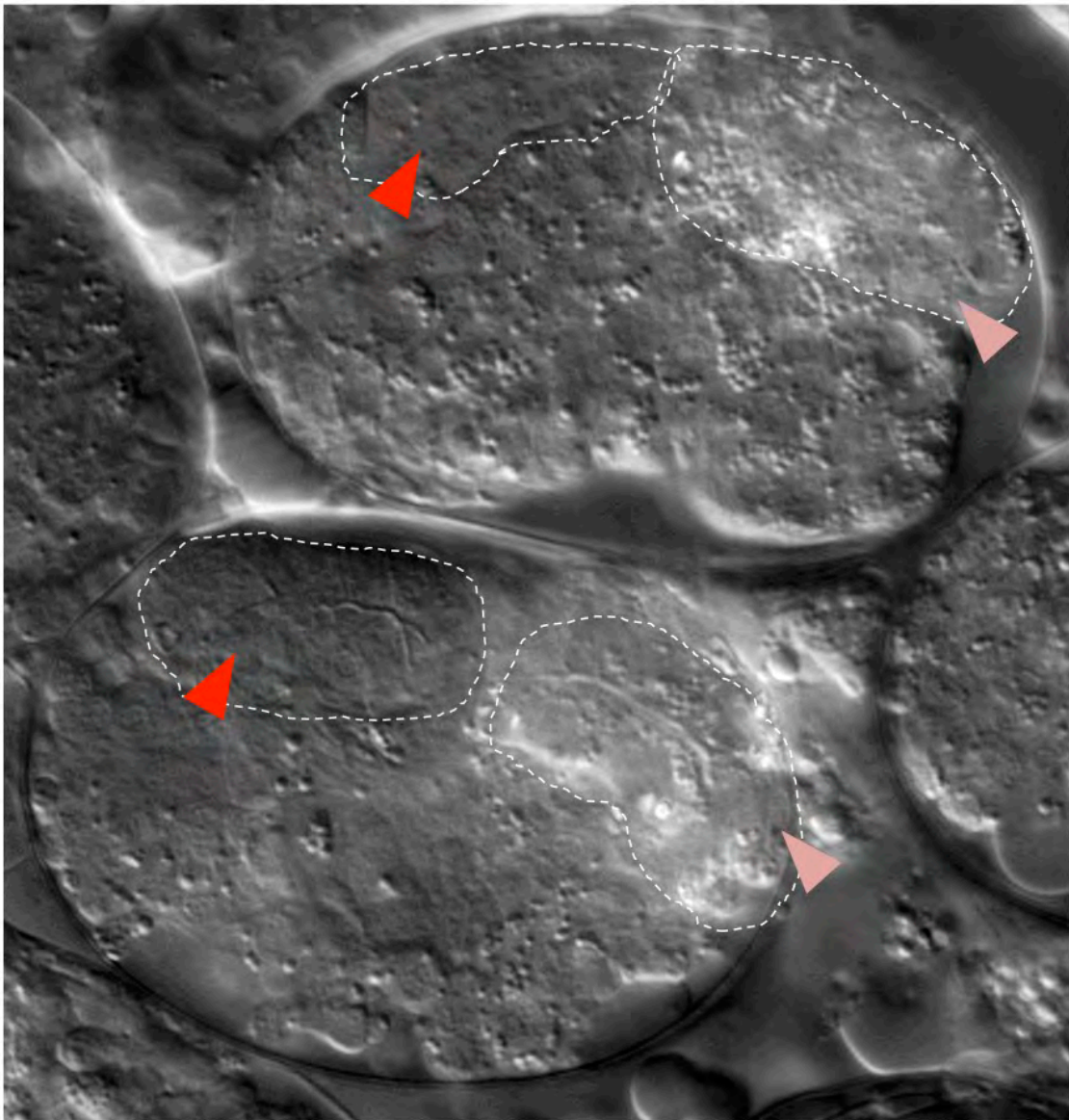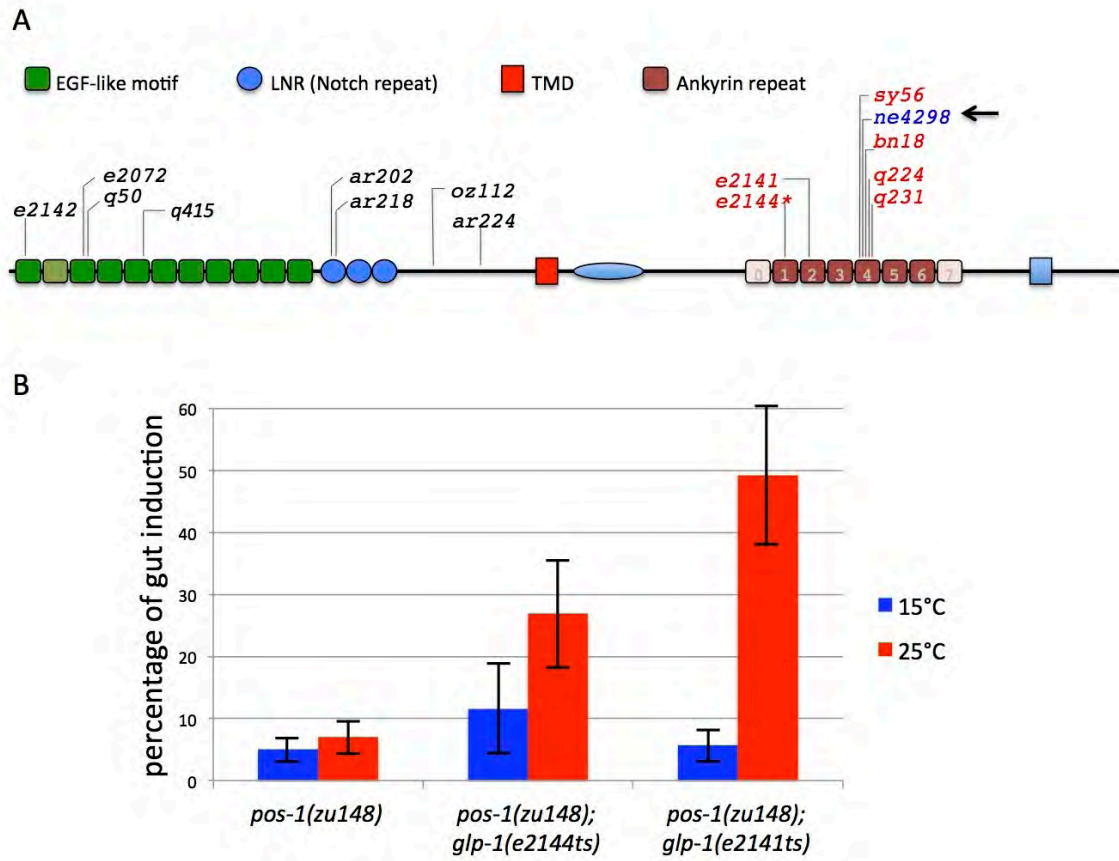
Figure 2.6

Figure 2.6. Embryos from *glp-1(ne4298ts); pos-1(RNAi)* worms using DIC microscopy. Embryos arrest with restored endoderm development (pink triangle) and improved pharyngeal differentiation (red arrows).

ne4298ts was mapped to *glp-1* and by sequencing was shown to harbor a lesion that results in an amino-acid substitution (G1031D) in the 4th (of 7) ankyrin repeats. Several other *glp-1*ts alleles exhibit amino acid substitutions in the same ankyrin repeat region (including *e2141*, *e2144*, *bn18* and *q231*) **(Figure 2.7)**. At least two of these apparent loss of function temperature sensitive alleles also suppress the endo-mesoderm defect of *pos-1(zu148)* **(Figure 2.7)**. Interestingly, the temperature sensitivity point for endoderm restoration occurs in the maternal gonad hours prior to fertilization suggesting that *glp-1* activity in the germline can interfere with endoderm specification in *pos-1* mutants much later in the embryo **(Figure 2.8)**.

glp-1ts alleles with mutations in ankyrin repeat region suppress pos-1

Figure 2.7

Figure 2.7. (A) A schematic of GLP-1 protein identifying its key structural features. Several temperature sensitive loss of function alleles including *ne4298* have mutations in the ankyrin repeat region. (B) Two other temperature sensitive loss of function *glp-1* alleles suppress the *pos-1* gutless phenotype at 25°C. *pos-1(zu148), pos-1(zu148); glp-1(e2144ts)* and *pos-1(zu148); glp-1(e2141ts)* hermaphrodites were grown at 15°C (n = 5, 11 and 6 respectively) and the percentage of gut induction in there progeny was scored (blue columns). Siblings from the same genetic backgrounds were upshifted as L4s to 25°C (n = 5, 11 and 6 respectively) and gut induction in their progeny was scored (red coloumns). Error bars reflect standard deviation.

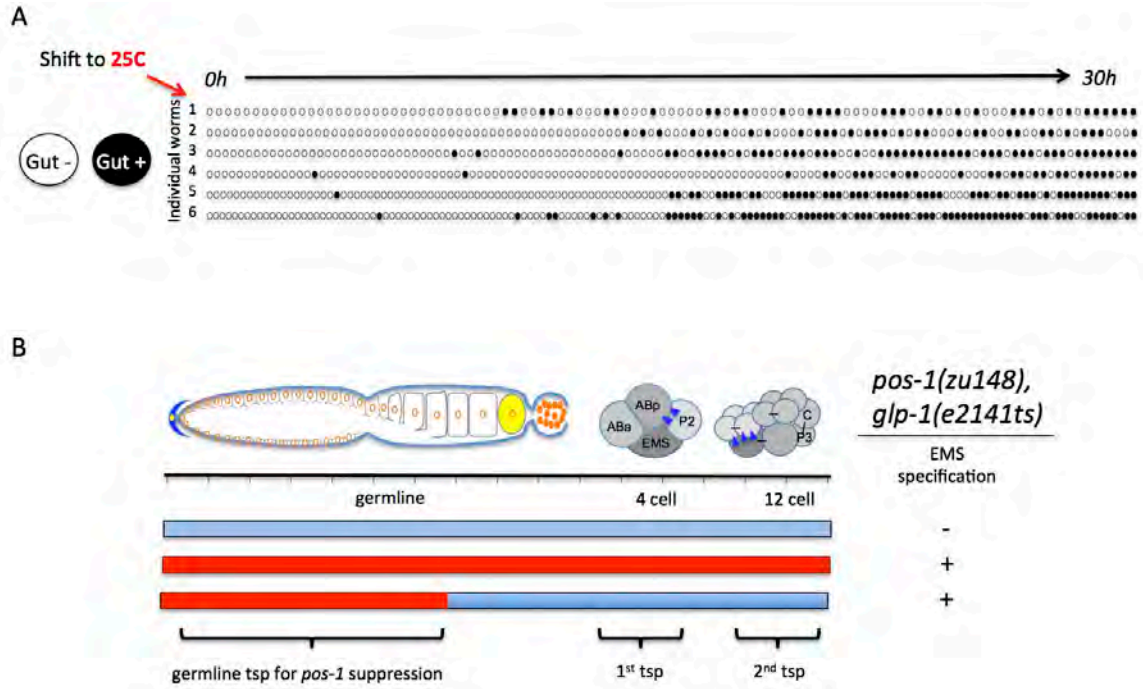**Temperature sensitive point (tsp) of *glp-1ts* suppression is in maternal germline**

Figure 2.8

Figure 2.8. The temperature sensitivity point of *glp-1ts* suppression is in the maternal germline. (A) Adult *glp-1(ne2141ts); pos-1(zu148)* worms reared at 15°C (permissive temperature) were upshifted to 25°C and embryos were scored for gut development in the order they were laid. Endoderm restoration ensues after 10 hours of upshift, indicating that the loss of *glp-1* function in the developing maternal germ cells and not in the embryos suppresses *pos-1*. (B) *glp-1* functions in three different contexts, germline proliferation in the distal end of the gonad, P2 to ABp signaling (1st tsp) and MS to AB8 signaling in the early embryo (2nd tsp). The first context is the site of *glp-1* temperature sensitive suppression of *pos-1*.

Summary

pos-1(-) embryos die gutless. To understand why, I conducted an RNAi screen using ~1000 genes required for embryonic development to search for suppressors of *pos-1* that would restore gut development. *cye-1* and *neg-1* were among the strong suppressors and were chosen for further study. Immunofluorescent staining did not reveal an increase in CYE-1 levels the absence of *pos-1*, ruling out the hypothesis that *cye-1* is a gut antagonizer repressed by POS-1. *neg-1* is the subject of the following chapter. In parallel, *glp-1(ne4298ts)* was isolated in a forward mutagenesis screen and found to suppress *pos-1*. Interestingly, the temperature sensitive point for gut restoration was in the maternal germline hours before the onset of embryogenesis.

MATERIALS AND METHODS

RNAi screen

Adult *pos-1(zu148) unc-42(e270)* heterozygous worms were fed *E. coli* HT115(DE3) expressing double stranded RNA targeting a specific gene expressed by the *C. elegans* genome. A quarter of the progeny were *pos-1* homozygous worms identified by their Unc phenotype. These worms were singled on to fresh plates and their dead embryos examined and scored for gut induction; evident by birefringent gut granules (Laufer et al., 1980). The screen focused on ~1000 genes essential for embryogenesis. Since many of these genes were also required for germline development, prolonged exposure to RNAi caused sterility in approximately one quarter of knockdowns. To avoid this limitation, *pos-1* homozygous worms in the L3 stage were placed on the RNAi food. This shorter exposure allowed for the production of examinable progeny.

Chapter III

neg-1

and its regulation

PREFACE

C13F10.7 was identified as an interactor of NEG-1 in collaboration with **Alex Tamburino** (Laboratory of Marian Walhout, University of Massachusetts Medical School, MA).

Cell-lineaging of *neg-1(tm6077)* was done in collaboration with **Zhuo Du** (Laboratory of Zhirong Bao, Memorial Sloan Kettering Cancer Center, NY).

Behavioral assays were done in collaboration with **Chris Chute** (Laboratory of Jagan Srinivasan, Worcester Polytechnic Institute, MA).

In vitro binding assays between *neg-1* 3'UTR and MEX-3, MEX-5 and POS-1 were done in collaboration with **Ebru Kaymak** (Laboratory of Sean Ryder, University of Massachusetts Medical School, MA).

Extension polyA tail assays were done in collaboration with **Traude Beilharz** (Monash University, Australia)

GFP tagged *neg-1* worm strains and *neg-1* 3'UTR GFP reporter strains were engineered in collaboration with **Masaki Shirayama**.

ABSTRACT

neg-1 was identified as a suppressor of the *pos-1* gutless phenotype. The product of this gene is a novel protein with no sequence homology beyond the *Caenorhabditis* genus. Importantly, this gene is required for proper embryogenesis since >70% of *neg-1* knockout embryos die with defects in anterior development.

neg-1 mRNA is maternally contributed to the embryo where it is expressed in the anterior blastomeres ABa and ABp but not the posterior blastomeres EMS and P2. In the absence of *pos-1* function, *neg-1* is ectopically expressed in P2 and EMS. Poly(A) tails of *neg-1* transcripts in the early embryo are extended beyond wildtype lengths in the absence of *pos-1* function and are shorter in the absence of *gld-3*. We hypothesized that POS-1 binds the 3'UTR of *neg-1* mRNA to block *gld-3* mediated cytoplasmic polyadenylation and subsequent translation. Indeed, POS-1 binds the 3'UTR of *neg-1 in vitro* as do the anterior blastomere determinants MEX-3 and MEX-5. *mex-5(RNAi)* leads to a loss of *neg-1* expression. I propose a model whereby POS-1 guards EMS identity by blocking GLD-3 mediated *neg-1* expression. Conversely, MEX-5 protects the fate of the anterior blastomere AB by countering POS-1 repression of *neg-1* expression.

Characterization of *neg-1* function remains preliminary. Interestingly, the embryonic lethality of *neg-1* is suppressed in *neg-1; med-1* double mutants. *med-1* is a GATA transcription factor implicated in EMS specification and subsequent mesoderm and endoderm differentiation. This genetic interaction suggests that *neg-1* and *med-1* could be involved in a mutually inhibitive regulatory network that segregates the fates of AB and EMS.

75

neg-1 is a suppressor of *pos-1*

RNAi knockdown of F32D1.6 restored EMS specification and gut development in the otherwise gutless *pos-1(zu148)* genetic background (>95% of *pos-1; F32D1.6(RNAi)* embryos were gut+) **(Figure 3.1)**. We refer to this gene as *neg-1* (<u>N</u>egative <u>E</u>ffect on <u>G</u>ut). Interestingly, two independent groups using single cell transcriptomics have demonstrated that F32D1.6 transcripts are enriched in the AB blastomere (Hashimshony et al., 2012) (Osborne-Nishimura personal communication). Moreover, F32D1.6 was among 968 genes found to be enriched in 13 embryonic motor neurons (*unc-4::GFP* neurons), suggesting an association with neuronal development (Fox et al., 2005).

neg-1(-) restores EMS specification and gut development in _pos-1(-)_ embryos
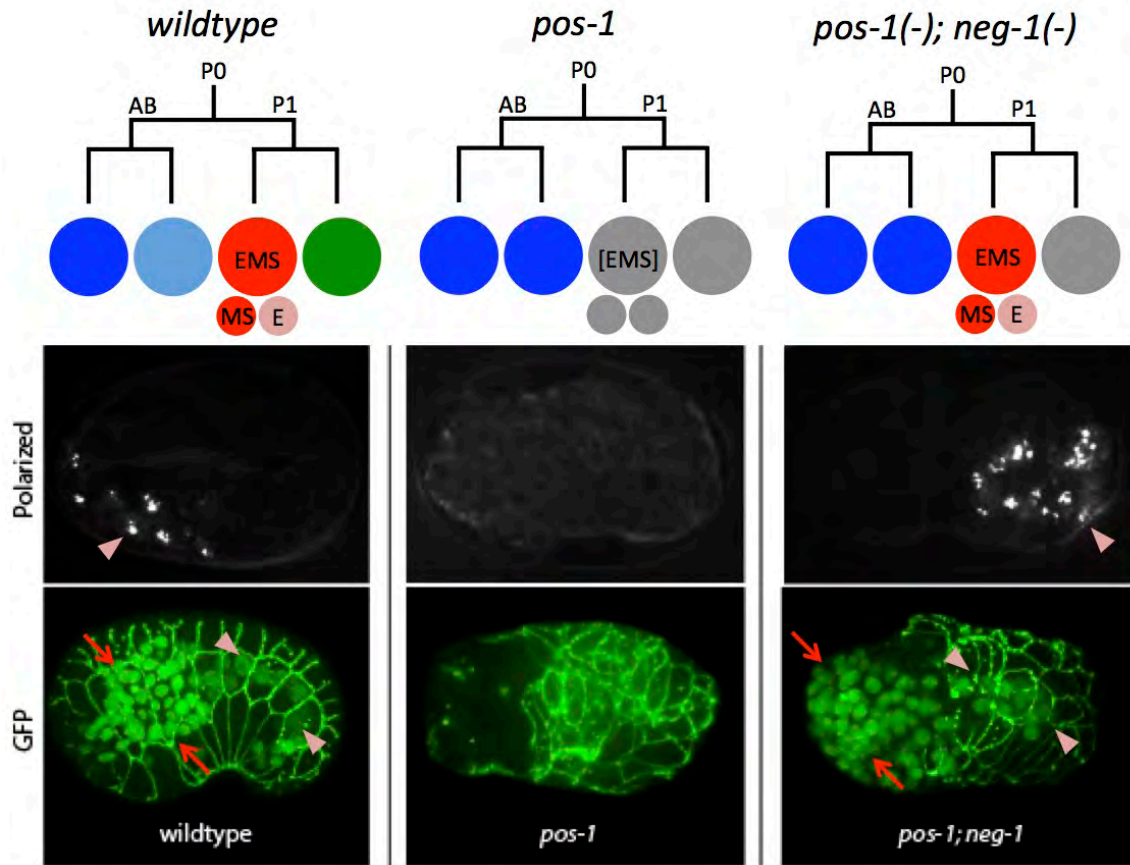

Figure 3.1

Figure 3.1. Birefringent gut granules (pink triangle) are restored in *pos-1(zu148); neg-1(RNAi)* embryos. Two GFP markers demonstrate the *pos-1* phenotype. *ajm-1::gfp* marker is expressed in cell-cell junctions of hypodermal cells surrounding the embryo. *pha-4::gfp* is expressed in nuclei of pharynx (between red arrows) and intestine (between pink triangles). *pos-1* embryos have incomplete hypodermis due to misspecification of ABp (light blue circle in wildtype) into ABa (blue blastomere). These embryos also lack pharynx and intestine due to failure in EMS identity determination (no GFP). *neg-1* loss of function restores EMS identity and endo-mesoderm differentiation (pharyngeal and intestinal development) but does not rescue the hypodermal defect.

NEG-1 sequence is poorly conserved

Local alignment searching using NCBI BLAST revealed a single homologue of NEG-1 in the nematode *C. brenneri* (31% identity, e-value 4×10^{-07}). A second query using CBN07554 solely retrieved NEG-1 (33% identity, e-value 5×10^{-09}). However, the public nematode database Wormbase annotated the *C. japonica* protein CJA09859 as a homologue of CBN07554 (identity 30%, e-value 2×10^{-08}). Oddly, a third BLAST query using CJA09859 did not retrieve CBN07554, but instead detected NEG-1 (34% identity, e-value 0.04). A multiple alignment revealed that homology was highest at the C termini of these three proteins **(Figure 3.2)**. The three nematode genes encoding NEG-1, CBN0775 and CJA09859 reside in corresponding genomic neighborhoods spanning approximately 25 kbp and demarcated by the genes *grl-13* and the uncharacterized gene F13A2.9 **(Figure 3.3)**. Given the sequence homology and the similar genomic location, I consider these genes to be homologues.

The initial BLAST query also retrieved a 7459 amino acid long predicted uncharacterized protein from the Acorn worm *Saccoglossus kowalevskii*. The last subject retrieved from the query was a 240 amino acid long hypothetical protein from the protist *Perkinsus marinus*. However, the poor e-value of these results (2.9 and 4.7 respectively) discouraged further pursuit.

BLASTP searches protein sequence databases. On the other hand, Domain Enhanced Lookup Time Accelerated BLAST, DELTA-BLAST first searches a database of position-specific score matrices (PSSMs) before searching protein sequences (Boratyn et al., 2012). This approach allows the recovery of distant homologues that may otherwise escape BLAST queries. A DELTA-BLAST query using *C. elegans* NEG-1

identified itself, the *C. brenneri* homolog and numerous 50S ribosomal subunit 15 (RPL15) proteins from myriad prokaryotes. Position specific iterated BLAST (PSI-BLAST) iteratively searches a protein sequence database, using the matches in one round to construct a position-specific score matrix (PSSM) for searching the database in next round (Altschul et al., 1997). Performing PSI-BLAST by selecting the *C. elegans* and *C. brenneri* NEG-1 sequences from the first DELTA-BLAST, but not the 50S ribosomal protein 15 subjects, retrieved NEG-1 and a mixture of plant 60S and prokaryotic 50S ribosomal protein 15 subjects. The identity shared between NEG-1 and these RPL15 averaged around 20% with an e-value of less than 10^{-30}. Conducting the same DELTA-BLAST followed by an iteration of PSI-BLAST using the *C. brenneri* homolog of NEG-1 did not retrieve RPL15. Instead, the large subunit of DNA-directed RNA polymerase II was the main result (identity ~24% and e-value < 10^{-24}. Despite low e-values, inspecting the alignments was not satisfying. Furthermore, this result was not arrived at when beginning the query with *C. brenneri* NEG-1.

Pattern Hit Initiated BLAST (PHI-BLAST) minimizes false positives by filtering those that do not include a predefined amino acid pattern. The C terminus of *C. elegans* NEG-1 and the homologs in *C. brenneri* and *C. japonica* share the following pattern LGTXXXRXXXLKLM, where X is any amino acid. I therefore used a more flexible version of this pattern (LGTX$_{1-10}$LKLM) to search for further NEG-1 homologs. However, this was unfruitful.

Furthermore, protein databases Pfam, InterPro and PROSITE detected no motifs or conserved region nor affiliation with any known protein family. Moreover, protein structure prediction services offered predicted structures with low confidence levels (I-

TASSER) or declined to predict all together (Phyre) (data not shown). However, NCBI

Concerved Domains identified a portion of the 50S ribosomal protein L15e domain in

NEG-1 but not in the *C. brenneri* or *C. japonica* homologs **(Figure 3.2)**.

Multiple alignment of NEG-1 and two homologs

A
```
NEG-1      MYSHPALLQFPLMSTD-AFYSTAFPTIPGTPNLLGYTAYSPQSIFND----SRPASSVNN 55
CBN07554   MYPHLSNYQFPMFSASPTFYPPAFPAMPISPIFFNQ---SPQSGFNSPTLYPFSLASTPP 57
CJA09859   MYMN---TQFPLTSSAPPITMMAHPQMANYLSAFSTNQQTPSGFFMTPLISPLAFSPMFL 57
           ** :    ***: *:  .:   *.* :.      :.    :*.. *      .  . :.

NEG-1      TPSTSFNSS-------IHSRNRSIAMLHSISHILSNDDSLGST-----TESSTPSTPEIH 103
CBN07554   TPFSSFNSASDF---GNWSFPRSMSDSFSLPSVHSASGSATSS-----STESATSSPEQQ 109
CJA09859   GGSFQTQSPICIGDVPRFKEEQTPIRSFSLPESESTIDVLNTSPTSEYSTTSSPACSDFC 117
           . :*.         . ::    .*:.   *  .   ::     : *:.: .:

NEG-1      PSNAPAKKT---IPLVLISDTEECDKKSRKRRIRSK--SSKKSRG-----------IR 146
CBN07554   QQRQTAPDSSCEISVPSSSSGSGDSPKQENRRFRRGR--MAKRSDG-----------VK 155
CJA09859   SLIISRPDQHIPLVHLTDSDDTENTPKTIRKRKNRSERPCAKRSKGKLSPCTPEVVEVVR 177
           .  .   :    *.  :  *  .:*  *..   :*:* *            ::

NEG-1      KSRK---------HEASDYR-----KKIMLGTFEGRKIVLKLMK------ 176
CBN07554   RRRM---------TSESEECDESEPRRIPLGTFNGRRVVLKLMGPSKPLQ 196
CJA09859   TSRASPIECPVENSPEVLLEEKRPRKKLVLGTLGKRKFTLKLMSPE---- 223
           *              ::: ***:  *:..****
```

B

NEG-1 ▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬ [Ribosomal L15e superfamily] ▬▬▬▬▬

PRK04243[PRK04243], 50S ribosomal protein L15e; Validated

Cd Length: 196 **Bit Score:** 34.51 **E-value:** 7.44e-03

```
                     10        20        30
            ....*....|....*....|....*....|....*...
NEG-1     125 CDKKSRKRRIRSKSS--KKSRGIRKSRKhEASDYRKKI 160
Cdd:PRK04243 152 CDKSHRGRVFRGLTSagKKGRGLRKKGK-GTEKVRPSI 188
```

Figure 3.2

Figure 3.2. (A) Two homologs from *C. brenneri* (CBN07554) and *C. japonica* (CJA09859) are aligned with NEG-1. The C-terminus exhibits the highest conservation. Putative nuclear localization sequences are underlined. (B) A predicted Ribosomal protein L15e superfamily domain detected by NCBI Conserved domains.

neg-1 and its homologs reside in the same genomic neighborhood
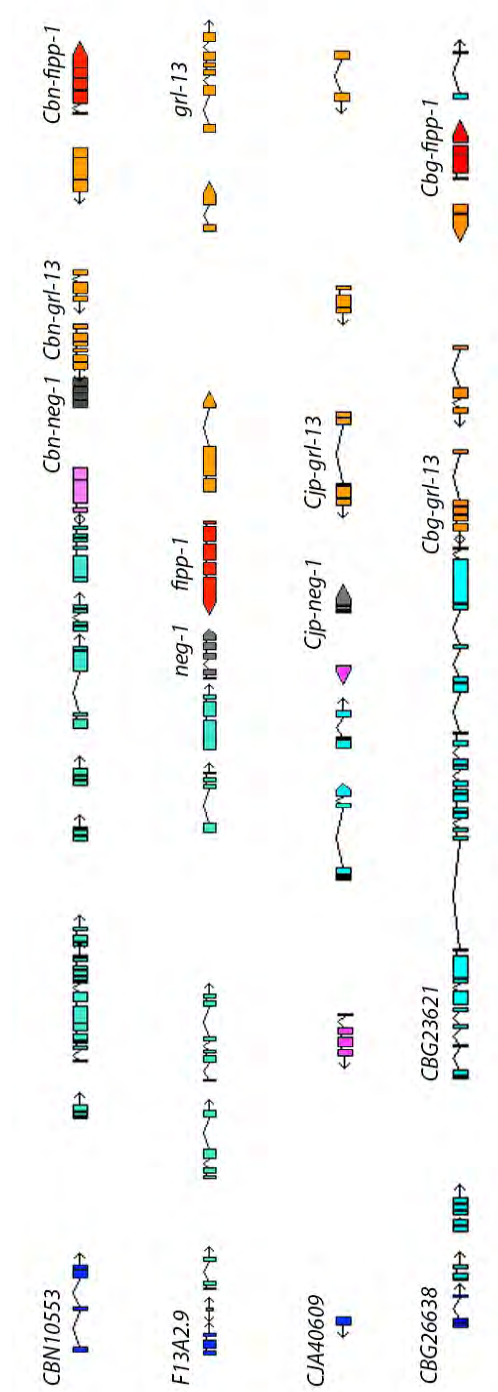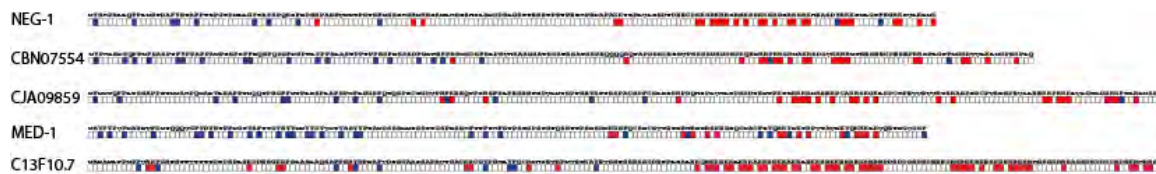

Figure 3.3

Figure 3.3. The genomic neighborhood of *neg-1* (grey) and its homologs *Cbn-neg-1* and *Cjp-neg-1* is heavily rearranged. Common features of their loci include the presence of *grl-13* and two adjacent genes (orange). A homolog of *fipp-1* (red) is present in near *neg-1* in *C. brenneri* but not in *C. japonica*, perhaps due to the contig ending near this region. F13A2.9 (blue) and its homologs demarcate one end of the genomic neighborhood. The locus in *C. brigssae* includes all landmarks but not *neg-1* homolog. Instead a larger predicted gene (CBG23621) resides where *neg-1* would be expected.

NEG-1 protein exhibits an skewed amino acid distribution

I noticed a striking asymmetry in the distribution of amino acids in NEG-1 and its two nematode homologues **(Figure 3.4)**. Aromatic hydrophobic amino acids tryptophan, phenylalanine, and tyrosine were concentrated in the first half of the protein sequence (9 in the N-terminal half and 2 in the C-terminal half, 9:2) whereas the basic amino acids arginine and lysine were concentrated in the second half (3:24). I wondered about the prevalence of such sequence asymmetry and searched the *C. elegans* proteome (WS234) for other skewed proteins. A query for proteins that had at least 3 times more aromatic hydrophobic residues (W/Y/F) in its N-terminal half and 3 times more basic residues (R/K) in its C-terminal half retrieved 41 proteins. Surprisingly, transcription factors including endo-mesoderm regulators MED-1, MED-2 and END-3 were over represented in this group of proteins with skewed sequences. For example, MED-1 sequence contains 22 residues of W, Y, or F in the N-terminal half and only 7 in the C-terminal. On the other hand, R and K residues do not exist in the N-terminal whereas 15 exist in the C-terminal half. Extending the query to *S. cerevisiae* retrieved the bZIP transcription factor GCN4. The Fly and Human proteome also include a number of skewed proteins that included transcription factors. A list of these results is compiled in **Table 3.1**.

Skewed distribution of aromatic and polar basic amino acids in NEG-1, its

homologues, MED-1 and C13F10.7

Figure 3.4

Figure 3.4. A diagram showing protein sequences as boxes each representing an amino acid. Aromatic amino acids Tryptophan, Tyrosine and Phenylalanine are skewed to the N-terminal half (blue). Basic amino acids Lysine and Arginine are skewed to the C-terminal half (red).

**Table 3.1. Proteins with skewed WFY/RK distribution (continued in appendix)**

Species	Name	length	N' FWY	C' FWY	FWY ratio	N' RK	C' RK	RK ratio	Functional Annotation
C. elegans	B0205.10	455	6	1	6.0	15	49	0.3	Uncharacterized protein
C. elegans	B0379.7	409	6	1	6.0	16	49	0.3	Uncharacterized protein
C. elegans	C04G6.10	166	5	0	5.0	5	17	0.3	Uncharacterized protein
C. elegans	C13F10.7	222	9	0	9.0	11	47	0.2	Putative zinc finger
C. elegans	C40H5.2B	115	12	3	4.0	1	6	0.2	Uncharacterized protein
C. elegans	D2062.4A	132	13	4	3.3	3	12	0.3	Uncharacterized protein
C. elegans	**END-1**	**221**	**16**	**5**	**3.2**	**1**	**19**	**0.1**	**ENDoderm determining**
C. elegans	F07C6.2	115	7	2	3.5	1	14	0.1	Uncharacterized protein
C. elegans	F11C1.7a	132	7	1	7.0	3	12	0.3	Uncharacterized protein
C. elegans	F13H8.5a	439	28	9	3.1	9	29	0.3	Uncharacterized protein
C. elegans	**F32D1.6**	**176**	**9**	**2**	**4.5**	**3**	**24**	**0.1**	**NEG-1**
C. elegans	F35C5.12	126	14	4	3.5	4	16	0.3	Uncharacterized protein
C. elegans	F36D1.7	113	6	0	6.0	1	4	0.3	Uncharacterized protein
C. elegans	F42A9.6c	101	16	5	3.2	2	7	0.3	Uncharacterized protein
C. elegans	F44F4.9	118	8	2	4.0	6	19	0.3	Uncharacterized protein
C. elegans	F55F10.3	181	10	1	10.0	6	21	0.3	Uncharacterized protein
C. elegans	F56A11.6	419	18	1	18.0	22	92	0.2	Uncharacterized protein
C. elegans	F59B10.6	105	10	1	10.0	2	7	0.3	Uncharacterized protein
C. elegans	H12D21.5	160	11	2	5.5	5	20	0.3	Uncharacterized protein
C. elegans	HRP-2	108	10	3	3.3	3	11	0.3	human HnRNP A1 homolog
C. elegans	**JUNE-1E**	**164**	**12**	**1**	**12.0**	**2**	**21**	**0.1**	**JUN transcription factor homolog**
C. elegans	K02E10.5	132	10	2	5.0	3	10	0.3	Uncharacterized protein
C. elegans	K03B4.6	134	10	1	10.0	1	10	0.1	Uncharacterized protein
C. elegans	**MED-1**	**174**	**22**	**7**	**3.1**	**0**	**15**	**0.1**	**Mesoderm and Endoderm Determination**
C. elegans	**MED-2**	**174**	**22**	**6**	**3.7**	**0**	**14**	**0.1**	**Mesoderm and Endoderm Determination**
C. elegans	MIG-2	195	11	3	3.7	5	17	0.3	abnormal cell MIGration
C. elegans	PQN-75	539	28	4	7.0	9	35	0.3	Prion-like-(Q/N-rich)-domain-bearing protein
C. elegans	RSP-7	452	19	2	9.5	16	92	0.2	SR Protein (splicing factor)
C. elegans	SST-20	231	14	3	4.7	1	28	0.0	Sperm Specific Transcript
C. elegans	T22C1.4	111	4	1	4.0	2	12	0.2	Uncharacterized protein
C. elegans	T25G12.3	204	15	1	15.0	4	23	0.2	Uncharacterized protein
C. elegans	Y38H6A.3	247	13	4	3.3	5	16	0.3	Uncharacterized protein
C. elegans	Y46G5A.23	108	4	0	4.0	1	8	0.1	Uncharacterized protein
C. elegans	Y57G7A.2	177	10	3	3.3	4	17	0.2	Uncharacterized protein
C. elegans	Y67D8B.3	119	11	0	11.0	4	15	0.3	Uncharacterized protein
C. elegans	Y69A2AR.8b	120	7	0	7.0	1	23	0.0	Uncharacterized protein
C. elegans	Y73B3A.1	659	27	7	3.9	18	62	0.3	Uncharacterized protein
C. elegans	Y95B8A.2	128	9	1	9.0	2	8	0.3	Uncharacterized protein
C. elegans	ZC250.4	103	5	0	5.0	2	7	0.3	Uncharacterized protein
C. elegans	ZC416.2	104	9	2	4.5	0	5	0.2	Uncharacterized protein
C. elegans	**ZIP-11**	**228**	**12**	**3**	**4.0**	**4**	**21**	**0.2**	**bZIP transcription factor**

NEG-1 interactors via yeast two hybrid

In order to gain insight into the role played by NEG-1, I conducted a yeast two hybrid screen using NEG-1 as bait and a library of *C. elegans* cDNA as preys (Walhout and Vidal, 2001). The results of this screen are summarized in **Table 3.2**. Of note are the interactions with the bromodomain protein BET-1, which appeared both in the pilot screen aimed at optimizing my transformation conditions and in the main screen conducted afterwards **(Figure 3.5)**.

EEF-1G and BET-1 interact with NEG-1 in a yeast-two-hybrid screen

-Leu -Trp -His + 40mM 3AT   -Leu -Trp -Ura   XGal
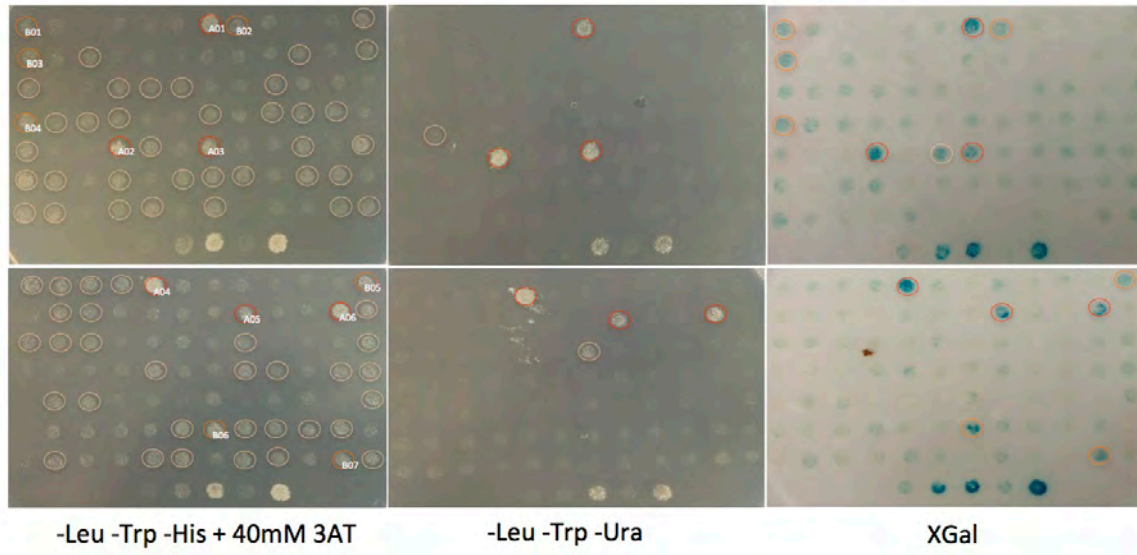
Figure 3.5

Figure 3.5. Three selection assays were used to confirm yeast-2-hybrid protein interactions: the ability to yeast cells grow in absence of histidine and presence of 40mM 3-amino-1,2,4-triazole (3-AT), the ability to grow in absence of uracil and the expression of LacZ. Red circles identify colonies that scored positive on all three tests, orange circles identify colonies that scored positive on two tests only. Tan circles scored positive on – His 40mM 3AT alone.

Table 3.2. Results from NEG-1 yeast-two-hybrid screen

Bait	Prey	-His	-Ura	XGal	N	Pilot	N	Annotation
NEG-1A	EEF-1G	✓	✓	✓	7	✓	1	Eukaryotic elongation factor gamma
NEG-1A	BET-1	✓	✗	✓	1	✓	2	Bromodomain protein
NEG-1A	RPL-7	✓	✗	✓	1	✗	0	Ribosomal protein
NEG-1A	GPD-2/3	✓	✗	✓	1	✗	0	Glyceraldehyde 3-phosphate dehydrogenase
NEG-1A	CEY-2	✓	✗	✓	1	✗	0	*C. elegans* Y-box
NEG-1A	ASP-1	✓	✗	✓	1	✗	0	Aspartyl protease
NEG-1A	ACT-5	✓	✗	✓	1	✗	0	Actin
NEG-1A	PDI-1	✓	✗	✓	1	✗	0	Protein Disulfide Isomerase
NEG-1A	COL-152	✓	✗	✓	1	✗	0	Collagen

In a parallel effort, we tested if NEG-1 interacts with transcription factors or RNA binding proteins by using a direct Y2H mating assay. NEG-1 was used as a prey against an array of RNA binding protein baits (Tamburino et al., 2013). In this assay NEG-1 interacted with C13F10.7, a putative zinc finger with a zinc knuckle domain (CX2CX4HX4C where X can be any amino acid) and a protein with skewed sequence like NEG-1 (**Figure 3.4** and **Table 3.1**). Interestingly, C13F10.7 interacted with the chromodomain protein CEC-8 in a proteome-wide Y2H screen (Li et al 2004). Moreover, in a computational study, C13F10.7 is predicted to interact with histone acetyl-transferases MYS-1 and MYS-2 (Zhong and Sternberg, 2006). *mys-1/2* have recently been shown to affect the nuclear localization of BET-1 (Shibata et al., 2010).

We also used NEG-1 as a bait against an array of transcription factors as preys (Reece-Hoyes et al., 2011). No interactions were detected from this assay.

Taking together, our yeast two hybrid screens and previous published physical and predicted interactions. NEG-1 appears to be nested in a network of interactions related to chromatin remodeling (**Figure 3.6**).

The putative zinc finger C13F10.7 is a NEG-1 interactor and interacts with

chromatin remodeling factors

A

>C13F10.7
MANANLTPMPFVRKFGSTSTTVTTTTDTDSPLSKDDESGKKPSLASLAQSAFPSRYDP
NAFVDISGAAISAPAVTGA**CKRC**GY**P**G**H**LYF**Q**CRNHIEVRPNVSTKAYEVSSTSSESS
DDETPLIALEKQRKKEKKLKKKAKKEKKKAKKLEKKERKKERKEKRRRRDDSDDSSDS
DEEEDRKRKKSKESKKEKRKRRHSSSDDRSASSDSDDDCRDSKRHRRS

B

— Y2H, this thesis
— Predicted, Zhong & Sternberg, 2006
— Physical, Li *et al.*, 2004
→ Functional, Shibata *et al.*, 2010

BET-1 *BET (two bromodomains) family protein*
CEC-8 *C.Elegans Chromodomain protein*
MYS-1/2 *MYST family histone acetyltransferase-like*
CIR-1 *CIR (transcription factor CBF1 Interacting coRepressor) homolog*
SIG-7 *Silencing In Germline defective*
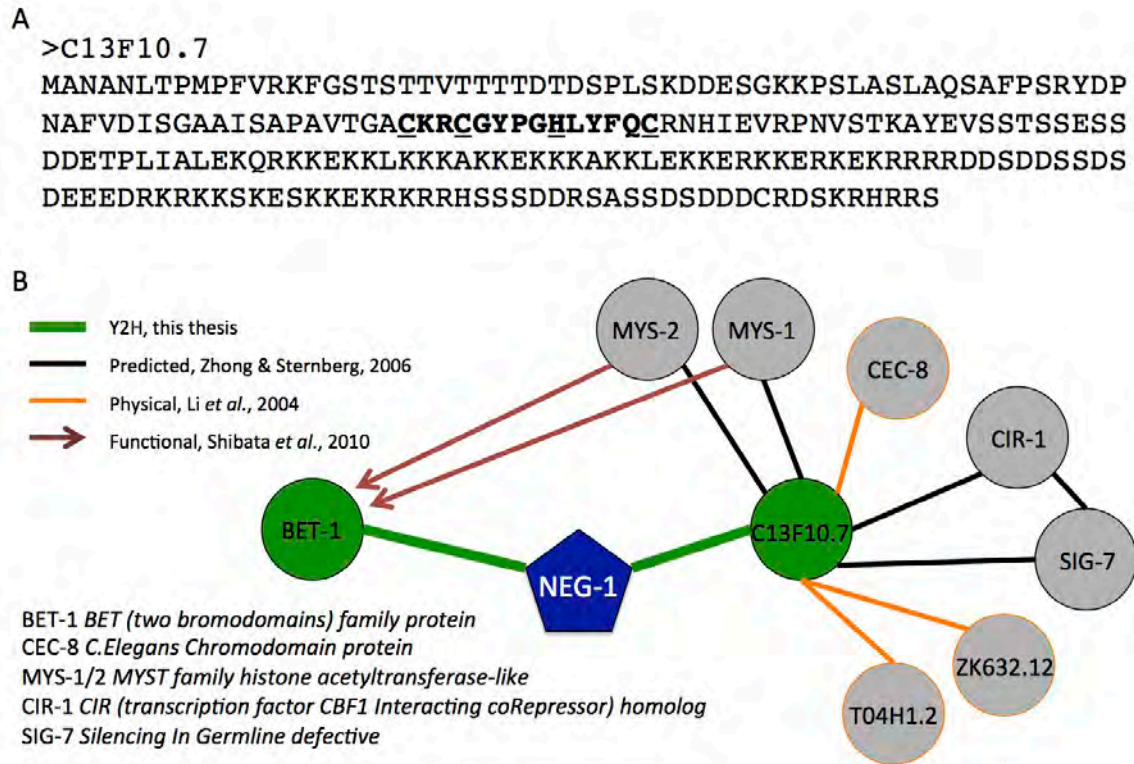

Figure 3.6

Figure 3.6. (A) C13F10.7 is a putative CCHC zinc finger with no known function. (B) Yeast two hybrid screens suggest that NEG-1 physically interacts with bromodomain protein BET-1 and the C13F10.7 (green). Additional protein interaction (orange edges), predicted functional interactions (black edges) and experimentally validated functional interactions (arrows) are shown.

Characterization of *neg-1* loss of function

We obtained a knockout allele of *neg-1* that removed a region encoding the last 97 amino acids. The deletion also removed the terminating stop codon of the gene thus fusing the coding region with the 3'UTR and adding 11 codons to the predicted gene product before encountering an in-frame stop codon **(Figure 3.7)**. Similar to *neg-1(RNAi)*, *neg-1(tm6077)* restored endoderm specification in *pos-1* dead embryos (97%). Furthermore, *neg-1(tm6077)* was essential for embryonic development as 79% of *neg-1(6077)* embryos died before hatching with defects in anterior morphogenesis **(Figure 3.7)**. Whereas the posterior half of *neg-1* embryos enclose and elongate, the anterior does not, due to incomplete hypodermal development. Furthermore, in 11% of dead embryos, a Glp-like incomplete pharynx was observed (data not shown).

neg-1 is required for embryogenesis
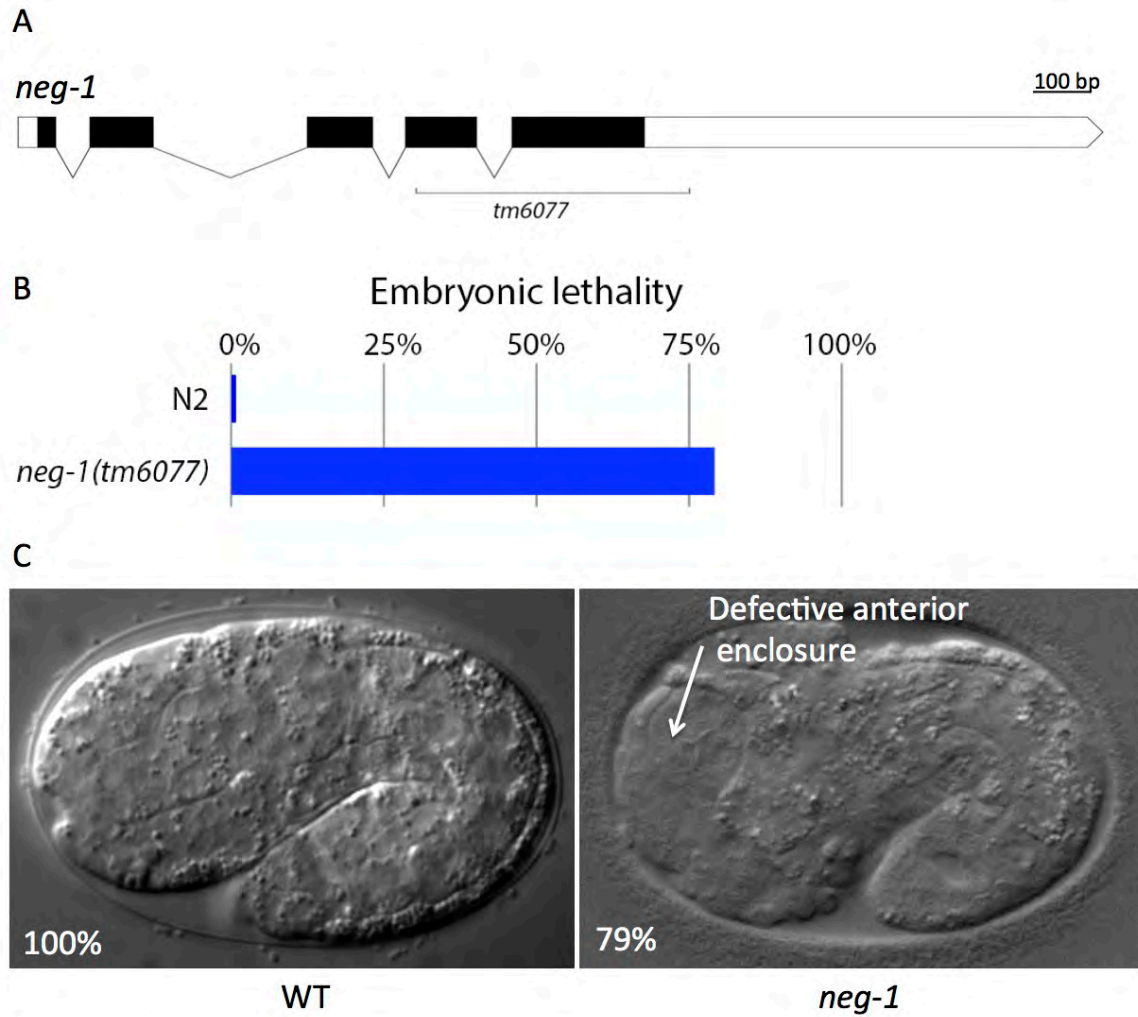
A

B

C

Figure 3.7

Figure 3.7. (A) The *neg-1* gene consists of five exons (black boxes) and a long annotated 3'UTR (white). The region deleted by *tm6077* is shown. (B) More than 75% of *neg-1(tm6077)* embryos die during embryogenesis. (C) *neg-1* dead embryos exhibit anterior morphological defects.

The mesoderm marker *pha-4* is misexpressed in *neg-1* mutant embryos

The expression of FoxA transcription factor *pha-4* serves as a marker for mesoderm development (Horner et al., 1998). We noticed that *pha-4* was inappropriately expressed in the anterior AB lineage **(Figure 3.8)**. In wildtype embryos only two of the eight AB granddaughters (ABalp and ABara) express a mesoderm fate in response to MS induction. In *apx-1* mutants, three additional AB granddaughters ABpra, ABprp and ABplp also express a mesoderm fate in response to MS induction (Mello et al., 1994). However, in *neg-1(6077)* embryos, the ABala, ABarp and ABpla granddaughters express *pha-4*. This means that three additional AB granddaughters express a mesoderm fate, but intriguingly are not the three granddaughters that undergo the Apx misspecification. Therefore, the *neg-1* embryos exhibit ectopic mesoderm differentiation in the AB lineage different from that observed in Apx embryos and more in tune with a subtle Mex phenotype.

In addition to this gain of ectopic expression, *pha-4* is also reduced in the ABalp and ABara that normally express this marker in wildtype embryos **(Figure 3.8)**.

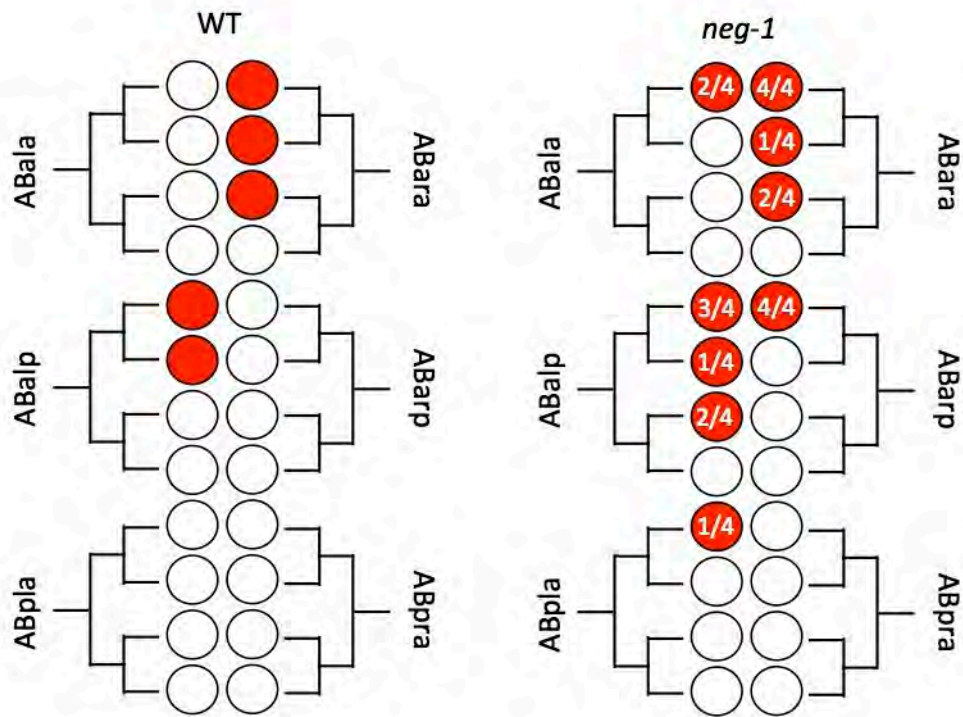**The mesoderm marker *pha-4* is misexpressed in *neg-1* mutant embryos.**

Figure 3.8

Figure 3.8. PHA-4::GFP (solid red) is expressed in the ABalp and ABara lineages of wildtype embryos. *neg-1(tm6077)* embryos ectopically express *pha-4* in ABala, ABarp and ABpla lineages. Wildtype embryos express *pha-4* in the anterior sublineage of ABalp. *neg-1(tm6077)* exhibit a gain of *pha-4* expression in the posterior sublineage of ABalp (2/4) or a loss of expression in both sublineages (1/4). ABara expresses *pha-4* in both anterior and posterior sublineages. ABara *pha-4* expression is reduced in *neg-1(tm6077)*. Numbers reflect the number of times *pha-4::gfp* expression was detected in a given cell in a lineaged embryo.

neg-1 worms exhibit behavioral defects

Since *neg-1* embryos exhibit defects in anterior ectoderm development including ectopic mesoderm differentiation, I wondered if the embryos that do hatch lived with a defective nervous system. Such a subtle defect would not be lethal but would alter behavior. No gross behavioral defects are observed in *neg-1* adults. However, a systematic behavioral survey revealed a defect in osmotic avoidance.

The glycerol drop test is a behavioral assay that tests osmotic avoidance. In this assay, a drop of glycerol is delivered near the tail of a worm as it moves forward. The glycerol drop instantly surrounds the worm and reaches the anterior sensory organs. Wildtype worms immediately sense the glycerol as a repellent and start moving backward. However, *neg-1* larvae showed an approximately 2-fold decrease in their avoidance index compared to wildtype worms **(Figure 3.9).**

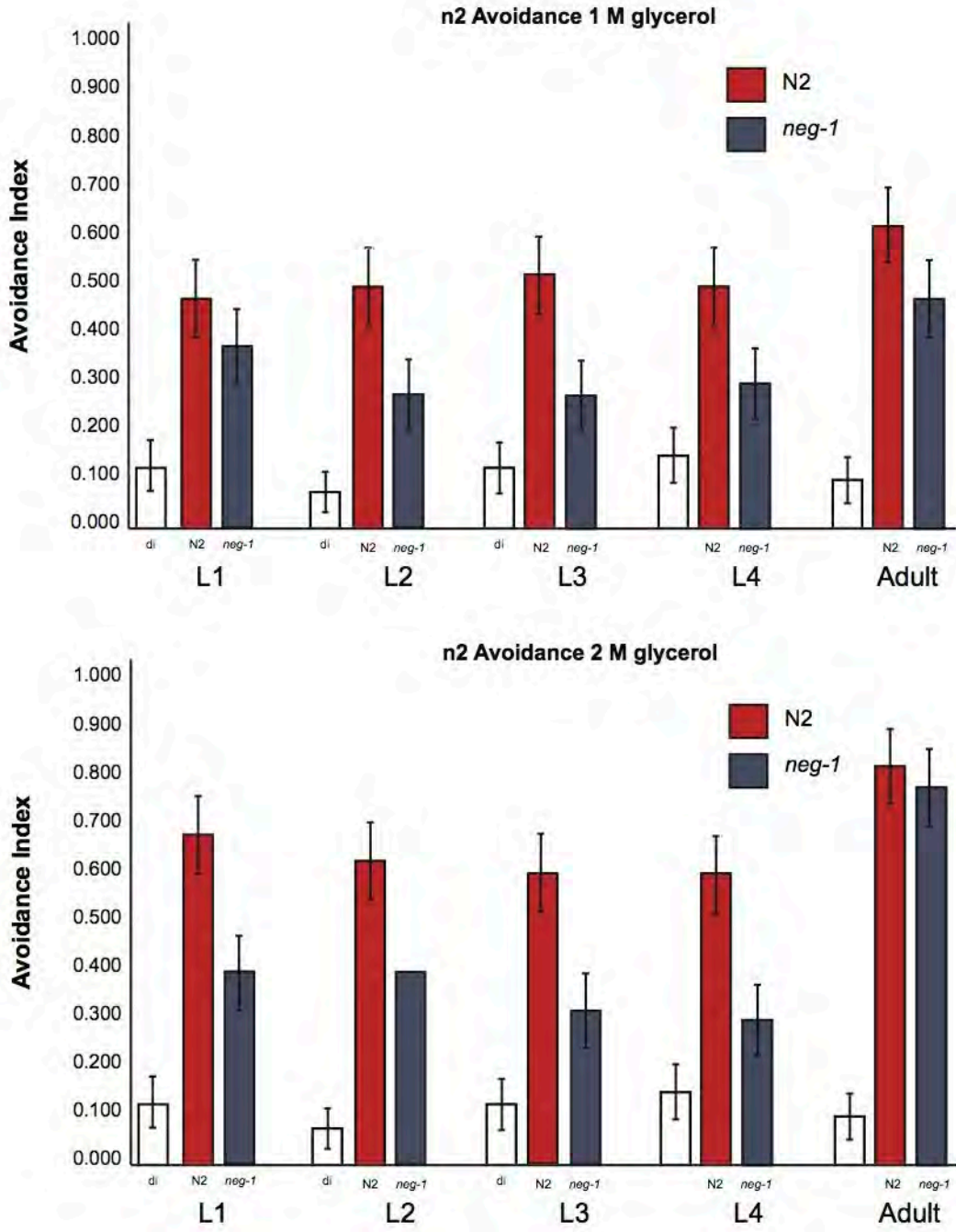**neg-1 worms are defective in osmotic avoidance**

Figure 3.9

Figure 3.9. Glycerol avoidance assay. A drop of glycerol is delivered near the tail of a worm as it moves forward. Wildtype respond by moving backwards, which is considered a positive response. The avoidance index is the number of positive responses divided by the total number of trials (n = 40 trails per genotype stage). Deionized water (di). Error bars are standard error of the mean.

NEG-1 is asymmetrically expressed in the early embryo

In order to investigate the spatial-temporal expression of NEG-1, we engineered a *neg-1::gfp* transgenic worm using MosSci transgenesis (Frokjaer-Jensen et al., 2008) and found that its expression was patterned in the early embryo **(Figure 3.10)**. NEG-1::GFP was detected in the zygotic nucleus and at equal levels in both nuclei of the two cell stage embryo (23/23). At the four-cell stage, NEG-1::GFP expression was higher in nuclei of the anterior ABa and ABp blastomeres compared to the nuclei of EMS and P2 (31/34). Following the four-cell stage, NEG-1::GFP was detected in the granddaughters of the AB blastomere and to a lesser degree in the great granddaughters before henceforth going undetected (data not shown). We also observed NEG-1::GFP in the nuclei of distal germ cells as well as nuclei of maturing oocytes **Figure 3.10**. We also noticed an intense sub-nucleur localization of NEG-1:GFP suggesting its localization along condensed chromatin **(Figure 3.10)**.

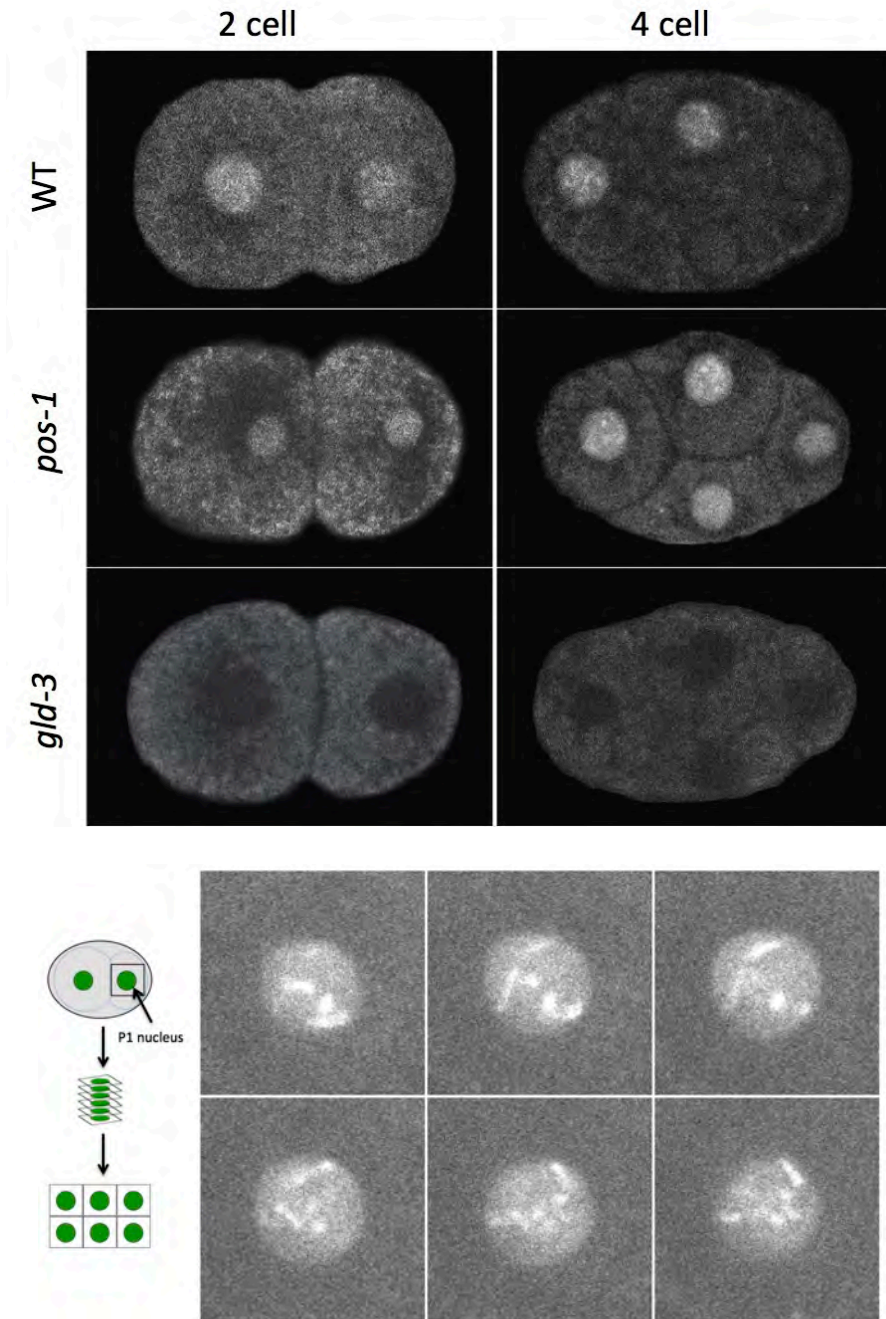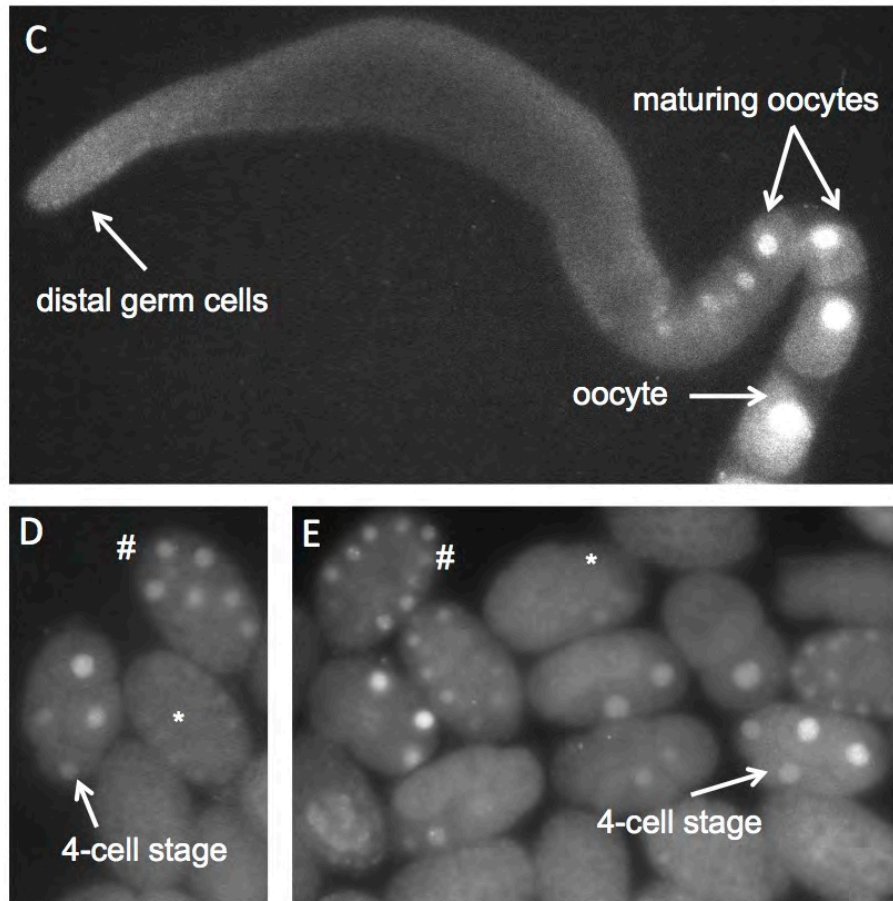**NEG-1 is asymmetrically expressed in the early embryo and co-localizes with chromatin**

Figure 3.10

Figure 3.10 continued

Figure 3.10. (A) NEG-1::GFP is detected in the both nuclei in the two-cell stage but only in ABa and ABp in the 4 cell stage. *pos-1(RNAi)* embryos show NEG-1::GFP expression in all four cells. NEG-1::GFP expression is lost in *gld-3(RNAi)* embryos. (B) Subnuclear NEG-1::GFP localization is most intense along chromatin. (C) NEG-1::GFP is expressed in germline with localization restricted to distal germ cells and maturing and mature oocytes. (D and E). GFP::*neg-1* 3'UTR reporter recapitulates asymmetric expression of NEG-1::GFP. Arrows highlight asymmetric expression in 4-cell stage. (#) highlight expression in early embryos (AB8 and AB16 stages), whereas (*) highlight lack of expression in later embryos (post AB16 stage).

neg-1 asymmetric expression is dependent on *pos-1* and *gld-3*

To test our prediction that *neg-1* is a target of POS-1 repression, we examined the effect of *pos-1* knockdown on NEG-1 localization. Ectopic NEG-1::GFP was observed in the posterior blastomeres EMS and P2 upon *pos-1* RNAi (14/14) **(Figure 3.10)**. A similar effect was observed in the *pos-1(zu148)* background (6/6). Our initial model postulated that the production of NEG-1 would be dependent on GLD-3 and GLD-2. No NEG-1::GFP was detected in the nuclei of one-cell or two-cell stage embryos following *gld-3* RNAi (0/7 and 0/11 respectively) **(Figure 3.10)**. During the four-cell and 6-to-8 cell stages, weak NEG-1::GFP could be detected in the AB lineage but not in that of P1 (4/11 and 6/12 respectively). The same loss of embryonic NEG-1::GFP was observed following *gld-2* RNAi, albeit to a lesser degree (2-cell: 4/14, 4-cell: 6/8, 6-to-8 cell: 4/5). However, this difference is likely due to shorter exposure to *gld-2* dsRNA, intentionally curtailed to avoid sterility. Taken together, we conclude that *pos-1* down-regulates the expression of NEG-1 in posterior blastomeres and that *gld-3* and *gld-2* are required for the expression of *neg-1* during the one- and two-cell stages and partially thereafter.

gld-3 is downstream of *pos-1* in the regulation of *neg-1* expression

Importantly, *pos-1(zu148)* worms expressing *neg-1::gfp* and growing on *gld-3* RNAi food developed embryos with no NEG-1::GFP expression (0/6). This means that with regards to *neg-1* expression, loss of both *pos-1* and *gld-3* resemble the phenotype of *gld-3* and that *gld-3* is therefore downstream of *pos-1* in the regulation of *neg-1* expression.

neg-1 mRNA in early embryos have shorter poly(A) tails after *gld-3* or *gld-2* knockdown but longer tails after *pos-1* knockdown

Since NEG-1::GFP is reduced in *gld-3(RNAi)* and *gld-2(RNAi)* embryos, we asked if *gld-3* and *gld-2* were required for the polyadenylation of *neg-1* mRNA. Several *C. elegans* transcripts expressed in the germline have been shown to be dependent on GLD-2 for their polyadenylation, which in turn precedes translation activation (Janicke et al., 2012; Kim et al., 2010). We assayed the length of *neg-1* poly(A)-tails using extension Poly-A Test (ePAT), which allows an efficient and rapid assessment of poly(A)-tail lengths with a sensitivity comparable to Northern blots but without the need of radiolabeled probes (Janicke et al., 2012). 1µg of total RNA isolated from wild-type (N2), *pos-1(RNAi)*, *gld-3(RNAi)* or *gld-2(RNAi)* early embryos was subjected to ePAT. Compared to *neg-1* mRNA from wildtype worms, the range of poly(A) tail length was longer in and *pos-1(RNAi)* early embryos and shorter after *gld-3* and *gld-2* knockdown **(Figure 3.11)**.
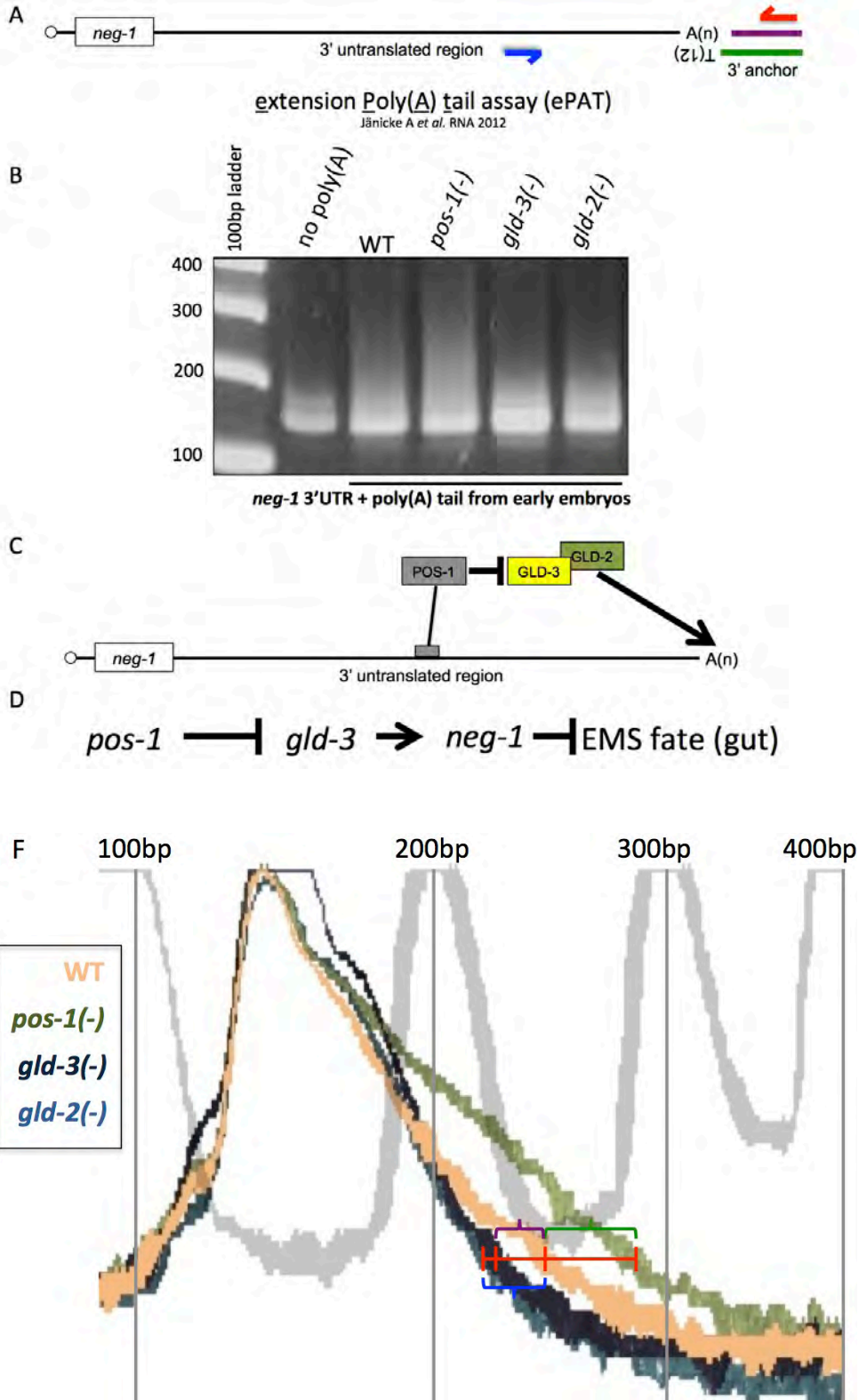
neg-1 poly(A) tail is affected by *pos-1* and *gld-3/2*

Figure 3.11

Figure 3.11. (A) Extension Poly(A) Test. 1μL of PAT anchor 5′-GCGAGCTCCGCGGCCGCGTTTTTTTTTTTTT (green) is added to 1μg of total RNA and Klenow Polymerase extends the poly(A)tail with a sequence complementary to the anchor (purple). Reverse transcription using a primer (red) that recognizes the extension (purple) is conducted to obtain cDNA. Then, 25 cycles of PCR amplification using a 3'UTR specific forward primer (blue) amplifies the poly(A) tails. (B) Smears reflect the extent of *neg-1* poly(A) tail length in each genetic background. The range of *neg-1* poly(A)-tails is longer in *pos-1(-)* early embryos and shorter in *gld-3(RNAi)* or *gld-2(RNAi)* early embryos. RNA was isolated by Ahmed Elewa ePAT reaction and gel were done by Traude Beilharz. (C) A model of *pos-1* repression of *neg-1* expression: POS-1 binds the 3'UTR of *neg-1* and prevents GLD-3/2 dependent polyadenylation. (D) A genetic pathway for EMS specification: *pos-1* represses *gld-3*, which promotes *neg-1*, which in turn inhibits EMS specification. (F) ePAT provides a semi-quantitative readout for poly(A)-tail length. The intensity of smears in (B) was quantified using LabWorks gel documentation software. Intensity curves were overlaid in comparison to a 100bp ladder (grey). Poly(A)-tails from *pos-1(RNAi)* embryos (green) exhibit a longer range of intensity compared to poly(A)-tails from wildtype (yellow). Conversely, poly(A)-tails from *gld-3(RNAi)* (purple) or *gld-2(RNAi)* embryos (blue) exhibit a shorter range of intensity.

The 3'UTR of *neg-1* is sufficient to confer asymmetric expression in the early embryo

3' untranslated regions regulate the majority of patterned gene expression in the germline and early embryo (Merritt et al., 2008). To confirm that the asymmetric expression of NEG-1::GFP was regulated by the *neg-1* 3'UTR, we expressed GFP with a nuclear localization sequence under the promoter of *oma-1* and the 3'UTR of *neg-1*. Expression under the *oma-1* promoter allows for germline transcription of the *gfp::neg-1 3'UTR*, the transcripts of which would be deposited in oocytes. We observed a pattern of GFP expression in the early embryo identical to that of NEG-1::GFP **(Figure 3.10)**. We therefore conclude that the 3'UTR of *neg-1* is sufficient to confer asymmetric expression in the early embryo.

POS-1, MEX-3 and MEX-5 bind the 3'UTR of *neg-1 in vitro*

The 3'UTR of *neg-1* includes a cluster of 3 overlapping predicted RNA binding protein (RBP) elements, referred to here as the RBP cluster. This RBP cluster begins with a uracil rich sequence that qualifies as a MEX-5 binding region and which becomes part of overlapping MEX-3 and POS-1 predicted binding elements (Farley et al., 2008; Pagano et al., 2009; Pagano et al., 2007). To address whether these proteins physically bind to the RBP cluster we conducted electrophoretic mobility shift (EMSA) and/or fluorescence polarization (FP) assays to determine their affinity to this sequence. POS-1, MEX-3 and MEX-5 bind the RBP cluster ($K_{d,app}$ = 76 nM, 40nM and 88 nM respectively). These affinities are comparable to those determined between each protein and confirmed biological targets. Mutating the MEX-3 or POS-1 binding elements

prevented their corresponding proteins from interacting ($K_{d,app}$ = 532 and >1000 nM respectively) (**Figure 3.12**).

A second MEX-5 binding region (M5FW) is predicted to exist upstream of the RBP cluster. MEX-5 interacted with this region with an affinity of $K_{d,app}$ = 125 nM. An additional POS-1 binding element overlapping with a MEX-3 consensus sequence is predicted to reside downstream of the RBP cluster (OLWT). However, neither site showed binding to its corresponding protein *in vitro* (**Table 3.3**). A summary of interactions is presented in **Figure 3.13**.

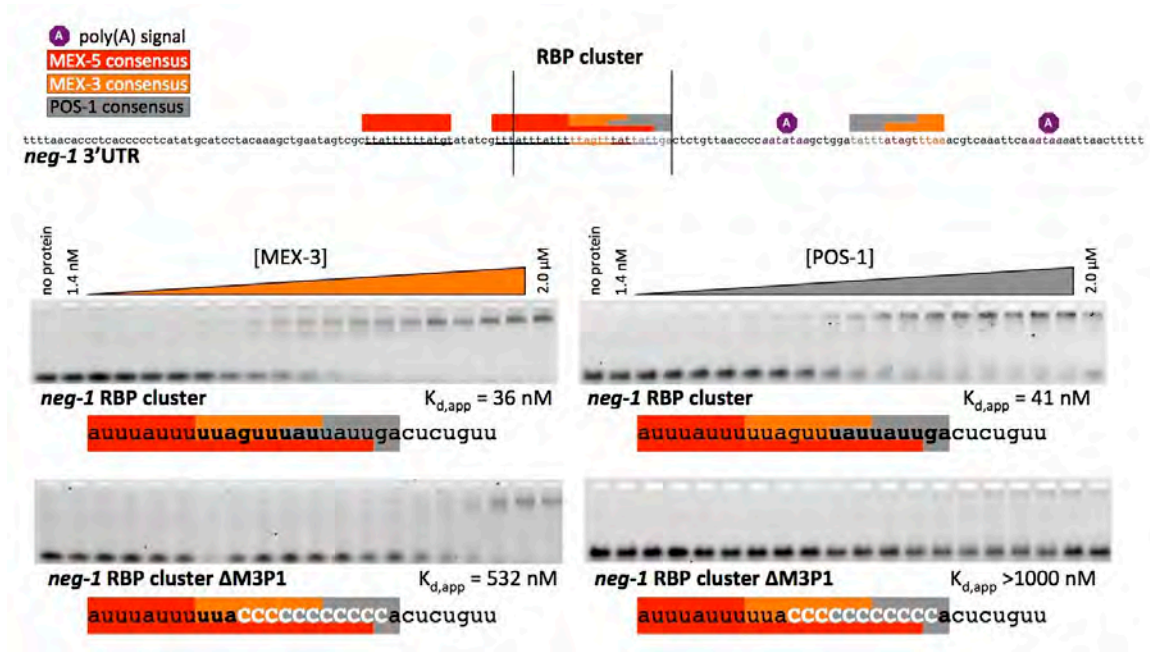# POS-1, MEX-3 and MEX-5 bind the *neg-1* 3'UTR

Figure 3.12

Figure 3.12. Predicted binding sites for POS-1 (grey), MEX-3 (orange) and MEX-5 (red) are mapped onto the *neg-1* 3'UTR. Predicted polyadenylation signals are in purple hexagons. The RBP cluster includes overlapping sites of the three proteins. Electrophoretic mobility shift assay gels demonstrate that MEX-3 and POS-1 bind the RBP cluster *in vitro*. This binding is lost upon mutating their predicted binding sites.

Table 3.3. Summary of in vitro binding assays between POS-1, MEX-5, MEX-3 and the *neg-1* 3'UTR

Protein	RNA	Sequence	Assay	Ave $K_{d,app}$ (nM)	SD	N
POS-1	RBPc WT	auuuauuu**uuaguuuauuauuga**cucuguu	EMSA	**76**	35	5
POS-1	RBPc Mut	auuuauuu**uuaCCCCCCCCCCa**cucuguu	EMSA	>1000	n/a	3
POS-1	OLWT	uaagcugga**uauuuauaguuuaa**acgucaa	EMSA	>1000	n/a	1
POS-1	RBPc WT	auuuauuu**uuaguuuauuauuga**cucuguu	FP	**59**	6	3
MEX-3	RBPc WT	auuuauuu**uuaguuuauuauuga**cucuguu	EMSA	**37**	5	6
MEX-3	RBPc Mut	auuuauuu**uuaCCCCCCCCCCa**cucuguu	EMSA	532	28	3
MEX-3	OLWT	uaagcugga**uauuuauaguuuaa**acgucaa	EMSA	1018	n/a	2
MEX-3	RBPc WT	auuuauuu**uuaguuuauuauuga**cucuguu	FP	**117**	19	6
MEX-3	RBPc Mut	auuuauuu**uuaCCCCCCCCCCa**cucuguu	FP	530	28	3
MEX-3	OLWT	uaagcugga**uauuuauaguuuaa**acgucaa	FP	>1000	n/a	3
MEX-5	M5FW	agucgc**uuauuuuuuaugu**auaucg	FP	**125**	6	3
MEX-5	RBPc WT	auuuauuu**uuaguuuauuauuga**cucuguu	FP	**88**	60	3

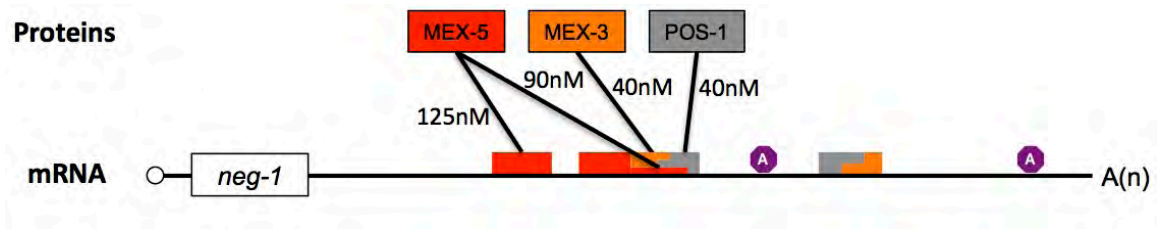**Summary of Protein-RNA interactions**

Figure 3.13

Figure 3.13. A diagram summarizing in vitro assays confirming the MEX-5 (red), MEX-3 (orange) and POS-1 (grey) interact with the 3'UTR of *neg-1*. Dissociation constants are averaged and rounded from both EMSA and FP assays.

### *mex-5* but not *mex-3* is required for *neg-1* expression in the early embryo

The binding of MEX-5 and MEX-3 to the *neg-1* 3'UTR raised the question of whether these proteins are required for *neg-1* expression. *mex-3(RNAi)* had no affect on the NEG-1::GFP expression in the one and two-cell stages (2/2 and 9/9 respectively). Furthermore, the asymmetric expression of NEG-1::GFP during the four-cell stage was similar to that observed in wildtype embryos (9/9). However, NEG-1::GFP was lost upon *mex-5* knockdown (0/30) **(Figure 3.14)**.

mex-5 is upstream of *pos-1* in the regulation of *neg-1* expression

Loss of NEG-1 expression in *mex-5(RNAi)* early embryos is opposite to the consequence of *pos-1* loss of function, whereby NEG-1::GFP is observed in all four cells of the 4-cell stage. *mex-5* knockdown in a *pos-1(zu148); neg-1::gfp* background resulted in an NEG-1::GFP expression pattern identical to that observed upon loss of *pos-1* function alone. All *pos-1(zu148); mex-5(RNAi)* early embryos expressed NEG-1::GFP (10/10) and 4 out of 4 embryos at the 4-cell stage showed NEG-1::GFP expression in all four cells. Therefore, *pos-1* is epistatic and genetically downstream to *mex-5* in the regulation of *neg-1* expression **(Figure 3.14)**.

mex-5 positively regulates _neg-1_ expression
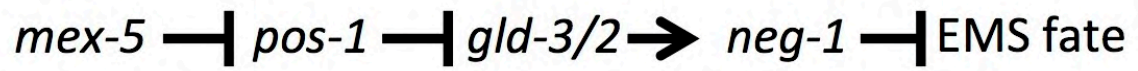
A
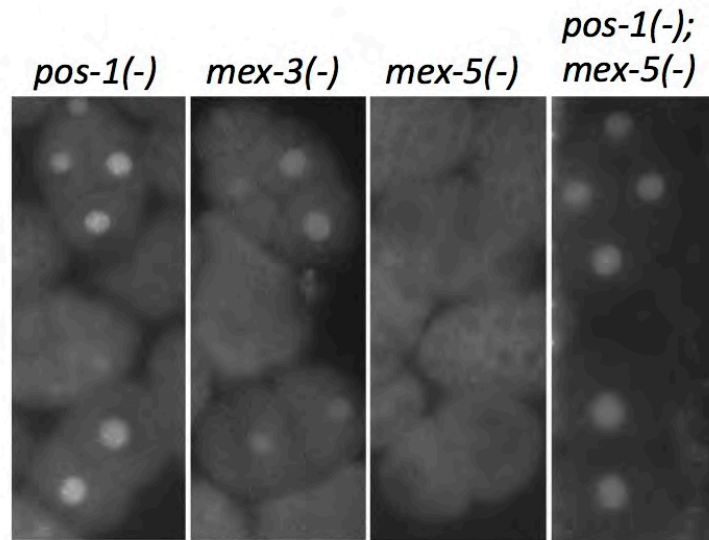
B

Figure 3.14

Figure 3.14. (A) NEG-1::GFP is observed equally in all four cells of *pos-1(RNAi)* embryos. *mex-3(RNAi)* does not affect *neg-1* expression pattern. *mex-5(RNAi)* prevents NEG-1::GFP expression. *pos-1(RNAi); mex-5(RNAi)* embryos express NEG-1::GFP equally in all four cells. (B) *mex-5* is upstream of *pos-1* in the regulation of *neg-1* expression and functions to counter *pos-1* repression.

neg-1 lethality is suppressed by loss of _med-1_ function

        The ectopic _pha-4_ expression observed in the AB lineage suggests inappropriate mesoderm differentiation. To test if this was the cause behind _neg-1_ embryonic lethality we built a _neg-1; med-1_ double mutant. _med-1_ is a key endo-mesoderm transcription factor that acts downstream of SKN-1 and upstream of END-3, an endoderm promoting transcription factor. _med-1(ok804)_ knockout worms are viable. It is only in conjunction with loss of _med-2_ function that endo-mesoderm differentiation is impaired (Maduro et al., 2001). _neg-1; med-1_ double mutants showed a dramatic reduction in embryonic lethality **(Figure 3.15)**. The percentage of dead embryos in the double mutant was reduced to 35.6% compared to 76.2% in _neg-1(tm6077)_. Therefore, _med-1_ is a suppressor of _neg-1_ embryonic lethality.

neg-1 embryonic lethality is reduced by med-1 loss of function
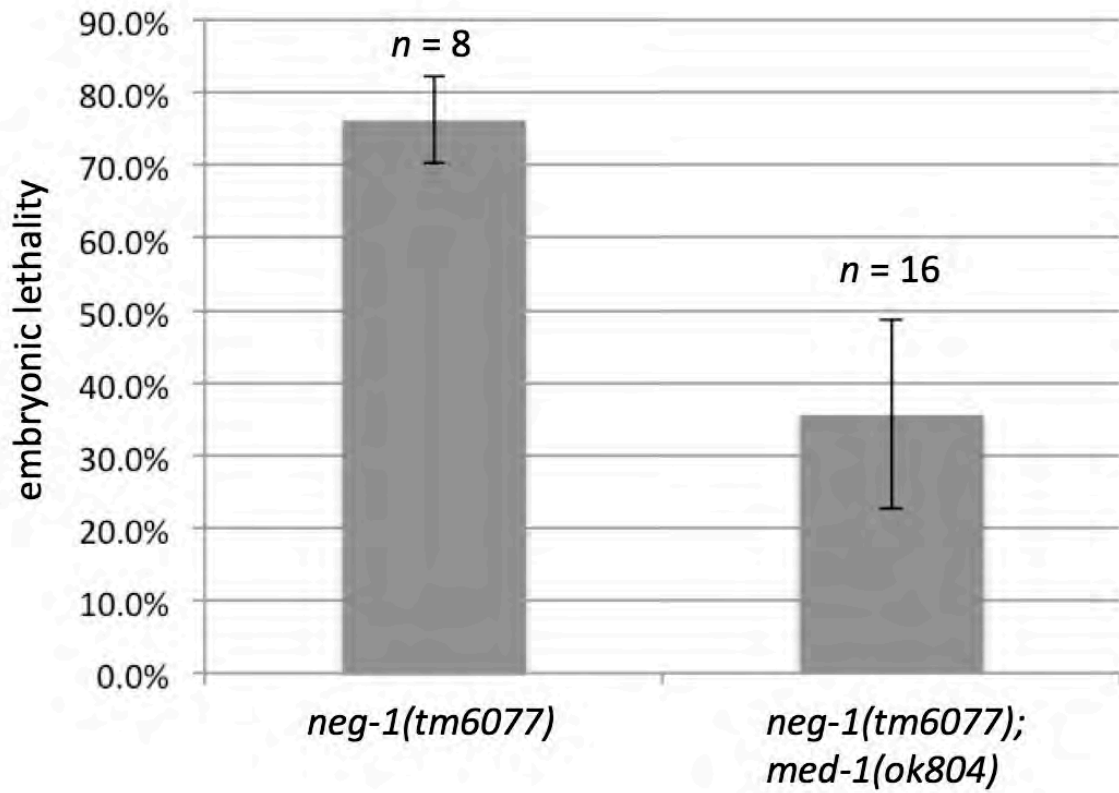

Figure 3.15

Figure 3.15. *neg-1* embryonic lethality is reduced from 76.2% to 35.6% of *neg-1(tm6077); med-1(ok804)* embryos. L4 hermaphrodite worms were singled and their progeny scored for embryonic lethality. N = number of hermaphrodites assayed. Error bars reflect standard deviation (6% and 13% respectively).

Model: *neg-1* expression is restricted by MEX-5 and POS-1 to the anterior blastomere lineage where it protects ectoderm fate by repressing endo-mesoderm differentiation.

Three lines of evidence suggest that *neg-1* is a repressor of endo-mesoderm fate. First, the mesoderm transcription factor *pha-4* is ectopically expressed in AB sublineages that normally differentiate into ectodermal fates (neurons and hypodermis). Second, the embryonic lethality of *neg-1* is suppressed by loss of *med-1* function, a primary mesoderm transcription factor. Third, ectopic *neg-1* expression in *pos-1* mutants represses endo-mesoderm fate of the EMS blastomere rendering *pos-1* embryos gutless. Overexpression of *med-1* restores endo-mesoderm differentiation.

neg-1 may be considered a protector of ectoderm fate for two reasons. First, *neg-1(-)* embryos die with defects in anterior development suggesting poor ectoderm differentiation and surviving worms exhibit behavioral defects indicative of neuronal mal-development. The lethality is reduced in the *neg-1; med-1* double mutant. Second, *neg-1* is normally expressed in the AB lineage, which gives rise to ectodermal cells and it is in that context that *neg-1* performs its essential mesoderm repressing function.

neg-1 expression in the anterior blastomere AB and its descendants is promoted by MEX-5, which can bind the *neg-1* 3'UTR and counter POS-1 repression. POS-1 in turn represses *neg-1* expression in posterior blastomeres EMS and P2 by preventing GLD-3/2 mediated polyadenylation of *neg-1* transcripts and thereby guards the endo-mesoderm fate of EMS (**Figure 3.16**).
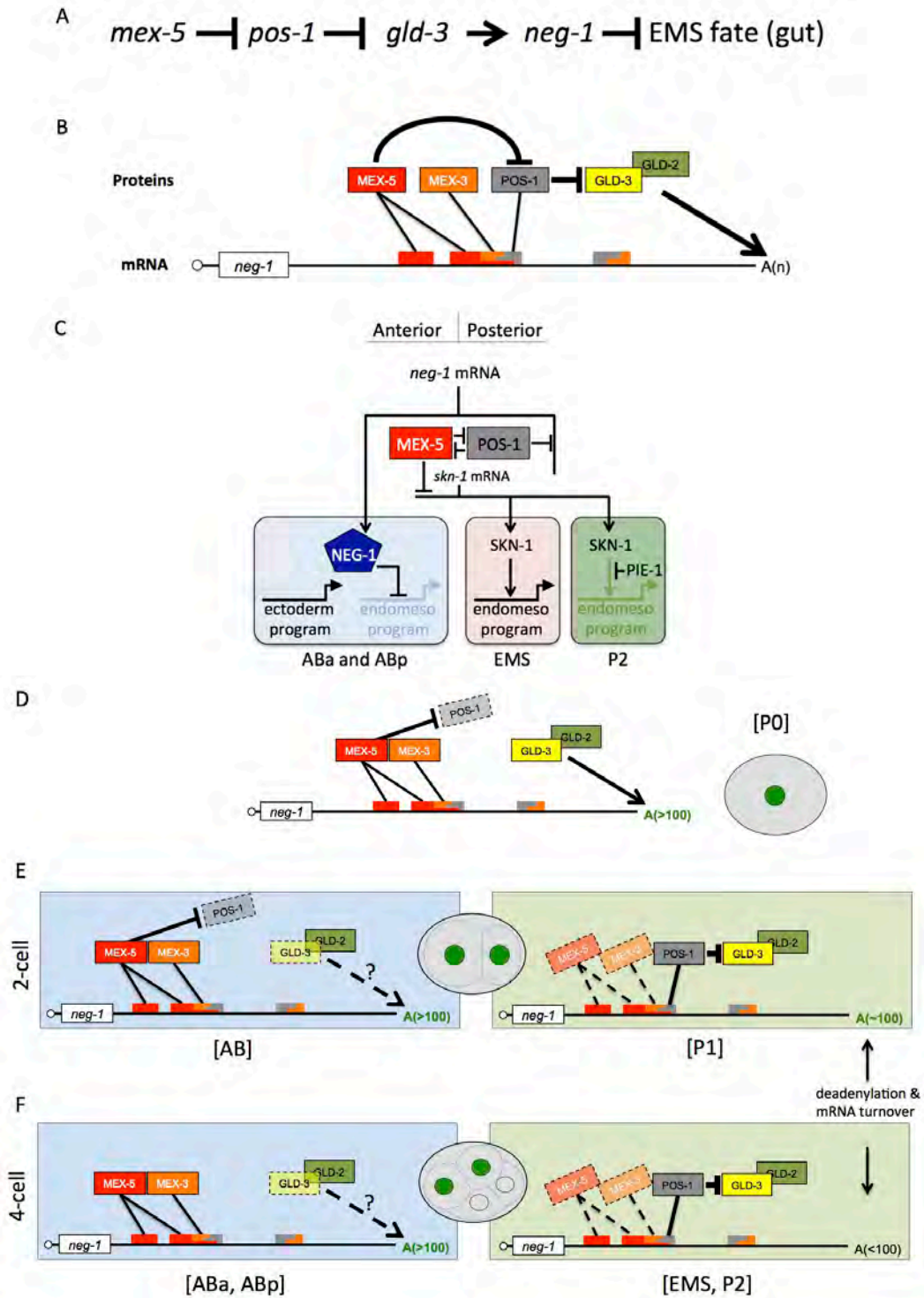
Model of *neg-1* regulation and function

Figure 3.16

Figure 3.16. (A) The genetic regulation of *neg-1* expression. *mex-5* repression of *pos-1* repression allows *gld-3* to positively regulate *neg-1*. *neg-1* expression represses the endo-mesoderm fate. (B) Molecular regulation of *neg-1* expression. MEX-5, MEX-3 and POS-1 bind the 3'UTR of *neg-1*. POS-1 prevents GLD-3/2 polyadenylation of *neg-1* and MEX-5 counters POS-1 repression. NEG-1 represses endo-mesoderm differentiation in anterior blastomeres. (C) NEG-1 protects ectoderm fate by repressing endo-mesoderm differentiation in AB lineage. MEX-5 both promotes *neg-1* and represses *skn-1* expression. SKN-1 protein is therefore reduced in AB daughter cells. Although SKN-1 is present in P2, transcription repressor PIE-1 represses the activation of somatic differentiation genes. (D) In the zygote P0, MEX-5 counters POS-1 repression and GLD-3/2 elongate *neg-1* poly(A) tails (**A(>100)**). (E) In the 2 cells stage, high MEX-5 and low POS-1 levels ensure *neg-1* expression in AB. It is not clear if GLD-3/2 act on *neg-1* in AB or if transcripts have extended poly(A) tails from the zygote stage, or if new GLD-3 independent polyadenylation occurs. *neg-1* transcripts with extended poly(A) tails from the zygote stage and/or residual MEX-5 may explain NEG-1 expression in P1 where POS-1 is expressed. (F) At the four cells stage high MEX-5 levels and no POS-1 ensure *neg-1* expression in the AB daughter cells. High POS-1 levels in EMS and P2 prevent new polyadenylation and continue to repress *neg-1* expression despite the presence of GLD-3/2. A background level of poly(A) deadenylation and/or mRNA turnover may explain why *neg-1* transcripts with tails elongated in the zygote stage are no longer expressed.

MATERIALS AND METHODS

neg-1::gfp strain construction

A region of cosmid F32D1 flanked by AflII and NotI digestion sites was subcloned and an ORF encoding GFP was introduced upstream of F32D1.6 *(neg-1)*. The engineered transgene preserved *neg-1* in its predicted operon and its downstream neighbor *fipp-1* and was introduced using MosSci transgenesis (Frokjaer-Jensen et al., 2008).

Cloning and of purification of POS-1 80-180 and MEX-5 236-350

Cloning and purification of recombinant POS-1 and MEX-5 was done as described in (Farley et al., 2008) and (Pagano et al., 2007) respectively.

Preparation of fluorescein-5-thiosemicarbazide labeled RNA

All RNA oligonucleotides used in this study were chemically synthesized and fluorescently labeled with fluorescein-5-thiosemicarbazide by Integrated DNA Technologies (IDT).

Fluorescent electrophoretic mobility shift and fluorescence polarization assays

Fluorescent EMSA and florescence anisotropy assays using recombinant POS-1 and MEX-5 and calculation of the apparent dissociation constant was done as described in (Farley et al., 2008) and (Pagano et al., 2007) respectively.

Protein sequence analysis

Using PERL programing, the number of W, F, Y, R and K residues in the each protein over 100 amino acids long and present in the C. elegans proteome (WS234) were

counted. The number of W, Y, and Fs in the N-terminal half was divided by the count in the C-terminal half. The reverse was done for R and K. Proteins with 3-fold or more WYF in the N-terminus and 3-fold or more R and K in the C-terminus were considered "skewed proteins". The same process was repeated with proteomes of *S. cerevisiae*, *D. melanogaster* and *H. sapiens*.

Yeast 2 Hybrid

Yeast-2-hybrid was done as described in (Walhout and Vidal 2001). Briefly NEG-1A-DB was constructed using Gateway technology. 2×10^{-6} yeast colonies were screened on –Leu –Trp –His +40mM 3AT upon transformation with a *C. elegans* AD-cDNA library (Gift from Laboratory of Marian Walhout, UMass Medical School). Growing colonies were retested on –Leu –Trp –His +40mM 3AT, -Uracil plates and by x-Gal staining to test for LacZ expression.

Chapter IV

Summary,

Conclusions and Discussion

The importance of repressing alternative identities during blastomere specification

The transition from totipotency to pluripotency is a literal reduction in potential fates. During nematode embryogenesis, the transition from germline totipotency to somatic stem cell pluripotency requires restricted expression of fate determinants. Classical *C. elegans* mutants with blastomere identity mishaps such of Pharynx and Intestine in Excess (*pie-1*) or Muscle EXcess (*mex-5/6* and *mex-3*), involve a repressor restricting the expression of a master transcription factor that activates a differentiation program (*skn-1* or *pal-1*). Here I describe a new theme: a translational repressor restricting the expression of a fate antagonist.

In the case of somatic stem cell EMS, SKN-1 promotes the endo-mesoderm fate. Yet the potential is not realized unless POS-1 represses *neg-1*, an endo-mesoderm antagonist. On the other side of the embryo where ectodermal differentiation occurs, MEX-5 represses SKN-1. In *mex-5(-)* embryos the anterior ectoderm lineage transforms to mesoderm due to ectopic SKN-1. I propose that this is not the only reason and that failure to express the mesoderm antagonist *neg-1* is also responsible. Similarly, the P2 to EMS transformation in *pie-1(-)* embryos has been attributed to ectopic SKN-1 activity. However, I propose that this transformation would not occur without POS-1 faithfully repressing *neg-1* expression. Indeed, *pie-1(-); pos-1(-)* embryos die with P2 transforming to an EMS that does not differentiate (Craig Mello personal communication). *neg-1* activity also explains the intriguing observation that reducing *pos-1* levels suppresses the excess muscle in *mex-5(-)* embryos, which become viable in a *pos-1* heterozygous

background (Tenlen et al., 2006). POS-1 reduction may increase *neg-1* expression in the zygote, which would then repress mesoderm differentiation in the anterior lineage.

Form this interplay of determinants and antagonists it seems that the switch from germline to a somatic fate involves three requirements:

1) Cessation of germline identity,

2) Positive adoption of a somatic identity,

3) Repression of antagonists to adopted identity and/or repression of alterative somatic identities.

The classical mutants *pie-1*, *mex-3* and *mex-5* have been interpreted in light of the first two requirements. *pos-1* mutants illuminate the consequences of failing the third requirement.

POS-1 and regulation of posterior lineages

EMS is not the only blastomere affected in *pos-1(-)* embryos. Anterior pharyngeal development from ABp and the germline blastomere P4 are also affected. It is reasonable to think that POS-1 regulates the expression of many maternal transcripts that affect these three blastomeres. Exactly how many would require transcriptome wide analyses to identify transcripts that (1) have a different poly(A) tail length, (2) change in abundance in the absence of POS-1, or (3) are not translated in the absence of POS-1. Based on the presence of POS-1 binding elements, 2902 transcripts are predicted to be POS-1 targets, 227 of which are expressed in the early embryo (Farley et al., 2008).

In order to gain further insight into the role of *pos-1* in regulating early embryogenesis, we are conducting a transcriptome-wide ePAT analysis to identify transcripts affected by loss of *pos-1*. Several questions could be answered using this data

set. For example, those transcripts that behave similar to *neg-1* (longer poly(A) tails) will be knocked down by RNAi to examine whether they are required for embryogenesis or whether they can correct any of the defects observed in *pos-1(-)* embryos.

It is not known why P2 to ABp signaling fails in *pos-1(-)* embryos. The regulation of *neg-1* encourages the hypothesis that *mex-5* represses an AB fate antagonist, which is promoted by *pos-1* to protect posterior lineages. In *pos-1(-)* embryos this antagonist is ectopically expressed in the zygote, AB and/or its daughters where it undermines the effect of P2 signaling. In fact, this proposed antagonist may be the cause behind failed MS to AB8 signaling. In other words, one third of *pos-1(-)* embryos may fail to develop anterior pharynx not because P2 and MS fails to signal, but because the AB lineage fails to respond.

glp-1/Notch, *cye-1*/Cyclin E and maternal physiology affecting the behavior of offspring.

The manner in which *glp-1* suppresses *pos-1* is quite interesting. By perturbing *glp-1* function in the maternal germline, these germ cells later proceed to become embryos that can develop gut even though *pos-1* is absent. A simple explanation could be that maternal *neg-1* expression is *glp-1* dependent. At restrictive temperature, *neg-1* expression is reduced in germ cells. When these germ cells become gametes and then zygotes, the reduced *neg-1* level is not hazardous to EMS specification. Thus embryos from *glp-1ts* moms could specify endo-mesoderm identity without the protection of POS-1. This may also explain why reducing the activity of transcription factors *efl-1ts* and *dpl-1ts* in the maternal germline restores endo-mesoderm development in *pos-1(-)* embryos (Tenlen et al., 2006).

cye-1/CyclinE functions with *glp-1* in the germline to promote germ cell proliferation and maintain the mitotic zone (Biedermann et al., 2009; Fox et al., 2011; Jeong et al., 2011). *cye-1* suppression of *pos-1* may be a due to reduced *neg-1* expression in the germline, which then reduces the extent of endo-mesoderm repression in the embryo. However, it remains to be discerned whether *cye-1* suppression stems from its reduction in the embryo or the germline. Anecdotally, onset of *cye-1* supression of *pos-1* is delayed compared to *gld-3* suppression (24h compared to 9h on RNAi food). This may be explained by *cye-1* suppression taking place in the germline the same way *glp-1* does, whereas *gld-3* suppression takes place in the embryo. I am currently testing whether germline NEG-1::GFP levels are affected in the *glp-1ts* and *cye-1(RNAi)* backgrounds.

The behavioral defect observed in *neg-1* worms brings to our attention the behavioral consequences of maternal mal-nutrition or disease. mRNA transcripts loaded into the embryo are affected by the state of the maternal germline. While perturbed levels of key factors such as *neg-1* mRNA may not abort embryogenesis, subtle neuronal deficiencies may occur. An exciting possibility is that starved worms produce embryos that develop their nervous system differently than progeny of well-fed worms. This difference may be the simple consequence of leaky mesoderm differentiation that impairs neuronal development. Embryonic endoderm may also be influenced by the state of the maternal germline. The same reduced *neg-1* that would impair neuronal development may allow for more robust endo-mesoderm differentiation. Many possibilities exist and the biological dots await our connections.

What is *neg-1*?

How NEG-1 represses endo-mesoderm genes is still unclear. However, its nuclear localization along chromatin suggests a role in gene regulation. Due to its lack of homology, predictable tertiary structure and disorganized nature, traditional methods of protein sequence analysis have not been useful in recognizing this gene or its protein product. In fact the norm seems to be that critical components of the transcriptional apparatus such as transcriptional activation factors, co-activators, chromatin remodelers and core-promoter factors lack obvious domains, motifs or structures (Levine et al., 2014). I have presented the skewed sequence of NEG-1 as a distinctive characteristic that is common with a number of transcription regulation factors. There exists a literature concerned with proteins with such "compositionally biased regions" (CBRs) (Wootton, 1994; Wootton and Federhen, 1996). Interestingly, more than one third of proteins include at least one CBR, which are usually neglected in favor of structured identifiable regions. NEG-1, however, does not offer us this distraction. The skewed presence of aromatic amino acids W, F and Y in the N-terminal half and of basic amino acids R and K in the C-terminal half may be the rudiments of a protein structure. DNA binding domains are rich in basic amino acids and proteins frequently utilize hydrophobic domains to interact with other proteins. Importantly, several forms of cancer involve the translocation of a low complexity domain to a variety of different DNA-binding domains.

One possibility is that *neg-1* is a product of horizontal gene transfer from a different organism. Using DELTA-BLAST, an algorithm more sensitive to distant homology, NEG-1 is homologous with ribosomal subunit 15 in myriad plants and to lesser degree parasitic protozoa. However, performing a similar search using the *C. brenneri* homolog

of NEG-1 suggests a region homologous to the large subunit of RNA polymerase II. An intriguing possibility could be that a foreign RPL15 or RNA Pol II gene (or a fusion of both) was incorporated into an ancestral nematode genome and transmitted to subsequent generations. Interestingly, RPL15 is overexpressed in forms of gastric and esophageal cancer(Hsu et al., 2011; Wang et al., 2006a; Zhang et al., 2004). Whether *neg-1* originates from a foreign RPL15 detrimental to gut development and homeostasis or whether its residual RNA polymerase resemblance allows it to interfere with *bona fide* transcription is highly speculative. However, excavating the embryo to discover the logic of life has never ceased to surprise.

Carnarvon: *Can you see anything?*

Carter: *Yes, it is wonderful!*[1]

---

[1] Lord Carnarvon and Howard Carter at the tomb of King Tutankhamen on November 26th 1922, Valley of the Kings, Egypt.

Appendix

Table 3.1. continued

Species	Name	length	N' FWY	C' FWY	FWY ratio	N' RK	C' RK	RK ratio	Functional Annotation
S. cerevisiae	YEL009C	281	13	3	4.3	8	28	0.3	General control protein GCN4
S. cerevisiae	YER190C-A	191	5	1	5.0	1	5	0.2	Uncharacterized protein
S. cerevisiae	YPL283W-A	191	5	1	5.0	1	5	0.2	Uncharacterized protein
S. cerevisiae	YMR135W-A	177	15	1	15.0	6	25	0.2	Uncharacterized protein
S. cerevisiae	YGR296C-A	191	5	1	5.0	1	5	0.2	Uncharacterized protein
S. cerevisiae	YJR023C	133	13	4	3.3	0	11	0.1	Uncharacterized protein
S. cerevisiae	YML133W-A	191	5	1	5.0	1	5	0.2	Uncharacterized protein
S. cerevisiae	YMR123W	122	11	2	5.5	1	13	0.1	V-type ATPase assembly factor PKR1
S. cerevisiae	YNL339W-A	191	5	1	5.0	1	5	0.2	Uncharacterized protein
S. cerevisiae	YIL100W	117	10	3	3.3	3	12	0.3	Uncharacterized protein
S. cerevisiae	YKR047W	101	11	0	11.0	3	14	0.2	Uncharacterized protein
S. cerevisiae	YLR204W	111	4	0	4.0	6	21	0.3	Mitochondrial mRNA-processing protein COX24
D. melano.	**FBpp0304505**	**710**	**20**	**4**	**5.0**	**7**	**38**	**0.2**	**daughterless**
D. melano.	FBpp0302697	103	9	1	9.0	4	13	0.3	Uncharacterized protein
D. melano.	FBpp0297202	119	8	2	4.0	1	9	0.1	Dmel_CG15219
D. melano.	FBpp0071997	181	4	1	4.0	1	23	0.0	Uncharacterized protein
D. melano.	FBpp0290488	110	4	0	4.0	3	20	0.2	Uncharacterized protein
D. melano.	**FBpp0099694**	**381**	**10**	**3**	**3.3**	**9**	**34**	**0.3**	**Activating transcription factor-2**
D. melano.	FBpp0305238	153	6	0	6.0	1	26	0.0	Dmel_CG17362
D. melano.	FBpp0075507	161	8	0	8.0	1	26	0.0	Dmel_CG9040
D. melano.	FBpp0075435	113	5	1	5.0	2	7	0.3	Uncharacterized protein
D. melano.	FBpp0081955	514	11	2	5.5	8	36	0.2	Uncharacterized protein
D. melano.	FBpp0083100	218	11	0	11.0	5	16	0.3	Dmel_CG7715
D. melano.	FBpp0084113	653	26	6	4.3	31	110	0.3	Dmel_CG5808
D. melano.	FBpp0084347	158	20	6	3.3	2	9	0.2	Dmel_CG14550
D. melano.	**FBpp0288515**	**755**	**24**	**6**	**4.0**	**8**	**37**	**0.2**	**kayak**
D. melano.	FBpp0070527	272	6	1	6.0	10	35	0.3	Dmel_CG14422
D. melano.	FBpp0070947	318	10	2	5.0	4	39	0.1	Dmel_CG14440

Table 3.1. continued

Species	Name	length	N' FWY	C' FWY	FWY ratio	N' RK	C' RK	RK ratio	Functional Annotation
H. sapiens	ENSP00000302665	195	12	3	4.0	8	28	0.3	insulin-like growth factor 1 (somatomedin C)
H. sapiens	ENSP00000402118	190	12	1	12.0	7	22	0.3	proteasome (prosome, macropain) subunit, alpha type, 4
H. sapiens	ENSP00000419768	147	18	2	9.0	0	10	0.1	ATPase, class VI, type 11B
H. sapiens	ENSP00000438313	218	4	1	4.0	1	16	0.1	WNK lysine deficient protein kinase 1
H. sapiens	ENSP00000439957	320	18	2	9.0	9	30	0.3	splicing factor, arginine/serine-rich 8
H. sapiens	ENSP00000411516	325	11	2	5.5	2	14	0.1	sterol regulatory element binding transcription factor 1
H. sapiens	ENSP00000461971	133	7	2	3.5	2	12	0.2	nuclear factor (erythroid-derived 2)-like 1
H. sapiens	ENSP00000392000	166	5	0	5.0	3	22	0.1	GNAS complex locus
H. sapiens	ENSP00000399644	677	26	8	3.3	14	53	0.3	RAB11 family interacting protein 3 (class II)
H. sapiens	ENSP00000472176	182	12	3	4.0	6	19	0.3	mucolipin 1
H. sapiens	ENSP00000397588	134	5	1	5.0	3	12	0.3	protein interacting with PRKCA 1
H. sapiens	ENSP00000457401	166	7	1	7.0	0	4	0.3	TOX high mobility group box family member 3
H. sapiens	**ENSP00000470027**	**253**	**16**	**5**	**3.2**	**4**	**13**	**0.3**	**mediator complex subunit 25**
H. sapiens	**ENSP00000429628**	**173**	**8**	**2**	**4.0**	**4**	**18**	**0.2**	**zinc finger protein 419**
H. sapiens	ENSP00000395180	105	4	1	4.0	1	15	0.1	TCF3 (E2A) fusion partner (in childhood Leukemia)
H. sapiens	ENSP00000471612	134	7	1	7.0	3	11	0.3	transmembrane protein 59-like
H. sapiens	ENSP00000405308	106	7	1	7.0	1	4	0.3	claudin 15
H. sapiens	ENSP00000361148	147	12	1	12.0	3	14	0.2	vascular endothelial growth factor A
H. sapiens	ENSP00000388105	101	4	1	4.0	2	10	0.2	family with sequence similarity 176, member A
H. sapiens	ENSP00000378568	389	7	1	7.0	8	78	0.1	splicing factor, arginine/serine-rich 11
H. sapiens	ENSP00000358094	121	13	2	6.5	1	9	0.1	chromosome 1 open reading frame 54
H. sapiens	ENSP00000451040	170	8	2	4.0	5	21	0.2	placental growth factor
H. sapiens	**ENSP00000465552**	**237**	**10**	**2**	**5.0**	**1**	**25**	**0.0**	**FBJ murine osteosarcoma viral oncogene homolog B**
H. sapiens	ENSP00000465317	158	5	1	5.0	2	9	0.2	mitochondrial rRNA methyltransferase 1 homolog (S. cerevisiae)
H. sapiens	**ENSP00000426127**	**105**	**7**	**2**	**3.5**	**1**	**13**	**0.1**	**zinc finger protein 331**
H. sapiens	ENSP00000469221	162	13	3	4.3	4	18	0.2	DnaJ (Hsp40) homolog, subfamily B, member 1
H. sapiens	ENSP00000358242	805	29	8	3.6	39	130	0.3	splicing factor, arginine/serine-rich 18
H. sapiens	ENSP00000462200	200	16	4	4.0	2	11	0.2	solute carrier family 39 (metal ion transporter), member 11
H. sapiens	ENSP00000473322	101	6	1	6.0	1	7	0.1	lysyl oxidase-like 2
H. sapiens	**ENSP00000447828**	**107**	**6**	**1**	**6.0**	**1**	**4**	**0.3**	**PH domain and leucine zipper containing 2**
H. sapiens	**ENSP00000386288**	**117**	**7**	**0**	**7.0**	**2**	**10**	**0.2**	**signal transducer and activator of transcription 4**
H. sapiens	ENSP00000454110	121	7	0	7.0	2	8	0.3	dual oxidase maturation factor 2
H. sapiens	ENSP00000427040	128	6	0	6.0	2	11	0.2	DNA-damage-inducible transcript 4-like
H. sapiens	ENSP00000429229	197	15	2	7.5	5	17	0.3	ER lipid raft associated 2

Table 3.1. continued

Species	Name	length	N' FWY	C' FWY	FWY ratio	N' RK	C' RK	RK ratio	Functional Annotation
H. sapiens	ENSP00000435422	282	16	4	4.0	5	27	0.2	Wolf-Hirschhorn syndrome candidate 1-like 1
H. sapiens	ENSP00000417243	125	5	0	5.0	3	15	0.2	integrin, beta 1
H. sapiens	ENSP00000368116	145	9	1	9.0	4	22	0.2	coiled-coil domain containing 3
H. sapiens	**ENSP00000470111**	**148**	**8**	**2**	**4.0**	**2**	**9**	**0.2**	**zinc finger protein 547**
H. sapiens	ENSP00000385379	150	10	3	3.3	5	17	0.3	nuclear receptor subfamily 4, group A, member 2
H. sapiens	ENSP00000334538	624	19	1	19.0	28	126	0.2	splicing factor, arginine/serine-rich 12
H. sapiens	**ENSP00000416227**	**244**	**9**	**2**	**4.5**	**8**	**31**	**0.3**	**mediator complex subunit 19**
H. sapiens	**ENSP00000356998**	**282**	**11**	**3**	**3.7**	**5**	**25**	**0.2**	**upstream transcription factor 1**
H. sapiens	ENSP00000432600	127	7	1	7.0	0	7	0.1	von Willebrand factor A domain containing 5B1
H. sapiens	**ENSP00000455891**	**121**	**4**	**0**	**4.0**	**2**	**11**	**0.2**	**ring finger and SPRY domain containing 1**
H. sapiens	ENSP00000371820	112	5	0	5.0	3	13	0.2	C1q and tumor necrosis factor related protein 7
H. sapiens	ENSP00000295400	160	6	1	6.0	2	10	0.2	transforming growth factor, alpha
H. sapiens	ENSP00000469318	122	6	0	6.0	2	7	0.3	kallikrein-related peptidase 2
H. sapiens	ENSP00000476179	203	10	2	5.0	7	22	0.3	PX domain containing serine/threonine kinase
H. sapiens	ENSP00000447188	169	8	0	8.0	1	19	0.1	DNA-damage-inducible transcript 3
H. sapiens	ENSP00000472057	148	4	1	4.0	6	21	0.3	regulator of chromosome condensation 1
H. sapiens	ENSP00000324775	156	7	2	3.5	4	25	0.2	transmembrane inner ear
H. sapiens	ENSP00000329557	159	7	1	7.0	4	13	0.3	transmembrane protein 89
H. sapiens	ENSP00000410814	219	19	4	4.8	8	31	0.3	nuclear pore complex interacting protein
H. sapiens	ENSP00000436315	138	6	1	6.0	3	14	0.2	RNA-binding region (RNP1, RRM) containing 3
H. sapiens	**ENSP00000427067**	**142**	**7**	**2**	**3.5**	**2**	**10**	**0.2**	**zinc finger, DHHC-type containing 11**
H. sapiens	ENSP00000455285	125	11	3	3.7	2	8	0.3	methyltransferase like 9
H. sapiens	**ENSP00000466265**	**102**	**7**	**2**	**3.5**	**2**	**7**	**0.3**	**zinc finger protein 607**
H. sapiens	ENSP00000359566	103	4	1	4.0	2	9	0.2	hypothetical protein LOC286411
H. sapiens	ENSP00000448481	116	5	1	5.0	0	13	0.1	proline rich 13
H. sapiens	ENSP00000392933	107	4	0	4.0	3	14	0.2	HLA-B associated transcript 4
H. sapiens	ENSP00000392227	188	12	3	4.0	5	17	0.3	Uncharacterized protein
H. sapiens	ENSP00000406761	107	4	0	4.0	3	14	0.2	G patch domain and ankyrin repeats 1
H. sapiens	ENSP00000414033	104	5	1	5.0	0	10	0.1	Uncharacterized protein
H. sapiens	**ENSP00000433093**	**109**	**7**	**2**	**3.5**	**3**	**13**	**0.2**	**zinc finger, BED-type containing 5**
H. sapiens	ENSP00000475001	122	6	0	6.0	3	12	0.3	TLX1 neighbor
H. sapiens	ENSP00000411450	108	6	1	6.0	2	9	0.2	eukaryotic translation initiation factor 6
H. sapiens	ENSP00000454033	114	4	0	4.0	3	12	0.3	Uncharacterized protein
H. sapiens	ENSP00000463993	121	13	2	6.5	1	9	0.1	Uncharacterized protein
H. sapiens	ENSP00000473947	105	4	1	4.0	1	15	0.1	TCF3 (E2A) fusion partner (in childhood Leukemia)
H. sapiens	ENSP00000475235	160	6	1	6.0	2	10	0.2	transforming growth factor, alpha

BIBLIOGRAPHY

Ables, E.T., and Drummond-Barbosa, D. (2013). Cyclin E controls Drosophila female germline stem cell maintenance independently of its role in proliferation by modulating responsiveness to niche signals. Development *140*, 530-540.

Akiyama-Oda, Y., Hosoya, T., and Hotta, Y. (1999). Asymmetric cell division of thoracic neuroblast 6-4 to bifurcate glial and neuronal lineage in Drosophila. Development *126*, 1967-1974.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic acids research *25*, 3389-3402.

Austin, J., and Kimble, J. (1987). glp-1 is required in the germ line for regulation of the decision between mitosis and meiosis in C. elegans. Cell *51*, 589-599.

Batchelder, C., Dunn, M.A., Choy, B., Suh, Y., Cassie, C., Shim, E.Y., Shin, T.H., Mello, C., Seydoux, G., and Blackwell, T.K. (1999). Transcriptional repression by the Caenorhabditis elegans germ-line protein PIE-1. Genes & development *13*, 202-212.

Bei, Y., Hogan, J., Berkowitz, L.A., Soto, M., Rocheleau, C.E., Pang, K.M., Collins, J., and Mello, C.C. (2002). SRC-1 and Wnt signaling act together to specify endoderm and to control cleavage orientation in early C. elegans embryos. Developmental cell *3*, 113-125.

Berger, C., Pallavi, S.K., Prasad, M., Shashidhara, L.S., and Technau, G.M. (2005a). A critical role for cyclin E in cell fate determination in the central nervous system of Drosophila melanogaster. Nature cell biology *7*, 56-62.

Berger, C., Pallavi, S.K., Prasad, M., Shashidhara, L.S., and Technau, G.M. (2005b). Cyclin E acts under the control of Hox-genes as a cell fate determinant in the developing central nervous system. Cell cycle *4*, 422-425.

Biedermann, B., Wright, J., Senften, M., Kalchhauser, I., Sarathy, G., Lee, M.H., and Ciosk, R. (2009). Translational repression of cyclin E prevents precocious mitosis and embryonic gene activation during C. elegans meiosis. Developmental cell *17*, 355-364.

Blackwell, T.K., Bowerman, B., Priess, J.R., and Weintraub, H. (1994). Formation of a monomeric DNA binding domain by Skn-1 bZIP and homeodomain elements. Science *266*, 621-628.

Boratyn, G.M., Schaffer, A.A., Agarwala, R., Altschul, S.F., Lipman, D.J., and Madden, T.L. (2012). Domain enhanced lookup time accelerated BLAST. Biology direct *7*, 12.

Bowerman, B., Draper, B.W., Mello, C.C., and Priess, J.R. (1993). The maternal gene skn-1 encodes a protein that is distributed unequally in early C. elegans embryos. Cell *74*, 443-452.

Bowerman, B., Eaton, B.A., and Priess, J.R. (1992). skn-1, a maternally expressed gene required to specify the fate of ventral blastomeres in the early C. elegans embryo. Cell *68*, 1061-1075.

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

Brenner, S. (2003). Nobel lecture. Nature's gift to science. Bioscience reports *23*, 225-237.

Brooks, S.A., and Blackshear, P.J. (2013). Tristetraprolin (TTP): interactions with mRNA and proteins, and current thoughts on mechanisms of action. Biochimica et biophysica acta *1829*, 666-679.

Chitwood, B.G., and Chitwood, M.B.H. (1937). An Introduction to Nematology (Baltimore, Md.,: Monumental printing co.).

Christensen, S., Kodoyianni, V., Bosenberg, M., Friedman, L., and Kimble, J. (1996). lag-1, a gene required for lin-12 and glp-1 signaling in Caenorhabditis elegans, is homologous to human CBF1 and Drosophila Su(H). Development *122*, 1373-1383.

Crittenden, S.L., Eckmann, C.R., Wang, L., Bernstein, D.S., Wickens, M., and Kimble, J. (2003). Regulation of the mitosis/meiosis decision in the Caenorhabditis elegans germline. Philosophical transactions of the Royal Society of London Series B, Biological sciences *358*, 1359-1362.

Cuenca, A.A., Schetter, A., Aceto, D., Kemphues, K., and Seydoux, G. (2003). Polarization of the C. elegans zygote proceeds via distinct establishment and maintenance phases. Development *130*, 1255-1265.

D'Agostino, I., Merritt, C., Chen, P.L., Seydoux, G., and Subramaniam, K. (2006). Translational repression restricts expression of the C. elegans Nanos homolog NOS-2 to the embryonic germline. Developmental biology *292*, 244-252.

DeRenzo, C., Reese, K.J., and Seydoux, G. (2003). Exclusion of germ plasm proteins from somatic lineages by cullin-dependent degradation. Nature *424*, 685-689.

Draper, B.W., Mello, C.C., Bowerman, B., Hardin, J., and Priess, J.R. (1996). MEX-3 is a KH domain protein that regulates blastomere identity in early C. elegans embryos. Cell *87*, 205-216.

Du, Z., Santella, A., He, F., Tiongson, M., and Bao, Z. (2014). De novo inference of systems-level mechanistic models of development from live-imaging-based phenotype analysis. Cell *156*, 359-372.

DuBois, R.N., McLane, M.W., Ryder, K., Lau, L.F., and Nathans, D. (1990). A growth factor-inducible nuclear protein with a novel cysteine/histidine repetitive sequence. The Journal of biological chemistry *265*, 19185-19191.

Eckmann, C.R., Crittenden, S.L., Suh, N., and Kimble, J. (2004). GLD-3 and control of the mitosis/meiosis decision in the germline of Caenorhabditis elegans. Genetics *168*, 147-160.

Eckmann, C.R., Kraemer, B., Wickens, M., and Kimble, J. (2002). GLD-3, a bicaudal-C homolog that inhibits FBF to control germline sex determination in C. elegans. Developmental cell *3*, 697-710.

Edgar, L.G., Carr, S., Wang, H., and Wood, W.B. (2001). Zygotic expression of the caudal homolog pal-1 is required for posterior patterning in Caenorhabditis elegans embryogenesis. Developmental biology *229*, 71-88.

Evans, T.C., Crittenden, S.L., Kodoyianni, V., and Kimble, J. (1994). Translational control of maternal glp-1 mRNA establishes an asymmetry in the C. elegans embryo. Cell *77*, 183-194.

Farley, B.M., Pagano, J.M., and Ryder, S.P. (2008). RNA target specificity of the embryonic cell fate determinant POS-1. Rna *14*, 2685-2697.

Farley, B.M., and Ryder, S.P. (2012). POS-1 and GLD-1 repress glp-1 translation through a conserved binding-site cluster. Molecular biology of the cell *23*, 4473-4483.

Fox, P.M., Vought, V.E., Hanazawa, M., Lee, M.H., Maine, E.M., and Schedl, T. (2011). Cyclin E and CDK-2 regulate proliferative cell fate and cell cycle progression in the C. elegans germline. Development *138*, 2223-2234.

Fox, R.M., Von Stetina, S.E., Barlow, S.J., Shaffer, C., Olszewski, K.L., Moore, J.H., Dupuy, D., Vidal, M., and Miller, D.M., 3rd (2005). A gene expression fingerprint of C. elegans embryonic motor neurons. BMC genomics *6*, 42.

Francis, R., Barton, M.K., Kimble, J., and Schedl, T. (1995a). gld-1, a tumor suppressor gene required for oocyte development in Caenorhabditis elegans. Genetics *139*, 579-606.

Francis, R., Maine, E., and Schedl, T. (1995b). Analysis of the multiple roles of gld-1 in germline development: interactions with the sex determination cascade and the glp-1 signaling pathway. Genetics *139*, 607-630.

Frokjaer-Jensen, C., Davis, M.W., Hopkins, C.E., Newman, B.J., Thummel, J.M., Olesen, S.P., Grunnet, M., and Jorgensen, E.M. (2008). Single-copy insertion of transgenes in Caenorhabditis elegans. Nature genetics *40*, 1375-1383.

Goldstein, B. (1992). Induction of gut in Caenorhabditis elegans embryos. Nature *357*, 255-257.

Gomperts, M., Pascall, J.C., and Brown, K.D. (1990). The nucleotide sequence of a cDNA encoding an EGF-inducible gene indicates the existence of a new family of mitogen-induced genes. Oncogene *5*, 1081-1083.

Gonczy, P., and Rose, L.S. (2005). Asymmetric cell division and axis formation in the embryo. WormBook : the online review of C elegans biology, 1-20.

Hansen, D., Wilson-Berry, L., Dang, T., and Schedl, T. (2004). Control of the proliferation versus meiotic development decision in the C. elegans germline through regulation of GLD-1 protein accumulation. Development *131*, 93-104.

Hashimshony, T., Wagner, F., Sher, N., and Yanai, I. (2012). CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell reports *2*, 666-673.

Holway, A.H., Kim, S.H., La Volpe, A., and Michael, W.M. (2006). Checkpoint silencing during the DNA damage response in Caenorhabditis elegans embryos. The Journal of cell biology *172*, 999-1008.

Horner, M.A., Quintin, S., Domeier, M.E., Kimble, J., Labouesse, M., and Mango, S.E. (1998). pha-4, an HNF-3 homolog, specifies pharyngeal organ identity in Caenorhabditis elegans. Genes & development *12*, 1947-1952.

Hsu, Y.A., Lin, H.J., Sheu, J.J., Shieh, F.K., Chen, S.Y., Lai, C.H., Tsai, F.J., Wan, L., and Chen, B.H. (2011). A novel interaction between interferon-inducible protein p56 and ribosomal protein L15 in gastric cancer cells. DNA and cell biology *30*, 671-679.

Hunter, C.P., and Kenyon, C. (1996). Spatial and temporal controls target pal-1 blastomere-specification activity to a single blastomere lineage in C. elegans embryos. Cell *87*, 217-226.

Hutter, H., and Schnabel, R. (1994). glp-1 and inductions establishing embryonic axes in C. elegans. Development *120*, 2051-2064.

Ishidate, T., Elewa, A., Kim, S., Mello, C.C., and Shirayama, M. (2014). Divide and differentiate: CDK/Cyclins and the art of development. Cell cycle *13*.

Jadhav, S., Rana, M., and Subramaniam, K. (2008). Multiple maternal proteins coordinate to restrict the translation of C. elegans nanos-2 to primordial germ cells. Development *135*, 1803-1812.

Janicke, A., Vancuylenberg, J., Boag, P.R., Traven, A., and Beilharz, T.H. (2012). ePAT: a simple method to tag adenylated RNA to measure poly(A)-tail length and other 3' RACE applications. Rna *18*, 1289-1295.

Jeong, J., Verheyden, J.M., and Kimble, J. (2011). Cyclin E and Cdk2 control GLD-1, the mitosis/meiosis decision, and germline stem cells in Caenorhabditis elegans. PLoS genetics *7*, e1001348.

Jones, A.R., Francis, R., and Schedl, T. (1996). GLD-1, a cytoplasmic protein essential for oocyte differentiation, shows stage- and sex-specific expression during Caenorhabditis elegans germline development. Developmental biology *180*, 165-183.

Jones, A.R., and Schedl, T. (1995). Mutations in gld-1, a female germ cell-specific tumor suppressor gene in Caenorhabditis elegans, affect a conserved domain also found in Src-associated protein Sam68. Genes & development *9*, 1491-1504.

Kamath, R.S., and Ahringer, J. (2003). Genome-wide RNAi screening in Caenorhabditis elegans. Methods *30*, 313-321.

Kemphues, K.J., Priess, J.R., Morton, D.G., and Cheng, N.S. (1988). Identification of genes required for cytoplasmic localization in early C. elegans embryos. Cell *52*, 311-320.

Kim, K.W., Wilson, T.L., and Kimble, J. (2010). GLD-2/RNP-8 cytoplasmic poly(A) polymerase is a broad-spectrum regulator of the oogenesis program. Proceedings of the National Academy of Sciences of the United States of America *107*, 17445-17450.

Kimble, J., and Crittenden, S.L. (2005). Germline proliferation and its control. WormBook : the online review of C elegans biology, 1-14.

Kimble, J., and Simpson, P. (1997). The LIN-12/Notch signaling pathway and its regulation. Annual review of cell and developmental biology *13*, 333-361.

Lai, W.S., Stumpo, D.J., and Blackshear, P.J. (1990). Rapid insulin-stimulated accumulation of an mRNA encoding a proline-rich protein. The Journal of biological chemistry *265*, 16556-16563.

Laufer, J.S., Bazzicalupo, P., and Wood, W.B. (1980). Segregation of developmental potential in early embryos of Caenorhabditis elegans. Cell *19*, 569-577.

Levine, M., Cattoglio, C., and Tjian, R. (2014). Looping Back to Leap Forward: Transcription Enters a New Era. Cell *157*, 13-25.

Lin, R., Thompson, S., and Priess, J.R. (1995). pop-1 encodes an HMG box protein required for the specification of a mesoderm precursor in early C. elegans embryos. Cell *83*, 599-609.

Maduro, M.F., Broitman-Maduro, G., Mengarelli, I., and Rothman, J.H. (2007). Maternal deployment of the embryonic SKN-1-->MED-1,2 cell specification pathway in C. elegans. Developmental biology *301*, 590-601.

Maduro, M.F., Meneghini, M.D., Bowerman, B., Broitman-Maduro, G., and Rothman, J.H. (2001). Restriction of mesendoderm to a single blastomere by the combined action of SKN-1 and a GSK-3beta homolog is mediated by MED-1 and -2 in C. elegans. Molecular cell *7*, 475-485.

Marin, V.A., and Evans, T.C. (2003). Translational repression of a C. elegans Notch mRNA by the STAR/KH domain protein GLD-1. Development *130*, 2623-2632.

Mello, C.C., Draper, B.W., Krause, M., Weintraub, H., and Priess, J.R. (1992). The pie-1 and mex-1 genes and maternal control of blastomere identity in early C. elegans embryos. Cell *70*, 163-176.

Mello, C.C., Draper, B.W., and Priess, J.R. (1994). The maternal genes apx-1 and glp-1 and establishment of dorsal-ventral polarity in the early C. elegans embryo. Cell *77*, 95-106.

Mello, C.C., Schubert, C., Draper, B., Zhang, W., Lobel, R., and Priess, J.R. (1996). The PIE-1 protein and germline specification in C. elegans embryos. Nature *382*, 710-712.

Merlet, J., Burger, J., Tavernier, N., Richaudeau, B., Gomes, J.E., and Pintard, L. (2010). The CRL2LRR-1 ubiquitin ligase regulates cell cycle progression during C. elegans development. Development *137*, 3857-3866.

Merritt, C., Rasoloson, D., Ko, D., and Seydoux, G. (2008). 3' UTRs are the primary regulators of gene expression in the C. elegans germline. Current biology : CB *18*, 1476-1482.

Mickey, K.M., Mello, C.C., Montgomery, M.K., Fire, A., and Priess, J.R. (1996). An inductive interaction in 4-cell stage C. elegans embryos involves APX-1 expression in the signalling cell. Development *122*, 1791-1798.

Morgan, D.O. (1997). Cyclin-dependent kinases: engines, clocks, and microprocessors. Annual review of cell and developmental biology *13*, 261-291.

Moroy, T., and Geisen, C. (2004). Cyclin E. The international journal of biochemistry & cell biology *36*, 1424-1439.

Morton, D.G., Roos, J.M., and Kemphues, K.J. (1992). par-4, a gene required for cytoplasmic localization and determination of specific cell types in Caenorhabditis elegans embryogenesis. Genetics *130*, 771-790.

Nakamura, K., Kim, S., Ishidate, T., Bei, Y., Pang, K., Shirayama, M., Trzepacz, C., Brownell, D.R., and Mello, C.C. (2005). Wnt signaling drives WRM-1/beta-catenin asymmetries in early C. elegans embryos. Genes & development *19*, 1749-1754.

Nakano, S., Stillman, B., and Horvitz, H.R. (2011). Replication-coupled chromatin assembly generates a neuronal bilateral asymmetry in C. elegans. Cell *147*, 1525-1536.

Nusser-Stein, S., Beyer, A., Rimann, I., Adamczyk, M., Piterman, N., Hajnal, A., and Fisher, J. (2012). Cell-cycle regulation of NOTCH signaling during C. elegans vulval development. Molecular systems biology *8*, 618.

Ogura, K., Kishimoto, N., Mitani, S., Gengyo-Ando, K., and Kohara, Y. (2003). Translational control of maternal glp-1 mRNA by POS-1 and its interacting protein SPN-4 in Caenorhabditis elegans. Development *130*, 2495-2503.

Oldenbroek, M., Robertson, S.M., Guven-Ozkan, T., Gore, S., Nishi, Y., and Lin, R. (2012). Multiple RNA-binding proteins function combinatorially to control the soma-restricted expression pattern of the E3 ligase subunit ZIF-1. Developmental biology *363*, 388-398.

Oldenbroek, M., Robertson, S.M., Guven-Ozkan, T., Spike, C., Greenstein, D., and Lin, R. (2013). Regulation of maternal Wnt mRNA translation in C. elegans embryos. Development *140*, 4614-4623.

Pagano, J.M., Farley, B.M., Essien, K.I., and Ryder, S.P. (2009). RNA recognition by the embryonic cell fate determinant and germline totipotency factor MEX-3. Proceedings of the National Academy of Sciences of the United States of America *106*, 20252-20257.

Pagano, J.M., Farley, B.M., McCoig, L.M., and Ryder, S.P. (2007). Molecular basis of RNA recognition by the embryonic polarity determinant MEX-5. The Journal of biological chemistry *282*, 8883-8894.

Page, B.D., Guedes, S., Waring, D., and Priess, J.R. (2001). The C. elegans E2F- and DP-related proteins are required for embryonic asymmetry and negatively regulate Ras/MAPK signaling. Molecular cell *7*, 451-460.

Priess, J.R. (2005). Notch signaling in the C. elegans embryo. WormBook : the online review of C elegans biology, 1-16.

Priess, J.R., Schnabel, H., and Schnabel, R. (1987). The glp-1 locus and cellular interactions in early C. elegans embryos. Cell *51*, 601-611.

Priess, J.R., and Thomson, J.N. (1987). Cellular interactions in early C. elegans embryos. Cell *48*, 241-250.

Reece-Hoyes, J.S., Diallo, A., Lajoie, B., Kent, A., Shrestha, S., Kadreppa, S., Pesyna, C., Dekker, J., Myers, C.L., and Walhout, A.J. (2011). Enhanced yeast one-hybrid assays for high-throughput gene-centered regulatory network mapping. Nature methods *8*, 1059-1064.

Rocheleau, C.E., Downs, W.D., Lin, R., Wittmann, C., Bei, Y., Cha, Y.H., Ali, M., Priess, J.R., and Mello, C.C. (1997). Wnt signaling and an APC-related gene specify endoderm in early C. elegans embryos. Cell *90*, 707-716.

Rocheleau, C.E., Yasuda, J., Shin, T.H., Lin, R., Sawa, H., Okano, H., Priess, J.R., Davis, R.J., and Mello, C.C. (1999). WRM-1 activates the LIT-1 protein kinase to transduce anterior/posterior polarity signals in C. elegans. Cell *97*, 717-726.

Ryder, S.P., Frater, L.A., Abramovitz, D.L., Goodwin, E.B., and Williamson, J.R. (2004). RNA target specificity of the STAR/GSG domain post-transcriptional regulatory protein GLD-1. Nature structural & molecular biology *11*, 20-28.

Schierenberg, E. (2006). Embryological variation during nematode development. WormBook : the online review of C elegans biology, 1-13.

Schnabel, R., Weigner, C., Hutter, H., Feichtinger, R., and Schnabel, H. (1996). mex-1 and the general partitioning of cell fate in the early C. elegans embryo. Mechanisms of development *54*, 133-147.

Schubert, C.M., Lin, R., de Vries, C.J., Plasterk, R.H., and Priess, J.R. (2000). MEX-5 and MEX-6 function to establish soma/germline asymmetry in early C. elegans embryos. Molecular cell *5*, 671-682.

Seydoux, G., and Dunn, M.A. (1997). Transcriptionally repressed germ cells lack a subpopulation of phosphorylated RNA polymerase II in early embryos of Caenorhabditis elegans and Drosophila melanogaster. Development *124*, 2191-2201.

Seydoux, G., Mello, C.C., Pettitt, J., Wood, W.B., Priess, J.R., and Fire, A. (1996). Repression of gene expression in the embryonic germ lineage of C. elegans. Nature *382*, 713-716.

Shelton, C.A., and Bowerman, B. (1996). Time-dependent responses to glp-1-mediated inductions in early C. elegans embryos. Development *122*, 2043-2050.

Shibata, Y., Takeshita, H., Sasakawa, N., and Sawa, H. (2010). Double bromodomain protein BET-1 and MYST HATs establish and maintain stable cell fates in C. elegans. Development *137*, 1045-1053.

Strome, S. (2005). Specification of the germ line. WormBook : the online review of C elegans biology, 1-10.

Subramaniam, K., and Seydoux, G. (1999). nos-1 and nos-2, two genes related to Drosophila nanos, regulate primordial germ cell development and survival in Caenorhabditis elegans. Development *126*, 4861-4871.

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

Tabara, H., Hill, R.J., Mello, C.C., Priess, J.R., and Kohara, Y. (1999). pos-1 encodes a cytoplasmic zinc-finger protein essential for germline specification in C. elegans. Development *126*, 1-11.

Tamburino, A.M., Ryder, S.P., and Walhout, A.J. (2013). A compendium of Caenorhabditis elegans RNA binding proteins predicts extensive regulation at multiple levels. G3 *3*, 297-304.

Tenlen, J.R., Schisa, J.A., Diede, S.J., and Page, B.D. (2006). Reduced dosage of pos-1 suppresses Mex mutants and reveals complex interactions among CCCH zinc-finger proteins during Caenorhabditis elegans embryogenesis. Genetics *174*, 1933-1945.

Thorpe, C.J., Schlesinger, A., Carter, J.C., and Bowerman, B. (1997). Wnt signaling polarizes an early C. elegans blastomere to distinguish endoderm from mesoderm. Cell *90*, 695-705.

Walhout, A.J., and Vidal, M. (2001). High-throughput yeast two-hybrid assays for large-scale protein interaction mapping. Methods *24*, 297-306.

Wang, H., Zhao, L.N., Li, K.Z., Ling, R., Li, X.J., and Wang, L. (2006a). Overexpression of ribosomal protein L15 is associated with cell proliferation in gastric cancer. BMC cancer *6*, 91.

Wang, L., Eckmann, C.R., Kadyk, L.C., Wickens, M., and Kimble, J. (2002). A regulatory cytoplasmic poly(A) polymerase in Caenorhabditis elegans. Nature *419*, 312-316.

Wang, P., Zhao, J., and Corsi, A.K. (2006b). Identification of novel target genes of CeTwist and CeE/DA. Developmental biology *293*, 486-498.

Waring, D.A., and Kenyon, C. (1991). Regulation of cellular responsiveness to inductive signals in the developing C. elegans nervous system. Nature *350*, 712-715.

Wootton, J.C. (1994). Non-globular domains in protein sequences: automated segmentation using complexity measures. Computers & chemistry *18*, 269-285.

Wootton, J.C., and Federhen, S. (1996). Analysis of compositionally biased regions in sequence databases. Methods in enzymology *266*, 554-571.

Wright, J.E., Gaidatzis, D., Senften, M., Farley, B.M., Westhof, E., Ryder, S.P., and Ciosk, R. (2011). A quantitative RNA code for mRNA target selection by the germline fate determinant GLD-1. The EMBO journal *30*, 533-545.

Zhang, F., Barboric, M., Blackwell, T.K., and Peterlin, B.M. (2003). A model of repression: CTD analogs and PIE-1 inhibit transcriptional elongation by P-TEFb. Genes & development *17*, 748-758.

Zhang, J., Gao, F.L., Zhi, H.Y., Luo, A.P., Ding, F., Wu, M., and Liu, Z.H. (2004). Expression patterns of esophageal cancer deregulated genes in C57BL/6J mouse embryogenesis. World journal of gastroenterology : WJG *10*, 1088-1092.

Zhong, W., and Sternberg, P.W. (2006). Genome-wide prediction of C. elegans genetic interactions. Science *311*, 1481-1484.