Negative regulators of a growth factor-mediated signaling pathway in the nematode *Caenorhabditis elegans*

Thesis by

Junho Lee

In partial fulfillment of the requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1994

(Submitted November 1, 1993)

Acknowledgments

I thank my advisor, Dr. Paul Sternberg. he has provided advice, discussions, and a comfortable and yet passionate environment for my research. I thank my thesis committee members, Drs. David Anderson, Bill Dunphy, Elliot Meyerowitz, and Alexander Varshavsky, for their invaluable discussions and encouragement. I also thank Gregg Jongeward. Without his help, I could not have started my thesis projects on the negative regulators of vulval differentiation. I thank all the members of the Sternberg lab, who have been so kind to take time to explain in plain English about what I did not understand, and have been all part of pleasant environments of the lab. I also thank all the members of the Korean community at Caltech.

I specially thank my wife, Kyungran Choi, for her support and love. She has been taking care of Chan, our son, all by herself in Korea while I was finishing up my thesis projects. I also thank all the members of my family for their support. Most of all, I thank my mother.

Abstract

Vulval differentiation in *C. elegans* is mediated by an Epidermal growth factor (EGF) - EGF receptor (EGFR) signaling pathway. I have cloned *unc-101*, a negative regulator of vulval differentiation of the nematode *C. elegans*. *unc-101* encodes a homolog of AP47, the medium chain of the *trans*-Golgi clathrin-associated protein complex. This identity was confirmed by cloning and comparing sequence of a *C. elegans* homolog of AP50, the medium chain of the plasma membrane clathrin-associated protein complex. I provided the first genetic evidence that the *trans*-Golgi clathrin-coated vesicles are involved in regulation of an EGF signaling pathway. Most of the *unc-101* alleles are deletions or nonsense mutations, suggesting that these alleles severely reduce the *unc-101* activity. A hybrid gene that contains parts of *unc-101* and mouse AP47 rescued at least two phenotypes of *unc-101* mutations, the Unc and the suppression of vulvaless phenotype of *let-23(sy1)* mutation. Therefore, the functions of AP47 are conserved between nematodes and mammals.

unc-101 mutations can cause a greater than wild-type vulval differentiation in combination with certain mutations in *sli-1*, another negative regulator of the vulval induction pathway. A mutation in a new gene, *rok-1*, causes no defect by itself, but causes a greater than wild-type vulval differentiation in the presence of a *sli-1* mutation. The *unc-101; rok-1; sli-1* triple mutants display a greater extent of vulval differentiation than any double mutant combinations of *unc-101, rok-1* and *sli-1*. Therefore, *rok-1* locus defines another negative regulator of the vulval induction pathway.

I analyzed a second gene encoding an AP47 homolog in *C. elegans*. This gene, CEAP47, encodes a protein 72% identical to both *unc-101* and

mammalian AP47. A hybrid gene containing parts of *unc-101* and CEAP47 sequences can rescue phenotypes of *unc-101* mutants, indicating that UNC-101 and CEAP47 proteins can be redundant if expressed in the same set of cells.

Table of Contents

CHAPTER 1. INTRODUCTION

I. *C. elegans* as an experimental system for studying developmental biology

Developmental biology deals with one of the ultimate questions in the biological science: how a single-celled egg develops into a perfectly organized adult that has many types of differentiated cells, tissues, and organs? Many experimental organisms have been studied to elucidate the mystery of development: the nematode *C. elegans* is one such organism.
C. elegans is a free-living soil nematode found all around the world (Wood, 1988). Several aspects of the biology of this nematode make it an excellent experimental organism for research in developmental biology. Development of *C. elegans* is rather simple and invariant. That is, development of one individual organism is essentially identical to that of any other individual, making it easy to observe developmental phenomenon reproducibly in different individuals. Cell division timings and patterns are almost always invariant, so that every cell division event and its resulting daughters can be easily observed. Indeed, a complete cell lineage map starting with a single-celled fertilized egg has been established. In addition, *C. elegans* is transparent, permitting us to visualize every cell under high power microscopes such as Nomarski microscopes.

To establish its invariant cell lineages, *C. elegans* utilizes both cell-autonomous and intercellular signaling. Signaling between cells or tissues is prevalent in the development of higher organisms, and may be studied more easily in this less complex system. For example, differentiation of vulval precursor cells into vulval cells requires an inductive signal from outside of the precursor cells, as discussed below. One can study cell-cell interactions *in*

vivo using *C. elegans* cell ablation technique(Sulston and White, 1980). Any single cell can be removed by use of a laser microbeam, and the consequence of the absence of this particular cell in the development process can be observed and tested.

Genetics can be carried out with relative ease with *C. elegans* (Brenner, 1974). *C. elegans* has two sexes, hermaphrodites and males. Self-fertilization by the hermaphrodites is the most common means for reproduction, but males can mate with hermaphrodites to make cross progeny. This aspect facilitates genetic experiments using *C. elegans* , as both self-fertilization and cross-fertilization can be used in constructing strains and genetic mapping. Adult animals are about 1 mm long, and feed on bacteria such as *E. coli*, making culturing a good number of strains relatively easy. Due to the short life cycle time (~ 3.5 days), one can carry out an F2 screen in one week. Many phenotypes can be observed using a dissecting microscope, simplifying a mass screen of mutations of interest. For example, one can screen 10,000 worms for particular mutants in minutes to hours using dissection microscope. Mutagenesis can be carried out on a large scale using different kinds of mutagens as well as ethyl methane sulfonate(EMS) (Moerman and Baillie, 1981, Rosenbluth, et al., 1985, Stewart, et al., 1991), . *C. elegans* has five sets of autosomes and one set of sex chromosomes, thus genetic mapping of mutations is relatively easy. Also, duplications and deficiencies are available for fine mapping in many regions of each chromosome (Herman, et al., 1979, Johnsen and Baillie, 1991, Rogalski, et al., 1982, Sigurdson, et al., 1984). Many genetic tools such as suppression experiments and epistasis experiments, can be used to study interactions of genes involved in particular developmental processes (Avery

and Wasserman, 1992, Brenner, 1974, Hodgkin, et al., 1987, Wood, 1988). In suppression experiments, one can isolate mutations that suppress a mutation that causes a certain phenotype. For example, *let-23* loss-of-function mutations cause a vulvaless phenotype, thus a defect in egg-laying, and these mutations can be used in a screen in which one looks for animals that can lay eggs. These animals are expected to carry another mutation that suppress the vulvaless phenotype of *let-23* mutations (Jongeward and Sternberg, 1993). In epistasis experiments, one can deduce which gene acts upstream or downstream of a particular gene involved in the same genetic pathway. For example, mutations in many loci cause a vulvaless phenotype, and mutations in other loci cause a multivulva phenotype. One can determine the hierarchy of genes in this pathway by constructing double mutants carrying a vulvaless mutation and a multivulva mutation and observing the phenotype of the double mutant animals (Avery and Wasserman, 1992, Ferguson et al., 1987, Hodgkin et al., 1985). Genetic mosaic analysis can be carried out using duplications or transgenes to elucidate the focus of action of particular genes (Herman, 1984, Herman, 1987, Herman, 1989, Schedin, et al., 1991, Siddiqui and Babu, 1980).

The molecular biology of *C. elegans* has been greatly pursued in the last few years. The genome size of *C. elegans* is about 1x 10^8 bases, which is only five times of the yeast genome. The *C. elegans* genome project has focused on the physical mapping and sequencing of the entire genome of the animal. So far, an almost complete physical map defined by overlapping and ordering YACs and cosmids covers about 97% of the genome, and is available through a database called ACEDB (A *C. Elegans* Data Base) (Coulson, et al., 1991, Coulson, et al., 1988, Coulson, et al., 1986). A hybridization filter

containing YACs that cover the genome is available, and can be used to find physical map positions of cloned pieces of DNA. Furthermore, the physical map is well correlated with the genetic map, making it possible to clone genes which are only known by their genetic map positions, or to map genes which were identified by molecular methods. Cloning of genes known by their mutations is facilitated by transgenic animal techniques (Fire, et al., 1990, Mello, et al., 1991). Transgenic animals are constructed using microinjection of a DNA of interest into the gonad of adult hermaphrodites, whose developing germ cells take up the DNA and inherit it to their progeny. The phenotype of the resultant progeny are tested for rescue of the phenotype by the presence of extragenic copies of injected DNA. Transposon tagging and polymorphism mapping using transposons can also be utilized to facilitate cloning of the genes of interest (Emmons et al., 1983, Moerman and Waterston, 1984). There are two different wild-type strains of *C. elegans:* the Bergerac strain and the Bristol strain. The Bergerac strain has many copies of the Tc1 transposons (~300 copies) actively transposing around the genome, whereas the Bristol strain has fewer copies of the transposon, which are inactive (Emmons, 1988). A Bergerac strain can be used to obtain mutations in a gene of interest by F2 screens or by F1 non-complementation screens (transposon tagging) (Moerman and Waterston, 1984). *lin-3* was cloned using this method (Hill and Sternberg, 1992). Another way to utilize these two different strains is to make heterozygotes with one set of transposon-rich chromosomes marked with certain marker mutations, and one set of transposon-poor chromosomes marked with different marker mutations, and then to look for transposons closely linked to a gene of interest by making recombinants and using inverse PCR techniques (Ochman, et al., 1988,

Ruvkun, et al., 1989). One can map these transposon polymorphisms genetically and physically to correlate the genetic and physical mapping position of the gene of interest and the polymorphisms. *unc-101* was cloned using this technique (Chapter 2).

Many genes molecularly analyzed in *C. elegans* have homologs in other organisms (e.g., Aroian et al., 1990, Chen et al., 1993, Chisholm, 1991, Clark et al., 1993, Clark et al., 1992, Greenwald, 1985, Han et al., 1993, Han and Sternberg, 1990, Hill and Sternberg, 1992, Kenyon and Wang, 1991, Lochrie et al., 1991, Miller et al., 1993, Shackleford et al., 1993, Yochem et al., 1988), thus studying *C. elegans* homologs will help us understand aspects of development in higher organisms. Also it is possible that many genes will be more easily identified and analyzed in *C. elegans* , as compared to other higher organism. This hopefully will lead to the discovery of new genes essential for development of higher organisms including humans.

II. Vulval differentiation in *C. elegans* and genes involved in the vulval induction pathway.

One typical phenomenon in the development of multicellular organisms is that of 'induction', where developing cells receive signals from other cells to acquire an ability to undergo a particular developmental fate (Gurdon, et al., 1989, Hart, et al., 1990, Jessell and Melton, 1992, Smith, 1989, Sternberg and Horvitz, 1989, Zipursky, et al., 1992). This process requires the production of a signal, a signal transduction cascade within responding cells, and possibly other regulatory events from the surrounding environments. The development of the vulva in the nematode *C. elegans* provides an excellent system for studying both *in vivo* and *in vitro* signal

transduction and its regulation in an induction process.

Differentiation of the vulva : overview

Hermaphrodites and males of *C. elegans* are similar in terms of general structure and anatomy, but they have different reproductive organs. Males produce sperm as gametes, and have specific structures for copulation in their tails, called spicules and rays. Hermaphrodites do not have spicules nor rays, but they produce both oocytes and sperm, allowing for self-fertilization. For laying eggs, hermaphrodites have a vulva, which is also used for copulation with males to produce cross progeny. The vulva is not an essential organ for viability, because animals without vulvae still can self-fertilize their oocytes and sperm, and these eggs can hatch inside the mother, eventually making their way out of the mother.

In wild-type intact animals, the vulva is composed of 22 nuclei that are progeny of three ventral hypodermal vulval precursor cells (VPCs) called P5.p, P6.p, and P7.p (figure 1). Three other ventral hypodermal cells, called P3.p, P4.p, and P8.p, which normally divide once and fuse with the hypodermis, can produce vulval cells in some cases as discussed below. These six cells are collectively referred to as vulval precursor cells (VPCs) because they all have the potential to become vulval cells. The cell lineage generated by the progeny of P6.p and either P5.p or P7.p, which normally give rise to vulval cells, are different. The P6.p cell undergoes two rounds of longitudinal divisions to produce four grand daughters that all divide transversely (this pattern of division is called 'TTTT'), thus producing 8 nuclei. The P5.p or the P7.p cell undergoes two rounds of longitudinal divisions to produce four grand daughters, one of which does not divide, one of which divides transversely, and two of which divide longitudinally (this pattern is referred to as NTLL),

ending up with 7 nuclei. The P3.p, P4.p, and P8.p cells undergo only one round of cell division to give rise to two daughters that fuse with the hypodermis. The cell lineage of P6.p is considered the 'primary' fate because when the P6.p cell is ablated by a laser microbeam early enough in the development, a neighboring cell migrates in and generates the P6.p cell lineage (Sternberg and Horvitz, 1986, Sulston and White, 1980). The lineage of P5.p or P7.p is considered the 'secondary' fate because when these cells are ablated, neighboring cells (P4.p, or P8.p) will adopt the cell lineage of P5.p or P7.p, respectively. The P6.p cell does not adopt a secondary fate when P5.p or P7.p is ablated. If all three cells of P5.p, P6.p, and P7.p are ablated, the other three VPCs, P3.p, P4.p, and P8.p, can replace them to produce a functional vulva. The cell lineage of P3.p, P4.p, and P8.p are considered the 'tertiary' fate.

One of the reasons why vulval development is fascinating for studying developmental biology is that this process requires a signal from outside of the precursor cells, exemplifying the phenomenon of induction (Figure 2). The anchor cell (AC) in the somatic gonad is necessary and sufficient to induce the differentiation of vulval cells (Kimble, 1981, Sulston and White, 1980) When the anchor cell is ablated before any division of the VPCs occur, all six VPCs undergo the tertiary fate and fuse with the hypodermis (Figure 3B). When the entire gonad except the anchor cell is ablated, a normal vulva is induced (Figure 3C). After these cell biological observations of the vulval induction, genetic analysis of vulval induction mutants was the next step to study this phenomenon. Many genes have been identified by their mutant phenotypes associated with vulval development (Figure 4) (Ferguson and Horvitz, 1985, Horvitz and Sulston, 1980, Sulston and Horvitz, 1981).

Vulvaless (vul) genes such as *lin-3, let-23, lin-2, lin-7,* and *lin-10* were identified because loss-of-function mutations in any of these genes resulted in a vulvaless (Vul) phenotype where fewer than three VPCs are induced. Multivulva genes such as *lin-15* and *lin-1* were identified because loss-of-function mutations in these genes resulted in a multivulva phenotype where more than three VPCs are induced to become vulval cells. Other genes such as *lin-45, unc-101, sem-5, let-60,* and *sli-1* were identified via reverse genetics or suppression genetics (Han, et al., 1990, Han, et al., 1993, Jongeward and Sternberg, 1993) (Beitel, et al., 1990, Clark, et al., 1992). These genes involved in vulval induction have been placed in a genetic pathway by epistasis analysis (Figure 5). Molecular analysis of the genes in vulval induction revealed a basic similarity of the vulval induction pathway to a mammalian epidermal growth factor(EGF)- EGF receptor (EGFR) signal transduction pathway (table 1). The following sections will describe genetic and molecular characteristics of genes involved in vulval development.

lin-3 encodes the signal for vulval induction.

lin-3 is a gene required for vulval induction because reduction-of-function mutations of *lin-3* result in a vulvaless phenotype in which fewer than three VPCs are induced to generate vulval cells (Ferguson and Horvitz, 1985, Ferguson, et al., 1987, Horvitz and Sulston, 1980, Sulston and Horvitz, 1981). Epistasis experiments suggested that *lin-3* acts before *let-23* and *let-60*, genes required for vulval induction (Aroian and Sternberg, 1991, Ferguson, et al., 1987, Han and Sternberg, 1990). As discussed below, *let-23* encodes a homolog of Epidermal Growth Factor Receptor (EGFR), and *lin-3* is the only vulvaless gene acting upstream of *let-23,* making *lin-3* a good

candidate for the signal from the anchor cell. The *lin-3* gene was cloned by a transposon tagging method (Hill and Sternberg, 1992). The molecular analysis of the *lin-3* gene revealed that *lin-3* encodes a protein with an extracellular domain similar to an Epidermal Growth Factor (EGF) motif (Carpenter and Cohen, 1990, Hill and Sternberg, 1992). When the *lin-3* gene is overexpressed by extrachromosomal transgene, the transgenic animals display a gonad-dependent multivulva phenotype. The *lin-3* transgene cannot cause a multivulva phenotype in the absence of functional *let-23*, indicating that *lin-3* acts via *let-23*. The *lin-3::lacZ* fusion gene is expressed only in the anchor cell at the time of vulval induction. In the transgenic animals carrying a heatshock construct that allows for production of a secreted form of the Lin-3 protein, more than three VPCs undergo vulval differentiation in the absence of the entire gonad including the anchor cell. Therefore, the secreted form of the Lin-3 protein is sufficient to induce vulval fate (R. Hill, W. Katz, and P. Sternberg, in prep). When a VPC is isolated by ablation of all other VPCs in the heat-shocked animals carrying the heatshock construct of *lin-3* , the VPC adopts either a primary or secondary fate, indicating that the EGF motif itself is capable of inducing both the primary and secondary fate of the VPCs (R. Hill, W. Katz, and P. Sternberg, in prep.). Those results, taken together, indicate that *lin-3* encodes the signal for vulval induction made by the anchor cell.

let-23 encodes a putative receptor for the signal.

let-23 is a complex locus that is required in at least five different tissues, including the vulval cells (Aroian and Sternberg, 1991). *let-23* function is required for the proper induction of vulval cells. Complete loss-of-

function of *let-23* results in lethality and a vulvaless phenotype in which no VPCs are induced . Thus, *let-23* activity is required for vulval induction. Also, a gain-of-function mutation of *let-23* results in excessive vulval induction (W. Katz and G. Lesa, per. comm.). However, *let-23* is also thought to be required for some sort of negative regulation of vulval induction, because certain alleles (for example *n1045)* whose activity is between null and wild type can cause a hyperinduction of vulval cells so that more than three cells are induced to generate vulval cells (Aroian and Sternberg, 1991). Therefore, *let-23* plays two opposite roles in vulval induction.

let-23 was cloned by correlating the genetic and physical maps of its region (Aroian, et al., 1990). Molecular analysis of *let-23* showed that it-encodes a putative receptor tyrosine kinase of the EGFR family (Aroian, et al., 1990, Schlessinger and Ullrich, 1992). Unlike the *lin-3* transgene, the *let-23* transgene does not cause any multivulva phenotype (Aroian, et al., 1990). Thus, it is conceivable that *let-23* is not a limiting factor in this pathway, and that *lin-3* is a limiting factor. It is possible that in the *let-23* transgenic animals, more LET-23 proteins can not be activated due to limited quantities of the LIN-3 protein. Another possibility is that although the transgene is present as multicopy, they cannot be overexpressed, or if they can be overexpressed, they can not be activated, because of the presence of negative regulators of *let-23* expression.

Mosaic analysis using the fusion of a *let*-23 transgene and a free duplication (sDp3) bearing a cell-autonomous marker (*ncl-1*) suggests that *let-23* is required in the VPCs for vulval induction (M. Koga and Y. Oshima, personal comm.), consistent with the idea that *let-23* encodes a receptor for the signal.

Molecular analysis of mutations associated with *let-23* alleles showed that mutations in the kinase domain or ligand binding domain result in severe reduction of *let-23* function, and that two less severe mutations bearing C-terminal truncations result in tissue-specific defects (Aroian, et al., 1993). Specifically, *let-23(sy1)* is a mutation that results in a truncation of 6 amino acid residues at the C-terminus. *let-23(sy1)* causes a vulvaless phenotype, but does not show any lethality. *let-23(sy97)* is a mutation that alters a splicing acceptor site at the C-terminus, resulting in a shorter protein than *let-23(sy1)* . *let-23(sy97)* causes both lethality and a vulvaless phenotype. *let-23(sy1)* was the mutation that was used for isolating *unc-101* mutations as a suppressor of the vulvaless phenotype (chapter 2).

***lin-2, lin-7,* and *lin-10*: genes required for vulval induction:**

lin-2, lin-7, and *lin-10* are also required for vulval induction. Mutations in any of these genes cause a vulvaless phenotype (Ferguson and Horvitz, 1985, Horvitz and Sulston, 1980). However, unlike other genes (for example, *lin-3, let-23, let-60,* and *lin-45)* whose loss-of function mutations cause a lethality, mutations in these genes do not cause any other phenotype than a vulval defect, suggesting that these genes are specific to the vulval pathway. Like *let-23*, some reduction-of-function alleles of these genes cause a hyperinduced phenotype, suggesting that these genes are also required for some sort of negative regulation of vulval induction (G. Jongeward, per. comm.). It is not clear where in the vulval induction pathway these genes act, because the phenotype of mutation in *let-23, sem-5* and these genes are all vulvaless, making epistasis analysis impossible. Now that a gain-of-function allele of *let-23* is available, it might be possible to place these genes

relative to *let-23* in the pathway.

lin-10 was cloned and was shown to encode a novel protein (Kim and Horvitz, 1990). *lin-2* was cloned and shown to encode a GMP kinase (R. Hoskins and S. Kim, per. comm.). One way in which a GMP kinase might work is to activate *let-60* ras. The Lin-2 protein is similar to a p55 protein, a palmylalyated membrane protein that co-purifies with a serine kinase, and is predicted to contain a GLGF repeat, an SH3 domain, and a domain similar to yeast guanylate kinase. The function of GLGF repeat is unknown. *lin-7* was also cloned and shown to encode a protein with a GLGF domain without any apparent functional domains, unlike *lin-2* (J. Simske and S. Kim, per. comm.).

sem-5 encodes a protein that acts as an adapter between activated receptor tyrosine kinase and ras activator.

Mutations in *sem-5* cause at least two phenotypes: a sex myoblast migration defect and a vulvaless phenotype(Clark, et al., 1992). Each of these two phenotypes can be obtained independently by different alleles of *sem-5*, indicating that the SEM-5 protein may have separate domains required for each function. In the vulval induction pathway, *sem-5* is thought to act between *let-23* and *let-60*, based on the epistasis and molecular analysis in other systems. Molecular analysis of *sem-5* revealed that this gene encodes a protein that contains essentially nothing but one SH2 domain flanked by two SH3 domains. In *Drosophila* and mammalian cells, the SH2 domain was found to be responsible for binding to a phosphorylated EGF receptor, and SH3 domains for binding to a guanidine exchange factor for ras (Buday and Downward, 1993, Egan, et al., 1993, Li, et al., 1993, Oliver, et al., 1993,

Rozakisadcock, et al., 1993, Simon, et al., 1993), Therefore *sem-5* is believed to mediate the coupling of receptor tyrosine kinases to Ras signaling for vulval induction.

let-60, a gene downstream of _lin-3_ and _let-23_, encodes a _ras_ homolog.

let-60 is believed to act as a switch in the vulval induction pathway as loss-of-function mutations cause a vulvaless phenotype and gain-of-function mutations cause a signal-independent multivulva phenotype (Beitel, et al., 1990, Han, et al., 1990). There are also dominant negative mutations in *let-60* that cause a dominant vulvaless phenotype. Epistasis experiment showed that *let-60* acts downstream of *lin-3* and *let-23*. Molecular cloning of this gene showed that *let-60* encodes a *ras* homolog (Barbacid, 1987, Downward, 1992, Han and Sternberg, 1990, Lowy and Willumsen, 1993, Milburn, et al., 1990). The genetic conclusions about *let-60* acting downstream of *lin-3* and *let-23* is consistent with its molecular nature being a ras homolog.

A _raf_ homolog is involved in vulval development, and is encoded by _lin-45_.

raf is a serine/threonine kinase involved in signal transduction pathways (Heidecker, et al., 1992). A *C. elegans raf* homolog was cloned by its homology to mammalian *raf* and was found to be encoded by *lin-45*, a gene isolated as a suppressor of a *lin-15* multivulva mutation (Han, et al., 1993). A reduction-of function mutation of *lin-45* causes a vulvaless phenotype. A transgene that contains a presumed dominant negative mutation in *lin-45* can cause a dominant vulvaless phenotype. Therefore, a *raf* kinase is involved in the vulval induction pathway. Epistasis experiments suggested

that *lin-45* acts either together with or after *let-60 ras.*

A MAP kinase is involved in vulval induction pathway.

Mitogen Activated Protein (MAP) kinases are serine/threonine kinases that are activated by *ras*-mediated signaling pathways and other signaling pathways (Howe, et al., 1992, Kyriakis, et al., 1992, Pelech and Sanghera, 1992). Mutations in a MAP kinase gene were isolated as suppressors of a *let-60(gain-of-function: gf)* multivulva phenotype (Y. Wu and M. Han, K. Kornfeld, and R. Horvitz, per. comm.). A MAP kinase gene was also cloned by reverse genetics (M. Lackner and S. Kim, per. comm.). The gene is called *sur-1* or *mpk-1.* While a *let-60(gf)* mutation causes a multivulva phenotype, the double mutation of *let-60(gf)* and this MAP kinase gene results in a largely wild-type vulval induction. Epistasis experiment shows that this gene likely acts after *let-60.* The phenotype of a single mutation of this MAP kinase is wild type. A hybrid construct that contains the regulatory region of *sur-1* and a cDNA portion of the mammalian ERK2 gene can functionally complement a mutation of *sur-1* (Y. Wu and H. Min, per. comm.).

***lin-1*, a gene downstream of *lin-45*, may encode a transcription factor.**

A loss-of-function mutation in *lin-1* causes a gonad-independent multivulva phenotype. Epistasis experiments showed that *lin-1* is the most downstream gene characterized so far in the vulval induction pathway, because a mutation in *lin-1* can suppress the vulvaless phenotypes of *lin-3, let-23, sem-5,* or *lin-45*(Ferguson, et al., 1987)(Ferguson et al., 1987, Clark et al., 1992, Han et al., 1993, Han and Sternberg, 1990). Unlike other genes in the main pathway of vulval induction such as *lin-3, let-23, let-60, sem-5,* and

lin-45, lin-1 is required for preventing VPCs from adopting vulval fates. There is no gain-of-function mutation of *lin-1* available that can test whether hyperactivity of this gene can act to prevent vulval induction. *lin-1* was cloned and shown to encode a transcription factor of the Ets family (G. Beitel and R. Horvitz, per. comm.).

***lin-15,* a negative regulator, encodes two new proteins.**

Because only three out of six VPCs are induced to generate vulval cells in wild-type intact animals, there must exist mechanisms to regulate proper induction so that only three AC-proximal VPCs are induced. One way would be production of limited amount of signal. The production of the signal is well regulated so that the signal is produced in only the anchor cell at the time of vulval induction, and the level of production is limited. It is also possible that the processing of LIN-3 protein may be involved in the regulation. To get invariant development, however, the animals may need much more regulations. A way to make sure that only cells receiving the inductive signal can adopt vulval fate would be to prevent VPCs from adopting vulval fates in the absence of signal. One such pathway is defined by *lin-15.*

Loss-of-function mutations in the *lin-15* locus confer a gonad-independent multivulva phenotype in which more than three VPCs are induced even in the absence of the entire gonad. This phenotype is suppressed by *let-23* mutations, indicating that *lin-15* acts before *let-23,* and that *let-23* is negatively regulated by *lin-15.* The wild-type function of *lin-15* is thought to prevent activation of the LET-23 receptors in the absence of the signal (Huang, et al., 1993). The inductive signal is thought to override this

negative regulation and induce three proximal VPCs to generate vulval cells. As shown by mosaic analysis (Herman and Hedgecock, 1990), *lin-15* acts in a non-autonomous manner. *lin-15* is required for its function probably in the hypodermis that surrounds the VPCs. This result is consistent with the fact that *let-23* encodes a receptor and *lin-15* acts upstream of *let-23*.

Genetically, the *lin-15* locus is a complex locus that consists of two sub-complementation groups each of which complements the other. These groups are named class A and class B (Ferguson and Horvitz, 1989). To cause a multivulva phenotype, members of both classes should be mutated. Any mutations in the *lin-15* (A) locus combined with any mutations in the *lin-15 (B)* locus will give a multivulva phenotype. There are many other genes that act as either class A or class B genes. A mutation in any one of the class A genes can cause a multivulva phenotype in combination with a mutation in any of the class B genes. Therefore, *lin-15(A)* with other class A genes and *lin-15(B)* with other class B genes define two redundant pathways for negative regulation of vulval induction.

Molecular analysis of *lin-15* showed that indeed *lin-15* encodes two functional transcripts (L. Huang et al., in prep.). These two transcripts are under the control of a single promoter, demonstrating a polycistronic gene in eukaryotes (Spieth, et al., 1993). The gene products of *lin-15* do not show any similarity to any known proteins, including other class A or class B genes cloned so far. This result is not very surprising because in *in vitro* systems such as tissue culture system, cell non-autonomous genes can not be identified through biochemical or genetic methods.

There are more negative regulators: unc-*101*, *sli-1*, and *rok-1*.

Another way to regulate vulval induction would be to regulate the EGFR signaling pathway by negative regulators within VPCs. Many such candidate genes, for example *unc-101, sli-1* and *rok-1*, have been isolated and characterized (G. Jongeward, personal comm., studies in this thesis). A combination of all these regulatory actions will ensure proper induction of three VPCs to generate vulval cells in intact animals.

sli-1 and *unc-101* mutations were isolated as suppressors of a vulvaless phenotype of certain reduction-of-function alleles of the *let-23* receptor tyrosine kinase gene. Mutations in *sli-1* and *unc-101* can suppress the vulvaless phenotype of a *let-23(sy1)* mutation to a hyperinduced phenotype. A single mutation of either *sli-1* or *unc-101* does not cause any vulval defect. Double mutant animals for *sli-1* and *unc-101*, however, display greater than wild-type vulval induction, indicating that these two genes are partially redundant negative regulators of the vulval induction pathway. Mutations in *unc-101* or *sli-1* do not suppress null mutations of *let-23*, suggesting that these genes cannot bypass the absence of active LET-23 receptors. Therefore, the roles of these genes are thought to be in the fine regulation of active receptors. Genetic interaction experiments suggested that these genes act near or at the *let-23* step in the pathway (G. Jongeward and P. Sternberg, in prep., Chapter 2 of this thesis). *sli-1* mutations can suppress *let-23, sem-5, lin-2, lin-7,* and *lin-10* mutations, but not mutations in *lin-3, let-60,* or *lin-45. unc-101* mutations suppress certain alleles of *let-23, lin-2, lin-7,* and *lin-10* very well, and also partially suppress *lin-3* and *lin-45* reduction-of-function alleles.

Although *unc-101* and *sli-1* are partially redundant, it is conceivable that *unc-101* has its own functions that are not redundant, because *unc-101*

mutations confer pleiotropic phenotypes on its own. *unc-101* mutant animals are defective in coordinated movement (D. Riddle, per. comm.), defecation (Thomas, 1990), neuronal uptake of a dye (E. Hedgecock, per. comm.), male spicule structure (H. Chamberlin, per. comm.), and viability (Lee, et al., 1993), suggesting that *unc-101* is involved not only in the EGFR signaling process of vulval precursor cells, but also in other signaling process in other types of cells.

sli-1 was cloned and found to encode a protein similar to a proto-oncogene *c-cbl* (C. Yoon, G. Jongeward, J. Lee, and P. Sternberg, unpublished results). The function of *c-cbl* in the oncogenic process is not known. Further studies of *sli-1* will extend the understanding of the role of the *c-cbl* proto-oncogene in the EGFR signaling pathway.

unc-101 has been cloned and shown to encode a homolog of AP47, the medium chain of the *trans*-Golgi clathrin associated protein complex, suggesting that clathrin coated pits and vesicles are involved in the negative regulation of an EGF signaling pathway (See Chapter 2 for details). It is not clear whether clathrin coated pits and vesicles on the plasma membrane are also involved in the regulation of this pathway. Further study on coated vesicles in *C. elegans* will help extend the understanding of roles of clathrin coated vesicles in this signal transduction process.

A *rok-1* mutation was isolated by its phenotype of a hyperinduced vulva in the presence of a *sli-1* mutation (Chapter 4). Because *sli-1* mutations can cause a hyperinduced phenotype with *unc-101* mutations, and because we assumed that there are more negative regulators involved in this pathway, we carried out a mutagenesis where *sli-1* mutant animals were mutagenized and animals with hyperinduced vulvae were isolated. In

addition to a *rok-1* allele, alleles of other genes that had been expected to be isolated by this screen were also isolated. A single mutant for any of *unc-101, sli-1,* or *rok-1* displays wild-type vulval induction. However, double mutants of *unc-101; rok-1*, *rok-1; sli-1,* and *unc-101; sli-1* display greater than wild type vulval induction. Furthermore, the triple mutants of *unc-101; rok-1; sli-1* display even greater vulval induction than any of the double mutants mentioned above. Therefore, these three genes, *unc-101, rok-1,* and *sli-1* define redundant negative regulators of the vulval induction pathway. It is not clear whether there is any more negative regulators of vulval induction pathway as the screen using *sli-1* mutant animals has not been performed to saturation.

Are there any more genes involved in vulval induction?

Basically, the vulval induction pathway is an EGF-EGFR signaling pathway (Horvitz and Sternberg, 1991, Sternberg and Horvitz, 1991). Considering a parallel pathway in mammalian cells, some components that play important roles in mammalian pathway are still missing in the vulval induction pathway. For example, GAP protein, Protein Kinase C and PLC-γ, are not implicated in the vulval induction pathway yet (Parsons, 1990) (Margolis, et al., 1990, Simon, et al., 1991, Troppmair, et al., 1992). It is not clear whether these genes are not involved in vulval induction, or their functions are redundant in *C. elegans*, making it difficult to isolate mutations. Or mutations in these genes might be lethal. Reverse genetics to obtain these genes may help understand their roles in *C. elegans*.

Genetic and molecular analysis of the negative regulators revealed new components of this signal transduction pathway. The *lin-15* gene encodes

novel proteins (Huang, et al., 1993), and *unc-101* being a clathrin adaptor protein (Chapter 2) serves as the first genetic evidence that the *trans*-Golgi clathrin-coated vesicles are involved in the negative regulation of the EGFR signaling pathway. *sli-1* cloning revealed a new proto-oncogene as a component of this signal transduction pathway. It would not be surprising if new genes involved in the vulval induction pathway encode novel proteins. Further genetic screens and analysis using newly identified genes such as *unc-101, sli-1,* and *rok-1,* will likely identify more genes regulating the vulval induction pathway.

III. Clathrin coated vesicles

unc-101, a negative regulator of the vulval induction pathway, encodes a homolog of AP47, the medium chain of the *trans*-Golgi clathrin-associated complex AP-1 (Chapter 2). This is the first genetic evidence that *trans*-Golgi clathrin coated pits and vesicles are involved in negative regulation of the EGF signaling pathway. In this section I summarize the structure and function of clathrin coated vesicles and their associated protein complexes (APs).

Clathrin coated vesicles

Clathrin coated vesicles are organelles that originate from the plasma membrane and the *trans*-Golgi membrane and mediate intracellular protein trafficking (figure 6). The clathrin coated pits on the plasma membrane is involved in the endocytosis of signaling receptors such as the EGF receptor, and constitutively recycling receptors such as the LDL receptor and the mannose-6-phosphate receptor (Schmid, 1992). The endocytosis of signaling

receptors is different from that of constitutive receptors because receptors such as the LDL receptors are meant to be recycled for continuous function, and signaling receptors such as the EGF receptors are to be degraded after endocytosis. The endocytosis of activated receptors can serve as a way of down-regulating the signaling pathway. The cytoplasmic tails of many receptors contain signals for internalization via clathrin coated vesicles (Chen, et al., 1990, Johnson, et al., 1990, Prywes, et al., 1986, Riedel, et al., 1989, Sosa, et al., 1993).

One of the best known functions of the coated vesicles on the *trans*-Golgi is for intracellular sorting of the lysosomal enzymes (Kornfeld, 1990). Lysosomal enzyme sorting occurs via two routes. For some enzymes such as lysosomal acid phosphatase, enzymes are secreted, and then recaptured via plasma membrane clathrin-coated vesicles for packaging into lysosomes. For most enzymes, sorting occurs intracellularly via *trans*-Golgi clathrin coated vesicles after biosynthesis of lysosomal enzymes. Clathrin coated vesicles on the *trans*-Golgi collect mannose-6-phosphate receptors that recognize and bind lysosomal enzymes tagged with mannose-6-phosphate, and transfer these receptor/ lysosomal enzyme complexes to the lysosome. There are two kinds of mannose-6-phosphate receptors: cation-independent mannose-6-phosphate receptors (CIMPRs) and cation-dependent mannose-6-phosphate receptors (CDMPRs). CIMPR is thought to be involved in both the sorting mechanisms, and CDMPR is involved in only the intracellular sorting mechanism, although both the receptors are known to recycle between the Golgi and the plasma membrane.

Coated vesicles are composed of membrane, membrane proteins for

trafficking, clathrin triskelia, and an associated protein complex (also called adaptor complex, abbreviated as APs; Figure 6). A clathrin triskelion is composed of three clathrin heavy chains and three light chains (Figure 7). The Carboxyl termini of the heavy chains are located at the center of the triskelion, and the amino termini form the terminal domain. The light chains are bound to the heavy chains at heavy chain segments about 60kD from the carboxy terminus. There are two types of clathrin light chains, namely LC_a and LC_b, in mammalian cells (Brodsky, 1988).

Genes for the clathrin heavy chain have been cloned from rat, bovine, Drosophila, nematode, and yeast (Bazinet, et al., 1993, Kirchhausen, et al., 1987, Payne and Schekman, 1985, Waterston, et al., 1992). It seems that there is only one gene for the heavy chain in these species. Two clathrin light chains are encoded by separate genes in mammalian cells. In addition, clathrin light chain genes are transcribed in different spliced forms in different tissues (Jackson, et al., 1987) (Kirchhausen, et al., 1987). Neuron-specific transcripts of both LC_a and LC_b have extra nucleotides due to alternative splicing events. Together with separate genes encoding similar gene products, alternative splicing is a way of offering a diversity of proteins with important functions. It is known that the kinase activity associated with the EGF receptor can phosphorylate LC_a , not LC_b , on its tyrosine residues (Gliekman, et al., 1989), suggesting that phosphorylation on LC_a may be required for internalization of the receptors. In yeast, only one type of clathrin light chain is expressed (Silveira, et al., 1990).

Clathrin-associated protein complexes (APs)

The APs lie between the membrane proteins for transport and the

clathrin cage in the clathrin coated vesicles, probably interacting with both the proteins and the membrane. The APs are also known to promote clathrin assembly. While the clathrin triskelion is a common structural unit of both the plasma membrane and *trans*-Golgi membrane coated vesicles, the APs are the components that appear to confer specificity to the plasma membrane and *trans*-Golgi coated vesicles (reviewed in (Keen, 1990, Pearse and Robinson, 1990)).

AP-1 and AP-2 are tetramers of four different subunits (Figure 8), and have very similar structures. The components of the APs are different, but are molecularly related. AP-1, the associated complex specifically located on the *trans*-Golgi coated vesicles (Ahle, et al., 1988), consists of two heavy chains, β' and γ-adaptin, one medium chain, AP47, and one small chain, AP19 (Matsui and Kirchhausen, 1990). AP-2, the associated complex specifically located on the plasma membrane coated vesicles, consists of two heavy chains of α and β-adaptin, one medium chain, AP50, and one small chain, AP17. Amino acid sequences of α– and γ–adaptin revealed their homology, so did those of β– and β'-adaptins, (Kirchhausen, et al., 1989, Robinson, 1989, Robinson, 1990). AP50 and AP47 are about 40 % identical to each other (Nakayama, et al., 1991, Thurieau, et al., 1988). Small chains are also similar to each other (Kirchhausen, et al., 1991). All the genes for these components have been cloned from rat, mouse, or bovine tissues. α-adaptins are encoded by two different genes in mouse (Robinson, 1989). Although both of the genes are expressed in the same sets of cells, the expression level is different in different type of cells. Whether these two proteins have different functions is not known. Single genes are thought to encode all other components. However, it is not clear whether thorough search for additional

genes has been carried out for the other genes. In yeast, homologs of β-adaptin, small chains and medium chains have been cloned (Kirchhausen, 1990, Kirchhausen, et al., 1991, Nakayama, et al., 1991). In *C. elegans,* three homologs of medium chains and a homolog of γ-adaptin have been cloned (Chapter 2, Chapter 3, and appendix, this thesis).

Biochemical functions of the components are not clear, but there have been reports that the medium chains may have kinase activity, though sequence analysis did not reveal any homology to known kinases (Myers and Forgac, 1993, Nakayama, et al., 1991, Pauloin, 1982).

There is a third clathrin-associated protein that is different from AP-1 and AP-2 in terms of structure and localization (Murphy, et al., 1991) . While AP-1 and AP-2 are found almost ubiquitously regardless of cell types (of course their subcellular localization is the *trans*-Golgi and the plasma membrane, respectively), this third AP, namely AP-3, is found only in neuronal cells. AP-3 is not a tetramer, but a single polypeptide of 180 KD. AP-3 was found to be more active than the other APs in promoting clathrin assembly (Lindner and Ungewickell, 1992). AP-3 was cloned from rat brain and was shown to encode a protein essentially identical to a mouse phosphoprotein F1-20 whose function had not been known (Morris, et al., 1993, Zhou, et al., 1993). Because AP-3 is found only in the neuronal cells, and actively promotes clathrin assembly, it is thought that AP-3 is required for clathrin assembly function specialized for the needs of the synapses such as synaptic vesicles. This gene is also subject to alternative splicing.

Interaction of clathrin-associated protein complexes

The interaction of clathrin-associated protein complexes within

themselves and with other components of clathrin-coated vesicles have been studied in the mammalian cells. Most biochemical studies have been carried out using AP-2, because AP-2 is more abundant than AP-1, is easier to purify, and the structures of AP-1 and AP-2 are considered similar.

The AP complexes undergo a strong, rapid self-assembly. The interactions between the components of AP-1 and AP-2 within the APs are so strong and stable that most *in vitro* purification methods cannot separate these components. That is one of the reasons that biochemical studies on the roles of each component has been very difficult. A study by expressing cloned components and trying reconstitution experiment in different conditions are being carried on (T. Kirchhausen, per. comm.). Another good method for studying individual components of the APs is genetics in which one can study the consequences of mutations in each protein.

As mentioned above, APs lie between the plasma or Golgi membrane and clathrin in the coated vesicles. It is conceivable that APs interact with membranes, receptors, and clathrin molecules. The domains of AP-2 required for the interaction of AP-2 and clathrin is the trunk domain of AP-2 that contains the N-terminal half of α- and β-adaptin, and intact AP50 and AP17 (Peeler, et al., 1993) (Figure 8). There is a conflicting report that β-adaptin is responsible for this interaction (Schroder and Ungewickell, 1991). The ear domain of AP-2 (Figure 8), which contains C-terminal halves of α- and β-adaptin, is not required for this interaction (Peeler, et al., 1993).

Interaction of APs with receptors have also been studied. Clathrin associated protein complexes AP-1 and AP-2 can bind the cytoplasmic tail of CIMPR. Tyrosine residues in the cytoplasmic tail are required for the interaction with AP-2, but not for AP-1, indicating that endocytosis and

trans-Golgi sorting may utilize different signals. CDMPR cytoplasmic tails contain two signals for their rapid endocytosis: two phenylalanines and one tyrosine residue (Johnson, et al., 1990). It is known that the trunk domain of AP-2 preferentially bind activated EGF receptors(Sorkin and Carpenter, 1993). For this interaction, phosphorylated tyrosine residues in the cytoplasmic region of EGFR are required (Sorkin, et al., 1992). It is not known which subunit of AP-2 directly binds the receptor. The trunk domain of AP-2 is also capable of binding the cytoplasmic domain of lysosomal acid phosphatase (LAP) with high affinity (Sosa, et al., 1993). An internalization signal in the LAP may be tyrosine or phenylalanine residue. AP-1 is not capable of binding the LAP cytoplasmic tail.

Interaction of APs with the plasma or Golgi membrane fraction also requires the trunk domain of AP-2 (Peeler, et al., 1993). When dissociated subunits of AP-2 were tested, α-adaptin, not β-adaptin, was found to be able to bind the plasma membrane fraction (Chang, et al., 1993), suggesting that α-adaptin mediates the interaction between AP-2 and the plasma membrane. By analogy to the AP-2 result, it is conceivable that γ-adaptin might mediate the interaction between AP-1 and the *trans*-Golgi membrane. Although it is not known whether there are specific receptors for the APs on the membranes, it is possible that there are proteins that are specifically located on the plasma membrane and the *trans*-Golgi membrane, and interact with AP-2 and AP-1, respectively.

Genetic studies on the clathrin vesicle proteins

Clathrin heavy chains have been shown to play important roles in both single-celled yeast and multicellular organisms such as the fruit fly. In yeast,

mutants deficient in a clathrin heavy chain are defective in the retention of the Kex2p protein in the Golgi apparatus (Seeger and Payne, 1992, Seeger and Payne, 1992). Kex2p protein is a protease in the *trans*-Golgi that is required for maturation of the α-mating factor. With mis-localization of Kex2p proteins, premature mating factors are secreted from these mutant cells. In addition, clathrin heavy chain-deficient yeast cells are inviable in some strains; in other strains that have suppressor genes, these mutants cells are viable, but slow growing (Lemmon, et al., 1990, Munn, et al., 1991). Bazinet et al. (1993) have recently found that a clathrin heavy chain gene is essential in *Drosophila*. In *Dictyostelium*, clathrin heavy chain-deficient cells display slower growth, defective pinocytosis, defective osmoregulation, and inability to complete the starvation-induced development cycle (Ohalloran and Anderson, 1992). A clathrin heavy chain was cloned in *C. elegans* (Waterston, et al., 1992), but no genetic data is available as to what roles clathrin heavy chains play in the nematode. There is no genetic data available in any system on functions of clathrin light chains.

Deletion of yeast homologs of the small chains of the associated protein complexes has little effect on cell growth, protein export, or endocytosis (H. Phan & G. Payne, personal comm.). Similarly, the deletion of a yeast homolog of medium chains does not result in any obvious phenotypic consequences, although there might be subtle phenotypes associated with these mutations (S.K. Lemmon, personal comm.). On the contrary, mutations in the medium chain AP47 of *C. elegans* result in pleiotropic phenotypes in *C. elegans* (Chapter 2). An AP47 protein is encoded by *unc-101*. Mutations in the *unc-101* gene cause behavioral and developmental defects including subviability, an uncoordinated movement, a defecation defect, a male spicule

defect. Also, mutations in *unc-101* suppress the vulvaless phenotype of a *let-23(sy1) EGFR* mutation, suggesting that *unc-101* is involved in the negative regulation of the vulval induction pathway. This is the first genetic evidence that clathrin-coated vesicles are involved in an EGF signal pathway. There is another AP47 homolog (called CEAP47) in the nematode (Chapter 4). A hybrid gene containing a part of *unc-101* and a part of CEAP47 can complement the UNC-101 function when expressed under the control of the *unc-101* promoter, indicating that these two proteins share some redundant functions. No genetic mutations for the CEAP47 gene is available, thus it is difficult to genetically analyze the extent of redundancy of the two medium chain genes.

Figure 1. Development of the vulva of *C. elegans*: Cell lineages of VPCs in vulval differentiation. Three out of six vulval precursor cells (VPCs), P5.p, P6.p, and P7.p undergo vulval cell differentiation, and the other three VPCs, P3.p, P4.p, and P8.p, fuse with syncytial hypodermal cells. The P6.p cell adopts a primary fate, P5.p and P7.p, secondary fates (see the text for details). T indicates a transverse cell division, and L, a longitudinal cell division. N stands for no cell division.

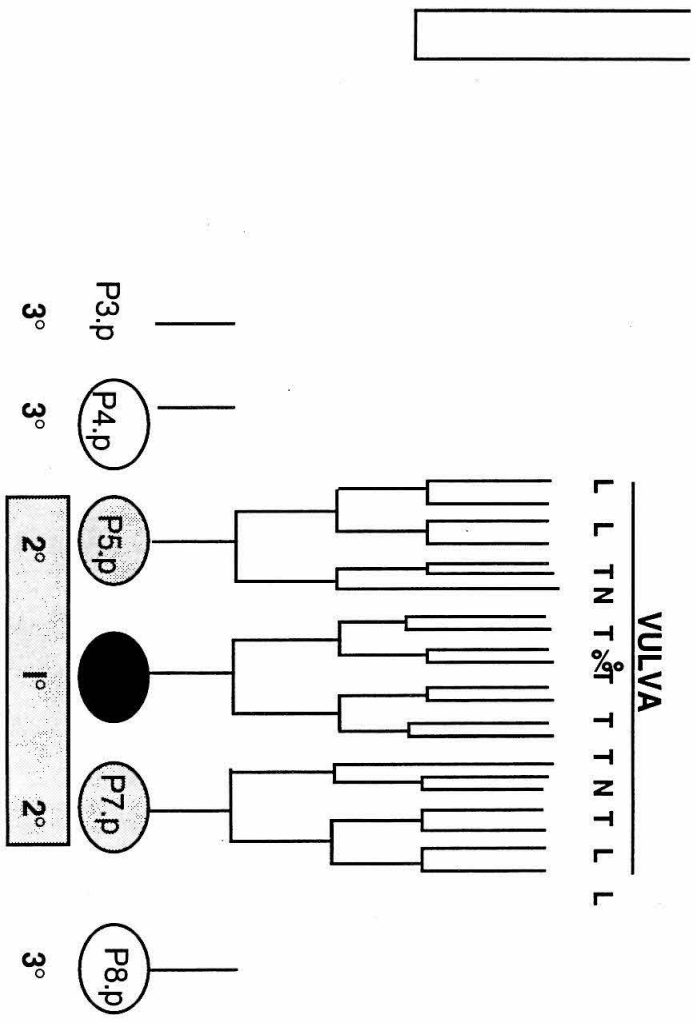
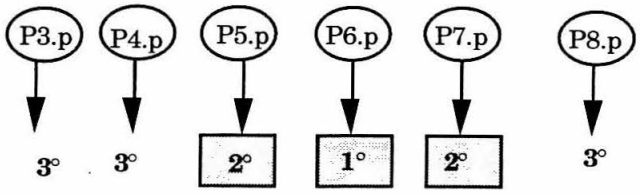
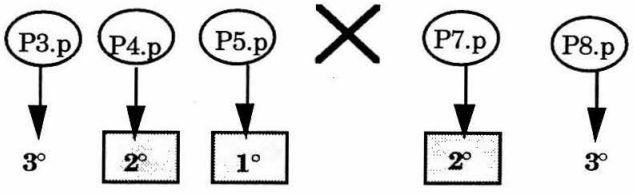

VULVA
L L TN T♂T TNTL L
P3.p
P4.p
P5.p
P7.p
P8.p
3°
3°
2°
1°
2°
3°

Figure 2. Three developmental fates of VPCs. (A) In intact animals, the P5.p, P6.p, and P7.p cells are induced to generate vulval cells. (B) If the P6.p cell is ablated by a laser microbeam, the P5.p or the P7.p cell migrates in to adopt 1° fate. (C) If the P5.p or the P7.p cell is ablated, the P4.p or the P8.p cell can replace these cells. (D) If the P5.p, the P6.p, and the P7.p cells are ablated, the other three VPCs can replace them to generate a normal vulva.

(1)


P3.p
P4.p
P5.p
P6.p
P7.p
P8.p
3°
3°
2°
1°
2°
3°

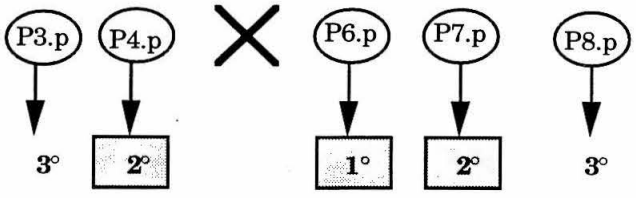
(2)

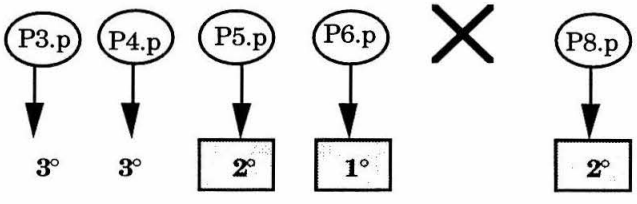

P3.p
P4.p
P5.p
P7.p
P8.p
3°
2°
1°
2°
3°

OR


3°
3°
2°
1°
2°

(3)


P3.p
P4.p
P6.p
P7.p
P8.p
3°
2°
1°
2°
3°


P3.p
P4.p
P5.p
P6.p
P8.p
3°
3°
2°
1°
2°

(4)

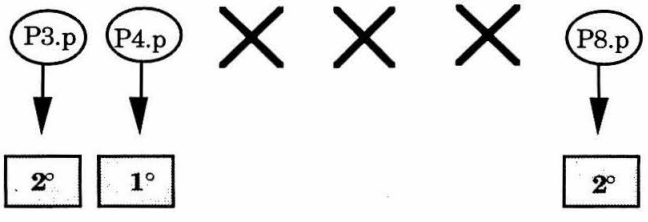
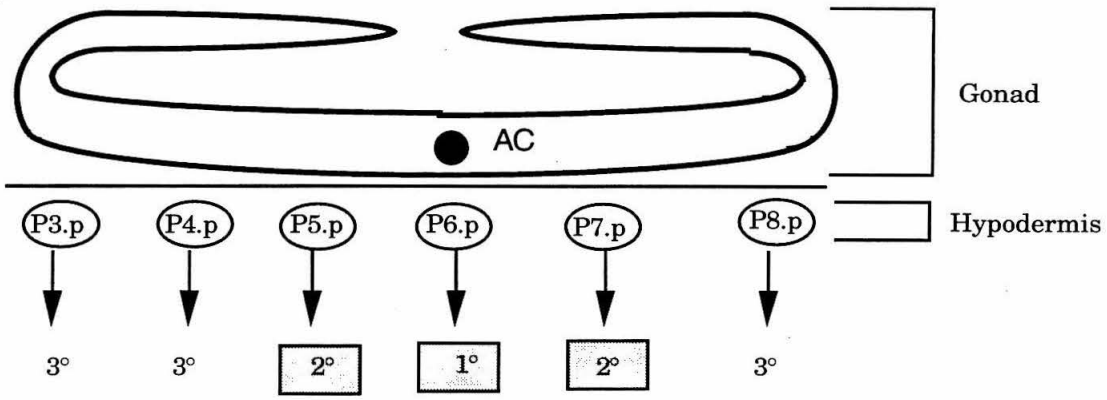
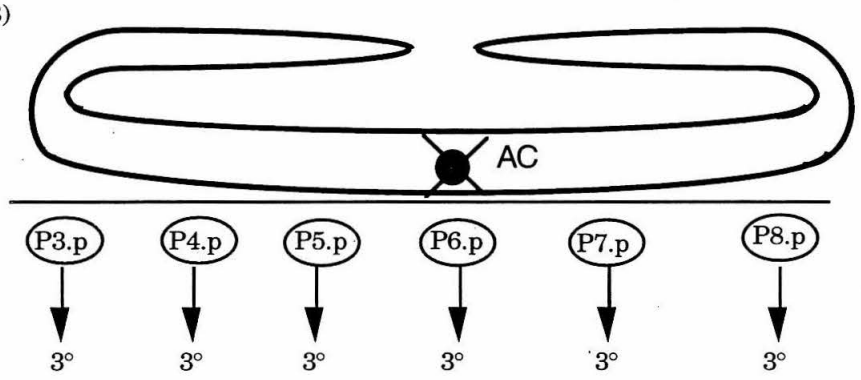

P3.p
P4.p
P8.p
2°
1°
2°

Figure 3. The anchor cell is required and sufficient for vulval induction. (A) The anchor cell induces three VPCs to generate vulval cells. (B) When the anchor cell is ablated, all VPCs adopt non-vulval fates, fusing with hypodermal cells. (C) When the entire gonad but the anchor cell is ablated, VPCs still can be induced.

(A)


Gonad
AC
Hypodermis
P3.p
P4.p
P5.p
P6.p
P7.p
P8.p
3°
3°
2°
1°
2°
3°

(B)


AC
P3.p
P4.p
P5.p
P6.p
P7.p
P8.p
3°
3°
3°
3°
3°
3°

(C)

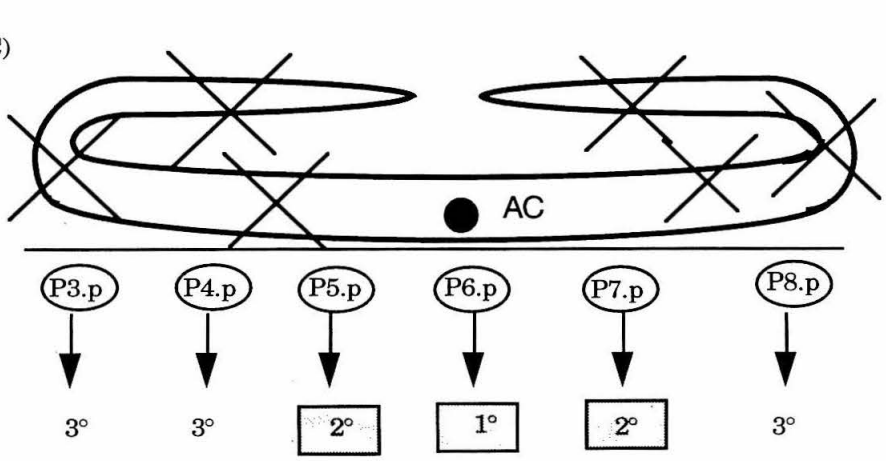
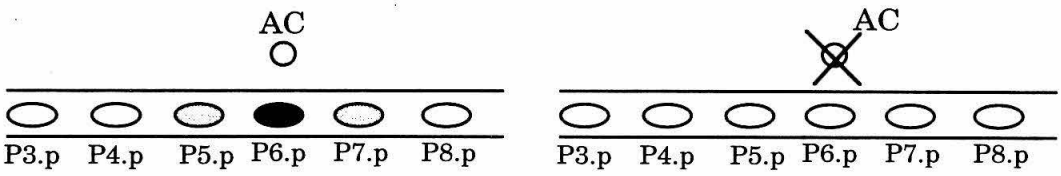
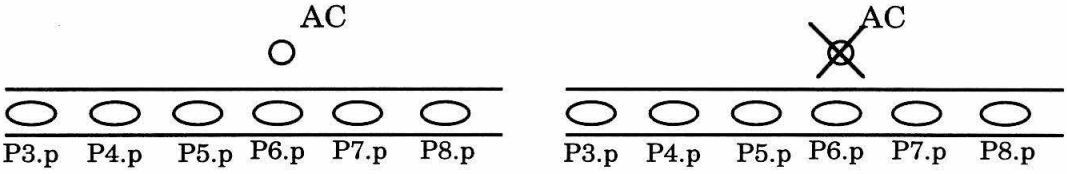

AC
P3.p
P4.p
P5.p
P6.p
P7.p
P8.p
3°
3°
2°
1°
2°
3°

Figure 4. Mutations that disrupt wild-type vulval induction. (A) Wild-type vulval induction. (B) Vulvaless phenotype. Even in the presence of the anchor cell, no VPCs are induced. This phenotype can be caused by loss-of-function mutations in the genes required for signal production or genes responsible for receiving and transducing signals. (C) Multivulva phenotype. Even in the absence of the anchor cell, more than three VPCs are induced. This phenotype can be caused by loss-of-function mutations in the genes required for inhibition or negative regulation of vulval induction or gain-of-function mutations in the genes required for responding to the signal. (D) Hyperinduced phenotype. More than three VPCs are induced. Vulval induction in this phenotype is dependent on the presence of the gonad. This phenotype can be caused by gain-of-function mutations in the genes required for the production of the signal, or loss-of-function mutations in the genes that are involved in negative regulation of the activated responding pathway.

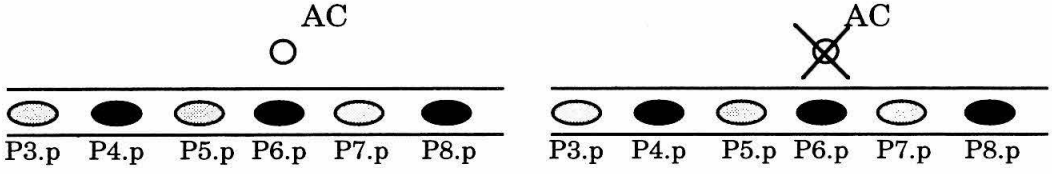
(A) WT



(B) Vulvaless: *lin-2, lin-7, lin-10, let-23, lin-3*



(C) Multivulva: *lin-1, lin-15*



(D) Hyperinduced: *lin-3(gf), let-23(n1045)*

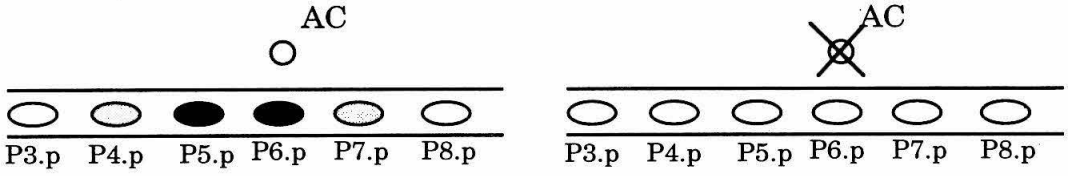


Figure 5. Diagram of genetic pathway of vulval induction. Genes have been placed in this pathway by epistasis analysis and molecular analysis. *lin-3* encodes the signal for the induction. *let-23* encodes a receptor tyrosine kinase that is likely the receptor of LIN-3 signal molecules. *sem-5* encodes an adapter protein with SH2 and SH3 domains. *let-60*, and *lin-45* encode ras and raf, respectively. *lin-15* encodes two new proteins. *unc-101* encodes a clathrin associated protein. *sli-1* encodes a homolog of *c-cbl* proto-oncogene. See text for details.

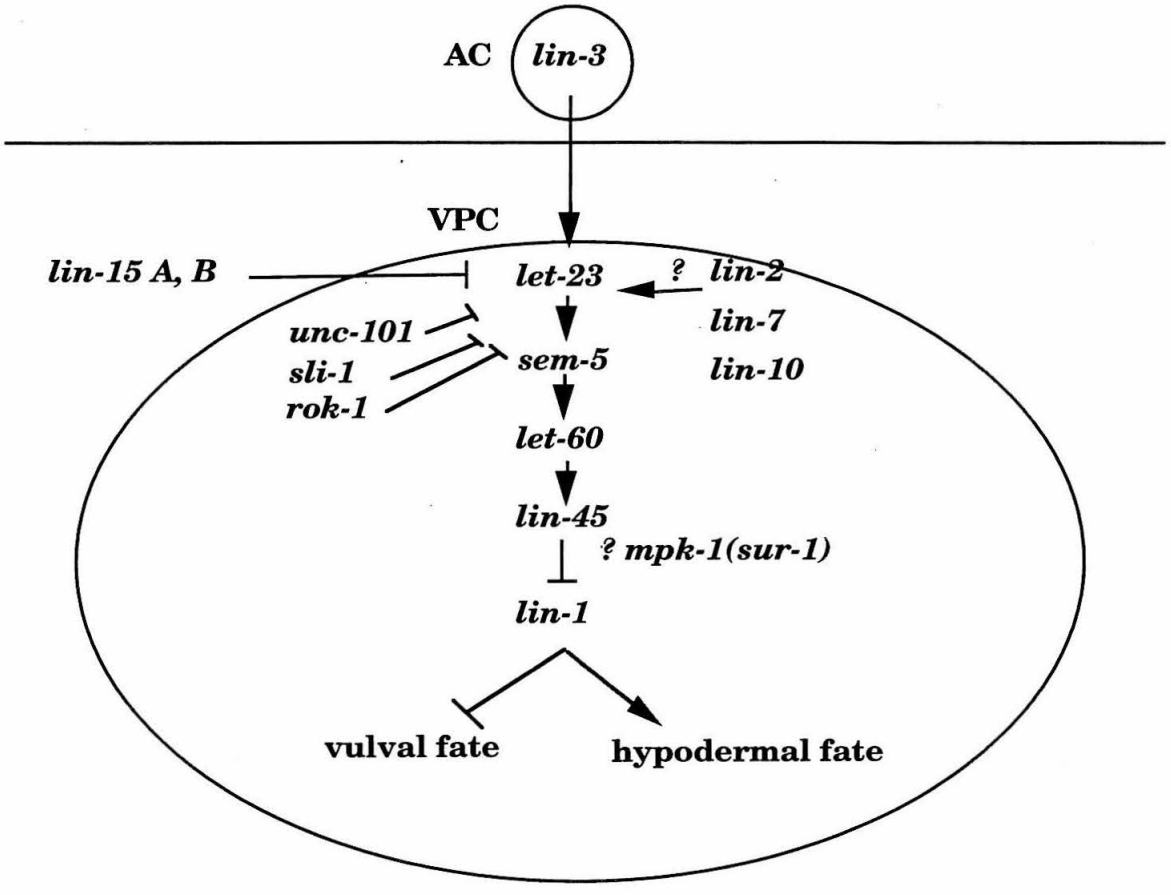

AC
lin-3
VPC
lin-15 A, B
let-23
?
lin-2
lin-7
lin-10
unc-101
sli-1
rok-1
sem-5
let-60
lin-45
? mpk-1(sur-1)
lin-1
vulval fate
hypodermal fate

Figure 6. Structure of clathrin-coated pits.

Clathrin-coated vesicles are composed of membrane, receptors, clathrin triskelions, and the associated protein complex (AP). Clathrin-coated pits are found on the plasma membrane and the *trans*-Golgi membrane.

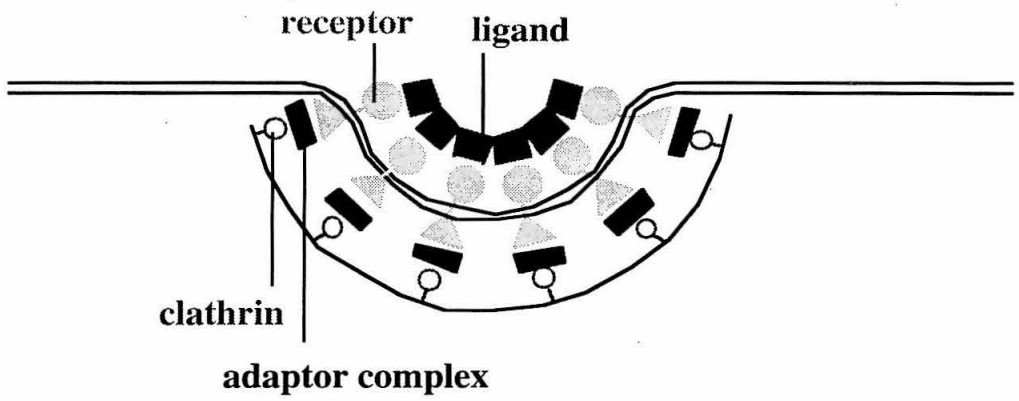
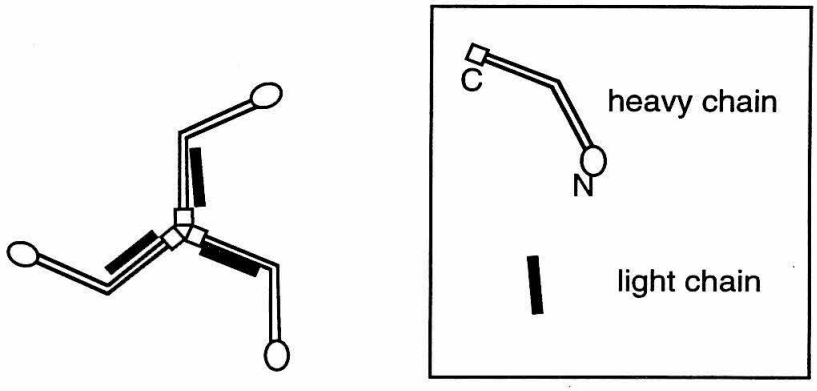

receptor
ligand
clathrin
adaptor complex

Figure 7. Structure of clathrin triskelion. . (A) A triskelion of clathrin is composed of three molecules of clathrin heavy chain and three molecules of light chains. At the center of each triskelion are located the N-termini of the heavy chains. The light chains are bound to the arms of the heavy chains. (B) The triskelions form a cage of triskelions that is a basic structure of the clathrin-coated pits.

(A)



(B)

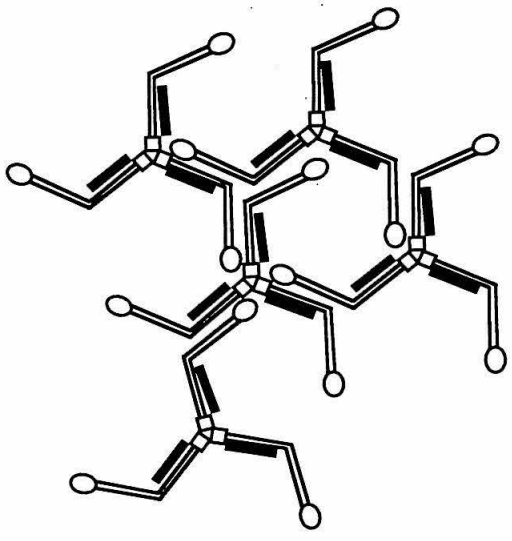
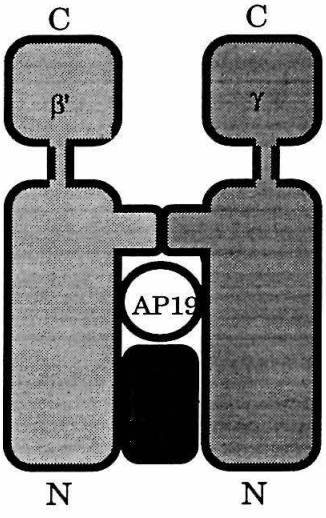
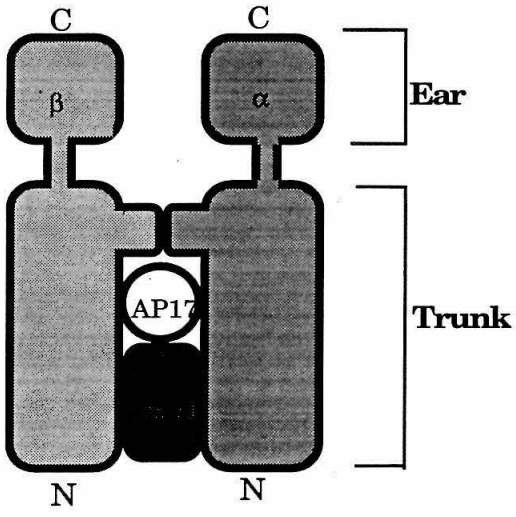


Figure 8. Structure of clathrin-associated protein complex AP-1 and AP-2. AP-1 is the clathrin-associated protein complex located on the Golgi membrane, and AP-2, the clathrin-associated protein complex located on the plasma membrane. Each AP is composed of four subunits: two large chains, one medium chain, and one small chain. The trunk domain of the APs include an intact medium chain, an intact small chain, and the N-terminal regions of the large chains; the ear domains include C-terminal regions of the large chains. The hinge region between these two domains is subject to protease treatment.



AP-1 = *trans*- Golgi form



AP-2 = plasma membrane form

Table 1. Summary of molecular analysis of genes involved in the vulval induction pathway.

Gene	Gene product	Reference
lin-3	protein with EGF motif: a signal	Hill & Sternberg, 1993
let-23	EGFR homolog	Aroian et al., 1990
sem-5	SH2-SH3- SH2 adapter	Clark et al., 1992
let-60	ras homolog	Han & Sternberg, 1990
lin-45	raf homolog	Han et al., 1993
sur-1 (mpk-1)	MAP kinase homolog	Y. Wu &M. Han; Lackner et al., per. comm.
lin-2	GMP kinase with GLGF repeat	R. Hoskins & S. Kim, per. comm.
lin-7	protein w/GLGF repeat	J. Simske & S. kim, per. comm.
lin-10	novel protein	Kim & Horvitz, 1990
lin-15	novel proteins	Huang et al., 1993
sli-1	*c-cbl* homolog	C. Yoon, G. Jongeward, J. Lee, P. Sternberg., unpublished results
unc-101	AP47 homolog	Lee et al., 1993 (Chapter 2)

Reference

Ahle, S., Mann, A., Eichelsbacher, U. and Ungewickell, E. (1988). Structural relationships between clathrin assembly proteins from the Golgi and the plasma membrane *EMBO J.* **7**, 919-929.

Aroian, R. V., Koga, M., Mendel, J. E., Ohshima, Y. and Sternberg, P. W. (1990). The *let-23* gene necessary for *Caenorhabditis elegans* vulval induction encodes a tyrosine kinase of the EGF receptor subfamily *Nature* **348**, 693-699.

Aroian, R. V., Lesa, G. and Sternberg, P. W. (1993). Mutations in the *Caenorhabditis elegans let-23* EGF receptor-like gene define elements important for cell-type specificity and function *In preparation*

Aroian, R. V. and Sternberg, P. W. (1991). Multiple functions of *let-23*, a *C. elegans* receptor tyrosine kinase gene required for vulval induction *Genetics* **128**, 251-267.

Avery, L. and Wasserman, S. (1992). Ordering gene function: the interpretation of epistasis in regulatory hierarchies. *TIG* **8**, 312-316.

Barbacid, M. (1987). *ras* genes. *Annu. Rev. Biochem.* **56**, 779-827.

Bazinet, C., Katzen, A. L., Morgan, M., Mahowald, A. P. and Lemmon, S. K. (1993). The Drosophila clathrin heavy chain gene: Clathrin function is essential in a multicellular organism. *Genetics* **134**, 1119-1134.

Beitel, G., Clark, S. and Horvitz, H. R. (1990). The *Caenorhabditis elegans ras* gene *let-60* acts as a switch in the pathway of vulval induction *Nature* **348**, 503-509.

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

Brodsky, F. M. (1988). Living with clathrin: its role in intracellular

membrane traffic *Science* **242**, 1396-1402.

Buday, L. and Downward, J. (1993). Epidermal growth factor regulates P21(ras) through the formation of a complex of receptor, GRB2 adaptor protein, and Sos nucleotide exchange factor. *Cell* **73**, 611-620.

Carpenter, G. and Cohen, S. (1990). Epidermal growth factor *J. Biol. Chem.* **265**, 7709-7712.

Chang, M. P., Mallet, W. G. and Mostov, K. E. (1993). Adapter self-aggregation, adapter receptor recognition and binding of adapter(AP2) components of clathrin-coated pits *EMBO J.* **12**, 2169-2180.

Chen, W., Lim, H. H. and Lim, L. (1993). The *CDC42* homologue from *Caenorhabditis elegans. J. of Biol. Chem.* **268**, 13280-13285.

Chen, W.-J., Goldstein, J. L. and Brown, M. S. (1990). NPXY, a sequence often found in cytoplasmic tails, is required for coated pit-mediated internalization of the low density lipoprotein receptor *Jour. of Biol. Chem.* **265**, 3116-3123.

Chisholm, A. (1991). Control of cell fate in the tail region of *C. elegans* by the gene *egl-5. Genes and Development* **111**, 921-932.

Clark, S. G., Chisholm, A. D. and Horvitz, H. R. (1993). Control of cell fates in the central body region of *C. elegans* by the homeobox gene *lin-39. Cell* **74**, 43-55.

Clark, S. G., Stern, M. J. and Horvitz, H. R. (1992). *C. elegans* cell-signalling gene *sem-5* encodes a protein with SH2 and SH3 domains. *Nature* **356**, 340-344.

Coulson, A., Kozono, Y., Lutterbach, B., Shownkeen, R., Sulston, J. and Waterson, R. (1991). YACs and the *C. elegans* Genome *BioEssays* **13**, 413-417.

Coulson, A., Waterston, R., Kiff, J., Sulston, J. and Kohara, Y. (1988). Genome linking with yeast artificial chromosomes *Nature* **335**, 184-186.

Coulson, A. R., Sulston, J., Brenner, S. and Karn, J. (1986). Toward a physical map of the genome of the nematode *Caenorhabditis elegans Proc. Natl. Acad. Sci. USA* **83**, 7821-7825.

Downward, J. (1992). Regulatory mechanisms for *ras* proteins *Bioessays* **14**, 177-184.

Egan, S. E., Giddings, B. W., Brooks, M. W. and Buday, L. (1993). Association of Sos ras exchange protein with GRB2 is implicated in tyrosine kinase signal transduction and transformation. *Nature* **363**, 45-51.

Emmons, S. W. (1988). The genome In *The nematode Caenorhabditis elegans* (ed. W. B. Wood), pp. 47-79. Cold Spring Harbor: Cold Spring Harbor Laboratory.

Ferguson, E. and Horvitz, H. R. (1985). Identification and characterization of 22 genes that affect the vulval cell lineages of *Caenorhabditis elegans Genetics* **110**, 17-72.

Ferguson, E. and Horvitz, H. R. (1989). The multivulva phenotype of certain *C. elegans* mutants results from defects in two functionally-redundant pathways *Genetics* **123**, 109-121.

Ferguson, E. L., Sternberg, P. W. and Horvitz, H. R. (1987). A genetic pathway for the specification of the vulval cell lineages of *Caenorhabditis elegans Nature* **326**, 259-267.

Fire, A., Kondo, K. and Waterston, R. (1990). Vectors for low copy transformation of *C. elegans Nucl. Acids Research* **18**, 4269-4270.

Gliekman, J. N., Conibear, E. and Pearse, B. M. (1989). Specificity of binding of clathrin adaptors to siganls on the mannose-6-phosphate/insulin-

like growth factor II receptor. *EMBO J.* **8**, 1041-1047.

Greenwald, I. S. (1985). *lin-12*, a nematode homeotic gene, is homologous to a set of mammalian proteins that includes epidermal growth factor *Cell* **43**, 583-590.

Gurdon, J. B., Mohun, T. J., Taylor, M. V. and Sharpe, C. R. (1989). Embryonic induction and muscle gene activation *TIG* **5**, 51-56.

Han, M., Aroian, R. and Sternberg, P. W. (1990). The *let-60* locus controls the switch between vulval and non-vulval cell types in *C. elegans Genetics* **126**, 899-913.

Han, M., Golden, A., Han, Y. and Sternberg, P. W. (1993). *C. elegans lin-45 raf* gene participates in *let-60 ras* stimulated vulval differentiation. *Nature* **363**, 133-140.

Han, M. and Sternberg, P. W. (1990). *let-60*, a gene that specifies cell fates during C. elegans vulval induction, encodes a ras protein *Cell* **63**, 921-931.

Hart, A. C., Kramer, H., Van Vactor, D. L., Paidhungat, M. and Zipursky, S. L. (1990). Induction of cell fate in the *Drosophila* retina - the *bride of sevenless* protein is predicted to contain a large extracellular domain and 7 transmembrane segments *Genes and Dev.* **4**, 1835-1847.

Heidecker, G., Kolch, W., Morrison, D. K. and Rapp, U. R. (1992). The role of Raf-1 phosphorylation in signal transduction *Adv. Cancer Res.* **58**, 53-73.

Herman, R. K. (1984). Analysis of genetic mosaics of the nematode *Caenorhabditis elegans Genetics* **108**, 165-180.

Herman, R. K. (1987). Mosaic analysis of two genes that affect nervous system structure in *Caenorhabditis elegans Genetics* **116**, 377-388.

Herman, R. K. (1989). Mosaic analysis in the nematode *Caenorhabditis*

elegans J. Neurogenetics **5**, 1-24.

Herman, R. K. and Hedgecock, E. M. (1990). The size of the *C. elegans* vulval primordium is limited by *lin-15* expression in surrounding hypodermis *Nature* **348**, 169-171.

Herman, R. K., Madl, J. E. and Kari, C. K. (1979). Duplications in *Caenorhabditis elegans Genetics* **92**, 419-435.

Hill, R. J. and Sternberg, P. W. (1992). The *lin-3* gene encodes an inductive signal for vulval development in *C. elegans Nature* **358**, 470-476.

Hodgkin, J., Doniach, T. and Shen, M. (1985). The sex determination pathway in the nematode *Caenorhabditis elegans*: Variations on a theme *Cold Spring Harbor Symp. Quant. Biol.* **50**, 585-594.

Hodgkin, J., Kondo, K. and Waterston, R. H. (1987). Suppression in the nematode *Caenorhabditis elegans Trends Genet.* **3**, 325-329.

Horvitz, H. R. and Sternberg, P. W. (1991). Multiple intercellular signalling systems control the development of the *C. elegans* vulva *Nature* **351**, 535-541.

Horvitz, H. R. and Sulston, J. E. (1980). Isolation and genetic characterization of cell-lineage mutants of the nematode *Caenorhabditis elegans Genetics* **96**, 435-454.

Howe, L. R., Leevers, S. J. G., N., Nakielny, S., Cohen, P. and Marshall, C. J. (1992). Activation of the MAP kinase pathway by the protein kinase *raf. Cell* **71**, 335-342.

Huang, L. S., Tzou, P. and Sternberg, P. W. (1993). The lin-15 locus of C. elegans... In preparation.

Jackson, A. P., Seow, H. F., Holmes, N., Drikamer, K. and Parham, P. (1987). Localization of clathrin light-chain sequences mediating heavy-chain

binding and coated vesicle diversity *Nature* **326**, 154-159.

Jessell, T. M. and Melton, D. A. (1992). Diffusible factors in vertebrate embryonic induction *Cell* **68**, 257-270.

Johnsen, R. C. and Baillie, D. L. (1991). Genetic analysis of a major segment [LGV(left)] of the genome of *Caenorhabditis elegans Genetics* **129**, 735-752.

Johnson, K. F., Chan, W. and Kornfeld, S. (1990). Cation-dependent mannose-6-phosphate receptor contains two internalization signals in its cytoplasmic domain. *Proc. Natl. Acad. Sci. USA* **87**, 10010-10014.

Jongeward, G. D. and Sternberg, P. W. (1993). The sli-1 locus of C. elegans *In preparation*

Keen, J. H. (1990). Clathrin and associated assembly and disassembly proteins *Annu. Rev. Biochem.* **59**, 415-438.

Kenyon, C. and Wang, B. (1991). A cluster of Antennapedia-class homeobox genes in a nonsegmented animal *Science* **253**, 516-517.

Kim, S. K. and Horvitz, H. R. (1990). The *Caenorhabditis elegans* gene *lin-10* is broadly expressed while required specifically for the determination of vulval cell fates *Genes & Devel.* **4**, 357-371.

Kimble, J. (1981). Lineage alterations after ablation of cells in the somatic gonad of *Caenorhabditis elegans Dev. Biol.* **87**, 286-300.

Kirchhausen, T. (1990). Identification of putative yeast homolog of the mammalian β chains of the clathrin-associated protein complexes *Mol. Cell. Biol.*

Kirchhausen, T., Davis, A. C., Frucht, S., Greco, B. O., Payne, G. S. and Tubb, B. (1991). AP17 and AP19, the mammalian small chains of the clathrin-associated protein complexes show homology to YAP17p, their

putative homolog in yeast *J. Biol. Chem* **266**, 1153-1157.

Kirchhausen, T., Harrison, S. C., Chow, E. P., Mattaliano, R. J., Ramachandran, K. L., Smart, J. and Brosius, J. (1987). Clathrin heavy-chain-molecular cloning and complete primary structure. *Proc. Natl. Acad. Sci. USA* **84**, 8805-8809.

Kirchhausen, T., Nathanson, K. L., Matsui, W., Vaisberg, A., Chow, E. P., Burne, Z., Keen, J. H. and Davies, A. E. (1989). Structural and functional division into two domains of the large (100- to 115kDa) chains of the clathrin-associated protein complex AP-2. *Proc. Natl. Acad. Sci. USA* **86**, 2612-2616.

Kirchhausen, T., Scarmoto, P., Harrison, S. C., Monroe, J. J., Chow, E. P., Mattaliano, R. J., Ramachandran, K. L. and al., e. (1987). Clathrin light-chains LCa and LCb are similar, polymorphic, and share repeated heptad motifs *Science* **236**, 320-324.

Kornfeld, S. (1990). Lysosomal enzyme targeting *Biochem. Soc. Trans.* **18**, 367-374.

Kyriakis, J. M., App, H., Zhang, X.-f., Banerjee, P., Brautigan, D. L., Rapp, U. R. and Avruch, J. (1992). Raf-1 activates MAP kinase-kinase. *Nature* **358**, 417-421.

Lee, J., Jongeward, G. D. and Sternberg, P. W. (1993). *unc-101*, a gene required for many aspects of *C. elegans* development and behavior, encodes a clathrin-associated protein *In preparation*

Lemmon, S. K., Freund, C., Conley, K. and Jones, E. W. (1990). Genetic instability of clathrin-deficient strains of *Saccharomyces cerevisiae Genetics* **124**, 27-38.

Li, N., Batzer, A., Daly, R. and Yajnik, V. (1993). Guanine-nucleotide-

releasing factor hSOS1 binds to GRB2 and links receptor tyrosine kinases to ras signaling. *Nature* **363**, 85-88.

Lindner, R. and Ungewickell, E. (1992). Clathrin-associated proteins of bovine brain coated vesicles - an analysis of their number and assembly-promoting activity *J. Biol. Chem.* **267**, 16567-16573.

Lochrie, M. A., Mendel, J. E., Sternberg, P. W. and Simon, M. I. (1991). Homologous and unique G protein alpha subunits in the nematode *Caenorhabditis elegans Cell Regul.* **2**, 135-154.

Lowy, D. R. and Willumsen, B. M. (1993). Function and regulation of ras *Annu. Rev. Biochem.* **62**, 851-891.

Margolis, B., Li, N., Koch, A., Mohammadi, M., Hurwitz, D. R., Zilberstein, A., Ullrich, A., Pawson, T. and Schlessinger, J. (1990). The tyrosine phosphorylated carboxyterminus of the EGF receptor is a binding site for GAP and PLC-g *The EMBNO Journal* **9**, 4375-4380.

Matsui, W. and Kirchhausen, T. (1990). Stabilization of clathrin coats by the core of the clathrin-associated protein complex AP-2 *Biochemistry* **29**, 10791-10798.

Mello, C. C., Kramer, J. M., Stinchcomb, D. and Ambros, V. (1991). Efficient gene transfer in *C. elegans* after microinjection of DNA into germline cytoplasm: extrachromosomal maintenance and integration of transforming sequences *EMBO J.* **10**, 3959-3970.

Milburn, M. V., Tong, L., DeVos, A. M., Brünger, A., Yamaizumi, Z., Nishimura, S. and Kim, S.-H. (1990). Molecular switch for signal transduction: Structural differences between active and inactive forms of protooncogenic *ras* proteins *Science* **247**, 939-945.

Miller, L. M., Gallegos, M. E., Morisseau, B. A. and Kim, S. A. (1993).

lin-31, a *Caenorhabditis elegans* HNF-3/*fork head* transcription factor homolog, specifies three alternative cell fates in vulval development *Genes & Dev.* **7**, 933-947.

Moerman, D. G. and Baillie, D. L. (1981). Formaldehyde mutagenesis in the nematode *Caenorhabditis elegans Mutat. Res.* **80**, 273-279.

Moerman, D. G. and Waterston, R. H. (1984). Spontaneous unstable *unc-22 IV* mutations in *Caenorhabditis elegans* var. Bergerac *Genetics* **108**, 859-877.

Morris, S. A., Schroder, S., Plessmann, U., Weber, K. and Ungewickell, E. (1993). Clathrin assembly protein AP180: primary structure, domain organization and identification of a clathrin binding site *EMBO J.* **12**, 667-675.

Munn, A. L., Silveira, L., Elgort, M. and Payne, G. S. (1991). Viability of clathrin heavy-chain-deficient *Saccharomyces cerevisiae* is compromised by mutations at numerous loci: implications for the suppression hypothesis *Mol. Cell Biol.* **11**, 3868-3878.

Murphy, J. E., Pleasure, I. T., Puszkin, S., Prasad, K. and Keen, J. H. (1991). Clathrin assembly protein AP-3 - the identity of the 155K protein, AP-180, and NP185 and demonstration of a clathrin binding domain *J. Biol. Chem.* **266**, 4401-4408.

Myers, M. and Forgac, M. (1993). The coated vesicle vacuolar (H+)-ATPase associates with and is phosphorylated by the 50 KDa polypeptide of the clathrin assembly protein AP-2. *J. Biol. Chem.* **268**, 9184-9186.

Nakayama, Y., Goebl, M., Greco, B. O., Lemmon, S., Chow, E. P. C. and Kirchhausen, T. (1991). The medium chains of the mammalian clathrin-associated proteins have a homolog in yeast. *Eur. J. Biochem.* **202**, 569-574.

Ochman, H., Gerber, A. S. and Hartl, D. L. (1988). Genetic applications of an inverse polymerase chain reaction *Genetics* **120**, 621-623.

Ohalloran, T. J. and Anderson, R. G. W. (1992). Clathrin heavy chain is required for pinocytosis, the presence of large vacuoles, and development in Dictyostelium. *J. Cell Biol.* **118**, 1371-1377.

Oliver, J. P., Raabe, T., Henkemeyer, M., Dickson, B., Mbamalu, G., Margolis, B., Schlessinger, J., Hafen, E. and Pawson, T. (1993). A *Drosophila* SH2-SH3 adaptor protein implicated in coupling the sevenless tyrosine kinase to an activator of *Ras* guanine nucleotide exchange, Sos. *Cell* **73**, 179-191.

Parsons, J. T. (1990). Closing the GAP in a signal transduction pathway *Trends in Genetics* **6**, 169-171.

Pauloin, C., Bernier,I., Jolles,P. (1982). Presence of cyclic nucleotide Ca++-independent protein kinase in bovine brain coated vesicles. *Nature* **298**, 574-576.

Payne, G. S. and Schekman, R. (1985). A test of clathrin function in protein secretion and cell-growth. *Science* **230**, 1009-1014.

Pearse, B. M. F. and Robinson, M. S. (1990). Clathrin, adaptors, and sorting *Annu. Rev. Biol.* **6**, 151-171.

Peeler, J. S., Donzell, W. C. and Anderson, R. G. W. (1993). The appendage domain of the AP-2 subunit is not required for assembly or invagination of clathrin-coated pits *J. Cell Biol.* **120**, 47-54.

Pelech, S. L. and Sanghera, J. S. (1992). MAP Kinases: Charting the regulatory pathways *Science* **257**, 1355-1356.

Prywes, R., Livneh, E., Ullrich, A. and Schlessinger, J. (1986). Mutations in the cytoplasmic domain of EGF receptor affect EGF binding and

receptor internalization. *EMBO J.* **5**, 2179-2190.

Riedel, H., Dull, T. J., Honegger, A. M., Schlessinger, J. and Ullrich, A. (1989). Cytoplasmic domains determine signal specificity, cellular routing characteristics and influence ligand binding of epidermal growth factor and insulin receptors *The EMBO Journal* **8**, 2943-2954.

Robinson, M. S. (1989). Cloning of cDNAs encoding two related 100-kD coated vesicle proteins (α-adaptins) *J. Cell Biol.* **108**, 833-842.

Robinson, M. S. (1990). Cloning and expression of γ-adaptin, a component of clathrin-coated vesicles associated with the Golgi apparatus. *J. Cell Biol.* **111**, 2319-2326.

Rogalski, T. M., Moerman, D. G. and Baillie, D. L. (1982). Essential genes and deficiencies in the *unc-22* IV region of *Caenorhabditis elegans* *Genetics* **102**, 725-736.

Rosenbluth, R. E., Cuddeford, C. and Baillie, D. L. (1985). Mutagenesis in Caenorhabditis elegans: II A spectrum of mutationa events induced with 1500 R of g-radiation *Genetics* **109**, 493-511.

Rozakisadcock, M., Fernley, R., Wade, J., Pawson, T. and Bowtell, D. (1993). The SH2 and SH3 domains of mammalian GRB2 couple the EGF receptor to the ras activator mSOS1. *Nature* **363**, 83-85.

Ruvkun, G., Ambros, V., Coulson, A., Waterston, R., Sulston, J. and Horvitz, H. R. (1989). Molecular genetics of the *Caenorhabditis elegans* heterochronic gene *lin-14 Genetics* **121**, 501-516.

Schedin, P., Hunter, C. and Wood, W. B. (1991). Autonomy and nonautonomy of sex determination in triploid intersex mosaics of *C. elegans* *Development* **112**, 863-879.

Schlessinger, J. and Ullrich, A. (1992). Growth factor signaling by receptor

tyrosine kinases. *Neuron* **9**, 383-391.

Schmid, S. L. (1992). The mechanism of receptor-mediated endocytosis - more questions than answers. *Bioessays* **14**, 589-596.

Schroder, S. and Ungewickell, E. (1991). Subunit interaction and function of clathrin-coated vesicle adaptors from the Golgi and the plasma membrane. *J. Biol. Chem.* **266**, 7910-7918.

Seeger, M. and Payne, G. S. (1992). A role for clathrin in the sorting of vacuolar proteins in the Golgi-complex of yeast. *EMBO. J.* **11**, 2811-1818.

Seeger, M. and Payne, G. S. (1992). Selective and immediate effects of clathrin heavy-chain mutations on Golgi membrane-protein retention in Saccharomyces-cerevisiae *J. Cell. Biol.* **118**, 531-540.

Shackleford, G. M., Shivakumar, S., Shiue, L., Mason, J., Kenyon, C. and Varmus, H. E. (1993). Two *wnt* genes in *Caenorhabditis elegans. Oncogene* **8**, 1857-1864.

Siddiqui, S. S. and Babu, P. (1980). Genetic mosaics of *Caenorhabditis elegans*: A tissue-specific fluorescent mutant *Science* **210**, 330-332.

Sigurdson, D. C., Spanier, G. J. and Herman, R. K. (1984). *Caenorhabditis elegans* deficiency mapping *Genetics* **108**, 331-345.

Silveira, L. A., Wong, D. H., Masiarz, F. R. and Schekman, R. (1990). Yeast clathrin has a distinct light chain that is important for cell-growth. *J. Cell Biol.* **111**, 1437-1449.

Simon, M. A., Bowtell, D. D. L., Dodson, G. S., Laverty, T. R. and Rubin, G. M. (1991). Ras1 and a putative guanine nucleotide exchange factor perform crucial steps in signaling by the sevenless protein tyrosine kinase *Cell* **67**, 701-716.

Simon, M. A., Dodson, G. S. and Rubin, G. M. (1993). An SH3-SH2-SH3

protein is required for p21ras1 activation and binds to Sevenless and Sos proteins in vitro *Cell* **73**, 169-177.

Smith, J. C. (1989). Mesoderm induction and mesoderm-inducing factors in early amphibian development. *Dev.* **105**, 665-677.

Sorkin, A. and Carpenter, G. (1993). Interaction of activated EGF receptors with coated pit adaptins *Science* **261**, 612-615.

Sorkin, A., Helin, K., Waters, C. M., Carpenter, G. and Beguinot, L. (1992). Multiple autophosphorylation sites of the epidermal growth factor receptor are essential for receptor kinase activity and internalization *J. Biol. Chem.* **267**, 8672-8678.

Sosa, M. A., Schmidt, B. and Vonfigura, K. (1993). *In vitro* binding of plasma membrane-coated vesicle adaptors to the cytoplasmic domain of lysosomal acid-phosphatase *J. Biol. Chem.* **268**, 12537-12543.

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

Sternberg, P. W. and Horvitz, H. R. (1986). Pattern formation during vulval development in *Caenorhabditis elegans Cell* **44**, 761-772.

Sternberg, P. W. and Horvitz, H. R. (1989). The combined action of two intercellular signalling pathways specifies three cell fates during vulval induction in *C. elegans Cell* **58**, 679-693.

Sternberg, P. W. and Horvitz, H. R. (1991). Signal transduction during *C. elegans* vulval induction *Trends Genet.* **7**, 366-371.

Stewart, H. I., Rosenbluth, R. E. and Baillie, D. L. (1991). Most ultraviolet irradiation induced mutations in the nematode Caenorhabditis elegans are chromosomal rearrangements. *Mutation Research* **249**, 37-54.

Sulston, J. E. and Horvitz, H. R. (1981). Abnormal cell lineages in mutants of the nematode *Caenorhabditis elegans Dev. Biol.* **82**, 41-55.

Sulston, J. E. and White, J. G. (1980). Regulation and cell autonomy during postembryonic development of *Caenorhabditis elegans Devel. Biol.* **78**, 577-597.

Thomas, J. H. (1990). Genetic analysis of defecation in *Caenorhabditis elegans Genetics* **124**, 855-872.

Thurieau, C., Brosius, J., Burne, C., Jolles, P., Keen, J., Mattaliano, R. J., Chow, E. P., Ramachandran, K. and Kirchhausen, T. (1988). Molecular cloning of complete amino acid sequence of AP50, an assembly protein associated with clathrin-coated vesicles *DNA* **7**, 663-669.

Troppmair, J., Bruder, J. T., App, H., Cai, H., Liptak, L., Szeberenyi, J., Cooper, G. M. and Rapp, U. R. (1992). Ras controls coupling of growth factor receptors and protein kinase C in the membrane to Raf-1 and B-Raf protein serine kinases in the cytosol *Oncogene* **7**, 1867-1873.

Waterston, R., Martin, C., Craxton, M., Huynh, C., Coulson, A., Hillier, L., Durbin, R., Green, P., Shownkeen, R., Halloran, N., Metzstein, M., Hawkins, T., Wilson, R., Berks, M., Du, Z., Thomas, K., Thierry-Mieg, J. and Sulston, J. (1992). A survey of expressed genes in *Caenorhabditis elegans nature genetics* **1**, 114-123.

Yochem, J., Weston, K. and Greenwald, I. (1988). *C. elegans lin-12* encodes a transmembrane protein similar to *Drosophila Notch* and yeast cell cycle gene products *Nature* **335**, 547-550.

Wood, W. B. (1988). The nematode *Caenorhabditis elegans* Cold Spring Harbor Laboratory.

Zhou, S., Tannery, N. H., Yang, J., Puszkin, S. and Lafer, E. M. (1993).

The synapse-specific phosphoprotein F1-20 is identical to clathrin assembly protein AP-3. *J. Biol. Chem.* **268**, 12655-12662.

Zipursky, S. L., Krämer, H., Cagan, R., Hart, A. and Van Vactor, D. L., Jr. (1992). Induction of the R7 neuron in the *Drosophila* compound eye: the *bride of sevenless* and *sevenless* interaction *Cold Spring Harbor Symp Quant Biol* **57**, 381-389.

CHAPTER 2. ANALYSIS OF *UNC-101*

unc-101, a gene required for many aspects of *C. elegans* development and behavior, encodes a clathrin-associated protein.

Junho Lee*, Gregg D. Jongeward*, and Paul W. Sternberg

HHMI and Division of Biology
California Institute of Technology
Pasadena, CA 91106

Running head: *C. elegans unc-101*, a clathrin-associated protein

Key words: *C. elegans, unc-101*, vulval induction, clathrin associated protein, AP47.

Editorial correspondence:
Paul W. Sternberg
HHMI and Division of Biology
California Institute of Technology
Pasadena, CA 91125
(818) 395-2181
Fax: (818) 568-8012

* J. Lee & G. Jongeward contributed equally to this paper.

Abstract

Our genetic analysis indicates that the *unc-101* gene of the nematode *C. elegans* is required for many aspects of development and behavior, including negative regulation of vulval induction. We have cloned *unc-101* and found that it encodes a homolog of the mammalian AP47 and AP50 proteins, medium chains of clathrin-associated protein complexes located at the *trans*-Golgi and the plasma membrane, respectively. Therefore, clathrin-mediated events contribute to the negative regulation of vulval differentiation. Comparison of sequences including a full length sequence of a *C. elegans* AP50 homolog reveals that UNC-101 is most closely related to AP47. Mouse AP47 and nematode UNC-101 proteins are functionally equivalent as assayed in transgenic nematodes. We have determined mutations of *unc-101* alleles from various genetic screens and shown that all but one allele are deletions or nonsense mutations, suggesting that these alleles severely reduce *unc-101* function.

Introduction

Clathrin coated pits and coated vesicles are organelles that originate from the plasma membrane and the *trans*-Golgi in eukaryotic cells and mediate intracellular trafficking of membrane proteins (Figure 1A; reviewed in Brodsky, 1988; Keen, 1990; Pearse & Robinson, 1990; Schmid, 1992). The main protein components in the coated vesicles are clathrin and their associated protein (AP) complexes. While clathrin triskelions are common structural units to both the plasma membrane and the *trans*-Golgi coated vesicles, the AP complexes differ in these compartments, probably conferring specific functions to differentially located clathrin vesicles. The *trans*-Golgi

associated protein complex AP-1 consists of two large chains, β' and γ, one medium chain, AP47, and one small chain, AP19 (Ahle et al., 1988; Keen, 1987; Matsui and Kirchhausen, 1990; Figure 1B). The plasma membrane associated protein complex AP-2 consists of two large chains, α and β, one medium chain, AP50, and one small chain, AP17. While the large chains β and β' are similar in sequence (Kirchhausen, 1989), the large chains α and γ have a more diverged primary structure. Both medium chains and small chains are homologous to their counterparts.

The coated vesicles of the plasma membrane are involved in the endocytosis of membrane proteins such as LDL receptor, transferrin receptor, and EGF receptor, while *trans*-Golgi coated vesicles are involved in sorting of proteins such as lysosomal enzymes, and are also thought to be involved in regulated secretion (for example, Keen, 1990; Pearse & Robinson, 1990; Sorkin & Carpenter, 1993).

Genetic analysis of a clathrin heavy chain in *Saccharomyces cerevisiae* showed that mutants deficient in a clathrin heavy chain are defective in the retention of the endonuclease Kex2p protein in the Golgi apparatus (Seeger and Payne, 1992). However, deletion of yeast homologs of the small chains of the APs has little effect on cell growth, protein export, or endocytosis (H. Phan & G. Payne, personal comm.). Similarly, the deletion of a yeast homolog of medium chains does not result in any obvious phenotypic consequences, although there might be subtle phenotypes associated with these mutations (S.K. Lemmon, personal comm.). Since yeast is a single-cell organism and yeast clathrin vesicles and their APs may have different functions than those in multicellular organisms, it is useful to examine the functions of coated vesicles in a multicellular organism. Indeed, Bazinet et

al. (1993) have recently found that clathrin heavy chain is essential in Drosophila.

The *unc-101* mutants of *C. elegans* were originally identified by virtue of their uncoordinated movement (D. Riddle, personal comm.), and by their abnormal uptake of dye by sensory neurons (E. Hedgecock, personal comm.). We first identified *unc-101* alleles as extragenic suppressors of a reduction-of-function mutation of *let-23*, an epidermal growth factor receptor-like tyrosine kinase (G. Jongeward and P. Sternberg, in prep.). Here, we determine the loss-of-function phenotype of *unc-101*, and provide evidence for its role in regulating vulval induction. We describe the cloning of *unc-101*, its sequence, and its functional equivalence with the mouse AP47 protein. *C. elegans* thus provides an opportunity to use molecular genetics to study clathrin-associated protein complexes in a multicellular organism.

Results

Pleiotropic effects of *unc-101* mutations

unc-101 mutations have pleiotropic effects on the behavior and development of *C. elegans*. Homozygous *unc-101* animals have uncoordinated movement. They are very sluggish, do not respond to a light touch, and tend to coil. *unc-101* animals also display abnormal uptake of the dye fluorescein-isothiocyanate (FITC) (E. Hedgecock, personal comm.). In wild-type animals six pairs of neurons in the amphid and two pairs of neurons in the phasmid are filled with FITC (Hedgecock et al., 1985). In *unc-101* animals, one pair of amphid neurons stains weakly, and the phasmid neurons are generally faint or unstained. *unc-101* mutant animals show irregularity in their defecation

cycle (Thomas, 1990). We have also isolated *unc-101* mutations as suppressors of the vulvaless phenotype of *let-23(sy1)* hemizygotes (G. Jongeward and P. Sternberg, in prep.). *unc-101* animals also have male tail defects such as abnormal ray and spicule structure (S. Emmons, H. Chamberlin, personal comms.). In addition, *unc-101* animals are subviable, as about half of *unc-101* animals do not survive to reach adulthood. These inviable animals arrest in late L1 stage, and in few cases in the L2, L3, or L4 stages.

Isolation of additional alleles of *unc-101*

The *unc-101* alleles *m1, rh6, sy108*, and *sy161* had previously been recovered from independent F2 screens for the different phenotypes of *unc-101* mutations. Although they were recovered in screens for different phenotypes, animals homozygous for these alleles display essentially identical phenotypes to those described above. Since the F2 screens could have failed to recover null alleles of *unc-101* had the null phenotype been lethal, we performed several genetic screens that could recover null alleles of *unc-101* (Figure 2).

We first screened for new alleles that failed to complement the Unc phenotype of *unc-101(sy108)* (non-complementation screen). Null alleles can be isolated in this screen because some animals carrying *unc-101(sy108)* in *trans* to a deletion of the *unc-101* locus are viable. Two alleles, *sy168* and *sy169*, were recovered in a screen of 15,000 EMS-mutagenized F1 gametes. The phenotypes animals homozygous for these two alleles were essentially identical to those of other alleles. As described below, *sy168* and *sy169* are likely either products of gene conversion events of *sy108* or the result of

recombination before recovery.

We recovered *sy216* by virtue of its failure to complement *unc-101(rh6)* in a similar non-complementation screen of 11,000 trimethylpsoralen (TMP)-mutagenized gametes. Animals homozygous for *sy216* arrest immediately after hatching without any apparent post embryonic divisions and live for several days before finally dying. This arrest phenotype is different from that associated with other *unc-101* alleles, since dying homozygotes of other alleles arrest and die rapidly. The lethality of *unc-101(sy216) / unc-101(sy108)* *trans*-heterozygotes is slightly enhanced. However, other phenotypes such as suppression of the *let-23(sy1)* mutation are not enhanced (see Methods). Based on the genetic and molecular analysis of *sy216,* we believe that *sy216* is a deletion of the entire *unc-101* locus and some essential gene(s) nearby (see Methods).

The non-complementation screens described above could have missed some of the new mutations because about half of the progeny from homozygous *unc-101* mothers die. We therefore designed another screen to avoid loss of new alleles due to subviability by providing a wild-type maternal copy of *unc-101* (Fig. 2c). All animals of the genotype *unc-101(sy108) / unc-101(sy216)* from an *unc-101(sy108) / +* mother are viable. Therefore, the maternal copy of *unc-101* (+) is sufficient to rescue the inviability of any new allele in *trans* to the visible allele *sy108.* We recovered two EMS induced alleles, *sy241* and *sy242,* from a screen of 20,000 F1 mutagenized gametes. *sy241* animals are viable and have phenotypes similar to those of animals of previously identified alleles. *sy242* was lost before it could be analyzed.

We conclude that the phenotypes represented by homozygotes of viable alleles are those associated with strong reduction-of function or null alleles.

We base this conclusion on two genetic arguments. First, the frequency of recovery, 1/7,500 ~ 1/10,000, is close to that of loss-of-function alleles (1/2,000 ~ 1/5,000, e.g., Brenner, 1974; Greenwald & Horvitz, 1980). Second, homozygotes for these alleles are phenotypically indistinguishable. The exception to the second criterion is *sy216*, which is a deletion of *unc-101* locus and adjacent gene(s). As discussed below, our molecular analysis is consistent with this argument.

***unc-101* is a negative regulator of the vulval induction pathway.**

In *C. elegans*, a signal from the anchor cell (AC) of the somatic gonad induces three out of six vulval precursor cells (VPCs) to generate vulval tissue. Proper vulval development requires genes that mediate the induction of vulval cells and also that prevent excessive induction. Many genes required for the production of vulval tissue have been genetically and molecularly characterized and shown to encode homologs of mammalian signaling molecules (*lin-3, let-23, sem-5, let-60,* and *lin-45;* reviewed by Horvitz and Sternberg, 1991). Reduction-of-function mutations in these genes result in a vulvaless (Vul) phenotype. Other genes (*lin-2, lin-7,* and *lin-10*) are also required for vulval induction. There are a number of genes that act as negative regulators of the induction pathway. For example, loss-of-function mutations of *lin-15* cause a gonad-independent multivulva phenotype (Ferguson and Horvitz, 1985; Ferguson et al., 1987; L. Huang, P. Tzou, P. Sternberg, unpublished results). *lin-15(+)* activity from cells other than anchor cell or VPCs antagonizes the inductive signal (Herman and Hedgecock, 1990).

We recovered two alleles of *unc-101* (*sy108* and *sy161*) as suppressors

of the Vulvaless (Vul) phenotype of the weak *let-23* allele, *sy1*. *let-23* is a *C. elegans* homolog of the epidermal growth factor (EGF) receptor tyrosine kinase that is involved in the development of several structures including the hermaphrodite vulva (Aroian et al., 1990). While mutations at the *unc-101* locus confer no defect in the extent of vulval differentiation in the absence of another mutation, averaging three VPCs forming vulval tissue per animal, several non-null mutations of *let-23* are suppressed by an *unc-101* mutation (Table 1). *let-23(sy1)*, the allele used in the original screen, is suppressed strongly. *let-23(sy1)* animals average one VPC per animal forming vulval tissue, whereas *unc-101; let-23(sy1)* animals are hyperresponsive to the signal, averaging 3.6 VPCs forming vulval tissue per animal. After ablation of the gonad of the *unc-101(sy108); let-23(sy1)* double mutant animals, no animal displayed vulval differentiation (n=7), suggesting that the vulval differentiation of the double mutants is still gonad-dependent and require the signal from the gonad as in *unc-101* mutant animals (n= 6) or in wild-type animals.

A more severe, but not a null, mutation, *let-23(sy12)* is suppressed from an average of less than one to nearly three VPCs per animal forming vulval tissue. Another severe allele, *let-23(sy97)*, is suppressed but only to a very slight extent. In contrast to suppression of these alleles, the allele *let-23(n1045)* is enhanced by an *unc-101* mutation at all temperatures tested. For instance, at 20°, *let-23(n1045)* animals average 2.5 VPCs forming vulval tissue, while *unc-101; let-23(n1045)* animals average only 0.7 VPCs forming vulval tissue. The nature of the *n1045* allele is quite complex as it produces multiple products, making interpretation of *unc-101; let-23* interaction difficult (Aroian et al., 1993).

let-23 is necessary for several aspects of development (Aroian & Sternberg, 1991). *unc-101* mutations suppress *let-23* mutant phenotypes in a subset of these cells. Approximately 85% of *let-23(sy97)* animals die as L1 larvae. This lethality is not suppressed by an *unc-101* mutation (data not shown). Complete loss-of-function alleles of *let-23* confer L1 larval lethality on all homozygotes. Nor does an *unc-101* mutation suppress the lethality of homozygotes for complete loss-of-function alleles of *let-23*. The sterility of the allele *let-23(sy12)* is partially suppressed. While all viable *let-23(sy12)* hermaphrodites are sterile, approximately 65% of *unc-101; let-23(sy12)* animals are at least slightly fertile.

To determine if the interaction of *unc-101* is limited to *let-23* or if *unc-101* mutations are capable of suppressing mutations of other genes required for vulval induction, we constructed a series of double mutants and examined their extent of vulval differentiation (Table 1). *lin-2, lin-7,* and *lin-10* are genes that act formally near *let-23* and are required for vulval induction (Figure 3). *lin-2, lin-7,* and *lin-10* mutants have defects exclusively in the vulval induction process (Horvitz & Sulston, 1980; Sulston & Horvitz, 1981; Ferguson & Horvitz, 1985; Ferguson et al., 1987; Sternberg & Horvitz, 1989; Kim & Horvitz, 1990). Strong reduction-of-function alleles of any of these loci cause a vulvaless phenotype with an average of 0.4 to 0.9 VPCs per animal undergoing vulval differentiation. *unc-101* mutations can suppress the vulvaless phenotype of each of these mutants to an average of 3.4 VPCs differentiating per animal.

lin-3 encodes a member of the EGF family of growth factors and is most likely the inductive signal for vulval differentiation (Hill and Sternberg, 1992). Reduction-of-function mutations at this locus are suppressed only

partially by an *unc-101* mutation. Specifically, *lin-3(n378)* homozygotes average less than one vulval precursor cell (VPC) per animal forming vulval tissue, while *unc-101; lin-3(n378)* double mutants average slightly more than two VPCs forming vulval tissue per animal (three VPCs form vulval tissue in wild-type animals). This partial suppression is also true for animals bearing the allele *lin-3(e1417)*.

let-60 encodes a *C. elegans* ras homolog. Loss-of-function mutations or dominant negative mutations (dn) of this locus result in lethality and a vulvaless phenotype (Beitel et al., 1990, Han et al., 1990, Han and Sternberg, 1990). Dominant negative mutations of the *let-60* gene are not suppressed by *unc-101*. Neither animals of genotype *let-60(dn)* nor *unc-101; let-60(dn)* display any vulval differentiation.

lin-45 encodes a raf homolog that likely acts downstream of *let-60* (Han et al., 1993). A reduction-of-function mutation of *lin-45* is partially suppressed by an *unc-101* mutation. *lin-45(sy96)* animals average one VPC forming vulval tissue, while *unc-101; lin-45(sy96)* animals average 1.9 VPCs forming vulval tissue.

To summarize, *unc-101* mutations suppress reduction-of-function mutations of several of the genes required for vulval differentiation. We conclude that *unc-101* is a negative regulator of vulval induction, required for the proper regulation of EGF-Receptor mediated signaling (Figure 3). *lin-3* (the putative inductive signal, Hill and Sternberg, 1992) and *lin-45* (Han, et al. 1993) are only partially suppressed. *let-60* dominant negative mutations are not suppressed. Therefore we propose that *unc-101* acts at or near the *let-23* step. The fact that *unc-101* mutations do not suppress any complete loss-of-function mutations for this pathway suggest that *unc-101* mutations

do not simply lead to a bypass of this signaling pathway.

Correlation of the genetic and physical maps near *unc-101*

To clone *unc-101*, we correlated the genetic and physical maps around *unc-101* by identifying and mapping transposon polymorphisms close to *unc-101* (Fig. 3). Multipoint restriction fragment length polymorphism (RFLP) mapping (Ruvkun et al., 1989) was used to map two transposon polymorphisms, TCUNC101A and TCUNC101E, with respect to the genetic markers *unc-75*, *ced-1*, *unc-101* and *unc-59*. The source of these RFLPs was the strain MT3618 (*unc-75 ced-1 unc-59*), which contains a transposon rich region between *unc-75* and *unc-59* (S. Glass, T. Gerber and R. Horvitz, personal comm.), relative to the *unc-101 (sy108)* strain, a typical Bristol strain. We recovered Unc-59 non Unc-75 recombinants from + + *unc-101(sy108)* + / *unc-75 ced-1* + *unc-59* heterozygotes. By inverse PCR of genomic DNA from different recombinant animals (Ochman et al., 1988), we identified the flanking region of two polymorphisms, TCUNC101A and TCUNC101E. We found that TCUNC101A was present in the congenic strain, but was not present either in the + + + *unc-59* or the + + *unc-101 unc-59* recombinants, indicating that this polymorphism is to the left of *unc-101* on the genetic map. TCUNC101E was present in + + + *unc-59*, but not in + + + unc-*101 unc-59* recombinants; therefore, TCUNC101E is near *unc-101* and to the right of TCUNC101A.

We further mapped these two polymorphisms physically and genetically. By hybridization to a yeast artificial chromosome (YAC) grid filter (Coulson et al., 1991), we located TCUNC101A on the physical map on the right arm of chromosome I. TCUNC101E was located about 600 kb to the

right of TCUNC101A on the same contig. To genetically map these polymorphisms, we recovered 75 more recombinants from the heterozygotes described above, and performed either Southern hybridization with TCUNC101E as probe or PCR with TCUNC101E primers. No recombination events were observed between TCUNC101E and *unc-101*, indicating that TCUNC101E is very close to *unc-101*. Using these recombinants, the relative genetic distance between *ced-1*, *unc-101*, and *unc-75* could be more precisely defined. The relative frequency of recombination was: *unc-75* (5/38) *ced-1* (12/38) *unc-101* (21/38) *unc-59*. The other 37 recombinants were recovered in a screen biased to detect recombination between *unc-75* and *unc-101*. Since the physical distance between the two polymorphisms is about 600 kb, and since TCUNC101E is inseparable from *unc-101*(+) marker, we proceeded to test genomic cosmid clones within 100 kb of TCUNC101E for the ability to complement the *unc-101* phenotype.

Rescue of *unc-101* mutations by DNA-mediated transformation

We identified a genomic cosmid that can rescue the uncoordinated phenotype of *unc-101* mutations by DNA-mediated transformation. We tested five cosmids within 100 kb of TCUNC101E for their ability to rescue the uncoordinated phenotype of *unc-101(sy108)* by introducing cosmid DNA as an extrachromosomal multicopy transgene (Figure 4). Only cosmid W05A3 was able to rescue *unc-101(sy108)*. A 6.3 kb subclone of W05A3, pJL5, is the smallest genomic fragment capable of rescuing the uncoordinated phenotype. The pJL5 subclone rescues at least three other phenotypes of *unc-101(sy108)*: lethality, defective uptake of FITC, and suppression of *let-23(sy1)* vulvaless phenotype. The partial lethality of *unc-*

101(sy108) was rescued in the transgenic animals carrying pJL5 (Table 2). The viability of non-transgenic animals is 47%. By contrast, 72% of the progeny of the transgenic parents were viable. pJL5 also suppresses the FITC staining defect of *unc-101(sy108)* animals (data not shown). The vulval differentiation of the transgenic animals of genotype *unc-101(sy108)*; *let-23(sy1)* ; *Ex[pJL5]* was lowered to 0.8 VPCs induced per animal (n=20), which is the level of *let-23 (sy1)* single mutant animals but unlike the vulval differentiation of *unc-101(sy108)*; *let-23(sy1)* double mutant animals (3.6 VPCs per animal). Therefore, pJL5 also rescues the suppression of the *let-23(sy1)* mutation. We did not test the rescue of the other phenotypes associated with *unc-101* mutations.

Genomic and cDNA structure of *unc-101*

We isolated a full length cDNA clone from a cDNA library (Barstead and Waterston, 1989) using the pJL2 plasmid as probe (Figure 4), and determined its nucleotide sequence (Figure 5A). There are three in-frame stop codons 5' to the putative ATG codon. Also, the 5' end of the cDNA sequence has seven nucleotides that are identical to the 3' end sequence of the *trans*-spliced leader sequence SL1 (Krause and Hirsh, 1987) indicating that this cDNA has a full length coding sequence and that *unc-101* is a *trans*-spliced gene. A polyadenylation signal, AATAAA, is present at nucleotide 1844. The nucleotide sequence predicts that the *unc-101* gene product is a protein of 422 amino acids.

We inferred the genomic structure of *unc-101* by partial genomic sequence data and PCR with primers specific to cDNA regions (Figure 5B). The *unc-101* cDNA is divided into seven exons by six introns. The smallest

rescuing plasmid, pJL5, has 95 nucleotides 5' to the SL1 acceptor sequence.

Physical basis of *unc-101* mutations

To confirm our identification of the *unc-101* coding region, we analyzed the lesions associated with *unc-101* alleles. We determined the locations of the mutations of eight alleles of *unc-101* (Figure 5B). We determined the mutations by directly sequencing PCR-amplified DNA preparations from genomic DNA or single mutant animals (Kretz et al., 1989). All but one mutation are predicted to result in truncated proteins due to either deletions or nonsense mutations. *sy108* is a deletion of 115 nucleotides in exon 3 and intron 3 and an insertion of 8 nucleotides at the deletion point. *sy168* and *sy169*, which were obtained in the non-complementation screen using *sy108,* were the same mutation as *sy108*, suggesting that these are results of gene conversion events, or recovery of the maternal allele. *sy237, sy241*, *m1* and *rh6* are nonsense mutations, encoding truncated proteins. *sy237* is a G to A transition at the nucleotide 550, making a TGG to a TAG stop codon. *sy241* is a C to T transition mutation at the nucleotide 1284, changing a CAA to a TAA stop codon. *m1* is a C to T mutation at the nucleotide 1314, changing a CAA to a TAA stop codon. *rh6* is another C to T mutation at the nucleotide 1086, changing CAA to TAA. *sy161*, the only missense mutation, is a C to T mutation at the nucleotide 552, changing a CGC (arginine) to a TGC (cysteine). This arginine residue is conserved in both AP47 and AP50 (See results below). We were unable to amplify by PCR any genomic DNA from *sy216* homozygotes, suggesting that this mutation is a deletion of the entire

gene (see Methods).

***unc-101* encodes a clathrin-associated protein.**

A database search with the translation of the *unc-101* cDNA sequence indicated high similarity to two mammalian proteins, mouse AP47, the medium chain of the *trans*-Golgi associated clathrin-associated complex AP-1, and rat AP50, the medium chain of plasma membrane associated clathrin-associated complex AP-2 (Fig. 7; Nakayama et al., 1991; Thurieau et al., 1988) UNC-101 is also similar to a yeast protein, Yap54, the yeast homolog of AP47 (Nakayama et al., 1991). UNC-101 protein is 74% identical to AP47, and 42% to AP50, suggesting that UNC-101 is a homolog of the mammalian AP47.

To confirm that UNC-101 is a homolog of AP47 protein, we sought to identify AP50 homologs in *C. elegans*. The *C. elegans* genome sequencing consortium has identified a cDNA encoding a homolog of AP50 (Waterston et al., 1992). Using this cDNA clone as a probe, we isolated three more cDNA clones, one of which contained a full length cDNA, and determined their sequences. All three are products of a single gene. One of the three cDNA clones had a full length coding sequence (Figure 6). The comparison of the amino acid sequence of this protein with those of other homologs revealed that this AP50 homolog protein is 81% identical to AP50 and 40-42% identical to AP47 homologs, indicating that this sequence is indeed an AP50 homolog (Fig. 7). Based on the extent of the amino acid sequence identity, we believe that *unc-101* encodes a *C. elegans* homolog of AP47 protein.

We physically mapped the AP50 homolog to the X chromosome by YAC grid hybridization, and identified a cosmid (T16D11) that contains the entire

genomic region of AP50 homolog in the middle of its insert (data not shown). When injected into *unc-101 (sy108)* animals, this cosmid did not rescue the Unc phenotype of *unc-101 (sy108)* animals. This negative result is consistent with *unc-101* being a homolog of AP47, but not AP50.

The function of AP47 clathrin-associated protein is conserved in mammals and nematodes.

Since the amino acid sequence of *unc-101* is similar to that of the mammalian homolog AP47, it is conceivable that their function has been conserved during evolution. To test this hypothesis, we examined whether the mammalian AP47 homolog can rescue the mutant phenotype of *unc-101* animals. We constructed a hybrid gene with *unc-101* genomic DNA and mouse AP47 cDNA (Figure 8). In the mouse/ nematode hybrid construct, 298 amino acid residues of 422 amino acids are from the mouse cDNA. We also constructed an *unc-101* cDNA hybrid gene that is identical to the AP47 hybrid except that it has *unc-101* cDNA portion instead of AP47 cDNA. Both of these hybrid genes rescue the Unc phenotype of *unc-101(sy108)* animals (Figure 9). The level of vulval differentiation of the transgenic *unc-101(sy108)*; *let-23(sy1)* animals carrying the AP47 hybrid gene was an average of 2.0 VPCs per animal (n= 10), compared with 3.6 VPCs of *unc-101(sy108)*; *let-23(sy1)* animals without the transgene (n= 20), indicating that the suppression of the *let-23(sy1)* vulvaless phenotype was also rescued (Figure 10). Thus, mouse AP47 and *C. elegans* UNC-101 are not only very similar in sequence, but also functionally exchangeable. Rescue of the lethality associated with *unc-101(sy108)* animals was not checked due to the low transmission of the transgene.

Discussion

We have analyzed the *C. elegans unc-101* gene genetically and molecularly. *unc-101* mutations suppress the vulvaless phenotype of weak alleles of *let-23*, a *C. elegans* EGF receptor homolog required for vulval induction. We have cloned the *unc-101* gene and shown that it encodes a homolog of AP47, the medium chain of the *trans*-Golgi associated clathrin-associated complex. We also determined a full length sequence of a homolog of AP50, the medium chain of plasma membrane associated complex. Sequence comparison clearly showed that UNC-101 is an AP47 homolog. This homology was confirmed by the fact that a nematode *unc-101*/ mouse AP47 hybrid gene could functionally replace the nematode *unc-101* gene. For now, *unc-101* is the only case in which mutations of the clathrin AP genes cause any visible phenotype. Since *unc-101* mutations have pleiotropic phenotypes in many different tissues, and since mammalian AP47 and nematode UNC-101 are functionally interchangeable, one can study the function of clathrin coated vesicles in the regulation of a signal transduction processes. Further study of *unc-101* and homologs of other components of the clathrin coated vesicles on the plasma membrane and the *trans*-Golgi compartment will help understand the nature of the coated vesicles and their roles in a well-characterized signal transduction pathway.

Null phenotype of *unc-101*

We analyzed alleles of *unc-101* recovered in several screens to define the null phenotype of this locus. We are convinced that the visible alleles represented by *sy108* severely reduce *unc-101* function for the following

reasons. First, the non-complementation screens we performed should be able to recover null alleles of *unc-101* and the recovery frequency was close to that for typical null alleles (Brenner, 1974; Greenwald and Horvitz, 1980). Second, with the exception of *sy216,* all alleles recovered from non-complementation screens display very similar phenotypes to the previously recovered alleles. The lethal phenotype of *sy216* is quite different from that of other alleles, indicating that this lethality is due to a mutation in other essential gene(s) near *unc-101*. Supporting this possibility are the facts that we were unable to amplify any *unc-101* genomic DNA from *sy216* homozygotes, that the rescuing cosmid could not rescue *sy216* despite its ability to rescue the lethality of *unc-101(sy108)*, and that the Southern analysis failed to show any polymorphism in *sy216 / hIn1* heterozygotes using the rescuing cosmid as probe (data not shown), indicating that the deletion of *sy216* is larger than 30 kb. Third, sequence analysis of the mutant alleles showed that all but one allele are deletions or nonsense mutations, encoding truncated, and probably non-functional proteins.

The subviability of *unc-101* mutant animals could be explained either 1) by the residual function of *unc-101* or 2) by the presence of a partially redundant homolog of AP47. This redundant homolog could be expressed at different levels in different individual animals, and a higher level of expression might take over some essential function of UNC-101 protein, enabling them to survive despite loss of *unc-101*. Our preliminary observations indicate that there is another homolog of AP47 in *C. elegans* (J. Lee, unpublished results).

It is also formally possible that the N-terminal residues of the predicted truncated mutant proteins provide some essential function.

Structure and function of AP50/ AP47 proteins.

It is not surprising that AP50 and AP47 sequences are similar (40% identical over 422 amino acids; Nakayama et al., 1991), since other components of the clathrin-associated complexes also are similar. Mammalian AP17 and AP19 small chains have 44% identity, and β and β' heavy chains show the highest degree of identity of 84% (Kirchhausen et al., 1991). The exception is α and γ heavy chains, which are only 29% identical. Given the extent of similarity between the components of the APs, it is conceivable that β and β' chains have common functions such as binding to clathrin trimers, and that the medium chains and small chains have some specific and some common functions. In contrast, α and γ heavy chains may have specific functions such as binding to specific membrane marker proteins or membrane receptor proteins.

Amino acid residues conserved in both AP47 and AP50 may have common functions such as interaction with clathrin trimers and membrane components. These residues are distributed throughout the peptide sequence. Amino acid residues specifically conserved in either AP47 or AP50 homologs may be important for their specific functions such as interaction with specific membrane proteins, and/or other components of their own type of associated complex. These residues are also scattered throughout the peptide sequence, making it difficult to predict the domain structure of the proteins.

We have so far failed to find residues important for AP47-specific function. Most alleles are deletions or nonsense mutations. *sy161* is the only missense mutation, and it occurs at an amino acid conserved in both the

AP47 and AP50 proteins. To identify AP47-specific residues, more extensive screens for *unc-101* mutations could be performed using the protocols we described here.

The *m1* and *rh6* alleles encode proteins missing just a few C-terminal residues but confer identical phenotypes to other alleles. It is possible that the C-terminal residues are important for its function, stability, or regulation.

Negative regulation of *C. elegans* vulval induction

We have shown that *unc-101* also acts as a negative regulator of vulval differentiation. Mutations at *unc-101* suppress defects associated with mutations in genes such as *let-23, lin-2, lin-7* and *lin-10* that are required for specification of vulval fates, suggesting that the wild-type function of *unc-101* is to negatively regulate the process of vulval induction. The loss of *unc-101* activity in an otherwise wild-type animal confers no vulval differentiation defect (*unc-101* homozygotes, however, do have defective vulval morphogenesis, resulting in an egg-laying defect), suggesting that *unc-101* acts to refine the response to the inductive signal, rather than to prevent cells from generating vulval cells. Consistent with this, *unc-101* mutations are not capable of bypassing the complete lack of either the inductive signal, LET-23 receptors, or LET-60 ras proteins, suggesting that an *unc-101* mutation (and therefore the removal of one pathway of negative regulation) is not sufficient to promote vulval fates in the absence of any inductive signal. The enhancement of the *let-23(n1045)* allele is difficult to interpret because the *n1045* mutation results in a variety of abnormal messenger RNAs (Aroian et al., 1993). We suspect that some species of proteins produced might have inhibitory negative effect on the signal transduction (see Aroian & Sternberg,

1991) and that this effect is enhanced by the lack of negative regulation by *unc-101*.

Roles of clathrin-associated complexes and *unc-101*

The AP complexes are thought to drive clathrin coat formation and to couple the clathrin lattice with distinct membrane proteins. The core of the complex (Figure 1), which contains N terminal halves of large chains and the medium and small chains, can bind to clathrin trimers (Matsui and Kirchhausen, 1990; Peeler et al., 1993). The function of each component of AP complex is not well understood. It is possible that the medium chains function as regulators of the clathrin-associated protein complexes.

Since *unc-101* mutations have many different effects on the behavior and development of *C. elegans*, the wild-type UNC-101 protein has important functions in many different types of cells such as neurons, vulval precursor cells, and male spicule cells. How would UNC-101 function in these cells? Besides a general role in regulating membrane trafficking, one function could be negative regulation of receptor-mediated signal transduction as in vulval differentiation. The involvement of *unc-101* in the *let-23* pathway may be exhibited in some subsets of cells such as vulval precursor cells, but it is also possible that *unc-101* may regulate different types of receptors in other cells. Further genetic analysis may help reveal the full spectrum of *unc-101* interactions.

Based on the results from the biochemical studies on coated vesicles and assuming that UNC-101 is a *trans*-Golgi clathrin-associated protein, there are many possible molecular mechanisms by which UNC-101 could act within the cells involved in the *C. elegans* vulval induction. UNC-101 might

regulate secretion of the LIN-3 signal within the anchor cell: *unc-101* might be involved in negative regulation of LIN-3 signal production by sequestering LIN-3 proteins from the secretion route after their synthesis. We do not favor this model because *unc-101* mutants in a *let-23*(+) background do not have any excessive vulval differentiation while overexpression of *lin-3* causes excessive vulval differentiation (Hill and Sternberg, 1992). Another possible role of UNC-101 is regulation of transport of LET-23 receptors after their biosynthesis: UNC-101 might regulate LET-23 receptors by maintaining an intracellular pool of the receptors after their synthesis. This action could provide a post-translational regulation of the protein activity. Hence, in the absence of intracellular pool of the receptors, more LET-23 receptors whose activity is reduced by a reduction-of-function mutation could reach the cell surface, being capable of transducing more signal. Sorkin and Carpenter (1993) have recently shown that α-adaptins of AP-2 complexes of the plasma membrane interact with EGF-R in A431 cells; AP-1 complexes containing UNC-101/AP47 might play an analogous role at other stages in intracellular trafficking. A third possible role of UNC-101 is attenuation of signal transduction after activation of the molecules in the pathway by directing intracellular degradation: One of the best known functions of the coated vesicles on the *trans*-Golgi is sorting of the lysosomal enzymes. This is accomplished by sorting mannose-6-phosphate receptors, which recognize lysosomal enzymes tagged with mannose-6-phosphate, and transferring these receptor/ ligand complexes to the lysosome. UNC-101 might be indirectly involved in the down-regulation of the signal transduction components through establishing lysosomal or pre-lysosomal compartments.

Methods

Strains and genetic methods

Methods for culturing and handling the nematode and general genetic methods were described by Brenner (Brenner, 1974). All genetic experiments were performed at 20 °C. Mutagenesis protocol using trimethyl psoralen was provided by M. Yandell and L. Edgar (personal comm.). The standard strain N2 was from Brenner (1974). The alleles for examining genetic interactions of *unc-101* were as follows: *lin-3: e1417, n378* (Horvitz and Sulston, 1980); *let-23: sy1, sy97, sy12,* (Aroian and Sternberg, 1991) *n1045* (Ferguson and Horvitz, 1985); *lin-2: n768* (Ferguson and Horvitz, 1985), *e1309* (Horvitz and Sulston, 1980); *lin-7: e1413* (Ferguson and Horvitz, 1985); *lin-10: e1439* (Ferguson and Horvitz, 1985); *let-60: sy100dn* (Han et al., 1990); *lin-45: sy96* (Han et al., 1993). The starting strain for identifying the polymorphisms associated with *unc-101*, MT3618 (*unc-75 (e950) ced-1 (n1506) unc-59 (e261)*), which has a portion of chromosome I (between *unc-75* and *unc-59*) from the mutator TR679 strain, was provided by S. Glass, T. Gerber and R. Horvitz. The *unc-101* alleles sequenced for localization of mutations were *m1* (D. Riddle), *rh6* (E. Hedgecock), *sy108*, *sy161* (G. Jongeward and P. Sternberg, in prep.), *sy168*, *sy169*, *sy241* (this study), and *sy237* (J. Lee, unpublished results).

Inverse PCR and mapping of the polymorphisms

The method for inverse PCR was described by Ochman et al. (1988). We picked recombinants from heterozygous progeny of + + *sy108* + / *unc-75 ced-1* + *unc-59*, and made them homozygous. We digested genomic DNAs of

these recombinants with HindIII, diluted, and self-ligated them, then used them as templates for PCR with Tc1 internal primers of Hill and Sternberg (1992). After finding two polymorphisms, TCUNC101A and TCUNC101E, 1.1 kb and 0.8 kb long, respectively, on an analytic agarose gel, we subcloned the fragments into a pBluescript vector. We used the inserts of these subclones as probe for physical and genetic mapping (Williams et al., 1992). We determined the DNA sequence of TCUNC101E polymorphism and made two PCR primers for further mapping.

EPR1: 5' GGTGA TAGCA CCATA TGGTT CC 3'

EPR2: 5' ATATA GTGCT GTGCG GAACT C 3'

We designed these primers so that if the recombinants had the polymorphic transposon, the PCR-amplified band using either of these primers and the Tc1 internal primer would be 125 bp long, and if not, an 80 bp band would be amplified from EPR1 and EPR2 as extending primers.

Cosmids and *C. elegans* physical map

We obtained all cosmids and the physical map data from A. Coulson and J. Sulston (MRC, Cambridge, UK; Coulson et al., 1988, Coulson et al., 1986).

DNA -mediated transformation of *unc-101* mutants

Microinjection of cosmids or subcloned DNAs was described by Mello et al. (1991). We used *unc-101 (sy108)* animals as host strain for rescuing the Unc and the lethality of the visible allele. We co-injected *unc-101 (sy108)* animals with pRF4, which bears a dominant *rol*-6 mutation that results in a rolling phenotype. Selection of this marker phenotype facilitates selecting

the transgenic animals. We used *unc-101(sy108)*; *let-23 (sy1)* animals for examining the rescue of the suppression of the *let-23 (sy1)* vulvaless phenotype. We injected these animals with the rescuing subclone pJL5 without any other marker, because the rescue of the Unc phenotype itself serves as a good marker for the presence of the transgene.

Genomic DNA and cDNA manipulations and sequencing

All procedures of handling genomic DNA and cDNA were as described (Sambrook et al., 1989).

Sequence data analysis

The Macvector software package of IBI (New Haven CT.) and the software of the Genetics Computer Group v7.0 (Devereux et al, 1984; GCG, 1991) was used to edit the genomic and cDNA sequences. Database search was performed using the BLAST program. Pileup and Prettyplot commands were used to generate the comparison of the amino acid sequences.

Determination of mutations

We made ten PCR primers from the intron sequences, 5' end and 3' nontranslated region as follows.

INT1: TTCCG CTAAT TTTCT CCGG

INT2: ATTGC GTCAT TTTCA ACGG

INT3: CGCTC CAATG ATAAA ACAC

INT4: GCATT TTCGC ATTGG AGCG

INT5: AAATG TGTTT TTCGA CTCG

INT6: AAAAA CTAGG CCACA TCAC

INT7: AAGTC AGGCC ATGCC TCAA

INT8: CATAA ATCTC ACATT GGGCA

INT9: GAGAA TTATG TGATT TTTTG

INT10: CTCGG CCACG GTCGT TTTT

With these PCR primers, we could amplify all the exons and their flanking intron sequences. We used genomic DNA or single worms of the homozygous mutant in PCR to amplify the exons and their boundaries. We directly sequenced the amplified DNAs using the same sets of primers (Kretz et al., 1989).

Construction of hybrid plasmid with *unc-101* and AP47

To construct an *unc-101*/ AP47 hybrid gene, we used the *unc-101* rescuing plasmid pJL2, the *unc-101* cDNA subclone, and mouse AP47 cDNA subclone as sources for the sequences. We kept the first two introns of the rescuing pJL2 plasmid in this construction because pJL2 has only 95 nucleotides 5' to the SL-1 acceptor sequence, and because expression of transgenes are more efficient in the presence of introns (e.g., Brinster et al., 1988, Buchman and Berg, 1988, Fire et al., 1990). We digested the rescuing plasmid pJL2 with NruI and EcoRV, and purified the 7.2 kb fragment from a low melting point agarose gel and ligated it with the 0.9 kb NruI/ EcoRV fragment from the *unc-101* cDNA subclone. The resulting plasmid was the *unc-101* cDNA hybrid gene. We made two PCR primers from the AP47 cDNA sequence to amplify the corresponding region of NruI/ EcoRV fragment of *unc-101* cDNA.

47-1: 5' CGACA ACTTT GTCAT CATCT A;

47-2 : 5' ATCCA CTCTT CTCAA TGATT TTC 3'

To facilitate subcloning, we replaced three nucleotides of the 5' ends of 47-1 and 47-2 primers with the recognition sequence of NruI and EcoRV, respectively. This replacement does not change the coding amino acid. We ligated the amplified DNA into the pJL2 NruI/ EcoRV fragment. We confirmed the correct reading frame by regeneration of the NruI and EcoRV sites. This construct has the 5' region of *unc-101* including the 5' region of the rescuing plasmid pJL2 and 5' coding region with two introns up to cDNA nucleotide 388, fused in frame to the AP47 cDNA from 389 to 1281, *unc-101* cDNA nucleotide from 1282 to the stop codon, and all the 3' untranslated region of *unc-101*.

Assay for the rescue of the phenotypes of *unc-101* animals by the nematode/mouse hybrid gene.

We obtained stable lines of transgenic animals after microinjection of the hybrid gene. For the Unc phenotype rescue, we transferred five transgenic animals to a spot on new plates, and after given time, we photographed the tracks that the animals created by moving on the bacterial lawn viewed in a Wild M5A stereomicroscope. For the rescue of the suppression of *let-23(sy1)* vulvaless phenotype, we examined the vulval induction of the transgenic animals bearing the hybrid gene in their late L3 stages using Nomarski optics.

Analysis of *sy216*

To phenotypically examine *sy216* and other *unc-101* alleles, we constructed *trans*-heterozygous animals bearing *sy216/ sy108*. The lethality of *sy216/ sy108 trans*-heterozygotes is enhanced. 74% of *sy216/ sy108*

heterozygous animals are inviable compared to 45% of *sy108/ sy108* animals. However, suppression of the vulvaless phenotype of *let-23(sy1)* is not enhanced. *sy216/ sy108*; *let-23(sy1)* animals averaged 3.6 VPCs undergoing vulval differentiation, which is the same level of vulval differentiation of *sy108/ sy108 ; let-23(sy1)* animals (n= 20, respectively).

Since TMP is known to induce high frequency of small deletions (L. Edgar, personal comm.), we suspected that *sy216* could be a small deficiency that deletes neighboring genes as well as *unc-101*. To address this issue genetically, we tested whether any of nearest genes are deleted by *sy216*. We constructed the *trans*-heterozygote *unc-101(sy216) / eDf3*. *eDf3* fails to complement the mutations that define the nearest genetically defined loci to the right, *unc-59*, *let-201, let-202,* and *let-203* (Figure 2A). This heterozygote is viable and wild-type. Therefore, *sy216* does not delete the nearest genetically defined loci to the right. *sy216* does not delete the nearest gene to the left, *ced-1*. Specifically, an animal of genotype *dpy-5 + sy216 / hIn1* was found among the cross-progeny of a *+ ced-1 sy216 / dpy-5 + unc-101(rh6)* hermaphrodite mated with *hIn1* males. Therefore, the *sy216* chromosome must contain a functional *ced-1* gene. However, it is likely that there are essential gene(s) in the interval between *ced-1* and *eDf3,* as this region is not saturated for lethal mutations.

To examine the rescue of the lethality of *sy216* by the rescuing cosmid W05A3, we used *sy216 / hIn1 unc-54* (*h1040*) heterozygotes as host for the transgene. *hIn1* is an inversion that suppresses recombination in the region between *unc-75* and *unc-59* (Zetka and Rose, 1992). We co-injected the W05A3 cosmid and the pRF4 marker DNAs, picked rolling F1 transgenic animals, and transferred twenty F2 individual animals to new plates and

checked their segregation of *unc-54* marker. All of them segregated *unc-54,* indicating that there was no animal of genotype *sy216/ sy216; [Ex W05A3 + rol-6]* .

To amplify genomic DNA from the *sy216 / sy216* arrested animals, we picked two or three arrested animals, and performed PCR as described above. As an internal positive control, we used a set of *let-23* PCR primers in the same PCR reactions with *unc-101* primers. We tried five different sets of *unc-101* primers, but only recovered an amplified band of *let-23* genomic DNA from the *let-23* primers.

To detect any polymorphisms linked to *sy216*, we used genomic DNA of *sy216 / hIn1* heterozygotes digested with various restriction enzymes in a Southern hybridization with W05A3 cosmid as probe. We did not detect any polymorphism compared with *hIn1* homozygotes as control.

Acknowledgments

We are grateful to D. Riddle and E. Hedgecock for sharing their *unc-101* alleles. We thank S. Lemmon and G. Payne for sharing their unpublished results. We thank L. Edgar for the trimethylpsoralen mutagenesis protocol; T. Kirchhausen for the AP47 cDNA; A. van der Bliek, H. Chamberlin, A. Golden, L. Huang, P. Kayne, J. Liu, T. Kirchhausen and R. Palmer for their comments. We also thank A. Coulson and the MRC and St. Louis group for all the cosmids, physical maps and YACs; S. Glass, T. Gerber and R. Horvitz for providing the MT3618 strain; M. Zetka and A. Rose for providing the *hIn1* strain. Some nematode strains were provided by *Caenorhabditis elegans* Genetic Center, which is funded by the NIH National

center for Research Resources. P.W.S. is an investigator of the Howard Hughes Medical Institute. This research has been supported by USPHS grant HD 23690 to P.W.S.

References

Ahle, S., Mann, A., Eichelsbacher, U. and Ungewickell, E. (1988). Structural relationships between clathrin assembly proteins from the Golgi and the plasma membrane. EMBO J. *7*, 919-929.

Aroian, R.V., Koga, M., Mendel, J.E., Ohshima, Y. and Sternberg, P.W. (1990). The *let-23* gene necessary for *Caenorhabditis elegans* vulval induction encodes a tyrosine kinase of the EGF receptor subfamily. *Nature* 348, 693-699.

Aroian, R.V., Levy, A.D., Koga, M., Ohshima, Y., Kramer, J.M. and Sternberg, P.W. (1993). Splicing in *Caenorhabditis elegans* does not require an AG at the 3' splice acceptor site. Mol. Cell. Biol. *13*, 626-637.

Aroian, R.V. and Sternberg, P.W. (1991). Multiple functions of *let-23*, a *C. elegans* receptor tyrosine kinase gene required for vulval induction. Genetics *128*, 251-267.

Barstead, R.J. and Waterston, R.H. (1989). The basal component of the nematode dense-body is vinculin. J. Biol. Chem. *264*, 10177-10185.

Bazinet, C., Katzen, A.L., Morgan, M., Mahowald, A.P. and Lemmon, S.K. (1993). The Drosophila clathrin heavy chain gene: Clathrin function is essential in a multicellular organism. Genetics *134*, 1119-1134.

Beitel, G., Clark, S. and Horvitz, H.R. (1990). The *Caenorhabditis elegans ras* gene *let-60* acts as a switch in the pathway of vulval induction. Nature *348*, 503-509.

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

Brinster, R.L., Allen, J.M., Behringer, R.R., Gelinas, R.E. and Palmiter, R.D. (1988). Introns increase transcriptional efficiency in transgenic mice. Proc. Natl. Acad. Sci. USA *85*, 836-840.

Brodsky, F.M. (1988). Living with clathrin: its role in intracellular membrane traffic. Science *242*, 1396-1402.

Buchman, A.R. and Berg, P. (1988). Comparison of intron-dependent and intron-independent gene expression. Mol. Cell. Biol. *8*, 4395-4405.

Clark, S.G., Stern, M.J. and Horvitz, H.R. (1992). *C. elegans* cell-signalling gene *sem-5* encodes a protein with SH2 and SH3 domains. Nature *356*, 340-344.

Coulson, A., Kozono, Y., Lutterbach, B., Shownkeen, R., Sulston, J. and Waterson, R. (1991). YACs and the *C. elegans* Genome. BioEssays *13*, 413-417.

Coulson, A., Waterston, R., Kiff, J., Sulston, J. and Kohara, Y. (1988). Genome linking with yeast artificial chromosomes. Nature *335*, 184-186.

Coulson, A.R., Sulston, J., Brenner, S. and Karn, J. (1986). Toward a physical map of the genome of the nematode *Caenorhabditis elegans*. Proc Natl Acad Sci USA *83*, 7821-7825.

Devereux, J., Haeberli, P. and Smithies, O. (1984). A comphrehensive set of sequence analysis programs for the VAX. Nucl Acids Res *12*, 387-395.

Ferguson, E. and Horvitz, H.R. (1985). Identification and characterization of 22 genes that affect the vulval cell lineages of *Caenorhabditis elegans*. Genetics *110*, 17-72.

Ferguson, E.L., Sternberg, P.W. and Horvitz, H.R. (1987). A genetic

pathway for the specification of the vulval cell lineages of *Caenorhabditis elegans*. Nature *326*, 259-267.

Fire, A., White-Harrison, S. and Dixon, D. (1990). A modular set of *lacZ* fusion vectors for studying gene expression in *Caenorhabditis elegans*. Gene *93*, 189-198.

GCG (1991). Program Manual for the GCG Package, Version 7.

Greenwald, I.S. and Horvitz, H.R. (1980). *unc-93(e1500)*: A behavioral mutant of *Caenorhabditis elegans* that defines a gene with a wild-type null phenotype. Genetics *96*, 147-164.

Han, M. and Sternberg, P.W. (1990). *let-60*, a gene that specifies cell fates during *C. elegans* vulval induction, encodes a ras protein. Cell *63*, 921-931.

Han, M., Golden, A., Han, Y. and Sternberg, P.W. (1993). *C. elegans lin-45 raf* gene participates in *let-60 ras* stimulated vulval differentiation. Nature *363*, 133-140

Herman, R.K. and Hedgecock, E.M. (1990). The size of the *C. elegans* vulval primordium is limited by *lin-15* expression in surrounding hypodermis. Nature *348*, 169-171.

Hill, R.J. and Sternberg, P.W. (1992). The *lin-3* gene encodes an inductive signal for vulval development in *C. elegans*. Nature *358*, 470-476.

Horvitz, H.R. and Sternberg, P.W. (1991). Multiple intercellular signalling systems control the development of the *C. elegans* vulva. Nature *351*, 535-541.

Horvitz, H.R. and Sulston, J.E. (1980). Isolation and genetic characterization of cell-lineage mutants of the nematode *Caenorhabditis elegans*. Genetics *96*, 435-454.

Keen, J.H. (1987a). Clathrin assembly proteins: affinity purification and a model for coat assembly. J. Cell Biol. *105*, 1989-1998.

Keen, J.H., Chesnut, M. H., Beck, K. A. (1987b). The clathrin coat assembly polypeptide complex: autophosphorylation and assembly activities. J. Biol. Chem *262*, 12182-12188.

Keen, J.H. (1990). Clathrin and associated assembly and disassembly proteins. Annn. Rev. Biochem. *59*, 415-438.

Kim, S.K. and Horvitz, H.R. (1990). The *Caenorhabditis elegans* gene *lin-10* is broadly expressed while required specifically for the determination of vulval cell fates. Genes & Devel. *4*, 357-371.

Kirchhausen, T., Davis, A.C., Frucht, S., Greco, B.O. and Payne, G.S. (1991). AP17 and AP19, the mammalian small chains of the clathrin-associated protein complexes show homology to YAP17p, their putative homolog in yeast. J. Biol. Chem *266*, 1153-1157.

Kirchhausen, T., Nathanson, K.L., Matsui, W., Vaisberg, A., Chow, E.P. (1989) Structural and functional division into 2 domains of the large (100-kD to 115-kD) chains of the clathrin-associated protein complex AP-2. P.N.A.S. *86,* 2612 -2616

Matsui, W., Kirchhausen, T. (1990) Stabilization of clathrin coats by the core of the clathrin-associated protein complex AP-2. Biochem. *29*, 10791 -10798

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

Kretz, K.A., Carson, G.S. and Obrien, J.S. (1989). Direct sequencing from low-melting agarose with sequenase. N. A. R. *17*, 5864.

Mello, C.C., Kramer, J.M., Stinchcomb, D. and Ambros, V. (1991). Efficient

gene transfer in *C. elegans* after microinjection of DNA into germline cytoplasm: extrachromosomal maintenance and integration of transforming sequences. EMBO J. *10*, 3959-3970.

Nakayama, Y., Goebl, M., Greco, B. O., and Lemmon, S., Chow, E. P. C., Kirchhausen, T. (1991). The medium chains of the mammalian clathrin-associated proteins have a homolog in yeast. Eur. Jour. of Biochem. *202*, 569-574.

Ochman, H., Gerber, A.S. and Hartl, D.L. (1988). Genetic applications of an inverse polymerase chain reaction. Genetics *120*, 621-623.

Pearse, B.M.F. and Robinson, M.S. (1990). Clathrin, adapters, and sorting. Annu. Rev. Biol. *6*, 151-171.

Peeler, J.S., Donzell, W.C., Anderson, R.G.W. (1993) The appendage domain of the AP-2 subunit is not required for assembly or invagination of clathrin-coated pits. J. Cell. Biol. *120,* 47-54

Ruvkun, G., Ambros, V., Coulson, A., Waterston, R., Sulston, J. and Horvitz, H.R. (1989). Molecular genetics of the *Caenorhabditis elegans* heterochronic gene *lin-14*. Genetics *121*, 501-516.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular cloning: a laboratory manual Second edition. Cold Spring Harbor Laboratory Press.

Seeger, M. and Payne, G.S. (1992). Selective and immediate effects of clathrin heavy-chain mutations on Golgi membrane-protein retention in *Saccharomyces cerevisiae.* J. Cell. Biol. *118*, 531-540.

Schmid, S.L. (1992). The mechanism of receptor-mediated endocytosis - more questions than answers. Bioessays *14*, 589-596.

Sorkin, A., Carpenter, G. (1993) Interaction of activated EGF receptors with coated pit adaptins. Science *261*, 612-615

Sternberg, P.W. and Horvitz, H.R. (1989). The combined action of two intercellular signalling pathways specifies three cell fates during vulval induction in *C. elegans*. Cell *58*, 679-693.

Sulston, J.E. and Horvitz, H.R. (1981). Abnormal cell lineages in mutants of the nematode *Caenorhabditis elegans*. Dev. Biol. *82*, 41-55.

Thomas, J.H. (1990). Genetic analysis of defecation in *Caenorhabditis elegans*. Genetics *124*, 855-872.

Thurieau, C., Brosius, J., Burne, C., Jolles, P., Keen, J., Mattaliano, R.J., Chow, E.P., Ramachandran, K. and Kirchhausen, T. (1988). Molecular cloning of complete amino acid sequence of AP50, an assembly protein associated with clathrin-coated vesicles. DNA *7*, 663-669.

Williams, B.D., Schrank, B., Huynh, C., Shownkeen, R. and Waterston, R.H. (1992). A genetic mapping system in *Caenorhabditis elegans* based on polymorphic sequence-tagged sites. Genetics *131*, 609-624.

Wood, W.B. (1988). The nematode *Caenorhabditis elegans*. Cold Spring Harbor Laboratory.

Zetka, M.C. and Rose, A.M. (1992). The meiotic behavior of an inversion in *Caenorhabditis elegans*. Genetics *131*, 321-332.

Table 1. Interaction of *unc-101* with other genes in the vulval induction pathway. Vulval differentiation in approximately twenty animals per each genotype was examined using Nomarski optics. A wild-type animal has three VPCs induced. Fewer than three VPCs induced result in a vulvaless phenotype. When more than three VPCs generate vulval tissue, a normal vulva and pseudovulval tissue are usually produced.

animals Genotype	*unc-101(+)* VPCs/animal forming vulval tissue	*unc-101(+)* % animals w/>3 VPCs	*unc-101(sy108)* VPCs/animal forming vulval tissue	*unc-101(sy108)* % w/>3VPCs
+	3	0	3	0
lin-3(e1417)	0.8	0	1.4	0
lin-3(n378)	0.8	0	2.1	0
let-23(sy97)	0	0	0.06	0
let-23(sy12)	0.02	0	2.9	13
let-23(sy1)	0.8	0	3.6	44
let-23(n1045) 15°	1.1	0	0.2	0
let-23(n1045) 20°	2.5	10	0.6	0
let-23(n1045) 25°	3.4	45	2.0	0
lin-2(n768)	2.8	0	3.2	20
lin-2(e1309)	0.5	0	3.4	35
lin-7(e1413)	0.9	0	3.4	40
lin-10(e1439)	0.5	0	3.4	30
let-60(sy100 dn)	0	0	0	0
lin-45(sy96)	0.9	0	1.9	7

Table 2. Complementation of *unc-101* mutation by pJL5 : Suppression of lethality. Eggs were transferred to new plates from hermaphrodites of three different genotypes, and the number of surviving adults and rolling animals were counted after three days. The number of rolling animals represent the stability of the transgene. pJL5 plasmid, when maintained in the transgenic animals, can enable the animals to survive (see text for details).

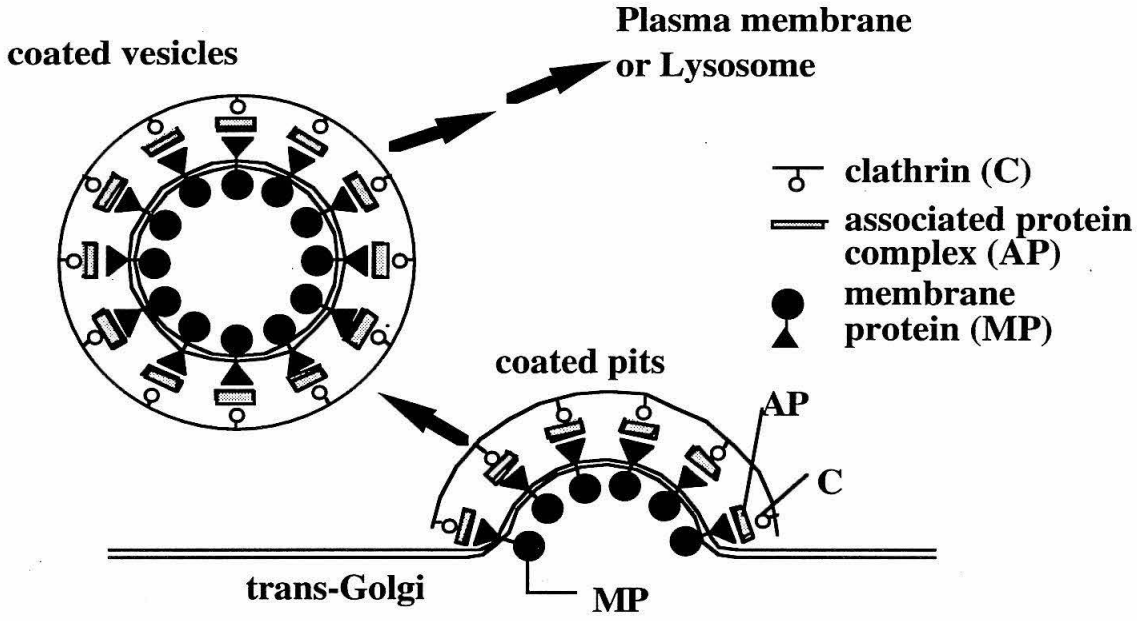
Genotype of parent	Transgene present	Number of eggs picked	Number of viable adults	Viability	Number of rollers	Stability of the transgene
+/+	Yes	390	354	91%	161	41%
sy108/+;	Yes	261	256	98%	79	31%
sy108/sy108	Yes	241	173	72%	76	32%
*sy108/sy108**	No	91	43	47%	0	N/A

* Unc animals from the transgenic line, that have lost the transgene.

Transgene derived from the microinjection of pJL5 with *rol-6(d)* as dominant marker.

Figure 1. Clathrin vesicles and their associated protein complexes

(A) coated pits and vesicles. A coated pit is composed of membrane fraction, collected membrane proteins, and clathrin triskelion cage with its associated protein complex. Coated pits are invaginated to form coated vesicles, which travel to their destinations such as lysosome and the plasma membrane. (Modified from Pearse and Robinson, 1990)


coated vesicles
Plasma membrane
or Lysosome
clathrin (C)
associated protein
complex (AP)
membrane
protein (MP)
coated pits
AP
C
trans-Golgi
MP

(B) Schematic structure of the clathrin associated protein complex. AP-1 complex is composed of four different peptide chains: two large chains, β, γ, one medium chain AP47, and one small chain AP19. AP-2, not shown in the figure, has similar structure to that of AP-1, and is composed of two large chains, α, β, one medium chain AP50, and one small chain AP17 (Modified from Nakayama et al., 1991).

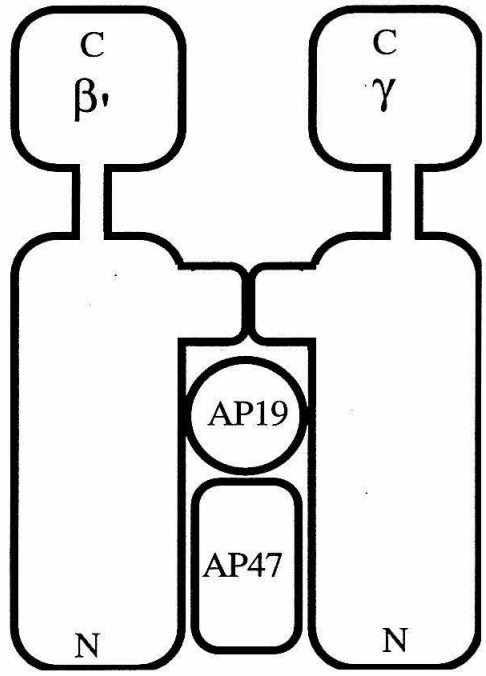

C
β'
C
γ
AP19
AP47
N
N

Figure 2. Isolation of null alleles of *unc-101*.

(A) genetic map around *unc-101*.

(B), (C) Non-complementation screens for the isolation of the null alleles of *unc-101*. EMS or trimethylpsoralen was used as mutagen. *dpy* is a marker for distinguishing the cross progeny from the self progeny. *eDf24* is a deficiency that deletes the rRNA genes on the chromosome I, conferring lethality to the homozygotes for this deficiency. *hIn1* is an inversion of chromosome I that suppresses recombination in the region between *unc-75* and *unc-54*. This balancer is marked with *unc-54,* facilitating the discrimination of the homozygotes.

Figure 2A.

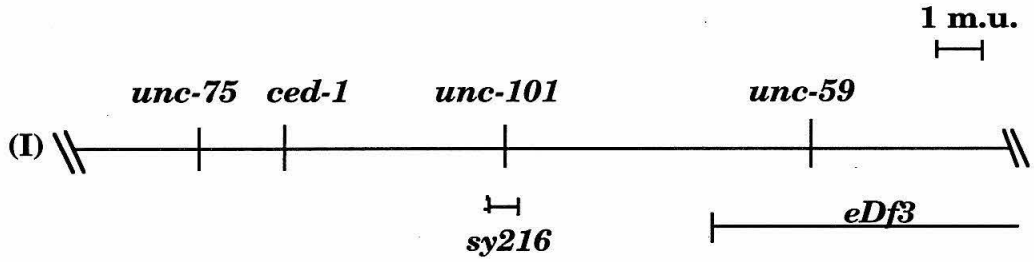


Figure 2B

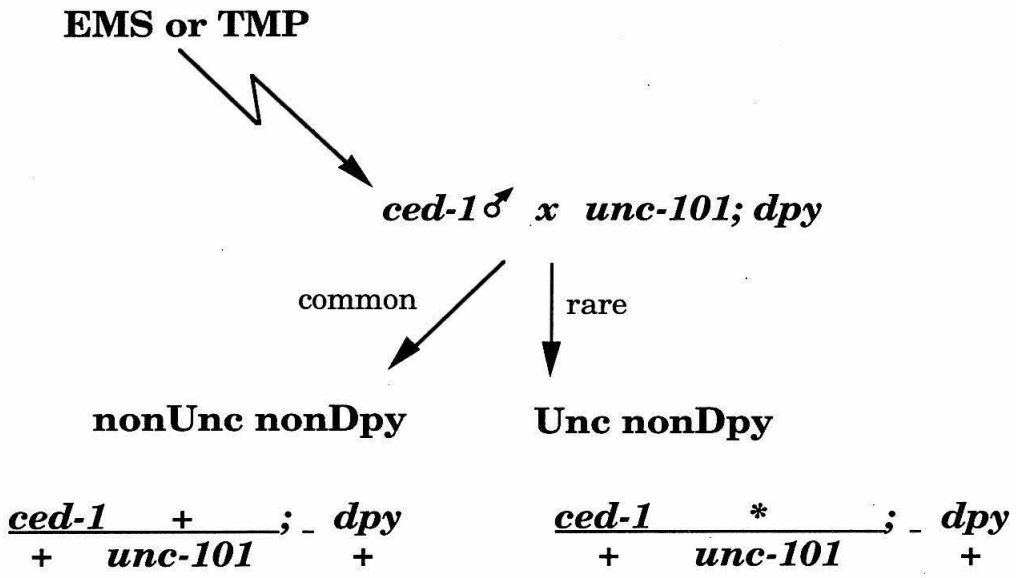

EMS or TMP
ced-1♂ x unc-101; dpy
common
rare
nonUnc nonDpy
Unc nonDpy
ced-1 + ; dpy
+ unc-101 +
ced-1 * ; dpy
+ unc-101 +

Figure 2C

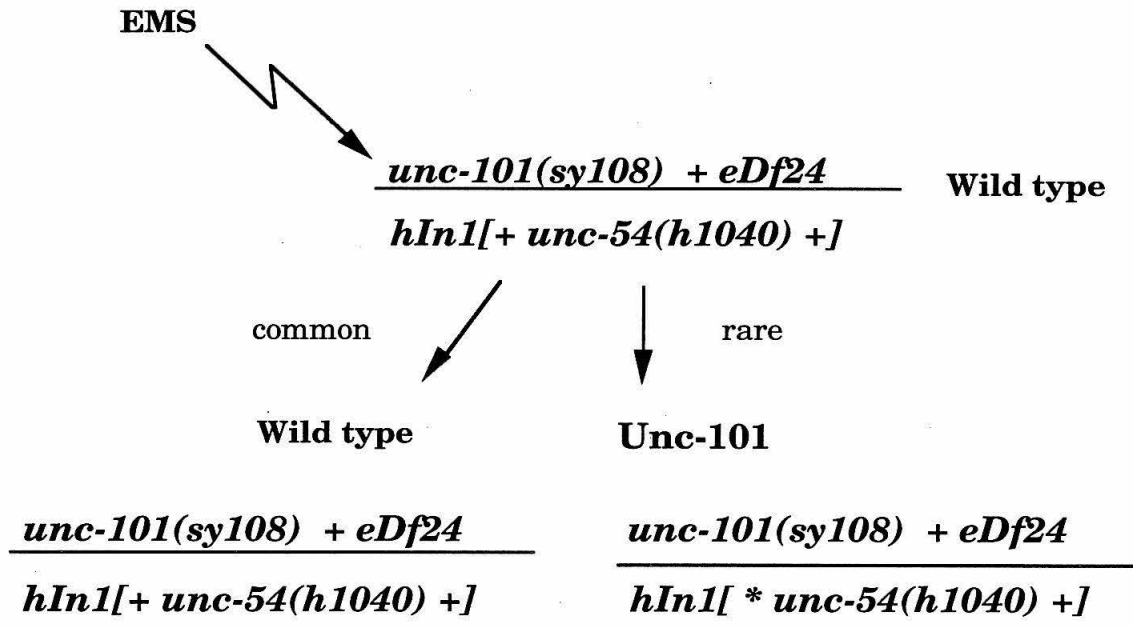

EMS
unc-101(sy108) + eDf24
hIn1[+ unc-54(h1040) +]
Wild type
common
rare
Wild type
Unc-101
unc-101(sy108) + eDf24
hIn1[+ unc-54(h1040) +]
unc-101(sy108) + eDf24
hIn1[* unc-54(h1040) +]

Figure 3. *unc-101* is a regulator of the vulval induction pathway.

The inductive signal produced by the anchor cell is likely to be encoded by *lin-3* (Hill and Sternberg, 1992). The signal encoded by *lin-3* is likely to be received by the *let-23* gene product, an EGF receptor homolog (Aroian et al., 1990). The genes *sem-5* and *let-60*, proposed to act downstream of *let-23*, encode a GRB2 homolog with SH2 and SH3 domains, and a *ras* homolog, respectively (Clark et al., 1992; Han and Sternberg, 1990). *lin-45*, which likely acts after *let-60*, encodes a *C. elegans raf* homolog (Han et al., 1993). *lin-2, lin-7,* and *lin-10* are also required for the vulval differentiation. *unc-101* mutations strongly suppress mutations of *lin-2, lin-7, lin-10,* and *let-23*, suggesting that *unc-101* may act at or near these genes (See text and table 1 for details).

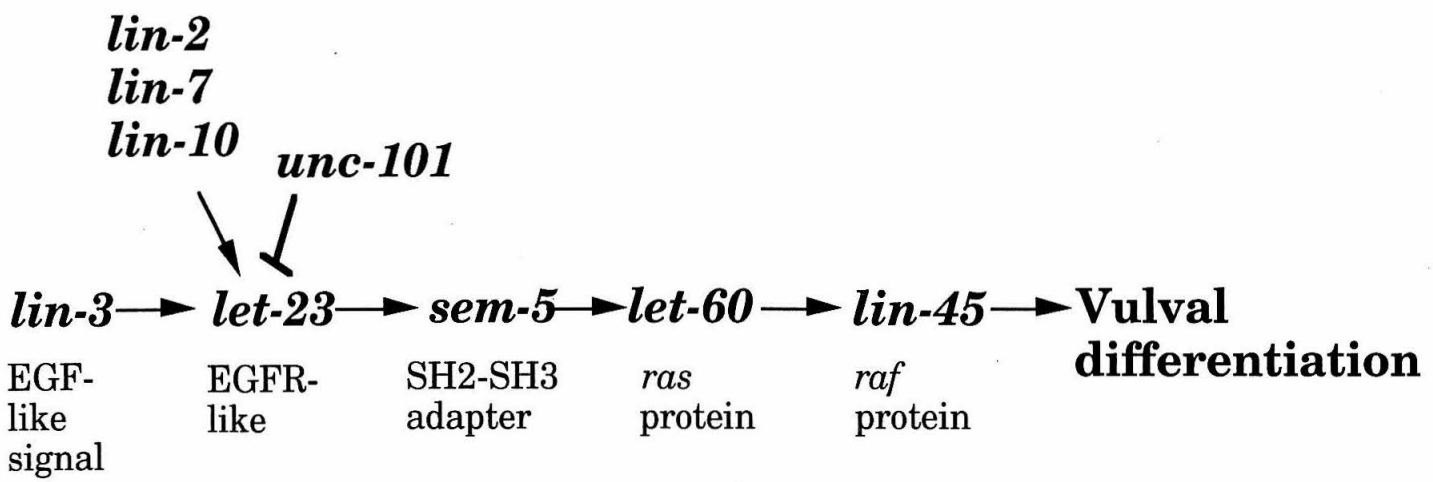
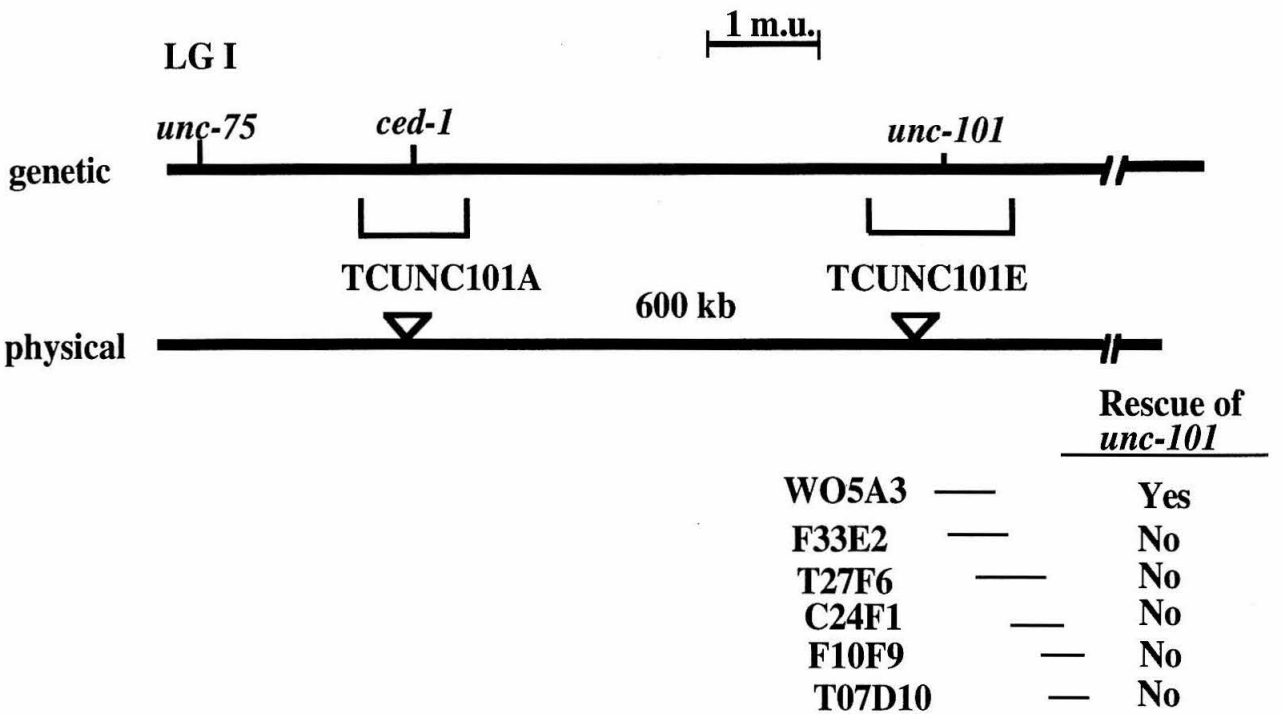

lin-2
lin-7
lin-10
unc-101
lin-3
let-23
sem-5
let-60
lin-45
Vulval differentiation
EGF-like signal
EGFR-like
SH2-SH3 adapter
ras protein
raf protein

Figure 4. Cloning of the *unc-101* locus

(A) Genetic and physical map near *unc-101*. The genetic distance between *ced-1* and *unc-101* is about 5 map units. The two RFLPs are marked as triangles. TCUNC101A is in the same YAC as the *ced-1* YAC. TCUNC101E is 600 kb from TCUNC101A, and genetically inseparable from *unc-101(+)* marker. The W05A3 cosmid, but not the other 4 cosmids, had the ability to rescue Unc phenotype of *unc-101(sy108)*.


1 m.u.
LG I
unc-75
ced-1
unc-101
genetic
TCUNC101A
TCUNC101E
600 kb
physical
Rescue of unc-101
WO5A3
Yes
F33E2
No
T27F6
No
C24F1
No
F10F9
No
T07D10
No

(B) Map of W05A3 cosmid and its subclones. pJL5, a 6.3 kb subclone, is the smallest genomic region capable of rescuing. pJL3 does not rescue the phenotype, suggesting that this restriction enzyme site disrupts the *unc-101* gene. Restriction sites shown are as follows: S, SpeI; N, NdeI; H, HindIII; P, PstI; and A, ApaI.

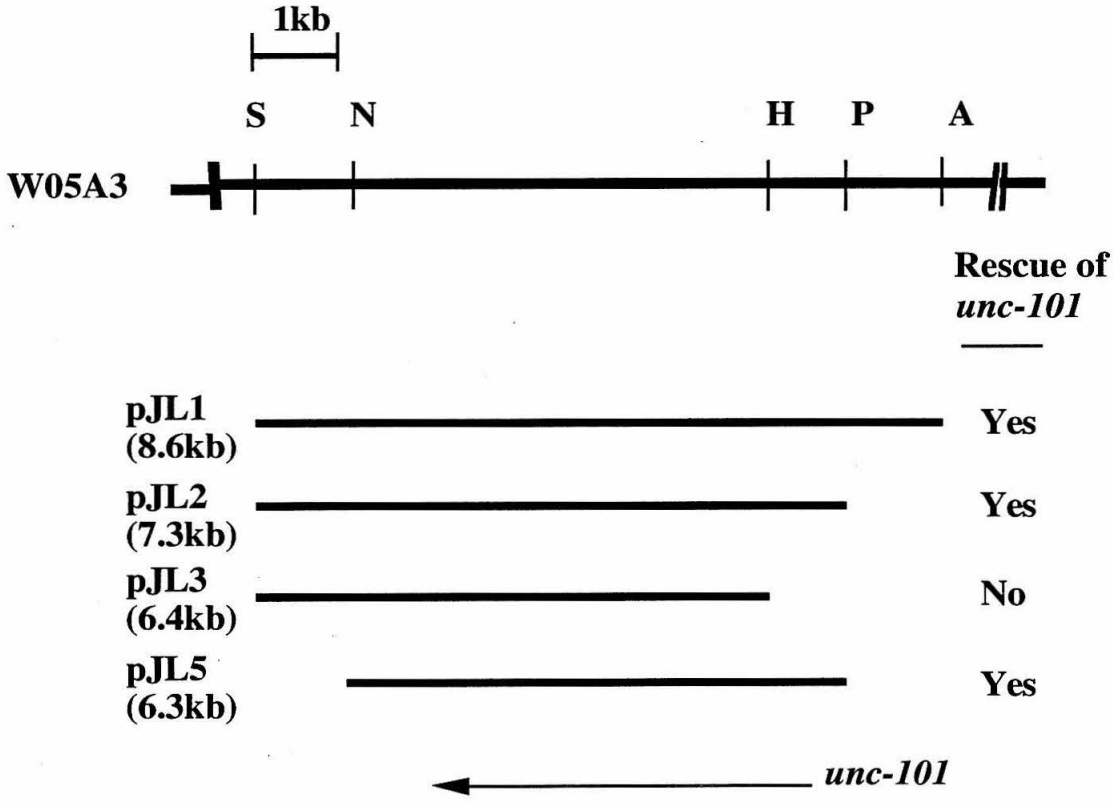

1kb
S
N
H
P
A
W05A3
Rescue of
unc-101
pJL1
(8.6kb)
Yes
pJL2
(7.3kb)
Yes
pJL3
(6.4kb)
No
pJL5
(6.3kb)
Yes
unc-101

Figure 5. cDNA sequence and genomic structure of *unc-101*.

(A) *unc-101* cDNA sequence and its predicted translation product. The region identical to the 3' end of SL-1 sequence is underlined. The start and stop codons, and a polyadenylation signal sequence are also underlined. The sites for the construction of the mammalian AP47 hybrid gene are marked as lines with an arrowhead. The first nucleotide G is the last nucleotide of the vector.

```
GAAGTTTGAGTATTTTCCAGTAGCTGCCACGTGGAATTTGACGATTTTAACGAGAAAATC 60
GCAAAAAATCGTCGAAAATGGCGACTTCCGCCATGTTTATACTGGATTTGAAGGGAAAAA 120
                    M  A  T  S  A  M  F  I  L  D  L  K  G  K
CGATAATTTCTCGAAATTATCGCGGAGACATCGACATGACGGCAATCGATAAATTCATTC 180
T  I  I  S  R  N  Y  R  G  D  I  D  M  T  A  I  D  K  F  I
ATTTACTCATGGAAAAAGAGGAAGAAGGCTCGGCAGCGCCCGTTTTGACCTATCAGGACA 240
H  L  L  M  E  K  E  E  E  G  S  A  A  P  V  L  T  Y  Q  D
CGAATTTCGTGTTTATCAAGCACACAAATATTTATTTGGTCTCAGCATGCCGTTCAAACG 300
T  N  F  V  F  I  K  H  T  N  I  Y  L  V  S  A  C  R  S  N
TCAACGTCACAATGATTTTGTCATTTTTGTACAAATGCGTCGAAGTTTTCTCCGAATATT 360
V  N  V  T  M  I  L  S  F  L  Y  K  C  V  E  V  F  S  E  Y
TCAAAGATGTCGAAGAGGAGTCGGTTCGCGACAATTTTGTCGTTATCTATGAACTTTTGG 420
F  K  D  V  E  E  E  S  V  R  D  N  F  V  V  I  Y  E  L  L
ACGAAATGATGGATTTCGGGTTCCCACAGACGACTGAGAGTCGAATTCTACAAGAATACA 480
D  E  M  M  D  F  G  F  P  Q  T  T  E  S  R  I  L  Q  E  Y
TCACACAAGAAGGTCAAAAACTAATTTCGGCACCCCGTCCCCCGATGGCAGTGACAAATG 540
I  T  Q  E  G  Q  K  L  I  S  A  P  R  P  P  M  A  V  T  N
CCGTCTCATGGCGCTCTGAAGGCATAAAATACCGAAAAAACGAGGTTTTCCTGGACGTAA 600
A  V  S  W  R  S  E  G  I  K  Y  R  K  N  E  V  F  L  D  V
TCGAAAGTGTGAACATGTTGGCCAGCGCCAACGGTACCGTACTTCAATCGGAAATTGTTG 660
I  E  S  V  N  M  L  A  S  A  N  G  T  V  L  Q  S  E  I  V
GAAGCGTTAAAATGCGTGTCTATCTTACCGGAATGCCTGAACTTCGGCTGGGTCTTAACG 720
G  S  V  K  M  R  V  Y  L  T  G  M  P  E  L  R  L  G  L  N
ATAAAGTACTTTTTGAGGGCAGTGGACGCGGAAAAAGCAAATCTGTGGAACTGGAAGACG 780
D  K  V  L  F  E  G  S  G  R  G  K  S  K  S  V  E  L  E  D
TGAAATTTCATCAATGTGTACGCCTGTCGCGTTTTGACACGGATCGAACGATCTCCTTCA 840
V  K  F  H  Q  C  V  R  L  S  R  F  D  T  D  R  T  I  S  F
TACCGCCCGACGGAGCATTTGAGCTTATGAGCTATCGATTAACAACCGTGGTGAAGCCGC 900
I  P  P  D  G  A  F  E  L  M  S  Y  R  L  T  T  V  V  K  P
TGATCTGGATCGAGACAAGCATCGAACGTCACAGTCACAGCCGTGTCTCGTTTATAATCA 960
L  I  W  I  E  T  S  I  E  R  H  S  H  S  R  V  S  F  I  I
AAGCGAAATCACAATTCAAACGGCGCTCCACTGCTAATAACGTGGAAATCATTATTCCAG 1020
K  A  K  S  Q  F  K  R  R  S  T  A  N  N  V  E  I  I  I  P
TCCCGTCGGACGCTGATTCACCGAAATTCAAGACAAGCATCGGTTCGGTGAAGTATACGC 1080
V  P  S  D  A  D  S  P  K  F  K  T  S  I  G  S  V  K  Y  T
CCGAGCAATCGGCCTTCGTATGGACTATTAAGAATTTTCCCGGCGGAAAAGAGTACCTTT 1140
P  E  Q  S  A  F  V  W  T  I  K  N  F  P  G  G  K  E  Y  L
TGACCGCCCATCTATCTCTACCGTCTGTGATGAGTGAAGAGTCTGAAGGACGGCCGCCGA 1200
L  T  A  H  L  S  L  P  S  V  M  S  E  E  S  E  G  R  P  P
TTAAAGTCAAATTTGAAATTCCGTATTTTACGACCAGCGGCATTCAGGTCCGTTATCTGA 1260
I  K  V  K  F  E  I  P  Y  F  T  T  S  G  I  Q  V  R  Y  L
AAATCATCGAGAAAAGAGGATATCAAGCATTGCCGTGGGTCCGCTACATTACTCAAAATG 1320
K  I  I  E  K  R  G  Y  Q  A  L  P  W  V  R  Y  I  T  Q  N
GAGAATACGAGATGCGGATGAAATAATTCCTGAAAAAATTACCTAAATTCATATTTTATT 1380
G  E  Y  E  M  R  M  K  *
GTATTTTATTCCCAATTTTACTCTTAATTTTTGGAATTTTTTATGAAAAATTGGTGAAAA 1440
ACGACCGTGGCCGAGTTTTTGGAAAATTTGGAGGCTAGGCCACCATGCTTCCAGTGGTGG 1500
GCTAACTTTTCGAAAATCCTAGCCACGGCCCCGTTTTCCATCAATTTTTGCCTCTTTTAA 1560
TGTCAAACATCTCCAATTTTTTCTGTGAAAATTTAATGCTCCGCGAGCTGCTCCCCGGCT 1620
CAAAACATGTGTTGTGTGTGCTCTTTCCCCTGATGACCCCGAACTCTATTTTTTTTTGTT 1680
CGAAAATTTTTATTTTATTTTTCCCACCGATTTATTTATTGATTTTTATCGCATAATTTA 1740
GATTTTTTCCAGAAAAACGAGTTTTTTCCCTTTCCCACCAGCATAATTTTCTCACCACAT 1800
GGATCCTCATCAATTTTTCCCGTTTTCCTTTTTTCAGTAAATCAATAAAATTTTCTGTCA 1860
TAATTAAAAAAAAAAAAAAAAC 1883
```

(B) Genomic structure of *unc-101*. *unc-101* has seven exons and six introns. The size of introns varies from 50 bp to 1.0 kb. The numbers above the structures represent the first nucleotide of the cDNA after SL1 sequence, the translation starting point, and the last nucleotides of exons. The locations and characteristics of mutations are marked as such.

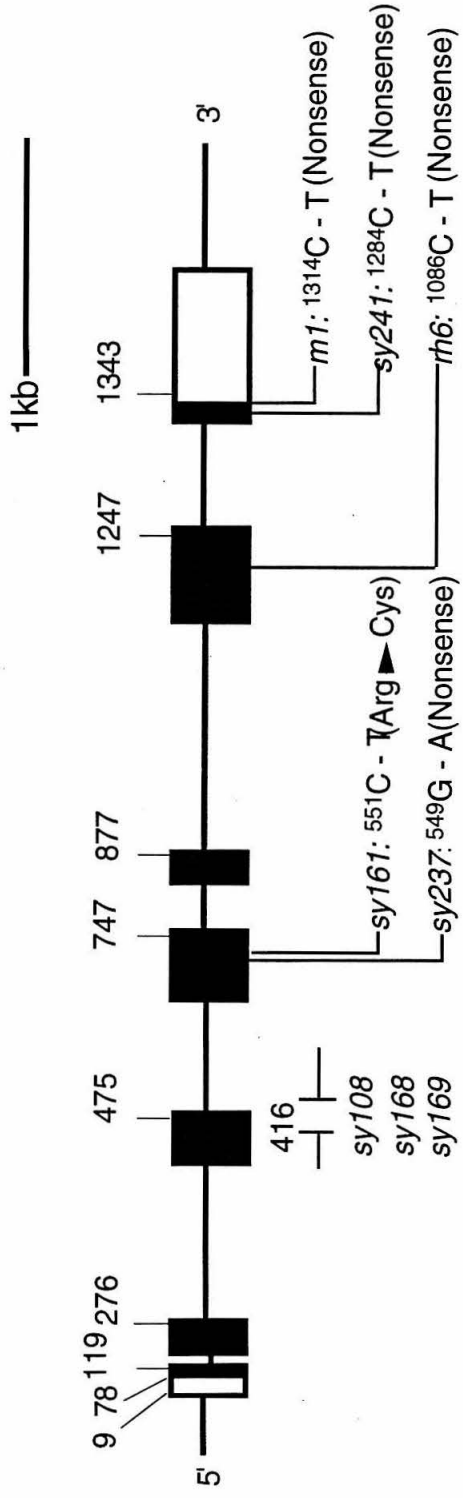

1kb
5'
9
78
119
276
475
416
747
877
1247
1343
3'
sy108
sy168
sy169
sy161: 551C - T(Arg ▶ Cys)
sy237: 549G - A(Nonsense)
m1: 1314C - T (Nonsense)
sy241: 1284C - T (Nonsense)
rh6: 1086C - T (Nonsense)

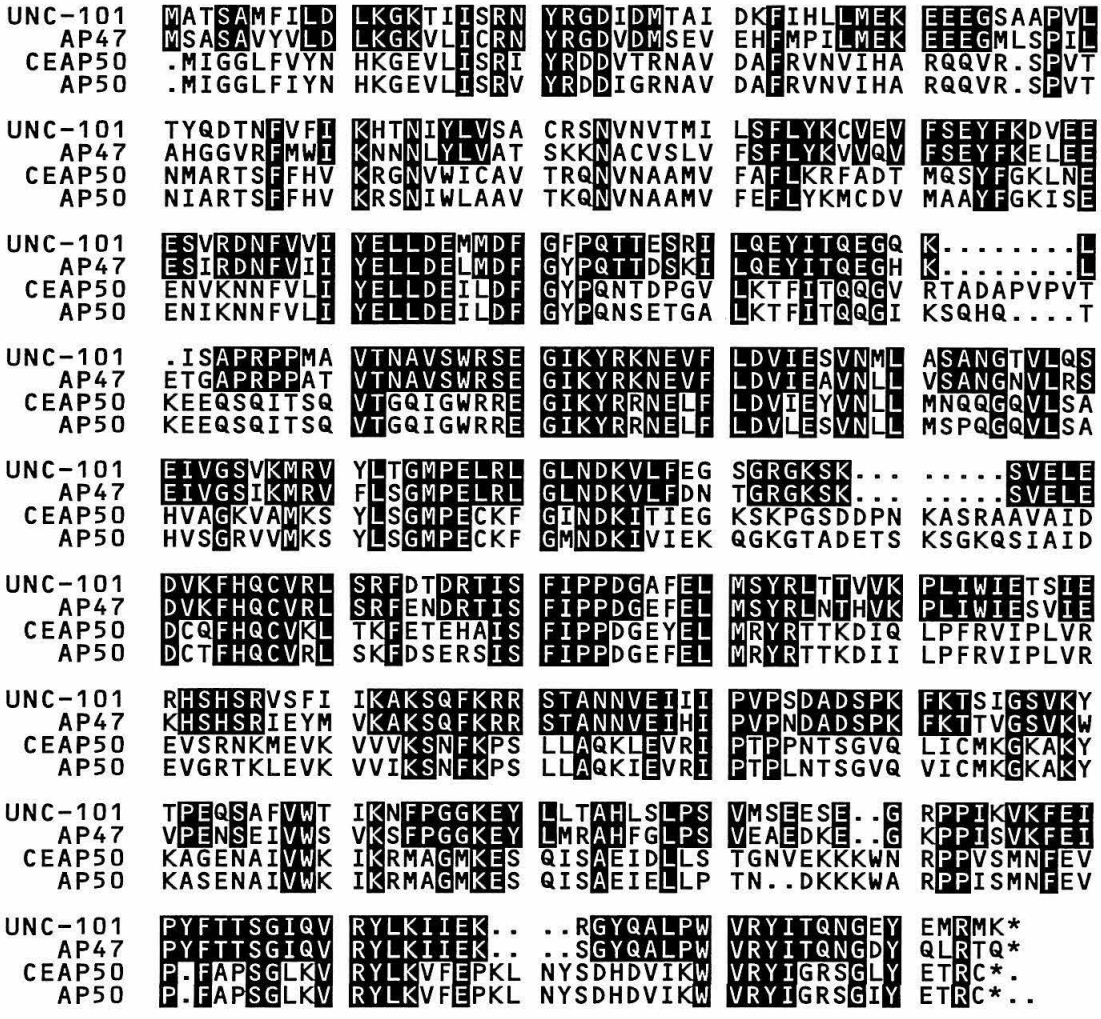
Figure 6. AP50 homolog in *C. elegans*:

cDNA sequence of an AP50 homolog in *C. elegans* and its predicted translation product is shown. The start and stop codons, and a putative polyadenylation signal are underlined.

```
CGGTTTGAGCAAACTCGGGGATGATTGGTGGATTGTTCGTTTACAATCACAAAGGAGAAG 60
                  M  I  G  G  L  F  V  Y  N  H  K  G  E
TGCTCATTTCGAGAATCTATCGAGACGATGTAACCCGGAACGCAGTCGACGCCTTCCGAG 120
V  L  I  S  R  I  Y  R  D  D  V  T  R  N  A  V  D  A  F  R
TCAACGTCATCCATGCCCGACAGCAAGTTCGCTCGCCAGTCACCAACATGGCTCGTACTT 180
V  N  V  I  H  A  R  Q  Q  V  R  S  P  V  T  N  M  A  R  T
CGTTCTTCCATGTGAAGCGTGGCAACGTCTGGATTTGTGCGGTGACACGTCAAAATGTCA 240
S  F  F  H  V  K  R  G  N  V  W  I  C  A  V  T  R  Q  N  V
ACGCTGCCATGGTTTTTGCGTTCTTGAAACGCTTCGCCGACACCATGCAGTCTTACTTTG 300
N  A  A  M  V  F  A  F  L  K  R  F  A  D  T  M  Q  S  Y  F
GAAAACTGAACGAGGAGAATGTGAAGAACAACTTTGTGTTGATTTATGAGTTGCTCGACG 360
G  K  L  N  E  E  N  V  K  N  N  F  V  L  I  Y  E  L  L  D
AGATTCTCGACTTTGGATACCCCCAGAATACGGACCCTGGTGTGCTGAAAACTTTCATCA 420
E  I  L  D  F  G  Y  P  Q  N  T  D  P  G  V  L  K  T  F  I
CCCAGCAAGGAGTCAGAACAGCTGATGCTCCTGTCCCAGTGACCAAAGAGGAGCAGTCAC 480
T  Q  Q  G  V  R  T  A  D  A  P  V  P  V  T  K  E  E  Q  S
AAATCACGTCTCAAGTGACTGGCCAAATTGGATGGCGTCGGGAGGGTATTAAGTACCGCC 540
Q  I  T  S  Q  V  T  G  Q  I  G  W  R  R  E  G  I  K  Y  R
GAAATGAGCTCTTCCTGGATGTTATTGAATATGTCAACTTGCTCATGAATCAACAAGGAC 600
R  N  E  L  F  L  D  V  I  E  Y  V  N  L  L  M  N  Q  Q  G
AAGTATTATCTGCTCATGTTGCCGGAAAAGTTGCGATGAAATCCTATTTGAGTGGAATGC 660
Q  V  L  S  A  H  V  A  G  K  V  A  M  K  S  Y  L  S  G  M
CGGAGTGCAAATTTGGCATCAACGACAAAATTACCATCGAAGGAAAGTCGAAGCCAGGAA 720
P  E  C  K  F  G  I  N  D  K  I  T  I  E  G  K  S  K  P  G
GTGATGATCCAAACAAAGCAAGCCGTGCCGCAGTGGCCATTGATGACTGTCAATTCCACC 780
S  D  D  P  N  K  A  S  R  A  A  V  A  I  D  D  C  Q  F  H
AATGCGTGAAGCTGACGAAATTTGAGACGGAGCACGCGATTTCTTTTATCCCACCGGACG 840
Q  C  V  K  L  T  K  F  E  T  E  H  A  I  S  F  I  P  P  D
GCGAGTACGAGCTGATGAGATACCGTACCACTAAGGATATCCAACTGCCATTCCGTGTGA 900
G  E  Y  E  L  M  R  Y  R  T  T  K  D  I  Q  L  P  F  R  V
TCCCATTGGTTCGTGAAGTGTCTCGTAACAAGATGGAAGTTAAGGTTGTCGTCAAGTCTA 960
I  P  L  V  R  E  V  S  R  N  K  M  E  V  K  V  V  V  K  S
ACTTCAAGCCATCCCTTCTTGCTCAAAAGCTCGAAGTTCGCATTCCAACCCCACCAAATA 1020
N  F  K  P  S  L  L  A  Q  K  L  E  V  R  I  P  T  P  P  N
CATCCGGCGTTCAACTTATTTGCATGAAGGGAAAAGCCAAGTACAAGGCAGGCGAGAATG 1080
T  S  G  V  Q  L  I  C  M  K  G  K  A  K  Y  K  A  G  E  N
CCATTGTGTGGAAAATAAAGCGTATGGCCGGAATGAAGGAAAGCCAAATTTCTGCGGAAA 1140
A  I  V  W  K  I  K  R  M  A  G  M  K  E  S  Q  I  S  A  E
TCGATCTTCTCTCAACTGGAAACGTTGAGAAGAAGAAATGGAATCGCCCACCGGTCAGCA 1200
I  D  L  L  S  T  G  N  V  E  K  K  K  W  N  R  P  P  V  S
TGAACTTTGAGGTTCCGTTTGCTCCATCTGGACTCAAAGTTCGCTACTTGAAGGTGTTTG 1260
M  N  F  E  V  P  F  A  P  S  G  L  K  V  R  Y  L  K  V  F
AGCCAAAACTGAACTATTCGGATCATGACGTCATCAAATGGGTTCGTTACATTGGAAGAT 1320
E  P  K  L  N  Y  S  D  H  D  V  I  K  W  V  R  Y  I  G  R
CGGGACTGTATGAAACCAGATGCTAGAACTTCACCCAACCCTTTCTATCTTCATTTGCTC 1380
S  G  L  Y  E  T  R  C  *
CCCAGCCGTCAAGCTTGAACATTTGCTCATTCTCGGTTCCAAGTGTTAAATATTATTTAT 1440
TTGCTTCATACAATTTTTAATTTTTTGTAATATTTTTGTATTATTTTCCAAATTTCATAA 1500
TGTAGAGTCATCGCAGCATTACAGTACTGTAAGAGTTGTTTTCAATTTATAAGAACTATT 1560
AGATCTTTAGACGTCTCGATAAGCATTTCGCATGCTTATAATTCCATGCATTTTCCGTGC 1620
CAAAATTCAAAACCCCCGCCCCACCCCGTTCCATCTTTATTTATTGTCAGTGTAAAATTT 1680
CAAAATTTCTGGAATCTTTTTCCATCATAAAATTCACTTCAAACGTCTTCTGACCCGTTG 1740
GAGCAATAACATTTTTAAAATATTTTTTTCATTTACCTTTTCTTCATAATTCATCATAAT 1800
TTATTGTAACGGCTTAATTTGATATATTGTTCTTTCTACATTCCATTTTTTGTCTCTCTC 1860
TCAAAATTTTCAAATTTCTATTGTTCATTTCTCATTTCTAGGTCGCAATTCAATAAGTTT 1920
GTAATTCAAATGAATAGCATAATATTTATTTTAAATAAAAGTTTTTTTTATTAAAGCGGA 1980
ATTCC 1985
```

Fig 7. UNC-101 is a clathrin-associated protein.

(A) Amino acid sequence alignment of homologs of UNC-101. AP47 and AP50 are the medium chains of *trans*-Golgi clathrin-associated protein complex and plasma membrane clathrin-associated protein complex, respectively (Nakayama et al., 1991; Thurieau et al., 1988). CEAP50 is an AP50 homolog in *C. elegans* (this study). This alignment only highlights the residues that are identical between UNC-101 and AP47 or between all four homologs.


UNC-101 MATSAMFILD LKGKTIISRN YRGDIDMTAI DKFIHLLMEK EEEGSAAPVL
AP47 MSASAVYVLD LKGKVLICRN YRGDVDMSEV EHFMPILMEK EEEGMLSPIL
CEAP50 .MIGGLFVYN HKGEVLISRI YRDDVTRNAV DAFRVNVIHA RQQVR.SPVT
AP50 .MIGGLFIYN HKGEVLISRV YRDDIGRNAV DAFRVNVIHA RQQVR.SPVT
UNC-101 TYQDTNFVFI KHTNIYLVSA CRSNVNVTMI LSFLYKCVEV FSEYFKDVEE
AP47 AHGGVRFMWI KNNNLYLVAT SKKNACVSLV FSFLYKVVQV FSEYFKELEE
CEAP50 NMARTSFFHV KRGNVWICAV TRQNVNAAMV FAFLKRFADT MQSYFGKLNE
AP50 NIARTSFFHV KRSNIWLAAV TKQNVNAAMV FEFLYKMCDV MAAYFGKISE
UNC-101 ESVRDNFVVI YELLDEMMDF GFPQTTESRI LQEYITQEGQ K.....L
AP47 ESIRDNFVII YELLDELMDF GYPQTTDSKI LQEYITQEGH K.....L
CEAP50 ENVKNNFVLI YELLDEILDF GYPQNTDPGV LKTFITQQGV RTADAPVPVT
AP50 ENIKNNFVLI YELLDEILDF GYPQNSETGA LKTFITQQGI KSQHQ....T
UNC-101 .ISAPRPPMA VTNAVSWRSE GIKYRKNEVF LDVIESVNML ASANGTVLQS
AP47 ETGAPRPPAT VTNAVSWRSE GIKYRKNEVF LDVIEAVNLL VSANGNVLRS
CEAP50 KEEQSQITSQ VTGQIGWRRE GIKYRRNELF LDVIEYVNLL MNQQGQVLSA
AP50 KEEQSQITSQ VTGQIGWRRE GIKYRRNELF LDVLESVNLL MSPQGQVLSA
UNC-101 EIVGSVKMRV YLTGMPELRL GLNDKVLFEG SGRGKSK... .....SVELE
AP47 EIVGSIKMRV FLSGMPELRL GLNDKVLFDN TGRGKSK... .....SVELE
CEAP50 HVAGKVAMKS YLSGMPECKF GINDKITIEG KSKPGSDDPN KASRAAVAID
AP50 HVSGRVVMKS YLSGMPECKF GMNDKIVIEK QGKGTADETS KSGKQSIAID
UNC-101 DVKFHQCVRL SRFDTDRTIS FIPPDGAFEL MSYRLTTVVK PLIWIETSIE
AP47 DVKFHQCVRL SRFENDRTIS FIPPDGEFEL MSYRLNTHVK PLIWIESVIE
CEAP50 DCQFHQCVKL TKFETEHAIS FIPPDGEYEL MRYRTTKDIQ LPFRVIPLVR
AP50 DCTFHQCVRL SKFDSERSIS FIPPDGEFEL MRYRTTKDII LPFRVIPLVR
UNC-101 RHSHSRVSFI IKAKSQFKRR STANNVEIII PVPSDADSPK FKTSIGSVKY
AP47 KHSHSRIEYM VKAKSQFKRR STANNVEIHI PVPNDADSPK FKTTVGSVKW
CEAP50 EVSRNKMEVK VVVKSNFKPS LLAQKLEVRI PTPPNTSGVQ LICMKGKAKY
AP50 EVGRTKLEVK VVIKSNFKPS LLAQKIEVRI PTPLNTSGVQ VICMKGKAKY
UNC-101 TPEQSAFVWT IKNFPGGKEY LLTAHLSLPS VMSEESE..G RPPIKVKFEI
AP47 VPENSEIVWS VKSFPGGKEY LMRAHFGLPS VEAEDKE..G KPPISVKFEI
CEAP50 KAGENAIVWK IKRMAGMKES QISAEIDLLS TGNVEKKKWN RPPVSMNFEV
AP50 KASENAIVWK IKRMAGMKES QISAEIELLP TN..DKKKWA RPPISMNFEV
UNC-101 PYFTTSGIQV RYLKIIEK.. ..RGYQALPW VRYITQNGEY EMRMK*
AP47 PYFTTSGIQV RYLKIIEK.. ..SGYQALPW VRYITQNGDY QLRTQ*
CEAP50 P.FAPSGLKV RYLKVFEPKL NYSDHDVIKW VRYIGRSGLY ETRC*.
AP50 P.FAPSGLKV RYLKVFEPKL NYSDHDVIKW VRYIGRSGIY ETRC*..

(B) Diagram of the comparison of identity among the homologs of medium chains of clathrin-associated complexes. The numbers represent the identity at amino acid sequence level. The most prominent homologies are marked by bold lines. This shows that UNC-101 and AP47 are of a group, and AP50 and CEAP50 are of another.

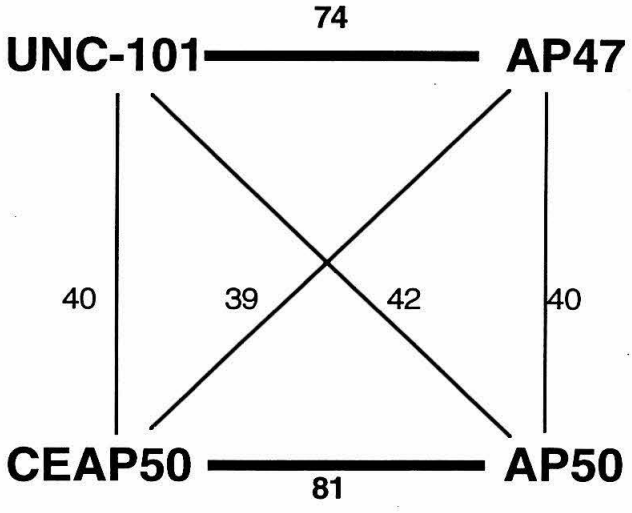

74
UNC-101
AP47
40
39
42
40
CEAP50
AP50
81

Figure 8. Schematic of mouse-nematode hybrid gene construction.

The NruI/ EcoRV fragment of the rescuing plasmid pJL2 was replaced with the corresponding cDNA fragments from *unc-101* cDNA clone or mouse AP47 cDNA clone in the *unc-101* hybrid gene and in the *unc-101/AP47* hybrid gene, respectively.

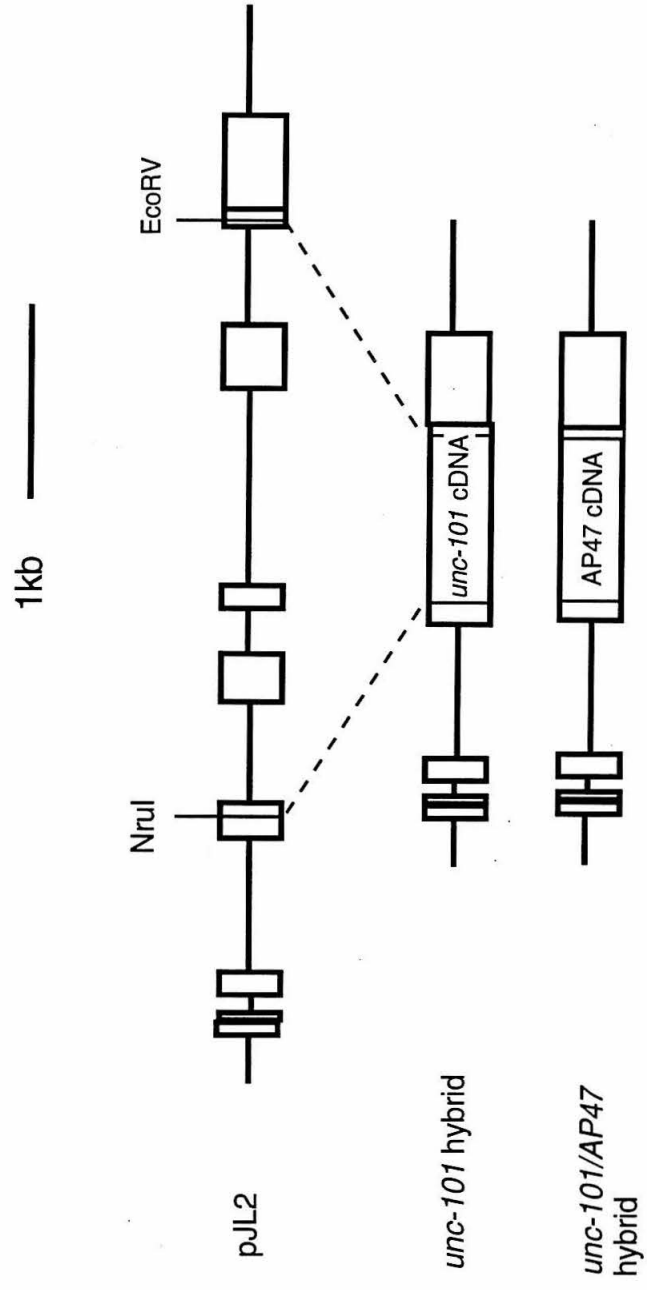

1kb
EcoRV
NruI
unc-101 cDNA
AP47 cDNA
pJL2
unc-101 hybrid
unc-101/AP47
hybrid

Figure 9. Rescue of Unc phenotype by the mouse-nematode hybrid gene.

The tracks that the worms create by moving on the bacterial lawn are shown. A, N2 wild type, 5 minutes after transfer; B, *unc-101 (sy108),* 40 minutes after transfer; C, transgenic animal of *unc-101 (sy108)*; *Ex[AP47 hybrid],* 5 minutes after transfer. The uncoordinated movement of *unc-101 (sy108)* is suppressed by the transgene.

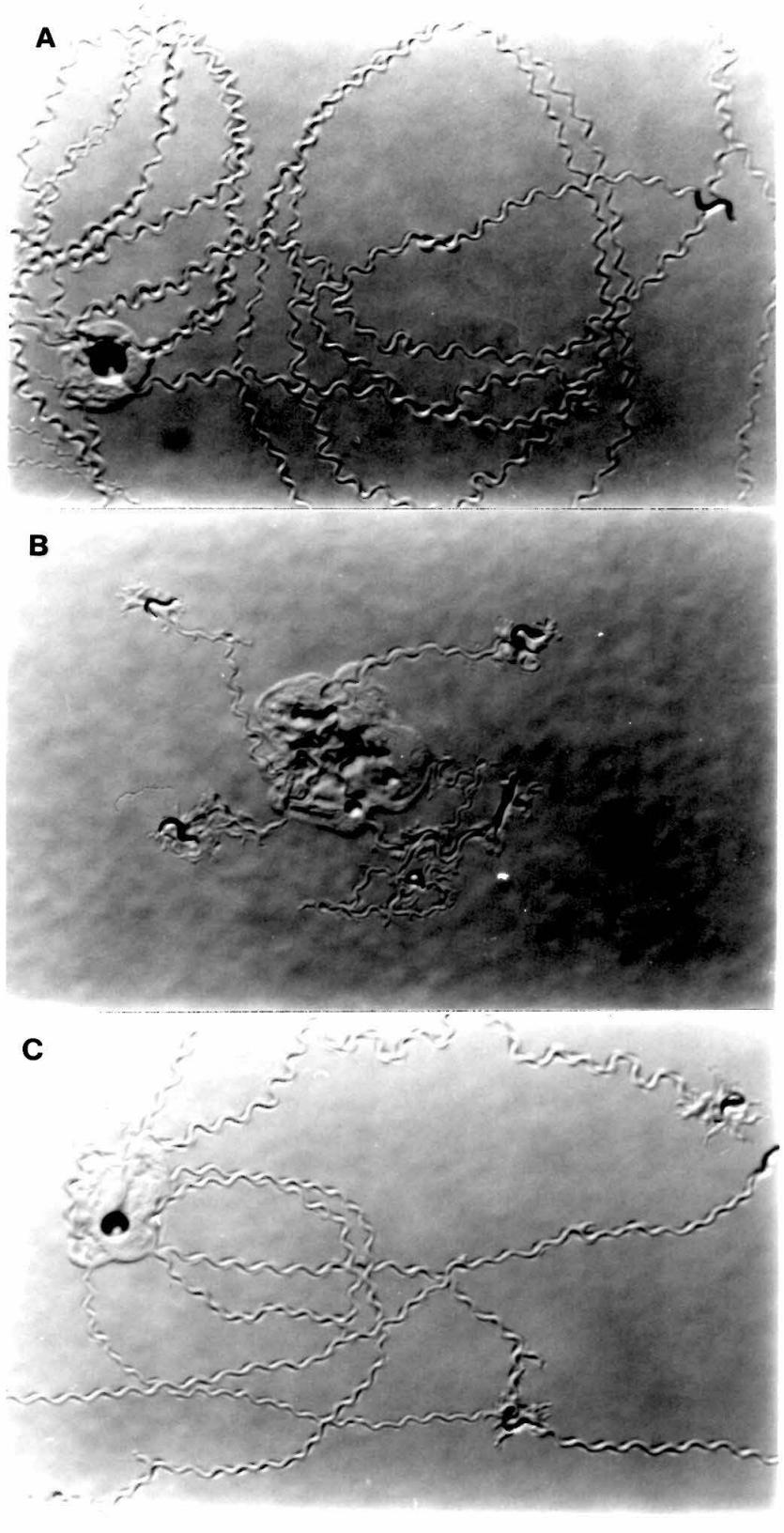
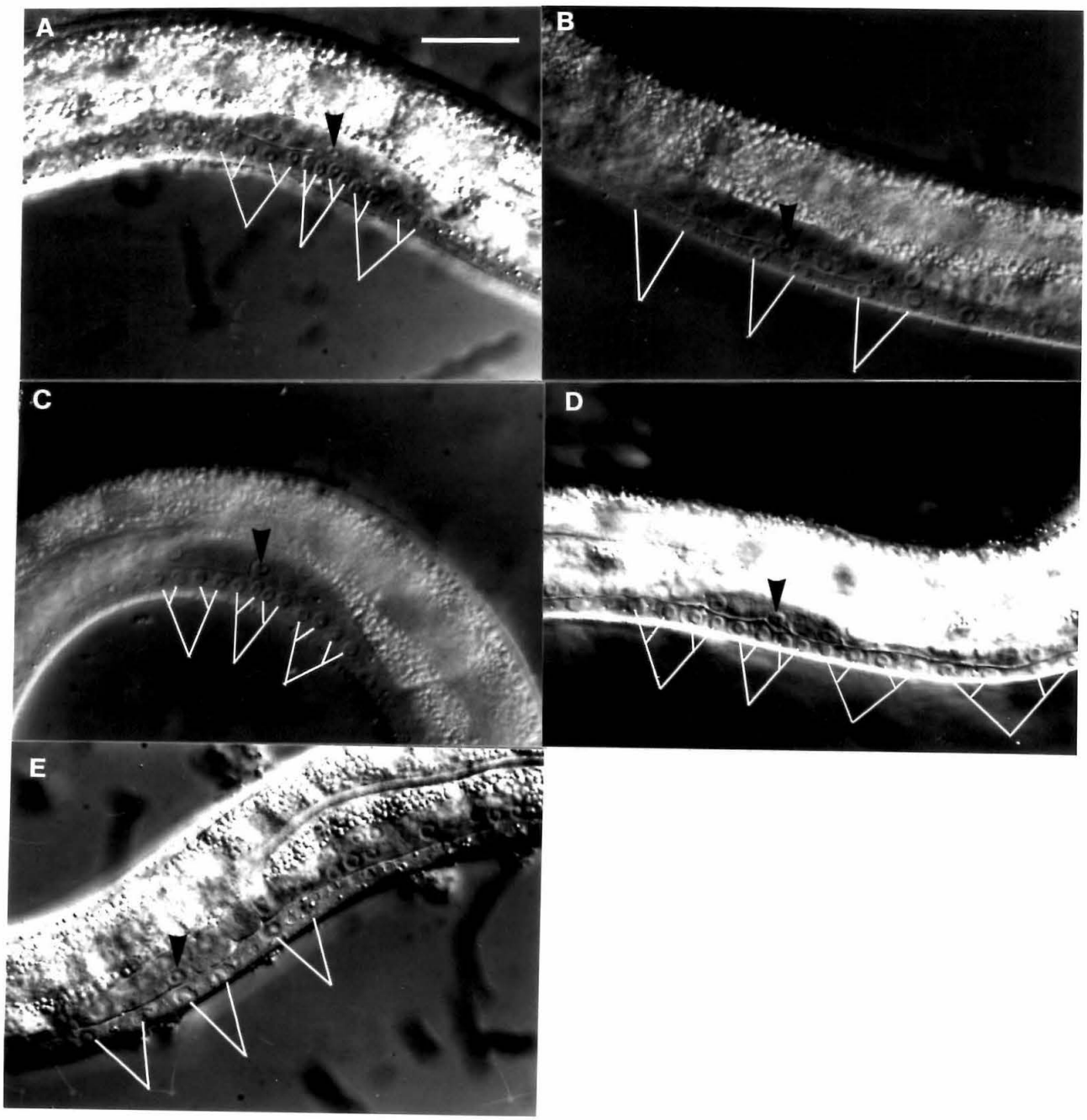

A
B
C

Figure 10. Rescue of suppression of the vulvaless phenotype of *let-23(sy1)* by the mouse-nematode hybrid gene.

The VPCs in the L3 molt stage, when the VPCs should have divided twice to generate four progeny, are shown. The triangles represent the anchor cell. A, N2, three VPCs each generated four progeny ; B, *let-23(sy1)*, no VPC divided to generate four progeny; C, *unc-101(sy108)*, same as N2; D, *unc-101(sy108); let-23(sy1),* four VPCs divided to generate four progeny; E, *unc-101(sy108); let-23(sy1); Ex[AP47 hybrid]*, no VPC divided to generate four progeny. The scale bar is 20 μm.


A
B
C
D
E

CHAPTER 3. ANALYSIS OF *ROK-1*, A NEGATIVE REGULATOR OF VULVAL DIFFERENTIATION

rok-1 is a negative regulator of vulval differentiation in the nematode *C. elegans*

Junho Lee and Paul W. Sternberg

HHMI, Division of Biology

California institute of Technology

(818) 395-2182

Fax: (818) 568-8012

ABSTRACT

We analyzed the *rok-1* (*r*egulator *o*f *k*inase) locus (IV) that was isolated in a genetic screen for new negative regulators of the vulval induction pathway of the nematode *C. elegans*. *rok-1(sy247)* causes essentially no phenotypes in an otherwise wild-type background. However, in combination with mutations in other negative regulator genes, *unc-101*, and *sli-1*, it causes subviability, a male spicule defect, and excessive vulval differentiation. Animals defective in all three negative regulatory genes have a greater extent of vulval differentiation than any single or double mutants. We propose that *rok- 1* defines a new redundant negative regulator in the vulval induction pathway. These genes display different redundancy patterns for vulval differentiation and viability.

Introduction

Vulval induction in the nematode *C. elegans* is an excellent genetic system for studying intercellular and intracellular signal transduction. The anchor cell (AC) in the gonad induces three out of six equipotential epidermal vulval precursor cells (VPCs) to generate vulval cells. Genes required for the induction of the VPCs have been genetically and molecularly studied and many of them have been shown to encode homologs of the proteins involved in the mammalian epidermal growth factor (EGF) signal transduction pathway (reviewed in Horvitz and Sternberg, 1991). *lin-3* encodes a protein with a single EGF motif on its extracellular region that is the induction signal produced by the gonadal AC (Hill, et al., in preparation, Hill and Sternberg, 1992). *let-23* encodes an EGF receptor homolog that is thought to be the receptor for the Lin-3 signal(Aroian, et al., 1990). *let-60, lin-45,* and *sem-5*, genes acting downstream of *let-23,* encode *ras, raf,* and GRB2 homologs, respectively (Clark, et al., 1992, Han, et al., 1993, Han and Sternberg, 1990).

The *C. elegans* system proves to be very useful in genetic studies of negative regulators of the EGF mediated signal transduction pathway. For example, several genetic loci have been characterized thus far as negative regulators of the vulval induction pathway, an EGF mediated signal transduction pathway. Mutations in the *lin-15* locus cause the activation of this pathway. *lin-15* products are thought to be acting in cells other than the AC or the VPCs, probably in the hypodermal hyp7 cell (Herman and Hedgecock, 1990). *lin-15* encodes two novel negative regulator proteins (L. Huang, P. Tzou and P. Sternberg, in prep.). Additional negative regulators of

vulval differentiation are defined by the genes *unc-101* and *sli-1*. *unc-101* and *sli-1* mutations were isolated as suppressors of *let-23* vulvaless mutations (G. Jongeward and P. Sternberg, in prep.). Mutations in either *unc-101* or *sli-1* suppress a vulvaless phenotype of certain reduction-of-function alleles of *let-23*. For example, double mutant animals of the genotype *unc-101(sy108); let-23(sy1)* or *let-23(sy1); sli-1(sy143)* display greater than wild-type vulval differentiation (hyperinduced phenotype; Hin), whereas *let-23(sy1)* single mutant animals have a vulvaless phenotype. Vulval differentiation in these doubly mutant animals is dependent on the presence of the gonad, unlike *lin-15* mutations which cause a multivulva (Muv) phenotype independent of the presence of the signal. The *unc-101* and *sli-1* loci are different from *lin-15* in that single mutants of *unc-101* or *sli-1* do not display any vulval defects; *lin-15* mutations confer a multivulva (Muv) phenotype where more than three VPCs are induced to generate vulval cells. More complicated is the fact that double mutants of *unc-101* and *sli-1* show a partially gonad-independent multivulva phenotype, suggesting that these two genes are partially redundant negative regulators of the vulval induction pathway (G. Jongeward and P. Sternberg, in prep.).

Molecular analysis of *unc-101* revealed that this gene encodes a homolog of mammalian AP47 protein, the medium chain of the *trans*- Golgi clathrin-associated protein complex (Lee et al., in prep., Chapter 2). *sli-1* encodes a homolog (30 % identical at amino acid level) of the proto-oncogene *c-cbl* (C. Yoon, G. Jongeward, J. Lee, P. Sternberg, unpublished results). SLI-1 and *c-cbl* proteins have a putative Ring-type Zn finger, possibly acting as a DNA binding protein. Molecular analyses of *unc-101* and *sli-1* suggest that the interaction of these two genes may not be direct and that more genes may

be involved in this interaction. In this study, we took advantage of the fact that *sli-1* mutations can cause a synthetic Muv phenotype with *unc-101* mutations. We performed genetic screens where we isolated hyperinduced or Muv mutants after mutagenizing *sli-1* mutant animals. Here we report the isolation and genetic analysis of *rok-1,* a new negative regulatory gene.

Materials and Methods

Strains and general methods

Methods for culturing and handling the nematode and general genetic methods were described by Brenner (Brenner, 1974). All genetic experiments were performed at 20 °C. The standard strain N2 was from Brenner (1974). The markers for linkage mapping and three factor crosses of *rok-1(sy247)* are as follows:

LG I: *dpy-5(e61)*

LG II: *dpy-10(e128)*

LG III: *dpy-17(e164)*

LG IV: *dpy-20(e1282) unc-22(s7), unc-31(e169), unc-24(e138), dpy-13(e184), dpy-4(e1166)*

LG V: *dpy-11(e224)*

LG X: *dpy-3(e27) unc-1(e719), sli-1(sy143)*

The alleles for examining genetic interactions of *rok-1(sy247)* are as follows: *let-23(sy1)* (Aroian and Sternberg, 1991) , *unc-101(sy108), sli-1(sy143),* and *lin-15(n744).*

Genetic screen and backcross

We mutagenized 10,000 F1 chromosome sets of *sli-1(sy143)* by ethyl methane sulfonate (EMS), and screened F2 progeny for mutants with more than wild-type vulval induction (hyperinduced; Hin), which is indicated by the presence of additional bump(s) on the ventral side of the animals which is visible under a dissecting microscope. From the first round of mutagenesis, we analyzed the candidate mutants, and found that *lin-2* and *lin-10* mutations had the most frequent recovery rates. Thereafter, we performed the complementation tests with *lin-2* and *lin-10* alleles as soon as we isolated candidate Hin animals to see if they are alleles of these genes. If any candidate complemented *lin-2* and *lin-10* mutations, we proceeded to further map the mutations. We performed backcrosses with either a wild-type (N2) or a *sli-1* strain (Figure 2). Three to four days after crossing N2 or *sli-1(sy143)* males with the Hin animals from the mutagenesis, we transferred many L4 hermaphrodites to a new plate, and the next day we transferred single wild-type animals to new plates. We scored F2 animals for the segregation of wild type, hyperinduced, and vulvaless "bag" phenotypes. If we observed 3/16 of the F2 progeny of N2 backcross were egg laying-defective (Egl) due to a vulvaless phenotype, we concluded that the new mutation is dependent on the presence of *sli-1* mutations and is an allele of a vulvaless gene unlinked to X. If we observed about the same number of F2 Hin animals (1/4) from both N2 and *sli-1* backcrosses and few or no vulvaless animals in N2 backcross, we concluded that either the mutation is dependent on the presence of *sli-1* and linked to X, or the mutation is independent of *sli-1*. If 1/4 of *sli-1* backcross F2 and 1/16 of N2 backcross F2 are Hin, and no F2 animals in N2 backcross are vulvaless, then this mutation is dependent on the presence of *sli-1*, and is not an allele of a vulvaless gene.

Genetic mapping of *rok-1(sy247)*

We used *marker; sli-1* double mutants to map *rok-1(sy247)* by linkage mapping, because *rok-1* needs the presence of *sli-1* mutations to show a Hin phenotype (see Results). We followed a standard mapping method as described in (Brenner, 1974). We then used the triple mutants of *dpy-13 unc-24; sli-1, unc-22 dpy-4;sli-1, unc-22 unc-31; sli-1,* and *unc-31 dpy-4; sli-1* strains to map *rok-1(sy247)* by three factor crosses. We mated N2 males with these marker strains, and the male progeny of genotype *marker/+; sli-1/ø* were mated into *rok-1(sy247); sli-1(sy143)* hermaphrodites. We isolated many L4 progeny from this mating, isolated recombinant animals in the next generation from the plates that segregated parental markers, and checked the segregation of the Hin phenotype in the next generation. We also performed complementation tests with deficiencies to find any deficiency that deletes *rok-1*. The deficiencies used were *sDf22* and *sDf60*.

Nomarski microscopy

We observed the extent of vulval differentiation using Nomarski optics as described in (Han and Sternberg, 1990). We observed the vulva differentiation of the animals in their L3 molt stage, when the induced VPCs should have divided twice to generate four daughter cells and uninduced VPCs have divided once to generate two daughter cells. The average number of induced VPCs were calculated as the total number of induced VPCs divided by the number of animals observed. The average number of induced VPCs is in wild-type animals is three, and that in Hin animals is greater than three. We observed the male spicules of N2, *rok-1(sy247),* and *rok-1(sy247); sli-*

1(sy143) males with the help of H. Chamberlin.

Microsurgery

To ablate the anchor cell, we performed microsurgery using a laser microbeam. The procedure for microsurgery was described in (Avery and Horvitz, 1987, Sulston and White, 1980). We ablated the Z1, Z2, Z3, and Z4 cells, which include the precursors of the germ line, the gonad and the anchor cell, in the very early L1 stage animals. When the ablated animals are in their L3 molt, we observed their vulval differentiation using Nomarski optics as described above.

Examination of viability of mutant animals

After transferring eggs from a culture plate to a new plate, we counted the number of adults after three days. The viability of a mutant strain is calculated as follows.

viability (%) = number of adult animals / number of eggs x 100

lethality (%) = 100 - viability (%)

We used N2 wild-type strain as a control strain for viability. The viability of N2 is theoretically 100 %, but due to damage during egg transfer, or animals that escape from the plate, the viability of N2 was calculated as 94 %.

Strain construction

For *marker; sli-1* strains, we mated *marker* / + males with *sli-1(sy143)* hermaphrodites, and obtained *marker* or + / + ; *sli-1* / ø males. These males were mated with *sli-1* hermaphrodites, and their hermaphrodite progeny were transferred to individual plates and checked for segregation of marker

phenotypes. The animals with the marker phenotype are of the genotype *marker; sli-1*.

For *sy247* single mutant construction, we used an *unc-31 dpy-4* chromosome to balance *sy247*, as it maps within 0.2 m.u. of *unc-31*. Single males from the mating between *unc-31 dpy-4* / + males and *sy247; sli-1(sy143)* hermaphrodites were mated with *unc-31 dpy-4* hermaphrodites. Males of the genotype *unc-31 dpy-4* / *sy247; sli-1* / ø will generate Unc Dpy male cross progeny from this mating, and males of the genotype *unc-31 dpy-4* / +; *sli-1* / ø will not generate any Unc Dpy male cross progeny. From a mating plate with Unc Dpy males, we picked wild-type males whose genotype should be *unc-31 dpy-4* / *sy247; sli-1* / ø , and mated them with *unc-31 dpy-4* hermaphrodites. We then mated the wild-type male progeny of this mating, whose genotype should be *unc-31 dpy-4* / *sy247;* + / ø , with *unc-31 dpy-4* hermaphrodites. Wild-type cross progeny of from this mating have the genotype of *unc-31 dpy-4* / *sy247;* + / + . Among the progeny of these mothers, we picked many wild-type L4 animals to individual plates. Animals that did not segregate any Unc, Dpy, or Unc Dpy animals were of the genotype *sy247;* +. We confirmed the presence of *sy247* in this final strain by mating *sy247* hermaphrodites with *sli-1* males and checking segregation of Hin animals in the F2.

For *unc-101; sy247; sli-1* triple mutants , from the mating of *unc-101* / + males with *sy247; sli-1* hermaphrodites, we isolated *unc-101 or* + / +; *sy247* / +; *sli-1* / + progeny. In the next generation, we selected nonUnc Hin animals from the plate with Unc segregants, and among the progeny of these Hin animals, we isolated Unc animals whose genotype should be *unc-101; sy247; sli-1*.

For *unc-101; sy247* double mutants, we isolated nonUnc progeny from the mating of *sy247* males with *unc-101* hermaphrodites, whose genotype is *unc-101/+; sy247/+*. We observed the Unc progeny of these mothers by Nomarski optics, and found that some of them are Hin. These Hin Unc animals are of the genotype of *unc-101; sy247*.

For *let-23(sy1); rok-1(sy247)* double mutants, we first constructed *sy1; unc-31 dpy-4* strains. From the mating of *sy247* males with *sy1; unc-31 dpy-4* hermaphrodites, we isolated cross-progeny whose genotype is *sy1/+; sy247/unc-31 dpy-4*. We picked many individual non-Unc-31 non-Dpy-4 Vul animals in the next generation, whose genotype could be one of the two classes: *sy1/sy1; sy247/sy247* or *sy1/sy1; sy247/unc-31 dpy-4,* depending on whether sy247 suppresses the vulvaless phenotype. If *sy247* does not suppress the vulvaless phenotype to wild type or Hin, 1/3 of the non-Unc-31 non-Dpy-4 Vul animals would be of the first class, which is the wanted genotype, and segregate no Dpy Unc animals. The other 2/3 of the non-Unc-31 non-Dpy-4 Vul animals would be of the second class, and segregate Dpy Unc animals. If *sy247* does suppress the vulvaless phenotype, then all the non-Unc-31 non-Dpy-4 Vul animals would be of the second class, and segregate 1/4 of Unc Dpy Vul animals, 2/4 of non-Unc non-Dpy Vul animals that segregate Unc Dpy progeny, and 1/4 of non-Unc non-Dpy non-Vul animals that do not segregate any Unc Dpy animals. Animals of the last category are of the genotype *sy1; sy247*.

For *sy247; lin-15 (A)* double mutants, we used *unc-31 dpy-4; lin-15 (A or B)* strains in the construction. We mated the *sy247* males with hermaphrodites of the genotype *unc-31 dpy-4; lin-15(n744).* The male progeny from this mating, whose genotype should be *unc-31 dpy-4/ sy247;*

n744/ø, were mated again with hermaphrodites of the genotype *unc-31 dpy-4; lin-15(n744).* The non-Unc non-Dpy hermaphrodite cross-progeny from this mating are of the genotype *unc-31 dpy-4 / sy247; n744/n744.* Among the progeny of these hermaphrodites, the non-Unc non-Dpy animals that do not segregate any Unc Dpy progeny were of the genotype *sy247/sy247; n744/n744*.

Results

Recovery of *rok-1(sy247)* as a synthetic Hin with *sli-1.*

To identify more negative regulators of vulval differentiation, we performed a genetic screen in which we mutagenized *sli-1(sy143)* animals, and looked for animals with greater than wild-type vulval differentiation. Since *sli-1(sy143)* suppresses the vulvaless phenotype of *lin-2, lin-7, lin-10,* and *let-23* mutations to a Hin phenotype, and causes a multivulva phenotype with *unc-101* mutations, we expected to recover alleles of these genes from our mutagenesis as well as mutations in new genes. Also we expected to recover mutations in any gene downstream in the vulval induction pathway that cause a multivulva phenotype, for example, *lin-1*. Indeed, we recovered four *lin-2* alleles, five *lin-10* alleles, one *unc-101* allele, and one *lin-1* allele as well as one new mutation, *sy247*.

sy247 causes a Hin phenotype in the presence of a *sli-1* mutation. By Nomarski optics, the average number of VPCs induced in the *sy247; sli-1* double mutants is 3.8 VPCs, compared with 3.0 of N2 wild type or *sli-1* single mutants. A gonad ablation experiment showed that vulval induction in these

Hin animals is still dependent on the presence of the gonad; in gonad-ablated animals, no vulval induction was observed (n=4). This gonad-dependency is different from that of *unc-101; sli-1* double mutants, since the *unc-101; sli-1* double mutants still display some vulval induction after gonad ablation (G. Jongeward and P. Sternberg, in prep.).

We mapped *sy247* using linkage mapping and three factor crosses (described in Materials and Methods; Figure 4, Table 1) to the left of *unc-22* and to the right of *dpy-4* on the chromosome IV. *rok-1* maps very close to *unc-31,* because from the three factor crosses using *unc-22 unc-31; sli-1* and *unc-31 dpy-4; sli-1*, *rok-1* was inseparable from the *unc-31* locus. None of 3 Unc-31 non-Unc-22 recombinants from *unc-22 unc-31/ rok-1* heterozygous mothers, and none of 27 Unc-31 nonDpy-4 recombinants from *unc-31 dpy-4/ rok-1* mothers segregated Hin progeny. We cannot tell whether *rok-1* is to the left or to the right of *unc-31* at this point. We found that the deficiencies *sDf22* and *sDf60* delete the *rok-1* locus (see the following result).

The *sy247* mutation reduces the *rok-1* activity.

We compared the extent of vulval differentiation in *rok-1(sy247) / rok-1(sy247); sli-1(sy143)/ sli-1(sy143)* animals with that in *sDf22/ rok-1(sy247) ; sli-1(sy143)/ sli-1(sy143)* animals. *sDf22/nT1* heat-shocked males were mated with *dpy-13 unc-24;sli-1* hermaphrodites, and individual male progeny, whose genotype are either *sDf22/ dpy-13 unc-24* or *nT1/ dpy-13 unc-24*; *sli-1/ø* , from this mating were mated with *dpy-13 sy247; sli-1* hermaphrodites. Non-Dpy cross-progeny of each mating, whose genotype is either *nT1/ dpy-13 sy247; sli-1* or *sDf22/ dpy-13 sy247; sli-1,* were examined

for their vulval induction. From one plate, 9 nonDpy cross-progeny were obtained, and they showed an average of 3.8 VPCs induced, which is virtually indistinguishable from *rok-1(sy247) / rok-1(sy247); sli-1(sy143) / sli-1(sy143)* animals. Moreover, four of nine animals had more than three VPCs induced (44 %), similar to *rok-1(sy247) / rok-1(sy247); sli-1(sy143) / sli-1(sy143)* animals. Therefore, the vulval phenotype of *rok-1(sy247) / rok-1(sy247); sli-1(sy143) / sli-1(sy143)* animals is as severe as that of *sDf22 / rok-1(sy247) ; sli-1(sy143) / sli-1(sy143)* animals. Another plate had three nonDpy progeny with wild-type vulval differentiation, indicating that the genotype of these animals was *nT1 / dpy- 13 sy247; sli-1.*

We also measured the lethality of the heterozygotes of *sDf22 / rok-1(sy247); sli-1(sy143)* compared with that of *rok-1(sy247) / rok-1(sy247); sli-1(sy143)* animals. We scored the heterozygotes of the genotype *sDf22 / dpy-13 rok-1(sy247); sli-1* for their segregation of Dpy and nonDpy progeny. From the heterozygotes of the genotype *sDf22 / dpy-13 rok-1(sy247); sli-1*, 271 animals were Dpy, and 139 animals, nonDpy. The ratio of Dpy animals to nonDpy animals is 2:1. If the lethality of heterozygotes of the genotype *sDf22 / dpy-13 rok-1(sy247); sli-1* had been enhanced, the ratio of Dpy/nonDpy would have been less than 2. Therefore, the homozygotes of the genotype *dpy-13 rok-1(sy247) / dpy-13 rok-1(sy247); sli-1(sy143)* has as severe a lethality phenotype as the heterozygotes of the genotype *sDf22 / dpy-13 rok-1(sy247); sli-1* We conclude that the *sy247* mutation reduces the *rok-1* activity.

_sy247_ causes additional phenotypes in the presence of a _sli-1_ mutation

We constructed *sy247; him-5(e1490); sli-1(sy143)* triple mutants to examine the male phenotype of *sy247* in the presence of *sli-1* mutations. Under Nomarski optics, the males display an abnormal spicule structure (Figure 3E, marked with an arrowhead), and the mating capability is abolished in these males (data not shown).

The double mutants of *sy247; sli-1* genotype also show partial lethality, as 55.6 % of the progeny of the homozygous mothers died as embryos or early larvae (n=261).

_sy247_ is a silent mutation.

We constructed *sy247* single mutants as described in Materials and Methods, and found that they display wild-type vulval induction. We obtained *sy247/ sy247* males by heatshock at 30° C for six hours, and found that they had wild-type spicules and normal mating capability. The lethality is essentially not obvious in this single mutant animals, because the viability was 91 % in a viability test, which is similar to the N2 viability (94%) (Table 3). Therefore, *sy247* mutation alone does not cause any visible phenotype.

Interaction with other negative regulators (1): vulval induction

So far, there have been four redundant negative regulators of the vulval induction pathway: *lin-15(A), lin-15(B), sli-1,* and *unc-101*. To examine how *sy247* interact with other negative regulators, we constructed double mutants or triple mutants containing mutations in these negative regulatory genes. The extent of vulval differentiation of these mutants are summarized in Table 2. *unc-101; sy247* double mutants have an average of 3.5 VPCs induced (n=19). As described above, *rok-1; sli-1* double mutants

have an average of 3.8 VPCs induced. *unc-101; sli-1* double mutants show an average of 3.7 VPCs induced (G. Jongeward, per. comm.). Interestingly, *unc-101; sy247; sli-1* triple mutants show an even greater extent of vulval differentiation (an average of 4.9 VPCs) than any of the three double mutant combinations. When comparing the percentage of Hin animals in these mutants, the enhanced phenotype becomes more obvious. In any single mutants, no animal shows greater than wild-type induction. In double mutants, 47 to 65 % of animals in the populations show greater than wild-type induction, and in the triple mutants, 100 % of animals in the population show greater than wild-type induction. Therefore, *rok-1* defines a new redundant negative regulator of vulval induction, acting together with *unc-101* and *sli-1*.

We examined whether vulval differentiation in the double and triple mutants of *unc-101, sli-1,* and *rok-1* is dependent on the presence of the anchor cell. We ablated the precursor cells of the gonad and the anchor cell as described in Materials and Methods, and observed the animals at their L3 molt stages to examine the extent of vulval differentiation (Table 1). In the double mutants of *rok-1; sli-1* and *rok-1; unc-101* , none of the ablated animals displayed vulval differentiation (n=4, and 10, respectively). In the ablated triple mutants of *unc-101; rok-1; sli-1*, the average number of VPCs induced was 2.3 VPCs per animal (n=9), which is virtually identical to that of *unc-101; sli-1* double mutants (G. Jongeward, per. comm.).

The *lin-15* locus contains two functional groups, class A and class B. Certain mutations of *lin-15* abolish functions of both the class A and class B, resulting in a multivulva phenotype. Certain mutations in *lin-15*, however, abolish either class A or class B function, resulting in no vulval

differentiation defect. To display a multivulva phenotype, both class A and class B functions must be mutated. No double mutants of the same classes cause vulval defects. *sli-1* can act as a class B multivulva gene at 25°C, as double mutants of genotype *lin-15(A); sli-1(sy143)* display a gonad-independent Muv phenotype at 25°C (Jongeward and Sternberg, 1993). To examine whether *rok-1* can interact with a *lin-15(B)* gene, we constructed a double mutant strain of *rok-1(sy247); lin-15(B)* . *lin-15(n744)* is a class B gene. The average number of induced VPCs of the double mutant animals was 3 at 20°C, and 3.2 (n=21) at 25°C. Thus, *rok-1* can act as a class A gene at 25°C.

Interaction with other negative regulators (2): lethality

We also compared the lethality associated with single, double, and triple mutants (Table 3) The *sli-1* and *rok-1* mutations do not cause any lethality alone, whereas *unc-101* mutations result in about 45 % lethality: 45 % of the progeny of *unc-101* homozygous mother die. As described above, 56 % of *rok-1; sli-1* double mutants die before adulthood. Double mutants defective in both *unc-101* and *sli-1* show 54% lethality, which is not very different from *unc-101* single mutants. Therefore, *sli-1* mutations do not enhance the lethality caused by *unc-101* mutations. On the contrary, double mutants of *unc-101; rok-1* show enhanced lethality: 82 % lethality (n= 185). Therefore, a *rok-1* mutation enhances the lethality caused by *unc-101* mutations. The lethality of the triple mutants was about 79 % (n= 109), which is about the same as that of *unc-101; rok-1* double mutants.

rok-1(sy247)* suppresses the vulvaless phenotype of a *let-23(sy1)

mutation.

sli-1 was identified by its mutations that suppress the vulvaless phenotype of *let-23* mutations. To test whether *rok-1* can also suppress a mutation in *let-23* gene, an EGFR homolog, we constructed a *let-23(sy1); rok-1(sy247)* double mutant strain as described in Materials and Methods. *sy1* is an allele that is suppressed by *unc-101* or *sli-1* mutations from an average of 1 VPC induced per animal to 3.5 VPCs per animal. The *let-23(sy1); rok-1(sy247)* double mutants displayed an average of 2.3 VPCs induced per animal. Therefore, *rok-1(sy247)* can suppress the *let-23(sy1)* mutation, though not as well as the other two negative regulator mutations.

Discussion

rok-1 defines a new negative regulator of the vulval induction pathway.

Based on the results, we propose that *sy247* defines a fifth negative regulator of the vulval induction pathway. First, the triple mutants of *unc-101; rok-1; sli-1* display a greater extent of vulval differentiation than any double mutant. In these triple mutants, all the animals observed under Nomarski optics showed greater than wild-type differentiation, while in double mutants some animals showed wild-type induction, and some, greater than wild-type differentiation. In the triple mutants, the sensitivity of VPCs are greater than VPCs in the double mutants, where they show fluctuations of induction level in individual animals. Second, *sy247* mutation suppresses the vulvaless phenotype of *let-23(sy1)*, though not as well as the other two negative regulators. Further analysis of the interactions of *rok-1* and other

genes in the vulval induction pathway will help to understand the roles of the *rok-1* gene.

The ablation experiment showed that the double mutants of *unc-101; sli-1* and the triple mutants of *unc-101; rok-1; sli-1* displayed some extent of vulval differentiation (~2 VPCs per animal) after ablation of the entire gonad, whereas the double mutants of *unc-101; rok-1* and *rok-1; sli-1* did not display any vulval differentiation after ablation. This suggests that the combination of the *unc-101* and *sli-1* gene activity may negatively regulate the basal activity of the vulval induction pathway as well as the activated activity of the pathway, while *rok-1* does not seem to be involved in negative regulation of the basal activity. The *rok-1* activity may be involved in negative regulation of the activity of the vulval induction pathway only after activation by the signal.

In addition to *unc-101, rok-1*, and *sli-1*, there are other negative regulator genes in the vulval induction pathway. *lin-15* is a locus consisting of two sub-complementation groups, namely class A and class B(Ferguson and Horvitz, 1989). Mutations in both the classes of the sub-complementation groups cause a gonad-independent multivulva phenotype, while mutations in one group do not cause any vulval defect. There are other genes that act as either class A or class B groups. Any mutations in class A genes can cause a multivulva phenotype in combination with any mutations in class B genes. Thus, there are at least five different pathways for the negative regulation for proper vulval induction in the nematode. A combination of activities of all these negative regulators will ensure that only three VPCs out of six VPCs, which all have the potentials to generate vulval cells, are induced by the LIN-3 signal. It is difficult at this point to explain

how *rok-1* and *sli-1* can act as a class A and class B genes, respectively, at 25°C, while *unc-101* does not act as class A or class B gene.

The fact that any single mutation in these five classes of negative regulator genes does not cause excessive vulval differentiation alone indicates that the activities of these genes have some redundant functions to make sure this system works properly. This seems to be true for the cells in the higher organisms. In the carcinogenesis process in humans, any single mutation hardly causes cancerous cells. Instead, the cancer cells undergo many events of mutations in a variety of regulating genes including protooncogenes and tumor suppressor genes. In higher organisms, activation of an EGFR pathway by oncogenic mutations is one of the causes of cancer. Similarly, in *C. elegans,* activation of an EGFR pathway by gain-of-function mutations of *lin-3* (EGF homolog), *let-23* (EGFR homolog), and *let-60* (ras homolog) result in excessive vulval differentiation. Thus, these mutations can be considered 'oncogenic' to some extent. It is possible that the negative regulatory genes in vulval induction may help to define novel tumor suppressor genes.

Three negative regulator genes show differential redundancy for different functions.

As described in the results, all three double mutant combinations of *unc-101, rok-1,* and *sli-1* are hyperinduced, and the triple mutants show an even higher hyperinduction, suggesting that the activity of each gene is partially independent of the others so that the loss of activities is additive. On the contrary, the lethality associated with *unc-101* single mutants or *unc-101; rok-1* double mutants is not enhanced by the presence of a *sli-1* mutation. A

sli-1 mutation does enhance the lethality of a *rok-1* mutation, suggesting that activity of *sli-1* for survival may be independent of that of *rok-1,* but may be overlapping with that of *unc-101*. The causes of the lethality associated with *sli-1* and *unc-101* mutations and that with a *rok-1* mutation seem to be independent: if these genes are in the same pathway for survival, the lethality should not be increased by adding a *rok-1* mutation. Therefore, these three genes show different redundant aspects in different gene activities. A possible reason for the different redundancies is that different cellular processes may have different sensitivities to a loss of activities of these three genes. In another simple model, three activities are required for both prevention of excessive vulval induction and survival of animals, and two of the activities may be sufficient for proper function of these systems. In vulval induction, all three genes may act as different activities. Then any double mutants would be defective in two of the three activities, causing a greater than wild-type vulval differentiation. For survival of an animal, *unc-101* may be involved in two of the activities, *rok-1* in a third, and *sli-1* in one of the activities *unc-101* is involved in. Then, *unc-101* single mutants would show some lethality, while the other two single mutants would not.

Figure 1. The *sli-1* mutagenesis. (A) Schematic of the mutagenesis procedure. This scheme is for isolating recessive mutations that cause hyperinduction of VPCs in the presence of a *sli-1* mutation. *sli-1(sy143)* animals are mutagenized with EMS as chemical mutagen. Mutated chromosomes exist as heterozygotes in the F1 progeny, and show phenotypes in the F2. EMS = ethyl methane sulfonate, Hin = hyperinduced, WT = wild type, and m= a new mutation.

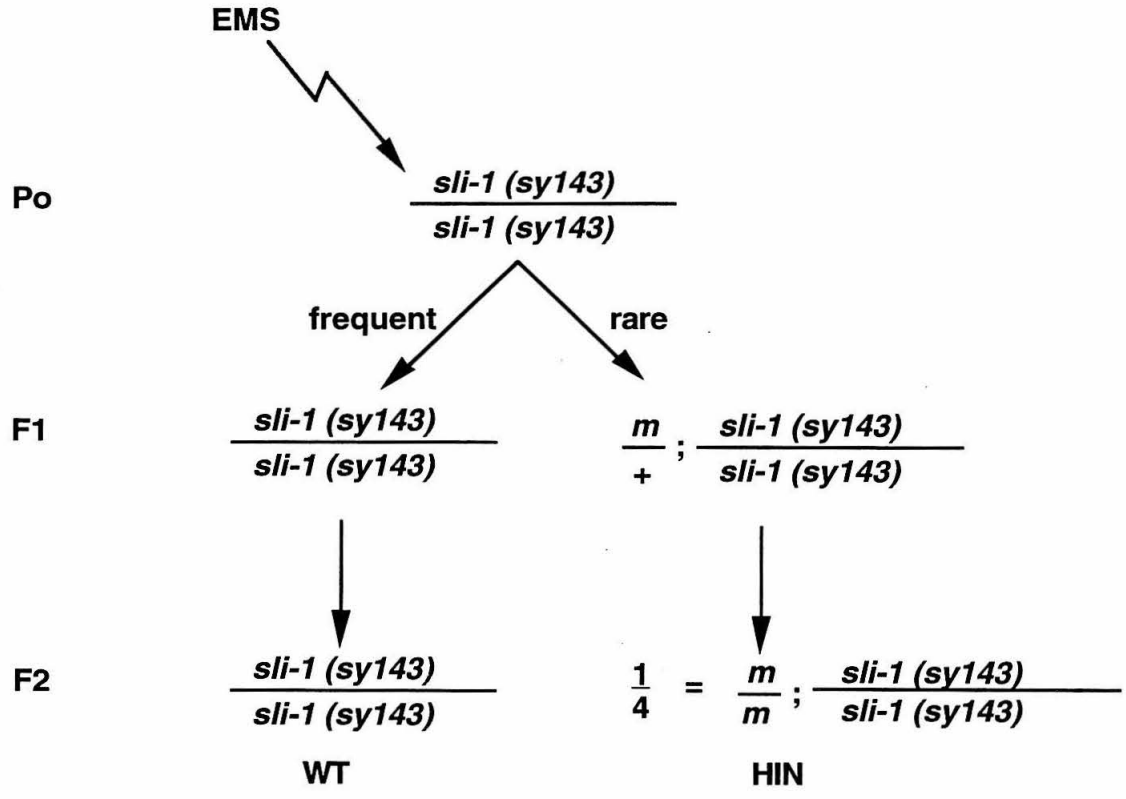

EMS
Po
sli-1 (sy143)
sli-1 (sy143)
frequent
rare
F1
sli-1 (sy143)
sli-1 (sy143)
m
+
sli-1 (sy143)
sli-1 (sy143)
F2
sli-1 (sy143)
sli-1 (sy143)
1
4
=
m
m
sli-1 (sy143)
sli-1 (sy143)
WT
HIN

Figure 1(B). Results of the *sli-1* mutagenesis. *lin-2* and *lin-10* mutations were the most frequent mutations recovered. One allele of *lin-1* and *unc-101* were recovered in addition to a mutation in a new gene, *rok-1*. The recovery frequency for the genes are shown.

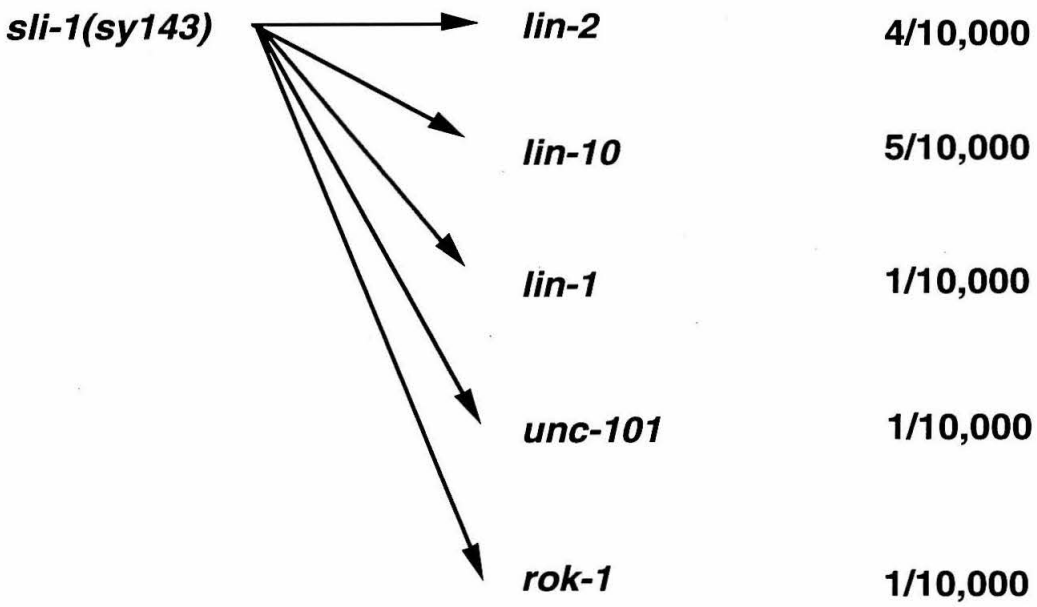

sli-1(sy143)
lin-2
4/10,000
lin-10
5/10,000
lin-1
1/10,000
unc-101
1/10,000
rok-1
1/10,000

Figure 2. Backcross of new mutations

If 3/16 of the F2 progeny of an N2 backcross are Egl, while 1/4 of the F2 of a *sli-1* backcross are Hin, the new mutation is an allele of a vulvaless gene unlinked to X that is dependent on the presence of a *sli-1* mutation for the Hin phenotype. If there are about 1/4 of Hin animals in the F2 progeny of both N2 and *sli-1* backcrosses and few or no vulvaless animals in an N2 backcross, it can either be that the mutation is dependent on the presence of *sli-1* and linked to X, or that the mutation is independent of *sli-1*. If 1/4 of the F2 progeny of a *sli-1* backcross and 1/16 of the F2 progeny of an N2 backcross are Hin, and there is no vulvaless animals in the F2 progeny of the N2 backcross , then this mutation is dependent on the presence of *sli-1*, and is not an allele of a vulvaless gene. N2 is a wild-type strain of *C. elegans*; *m,* a new mutation; Hin, hyperinduced phenotype of VPCs; WT, wild type; Vul, a vulvaless phenotype.

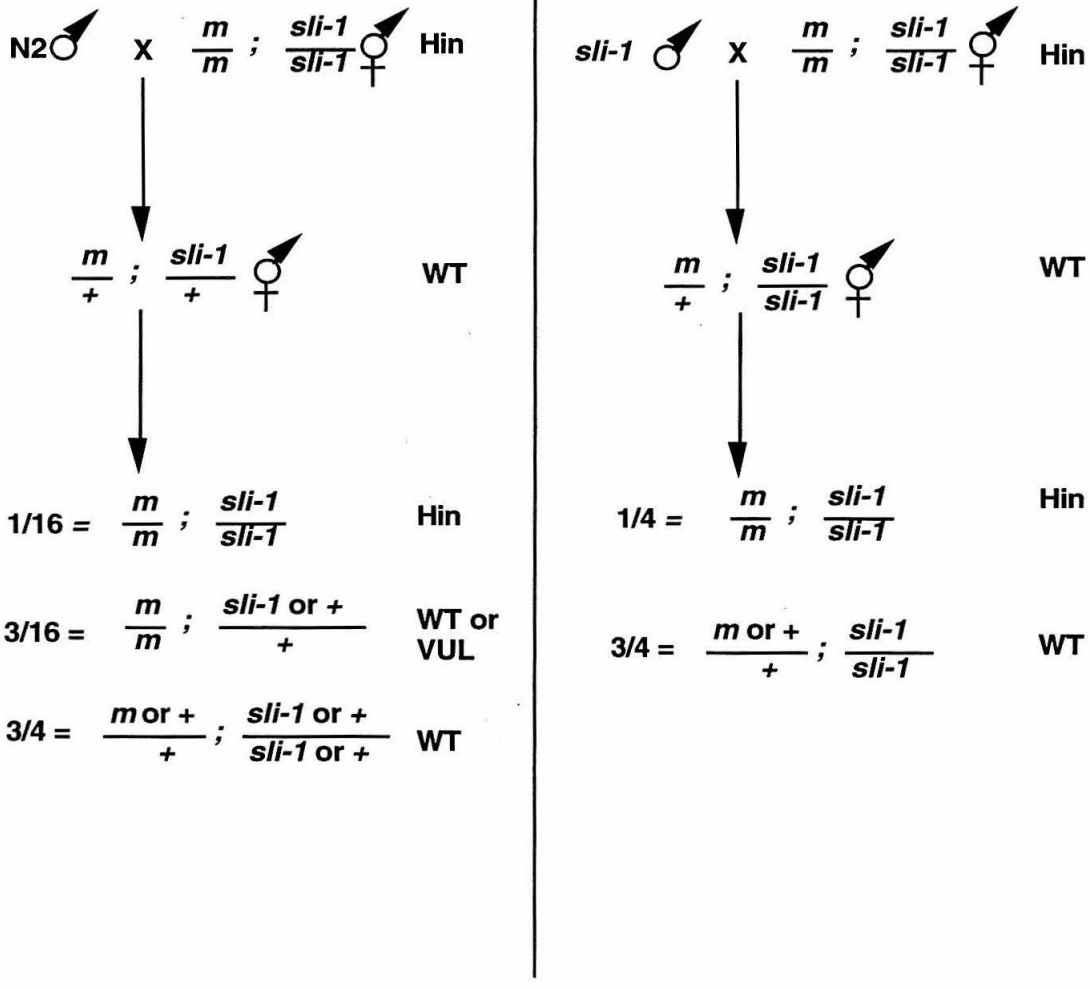
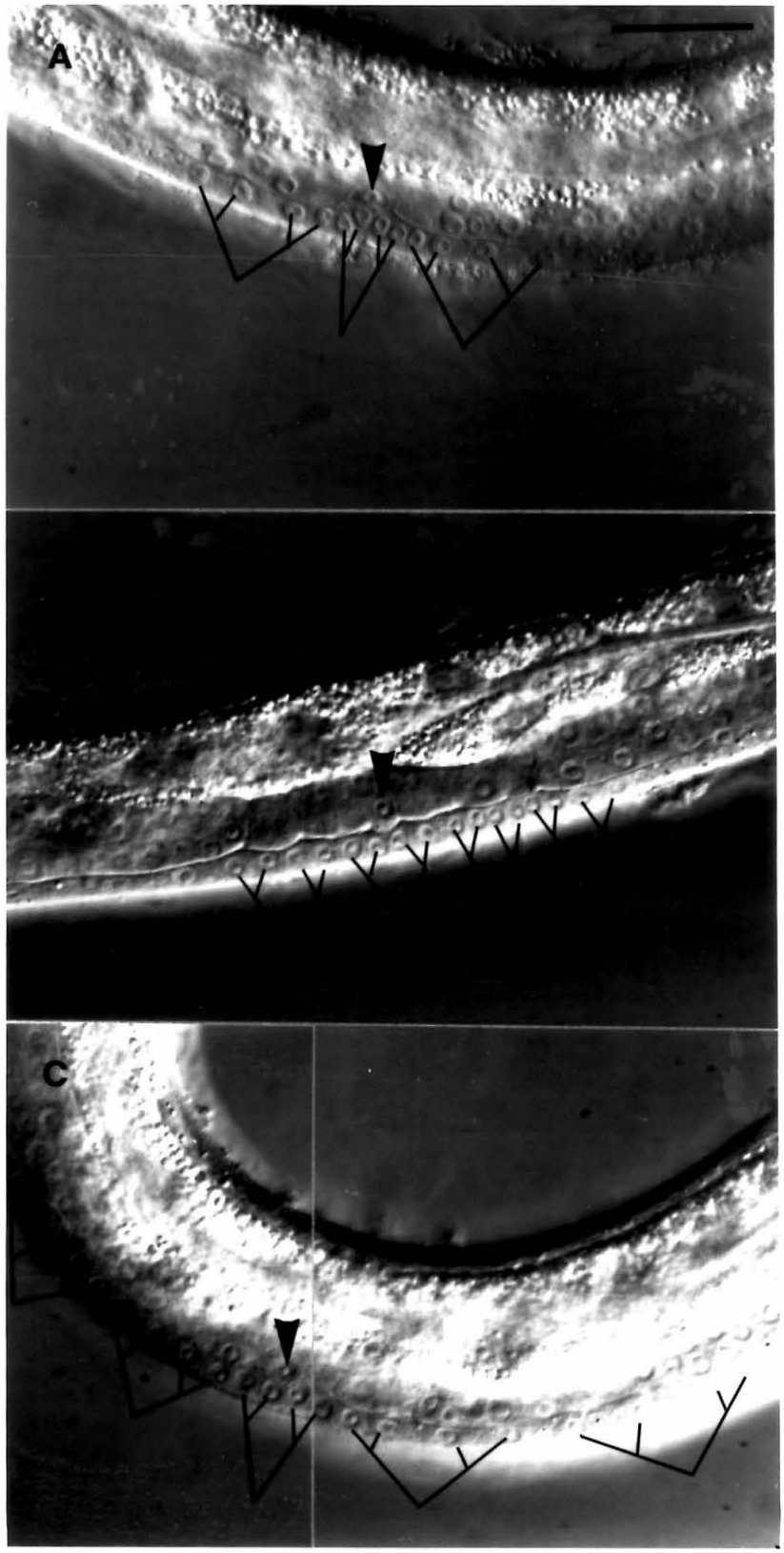

N2 ♂ X m/m ; sli-1/sli-1 ♀ Hin
m/+ ; sli-1/+ ♀ WT
1/16 = m/m ; sli-1/sli-1 Hin
3/16 = m/m ; sli-1 or + / + WT or VUL
3/4 = m or + / + ; sli-1 or + / sli-1 or + WT
sli-1 ♂ X m/m ; sli-1/sli-1 ♀ Hin
m/+ ; sli-1/sli-1 ♀ WT
1/4 = m/m ; sli-1/sli-1 Hin
3/4 = m or + / + ; sli-1/sli-1 WT

Fig 3 Phenotype of *rok-1(sy247)*. (A) Vulval induction of a *rok-1* single mutant animal with three VPCs induced to generate four daughters , (B) a *rok-1; sli-1* double mutant animal with four VPCs induced to generate four daughters, and (C) an *unc-101; rok-1; sli-1* triple mutant with five VPCs induced. These photos show only typical examples of vulval differentiation in the double and triple mutants. The extent of vulval differentiation varies in individual animals. The anchor cell is indicated by an arrowhead.
(D, E) Male spicule of a wild-type(D) and a *rok-1; him-5; sli-1* mutant animal(E). The *rok-1; him-5; sli-1* mutant animal has an abnormal spicule structure as indicated by an arrowhead.


A
C

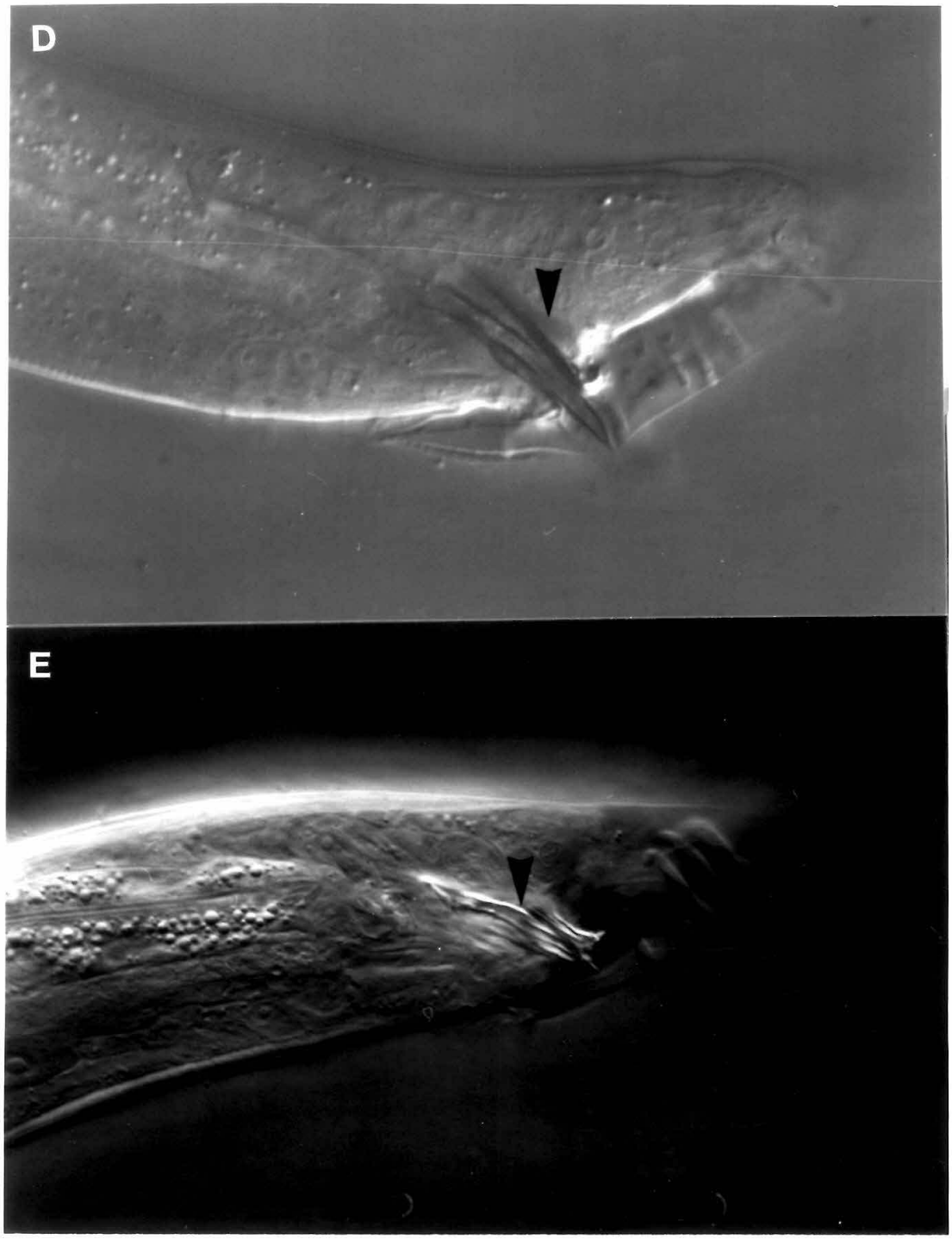

D
E

Figure 4. Genetic map of *rok-1*. *rok-1* is mapped to chromosome IV near *unc-31*. The scale bar is 1 map unit. The deficiencies, *sDf22* and *sDf60*, delete the *rok-1* locus.

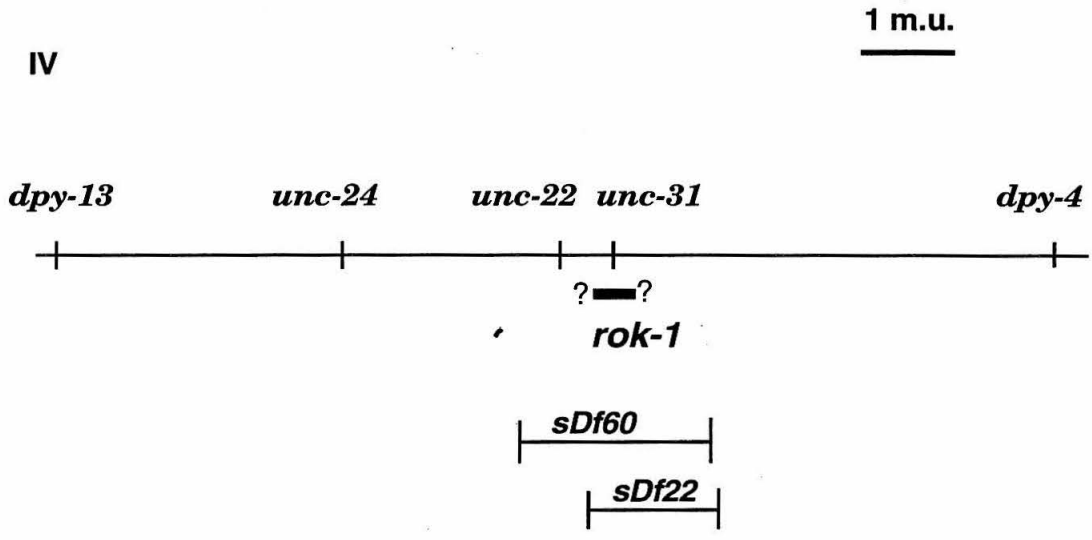

1 m.u.
IV
dpy-13
unc-24
unc-22
unc-31
dpy-4
?
?
rok-1
sDf60
sDf22

Table 1. Genetic mapping of *rok-1(sy247)*: Three factor crosses. See the material and methods for detail. Dpy is a dumpy phenotype; nDpy, non-Dpy; Unc, uncoordinated; nUnc, non-uncoordinated.

Genotype of heterozygotes	Phenotype of recombinants selected	Number of recombinants segregating of Hin phenotype
dpy-13 unc-24 / ***rok-1***	**Dpy nUnc**	**3/3**
unc-22 unc-31 / ***rok-1***	**Unc-31 nUnc-22**	**0/3**
unc-31 dpy-4 / ***rok-1***	**Unc nDPy**	**0/27**
unc-22 dpy-4 / ***rok-1***	**Dpy nUnc**	**4/5**

Table 2 Vulval differentiation of single, double, and triple mutants carrying *unc-101, rok-1,* or *sli-1*. Vulval induction of each animal was examined using Nomarski optics. For intact animals, Average numbers of induced VPCs per animal, and percentage of the animals having more than three induced VPCs are also shown. Wild-type vulval induction is three VPCs per animal. The maximum number of VPCs that can be induced is six. For the gonad-ablated animals, the average numbers of induced VPCs per animal after ablation are shown. The numbers shown in the parentheses are the numbers of animals examined. Data with asterisks are from Jongeward and Sternberg, (1993).

Genotype	Intact		Gonad-ablated
	Average # of induced VPCs	% animals with >3 VPCs induced	Average # of induced VPCs
+	3	0 (>20)	0 (>10)
sli-1	3	0 (>20)	0 (6)*
rok-1	3	0 (>20)	0 (6)
unc-101	3	0 (>20)	0 (6)*
rok-1; sli-1	3.8	64 (22)	0 (4)
unc-101; sli-1	3.8	65 (20)	1.9 (7)*
unc-101; rok-1	3.5	47 (19)	0 (10)
unc-101; rok-1; sli-1	4.9	100 (13)	2 (9)

Table 3. Viability table of single, double, and triple mutants carrying *unc-101, rok-1,* or *sli-1*. Viability was calculated as number of adult animals divided by number of eggs picked for examination.

Genotype	viability % (number of animals examined)
+	94 (96)
sli-1	92 (278)
rok-1	91 (45)
unc-101	55 (521)
rok-1; sli-1	44 (261)
unc-101; sli-1	46 (134)
unc-101; rok-1	18 (185)
unc-101; rok-1; sli-1	21 (109)

References

Aroian, R. V., Koga, M., Mendel, J. E., Ohshima, Y. and Sternberg, P. W. (1990). The *let-23* gene necessary for *Caenorhabditis elegans* vulval induction encodes a tyrosine kinase of the EGF receptor subfamily *Nature* **348**, 693-699.

Aroian, R. V. and Sternberg, P. W. (1991). Multiple functions of *let-23*, a *C. elegans* receptor tyrosine kinase gene required for vulval induction *Genetics* **128**, 251-267.

Avery, L. and Horvitz, H. R. (1987). A cell that dies during wild-type *C. elegans* development can function as a neuron in a *ced-3* mutant *Cell* **51**, 1071-1078.

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

Clark, S. G., Stern, M. J. and Horvitz, H. R. (1992). *C. elegans* cell-signalling gene *sem-5* encodes a protein with SH2 and SH3 domains. *Nature* **356**, 340-344.

Ferguson, E. and Horvitz, H. R. (1985). Identification and characterization of 22 genes that affect the vulval cell lineages of *Caenorhabditis elegans Genetics* **110**, 17-72.

Ferguson, E. and Horvitz, H. R. (1989). The multivulva phenotype of certain *C. elegans* mutants results from defects in two functionally-redundant pathways *Genetics* **123**, 109-121.

Han, M., Golden, A., Han, Y. and Sternberg, P. W. (1993). *C. elegans lin-45 raf* gene participates in *let-60 ras* stimulated vulval differentiation. *Nature* **363**, 133-140.

Han, M. and Sternberg, P. W. (1990). *let-60*, a gene that specifies cell fates

during C. elegans vulval induction, encodes a ras protein *Cell* **63**, 921-931.

Herman, R. K. and Hedgecock, E. M. (1990). The size of the *C. elegans* vulval primordium is limited by *lin-15* expression in surrounding hypodermis *Nature* **348**, 169-171.

Hill, R. J., Katz, W. S. and Sternberg, P. W. (in preparation). The EGF domain of Lin-3 is sufficient to induce vulval development.

Hill, R. J. and Sternberg, P. W. (1992). The *lin-3* gene encodes an inductive signal for vulval development in *C. elegans Nature* **358**, 470-476.

Horvitz, H. R. and Sternberg, P. W. (1991). Multiple intercellular signalling systems control the development of the *C. elegans* vulva *Nature* **351**, 535-541.

Horvitz, H. R. and Sulston, J. E. (1980). Isolation and genetic characterization of cell-lineage mutants of the nematode *Caenorhabditis elegans Genetics* **96**, 435-454.

Jongeward, G. D. and Sternberg, P. W. (1993). The *sli-1* locus of *C. elegans In preparation*

Sulston, J. E. and White, J. G. (1980). Regulation and cell autonomy during postembryonic development of *Caenorhabditis elegans Devel. Biol.* **78**, 577-597.

CHAPTER 4. TWO DISTINCT AP47 HOMOLOGS IN THE NEMATODE

There exist two distinct AP47 proteins in the nematode *C. elegans*.

Junho Lee and Paul W. Sternberg

HHMI Division of Biology

Caltech, CA 91125

Tel: (818) 395-2181

Fax: (818) 568-8012

ABSTRACT

The AP47 protein is the medium chain of the clathrin-associated protein complex AP-1 located on the *trans*-Golgi membrane in mammals. Here we report that the nematode *C. elegans* has at least two genes encoding AP47 proteins. One AP47 protein is encoded by *unc-101*, mutations in which gene cause pleiotropic effects (Lee, et al., 1993). Another AP47 is encoded by the second gene CEAP47. Amino acid comparison shows that the CEAP47 protein is 72% identical to UNC-101, and 72% identical to the mammalian AP47. We predict that the mammalian cells will also have at least two distinct AP47 proteins. A hybrid protein of UNC-101 and CEAP47 can complement *unc-101* function when expressed under the control of the *unc-101* promoter. We also discuss the redundancy of CEAP47 and UNC-101 in the nematode.

Introduction

Clathrin coated pits and vesicles are ubiquitous organelles found in all the eukaryote cells that mediate intracellular protein trafficking (Brodsky, 1988, Keen, 1990, Pearse and Robinson, 1990). Clathrin coated vesicles are composed of membrane fraction, selected receptors on the membrane, clathrin, and clathrin-associated proteins (APs) (Figure 1). While clathrin is a common structural unit to all the clathrin-coated vesicles (hence name of vesicles is clathrin-coated vesicles), APs vary depending on the localization of the vesicles at the cellular and subcellular level (Ahle, et al., 1988). Most cells contain AP-1 and AP-2 protein complexes on the Golgi and the plasma membrane as clathrin-associated protein, respectively, and neuronal cells contain AP-3 protein as neuronal cell-specific APs in addition to AP-1 and AP-2 (Morris, et al., 1993, Murphy, et al., 1991, Zhou, et al., 1993). AP-1 and AP-2 are similar in their composition and structure in that they are hetero-tetramer of two large chains, one small chain, and one medium chain (Matsui and Kirchhausen, 1990). AP-1 complex contains b'- and r-adaptin as large chains, AP47 as a medium chain , and AP19 as a small chain . AP-2 has a- and b-adaptin as large chains, AP50 as a medium chain, and AP17 as a small chain. The large chains share some similarity in their amino acid sequence (Kirchhausen, et al., 1989, Robinson, 1989, Robinson, 1990), so do those of medium chains and small chains (Kirchhausen, et al., 1991, Nakayama, et al., 1991, Thurieau, et al., 1988). AP-3, however, is unique, because it is a single peptide of 180 kD, and does not show any homology to any subunits of AP-1 or AP-2 (Morris, et al., 1993, Murphy, et al., 1991, Zhou, et al., 1993).

Since clathrin vesicles are present in almost every type of cell, one question rises: Do they perform same function in every cell? If they do not,

how do they perform different functions? There are a few indications of clathrin-coated vesicles showing some tissue-specificity, suggesting that they can function differently in different cells. For instance, clathrin light chains have two different versions: LCa and LCb(Brodsky, et al., 1991). Although they are expressed in the same cells, their expression level varies in different tissues. In addition, both the genes encoding clathrin light chains are subject to alternative splicing that produces neuron-specific transcripts with additional stretch of nucleotides (Jackson, et al., 1987, Kirchhausen, et al., 1987). Clathrin light chains are thought to have a regulatory function for clathrin assembly. It is conceivable that while structural units of the clathrin-coated vesicles (e.g., clathrin heavy chains) are ubiquitously present, molecules with regulatory functions vary in different compartments, cells, or tissues, conferring specificity's to the clathrin-coated vesicles. The medium chains of the clathrin-associated protein complexes were reported to have kinase activity, although their amino acid sequences do not show any homology to known kinases (Myers and Forgac, 1993, Nakayama, et al., 1991, Pauloin, 1982), suggesting that they might also have regulatory functions for the AP assembly or other interactions that APs are involved in, for example, interaction with receptors, clathrin triskelions, or membranes. Another example for tissue-specificity is the presence of AP-3 in the neuronal cells, which can provide additional specific function for neuronal cells such as synaptic vesicle assembly. AP-3 is known to promote clathrin assembly more efficiently than AP-1 or AP-2 (Lindner and Ungewickell, 1992).

α-adaptin, a large subunit of the plasma membrane clathrin-associated protein complex AP-2, is encoded by two separate genes (Robinson, 1989). Also, Robinson (1990) reported that there might exist more than one gene for

the g-adaptin. For the medium chains and small chains of AP-1 and AP-2, it is not clear whether they are encoded by a single gene or whether they have different versions.

Genetic analysis of *unc-101*, a gene encoding a medium chain (AP47) of AP-1 in the nematode *C. elegans* showed that mutations in this protein cause pleiotropic effects including subviability, an uncoordinated movement (D. Riddle, pers. comm.), a defect in dye uptake of certain neuron cells (E. Hedgecock, per. comm.), a male spicule defect (H. Chamberlin, per. comm.), a defecation defect (Thomas, 1990), and suppression of a reduction-of-function mutation in the epidermal growth factor receptor (EGFR) gene (Lee, et al., 1993). Interestingly, putative null mutations of *unc-101* do not cause 100 % lethality, but 50 % lethality: 50 % of progeny from homozygous mother of *unc-101/unc-101* genotype will survive (Lee, et al., 1993), suggesting that there might exist another gene encoding the medium chain of AP-1 that is partially redundant with *unc-101* and can replace the essential function of *unc-101* in some individuals when *unc-101* is defective.

Here we report identification of a second homolog of AP47, the medium chain of the *trans*-Golgi-associated protein complex AP-1, that might provide tissue-specific functions for the clathrin vesicles. We also show that the protein encoded by this gene can complement *unc-101* when expressed under the control of *unc-101* promoter. We discuss evolutionary concerns about these genes. We propose that in wild-type animals, these two genes are expressed in non-overlapping patterns, performing different functions in different compartments or cells.

Results

Molecular cloning of CEAP47, a second homolog of AP47, the medium chain of *trans*-Golgi clathrin-associated protein complex AP-1

A *C. elegans* cDNA sequencing project identified a cDNA clone (CEED20) containing a partial sequence similar to *unc-101* (Kervalage, per. comm.). We first determined the sequence of the CEED20 clone, and found that this clone contains a sequence of 817 nucleotides. A sequence comparison showed that the translation of this sequence has a homology to the AP47 protein from the amino acid #238, and to the CEAP50 protein from the amino acid #251. We compared this partial cDNA sequence with that of *unc-101* and other medium chain homologs. Amino acid sequence encoded by this partial sequence was 73 % identical to the corresponding *unc-101* sequence, and 42 % identical to AP50, indicating that this sequence encodes a protein more similar to AP47. A nucleotide sequence comparison showed that 2/3 of the differences between CEAP47 and *unc-101* were at the base of the codons (data not shown). We confirmed that this sequence is a sequence from *C. elegans*, not from any contaminant organism, by two means. First, the Southern hybridization of *C. elegans* genomic DNA digested with various restriction enzyme with radiolabeled probe made of this cDNA insert under highly stringent condition showed single bands hybridized, indicating that this sequence is from *C. elegans* (data not shown). Second, when we performed a PCR reaction of *C. elegans* genomic DNA with two primers made from the cDNA sequence, we found an intron in the genomic DNA sequence (Figure 2). This intron is 50 bp long, which is typical of *C. elegans* introns, and contains sequences mostly conserved in *C. elegans* introns (Wood, 1988).

We then screened a cDNA library provided by Barstead and probes made from the cDNA of CEED20. We isolated three full length cDNA clones, and determined one of the sequences (Figure 3).

Sequence analysis of CEAP47 gene

The cDNA sequence of the clone D20-3 is 1549 bases long, and contains an open reading frame from the second nucleotide through the nucleotide 1294. The first methionine codon ATG is likely at the nucleotide 14, based on the homology with other medium chain homologs. The putative translation from this sequence is 426 amino acid long. Comparison of the sequence of CEAP47 with other medium chain homologs indicated that this protein is more related to AP47 than to AP50 (Figure 4). Surprisingly, CEAP47 is as similar to the mammalian AP47 as to UNC-101 in *C. elegans* (72 % identity in both cases). The DNA sequence comparison between CEAP47 and *unc-101* showed that the discrepancies are biased toward the third bases of codons. There was not much homology in the 5' non-translated region, nor in the 3' non-translated region(data not shown), suggesting that these two genes might be subject to different regulation.

CEAP47 protein can complement UNC-101 protein.

Because mutations in the *unc-101* gene, which encodes a protein 72% identical to CEAP47, show pleiotropic phenotypes when the CEAP47 gene is intact, it is conceivable that CEAP47 gene has different functions from those of *unc-101*. However, we have shown that a hybrid construct, which contains the *unc-101* promoter region, and most of the mammalian AP47 (which is again 72 % identical to *unc-101*), and 3' region of *unc-101*, can rescue at least

two phenotypes associated with an *unc-101* mutation when introduced into the mutant animals by microinjection (Lee, et al., 1993). To test whether CEAP47 protein also can complement functions of UNC-101 protein, we constructed and examined a hybrid gene. The hybrid gene contains exactly the same contents as other constructs used for mammalian AP47 experiments (Lee, et al., 1993), except that this hybrid gene contains the cDNA portion of CEAP47 (figure 5). We microinjected this hybrid gene into the gonad of the *unc-101(sy108); let-23(sy1); dpy-20(e1284)* animals with a cloned DNA of *dpy-20(+)* as selection marker. The resultant transgenic animals were examined for their extent of movement and suppression of the vulvaless phenotype of the *let-23(sy1)* mutation. The uncoordinated movement was rescued, because the transgenic animals could move as well as wild-type animals. The suppression of the vulvaless phenotype of the *let-23(sy1)* mutation was also rescued, because 4 out of 8 transgenic animals had restored the vulvaless phenotype. The uncoordinated movement (Unc) phenotype was an easier phenotype to rescue by the transgene, because we observed the non-uncoordinated (non-Unc) animals from the F1 transient transgenic animals as well as stable lines of transgenic animals. On the contrary, we did not observe the rescue of the suppression of the vulvaless phenotype of the *let-23(sy1)* mutation until we established stable lines of transgenic animals.

DISCUSSION

Two distinct AP47 genes in the nematode *C. elegans:* indication of the second type of the *trans* Golgi clathrin-coated vesicles?

We have identified a second homolog of AP47, the medium chain of the

trans-Golgi clathrin-associated protein complex AP-1. Only one gene has been identified in the mammalian cells that encodes an AP47 protein. As described above, CEAP47 is 72 % identical to UNC-101 and also 72 % identical to the mammalian AP47. This extent of identity among the AP47 homologs suggests that there may exist a second gene for a mammalian AP47. If AP47 genes were duplicated after the nematode and the higher organisms such as mammalian lineages diverged during evolution, CEAP47 and UNC-101 should show higher homology to each other than to mammalian AP47. It is conceivable that these AP47 genes were duplicated very early during evolution, and remained homologous to each other because of their redundant functions. It is also conceivable that one copy of these genes were not lost during evolution because each of the genes has its own unique function.

Clathrin-coated vesicles on the *trans*-Golgi membrane are thought to be involved in 1) sorting and targeting mannose-6-phosphate receptors loaded with lysosomal enzymes into the lysosomes, and 2) targeting proteins to secretory granules (Keen, 1990, Kornfeld, 1990, Pearse and Robinson, 1990). How could the clathrin-coated vesicles with the same APs perform two different functions on the same membrane compartment? Could it be that there are two distinct coated vesicles on the *trans*-Golgi membrane containing different, yet highly related, APs? Robinson (1990) suggested that γ-adaptin might be encoded by more than one gene, and that the presence of different forms of γ-adaptin in the mammalian cells could provide specificity for the clathrin-coated vesicles on the *trans*-Golgi compartment. In this report, we showed that in *C. elegans,* there is a gene encoding the second homolog of the medium chain of the *trans*-Golgi clathrin-associated protein complexes.

Though the indications on γ-adaptin and AP47 homologs are from different species, it is conceivable, based on their high similarity to their homologs and conserved functions, that in both the species, there might exist second homologs of the other components of AP-1, and that there might exist the second type of the *trans*-Golgi clathrin-coated vesicles. Further analysis of the genes of the subunits of the *trans*-Golgi clathrin-associated protein complexes in both the species will solve this question.

Redundancy of CEAP47 and *unc-101*

Because mutations in the *unc-101* locus cause pleiotropic effects, *unc-101* seems to be not redundant with CEAP47. However, the phenotypes of *unc-101* putative null mutations are not identical in all the animals bearing the same mutations in *unc-101*. For example, the lethality associated with the putative null mutations is not completely penetrant. Only 50 % of the progeny of homozygous hermaphrodites will die, while the other 50 % will survive. The defecation defect is more fluctuating even in a single animal (Thomas, 1990). Each defecation cycle in *C. elegans* is composed of an anterior body muscle contraction (aBoc), a posterior body muscle contraction (pBoc), and an expulsion (Exp) step. In *unc-101* mutant animals, the aBoc step is missing in half of the defecation cycles. In the other half cycles of defecation, the aBoc is normal. Therefore it is conceivable that there may exist a gene that shares partial redundancy with *unc-101*. CEAP47 may be one such gene. We showed that hybrid proteins, 2/3 of which is from the CEAP47 gene and 1/3 from *unc-101*, can complement the defective UNC-101 protein function, if they are expressed under the control of the *unc-101* promoter and regulation. This result suggests that CEAP47 may be

redundant with *unc-101* in the wild-type animals. The extent of redundancy of the CEAP47 and the *unc-101* gene in the wild-type animals is not known. There are a few possibilities. One possibility is that *unc-101* and CEAP47 have identical functions, and that a full expression of both the genes are required for production of sufficient quantity of proteins. This is unlikely because if this is the case, the mutations in the *unc-101* locus would be dosage-dependent. But they are fully recessive, indicating that loss of one copy of *unc-101* does not cause any defect. However, one still cannot exclude the possibility that there is a threshold for the expression level of these genes to execute proper functions. From this hypothesis, one can predict 1) that the CEAP47 mutant animals will have the same phenotype as *unc-101* mutant animals, 2) that the double mutants for both the genes will be 100 % lethal, and 3) that the heterozygotes of *CEAP47* +/ + *unc-101* trans-heterozygotes will have the same phenotype as *unc-101.*

Another possibility is that although the proteins of CEAP47 and UNC-101 have the same functions, these proteins are expressed in different types of cells. When *unc-101* is mutated, a basal level of CEAP47 expression might be somehow activated in the cells that normally produce UNC-101 protein, and compensate for a defect in the Unc-101 protein. If this is the case, 1) mutations in the CEAP47 locus will have different phenotypes than *unc-101,* 2) the heterozygotes of *CEAP47* +/ + *unc-101* will have a wild-type phenotype. It would be hard to predict whether the double mutant will be completely lethal.

A third hypothesis is that although these two proteins are redundant, CEAP47 is not actively expressed in the wild-type animals, and is expressed only when *unc-101* is mutated. If this is the case, 1) mutations in the CEAP47

will be silent, and 2) The double mutants will be 100% lethal.

If one assumes the presence of the second clathrin coated-vesicles on the *trans*-Golgi membrane as hypothesized above, one plausible hypothesis is that CEAP47 is a component of the second type of the *trans*-Golgi clathrin-associated protein complex as hypothesized above, performing different functions from the UNC-101-containing clathrin-associated protein complex, and that CEAP47 can somehow compensate for the UNC-101 functions only when CEAP47 is overexpressed and/ or *unc-101* is mutated. If this is the case, one can predict 1) that singly mutants for CEAP47 will probably have different phenotypes than *unc-101*, and 2) that doubly mutants for *unc-101* and CEAP47 will be 100% lethal,

At this point, we cannot distinguish among these hypotheses because no genetic mutations are available for the CEAP47 gene yet. Studies of expression patterns by transgenic animal techniques using hybrid genes with reporter molecules such as β-galactosidase gene might help solve some of redundancy problem (Mello, et al., 1991). However, this technique has some limitations. Because the transgenes have multicopy of the gene of interest, the expression pattern by the transgenes may not reflect the intact expression patterns. Also, the transgenic animals tend to lose the transgenes randomly during cell divisions, making it difficult to interpret the expression patterns exhibited by the transgenic animals.

Materials and Methods

cDNA screening

We used the cDNA clone CEED20 (accession number T00259) from the Genbank database as a start point for cloning full length cDNA clones. We used the CEED20 DNA as probe in a cDNA screening for full length clones. We used a standard hybridization procedure (Sambrook, et al., 1989). We isolated three cDNA clones from a *C. elegans* cDNA library (Barstead and Waterston, 1989), all of which contained inserts of the same length. We determined the sequence of one of the clones, CEED20-3. Sequencing reactions were performed using Sequenase 2.0 and reagents from United States Biochemical.

Sequence analysis

Compiling of DNA and amino acid sequences were carried out using Macvector™ program (IBI) and the GCG package v7.0, a software of the Genetics Computer Group (Devereux, et al., 1984, GCG, 1991). The BLAST program of the GCG software was used to search and compare homologies of the sequences. The Pileup and Gap programs were used to generate the comparisons of the amino acid sequences.

PCR reactions for comparing cDNA and genomic sequence of CEAP47

We made two primers for PCR from the cDNA clone CEED20: CE-1, 5' CGTTTATCTCGCTTTGATTC and CE-4, TTTTGCTGTACCTGCTCCAG. These primers were also used as primers for sequencing the PCR product. We used genomic DNA of N2 strain as template for PCR. PCR reactions were

performed by a standard procedure (Ausubel, et al., 1992, Sambrook, et al., 1989).

Southern hybridization

A southern blot of genomic DNAs of N2 and *lin-3* strains was a gift from R. Hill. We used the CEED20 DNA as template for making random-primed probe in the hybridization. Hybridization method was standard and washing condition was high stringency (Sambrook, et al., 1989).

Construction of hybrid genes.

Construction of the *unc-101* hybrid genes was described (Lee, et al., 1993). To construct the CEAP47 hybrid gene, we amplified the CEAP47 cDNA from nucleotide #325 through #1229 using two PCR primers. Both the 5/ and 3' subcloning sites, NruI and EcoRV, are conserved in CEAP47. The two PCR primers are: CE-6, 5'CGATAATTTCGTTATTATTTATG3', and CE-7, 5'ATCCAGATTTCTCTATGATTTT.

The amplified DNA was ligated to the 7.2kb NruI/EcoRV fragment of pJL2. The resulting plasmid is the CEAP47 hybrid gene. This construct contains the 5' promoter region of *unc-101*, the 5' coding region of *unc-101 up* to the *unc-101* cDNA nucleotide #388, the CEAP47 cDNA from NruI site to EcoRV site (corresponding nucleotides in *unc-101* are #389 to #1281, the *unc-101* 3' region from #1282 to the end of cDNA , and the untranscribed 3' region of *unc-101.* The protein predicted to be made from this construct contains 301 amino acid residues from CEAP47, and 123 amino acid residues from *unc-101.*

Microinjection experiments

Microinjection of DNA into the gonad of *C. elegans* hermaphrodite adults was described by Mello et al.(1991). We used *unc-101(sy108); let-23(sy1); dpy-20(e1282)* animals as the host for microinjection. We injected the hybrid gene with a *dpy-20(+)* clone as a selection marker. After microinjection we selected nonDpy transgenic animals, established stable lines that inherited the transgenes, and examined the phenotype of Unc and vulval differentiation.

Figure 1. Structure of clathrin-coated vesicles and their associated protein complexes. (A) clathrin-coated pits and vesicles. Clathrin coated pits and vesicles are composed of membrane, proteins such as receptors, clathrin triskelions, and associated protein complexes (APs).

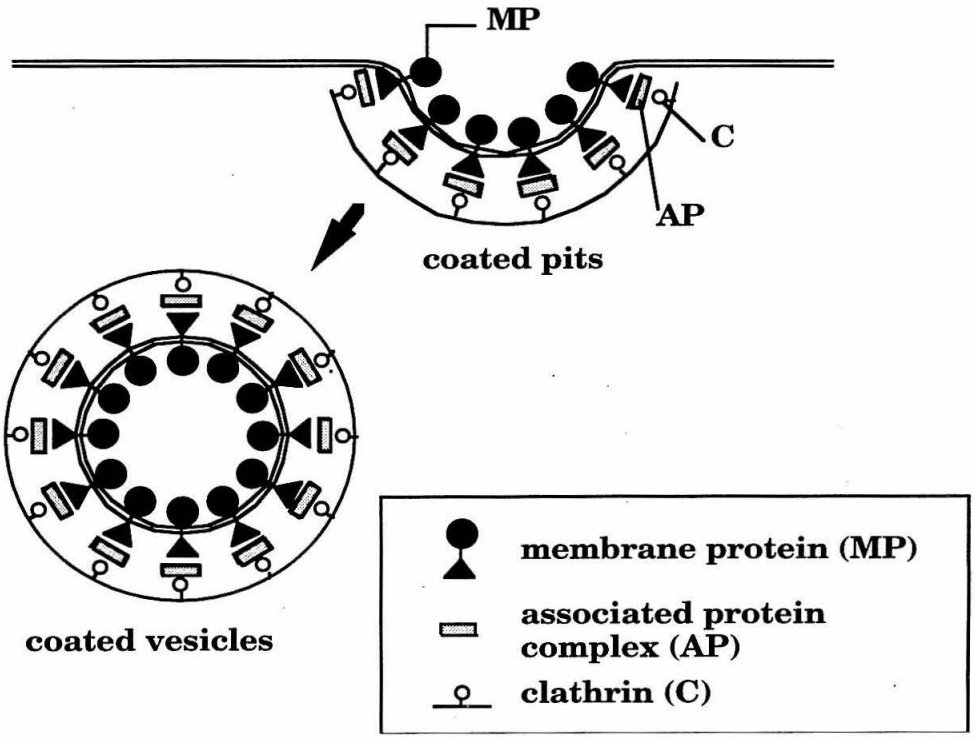
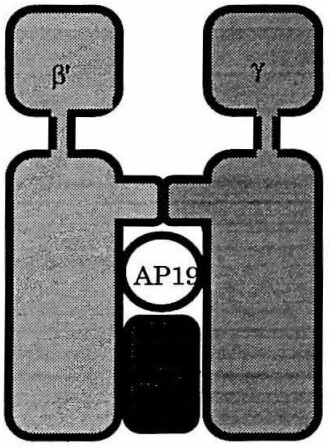
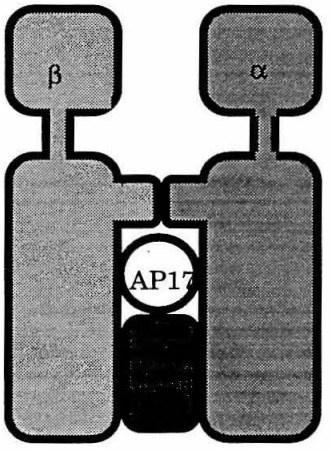

MP
C
AP
coated pits
coated vesicles
membrane protein (MP)
associated protein complex (AP)
clathrin (C)

Figure 1. (B) Structure of clathrin-associated protein complexes AP-1 and AP-2 complexes. Both the APs are composed of four subunits: two large chains, one medium chain, and one small chain. AP-1, the complex located on the *trans*-Golgi, contains γ- and β'-adaptin, AP47, and AP19. AP-2, the complex located on the plasma membrane, contains α- and β-adaptin, AP50, and AP17.



AP-1 = *trans*- Golgi form



AP-2 = plasma membrane form

Figure 2. Comparison of partial genomic sequence of CEAP47 and its corresponding cDNA sequence. The upper lines of sequence is partial cDNA sequence of the CEED20 clone from a database. The bottom lines of sequence is the genomic DNA sequence obtained from a genomic PCR reaction using primers from the CEED20 sequence. Two sequences are identical except for an intron inserted in the genomic sequence . The arrowhead indicates the location of an intron in the *unc-101* gene.

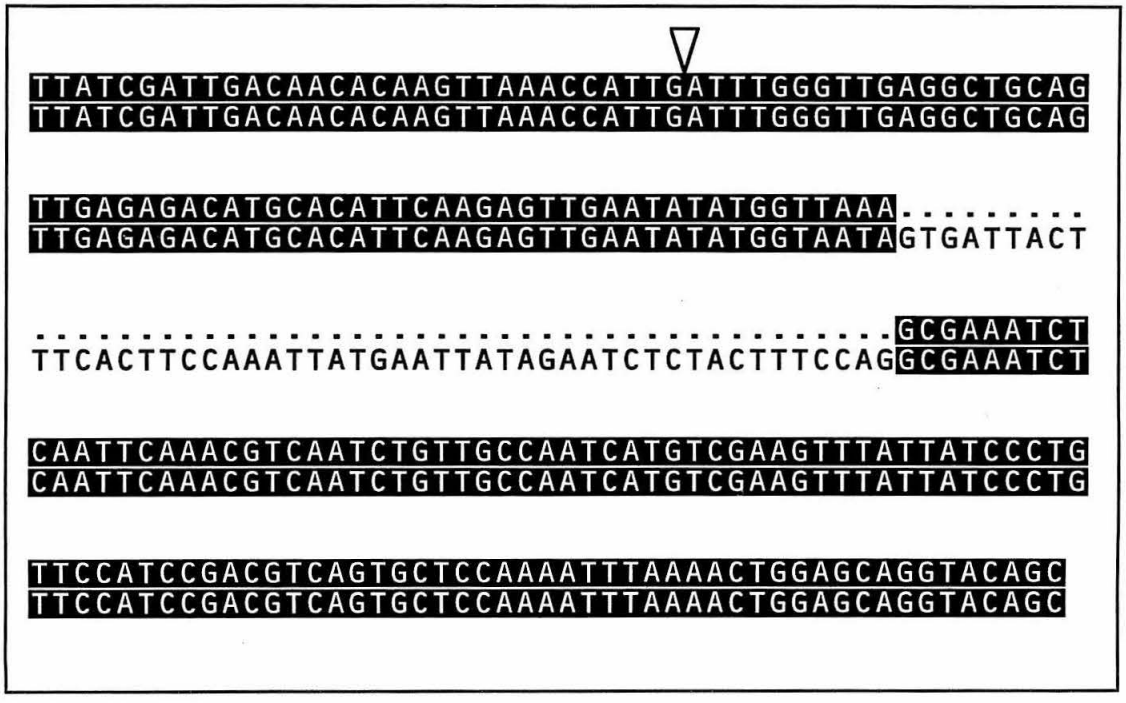
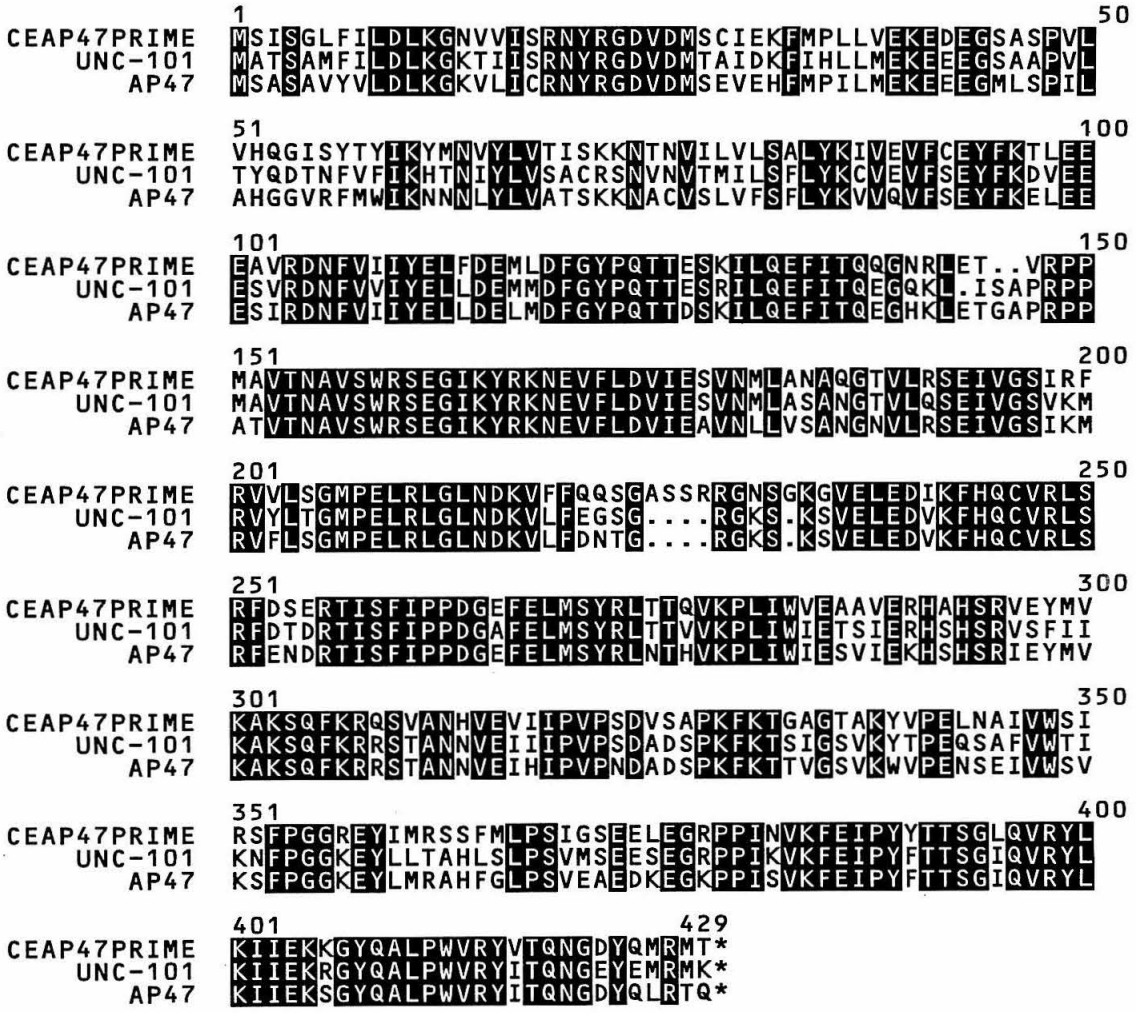

TTATCGATTGACAACACAAGTTAAACCATTGATTTGGGTTGAGGCTGCAG
TTATCGATTGACAACACAAGTTAAACCATTGATTTGGGTTGAGGCTGCAG
TTGAGAGACATGCACATTCAAGAGTTGAATATATGGTTAAA.....
TTGAGAGACATGCACATTCAAGAGTTGAATATATGGTAATAGTGATTACT
.....GCGAAATCT
TTCACTTCCAAATTATGAATTATAGAATCTCTACTTTCCAGGCGAAATCT
CAATTCAAACGTCAATCTGTTGCCAATCATGTCGAAGTTTATTATCCCTG
CAATTCAAACGTCAATCTGTTGCCAATCATGTCGAAGTTTATTATCCCTG
TTCCATCCGACGTCAGTGCTCCAAAATTTAAAACTGGAGCAGGTACAGC
TTCCATCCGACGTCAGTGCTCCAAAATTTAAAACTGGAGCAGGTACAGC

Figure 3. cDNA sequence of CEAP47 and its translation. The first methionine and the stop codon was assigned based on the homology between the translated sequence of this gene and the AP47 protein sequence. The DNA sequences used for designing PCR primers are underlined. The CEED20 clone starts at the nucleotide #733.

```
CGGAGAAATCGACATGTCGATTTCCGGTCTCTTCATCTTGGACCTCAAAGGAAATGTAGT 60
                M  S  I  S  G  L  F  I  L  D  L  K  G  N  V  V
AATATCGAGGAATTACCGCGGAGATGTGGATATGTCGTGTATCGAAAAGTTTATGCCACT 120
  I  S  R  N  Y  R  G  D  V  D  M  S  C  I  E  K  F  M  P  L
GTTGGTTGAAAAGGAAGACGAAGGAAGTGCATCTCCAGTATTAGTGCATCAAGGAATTAG 180
  L  V  E  K  E  D  E  G  S  A  S  P  V  L  V  H  Q  G  I  S
TTATACATATATTAAATATATGAATGTTTACTTGGTGACGATTTCGAAGAAAAACACGAA 240
  Y  T  Y  I  K  Y  M  N  V  Y  L  V  T  I  S  K  K  N  T  N
TGTGATTCTTGTATTGTCGGCTCTTTACAAAATTGTCGAGGTTTTCTGCGAATATTTCAA 300
  V  I  L  V  L  S  A  L  Y  K  I  V  E  V  F  C  E  Y  F  K
AACATTGGAGGAAGAAGCTGTTCGCGATAATTTCGTTATTATTTATGAACTTTTCGACGA 360
  T  L  E  E  E  A  V  R  D  N  F  V  I  I  Y  E  L  F  D  E
AATGCTCGATTTTGGATATCCACAGACAACGGAGAGCAAGATTCTTCAAGAATTCATAAC 420
  M  L  D  F  G  Y  P  Q  T  T  E  S  K  I  L  Q  E  F  I  T
ACAACAAGGTAATCGTTTGGAGACAGTTCGTCCACCCATGGCAGTCACAAACGCGGTTTC 480
  Q  Q  G  N  R  L  E  T  V  R  P  P  M  A  V  T  N  A  V  S
ATGGAGATCAGAAGGGATCAAATATCGGAAAAATGAAGTTTTTCTCGATGTTATTGAAAG 540
  W  R  S  E  G  I  K  Y  R  K  N  E  V  F  L  D  V  I  E  S
TGTTAATATGCTGGCAAATGCTCAAGGAACCGTACTCCGTTCAGAAATCGTTGGTTCAAT 600
  V  N  M  L  A  N  A  Q  G  T  V  L  R  S  E  I  V  G  S  I
TCGATTCCGTGTTGTTCTTTCTGGAATGCCGGAACTTCGACTTGGACTTAATGACAAGGT 660
  R  F  R  V  V  L  S  G  M  P  E  L  R  L  G  L  N  D  K  V
GTTCTTCCAGCAATCTGGTGCAAGTTCTAGACGTGGTAACAGTGGAAAAGGAGTCGAATT 720
  F  F  Q  Q  S  G  A  S  S  R  R  G  N  S  G  K  G  V  E  L
GGAAGATATCAAATTCCATCAATGTGTCCGTTTATCTCGCTTTGATTCCGAAAGAACCAT 780
  E  D  I  K  F  H  Q  C  V  R  L  S  R  F  D  S  E  R  T  I
TTCTTTTATTCCACCGGATGGAGAATTTGAGCTGATGAGTTATCGATTGACAACACAAGT 840
  S  F  I  P  P  D  G  E  F  E  L  M  S  Y  R  L  T  T  Q  V
TAAACCATTGATTTGGGTTGAGGCTGCAGTTGAGAGACATGCACATTCAAGAGTTGAATA 900
  K  P  L  I  W  V  E  A  A  V  E  R  H  A  H  S  R  V  E  Y
TATGGTTAAAGCGAAATCTCAATTCAAACGTCAATCTGTTGCCAATCATGTCGAAGTTAT 960
  M  V  K  A  K  S  Q  F  K  R  Q  S  V  A  N  H  V  E  V  I
TATCCCTGTTCCATCCGACGTCAGTGCTCCAAAATTTAAAACTGGAGCAGGTACAGCAAA 1020
  I  P  V  P  S  D  V  S  A  P  K  F  K  T  G  A  G  T  A  K
AATATGTTCCAGAGCTTAATGCTATTGTTTGGAGTATTCGAAGCTTCCCAGGAGGCCGTGA 1080
  Y  V  P  E  L  N  A  I  V  W  S  I  R  S  F  P  G  G  R  E
ATATATAATGAGATCTTCGTTCATGCTCCCATCAATTGGCTCCGAAGAGCTGGAAGGCCG 1140
  Y  I  M  R  S  S  F  M  L  P  S  I  G  S  E  E  L  E  G  R
TCCGCCAATCAATGTTAAATTTGAAATTCCTTACTACACAACTTCCGGACTACAAGTTCG 1200
  P  P  I  N  V  K  F  E  I  P  Y  Y  T  T  S  G  L  Q  V  R
ATATTTAAAAATCATAGAGAAATCTGGATATCAAGCACTTCCATGGGTACGATATGTTAC 1260
  Y  L  K  I  I  E  K  S  G  Y  Q  A  L  P  W  V  R  Y  V  T
TCAAAACGGAGATTATCAAATGAGAATGACCTAATAAGAAGCTTTCTGTTCCCAATATCC 1320
  Q  N  G  D  Y  Q  M  R  M  T  *
TTCAAGCCACCTTCAAGCCAAAACAGTCATCCCTGTAATTTTTGTTCTATTCAATTCCCC 1380
CGCCATTCGGTCCCTTGGGTGCCTTATATTTTTTTTTTAATGTTTAATTTTATTCCGAGA 1440
GAGAAAAAAACTAGTAATTTATTAACTGTGATATTGTCTTTTGTTGCTCTTTTCAACGCT 1500
TTTCATCCCCTGGTTTCTTGTATCTTTCCTTTTTTTTGTTTCGTTTTTC             1549
```

Figure 4. (A) Comparison of amino acid sequences of CEAP47, UNC-101, and mammalian AP47. The amino acid residues that are identical in all three homologs are highlighted.


1 50
CEAP47PRIME MSISGLFILDLKGNVVISRNYRGDVDMSCIEKFMPLLVEKEDEGSASPVL
UNC-101 MATSAMFILDLKGKTIISRNYRGDVDMTAIDKFIHLLMEKEEEGSAAPVL
AP47 MSASAVYVLDLKGKVLICRNYRGDVDMSEVEHFMPILMEKEEEGMLSPIL
51 100
CEAP47PRIME VHQGISYTYIKYMNVYLVTISKKNTNVILVLSALYKIVEVFCEYFKTLEE
UNC-101 TYQDTNFVFIKHTNIYLVSACRSNVNVTMILSFLYKCVEVFSEYFKDVEE
AP47 AHGGVRFMWIKNNNLYLVATSKKNACVSLVFSFLYKVVQVFSEYFKELEE
101 150
CEAP47PRIME EAVRDNFVIIYELFDEMLDFGYPQTTESKILQEFITQQGNRLET..VRPP
UNC-101 ESVRDNFVVIYELLDEMMDFGYPQTTESRILQEFITQEGQKL.ISAPRPP
AP47 ESIRDNFVIIYELLDELMDFGYPQTTDSKILQEFITQEGHKLETGAPRPP
151 200
CEAP47PRIME MAVTNAVSWRSEGIKYRKNEVFLDVIESVNMLANAQGTVLRSEIVGSIRF
UNC-101 MAVTNAVSWRSEGIKYRKNEVFLDVIESVNMLASANGTVLQSEIVGSVKM
AP47 ATVTNAVSWRSEGIKYRKNEVFLDVIEAVNLLVSANGNVLRSEIVGSIKM
201 250
CEAP47PRIME RVVLSGMPELRLGLNDKVFFQQSGASSRRGNSGKGVELEDIKFHQCVRLS
UNC-101 RVYLTGMPELRLGLNDKVLFEGSG....RGKS.KSVELEDVKFHQCVRLS
AP47 RVFLSGMPELRLGLNDKVLFDNTG....RGKS.KSVELEDVKFHQCVRLS
251 300
CEAP47PRIME RFDSERTISFIPPDGEFELMSYRLTTQVKPLIWVEAAVERHAHSRVEYMV
UNC-101 RFDTDRTISFIPPDGAFELMSYRLTTVVKPLIWIETSIERHSHSRVSFII
AP47 RFENDRTISFIPPDGEFELMSYRLNTHVKPLIWIESVIEKHSHSRIEYMV
301 350
CEAP47PRIME KAKSQFKRQSVANHVEVIIPVPSDVSAPKFKTGAGTAKYVPELNAIVWSI
UNC-101 KAKSQFKRRSTANNVEIIIPVPSDADSPKFKTSIGSVKYTPEQSAFVWTI
AP47 KAKSQFKRRSTANNVEIHIPVPNDADSPKFKTTVGSVKWVPENSEIVWSV
351 400
CEAP47PRIME RSFPGGREYIMRSSFMLPSIGSEELEGRPPINVKFEIPYYTTSGLQVRYL
UNC-101 KNFPGGKEYLLTAHLSLPSVMSEESEGRPPIKVKFEIPYFTTSGIQVRYL
AP47 KSFPGGKEYLMRAHFGLPSVEAEDKEGKPPISVKFEIPYFTTSGIQVRYL
401 429
CEAP47PRIME KIIEKKGYQALPWVRYVTQNGDYQMRMT*
UNC-101 KIIEKRGYQALPWVRYITQNGEYEMRMK*
AP47 KIIEKSGYQALPWVRYITQNGDYQLRTQ*

(B) Diagram showing the identity of amino acid sequences among the homologs of the medium chains of the clathrrin-coated vesicles. CEAP47 is more similar to AP47 than AP50 homologs.

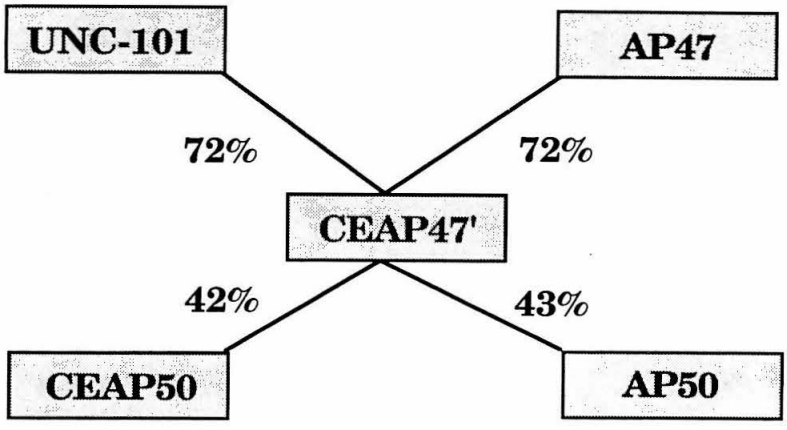
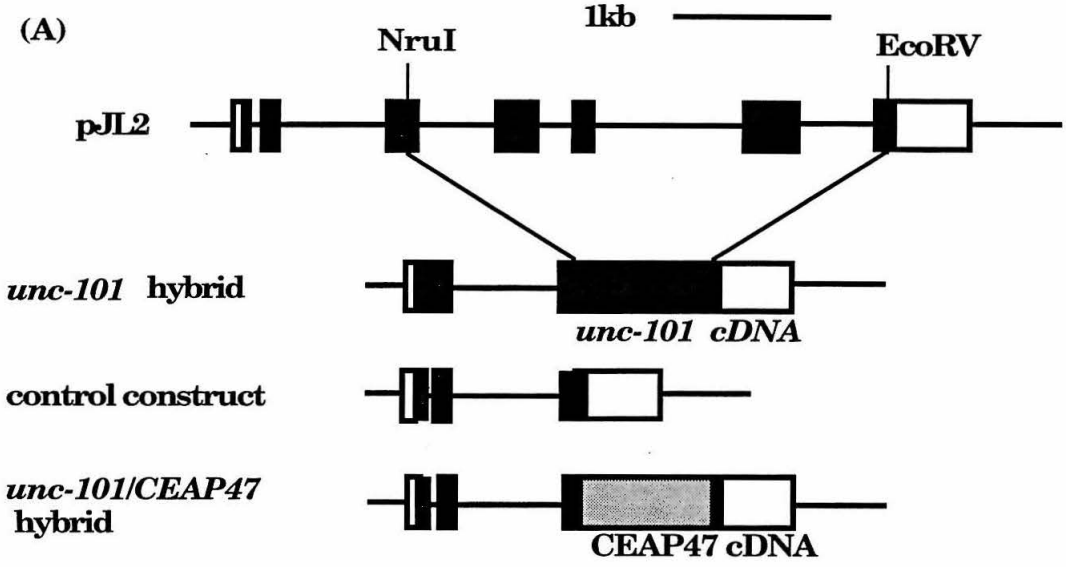

UNC-101
AP47
72%
72%
CEAP47'
42%
43%
CEAP50
AP50

Figure 5. CEAP47proteins can complement UNC-101 proteins. (A) The CEAP47 hybrid gene contains the 5' putative promoter region of the *unc-101* gene, the 5' coding region of *unc-101* up to the NruI restriction site, CEAP47 coding region to the EcoRV site, and the 3' region of the *unc-101* gene. The resultant protein predicted to be produced from this hybrid gene contains 301 amon acids of CEAP47, and 123 amino acids *unc-101*. A positive control construct, the *unc-101* hybrid gene, contains all amino acids from *unc-101*. A negative control contstruct contains only 123 amino acids from *unc-101*, and no amino acids from CEAP47. (B) The result of microinjection of the hybrid genes. The positive control gene and the CEAP47 hybrid gene rescued at least two phenotypes of *unc-101* mutations, while the negative control did not rescue any. The *unc-101/CEAP47* hybrid gene rescued both the phenotypes (see the text for details).



(B) Rescue of *unc-101* phenotypes

	Unc phentype	Suppression of vulvaless phenotype of let-23(sy1)
unc-101 hybrid	Yes	Yes
control construct	No	No
unc-101/CEAP47 hybrid	Yes	Yes

References

Ahle, S., Mann, A., Eichelsbacher, U. and Ungewickell, E. (1988). Structural relationships between clathrin assembly proteins from the Golgi and the plasma membrane *EMBO J.* **7**, 919-929.

Barstead, R. J. and Waterston, R. H. (1989). The basal component of the nematode dense-body is vinculin *J. Biol. Chem.* **264**, 10177-10185.

Brodsky, F. M. (1988). Living with clathrin: its role in intracellular membrane traffic *Science* **242**, 1396-1402.

Brodsky, F. M., Hill, B. L., Acton, S. L., Nathke, I., Wong, D. H., Ponnambalam, S. and Parham, P. (1991). Clathrin light chains: arrays of protein motifs that regulate coated-vesicle dynamics. *TIBS* **16**, 208-213.

Devereux, J., Haeberli, P. and Smithies, O. (1984). A comphrehensive set of sequence analysis programs for the VAX *Nucl Acids Res* **12**, 387-395.

GCG (1991). Program Manual for the GCG Package, Version 7

Jackson, A. P., Seow, H. F., Holmes, N., Drikamer, K. and Parham, P. (1987). Localization of clathrin light-chain sequences mediating heavy-chain binding and coated vesicle diversity *Nature* **326**, 154-159.

Keen, J. H. (1990). Clathrin and associated assembly and disassembly proteins *Annn. Rev. Biochem.* **59**, 415-438.

Kirchhausen, T., Davis, A. C., Frucht, S., Greco, B. O., Payne, G. S. and Tubb, B. (1991). AP17 and AP19, the mammalian small chains of the clathrin-associated protein complexes show homology to YAP17p, their putative homolog in yeast *J. Biol. Chem* **266**, 1153-1157.

Kirchhausen, T., Harrison, S. C., Chow, E. P., Mattaliano, R. J., Ramachandran, K. L., Smart, J. and Brosius, J. (1987). Clathrin

heavy-chain-molecular cloning and complete primary structure. *proc. Natl. Acad. Sci. USA* **84**, 8805-8809.

Kirchhausen, T., Nathanson, K. L., Matsui, W., Vaisberg, A., Chow, E. P., Burne, Z., Keen, J. H. and Davies, A. E. (1989). Structural and functional division into two domains of the large (100- to 115kDa) chains of the clathrin-associated protein complex AP-2. *Proc. Natl. Acad. Sci. USA* **86**, 2612-2616.

Kirchhausen, T., Scarmoto, P., Harrison, S. C., Monroe, J. J., Chow, E. P., Mattaliano, R. J., Ramachandran, K. L. and al., e. (1987). Clathrin light-chains LCa and LCb are similar, polymorphic, and share repeated heptad motifs *Science* **236**, 320-324.

Kornfeld, S. (1990). Lysosomal enzyme targetting *Biochem. Soc. Trans.* **18**, 367-374.

Lee, J., Jongeward, G. D. and Sternberg, P. W. (1993). *unc-101*, a gene required for many aspects of *C. elegans* development and behavior, encodes a clathrin-associated protein *In preparation*

Lindner, R. and Ungewickell, E. (1992). Clathrin-associated proteins of bovine brain coated vesicles - an analysis of their number and assembly-promoting activity *J. Biol. Chem.* **267**, 16567-16573.

Matsui, W. and Kirchhausen, T. (1990). Stabilization of clathrin coats by the core of the clathrin-associated protein complex AP-2 *Biochemistry* **29**, 10791-10798.

Mello, C. C., Kramer, J. M., Stinchcomb, D. and Ambros, V. (1991). Efficient gene transfer in *C. elegans* after microinjection of DNA into germline cytoplasm: extrachromosomal maintenance and integration of transforming sequences *EMBO J.* **10**, 3959-3970.

Morris, S. A., Schroder, S., Plessmann, U., Weber, K. and Ungewickell, E. (1993). Clathrin assembly protein AP180: primary structure, domain organization and identification of a clathrin binding site *EMBO J.* **12**, 667-675.

Murphy, J. E., Pleasure, I. T., Puszkin, S., Prasad, K. and Keen, J. H. (1991). Clathrin assembly protein AP-3 - the identity of the 155K protein, AP-180, and NP185 and demonstration of a clathrin binding domain *J. Biol. Chem.* **266**, 4401-4408.

Myers, M. and Forgac, M. (1993). The coated vesicle vacuolar (H+)-ATPase associates with and is phosphorylated by the 50 KDa polypeptide of the clathrin assmbly protein AP-2. *J. Biol. Chem.* **268**, 9184-9186.

Nakayama, Y., Goebl, M., Greco, B. O., Lemmon, S., Chow, E. P. C. and Kirchhausen, T. (1991). The medium chains of the mammalian clathrin-associated proteins have a homolog in yeast. *Eur. J. Biochem.* **202**, 569-574.

Pauloin, C., Bernier,I., Jolles,P. (1982). Presence of cyclic nucleotide Ca++-independent protein kinase in bovine brain coated vesicles. *Nature* **298**, 574-576.

Pearse, B. M. F. and Robinson, M. S. (1990). Clathrin, adaptors, and sorting *Annu. Rev. Biol.* **6**, 151-171.

Robinson, M. S. (1989). Cloning of cDNAs encoding two related 100-kD coated vesicle proteins (a-adaptins) *J. Cell Biol.* **108**, 833-842.

Robinson, M. S. (1990). Cloning and expression of g-adaptin, a component of clathrin-coated vesicles associated with the Golgi apparatus. *J. Cell Biol.* **111**, 2319-2326.

Thomas, J. H. (1990). Genetic analysis of defecation in *Caenorhabditis*

elegans Genetics **124**, 855-872.

Thurieau, C., Brosius, J., Burne, C., Jolles, P., Keen, J., Mattaliano, R. J., Chow, E. P., Ramachandran, K. and Kirchhausen, T. (1988). Molecular cloning of complete amino acid sequence of AP50, an assembly protein associated with clathrin-coated vesicles *DNA* **7**, 663-669.

Zhou, S., Tannery, N. H., Yang, J., Puszkin, S. and Lafer, E. M. (1993). The synapse-specific phosphoprotein F1-20 is identical to clathrin assembly protein AP-3. *J. Biol. Chem.* **268**, 12655-12662.

Appendix I: A *C. elegans* homolog of γ-adaptin, a large chain of the *trans*-Golgi clathrin-associated protein complex AP-1

Clathrin coated pits and vesicles are involved in intracellular protein trafficking such as endocytosis, lysosomal enzyme sorting, and retention of Golgi proteins (Brodsky, 1988, Keen, 1990, Payne and Schekman, 1989, Pearse and Robinson, 1990, Schmid, 1992, Seeger and Payne, 1992). Clathrin coated vesicles are composed of membrane, membrane proteins such as the receptors to be transported, clathrin, and its associated protein complex. There are at least two different classes of coated pits, one on the plasma membrane, and another on the *trans*-Golgi (Ahle, et al., 1988, Keen, 1987, Kirchhausen, et al., 1989). Clathrin is a structural component common to both the coated pits, and their associated protein complexes are the components that provide localization specificity for either the plasma membrane or the *trans*-Golgi coated pits. Each complex is composed of four different subunits: two large chains, one medium chain, and one small chain (Kirchhausen, 1990, Kirchhausen, et al., 1991, Kirchhausen, et al., 1989, Nakayama, et al., 1991, Robinson, 1990).

Mutations in *unc-101,* which we have shown to encode a medium chain of the *trans*-Golgi-associated protein complex, cause many different phenotypes, for example, uncoordinated movement, a FITC staining defect, subviability, and suppression of vulvaless phenotype of a mutation in the receptor tyrosine kinase *let-23* (Chapter 2). To understand the roles of clathrin associated protein complexes, it is necessary to study the other components. For example, would elimination of a γ-adaptin have the same phenotypes as *unc-101* mutants?

A cDNA clone encoding a homolog of γ-adaptin, the large chain of the *trans*-Golgi associated protein complex AP-1, was isolated by a random sequencing of a cDNA library (A. R. Kerlavage et al., personal comm.). Using

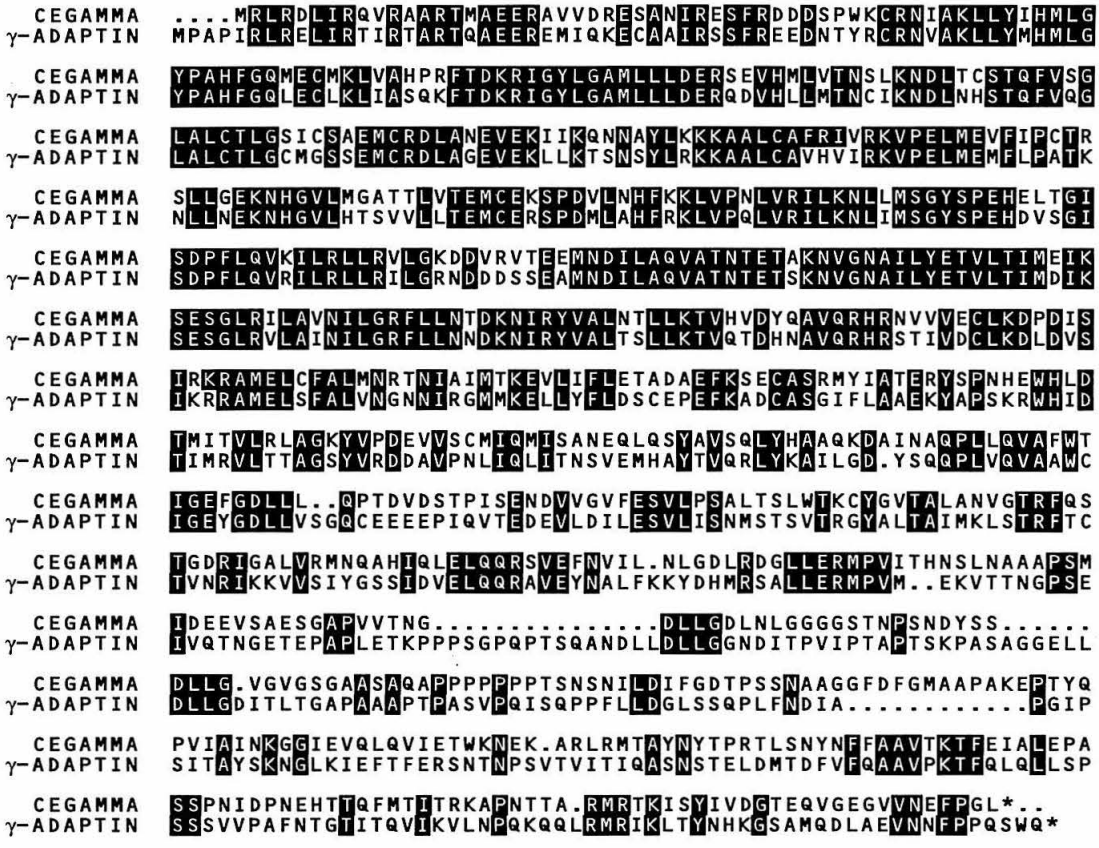
this cDNA clone, which has only the middle part of the gene, we identified additional cDNA clones. One clone contained six nucleotides on its 5' end that are identical to SL-1, indicating that this clone indeed contains a full length transcribed region (Spieth, et al., 1993). We determined a full length cDNA sequence (Figure 1). The cDNA is 3,040 nucleotides long, and is predicted to encode a single peptide of 803 amino acids. An amino acid sequence comparison reveals that *C. elegans* γ-adaptin is 52% identical to mammalian γ-adaptin (Robinson, 1990)(Figure 2), 29 to 30 % identical to α-adaptins (Robinson, 1989), and only 20-21 % identical to β- and β'-adaptin (Kirchhausen, 1990). In addition, we found that the N-terminal half of the proteins is more homologous among these adaptins than the C-terminal half. We have mapped this cDNA to the right end of chromosome IV by hybridization to a YAC grid filter provided by A. Coulson (Coulson, et al., 1991, Coulson, et al., 1988).

Figure 1. The cDNA sequence and its translation of a γ-adaptin homolog in *C. elegans*. The putative start and stop codons are underlined. A putative poly adenylation signal is also underlined.

```
GTTTGAGGGTGAGGCTCCAAGGATGGCGAC GGCTGTCGAGCTAGCCATTGAGAAGATCGA 60
TGAGTATAAGTCGAAGATTGGAAGAACTCT CGGCACTCCGATGCGGCTTCGAGACCTAAT 120
                                    M  R  L  R  D  L  I
TCGGCAGGTCAGAGCGGCACGAACGATGGC TGAAGAGCGAGCAGTGGTGGATAGAGAAAG 180
  R  Q  V  R  A  A  R  T  M  A   E  E  R  A  V  V  D  R  E  S
TGCGAATATTCGAGAAAGCTTCCGAGACGA TGATTCTCCGTGGAAATGTAGAAATATCGC 240
  A  N  I  R  E  S  F  R  D  D   D  S  P  W  K  C  R  N  I  A
GAAATTGCTCTATATTCATATGCTCGGCTA CCCAGCGCATTTTGGACAGATGGAGTGCAT 300
  K  L  L  Y  I  H  M  L  G  Y   P  A  H  F  G  Q  M  E  C  M
GAAACTGGTGGCTCACCCCCGATTCACTGA CAAACGAATTGGGTATCTTGGAGCTATGCT 360
  K  L  V  A  H  P  R  F  T  D   K  R  I  G  Y  L  G  A  M  L
TCTTCTCGATGAACGATCGGAAGTTCATAT GCTCGTCACTAATTCGCTTAAAAACGACCT 420
  L  L  D  E  R  S  E  V  H  M   L  V  T  N  S  L  K  N  D  L
CACCTGTTCCACCCAATTCGTCAGCGGACT AGCTCTCTGCACTCTGGGCTCAATTTGCTC 480
  T  C  S  T  Q  F  V  S  G  L   A  L  C  T  L  G  S  I  C  S
AGCGGAAATGTGTCGAGACCTGGCCAACGA GGTTGAGAAGATCATCAAGCAGAACAATGC 540
  A  E  M  C  R  D  L  A  N  E   V  E  K  I  I  K  Q  N  N  A
GTATTTAAAGAAAAAAGCGGCGCTCTGTGC GTTCCGTATCGTTCGAAAAGTGCCGGAGCT 600
  Y  L  K  K  K  A  A  L  C  A   F  R  I  V  R  K  V  P  E  L
GATGGAAGTCTTCATTCCATGTACCCGATC GCTTCTGGGAGAGAAAAATCACGGAGTTTT 660
  M  E  V  F  I  P  C  T  R  S   L  L  G  E  K  N  H  G  V  L
GATGGGCGCAACGACTTTGGTCACAGAAAT GTGCGAGAAATCTCCGGATGTCTTGAATCA 720
  M  G  A  T  T  L  V  T  E  M   C  E  K  S  P  D  V  L  N  H
TTTTAAGAAATTGGTGCCGAATTTGGTACG GATTCTAAAGAATCTTCTAATGAGTGGATA 780
  F  K  K  L  V  P  N  L  V  R   I  L  K  N  L  L  M  S  G  Y
TTCGCCGGAACACGAGCTCACCGGCATCTC TGACCCGTTCCTTCAGGTCAAAATTCTGAG 840
  S  P  E  H  E  L  T  G  I  S   D  P  F  L  Q  V  K  I  L  R
ATTATTGAGAGTTTTGGGNAAGGATGACGT ACGTGTCACTGAAGAGATGAATGATATTCT 900
  L  L  R  V  L  G  K  D  D  V   R  V  T  E  E  M  N  D  I  L
GGCGCAAGTGGCAACGAATACGGAAACGGC GAAAAATGTCGGAAACGCGATTCTCTATGA 960
  A  Q  V  A  T  N  T  E  T  A   K  N  V  G  N  A  I  L  Y  E
GACTGTACTCACGATTATGGAGATTAAGAG CGAGAGTGGCCTGCGCATTTTGGCTGTGAA 1020
  T  V  L  T  I  M  E  I  K  S   E  S  G  L  R  I  L  A  V  N
CATTCTTGGACGATTCCTTTTGAATACGGA TAAGAATATTCGATATGTGGCGTTGAATAC 1080
  I  L  G  R  F  L  L  N  T  D   K  N  I  R  Y  V  A  L  N  T
GCTTCTGAAGACTGTTCATGTTGATTATCA GGCCGTACAACGTCACCGTAACGTAGTCGT 1140
  L  L  K  T  V  H  V  D  Y  Q   A  V  Q  R  H  R  N  V  V  V
TGAATGTCTCAAAGATCCGGATATTTCGAT CAGAAAACGTGCAATGGAGCTCTGCTTTGC 1200
  E  C  L  K  D  P  D  I  S  I   R  K  R  A  M  E  L  C  F  A
TCTGATGAACCGTACAAATATCGCAATAAT GACGAAAGAAGTGCTCATTTTCCTCGAGAC 1260
  L  M  N  R  T  N  I  A  I  M   T  K  E  V  L  I  F  L  E  T
CGCCGACGCTGAATTCAAATCGGAATGTGC TTCACGAATGTATATTGCAACGGAAAGATA 1320
  A  D  A  E  F  K  S  E  C  A   S  R  M  Y  I  A  T  E  R  Y
TTCGCCGAATCATGAATGGCATTTGGATAC GATGATTACTGTACTTAGATTGGCCGGAAA 1380
  S  P  N  H  E  W  H  L  D  T   M  I  T  V  L  R  L  A  G  K
ATACGTTCCTGACGAGGTGGTCTCCTGTAT GATCCAAATGATTTCGGCGAATGAGCAACT 1440
  Y  V  P  D  E  V  V  S  C  M   I  Q  M  I  S  A  N  E  Q  L
GCAGAGCTATGCAGTATCCCAGCTTTACCA TGCCGCTCAAAAAGATGCAATCAATGCTCA 1500
  Q  S  Y  A  V  S  Q  L  Y  H   A  A  Q  K  D  A  I  N  A  Q
ACCACTGCTTCAAGTAGCTTTCTGGACAAT TGGAGAATTTGGAGATTTGTTACTTCAACC 1560
  P  L  L  Q  V  A  F  W  T  I   G  E  F  G  D  L  L  L  Q  P
AACAGACGTTGATAGTACTCCAATCTCTGA AAATGATGTTGTCGGAGTATTTGAATCGGT 1620
  T  D  V  D  S  T  P  I  S  E   N  D  V  V  G  V  F  E  S  V
TCTTCCGTCTGCTCTAACCTCACTTTGGAC AAAATGTTATGGAGTTACTGCTCTTGCGAA 1680
  L  P  S  A  L  T  S  L  W  T   K  C  Y  G  V  T  A  L  A  N
CGTGGGAACCAGGTTCCAGTCGACTGGAGA TCGAATTGGAGCACTTGTTCGTATGAATCA 1740
  V  G  T  R  F  Q  S  T  G  D   R  I  G  A  L  V  R  M  N  Q
```

```
GGCTCATATTCAGTTGGAGTTACAGCAGCG ATCTGTGGAATTCAATGTGATTTTGAATTT 1800
  A  H  I  Q  L  E  L  Q  Q  R   S  V  E  F  N  V  I  L  N  L
GGGAGATTTGAGAGACGGACTGCTCGAACG AATGCCAGTAATCACCCATAATTCCCTAAA 1860
  G  D  L  R  D  G  L  L  E  R   M  P  V  I  T  H  N  S  L  N
TGCCGCTGCTCCATCAATGATCGATGAGGA AGTTTCTGCAGAATCGGGAGCACCAGTAGT 1920
  A  A  A  P  S  M  I  D  E  E   V  S  A  E  S  G  A  P  V  V
AACCAATGGAGATCTCCTTGGTGATCTGAA CCTTGGTGGTGGTGGTTCAACGAATCCATC 1980
  T  N  G  D  L  L  G  D  L  N   L  G  G  G  G  S  T  N  P  S
AAATGACTATAGTTCTGATCTTTTGGGTGT TGGAGTTGGTAGTGGAGCAGCTTCTGCTCA 2040
  N  D  Y  S  S  D  L  L  G  V   G  V  G  S  G  A  A  S  A  Q
GGCACCACCACCACCACCACCACCTACCTC AAATTCGAATATTTTGGACATTTTCGGCGA 2100
  A  P  P  P  P  P  P  P  P  T  S   N  S  N  I  L  D  I  F  G  D
TACACCATCTTCAAATGCTGCTGGTGGTTT TGACTTTGGAATGGCTGCTCCAGCGAAGGA 2160
  T  P  S  S  N  A  A  G  G  F   D  F  G  M  A  A  P  A  K  E
GCCCACTTATCAGCCGGTAATCGCTATCAA CAAGGGCGGAATCGAAGTTCAACTGCAAGT 2220
  P  T  Y  Q  P  V  I  A  I  N   K  G  G  I  E  V  Q  L  Q  V
CATCGAAACGTGGAAAAACGAGAAGGCTCG TCTGAGAATGACGGCTTACAACTATACACC 2280
  I  E  T  W  K  N  E  K  A  R   L  R  M  T  A  Y  N  Y  T  P
AAGAACTCTTTCGAACTATAATTTCTTCGC AGCCGTCACAAAAACGTTTGAAATTGCACT 2340
  R  T  L  S  N  Y  N  F  F  A   A  V  T  K  T  F  E  I  A  L
GGAACCGGCTTCATCGCCAAATATTGATCC AAATGAGCATACTACTCAATTTATGACTAT 2400
  E  P  A  S  S  P  N  I  D  P   N  E  H  T  T  Q  F  M  T  I
TACAAGAAAAGCACCTAATACTACTGCTCG CATGCGCACCAAAATCTCGTACATTGTCGA 2460
  T  R  K  A  P  N  T  T  A  R   M  R  T  K  I  S  Y  I  V  D
CGGAACGGAACAAGTCGGCGAAGGAGTTGT CAACGAGTTTCCCGGATTGTAAAAATCTAA 2520
  G  T  E  Q  V  G  E  G  V  V   N  E  F  P  G  L  *
TTTTCCCCCAAAAAAAATTTCCCAAATGCC ACGTTTTTTCCTAATTTATTGATTTTTTTT 2580
TCTATGTGTTTCTAAAAATACCATTTTTGT TCTTATTTTAGTGTGTGTGTGTCTATTGTC 2640
AAATAATGTGAGAAAACTCTAATGGTTGTG ATTATTTTTATTTAATTTTTTTTCTTCAAA 2700
TTCCTGCTGCTAGGGGTTACTGTGATAATA TTTTCCCCTATTTACCCGAAAAAATTCCAA 2760
AAAAATTGCACTCCCGACCCCCACTTTTCC ACCACCAAAAAATCCGCTTTTTCATGTATT 2820
TACCCCTAAAATTGCACAAATAGGCACCTG CCTGCCCACATACCTACAGGCAGTGCCTTG 2880
TAGGTAGGCACGCAGGTAGCAAGCCTACTA TCCTCAATAATTAGTTTTTTATTGGCTAAA 2940
AAANNCAAAAAAATATCGTCTTGGCAGAGT AGAGATTCCCCATTTTTTTTTATTGGATTT 3000
TATTAATAAACCCCCACTTTTTGAGCAAAA AAAAAAAAAA 3040
```

Figure 2. Comparison of the amino acid sequences of the *C. elegans* γ-adaptin and its mammalian homolog, γ-adaptin. The amino terminal half of the proteins is more conserved than the carboxy terminal half.


CEGAMMA ....MRLRDLIRQVRAARTMAEERAVVDRESANIRESFRDDDSPWKCRNIAKLLYIHMLG
γ-ADAPTIN MPAPIRLRELIRTIRTARTQAEEREMIQKECAAIRSSFREEDNTYRCRNVAKLLYMHMLG
CEGAMMA YPAHFGQMECMKLVAHPRFTDKRIGYLGAMLLLDERSEVHMLVTNSLKNDLTCSTQFVSG
γ-ADAPTIN YPAHFGQLECLKLIASQKFTDKRIGYLGAMLLLDERQDVHLLMTNCIKNDLNHSTQFVQG
CEGAMMA LALCTLGSICSAEMCRDLANEVEKIIKQNNAYLKKKAALCAFRIVRKVPELMEVFIPCTR
γ-ADAPTIN LALCTLGCMGSSEMCRDLAGEVEKLLKTSNSYLRKKAALCAVHVIRKVPELMEMFLPATK
CEGAMMA SLLGEKNHGVLMGATTLVTEMCEKSPDVLNHFKKLVPNLVRILKNLLMSGYSPEHELTGI
γ-ADAPTIN NLLNEKNHGVLHTSVVLLTEMCERSPDMLAHFRKLVPQLVRILKNLIMSGYSPEHDVSGI
CEGAMMA SDPFLQVKILRLLRVLGKDDVRVTEEMNDILAQVATNTETAKNVGNAILYETVLTIMEIK
γ-ADAPTIN SDPFLQVRILRLLRILGRNDDDSSEAMNDILAQVATNTETSKNVGNAILYETVLTIMDIK
CEGAMMA SESGLRILAVNILGRFLLNTDKNIRYVALNTLLKTVHVDYQAVQRHRNVVVECLKDPDIS
γ-ADAPTIN SESGLRVLAINILGRFLLNNDKNIRYVALTSLLKTVQTDHNAVQRHRSTIVDCLKDLDVS
CEGAMMA IRKRAMELCFALMNRTNIAIMTKEVLIFLETADAEFKSECASRMYIATERYSPNHEWHLD
γ-ADAPTIN IKRRAMELSFALVNGNNIRGMMKELLYFLDSCEPEFKADCASGIFLAAEKYAPSKRWHID
CEGAMMA TMITVLRLAGKYVPDEVVSCMIQMISANEQLQSYAVSQLYHAAQKDAINAQPLLQVAFWT
γ-ADAPTIN TIMRVLTTAGSYVRDDAVPNLIQLITNSVEMHAYTVQRLYKAILGD.YSQQPLVQVAAWC
CEGAMMA IGEFGDLLL..QPTDVDSTPISENDVVGVFESVLPSALTSLWTKCYGVTALANVGTRFQS
γ-ADAPTIN IGEYGDLLVSGQCEEEEPIQVTEDEVLDILESVLISNMSTSVTRGYALTAIMKLSTRFTC
CEGAMMA TGDRIGALVRMNQAHIQLELQQRSVEFNVIL.NLGDLRDGLLERMPVITHNSLNAAAPSM
γ-ADAPTIN TVNRIKKVVSIYGSSIDVELQQRAVEYNALFKKYDHMRSALLERMPVM..EKVTTNGPSE
CEGAMMA IDEEVSAESGAPVVTNG.....DLLGDLNLGGGGSTNPSNDYSS.....
γ-ADAPTIN IVQTNGETEPAPLETKPPPSGPQPTSQANDLLDLLGGNDITPVIPTAPTSKPASAGGELL
CEGAMMA DLLG.VGVGSGAASAQAPPPPPPPTSNSNILDIFGDTPSSNAAGGFDFGMAAPAKEPTYQ
γ-ADAPTIN DLLGDITLTGAPAAAPTPASVPQISQPPFLLDGLSSQPLFNDIA.....PGIP
CEGAMMA PVIAINKGGIEVQLQVIETWKNEK.ARLRMTAYNYTPRTLSNYNFFAAVTKTFEIALEPA
γ-ADAPTIN SITAYSKNGLKIEFTFERSNTNPSVTVITIQASNSTELDMTDFVFQAAVPKTFQLQLLSP
CEGAMMA SSPNIDPNEHTTQFMTITRKAPNTTA.RMRTKISYIVDGTEQVGEGVVNEFPGL*..
γ-ADAPTIN SSSVVPAFNTGTITQVIKVLNPQKQQLRMRIKLTYNHKGSAMQDLAEVNNFPPQSWQ*

Referennces

Ahle, S., Mann, A., Eichelsbacher, U. and Ungewickell, E. (1988). Structural relationships between clathrin assembly proteins from the Golgi and the plasma membrane *EMBO J.* **7**, 919-929.

Brodsky, F. M. (1988). Living with clathrin: its role in intracellular membrane traffic *Science* **242**, 1396-1402.

Coulson, A., Kozono, Y., Lutterbach, B., Shownkeen, R., Sulston, J. and Waterson, R. (1991). YACs and the *C. elegans* Genome *BioEssays* **13**, 413-417.

Coulson, A., Waterston, R., Kiff, J., Sulston, J. and Kohara, Y. (1988). Genome linking with yeast artificial chromosomes *Nature* **335**, 184-186.

Keen, J. H. (1987). Clathrin assembly proteins: affinity purification and a model for coat assembly *J. Cell Biol.* **105**, 1989-1998.

Keen, J. H. (1990). Clathrin and associated assembly and disassembly proteins *Annn. Rev. Biochem.* **59**, 415-438.

Kirchhausen, T. (1990). Identification of putative yeast homolog of the mammalian b chains of the clathrin-associated protein complexes *Mol. Cell. Biol.*

Kirchhausen, T., Davis, A. C., Frucht, S., Greco, B. O., Payne, G. S. and Tubb, B. (1991). AP17 and AP19, the mammalian small chains of the clathrin-associated protein complexes show homology to YAP17p, their putative homolog in yeast *J. Biol. Chem* **266**, 1153-1157.

Kirchhausen, T., Nathanson, K. L., Matsui, W., Vaisberg, A., Chow, E. P., Burne, Z., Keen, J. H. and Davies, A. E. (1989). Structural and functional division into two domains of the large (100- to 115kDa) chains of the clathrin-associated protein complex AP-2. *Proc. Natl. Acad. Sci. USA* **86**,

2612-2616.

Nakayama, Y., Goebl, M., Greco, B. O., Lemmon, S., Chow, E. P. C. and Kirchhausen, T. (1991). The medium chains of the mammalian clathrin-associated proteins have a homolog in yeast. *Eur. J. Biochem.* **202**, 569-574.

Payne, G. S. and Schekman, R. (1989). Clathrin: a role in the intracellular retention of a Golgi membrane protein *Science* **245**, 1358-1365.

Pearse, B. M. F. and Robinson, M. S. (1990). Clathrin, adaptors, and sorting *Annu. Rev. Biol.* **6**, 151-171.

Robinson, M. S. (1989). Cloning of cDNAs encoding two related 100-kD coated vesicle proteins (a-adaptins) *J. Cell Biol.* **108**, 833-842.

Robinson, M. S. (1990). Cloning and expression of g-adaptin, a component of clathrin-coated vesicles associated with the Golgi apparatus. *J. Cell Biol.* **111**, 2319-2326.

Schmid, S. L. (1992). The mechanism of receptor-mediated endocytosis - more questions than answers. *Bioessays* **14**, 589-596.

Seeger, M. and Payne, G. S. (1992). Selective and immediate effects of clathrin heavy-chain mutations on Golgi membrane-protein retention in Saccharomyces-cerevisiae *J. Cell. Biol.* **118**, 531-540.

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

Appendix II: Interaction of *unc-101* with *daf-1*, a gene encoding a putative receptor serine/threonine kinase involved in dauer formation in *C. elegans*

Because *unc-101* encodes a clathrin-associated protein on the *trans*-Golgi membrane, it is possible that the UNC-101 protein may be involved in other signal transduction processes than the vulval induction pathway. We tested whether *unc-101* interacts with a gene encoding a putative receptor involved in another signaling pathway. Dauer is an alternative, dormant developmental stage of the nematode, in which the nematodes maintain a very low metabolism, and are resistant to harsh environmental conditions such as limited food supply and the lack of moisture (Cassada, 1975, Cassada and Russell, 1975, Golden and Riddle, 1984, Riddle, 1988). Dauers can survive longer than normally developed nematodes. Dauer formation is induced when food supply is limited or there are too many nematodes around. The dauer formation is regulated by many genes that are either required for dauer formation, or required for preventing dauer formation (Riddle, et al., 1981, Vowels and Thomas, 1992). *daf-1* is one of the genes required for preventing dauer formation. Loss-of-function mutations in *daf-1* causes a temperature-sensitive dauer-constitutive phenotype. At 25°C, the mutants animals are 100% dauers, and at 15°C, they are wild type. We chose *daf-1* as our first gene to test for interaction with *unc-101,* because this gene encodes a putative receptor serine/ threonine kinase (Georgi, et al., 1990).

We used three alleles of *daf-1*, *m40, m122,* and *m213* for constructing double mutants with an *unc-101(sy108)* mutation. *m40* is a genetic null allele, and *m213* is a mutation in the kinase domain of the DAF-1 protein, and probably is a molecular null. *m122* is a reduction-of-function allele (D. Riddle, per. comm.). We used the *dpy-9* marker to balance *daf-1.* The genetic distance between the two genes is 0 with error rate of 0.05. We mated *dpy-9* males with *daf-1* hermaphrodites. The male progeny from this mating

were mated with *unc-101; dpy-9* hermaphrodites. In the next generation, we picked non-Dpy non-Unc hermaphrodite progeny, whose genotype is *unc-101/+; daf-1/dpy-9, and* transferred them to 15°C. From the progeny of these mothers, we picked 20 individual Unc non-Dpy animals, and check their genotype by the segregation of Dpy animals at 15°C. The animals that did not segregate any Dpy progeny are of the genotype *unc-101/unc-101; daf-1/daf-1*.

The *unc-101; m40* double mutants and the *unc-101; m213* double mutants had the same dauer phenotype as the *m40* or *m213* single mutants, because at 15°C they are wild type, and at 25°C they are all dauers. However, the *unc-101*; *m122* double mutants displayed an enhanced phenotype. Specifically, at 15°C, about 57 % of the double mutant animals were dauers (494 animals examined). This suggests that *unc-101* interacts with *daf-1* in an allele-specific manner, and that *unc-101* mutations do not suppress or enhance the null mutation of the *daf-1* gene. This result also suggests that *unc-101* is not just involved in an EGF signaling process, but also can be involved in other signaling pathways. The function of *unc-101* may be down-regulation of broad range of cellular signaling processes.

References

Cassada, R. C. and Russell, R. L. (1975). The dauer larva, a post-embryonic developmental variant of the nematode *Caenorhabditis elegans* *Dev. Biol.* **46**, 326-342.

Georgi, L. L., Albert, P. S. and Riddle, D. R. (1990). *daf-1*, a *C. elegans* gene controlling dauer larva development, encodes a novel receptor protein kinase *Cell* **61**, 635-645.

Golden, J. W. and Riddle, D. L. (1984). The *Caenorhabditis elegans* dauer larva: Developmental effects of pheromone, food, and temperature *Dev. Biol.* **102**, 368-378.

Riddle, D. L. (1988). The dauer larva In *The nematode Caenorhabditis elegans* (ed. W. B. Wood), pp. 393-412. Cold Spring Harbor: Cold Spring Harbor Press.

Riddle, D. L., Swanson, M. M. and Albert, P. S. (1981). Interacting genes in nematode dauer larva formation *Nature* **290**, 668-671.

Vowels, J. J. and Thomas, J. H. (1992). Genetic analysis of chemosensory control of dauer formation in *Caenorhabditis elegans Genetics* **130**, 105-123.