The *Caenorhabditis elegans* gene *sem-4* controls neuronal and mesodermal cell development and encodes a zinc finger protein

Michael Basson and H. Robert Horvitz

Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 USA

Neuronal and mesodermal cell types are generated in separate cell lineages during the larval development of *Caenorhabditis elegans*. Here we demonstrate that the gene *sem-4* is required in both types of lineages for the normal development of neuronal and mesodermal cell types. The *sem-4* gene encodes a protein containing seven zinc finger motifs of the C2H2 class, four of which are arranged in two pairs widely separated in the primary sequence of the protein. These pairs of zinc fingers are similar to pairs of zinc fingers in the protein encoded by the *Drosophila* homeotic gene *spalt* and in the human transcription factor PRDII-BF1. Analysis of *sem-4* alleles suggests that different zinc fingers in the SEM-4 protein may function differentially in neuronal and mesodermal cell types. We propose that *sem-4* interacts with different transcription factors in different cell types to control the transcription of genes that function in the processes of neuronal and mesodermal cell development.

[Key Words: *C. elegans*; zinc finger; neuronal development; mesodermal development]

Received March 29, 1996; revised version accepted June 18, 1996.

The large number of different types of neurons in both vertebrate and invertebrate nervous systems raises the question of how the development of a particular neuronal cell type is specified. To pursue this question, Desai et al. (1988; Desai and Horvitz 1989) initiated a genetic analysis of the HSN serotonergic motor neurons in *Caenorhabditis elegans*. This analysis identified genes that are required for a number of distinct aspects of HSN development and placed these genes in a pathway of gene action. Some of these genes encode types of proteins that have been shown to control neuronal development in other organisms. Examples of such genes are *egl-5*, which encodes a homeo domain protein (Wang et al. 1993), *unc-6*, which encodes a member of the netrin family (Serafini et al. 1994), *unc-33*, which encodes a protein that is related to CRMP-62, a chick collapsin response mediator protein (Goshima et al. 1995), and *unc-86*, which encodes a POU homeo domain protein (Finney and Ruvkun 1990). This correlation between the types of proteins that control HSN development and the types of proteins that control neuronal development in other organisms suggests that study of HSN development will reveal general principles of how neuronal cell type is specified.

The POU gene *unc-86* is required for the normal pattern of HSN axonal outgrowth, normal HSN nuclear morphology, and HSN serotonin expression (Desai et al. 1988). POU homeo domain proteins have been shown to be required for neuronal development in both *Drosophila* (Bhat et al. 1995; Yeo et al. 1995) and mammals (for review, see Sharp and Morgan 1996). Mutation of the *C. elegans* gene *sem-4* (sex muscle defective) causes the same spectrum of defects in HSN development as that caused by mutation of *unc-86* (Desai et al. 1988). Because POU homeo domain proteins seem likely to play an evolutionarily conserved role in neuronal development, we are interested in determining how the SEM-4 and UNC-86 proteins interact in controlling HSN development. To understand how *sem-4* functions in HSN development as well as in the development of other cell types, we have performed a genetic and molecular analysis of *sem-4*.

Results

Identification of putative null alleles of *sem-4*.

The first allele of *sem-4*, *n1378*, was isolated in a screen for animals with defective sex muscles (see Materials and methods; M. Stern and H.R. Horvitz, unpubl.). The sex muscle defect of *sem-4(n1378)* animals prevents egg laying. Unlaid eggs hatch inside their parent, resulting in

---

*Corresponding author.*
a bag-of-worms phenotype (Horvitz and Sulston 1980). To obtain additional alleles of sem-4, we isolated mutations that failed to complement n1378 for the bag-of-worms defect (see Materials and methods). We isolated two new alleles of sem-4, n1971, and n2087, and a deficiency that deletes sem-4 genomic DNA, nDf43 (see Materials and methods). Like n1378 animals, animals bearing either of the newly identified sem-4 alleles n1971 or n2087 formed bags-of-worms. Unlike n1378 animals, however, animals bearing either of these newly identified alleles exhibited a second defect: The intestines of these animals became distended as a result of a defect in their defecation behavior, with the consequence that these animals appeared to have a white stripe extending down the length of their bodies when viewed in a dissecting microscope. Animals with this appearance are considered to be constipated (McIntire et al. 1993b). These observations suggested that n1378 might be a weaker allele of sem-4 than the other two alleles, a hypothesis supported by data presented below. n1971 and n2087 are reasonable candidates for being null alleles of sem-4 because they were isolated in a screen for sem-4 mutations in which a deficiency that deletes the sem-4 locus was isolated. Furthermore, in sem-4 heteroallelic animals, n1971 appeared to cause the same reduction in sem-4 function as the deficiency nDf43 that deletes the sem-4 locus [data not shown].

**Figure 1.** Two mesodermal cell types normally generated in the M lineage are not generated in sem-4 animals. (A) M lineage in wild-type first-larval-stage animals. The M mesoblast generates 14 body muscles, two coelomocytes, and two sex myoblasts. (B) M lineage in sem-4(n1971) first-larval-stage animals. The lineage from M.d/v cells was observed in four animals, and the lineage from M.d/v/r cells was observed in four animals. In all eight lineages examined, M.d/v/r/paa cells underwent an extra round of cell division, generating two cells with the appearance of body muscles. M.v/l/r/paa cells, which normally become sex myoblasts, were generated. However, unlike sex myoblasts, they failed to migrate anteriorly during the second larval stage and failed to divide in the third larval stage; instead they retained the appearance of body muscles. In 12 second-larval-stage sem-4(n1971) animals, 12 of 12 M.v/l/r/paa cells were observed to fail to migrate. In 50 third-larval-stage sem-4(n1971) animals, no newly generated muscle cells were observed. Names of cells in bold correspond to those cells affected by sem-4 mutation. Unless otherwise indicated, cell divisions generated daughter cells in anterior and posterior positions relative to the plane of cell division, which are drawn to the left and right, respectively. d, v, l, and r indicate cells generated in dorsal, ventral, left, and right positions relative to the plane of cell division, respectively. For each cell that is generated and then divides, the length of the vertical line associated with that cell is proportional to the time elapsed between cell generation and cell division. A comparison of the timing of cell divisions in wild-type and sem-4 lineages supports the model that in wild-type animals, M.d/v/r/paa cells are transformed into M.v/l/r/paa cells, rather than the model that the M.d cell is transformed into the M.v cell [see text]. (SM) Sex myoblast; (cc) coelomocyte; (bm) body muscle.

**sem-4 is required for the generation of two mesodermal cell types**

sem-4(n1378) animals lack sex muscles (M. Stern and H.R. Horvitz, unpubl.). The sex muscles are derived from a pair of sex myoblasts that migrate anteriorly from their birthplaces in the posterior of the animal to positions near the middle of the animal, where they divide during the third larval stage to give rise to the sex muscles (Sulston and Horvitz 1977). The sex myoblasts, as well as a pair of coelomocytes and 14 body muscle cells, are in turn derived from the M mesoblast during the first larval stage [Fig. 1]. To determine the cause of the absence of sex muscles in sem-4 mutant animals, we examined whether sem-4 animals contained sex myoblasts, as viewed in Nomarski differential interference contrast [DIC] microscopy. We found that 98%–99% of sex myoblasts were absent in sem-4(n1378) and sem-4(n1971) animals (Table 1). Because sex muscles are required for egg laying (Trent et al. 1983), the absence of the sex myoblasts can account for the bag-of-worms defect of sem-4(n1378) and sem-4(n1971) animals. Moreover, we found that the coelomocytes derived from the M mesoblast were also missing in sem-4 mutant animals. Coelomocytes were absent at a significantly higher penetrance in sem-4(n1971) animals than in sem-4(n1378) animals (Table 1), supporting the classification of n1378 as a weak allele of sem-4.

To determine the cause of the lack of sex myoblasts and coelomocytes in sem-4(n1971) animals, we observed the M lineage in sem-4(n1971) animals [Fig. 1]. We found that the cells that normally become sex myoblasts were generated but failed to exhibit the appropriate characteristics of sex myoblasts. They did not migrate anteriorly and did not undergo cell division. Instead, these cells had the appearance of body muscle cells. The cells that would normally become coelomocytes were generated but underwent an extra round of cell division [Fig. 1]. Neither daughter cell acquired the distinctive appearance of a coelomocyte, in which the cell has a round shape, the nucleus is small, and vacuoles are present in the cytoplasm, as viewed with DIC microscopy. Instead, both daughter cells had the appearance of body muscle cells.

We propose that the defects observed in the M lineage...
Table 1. Penetrances of sem-4 defects

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Bag-of-worms(a)</th>
<th>Sex myoblasts(a)</th>
<th>Coelomocytes(a)</th>
<th>HSN serotonin(a)</th>
<th>Constipation(b)</th>
<th>DVB nucleus(b)</th>
<th>DVB GABA(b)</th>
<th>AVL axon(c)</th>
<th>Post-deirid(d)</th>
<th>PHC mec-3::lacZ(c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>0 (201)</td>
<td>0 (204)</td>
<td>1 (205)</td>
<td>0 (176)</td>
<td>0 (132)</td>
<td>2 (205)</td>
<td>0 (189)</td>
<td>0 (105)</td>
<td>1 (205)</td>
<td>0 (132)</td>
</tr>
<tr>
<td>n1378</td>
<td>99 (120)</td>
<td>98 (166)</td>
<td>30 (166)</td>
<td>77 (406)</td>
<td>2 (124)</td>
<td>41 (163)</td>
<td>54 (237)</td>
<td>0 (130)</td>
<td>2 (166)</td>
<td>N.D.</td>
</tr>
<tr>
<td>n1971</td>
<td>95 (120)</td>
<td>99 (165)</td>
<td>96 (166)</td>
<td>98 (200)</td>
<td>97 (120)</td>
<td>91 (157)</td>
<td>98 (241)</td>
<td>89 (83)</td>
<td>17 (164)</td>
<td>96 (94)</td>
</tr>
<tr>
<td>n2654</td>
<td>1 (283)</td>
<td>7 (111)</td>
<td>2 (111)</td>
<td>73 (220)</td>
<td>84 (285)</td>
<td>72 (110)</td>
<td>79 (194)</td>
<td>43 (147)</td>
<td>0 (111)</td>
<td>N.D.</td>
</tr>
</tbody>
</table>

For wild-type animals, 3% had two additional cells; for n1971 animals, 13% had one additional cell, 68% had two additional cells, 13% had three additional cells, and 2% had four additional cells. [N.D.] Not determined.

\(a\)Percentage of animals, sides of animals, or cells that exhibited the indicated defect is listed; the number examined is listed in parentheses.

\(b\)Percentage of animals that formed bags-of-worms.

\(c\)Percentage of sides of third-larval-stage animals that lacked an identifiable sex myoblast, as viewed with DIC microscopy.

\(d\)Percentage of sides of adult animals that did not have a neuron in the position of the HSN that was stained with anti-serotonin antibodies.

\(e\)Percentage of adult animals that had a distended intestine.

\(f\)Percentage of third-larval-stage animals in which the nucleus of K.p, which normally becomes the DVB neuron, had a hypodermal appearance, as viewed with DIC microscopy.

\(g\)Percentage of third-larval-stage or older animals that did not have a neuron in the position of DVB that was stained with anti-GABA antibodies.

\(h\)Percentage of AVL neurons whose axons failed to reach their target region near the anus, as determined by staining with anti-GABA antibodies. Strains contained an unc-30 mutation, which abolished staining of ventral cord motor neurons that would otherwise obscure staining of the AVL axon.

\(i\)Percentage of sides of third-larval-stage animals that lacked an identifiable sex myoblast, as viewed with DIC microscopy.

\(j\)Percentage of sides of third-larval-stage animals in which the postdeirid sensillum contained an extra nucleus with a neuronal appearance, as viewed with DIC microscopy.

\(k\)Percentage of third-larval-stage or older animals containing more than two cells that expressed the mec-3:: lacZ fusion transgene.

of sem-4(n1971) animals are a result of two cell fate transformations. First, the cells that would normally become coelomocytes, M.d/l/r/pa cells, are transformed into their ventral homologs, M.v/l/r/pa cells, which are the mothers of the sex myoblasts. This transformation accounts for the extra round of cell division by M.d/l/r/pa cells and is the same transformation that has been proposed to be caused by mutation of the gene lin-12 (Greenwald et al. 1983). Second, the sex myoblasts are transformed into body muscle cells. This transformation accounts for the observation that both M.v/l/r/paa cells, which normally become sex myoblasts, and M.d/l/r/paa cells, which would be expected to become sex myoblasts as a result of the transformation of M.d/l/r/pa cells described above, have the appearance of body muscle cells.

sem-4 is required for the differentiation of the HSN serotonergic motor neurons

The sem-4 mutation n1378 has been shown to affect the differentiation of the two HSN serotonergic motor neurons, which innervate the sex muscles and stimulate egg laying (Desai et al. 1988). sem-4(n1378) causes defects in three aspects of HSN differentiation. First, as viewed with DIC microscopy, the HSNs in the fourth larval stage fail to acquire a distinctive nuclear morphology in which the nucleus grows in size, the nucleolus becomes visible, and the nucleus becomes half-encircled by a structure that may correspond to the outline of the HSN cell body (hood formation in Desai et al. 1988). Second, most HSNs fail to express serotonin, as scored using anti-serotonin antibodies. Third, the HSNs that do express serotonin have an abnormal pattern of axonal outgrowth.

We found that the putative null allele of sem-4, n1971, also caused defects in HSN nuclear morphology and in serotonin expression. Whereas 100% \((n = 59)\) of the HSNs of wild-type animals had normal morphology in the fourth larval stage, only 5% \((n = 96)\) of the HSNs of sem-4(n1971) animals had normal morphology. sem-4(n1971) caused a more penetrant defect in HSN serotonin expression than did sem-4(n1378) (Table 1), again supporting the classification of n1378 as a weak allele of sem-4. Because sem-4(n1971) abolished serotonin expression, axonal morphology could not be examined in these animals.

Furthermore, we found that HSN nuclear positioning also was defective in sem-4(n1971) animals. The HSN nuclei were often found in positions dorsal to their normal positions in which they are closely apposed to the ventral muscle quadrants: Whereas 97% \((n = 59)\) of the HSNs of wild-type animals were located ventrally, only 25% \((n = 96)\) of the HSNs of sem-4(n1971) animals were located ventrally. The remaining HSNs were located in variable positions, ranging from a position slightly more dorsal than normal to a position dorsal to the midline of the animal.
Defecation by C. elegans hermaphrodites occurs as a result of a series of contractions by three muscle groups [the anterior body muscles, the posterior body muscles, and the enteric muscles]; these contractions are repeated in a cycle that occurs once every 45–50 sec [Thomas 1990; McIntire et al. 1993b]. Because sem-4(n1971) animals were constipated, we inspected the muscle contractions in the defecation cycle of sem-4(n1971) animals and found that both the anterior body and enteric muscle contractions were absent [data not shown].

Because the muscle contraction defects observed in sem-4(n1971) animals are like those produced in wild-type animals by killing two motor neurons that express the neurotransmitter -/aminobutyric acid (GABA), DVB, and AVL [McIntire et al. 1993b], we investigated whether DVB and AVL develop normally in sem-4(n1971) animals. In the case of DVB, we found that K.p, the cell that normally becomes DVB, was generated normally in the first larval stage in 10 of 10 sem-4(n1971) animals examined by staining with anti-GABA antibodies (Table 1). Furthermore, this cell failed to express GABA, as examined by staining with anti-GABA antibodies [Table 1; Fig. 2A–C]. The absence of GABA expression by the K.p cell and the hypodermal appearance of this cell suggested that DVB is at least partially transformed into a hypodermal cell in sem-4(n1971) animals.

In the case of AVL, we found that AVL axonal morphology was defective in sem-4(n1971) animals, as examined by staining with anti-GABA antibodies. In wild-type animals, the AVL axon circles the nerve ring, enters the ventral cord, and extends posteriorly along the length of the ventral cord to reach its target region near the anus [White et al. 1986] (Fig. 2D). In sem-4(n1971) animals, the initial trajectory of the AVL axon around the nerve ring and into the ventral cord was apparently normal. We could not examine the morphology of the AVL axon in the ventral cord of sem-4(n1971) animals, because staining of the DD and VD GABAergic motor neurons in the ventral cord obscures staining of the AVL axon [McIntire et al. 1992]. We therefore used the unc-30(e165) mutation, which abolishes staining of the DD and DD GABAergic motor neurons in the ventral cord [McIntire et al. 1993a]. We found that elongation of the AVL axon along the ventral cord in sem-4(n1971); unc-30(e165) animals was defective: The axon terminated prematurely, failing to reach its normal target region near the anus (Table 1; Fig. 2E,F).

The developmental defects of the DVB and AVL neurons observed in sem-4(n1971) animals can account for the enteric muscle contraction defect and the constipation of these animals. That AVL outgrowth was defec-

Figure 2. The GABAergic neurons DVB and AVL are defective in sem-4 animals. [A] Drawing of the morphology of the DVB neuron in wild-type animals [White et al. 1986]. The DVB cell body is located in the tail (arrow); [B] Posterior region of a wild-type animal stained with anti-GABA antibodies. The DVB cell body (arrow) and motor neurons in the ventral cord [McIntire et al. 1993b] were stained. Anterior is to the left. [C] Posterior region of a sem-4 animal stained with anti-GABA antibodies. Ventral cord motor neurons were stained, but no staining of the DVB cell body was observed. Anterior is to the left. [D] drawing of the morphology of the AVL neuron in wild-type animals [White et al. 1986]. Note that the AVL axon circles the nerve ring, enters the ventral cord, and extends posteriorly along the length of the ventral cord to the region near the anus. [E] unc-30 first-larval-stage animal stained with anti-GABA antibodies. The AVL axon [long arrow] extends along the ventral cord and terminates (arrowhead) in the region of the anus (short arrow). Staining of the DVB neuron was not observed because DVB is not generated until late in the first larval stage [Sulston and Horvitz 1977]. [F] sem-4; unc-30 fourth-larval-stage animal stained with anti-GABA antibodies. The AVL axon [long arrow] extends along the ventral cord but terminates prematurely [arrowhead], before it reaches the region of the anus (short arrow). Note that the sem-4 mutation blocked staining by DVB. The unc-30 mutation was used to abolish GABA expression by motor neurons in the ventral cord that would otherwise obscure visualization of the AVL axon.
tive in sem-4(n1971) animals but not in sem-4(n1378) animals (Table 1) is consistent with the observation that sem-4(n1971) animals, but not sem-4(n1378) animals, were constipated. Because killing DVB alone does not result in constipation (McIntire et al. 1993b; E. Jorgensen and H.R. Horvitz, unpubl.), the partially penetrant DVB defect displayed by sem-4(n1378) animals would not be expected to lead to constipation. On the other hand, the combination of the DVB and AVL defects in sem-4(n1971) animals would be expected to lead to an enteric muscle contraction defect and constipation. The addition of exogenous GABA to sem-4(n1971) animals rescued their enteric muscle contraction defect (E. Jorgensen and H.R. Horvitz, unpubl.), supporting the hypothesis that the defect in enteric muscle contraction results from defects in DVB and AVL function.

sem-4 is required for the development of three types of neurons generated by the T.pp neuroblast

The touch receptor neurons are mechanosensory cells required for an avoidance behavior in which animals that are lightly touched with an eyelash move away from the stimulus (Chalfie and Au 1989). The tails of wild-type animals contain two touch receptor neurons, PLM(L/R). Mitani et al. (1993) showed that the tails of sem-4 animals contain four touch receptor neurons by two criteria: [1] Four axons with the characteristic appearance of touch receptor axons were found in thin sections through the tail using electron microscopy; and [2] four cells were found that expressed two genes, mec-4 and mec-7, that are expressed specifically in the touch receptor neurons and several other types of sensory neurons. We confirmed that thin sections through the tails of sem-4(n1971) animals contained four or five axons with the appearance of touch receptor axons [E. Hartweig, M. Basson, and H.R. Horvitz, unpubl.]. We also examined the expression of the gene mec-3, which is expressed specifically in the touch receptor neurons and several other types of sensory neurons [Way and Chalfie 1989]. We found that four cells in the tails of most sem-4(n1971) animals expressed a mec-3::lacZ fusion transgene (Table 1).

To determine the origin of the extra touch receptor-like neurons in the tails of sem-4(n1971) animals, we took advantage of the observation that the nucleus of a PLM touch neuron in wild-type animals can be identified using DIC microscopy by its position and appearance: The PLM nucleus lies in an extreme lateral position closely apposed to the ventral muscle quadrant and appears to be half-encircled by a structure that may correspond to the outline of the PLM cell body. In wild-type animals, the two PLM neurons are born during embryogenesis. We found that as in wild-type animals, each side of the tail of newly hatched first-larval-stage sem-4(n1971) animals had a single nucleus in the normal position and with the normal appearance of a PLM nucleus. In the second larval stage, however, each side of the tail of sem-4(n1971) animals had two such nuclei, implying that one touch receptor-like neuron was generated on each side of the tail during the first larval stage. Analysis of the cell lineages that occur in the tail during the first larval stage revealed three abnormalities in the cell lineages derived from the T(L/R).pp neuroblasts (Fig. 3A,B). First, the two nuclei that normally belong to the PHC neurons, which have been classified as sensory neurons based on their process morphologies (White et al. 1986), acquired the characteristic appearance of PLM nuclei. This result suggests that the two cells that normally become PHC neurons were transformed into the two extra touch receptor-like neurons. We presume that these extra touch receptor-like neurons, identified by the positions and appearance of their nuclei with DIC microscopy, are responsible for the extra touch receptor axons observed in electron micrographs and correspond to the extra cells that express the mec-4 and mec-7 genes [Mitani et al. 1993] and the mec-3::lacZ fusion transgene. Second, we observed that the fates of the two presumptive PVW neurons varied: Of 20 cells examined, 9 had a normal appearance, 8 died with the morphology characteristic of programmed cell death, and 3 divided. Third, the presumptive PLN neurons, which normally migrate anteriorly, failed to migrate.

Figure 3. The lineage generated by the T(L/R).pp neuroblasts in sem-4 animals is similar to the lineage generated by the AB.p(L/R)apappp neuroblasts in wild-type animals. (A) The T(L/R).pp lineage in wild-type first-larval-stage animals does not generate touch receptor neurons. (PVW) Interneuron; (PHC) sensory neuron; (PLN) interneuron associated with PLM. (B) T(L/R).pp lineages in sem-4 first-larval-stage animals. The two T(L/R).ppa cells, which normally become the PHC neurons, were observed to adopt the appearance of touch receptor neurons in 12 of 12 cases examined. The two T(L/R).ppp cells, which normally become the PVW neurons, were observed to die in 8 of 20 cases examined (see text). The two T(L/R).ppap cells, which normally become the PLN neurons, failed to migrate anteriorly in 14 of 14 cases examined. (mec-3) Cells that have the appearance of touch receptor neurons are presumed to express the mec-3::lacZ fusion transgene (see text). (C) The AB.p(L/R)apappp lineage in wild-type animals. Lineage diagrams are drawn as in Fig. 1. (PLM) Touch receptor neuron that expresses the mec-3::lacZ fusion transgene in wild-type animals; (ALN) interneuron associated with ALM; (X) programmed cell death.
Basson and Horvitz

The abnormalities in the development of both the PHC neurons [their transformation into touch receptor-like neurons], the PVW neurons [their deaths], and the PLN neurons [their failure to migrate] make the lineages derived from the T(L/R).pp cells closely resemble the lineages derived from the AB.p(L/R).apappp cells, which give rise to the PLM neurons [Fig. 3C]. The cell fate transformations caused by mutation of sem-4 suggest that in wild-type animals, sem-4 may act at the level of the T(L/R).pp cells to specify their normal lineages. The AB.p(L/R).apappp cells are anterior homologs of the T(L/R).pp cells closely resemble the PLN neurons [their failure to migrate] make the lineages of the T(L/R).pp cells to specify their normal lineages. The AB.p(L/R).apappp cells are anterior homologs of the T(L/R).pp cells, which are the grandmothers of T(L/R).pp. Thus, sem-4 may act to distinguish the identity of two pairs of related neuroblasts, T(L/R).pp and AB.p(L/R).apappp. Alternatively, sem-4 may act independently in the PHC, PVW, and PLN neurons to specify their normal cell fates.

Three other cell types displayed developmental defects in sem-4 animals

We observed three additional developmental defects by inspecting sem-4(n1971) animals with DIC microscopy and by observing cell lineages that occur during larval development. First, 17% percent of sem-4(n1971) third-larval-stage animals had an extra nucleus with a neuronal appearance in the region of the postdeirid sensillum [Table 1]. To determine the origin of this nucleus, we followed the two lineages from the Q and VS.paa neuroblasts that give rise to the neurons in this region of the animal. Whereas the lineage from neuroblast Q was normal, the lineage from neuroblast VS.paa was abnormal: In 2 of 24 cases examined, the VS.paa cell, which normally undergoes programmed cell death in wild-type animals, instead survived.

Second, extra nuclei with a hypodermal appearance in the region of the pre-anal ganglion were observed in third-larval-stage sem-4(n1971) animals. Third-larval-stage wild-type animals have three nuclei with a hypodermal appearance in this region [the nuclei of P11.p, P12.p, and U]. Among 142 third-larval-stage sem-4(n1971) animals examined, 11 (8%) had three nuclei, 113 (80%) had four nuclei, and 18 (13%) had five nuclei with a hypodermal appearance in this region. Analysis of the lineages that occur in this region [from cells P11, P12, and Y] revealed that the lineages from cells P11 and P12 were normal but that the lineage from the Y cell was abnormal. In first-larval-stage wild-type animals, the nucleus of the Y cell has a hypodermal appearance. In the second larval stage, the Y cell differentiates into the PDA neuron, the nucleus of which has a neuronal appearance. In sem-4(n1971) animals, the Y cell failed to undergo neuronal differentiation, as judged by the failure of the nucleus of the Y cell to acquire a neuronal appearance as viewed with DIC microscopy: In all 11 sem-4(n1971) animals observed from the second larval stage into the third larval stage, the Y cell failed to acquire a neuronal appearance. Failure of the Y cell to undergo neuronal differentiation has also been observed in egl-5 mutants [Chisholm 1991]. The occasional presence of two additional nuclei with a hypodermal appearance in sem-4(n1971) animals may be caused by cell division of the Y cell, as was observed in egl-5 mutants [Chisholm 1991].

Third, we observed extra hypodermal nuclei in the ventral cords of third-larval-stage sem-4(n1971) animals. In wild-type animals, the 12 Pn.p ventral hypodermal cells are generated in the first larval stage; a subset of the Pn.p cells, P[3–8], divide in the third larval stage. In 10 of 13 sem-4(n1971) second-larval-stage animals in which the lineages of the 12 Pn.p cells were followed, at least one Pn.p cell divided precociously. The Pn.p cells that divided precociously in the 10 animals were as follows: P3.p [one animal]; P3.p and P4.p [one animal]; P3.p, P4.p, and P5.p [one animal]; P4.p [two animals]; P4.p and P5.p [two animals]; P4.p, P5.p, and P6.p [one animal]; P5.p [one animal]; and P5.p and P5.p [one animal]. Precocious Pn.p cell division has been reported for a number of other mutants, including lin-25 and lin-31 [Ferguson et al. 1987]. The pattern of precocious cell division in sem-4(n1971) animals, however, is unlike that observed in lin-25 or lin-31: in sem-4(n1971) animals, only P[3–6].p were observed to divide precociously in the second larval stage, whereas in lin-25 animals, P[1–11].p can divide precociously in the first larval stage and in lin-31 animals P[3–8].p can divide precociously in the second larval stage [Ferguson et al. 1987].

Two sem-4 mutations differentially perturb sem-4 function in neuronal and mesodermal cell types

In addition to the weak sem-4 allele n1378 and the putative null sem-4 alleles n1971 and n2087 described above, we received three sem-4 alleles from other investigators: n2477, n2654, and h769 [see Materials and methods]. Alleles n2477 and h769 caused animals to form bags-of-worms and to be constipated at a similar high penetrance as did alleles n1971 and n2087 [data not shown], suggesting that n2477 and h769 may also represent null alleles of sem-4. n2654, however, unlike the other sem-4 alleles, caused animals to be constipated but not to form bags-of-worms. This phenotype suggested that in n2654 animals the GABAergic neurons DVB and AVL are defective but that the sex myoblasts are not. In accordance with this expectation, in n2654 animals 79% of DVB neurons failed to express GABA and 43% of AVL neurons exhibited an axonal outgrowth defect, but only 7% of sex myoblasts were absent [Table 1].

In n2654 animals, neuronal cell types were affected much more severely than were mesodermal cell types [Table 1]: Three neuronal cell types, the HSN, DVB, and AVL neurons, all exhibited defects at moderate to high penetrances, whereas two mesodermal cell types, sex myoblasts and coelomocytes, exhibited defects at a low penetrance. A comparison of the defects caused by the two non-null sem-4 alleles, n1378 and n2654, reveals that whereas n1378 caused more severe defects than n2654 in two mesodermal cell types [sex myoblasts and coelomocytes], n2654 caused more severe defects than n1378 in two of three neuronal cell types [DVB and AVL]. In the third neuronal cell type [HSN], n1378 and n2654 caused nearly equivalent defects [Table 1].
complementary pattern of defects caused by the n1378 and n2654 alleles is illustrated most strikingly by the AVL neuron and the sex myoblasts: Whereas in n2654 animals 43% of AVL neurons exhibited an axonal outgrowth defect, in n2654 of sex myoblasts were absent, in this defect; conversely, whereas in n1378 animals 98% of sex myoblasts were absent, in n2654 animals only 7% of sex myoblasts were absent. These data show that in two mesodermal cell types, the n1378 mutation perturbs sem-4 function more severely than does the n2654 mutation and that in two neuronal cell types, the n2654 mutation perturbs sem-4 function more severely than does the n1378 mutation.

Molecular cloning of sem-4

We mapped sem-4 genetically to a region on LG I between dpy-14 and unc-14 (see Materials and methods). sem-4 is located within the same genetic interval as the restriction fragment length polymorphism hP9, which is detected by the cosmid F15C11 (Starr et al. 1989). We used cosmids from the genomic region around F15C11 (Coulson et al. 1988) in germ-line transformation experiments (data not shown). A 24-kb fragment containing genomic DNA from both cosmids C09F2 and E02H2 rescued the bag-of-worms defect of sem-4(n1378) animals (Fig. 4B; data not shown).

Northern blot experiments using probes derived from the 24-kb genomic fragment identified a 2.9-kb transcript that is likely to be derived from the sem-4 gene. We identified cDNA clones derived from the region of genomic DNA that encodes the 2.9-kb transcript (see Materials and methods). DNA sequence analysis of both cDNA clones and genomic DNA (see Materials and methods) demonstrated the existence of two classes of transcripts, 2864 and 2780 nucleotides in length (Fig. 4B). Both classes of transcripts contain the SL1 leader (Krause and Hirsh 1987) at their 5′ ends. The structures of the two classes of transcripts suggest that they are derived by trans-splicing of the SL1 leader to either of two splice acceptors: Splicing of the SL1 leader to the upstream splice acceptor generates the longer transcript, whereas splicing of the SL1 leader to the downstream splice acceptor generates the shorter transcript.

Sequence analysis of mutant sem-4 alleles confirms the identification of the sem-4 gene

To confirm that the two classes of transcripts described above are derived from the sem-4 gene, we determined the genomic DNA sequence of the region encoding these transcripts from strains containing mutant sem-4 alleles (see Materials and methods). We found a mutation in the genomic DNA of mutants containing each of six different sem-4 alleles (Fig. 4C). These mutations disrupt the predicted open reading frames encoded by sem-4. As described below, both sem-4 transcripts are predicted to encode proteins containing seven zinc finger motifs of the C2H2 class. The mutation n1971, which is a putative null allele by genetic criteria (see above), disrupts the splice donor consensus sequence following the second exon of the longer transcript or the first exon of the shorter transcript. Proteins encoded by the longer and shorter transcripts made from the n1971 allele are predicted to contain only the first 75 amino acids or first 46 amino acids of the wild-type sequence, respectively. Both types of truncated proteins would lack all seven zinc finger motifs. The lesion associated with the n1971 allele therefore supports the hypothesis that n1971 is a
null allele. The alleles n2087, n2477, and h769 all cause phenotypic defects similar to those caused by n1971 and correspond to nonsense mutations that are predicted to result in severely truncated SEM-4 proteins. These truncated proteins would lack all seven (n2477) or six of seven (n2087, h769) zinc fingers. These alleles are therefore likely to represent null alleles. The n1378 and n2654 alleles are discussed below.

The SEM-4 protein contains pairs of zinc fingers that are similar to those in the Spalt and PRDII-BF1 proteins

The two sem-4 transcripts are predicted to encode proteins of 744 and 715 amino acids in length, both of which contain seven zinc finger motifs of the C2H2 class (Klug and Schwabe 1995; Miller et al. 1985; Figs. 4A and 5). These seven zinc fingers are distributed widely in the primary sequence of the SEM-4 protein and occur either singly (fingers 1, 4, and 7) or in pairs (fingers 2 and 3 and fingers 5 and 6); the two fingers in each pair are separated by the conserved linker sequence found commonly between zinc fingers (Schuh et al. 1986). The spacing and sequence of zinc fingers 2–7 suggest that the sem-4 gene may have arisen by duplication of an ancestral gene containing three zinc fingers: The spacing between fingers 2–4 is similar to the spacing between fingers 5–7, and fingers 2 and 5 and fingers 3 and 6 are highly similar in sequence (Fig. 5).

The two pairs of zinc fingers in SEM-4 (fingers 2 and 3 and fingers 5 and 6) are highly similar in sequence to pairs of zinc fingers in the protein encoded by the Drosophila homeotic gene Spalt, which specifies pattern elements in both the head and tail regions (Jürgens 1988; Kühnlein et al. 1994) (Fig. 5B). The two pairs of zinc fingers in SEM-4 are, to a lesser degree, also similar to pairs of zinc fingers in the human protein PRDII–BF1 (Fan and Maniatis 1990), which has also been named MBP-1 (Baldwin et al. 1990) and HIV–EP1 (Mackawa et al. 1989) (Fig. 5C). PRDII–BF1 binds to the HIV-1 enhancer and related promoter and enhancer elements in several genes (e.g., Fan and Maniatis 1990) and can stimulate transcription from the HIV-1 long terminal repeat (Seeler et al. 1994). SEM-4, Spalt, and PRDII–BF1 are structurally similar in that all three proteins contain widely spaced pairs of zinc fingers (Fig. 5A). The two pairs of zinc fingers in PRDII–BF1 have been shown to bind to similar DNA sequences [Fan and Maniatis 1990]. Because the second finger in each pair of zinc fingers in the SEM-4, Spalt, and PRDII–BF1 proteins share sequence similarity in the region that, in other zinc fingers of the C2H2 class, has been shown to form an α-helix that makes sequence-specific contacts with DNA (Klug and Schwabe 1995), the pairs of zinc fingers in all three proteins may bind to similar sites. Except for the zinc fingers, no sequence similarity between SEM-4 and other proteins has been detected.

PRDII–BF1 is closely related to several other zinc-finger DNA-binding proteins that bind to similar target sequences and that contain widely spaced pairs of zinc finger motifs that are related in sequence. This family includes the human proteins MBP-2 (van’t Veer et al. 1992) and KBP-1 (van’t Veer et al. 1992), as well as the Drosophila protein Schnurri (Arora et al. 1995; Grieder et al. 1995), which acts in the signaling pathway mediated by the transformation growth factor-β (TGF-β)-related growth factor Decapentaplegic. The zinc fingers in SEM-4 and Spalt are more similar to each other than to the zinc fingers in PRDII–BF1, MBP-2, KBP-1, or Schnurri, indicating that SEM-4 and Spalt are more closely related to each other than to the other proteins.

Sequence analysis of two weak sem-4 alleles suggests that different zinc fingers function differentially in neuronal and mesodermal cell types

The weak allele n1378 corresponds to a nonsense mutation that is predicted to truncate the SEM-4 proteins, removing the carboxy-terminal 176 amino acids (Fig. 4). The region of the proteins removed by n1378 contains three zinc fingers, two of which (fingers 5 and 6) constitute the carboxy-terminal pair of zinc fingers in SEM-4. The defects caused by n1378 may therefore result, at least in part, from the absence of this pair of zinc fingers in the protein encoded by the n1378 allele and from the loss of DNA-binding activity that this pair of zinc fingers might normally confer.

The weak allele n2654 corresponds to a missense mutation affecting one of the histidine residues of the sec-
ond zinc finger [Fig. 4], which is in the amino-terminal pair of zinc fingers in SEM-4. Because this histidine residue is one of those in the zinc finger consensus sequence that is involved in binding zinc, n2654 is likely to destroy the function of the second zinc finger. As n2654 affects mesodermal cell types at a very low penetrance but neuronal cell types at a moderate to high penetrance (Table 1), the function of the second zinc finger may be required to a higher degree in neuronal cell types than in mesodermal cell types. Furthermore, as discussed above, n2654 causes more severe defects than does n1378 in neuronal cell types, and n1378 causes more severe defects than does n2654 in mesodermal cell types. These data suggest that the amino-terminal pair of zinc fingers may play a more important role in neuronal cell types than does the carboxy-terminal pair of zinc fingers, and that the reverse may be true in mesodermal cell types.

Discussion

sem-4 animals exhibit severe defects in both their egg-laying and defecation behaviors. We have shown that these behavioral defects are caused by underlying defects in the development of a mesodermal cell type, the sex myoblasts, and of two classes of neurons, the DVB and AVL GABAergic motor neurons. For these and other cell types, lack of sem-4 function has widely varying developmental consequences [Table 2]. In the case of neuronal cell development, sem-4 can act at several different levels in the hierarchy of developmental events required to generate a specific neuronal cell type: For the DVB motor neuron, to adopt a neuronal or a non-neuronal cell fate; for the PHC sensory neuron, to adopt the fate of a specific neuronal cell type; and for the HSN and AVL motor neurons, to execute the differentiation events appropriate to a specific neuronal cell type. We do not know whether sem-4 functions within the cells affected by mutation or in other cells that influence the development of the affected cells.

The cells affected by mutation of sem-4 do not share any obvious similarities that would suggest a single developmental role for sem-4. First, sem-4 is required for the normal development of a diverse set of cell types, including neuronal, mesodermal, and hypodermal cells (Table 2). Second, sem-4 is required for the development of cells found in several different regions of the body: Most cell types affected by sem-4 mutation are located in the posterior region of the animal and in the tail, but one cell type (AVL) is located in the head and one (Pn.p) is located in the mid-body region. Third, the embryonic cell lineages that give rise to the cells affected by mutation of sem-4 are not obviously related. Thus, the cells that require sem-4 function are not obviously related by cell type, position in the body, or position in the cell lineage.

In several cases, the developmental defects caused by mutation of sem-4 can be interpreted as the transformation of a cell into a cell resembling its sister cell. For example, the neuron DVB is transformed into a hypoder-

### Table 2. Summary of the phenotype of sem-4 animals

<table>
<thead>
<tr>
<th>Cell type</th>
<th>Defect or cell fate transformation caused by sem-4 mutation*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Neuronal cell types</strong></td>
<td></td>
</tr>
<tr>
<td>HSN serotonergic motor neuron</td>
<td>defective differentiation</td>
</tr>
<tr>
<td>DVB GABAergic motor neuron</td>
<td>transformed to hypodermal cell</td>
</tr>
<tr>
<td>AVL GABAergic motor neuron</td>
<td>defective axonal outgrowth</td>
</tr>
<tr>
<td>PHC sensory neuron</td>
<td>transformed to touch receptor neuron</td>
</tr>
<tr>
<td>PVW neuron</td>
<td>transformed to programmed cell death</td>
</tr>
<tr>
<td>PDA neuron</td>
<td>transformed to hypodermal cell</td>
</tr>
<tr>
<td><strong>B. Mesodermal cell types</strong></td>
<td></td>
</tr>
<tr>
<td>Sex myoblast</td>
<td>transformed to body muscle cell</td>
</tr>
<tr>
<td>Coelomocyte</td>
<td>transformed to myoblast</td>
</tr>
<tr>
<td><strong>C. Other cell types</strong></td>
<td></td>
</tr>
<tr>
<td>V5.paapp cell death</td>
<td>transformed to neuronal cell</td>
</tr>
<tr>
<td>Pn.p hypodermal cells</td>
<td>precocious cell division</td>
</tr>
</tbody>
</table>

The cell types affected by sem-4 mutation and the defects or cell fate transformations caused by sem-4 mutation in the development of those cell types are listed [see text for penetrances].

*The cell fate transformations listed are suggested by the following criteria, as described in the text: the appearance of the cell in DIC microscopy (DVB, PHC, PVW, PDA, sex myoblast, coelomocyte, V5.paapp); the appearance of the cell in electron microscopy (PHC); and the expression of cell-type specific markers (DVB, PHC). sem-4 mutation caused also two additional defects: locomotory defects and misplacement of the nuclei of the P11/12(L/R) hypodermal cells [see Materials and methods].

mal cell resembling its sister, the K.a. cell. Such a transformation may also be the case for sex myoblasts, which are transformed into body muscles resembling their sisters, and the V5.paapp cell, which fails to die, has a neuronal appearance, and expresses UNC-86 protein [M. Basson and H.R. Horvitz, unpubl.], thereby resembling its sister, the PVD neuron. Thus, sem-4 may play a role in processes that make sister cells different from each other, like a number of other C. elegans genes that have been proposed to act in these processes [Horvitz and Herskowitz 1992]. However, in other cases, mutation of sem-4 does not appear to cause the transformation of a cell into a cell resembling its sister cell. The HSN neuron fails to differentiate normally but is not obviously transformed into a cell resembling its sister cell, the PHB neuron. Coelomocytes appear to be transformed not into cells resembling their sisters but rather into cells resembling their ventral homologs, the mothers of the sex myoblasts. The transformations of the PVW and PHC neurons into cells that undergo programmed cell death and...
become a touch receptor neuron, respectively, are consistent not with sister–sister cell fate transformations but rather with transformations of T[L/R].pp cells into cells resembling their approximate anterior homologs, AB.p(1/r)apappp cells.

sem-4 is predicted to encode a protein containing seven zinc finger motifs of the C2H2 class. Two pairs of zinc fingers in SEM-4 show sequence similarity to pairs of zinc fingers in the Drosophila protein Spalt and in the PRDII–BF1 class of zinc finger-containing proteins. Kühnlein et al. (1994) noted three structural similarities between the Spalt protein and the human DNA-binding protein PRDII–BF1: (1) Both Spalt and PRDII–BF1 contain widely spaced pairs of zinc fingers; (2) pairs of fingers in each protein are highly similar in sequence to the other pairs of fingers within the same protein, and (3) the second finger in each pair is similar in sequence in both proteins. The SEM-4 protein shares all three of these structural features. SEM-4, Spalt, and the PRDII–BF1-related proteins may therefore constitute members of a subfamily of zinc finger proteins containing pairs of zinc fingers that are similar in sequence and that may bind to similar sites.

Both Spalt (Lecuit et al. 1996; Nellen et al. 1996) and the PRDII–BF1 family member Schnurri (Arora et al. 1995; Grieder et al. 1995) have been shown to be targets of signaling by the TGF-β-like protein Decapentaplegic. The sequence similarities among SEM-4, Spalt, and PRDII–BF1 family members raise the possibility that related proteins may therefore constitute members of a subfamily of zinc finger proteins containing pairs of zinc fingers that are similar in sequence and that may bind to similar sites.

The developmental defects observed in sem-4 mutant animals suggest that the putative SEM-4 transcription factor controls the transcription of genes that function in neuronal, mesodermal, and hypodermal cell development. The pleiotropic nature of the sem-4 phenotype raises the question of how SEM-4 acts in different cell types to control different developmental processes. One possibility is that the putative SEM-4 transcription factor is expressed in the neuronal, mesodermal, and hypodermal cells that require sem-4 function and that the outcome of SEM-4 expression in these cell types depends on the differential expression of other transcription factors. Analysis of two sem-4 mutations suggests that in the SEM-4 protein pairs of zinc fingers might function differentially in different cell types. For example, a mutation that is likely to abolish the function of a zinc finger in the amino-terminal pair of zinc fingers prevents function in the AVL neuron but not in the sex myoblasts, whereas a mutation that removes the carboxy-terminal pair of zinc fingers prevents function in the sex myoblasts but not in the AVL neuron. The observation that the two pairs of zinc fingers in SEM-4 are similar to each other suggests that they may bind to similar DNA sequences, as has been demonstrated for the two pairs of zinc fingers that are similar to each other in PRDII–BF1 (Fan and Maniatis 1990). Therefore, the differential function of the two pairs of zinc fingers in SEM-4 might lie not in their recognizing different binding sites but rather in their interaction with other proteins that are expressed in different cell types.

As zinc finger proteins have been shown to interact with homeo domain proteins to control transcription, such as in the cases of the zinc finger protein Zn-15 and the POU homeo domain protein Pit-1 (Lipkin et al. 1993) and the zinc finger protein Hunchback and the homeo domain protein Bicoid (Simpson-Brose et al. 1994), an attractive model for SEM-4 function is that it interacts with different homeo domain proteins in different cell types. One such cell type might be the PDA neuron. Because PDA differentiation requires both sem-4 function and egl-5 function (Chisholm 1991), SEM-4 might act with the homeo domain protein EGL-5 (Wang et al. 1993) in this neuron. A second such cell type might be the HSN neurons. HSN differentiation requires both sem-4 and unc-86 function (Desai et al. 1988), and the UNC-86 POU homeo domain protein has been shown to be expressed in HSN neurons (Finney and Ruvkun 1990). As a sem-4::lacZ fusion transgene is also expressed in HSN neurons (M.E. Basson and H.R. Horvitz, unpubl.), coexpression of these two transcription factors and perhaps other transcription factors as well in this neuronal cell type could together control the transcription of genes that function specifically to execute the HSN differentiation program.

Materials and methods

Genetic methods

Worms were maintained as described by Brenner (1974) at 20°C unless indicated otherwise. The following mutations were used: LG I, dpy-5(e61), dpy-14(e188ts), unc-14(e57), unc-13(e51); LG IV, unc-30(e165); LG X, lon-2(e678), unc-9(e101) (Hodgkin et al. 1988). We also used the mutations let-389(h428) (Howell et al. 1987), let-400(h269) I, let-520(h690) I (McKim et al. 1992), and xol-1(y9) X (Miller et al. 1988). We used the deficiency hDH8 (Starr et al. 1989) and the free duplications sDp2 (Rose et al. 1984), hDP58, and hDP62 (McKim and Rose 1990). The sem-4 alleles used are described below.

Isolation of sem-4 alleles

sem-4(n1378) was isolated in a screen for animals with defective sex muscles as assessed by polarized light microscopy [M. Stern and H.R. Horvitz, unpubl.]. n1791, n2087, and n2088 were isolated in two screens for mutations that failed to complement the bag-of-worms defect of n1378. In the first screen, sem-4(n1378) unc-13(e51) heterozygous males were mated with mutagenized unc-9(e101) hermaphrodites, and non-Unc progeny were screened for animals that formed bags-of-worms. One new allele, n1791, was recovered from a screen of ~4400 haploid genomes. In the second screen, wild-type males were mutagenized and mated with dpy-5(e61) sem-4(n1378) unc-13(e51); lon-2(e678) xol-1(y9) hermaphrodites, and non-Unc progeny were screened for animals that formed bags-of-worms. Two new sem-4 alleles, n2087 and n2088, were recovered from a screen of ~17,500 haploid genomes. In both screens, worms were muta-
genized with ethyl methanesulfonate [Brenner 1974]. n2088 was renamed as the deficiency nDf43, because it failed to complement mutations in four different complementation groups: sem-4(n1378), let-389(h769), let-400(h269), and let-520(h690). sem-4(n1378) complemented let-389(h248), let-400(h269), and let-520(h690). nDf43 complemented dpy-14(e188ts) and unc-14(e57). n2477 and n2654 were isolated in screens for animals with defective defecation behavior [E. Jorgensen and H.R. Horvitz, unpubl.]. h769 was isolated in a screen for lethal mutations on LG I [McKim et al. 1992]. Our analysis of h769 revealed that this mutation did not cause lethality; like other sem-4 alleles, h769 caused animals to form bags-of-worms and to be constipated.

Other aspects of the sem-4 phenotype

Two aspects of the phenotype of sem-4(n1971) animals were noted that are not discussed in Results. First, these animals were uncoordinated: They were slightly sluggish, and their bodies made deeper bends than is normal. These defects were observed also in animals containing each of the other putative null alleles. The cellular defects causing these locomotory defects are not known. Second, the nuclei of the P11/12(L/R) hypodermal cells in newly hatched first-larval-stage animals were in more posterior positions than in wild-type animals, as viewed with DIC microscopy. The cell lineages derived from P11/12(L/R), however, were normal. The positions of P11/12(L/R) and the lineages derived from them were not examined in animals containing other sem-4 alleles.

sem-4 genetic mapping

sem-4(n1378) had been shown to lie between dpy-5 and unc-13 on LG I [M. Stern and H.R. Horvitz, unpubl.]. We showed that sem-4(n1378) mapped between dpy-14 and unc-14: From a heterozygote of genotype dpy-14(e188ts) unc-14(e57)/sem-4(n1378), 21/29 Unc non-Dpy recombinants segregated sem-4(n1378). We then localized sem-4 using available deficiencies and duplications: We showed that sem-4 complemented the deficiency hDp8, which was not complemented by the duplication hDp62, and was complemented by the duplications hDp58 and SDp2.

DIC microscopy and lineage analysis

The positions and appearances of nuclei in wild-type animals and cell lineages in wild-type animals have been described [Sulston and Horvitz 1977; Sulston et al. 1983]. We observed cell lineages in living animals using DIC optics. We classified a nucleus as neuronal if it was small, had no obvious nucleolus, and had a granular nucleoplasm; we classified a nucleus as hypodermal if it was large, had a nucleolus, and had a smooth nucleoplasm.

Immunocytochemistry

HSN serotonin expression was scored by staining with anti-serotonin antibodies [McIntire et al. 1992]. GABA expression in the neuron DVB was scored, and axonal morphology of the neuron AVL was visualized by staining with anti-GABA antibodies [McIntire et al. 1992]. Antibodies were detected by staining with FITC-labeled goat-anti-rabbit antibodies (Cappel, Inc., West Chester, PA) followed by staining with FITC-labeled rabbit-anti-goat antibodies [Cappel Inc.]. Stained worms were mounted with an equal volume of 1 mg/ml of p-phenylenediamine and viewed with fluorescence optics.

Expression of the mec-3::lacZ fusion transgene

jeln2 is a mec-3::lacZ fusion transgene that is integrated into the genome [Way et al. 1991] and was generously provided by J. Way [Rutgers University, New Brunswick, NJ]. The expression pattern of the mec-3::lacZ transgene in wild-type animals has been described [Way and Chalfie 1989, Way et al. 1991]. The expression pattern of the transgene was determined by staining worms for β-galactosidase activity using the substrate X-gal [5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside] [Fire et al. 1990]. Strains containing jeln2 were grown at 25°C, which results in more reliable staining [Way et al. 1991].

nDf43 deletes sem-4 genomic DNA

To demonstrate that nDf43 deletes sem-4 genomic DNA, we showed that sem-4 genomic DNA could not be amplified from single nDf43 homozygous animals using the polymerase chain reaction (PCR) technique. nDf43 homozygous animals arrest development as young larvae. We picked either wild-type or homozygous dpy-5 nDf43 animals [derived from dpy-5 nDf43/ dpy-14 unc-14] and performed PCR on single animals [Williams et al. 1992] using two sets of oligonucleotide primers derived from the sem-4 cDNA sequence. For both sets of primers, DNA could be amplified from 6/6 wild-type animals and from 0/6 dpy-5 nDf43 animals. A set of primers from the gene ces-1 [Ellis and Horvitz 1991] served as a positive control.

Molecular biology of sem-4

sem-4 cDNA clones were obtained from a kgt1 cDNA library constructed from mixed-stage poly[A]⁺ RNA [Okkema and Fire 1994]. Eight different cDNA clones were isolated that defined four groups. The first group contained one clone that began with the final 7 nucleotides of the SL1 leader [Krause and Hirsh 1987] followed by a transcript containing 10 exons as depicted by the longer transcript shown in Figure 4B. The second group contained four clones that began within the first exon of the longer transcript. The third group contained one clone that began with the final 7 nucleotides of the SL1 leader followed by a transcript containing 9 exons as depicted by the shorter transcript shown in Figure 4B. The fourth group contained two clones that began within the second exon of the longer transcript.

Genomic DNA sequence was determined using oligonucleotide primers designed using the sequence of the sem-4 cDNA. The genomic DNA sequence positioned the open reading frame of the cDNA clone in the 24-kb SphI–BamHI fragment of genomic DNA and identified the exon–intron boundaries of the gene [Fig. 4B].

To determine the sequence of sem-4 mutant alleles, genomic DNA from wild-type animals and from animals containing one of the six mutant sem-4 alleles [n1378, n1971, n2087, n2477, n2654, or h769] was isolated as described [Williams et al. 1992]. Genomic DNA was amplified by PCR using five pairs of oligonucleotide primers designed to amplify the entire sem-4 open reading frame and all of the exon–intron boundaries. Primers were then used to determine the sequence of both strands of the entire sem-4 open reading frame in wild-type and mutant animals. The sequences of the oligonucleotide primers used are available upon request. The sequences of the long and short sem-4 transcripts have been deposited in the GenBank data base (accession nos. U60112 and U60113, respectively).

Acknowledgments

We are grateful to Michael Stern for generously giving us sem-4(n1378) to begin this work. We thank Ann Rose and Kim
Basson and Horvitz

McKim for giving us a semi-4 allele and for generously sharing strains and unpublished data concerning the genetic and physical maps of the cluster of LG I. Marty Chalfie and Shohei Mitani for sharing unpublished data concerning touch receptor neurons in semi-4 animals, Jeff Way for the mec-3::lacZ transgene, Erik Jorgensen for two semi-4 alleles, Mark Metzstein for ces-1 primers, Erika Hartwig for assistance with photography, Beth James for DNA sequencing, and Gian Garriga for poly[A]⁺ RNA and for advice throughout the course of this work. This work was supported by U.S. Public Health Service research grants GM34660 to H.R. Horvitz. E.J. was supported by postdoctoral fellowships from the Jane Coffin Childs Fellowship and the Charles A. King Trust. H.R.H. is an Investigator of the Howard Hughes Medical Institute.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

References


sem-4 controls neuronal and mesodermal development


The Caenorhabditis elegans gene sem-4 controls neuronal and mesodermal cell development and encodes a zinc finger protein.

M Basson and H R Horvitz

Genes Dev. 1996, 10: 
Access the most recent version at doi:10.1101/gad.10.15.1953