Development 130, 3831-3840

doi:10.1242/dev.00398

The *Caenorhabditis elegans spalt*-like gene *sem-4* restricts touch cell fate by repressing the selector Hox gene *egl-5* and the effector gene *mec-3*

Anne S. Toker[1], Yingqi Teng[2], Henrique B. Ferreira[2,*], Scott W. Emmons[2] and Martin Chalfie[1,†]

[1]Department of Biological Sciences, Columbia University, New York, NY 10027, USA
[2]Department of Molecular Genetics, Albert Einstein College of Medicine, Bronx, NY 10461, USA
*Present address: Departamento de Biologia Molecular e Biotecnologia and Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil
†Author for correspondence (e-mail: mc21@columbia.edu)

Accepted 13 January 2003

SUMMARY

Members of the spalt (sal) gene family encode zinc-finger proteins that are putative tumor suppressors and regulate anteroposterior (AP) patterning, cellular identity, and, possibly, cell cycle progression. The mechanism through which sal genes carry out these functions is unclear. The *Caenorhabditis elegans* sal gene *sem-4* controls the fate of several different cell types, including neurons, muscle and hypodermis. Mutation of *sem-4* transforms particular tail neurons into touch-neuron-like cells. In wild-type *C. elegans*, six touch receptor neurons mediate the response of the worm to gentle touch. All six touch neurons normally express the LIM homeobox gene *mec-3*. A subset, the two PLM cells, also express the Hox gene *egl-5*, an *Abdominal-B* homolog, which we find is required for correct *mec-3* expression in these cells. The abnormal touch-neuron-like-cells in *sem-4* animals express *mec-3*; we show that a subset also express *egl-5*.

We report: (1) that ectopic expression of *sem-4* in normal touch cells represses *mec-3* expression and reduces touch cell function; (2) that *egl-5* expression is required for both the fate of normal PLM touch neurons in wild-type animals and the fate of a subset of abnormal touch neurons in *sem-4* animals, and (3) that SEM-4 specifically binds a shared motif in the *mec-3* and *egl-5* promoters that mediates repression of these genes in cells in the tail. We conclude that *sem-4* represses *egl-5* and *mec-3* through direct interaction with regulatory sequences in the promoters of these genes, that *sem-4* indirectly modulates *mec-3* expression through its repression of *egl-5* and that this negative regulation is required for proper determination of neuronal fates. We suggest that the mechanism and targets of regulation by *sem-4* are conserved throughout the sal gene family: other sal genes might regulate patterning and cellular identity through direct repression of Hox selector genes and effector genes.

Key words: spalt, Repressor, Hox, LIM homeodomain, AP patterning

INTRODUCTION

Members of the spalt (sal) gene family encode zinc-finger proteins that both control normal development and appear to function as tumor suppressors in human beings and mice. *Drosophila sal* is required for many aspects of development, including the establishment of head and tail identity (Jurgens, 1988), and photoreceptor differentiation in the eye (Mollereau et al., 2001). In *Caenorhabditis elegans*, the sal gene *sem-4* controls the fate of several different cell types including neurons, muscle and hypodermis (Basson and Horvitz, 1996). The human gene *SALL1* (also known as *Hsal1*) is mutated in patients with Townes-Brocks syndrome (TBS), an autosomal dominant developmental disorder characterized by sensorineural hearing loss and by malformations of the limbs, anus, kidneys, heart and gonads (Kohlhase, 2000; Ma et al., 2001b; Surka et al., 2001). SALL1 and mouse sall1 act as repressors in cell culture assays (Kiefer et al., 2002; Netzer et al., 2001). Another human sal gene, *HSAL2* (also known as *SALL2*), maps to a region associated with human ovarian cancers, and alterations in the expression of *HSAL2* have been found in human ovarian carcinoma (Ma et al., 2001a). The mouse homologue of *HSAL2*, *msal-2*, is a target of polyoma tumor virus large T antigen (Li et al., 2001).

Three fundamental questions about the functions of *sem-4* and the other sal genes remain unresolved: (1) do sal genes bind DNA; (2) do they function as repressors in vivo; and (3) what are their targets? This work identifies the *C. elegans* sal gene *sem-4* as a regulator of *C. elegans* homeobox gene expression and examines the mechanism through which sal genes control cellular identity. Like other sal genes, *sem-4* is required for multiple aspects of development. The strongest alleles of *sem-4* are egg-laying defective (*Egl*), uncoordinated (*Unc*), partially sterile and constipated, and have deformed tails (Basson and Horvitz, 1996). Animals with *sem-4* mutations exhibit abnormalities in a range of different cell types, including neurons, muscle cells, coelomocytes and vulval cells (Basson and Horvitz, 1996; Grant et al., 2000). In particular, *sem-4* animals produce additional touch-neuron-like cells that

express the touch cell effector gene *mec-3* (Basson and Horvitz, 1996; Mitani et al., 1993; Mitani, 1995).

We report that, in *sem-4* animals, the abnormal touch-neuron-like cells and their precursors ectopically express the Hox gene *egl-5*, an *Abdominal-B* (*Abd-B*) homolog. We show that inappropriate expression of *egl-5* transforms the fates of these neuroblasts and neurons, that SEM-4 binds to a shared motif in the *mec-3* and *egl-5* promoters, and that ectopic *sem-4* represses *mec-3* expression in vivo. Our findings point to three conclusions: first, that *sem-4* and other sal genes are repressors that control cellular identity by restricting expression of Hox genes and effector genes; second, that sal genes independently regulate these genes at multiple stages in developmental pathways; and, third, that sal proteins bind directly to a shared motif in regulatory regions of their targets.

MATERIALS AND METHODS

Nematode strains and maintenance

Wild-type *C. elegans* (var. Bristol, N2) and mutant strains were grown as described by Brenner (Brenner, 1974). The following mutant strains were used: CB3531 [*mab-5(e1239)III*; *him-5(e1490)V*] (Kenyon, 1986); EM597 [*him-5*(*e1490*)*V*; *lin-15*(*n765*)*X*; *bxIs12*(*egl-5::gfp lin-15*(+))]; EM783 [*pha-1*(*e2123*)*III*; *him-5*(*e1490*)*V*; *bxEx87*(*egl-5::gfp pha-1*(+))]; EM784 [*pha-1*(*e2123*)*III*; *him-5*(*e1490*)*V*; *bxEx88* (*P*V6CRE*gfp pha-1*(+))]; EM785 [*pha-1*(*e2123*)*III*; *him-5*(*e1490*)*V*; *bxEx89* (*P*V6CREΔ100*gfp pha-1*(+))]; MH1346 [*unc-119*(*ed3*)*III*; *kuIs35* (*sem-4::gfp unc-119*(+))] (Grant et al., 2000); MT1081 [*egl-5(n486)III*] (Trent et al., 1983); MT1514 [*lin-39(n709)III*] (Li and Chalfie, 1990); MT3179 [*sem-4(n1378)I*] (Desai et al., 1988); MT5825 [*sem-4(n1971)I*] (Basson and Horvitz, 1996); MT5826 [*sem-4(n2087)I*] (Basson and Horvitz, 1996); MT6921 [*sem-4(n2654)I*] (Basson and Horvitz, 1996); TU2562 [*dpy-20*(*e1282*)*IV*; *uIs22*(*mec-3::gfp dpy-20*(+))] (Wu et al., 2001). Double mutant strains were created according to standard genetic methods (Brenner, 1974).

Phenotypic characterization

Cell lineages were followed as described by Sulston and Horvitz (Sulston and Horvitz, 1977). Expression of *gfp* reporters was observed at 1000× magnification. The touch sensitivity assay was modified from Hobert et al. (Hobert et al., 1999). Each animal was touched with an eyebrow hair ten times alternately at the head and tail. At least 100 animals were scored for each stable line or mutant strain. To ensure that observations of cells in temperature-sensitive *sem-4* and Hox mutant strains were not influenced by maternal effects, we scored worms from at least the third generation grown at each particular temperature. At least 45 gravid adults were scored for each strain at each temperature.

Expression studies

Expression studies using *sem-4* were carried out using a PCR-amplified fragment of the *sem-4* genomic sequence from cosmid F15C11 that was cloned into the *Pst*I and *Kpn*I sites of Fire vector pPD 95.75 (Fire et al., 1990). The fragment contained 5 kb upstream of the *sem-4* start site and the entire genomic sequence fused at the 3′ end to *gfp*. This construct was injected into N2 worms and two stable lines were obtained. It was difficult to produce stable lines containing this expression construct because the stable lines exhibited various elements of the *sem-4* phenotype, including sterility and deformed tails. In addition, in one instance, T.pppp underwent a necrotic death in a *sem-4::gfp* transformant, as it did in the *sem-4(n1971; mec-3::gfp)* strain. We also used MH1346, containing an integrated *sem-4::gfp* reporter, for *sem-4* expression studies.

egl-5 expression studies were carried out using the integrated *egl-5::gfp* array in strain EM597 (*bxIs12*), and the extrachromosomal arrays in strains EM783, EM784 and EM785. *bxIs12* was generated by integration of a transgene (EM#285) that was constructed to contain 16,027 nucleotides upstream of the *egl-5* AUG (beginning at an *Nru*I site at position 23,448 in cosmid C08C3), the full set of *egl-5* exons and introns, and 2639 nucleotides downstream of the *egl-5* stop codon (ending at position 43,981, the right end of C08C3). GFP was inserted at an *Apa*I site in the third *egl-5* exon at position 40,261 and disrupts the homeodomain, so that the expressed protein is expected to be non-functional. The arrays in EM783, EM784 and EM785 were constructed by insertion of *gfp* into the *Apa*I site in exon 3 of *egl-5*. EM783 (containing cosmid C08C3 bp 36249-43981), EM784 (C08C3 35986-36293) and EM785 (C08C3 36088-36293) were generated by PCR fusion (Hobert et al., 1999).

For ectopic *sem-4* expression studies, $P_{mec\text{-}7}$*sem-4* constructs were created by PCR amplification of portions of cosmid F15C11. For ease of cloning, we omitted the first, small exon from these clones, which began instead at the second start site (in exon 2) identified by Basson and Horvitz (Basson and Horvitz, 1996). The resulting DNA fragments were cloned into the *Xma*I and *Kpn*I sites of Fire vector pPD52.102 (Fire et al., 1990). The $P_{mec\text{-}7}$*sem-4* (*n1378*) construct contained an additional point mutation that produced a D506V change and the $P_{mec\text{-}7}$*sem-4* (*n2087*) construct contained an additional point mutation that produced a T301A change. For studies of the effect of ectopic *sem-4* on *mec-3* expression, $P_{mec\text{-}7}$*sem-4* constructs were injected into TU2562. For each construct, we scored three independent stable lines and, for each stable line, we scored at least 50 worms. Plasmids were injected with the dominant *rol-6*(*su1006*) marker plasmid pRF4 using standard methods (Mello et al., 1992). Injected DNA forms stable extrachromosomal arrays containing multiple copies of the injected construct.

Gel-mobility-shift assays

A PCR-amplified full-length *sem-4* cDNA was cloned into the *Bam*HI and *Not*I sites of pGEX 6P-1 (Amersham Pharmacia Biotech). The construct produced an N-terminal glutathione-*S*-transferase (GST) fusion. This GST::SEM-4 fusion was overproduced in BL21(DE3)pLysS cells and the fusion protein solubilized as described in Frangioni and Neel (Frangioni and Neel, 1993). SEM-4 was cleaved from the glutathione-bound GST moiety according to the manufacturer's instructions in the following cleavage buffer: 50 mM Tris-HCl pH 7.0, 150 mM NaCl, 1 mM EDTA, 0.1% bovine serum albumin (BSA), 10 mM DTT, 2% Triton X-100, 0.5 mM $ZnSO_4$.

The 96 bp *mec-3* promoter fragment used as a probe in the Fig. 4A gel-mobility-shift assays extended from positions 1714 to 1809 of the *mec-3* promoter (Way and Chalfie, 1988). The m3-1 sequence extends from 1759 to 1764, m3-2 from 1714 to 1718 and m3-3 from 1648 to 1654. The 100 bp *egl-5* promoter fragment used as a probe in the Fig. 4B gel-mobility-shift assays extended from position 38515 to 38615 of cosmid C08C3. The e5-1 sequence extends from position 38525 to 38531, e5-2 from 38551 to 38557 and e5-3 from 38600 to 38606. The 105 bp *egl-5* promoter fragment used as a probe in the Fig. 4D gel-mobility-shift assays extended from position 35986 to 36091 of cosmid C08C3. The e5-4 sequence extends from position 35997 to 36003, e5-5 from 36054 to 36060 and e5-T1 from 36070 to 36087.

DNA probes were end labeled using T4 polynucleotide kinase (US Biochemicals) according to the manufacturer's instructions. Labeled probes were separated from free nucleotide on an Amicon Millipore YM-3 column. The following reagents were added to the purified protein, probe and any cold competitor, at these final concentrations: 8.5 mM NaHEPES (pH 7.9), 30 mM KCl, 10.4 mM DTT, 0.3 mM PMSF, 2% Triton X-100, 1 mg ml^{-1} BSA, 4% Ficoll, 8 μg ml^{-1} poly-dIdC (Amersham Pharmacia Biotech), 0.5 mM $ZnSO_4$. The purified SEM-4 protein was preincubated in binding buffer with any cold competitor for 10 minutes at room temperature; the probe was then added and the reaction incubated for an additional 15 minutes at room

temperature. The reactions were run on 4% acrylamide gels (cross-link acrylamide:bisacrylamide 37.5:1) at 4°C for about 2 hours at 200 V in 0.5× TBE. Before the reactions were loaded, the gels were prerun at 4°C for 2 hours at 200 V in 0.5× TBE.

RESULTS

Loss of *sem-4* function causes proliferation of two distinct touch-cell fates

We have continued the characterization of the *sem-4* phenotype begun by Mitani et al. (Mitani et al., 1993) and Basson and Horvitz (Basson and Horvitz, 1996). They concluded that *sem-4* animals produced two additional touch-neuron-like cells (Basson and Horvitz, 1996; Mitani et al., 1993), instead of PHC neurons, possibly as the result of transformation of a neuronal sublineage (the T.pp lineage) (Fig. 1) into a PLM-like lineage (Basson and Horvitz, 1996). The identification of these abnormal cells as touch-neuron-like was based on their expression of several different touch-cell genes, including the touch-cell effector gene *mec-3*, and the resemblance of their processes to touch-cell processes by electron microscopy (Basson and Horvitz, 1996; Mitani, 1995; Mitani et al., 1993). Basson and Horvitz (Basson and Horvitz, 1996) also observed that, in *sem-4* mutants, the fate of the anterior daughter (T.ppa) of the T.pp neuroblast was variable: it sometimes died, sometimes differentiated and sometimes divided.

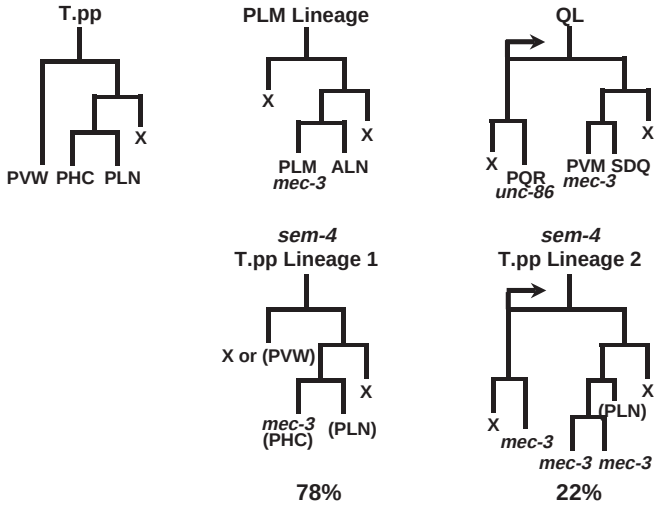

Fig. 1. Cell fate transformations in the T.pp lineage in *sem-4* animals. Lineage diagrams of the wild-type T.pp, AB.p(l/r)apappp and QL lineages (top) and the variable T.pp lineage in *sem-4* animals (bottom). Cell deaths are indicated with an 'X'. Cells in parentheses exhibit abnormal morphologies or migration patterns. Cells that expressed the *mec-3::gfp* reporter are indicated with a '*mec-3*'. The proportion of *sem-4* animals in which T.pp adopted each fate (only one side was examined per animal) is shown beneath the lineage diagrams. Of the 18 animals in which T.ppa did not divide, T.ppa died in nine and lived in nine. In four out of 23 animals, T.ppa divided and gave rise to an anterior daughter that died and a posterior daughter that migrated ventrally and then began to express *mec-3::gfp*. In a fifth animal, T.ppa divided and the posterior daughter migrated to the ventral side but the worm died before the lineage was completed. While examining these lineages, we found that loss of *sem-4* function not only induced apoptosis in a cell that should normally differentiate (T.ppa), but also sometimes induced deaths that appeared to be necrotic, both in T.ppa and in a cell that normally undergoes apoptosis (T.pppp). These necrotic deaths were characterized by the formation of large vacuoles, similar to those induced by gain-of-function mutations in degenerin genes (Chalfie and Wolinsky, 1990; Hall et al., 1997). T.pppp died as a vacuolated cell in one of the five lineages in which T.ppa divided and in two of the nine lineages in which T.ppa lived; in all nine lineages in which T.ppa died, T.pppp underwent apoptosis. T.ppa itself died as a vacuolated cell in three out of nine lineages. The *egl-5* gene was ectopically expressed both in cells undergoing apoptosis and in cells undergoing necrosis in *sem-4* animals.

We report three additional findings that clarify and extend the initial characterization of T-lineage defects in *sem-4* animals. First, we found that the number of ectopic touch-neuron-like cells, as defined by their expression of a *mec-3::gfp* fusion, ranged from one to five, with more produced at higher temperatures (Table 1): at 25°C, 44% of *sem-4* animals produced more than two additional touch-neuron-like cells. Second, the ectopic touch-neuron-like cells displayed two distinct morphologies (Fig. 2): some resembled PLM touch neurons, with cell bodies flattened against the muscle line and processes extending anteriorly and posteriorly; others resembled PVM touch neurons, with rounded cell bodies and processes extending ventrally and then anteriorly. Third, about 80% of the lineages were transformed into a PLM-like lineage (Lineage 1 in Fig. 1); the remaining lineages appear to be variants of the PVM lineage (Lineage 2 in Fig. 1).

We examined 18 type 1 lineages and all produced a single, *mec-3*-expressing cell (T.pppaa) that always resembled a PLM cell. In all of these lineages T.pppaa migrated to the ventral muscle and flattened against it (as do the PLM cells). By contrast, we observed three principal characteristics of Lineage 2 that suggested that it had been substantially but not completely transformed to a PVM-producing lineage (the QL lineage). First, the two type 2 lineages that we followed through the division of T.pppaa each generated three *mec-3*-expressing cells (T.ppap, T.pppaaa and T.pppaap). Five of these cells were PVM-like; only one (a T.pppaap cell) was PLM-like. The PVM-like cells neither migrated toward the ventral muscle nor took on a flattened appearance; these cells had rounded cell bodies and ventral processes similar to those of wild-type PVM cells. Second, the migration of T.ppa in the type 2 lineages resembled the migration of the equivalent cell (QL.a) in the PVM-producing lineage: these cells migrate posteriorly before dividing to produce an anterior daughter that dies and a posterior daughter (QL.ap) that survives. Third, in the QL lineage, QL.ap expresses *unc-86*; in the type 2 lineages, we conclude that the equivalent cell (T.ppap) also expresses *unc-86* because we observed expression of *mec-3*, which is dependent on *unc-86* expression (Way and Chalfie, 1989; Duggan et al., 1998). The transformation to a PVM-producing lineage, however, was incomplete: T.pppaa underwent an extra round of division relative to the equivalent cell in the QL lineage (QL.paa) and T.ppap expressed *mec-3* in addition to *unc-86* (expressed by QL.ap).

Because Lineage 2 produced more ectopic *mec-3::gfp*-expressing cells than Lineage 1, we hypothesized that animals with more ectopic *mec-3::gfp*-expressing cells should produce PVM-like cells. Conversely, animals with fewer ectopic *mec-3::gfp*-expressing cells should produce mostly PLM-like cells.

Table 1. Mutation of *egl-5* alters the fates of wild-type PLM cells and ectopic touch cells in *sem-4* animals

	Temperature									
	15°C			20°C			25°C			25°C
mec-3::gfp strains	<2	2	>2	<2	2	>2	<2	2	>2	PLM-like
Wild type	10	89	1	0	100	0	0	100	0	90% (90/100)
egl-5	78	22	0	49	51	0	3	97	0	77% (77/100)
	<4	4	>4	<4	4	>4	<4	4	>4	
sem-4	12	80	8	8	64	28	6	50	44	84% (283/338)
sem-4; *egl-5*	92	6	2	42	46	12	4	92	4	30% (64/213)

For wild-type and *egl-5* worms, at each temperature the percentage of worms with fewer than two, two, or greater than two *mec-3::gfp*-expressing tail cells is shown. For *sem-4* and *sem-4*; *egl-5* worms, at each temperature, the percentage of worms with fewer than four, four, or greater than four *mec-3::gfp*-expressing tail cells is shown. The percentage of PLM-like cells is shown for wild-type and *egl-5* worms at 25°C. The percentage of PLM-like cells is shown for *sem-4* and *sem-4*; *egl-5* worms at 25°C. At least 50 gravid adults were scored for each strain at each temperature.

At 25°C, in 51 *sem-4* adults with four *mec-3::gfp*-expressing tail cells, 93% (189/204) of the cells resembled PLM, 3% (6/204) resembled PVM and 4% (9/204) could not be unambiguously classified as either PLM or PVM. The cells in this last category generally had rounded cell bodies, like PVM, but had processes that were difficult to see or did not extend ventrally and then anteriorly. In 52 animals with more than four *mec-3::gfp*-expressing tail cells, 67% (181/269) resembled PLM, 10% (28/269) resembled PVM and 22% (60/269) could not be unambiguously classified.

sem-4 restricts proliferation of touch-cell fate through repression of *egl-5*

We found that loss of *sem-4* function caused transformation of the T.pp lineage into two distinct touch-cell lineages: a PLM-like lineage and a PVM-like lineage. The transformation of the posterior T.pp lineage into a mid-body PVM-like lineage is a transformation along the AP axis of the worm. AP transformations often result from defects in the expression or function of Hox genes that control AP patterning and several different aspects of cellular identity, including differentiation, growth and proliferation (Cillo et al., 2001; Veraksa et al., 2000).

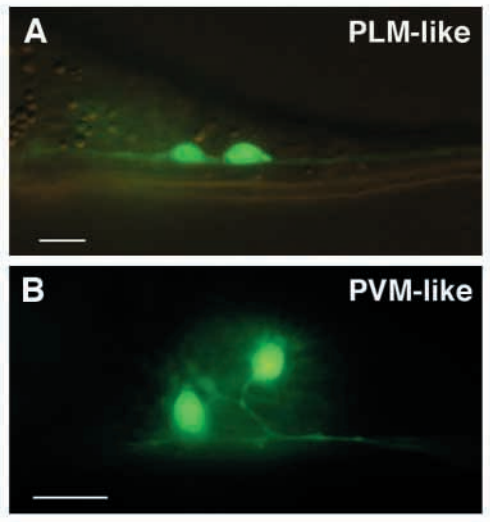

Fig. 2. PLM-like (A) and PVM-like (B) morphologies of *mec-3::gfp*-expressing tail cells in *sem-4* animals. Scale bars: 10 μm.

The *C. elegans* Hox gene that controls patterning of posterior structures is the *Abd-B* homolog *egl-5* (Chisholm, 1991; Salser and Kenyon, 1994). In the tail, *egl-5* is normally expressed in two pairs of neurons: the PLM touch neurons and the PVC interneurons (Ferreira et al., 1999). Although *egl-5* animals are touch insensitive in the tail (Chalfie and Au, 1989; Chisholm, 1991), this defect has been attributed to abnormal development of the PVC interneurons (Chisholm, 1991), which are part of the neural circuit for touch sensitivity (Chalfie et al., 1985). We find that *egl-5* is required for normal PLM development: PLM cells in *egl-5* animals showed substantially reduced expression of *mec-3* and abnormal morphologies (Table 1). We suggest that *egl-5* activates *mec-3* expression in PLM cells.

The findings that *egl-5* is required for correct determination of PLM fate in wild-type animals and that the T.pp lineage is sometimes transformed to a PLM-like lineage in *sem-4* animals suggest that loss of *sem-4* function produces ectopic expression of *egl-5* in the T.pp lineage. We observed ectopic expression in *sem-4* animals of an *egl-5::gfp* reporter first in the T.pp neuroblast and then in T.ppa, T.ppp, T.pppa and T.pppaa (which ectopically expresses *mec-3* in *sem-4* animals). Because we did not observe ectopic *egl-5* expression in T.pppap (which does not express *mec-3* in *sem-4* animals), ectopic *egl-5* expression in T.pppaa was probably not residual *gfp* expression from T.pppa. In wild-type worms, we observed expression of a *sem-4::gfp* reporter in T, T.a, T.p and all descendants of T.p. Thus, ectopic *egl-5* expression in the T lineage in *sem-4* animals began two cell divisions (in T.pp) after *sem-4* expression would normally begin (in T). This ectopic T lineage expression was observed with two different *egl-5::gfp* reporters in the *sem-4* background: one reporter contained a 12 kb region immediately upstream of the *egl-5* translation start site; the second reporter contained only a 3 kb region immediately upstream of the start site (Fig. 3). We propose that *sem-4* acts both in neuroblasts and neurons to restrict *egl-5* expression.

To investigate whether ectopically expressed *egl-5* promoted the PLM fate among the abnormal touch-neuron-like cells in *sem-4* mutants, we constructed *sem-4; egl-5* animals containing the integrated *mec-3::gfp* reporter and examined the number and morphology of the ectopic touch neurons produced. Using a relatively strong, partial loss-of-function *egl-5* allele, we found that loss of *egl-5* function decreased the number and significantly altered the morphology of the ectopic touch neurons (Table 1). In *sem-4* mutants, most *mec-3*-

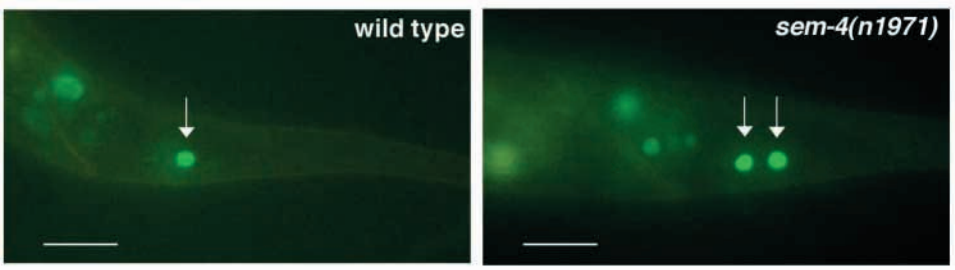

Fig. 3. T lineage expression in wild-type and *sem-4* animals of an *egl-5::gfp* fusion reporter gene containing ~3 kb upstream of the *egl-5* ATG fused to the *gfp* gene. Arrows indicate PLM (left) and PLM and T.pppaa (right). The same ectopic expression was observed in *sem-4* animals with an *egl-5::gfp* reporter containing 12 kb upstream of the *egl-5* ATG fused to the *gfp* gene. Scale bar: 10 μm.

expressing tail cells (84%) clearly resembled PLM cells. In *sem-4*; *egl-5* animals, the proportion of PLM-like cells decreased to 30% (Table 1).

We investigated whether mutation of the other *C. elegans* Hox genes, *mab-5* and *lin-39*, could affect the number or morphology of the ectopic touch neurons in *sem-4* animals. A partial loss-of-function allele of *lin-39* and a putative null allele of *mab-5* moderately decreased both the total number of ectopic touch neurons (data not shown) and the number of PLM-like cells: 70% (141/200) of *mec-3*-expressing tail cells in *sem-4; mab-5* animals and 74% (147/200) in *sem-4; lin-39* animals were PLM-like.

We conclude that *sem-4* normally prevents the expression of *egl-5* and possibly of *mab-5* and *lin-39* in the T lineage. Because *egl-5* appears to activate *mec-3* expression in normal PLM cells, we suggest that *sem-4* restricts *mec-3* expression in the T lineage indirectly through restriction of *egl-5*. We found that *egl-5* is required for correct expression of the PLM fate in wild-type PLM cells and for ectopic expression of the PLM fate in the abnormal touch cells in *sem-4* animals. Negative regulation of *egl-5* by *sem-4* is therefore necessary to restrict inappropriate proliferation of the PLM fate in wild-type animals.

SEM-4 binds to a shared motif in the *mec-3* and *egl-5* promoters

Mutations in the *mec-3* promoter at a 6 bp site (m3-1 in Fig. 4A), about 300 bp upstream of the translation start produced ectopic expression of a *mec-3::lacZ* reporter in additional cells in the tail (Xue, 1993). The m3-1 region is one of several in the *mec-3* promoter that is conserved between *C. elegans* and the related nematode *C. briggsae*. The ectopic *mec-3::lacZ* expression suggested that m3-1 might be a SEM-4 binding site. We found that, although purified SEM-4 protein did not shift a 24 bp fragment with m3-1 at its center (data not shown), it did shift a 96 bp fragment with m3-1 at its center (Fig. 4A). Four complexes were produced that were competed away by unlabeled specific competitor.

Mutation of m3-1 in the unlabeled competitor decreased, but did not eliminate, its ability to compete away complexes 2 and 4 (Fig. 4A), and had no effect on its ability to compete away complexes 1 and 3. An oligonucleotide composed of four tandem copies of the 24 bp fragment with m3-1 at the center was not as effective a competitor as the 96 bp wild-type competitor (data not shown). These results suggest that the 96 bp fragment used as a probe contained more than one SEM-4 binding site. The 5′ end of this fragment contains a sequence identical at five out of six positions to m3-1 (m3-2 in Fig. 4A) that we tested for SEM-4 binding. Mutations at both m3-1 and m3-2 substantially reduced the ability of the unlabeled oligonucleotide to compete away complexes 1, 2 and 4 (Fig. 4A). Complex 3 is probably formed by SEM-4 binding to a site other than m3-1 or m3-2.

We also looked for putative SEM-4 binding sites in the *egl-5* promoter. We found a cluster of three sites (e5-1, e5-2, e5-3), one identical to m3-1 (although on the opposite strand) and two identical to m3-2 (with one on the opposite strand), within an 81 bp region located about 900 bp upstream of the *egl-5* translation start site (Fig. 4B). Similar clusters, of three sites within 125 bp, occur infrequently in the region we analysed. Specifically, in the 31 kb separating the translation starts of *egl-5* and *mab-5* (these genes are transcribed in opposite directions, with their 5′ ends facing one another, separated by one predicted gene), only one other similar cluster occurs. This second cluster, of three sites within a 121 bp region, is located about 5 kb upstream of the *mab-5* translation start site. We re-examined the region of the *mec-3* promoter that contained the two sites we identified and found a third site (m3-3) about 60 bp upstream of the 5′ site (Fig. 4A). Thus, this region contains a cluster of three sites within 117 bp. A 206 bp region in the *lin-39* promoter, about 10 kb upstream of the translation start site, also contains a cluster of three SEM-4 sites.

SEM-4 shifted a 100 bp fragment containing e5-1, e5-2 and e5-3 (Fig. 4B). Two complexes were produced that were competed away by unlabeled competitor oligonucleotides. Mutation of all three sites in the wild-type competitor significantly decreased, but did not eliminate, its ability to compete away both complexes.

We identified a second region of the *egl-5* promoter to which SEM-4 binds. About 3.5 kb upstream of the *egl-5* start site, there is a 300 bp sequence (V6CRE) that mediates expression of *egl-5* in the V6 hypodermal lineage. We found that a reporter containing V6CRE fused to *gfp* was not expressed in the T lineage. Occasionally, some faint expression was detected in a couple of T lineage cells. In a *sem-4* background, however, expression of P_{V6CRE}*gfp* increased significantly (Fig. 4C). T lineage expression in wild-type worms of a reporter lacking the first 100 bp of P_{V6CRE}*gfp* ($P_{\text{V6CRE}\Delta 100}$*gfp*) was similarly strong (Fig. 4C).

The region deleted in $P_{\text{V6CRE}\Delta 100}$*gfp* contains a consensus TRA-1 binding site (e5-T1) and two sites that contain five out of six bases of SEM-4 binding sites (e5-4 and e5-5). We tested binding of SEM-4 to a probe composed of the first 105 bp of V6CRE. SEM-4 shifted this probe (Fig. 4D), forming four complexes that were competed away by unlabeled competitor oligonucleotides. Mutation of e5-4 and e5-5 in the wild-type competitor produced a small but consistent reduction in competition (Fig. 4D). Mutation of e5-T1 had no effect on the ability of the competitor to compete away the complexes (Fig. 4D).

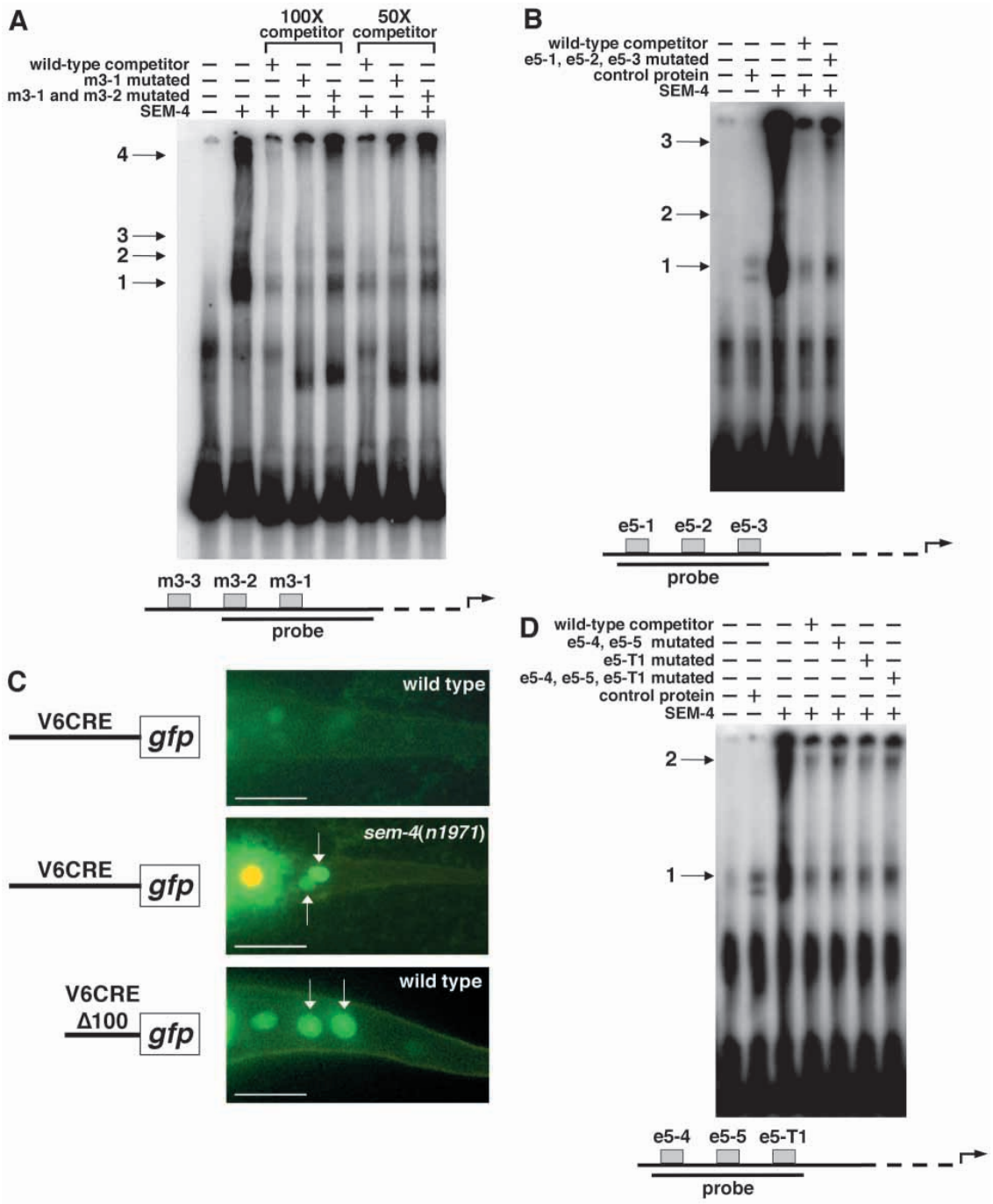

Fig. 4. Binding of SEM-4 to the *mec-3* and *egl-5* promoters. (A) Binding to the *mec-3* promoter. The indicated changes were made to inactivate the sites in the specific competitor oligonucleotides for the gel shift. The addition of SEM-4 to radiolabeled probe produced four complexes (arrows). Binding decreased substantially in the presence of 100-fold and 50-fold molar excess of cold specific competitor. Binding did not decrease as much when mutated competitor was added. Mutation of m3-1, from AGACAA to AGCTAG, restored some of the binding; mutation of both m3-1 (to AGCTAG) and m3-2 (from ACACAA to ACCTAG), restored more of the binding. The sequence of m3-3 is ACACAA. (B) Binding to a region of the *egl-5* promoter close to the translation start site. Control protein was prepared from cells transformed with the pGEX6P-1 vector, lacking the *sem-4* cDNA insert. The addition of SEM-4 to radiolabeled probe produced three complexes (arrows). Binding decreased substantially in the presence of 100-fold molar excess of cold specific competitor. Mutation of e5-1 (from TTGTGT to CTAGGT), e5-2 (from TTGTCT to CTAGCT) and e5-3 (from ACACAA to ACCTAG), in the specific competitor restored binding of complexes 1 and 3. (C) Ectopic T lineage expression of P_{V6CRE}*gfp* in *sem-4* animals and of $P_{V6CRE\Delta100}$*gfp* in wild-type animals. P_{V6CRE}*gfp* in wild-type animals shows only occasional, faint T lineage expression (top). Scale bar: 10 μm. Arrows indicate T.pa and T.pp cells expressing GFP. (D) Binding to V6CRE. The addition of SEM-4 to radiolabeled probe produced two complexes (arrows). Binding decreased substantially in the presence of 100-fold molar excess of cold specific competitor. Mutation of e5-4 and e5-5 as indicated in the specific competitor restored binding of complexes 1 and 3. Mutation of e5-T1 as indicated did not restore binding. Mutation of e5-4, e5-5 and e5-T1 produced the same restoration of binding as mutation of e5-4 and e5-5.

We conclude that SEM-4 binds to a shared motif in the *mec-3* and *egl-5* promoters and suggest that these interactions repress *mec-3* and *egl-5* expression in the T lineage. We propose that *sem-4* restricts *mec-3* expression both directly, by binding to the *mec-3* promoter, and indirectly, through repression of *egl-5*.

Ectopic *sem-4* represses *mec-3* expression *in vivo*

The *mec-3* gene is normally expressed in only ten cells: the six touch neurons, a pair of neurons in the head (the FLP cells) and a pair of mid-body neurons (the PVD cells). To test whether ectopic *sem-4* could repress *mec-3 in vivo*, we expressed *sem-4* in the touch cells under the control of the *mec-7* promoter ($P_{mec\text{-}7}$*sem-4*). *mec-7* encodes a β-tubulin that is expressed strongly in all six touch neurons during their terminal differentiation and less strongly in several other cells (Hamelin et al., 1992; Mitani et al., 1993; Savage et al., 1989). We analysed the effect of this ectopic *sem-4* activity on *mec-3* expression by transforming $P_{mec\text{-}7}$*sem-4* into worms containing an integrated *mec-3::gfp* reporter.

Transformation with $P_{mec\text{-}7}$*sem-4* decreased the proportion of *mec-3::gfp*-fluorescent PLM cells from 100% (102/102) in the control line to 48±20% in three transformed lines (350 cells) (Fig. 5A). (In *C. elegans*, transformed DNAs form extrachromosomal arrays that are often not present in all cells.) We also tested the ability of several truncated versions of SEM-4 to decrease expression of the *mec-3::gfp* reporter (Fig. 5B-F). The N-terminal half of SEM-4, truncated after zinc finger 3, was a relatively effective repressor: transformation with this construct decreased the proportion of fluorescent PLM cells to 74±18% in three transformed lines (312 cells) (Fig. 5B). When SEM-4 was truncated after zinc finger 2, however, the resulting fragment did not decrease *mec-3::gfp* expression (Fig. 5C).

We found that the reduction in *mec-3* expression produced by ectopically expressed *sem-4* was sufficient to decrease touch-cell function significantly. Wild-type animals transformed with $P_{mec\text{-}7}$*sem-4* were considerably less touch sensitive than animals transformed with $P_{mec\text{-}7}$*sem-4(Q321ocher)*, which encodes a null allele of *sem-4* (Fig. 5G). The null allele is an important control because transformation with genes driven by touch-cell promoters sometimes produces partial touch insensitivity, as it did in our experiments (Fig. 5G).

We tested the effect on touch-cell function of ectopic expression of two partial loss-of-function *sem-4* alleles (Fig. 5G). The *n2654* 'neuronal' allele (containing a mis-sense mutation that changes one of the zinc-chelating histidines in zinc finger 2 to a tyrosine) exhibits more defects in neurons than in mesoderm (Basson and Horvitz, 1996); the *n1378*

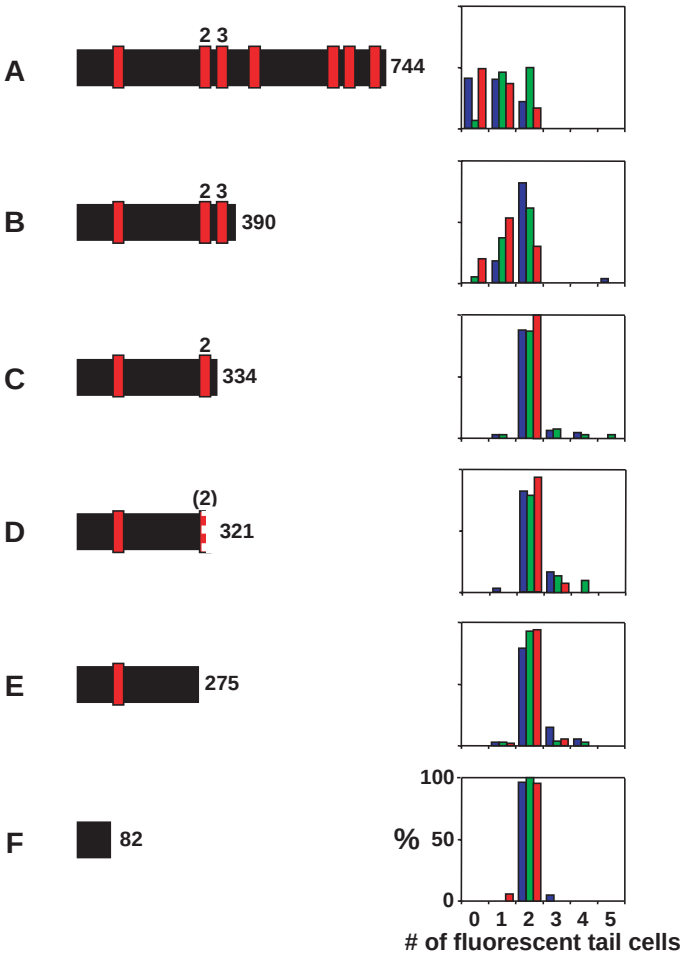

Fig. 5. Reduction of *mec-3* expression and touch sensitivity caused by ectopic expression of *sem-4*. (A-F) GFP fluorescence in tail cells of transformants containing different $P_{mec\text{-}7}$*sem-4* constructs. These constructs were injected into the *uIs22* strain, which contains an integrated *mec-3::gfp* reporter. When *uIs22* was transformed only with the *rol-6* marker, every animal had two fluorescent tail cells. Red boxes represent C2H2 zinc fingers; the number of the terminal residue in each fragment is given. Three stable lines are represented on each histogram and indicated by differently colored bars. (G) Effect of ectopic expression of different $P_{mec\text{-}7}$*sem-4* constructs on touch sensitivity. The proportion of responding worms is a cumulative measure: worms that responded to four touches were considered to have responded to three, two, one and zero touches. The squares show the touch response of wild type (N2). The circles show the average touch response of transformants from two independent stable lines containing $P_{mec\text{-}7}$*sem-4(Q321ocher)*. Each of the remaining curves represents the mean±s.d. The upright triangles show the touch response of transformants containing $P_{mec\text{-}7}$*sem-4(H323Y)*. The inverted triangles show the touch response of transformants containing $P_{mec\text{-}7}$*sem-4*. The diamonds show the touch response of transformants containing $P_{mec\text{-}7}$*sem-4(Q569ocher)*. The actual decrease in touch sensitivity produced by *sem-4* is the difference between the circles and the inverted triangles.

'mesodermal' allele (containing a nonsense mutation, Q569ocher, which truncates zinc fingers 5, 6 and 7) exhibits more defects in mesoderm than in neurons (Basson and Horvitz, 1996). We found that transformation with $P_{mec\text{-}7}$*sem-4(n1378)*, the Q569ocher 'mesodermal' allele, produced touch insensitivity, although not as effectively as wild-type *sem-4* (Fig. 5G). By contrast, transformation with $P_{mec\text{-}7}$*sem-4(n2654)*, the 'neuronal' allele, did not produce touch insensitivity (Fig. 5G). We also found that the neuronal allele was more touch insensitive than the mesodermal allele (Fig. 6). These results are consistent with the hypothesis that *sem-4(n2654)* does not function properly in neurons.

We conclude from these data that ectopically expressed *sem-4* represses *mec-3* in vivo. Two observations indicate that the zinc-finger pair 2 and 3 is crucial for this repression. First, a mutant version of SEM-4 truncated after zinc finger 3 was a relatively effective repressor, whereas a version truncated between zinc fingers 2 and 3 did not repress. Second, the mesodermal allele (truncated after zinc finger 4) produced touch insensitivity, whereas the neuronal allele, containing a defective zinc finger 2, did not.

Mutant versions of *sem-4* display a gain-of-function phenotype

Production of three mutant versions of SEM-4 under the control of the *mec-7* promoter produced *mec-3::gfp* expression in additional cells in the tail (Fig. 5C-E). Because *mec-7* is expressed in cells other than the touch cells (Hamelin et al., 1992; Mitani et al., 1993), the $P_{mec\text{-}7}$ constructs probably expressed the truncated SEM-4 proteins in these additional cells. These gain-of-function mutant proteins contained C2H2 zinc finger 1 and various portions of zinc finger 2 (Fig. 5C-E). These truncated SEM-4 fragments might have interfered either with endogenous SEM-4 or with other SEM-4 interacting partners in these additional cells.

We observed a gain-of-function phenotype in the *sem-4(n2087)* allele, which contains a nonsense mutation in zinc finger 2 (Basson and Horvitz, 1996). We found that *n2087* worms were more touch insensitive than *n1971* worms, which contain an early splice-site mutation N-terminal to zinc finger 1 (Fig. 6) (Basson and Horvitz, 1996). The SEM-4 protein fragment encoded by *n2087* is the same as that encoded by the $P_{mec\text{-}7}$*sem-4* construct (shown in Fig. 5D). Worms transformed with this construct had the most ectopic *mec-3::gfp*-expressing cells (Fig. 5D). The fact that homozygous *n2087* worms exhibit a gain-of-function phenotype suggests that truncated SEM-4 proteins interfere with the function of a protein other than SEM-4.

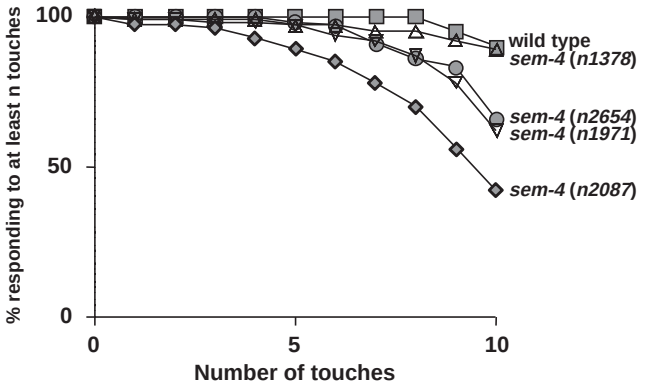

Fig. 6. Touch-response curves for different *sem-4* alleles. The touch assays and the representation of the results are as in Fig. 5G. The squares show the touch response of N2 (wild-type) worms. Circles show the touch response of *sem-4(n2654)* animals (H323Y), upright triangles show the touch response of *sem-4(n1378)* animals (Q569ocher), diamonds show the touch response of *sem-4(n2087)* animals (Q321ocher) and inverted triangles show the touch response of *sem-4(n1971)* animals (early splice donor mutation).

DISCUSSION

Function of sal genes is evolutionarily conserved

Several lines of evidence indicate that the function of sal genes has been evolutionarily conserved. The sal genes appear to function as cell fate determinants, regulators of Hox genes and AP patterning, and transcriptional repressors. The targets of and mechanism of regulation by sal genes appear to be conserved. The sal genes are involved in determination of precursor and differentiated cell fates. In *Drosophila*, *sal* is required in neuronal precursors and differentiated neurons to restrict neuronal fate to the proper cells. It restricts the fates of neuronal precursors during development of a sensory organ in the peripheral nervous system (Rusten et al., 2001). In the developing *Drosophila* eye, *sal* and *salr* (*spalt-related*) are required late in pupation for terminal differentiation of particular photoreceptor cells (Mollereau et al., 2001). Loss of *sal* and *salr* resulted in transformation of certain photoreceptor cells into other photoreceptor cells, as judged by rhabdomere morphology and opsin gene expression (Mollereau et al., 2001). The sal genes are probably involved in determining neural fates in mouse and human beings: human *SALL1* is expressed in specific areas of the fetal brain (thalamus) and adult brain (corpus callosum and substantia nigra) (Kohlhase et al., 1996); mouse *sall1* is expressed in particular embryonic neural tissues (tissues surrounding some of the ventricles and specific layers of the neural tube) (Buck et al., 2001).

Hox genes appear to be targets not only of *sem-4* but also of other sal genes. *Drosophila sal* might negatively regulate *Sex combs reduced* (*Scr*) and other *Drosophila* Hox genes. Loss of *sal* function in *Drosophila* BX-C^{-} embryos produced some limited ectopic expression of the Hox gene *Scr* (Casanova, 1989). Mutations in *sal* enhanced the phenotypes of *Polycomb* group (PcG) mutants. These genes are known to be negative regulators of Hox genes (Landecker et al., 1994). Loss of *sal* function affects AP patterning in *Drosophila*. Mutations in *sal* incompletely transform both head and tail structures into trunk-like structures; *sal* activity has been shown to promote head development (Jurgens, 1988). Hox genes in mammals might also be targets of sal family genes. Patients with TBS, which is caused by mutations in *SALL1*, display characteristic features of syndromes associated with mutations in HOX genes (Powell and Michaelis, 1999; Surka et al., 2001; Veraksa et al., 2000).

LIM homeobox genes, such as *mec-3*, might also be conserved targets of sal genes. The closest mammalian homolog to *mec-3* is the human LIM homeobox gene *Lhx5* (Zhao et al., 2000). *Lhx5* and the human *SALL1* gene appear to be expressed in different sets of cells in the developing

thalamus, which constitutes a very small portion of the entire brain (Kohlhase et al., 1996; Nakagawa and O'Leary, 2001). *SALL1* and *Lhx5* are not expressed in most other regions of the fetal brain. Their expression in separate thalamic cells could indicate that *SALL1* restricts *Lhx5* expression in the thalamus.

The mechanism through which *sem-4* negatively regulates its targets is probably conserved. SEM-4, SALL1 and mouse sall1 are transcriptional repressors. SALL1 and mouse sall1, fused to heterologous DNA binding domains, behaved as repressors in mammalian cell culture assays (Kiefer et al., 2002; Netzer et al., 2001). We suggest that these genes bind directly to regulatory regions of their targets.

Particular mutant versions of SALL1, like certain SEM-4 truncations, might act as gain-of-function proteins. TBS is an autosomal dominant disorder caused by mutations in *SALL1* (Kohlhase, 2000). No deletions of the entire gene or mutations that truncate the protein upstream of the first zinc finger have been detected. Thus, all 21 *SALL1* mutant alleles encode truncated protein products that contain at least the first zinc finger. We found that truncations of SEM-4 containing the first zinc finger acted as gain-of-function proteins. We suggest that TBS could result, at least in part, from interference by these truncated proteins with wild-type SALL1 or other proteins.

SEM-4 negatively regulates genes at multiple levels of a developmental hierarchy

Very little is known about the pathways that lead from Hox proteins to determination of the fates of particular structures or individual cells. Recent evidence has suggested that Hox proteins can act independently on genes that function at different points along a particular developmental pathway (Veraksa et al., 2000). For example, the *Drosophila* Hox gene *Ultrabithorax* (*Ubx*) negatively regulates diverse genes throughout haltere development (Weatherbee et al., 1998). These genes encode signaling molecules, their immediate targets (including *sal* and *salr*) and proteins further downstream, including transcription factors. We propose that negative regulators of Hox genes, like the Hox genes themselves, also function at different levels in a given developmental hierarchy. We found that *sem-4* negatively regulates both the Hox gene *egl-5* in precursors and differentiating cells and the LIM homeobox gene *mec-3* in differentiating cells. Furthermore, we discovered that *egl-5* positively regulates *mec-3* in normal PLM cells. Because SEM-4 binds to a shared motif in the promoters of both *mec-3* and *egl-5*, we conclude that *sem-4* negatively regulates each gene independently and also inhibits *mec-3* expression through inhibition of *egl-5*.

No evidence for a global repression system for Hox genes in *C. elegans* has been reported. *Polycomb* group (PcG) and *trithorax* group (trxG) genes were originally identified in *Drosophila* as repressors and activators, respectively, of Hox gene expression (Brock and van Lohuizen, 2001). PcG genes are now known to function together in a chromatin repressive complex (Francis and Kingston, 2001). Although the roles of Hox and trxG genes in patterning are conserved in *C. elegans*, a role for PcG genes has not yet been established. We speculate that *sem-4* might function as a member of a general repressive complex akin to the PcG complex.

Drosophila and mammalian studies have suggested that sal genes might function as PcG genes. Casanova (Casanova, 1989) found that *sal* mutations caused limited ectopic expression of the Hox genes *Ubx* and *Scr*, and Landecker et al. (Landecker et al., 1994) found that *sal* mutations enhanced mutations in the PcG genes *polyhomeotic* and *Polycomb-like*. Human SALL1 localizes to chromocenters in mammalian cells (Netzer et al., 2001) and mouse sall1 interacts with components of chromatin remodeling complexes (Kiefer et al., 2002). One additional speculation is that *Drosophila sal* might bind to a 138 bp silencing sequence in the *Polycomb* response element in *Abd-B*, the *egl-5* ortholog (Busturia et al., 2001). We have identified two sites that match the SEM-4 binding sequence in this *Drosophila* silencing element.

PcG genes might play a role in positive, in addition to negative, regulation of Hox genes (Brock and van Lohuizen, 2001): mutations in some PcG genes enhance trxG mutant phenotypes. *sem-4* also appears to have a positive regulatory role in Hox gene expression in certain tissues: *sem-4* might activate *lin-39* in vulval lineages (Grant et al., 2000) and *egl-5* in hypodermal lineages (Y. Teng et al., unpublished).

We thank M. Basson for *sem-4* cDNA clones, J. McDermott for advice on gel shift assays, R. Horvitz for the *sem-4(n2654)* strain, W. Hanna-Rose for MH1346 and D. Xue for identifying m3-1 as a possible SEM-4 binding site. Some nematode strains used in this work were provided by the *Caenorhabditis* Genetics Center, which is funded by the NIH National Center for Research Resources (NCRR). This research was funded by NIH grants GM30997 to M.C. and GM39353 to S.W.E., and by postdoctoral fellowships from the American Cancer Society to A.S.T. and from the Brazilian National Council of Scientific and Technological Development (CNPq) to H.B.F. S.W.E. is the Siegfried Ullmann Professor of Molecular Genetics.

REFERENCES

Basson, M. and Horvitz, H. R. (1996). The *Caenorhabditis elegans* gene *sem-4* controls neuronal and mesodermal cell development and encodes a zinc finger protein. *Genes Dev.* **10**, 1953-1965.

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

Brock, H. W. and van Lohuizen, M. (2001). The *Polycomb* group – no longer an exclusive club? *Curr. Opin. Genet. Dev.* **11**, 175-181.

Buck, A., Kispert, A. and Kohlhase, J. (2001). Embryonic expression of the murine homologue of SALL1, the gene mutated in Townes-Brocks syndrome. *Mech. Dev.* **104**, 143-146.

Busturia, A., Lloyd, A., Bejarano, F., Zavortink, M., Xin, H. and Sakonju, S. (2001). The MCP silencer of the *Drosophila Abd-B* gene requires both Pleiohomeotic and GAGA factor for the maintenance of repression. *Development* **128**, 2163-2173.

Casanova, J. (1989). Mutations in the *spalt* gene of *Drosophila* cause ectopic expression of *Ultrabithorax* and *Sex combs reduced*. *Roux's Arch. Dev. Biol.* **198**, 137-140.

Chalfie, M. and Au, M. (1989). Genetic control of differentiation of the *Caenorhabditis elegans* touch receptor neurons. *Science* **243**, 1027-1033.

Chalfie, M. and Wolinsky, E. (1990). The identification and suppression of inherited neurodegeneration in *Caenorhabditis elegans*. *Nature* **345**, 410-416.

Chalfie, M., Sulston, J. E., White, J. G., Southgate, E., Thomson, J. N. and Brenner, S. (1985). The neural circuit for touch sensitivity in *Caenorhabditis elegans*. *J. Neurosci.* **5**, 956-964.

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

Cillo, C., Cantile, M., Faiella, A. and Boncinelli, E. (2001). Homeobox genes in normal and malignant cells. *J. Cell. Physiol.* **188**, 161-169.

Desai, C., Garriga, G., McIntire, S. L. and Horvitz, H. R. (1988). A genetic pathway for the development of the *Caenorhabditis elegans* HSN motor neurons. *Nature* **336**, 638-646.

Duggan, A., Ma, C. and Chalfie, M. (1998). Regulation of touch receptor differentiation by the *Caenorhabditis elegans mec-3* and *unc-86* genes. *Development* **125**, 4107-4119.

Ferreira, H. B., Zhang, Y., Zhao, C. and Emmons, S. W. (1999). Patterning of *Caenorhabditis elegans* posterior structures by the Abdominal-B homolog, egl-5. *Dev. Biol*. **207**, 215-228.

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

Francis, N. J. and Kingston, R. E. (2001). Mechanisms of transcriptional memory. *Nat. Rev. Mol. Cell Biol.* **2**, 409-421.

Frangioni, J. V. and Neel, B. G. (1993). Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. *Anal. Biochem.* **210**, 179-187.

Grant, K., Hanna-Rose, W. and Han, M. (2000). *sem-4* promotes vulval cell-fate determination in *Caenorhabditis elegans* through regulation of lin-39 Hox. *Dev. Biol.* **224**, 496-506.

Hall, D. H., Gu, G., Garcia-Anoveros, J., Gong, L., Chalfie, M. and Driscoll, M. (1997). Neuropathology of degenerative cell death in *Caenorhabditis elegans*. *J. Neurosci*. **17**, 1033-1045.

Hamelin, M., Scott, I. M., Way, J. C. and Culotti, J. G. (1992). The *Mec7* β-tubulin gene of *Caenorhabditis elegans* is expressed primarily in the touch receptor neurons. *EMBO J.* **11**, 2885-2893.

Hobert, O., Moerman, D. G., Clark, K. A., Beckerle, M. C. and Ruvkun, G. (1999). A conserved LIM protein that affects muscular adherens junction integrity and mechanosensory function in *Caenorhabditis elegans*. *J. Cell Biol.* **144**, 45-57.

Jurgens, G. (1988). Head and tail development of the *Drosophila* embryo involves *spalt*, a novel homeotic gene. *EMBO J.* **7**, 189-196.

Kenyon, C. (1986). A gene involved in the development of the posterior body region of *C. elegans*. *Cell* **46**, 477-487.

Kiefer, S. M., McDill, B. W., Yang, J. and Rauchman, M. (2002). Murine sall1 represses transcription by recruiting a histone deacetylase complex. *J. Biol. Chem.* **277**, 14869-14876.

Kohlhase, J. (2000). SALL1 mutations in Townes-Brocks syndrome and related disorders. *Hum. Mutat.* **16**, 460-466.

Kohlhase, J., Schuh, R., Dowe, G., Kuhnlein, R. P., Jackle, H., Schroeder, B., Schulz-Schaeffer, W., Kretzschmar, H. A., Kohler, A., Muller, U. et al. (1996). Isolation, characterization and organ-specific expression of two novel human zinc finger genes related to the *Drosophila* gene *spalt*. *Genomics* **38**, 291-298.

Landecker, H. L., Sinclair, D. A. R. and Brock, H. W. (1994). Screen for enhancers of *Polycomb* and *Polycomblike* in *Drosophila melanogaster*. *Dev. Genet.* **15**, 425-443.

Li, C. and Chalfie, M. (1990). Organogenesis in *C. elegans*: positioning of neurons and muscles in the egg-laying system. *Neuron* **4**, 681-695.

Li, D., Dower, K., Ma, Y., Tian, Y. and Benjamin, T. L. (2001). A tumor host range selection procedure identifies $p150^{sal2}$ as a target of polyoma virus large T antigen. *Proc. Natl. Acad. Sci. USA* **98**, 14619-14624.

Ma, Y., Dawei, L., Chai, L., Luciani, A., Ford, D., Morgan, J. and Maizel, A. L. (2001a). Cloning and characterization of two promoters for the human *HSAL2* gene and their transcriptional repression by the Wilms tumor suppressor gene product. *J. Biol. Chem*. **276**, 48223-48230.

Ma, Y., Singer, D. B., Gozman, A., Ford, D., Chai, L., Steinhoff, M. M., Hansen, K. and Maizel, A. L. (2001b). *Hsal1* is related to kidney and gonad development and is expressed in Wilms tumor. *Pediatr. Nephrol*. **16**, 701-709.

Mello, C. C., Kramer, J. M., Stinchcomb, D. and Ambros, V. (1992). Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959-3970.

Mitani, S. (1995). Genetic regulation of *mec-3* gene expression implicated in the specification of the mechanosensory neuron cell types in *Caenorhabditis elegans*. *Dev. Growth Differ.* **37**, 551-557.

Mitani, S., Du, H., Hall, D., Driscoll, M. and Chalfie, M. (1993). Combinatorial control of touch receptor neuron expression in *Caenorhabditis elegans*. *Development* **119**, 773-783.

Mollereau, B., Dominguez, M., Webel, R., Colley, N. J., Keung, B., de Celis, J. F. and Desplan, C. (2001). Two-step process for photoreceptor formation in *Drosophila*. *Nature* **412**, 911-913.

Nakagawa, Y. and O'Leary, D. D. (2001). Combinatorial expression patterns of LIM-homeodomain and other regulatory genes parcellate developing thalamus. *J. Neurosci.* **21**, 2711-2725.

Netzer, C., Rieger, L., Brero, A., Zhang, C. D., Hinzke, M., Kohlhase, J. and Bohlander, S. K. (2001). *SALL1*, the gene mutated in Townes-Brocks syndrome, encodes a transcriptional repressor which interacts with TRF1/PIN2 and localizes to pericentromeric heterochromatin. *Hum. Mol. Genet*. **10**, 3017-3024.

Powell, C. M. and Michaelis, R. C. (1999). Townes-Brocks syndrome. *J. Med. Genet*. **36**, 89-93.

Raaphorst, F. M., Otte, A. P. and Meijer, C. J. L. M. (2001). Polycomb-group genes as regulators of mammalian lymphopoiesis. *Trends Immunol.* **22**, 682-690.

Rusten, E., Cantera, R., Urban, J., Technau, G., Kafatos, F. C. and Barrio, R. (2001). *spalt* modifies EGFR-mediated induction of chordotonal precursors in the embryonic PNS of *Drosophila* promoting the development of oenocytes. *Development* **128**, 711-722.

Salser, S. J. and Kenyon, C. (1994). Patterning *C. elegans*: homeotic cluster genes, cell fates and cell migrations. *Trends Genet.* **10**, 159-164.

Savage, C., Hamelin, M., Culotti, J. G., Coulson, A., Albertson, D. G. and Chalfie, M. (1989). *mec-7* is a β-tubulin gene required for the production of 15-protofilament microtubules in *Caenorhabditis elegans*. *Genes Dev.* **3**, 870-881.

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

Surka, W. S., Kohlhase, J., Neunert, C. E., Schneider, D. S. and Proud, V. K. (2001). Unique family with Townes-Brocks syndrome, *SALL1* mutation and cardiac defects. *Am. J. Med. Genet.* **102**, 250-257.

Trent, C., Tsung, N. and Horvitz, H. R. (1983). Egg-laying defective mutants of the nematode *C. elegans*. *Genetics* **104**, 619-647.

Veraksa, A., Del Campo, M. and McGinnis, W. (2000). Developmental patterning genes and their conserved functions: from model organisms to humans. *Mol. Genet. Metab*. **69**, 85-100.

Way, J. C. and Chalfie, M. (1988). *mec-3*, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in *C. elegans*. *Cell* **54**, 5-16.

Way, J. C. and Chalfie, M. (1989). The *mec-3* gene of *Caenorhabditis elegans* requires its own product for maintained expression and is expressed in three neuronal cell types. *Genes Dev.* **3**, 1823-1833.

Weatherbee, S. D., Halder, G., Kim, J., Hudson, A. and Carroll, S. (1998). Ultrabithorax regulates genes at several levels of the wing-patterning hierarchy to shape the development of the *Drosophila* haltere. *Genes Dev.* **12**, 1474-1482.

Wu, J., Duggan, A. and Chalfie, M. (2001). Inhibition of touch cell fate by Egl44 and Egl46 in *C. elegans*. *Genes Dev.* **15**, 789-802.

Xue, D., Tu, Y. and Chalfie, M. (1993). Cooperative interactions between the *Caenorhabditis elegans* homeoproteins UNC-86 and MEC-3. *Science* **261**, 1324-1328.

Zhao, Y., Hermesz, E., Yarolin, M. C. and Westphal, H. (2000). Genomic structure, chromosomal localization and expression of the human LIM-homeobox gene *LHX5*. *Gene* **260**, 95-101.