

# UNC-39, the *C. elegans* homolog of the human myotonic dystrophy-associated homeodomain protein Six5, regulates cell motility and differentiation

Judith L. Yanowitz,<sup>a,\*</sup> M. Afaq Shakir,<sup>b</sup> Edward Hedgecock,<sup>c</sup> Harald Hutter,<sup>d</sup>  
Andrew Z. Fire,<sup>a</sup> and Erik A. Lundquist<sup>b,\*</sup>

<sup>a</sup>Department of Embryology, Carnegie Institution of Washington, Baltimore, MD 21210, USA

<sup>b</sup>Department of Molecular Biosciences, University of Kansas, Lawrence, KS 66045, USA

<sup>c</sup>Department of Biology, Johns Hopkins University, Baltimore, MD 21218, USA

<sup>d</sup>Max-Planck-Institut für medizinische Forschung, 69120 Heidelberg, Germany

Received for publication 27 January 2004, revised 20 April 2004, accepted 2 May 2004

Available online 2 July 2004

## Abstract

Mutations in the *unc-39* gene of *C. elegans* lead to migration and differentiation defects in a subset of mesodermal and ectodermal cells, including muscles and neurons. Defects include mesodermal specification and differentiation as well as neuronal migration and axon pathfinding defects. Molecular analysis revealed that *unc-39* corresponds to the previously named gene *ceh-35* and that the UNC-39 protein belongs to the Six4/5 family of homeodomain transcription factors and is similar to human Six5, a protein implicated in the pathogenesis of type I myotonic dystrophy (DM1). We show that human Six5 and UNC-39 are functional homologs, suggesting that further characterization of the *C. elegans unc-39* gene might provide insight into the etiology of DM1.

© 2004 Elsevier Inc. All rights reserved.

**Keywords:** *C. elegans*; *unc-39*; Myotonic dystrophy

## Introduction

Many molecular and genetic pathways controlling cellular physiology and development are remarkably conserved among metazoans. Members of the “Six” class of homeodomain transcription factors have emerged as important regulators of conserved developmental events. Six family members are defined by a conserved Six domain, which is thought to interact with other proteins, and a characteristic homeodomain involved in DNA binding (Kawakami et al., 2000). Loss of function of *Six* genes in fruit flies and vertebrates leads to defects in anterior ectoderm (often leading to eyeless organisms) and mesoderm (Cheyette et

al., 1994; Klesert et al., 2000; Seimiya and Gehring, 2000; Serikaku and O’Tousa, 1994), whereas overexpression of *Six* family members in vertebrates can lead to hyperproliferation of forebrain and eye structures (Kobayashi et al., 1998; Lagutin et al., 2001) and can induce expression of the myogenic genes *MyoD* and *myogenin* (Heanue et al., 1999; Laclef et al., 2003; Spitz et al., 1998). Thus, *Six* family members are evolutionarily conserved players in anterior ectoderm and mesoderm development and likely act by controlling the expression of other genes.

Recently, human genetic syndromes have been associated with perturbation of *Six* family members, each of which is characterized by anterior ectoderm defects and, in some cases, mesoderm defects. Holoprosencephaly, a syndrome of craniofacial and brain abnormalities consistent with altered development of the anterior neural plate and eye, is due to mutations in the *Six3* gene (Pasquier et al., 2000; Wallis et al., 1999). Bilateral anophthalmia, a syndrome of forebrain and pituitary defects, is due to haploinsufficiency of the *Six6* gene (Gallardo et al., 1999).

\* Corresponding authors. Judith L. Yanowitz is to be contacted at Carnegie Institution of Washington, Department of Embryology, 115 West University Parkway, Baltimore, MD 21210. Erik A. Lundquist, Department of Molecular Biosciences, University of Kansas, 5049 Haworth Hall, 1200 Sunnyside Avenue, Lawrence, KS 66045. Fax: +1-785-8645294.

E-mail addresses: [jly@alum.mit.edu](mailto:jly@alum.mit.edu) (J.L. Yanowitz), [erikl@ku.edu](mailto:erikl@ku.edu) (E.A. Lundquist).

Human myotonic dystrophy type I (DM1), the most common form of muscular dystrophy, with an incidence of 1 in 8000, is associated with a decrease in *Six5* expression (Klesert et al., 1997; Thornton et al., 1997). DM1 is a multisystemic disorder with widely variable clinical presentation including myotonia, cataracts, gonadal atrophy, endocrine abnormalities, cardiac conduction defects, and mental impairment (Harper, 2001). DM1 maps to a CTG expansion in the region overlapping the 3' - untranslated region of the *DMPK* (myotonic dystrophy protein kinase) gene (Brook et al., 1992; Fu et al., 1993; Harley et al., 1992) and the promoter of the human *Six5* gene (Boucher et al., 1995). This expansion serves as the docking site for the CTCF protein which is involved in establishing a silenced chromatin domain affecting both *DMPK* and *Six5* (Filippova et al., 2001). Recent evidence in mice and tissue culture supports roles for *Six5* in muscle, heart, and eye (Wakimoto et al., 2002). However, a mouse knockout of the *Six5* gene produces only mild defects as compared to DM1 patients (Klesert et al., 2000; Sarkar et al., 2000). The lack of severity is likely to be due both to feedback mechanisms which upregulate expression of other *Six* family members as well as to genetic background effects in the inbred strains of mice. Recently, a mutation in the gene encoding the *Drosophila Six5* ortholog *DSix-4* was shown to affect both muscle and gonad formation (Kirby et al., 2001).

In this work, we describe a genetic and molecular characterization of the *C. elegans unc-39* gene. *unc-39*, which corresponds to the previously named *ceh-35* gene, encodes a *C. elegans* member of the Six4/5 family, and *unc-39* mutant animals exhibit a broad range of defects in mesodermal and anterior neuronal specification and differentiation. We show that the Six domain and homeodomain from human *Six5* can complement mutations in *unc-39*, indicating that human *Six5* and *C. elegans UNC-39* are functional homologs.

## Materials and methods

### *C. elegans* mutations and techniques

*C. elegans* strains were cultured using standard techniques (Brenner, 1974). The wild-type strain was N2. All analyses were performed at 20°C unless otherwise noted. The following mutations were used in this work: LG V: *him-5*(e1467), *itDf2*, *nDf42*, *unc-39*(ct73, e257, and rh72), and *yDf8*.

The following integrated transgenes were used to score mesodermal cell types: coelomocytes, *arIs39*[*myo-3::secreted gfp*] *X* (*myo-3::secreted gfp* produces GFP which is secreted into the pseudocoelom and is endocytosed by the coelomocytes), and *ccls4438* [*hlh-8::gfp*] *IV* (a cell-intrinsic marker that consists of an enhancer from the *twist* gene promoter that drives cc-specific *gfp* expression) (Harfe

et al., 1998b); body wall muscles, *ccls4251*[*myo-3::gfp*] *I* (Kostas and Fire, 2002); daughter cells of M, *ayIs6*[*hlh-8::gfp*] *X* (Harfe et al., 1998a); head mesodermal cell and vulval muscles, *ccls4443*[*arg-1::gfp*] *IV* (Kostas and Fire, 2002); vm1 vulval muscles, *ayIs2*[*egl-15::gfp*] *IV* (Kostas and Fire, 2002); somatic gonad, *otEx105*[*lim-7::gfp*]. The following integrated transgenes were used to score neuronal cell types: amphid neurons, *lqIs2*[*osm-6::gfp*] *X* (Collet et al., 1998); CAN neurons, *lqIs27*[*ceh-23::gfp*] (Forrester et al., 1998); command interneurons, *zdlIs3*[*glr-1::gfp*] (Maricq et al., 1995).

*unc-39* was mapped relative to *him-5* on chromosome V by determining if *unc-39* and *him-5* were removed by a series of deficiencies in the region (Fig. 4A). Deficiencies *nDf42*, *yDf8*, and *itDf2* all failed to complement *unc-39*(e257), whereas only *nDf42* and *itDf2* failed to complement *him-5*.

Germline transformation of *C. elegans* was performed by standard techniques (Mello and Fire, 1995). A plasmid harboring the dominant *rol-6* mutation was used as a cotransformation marker. Details about transgenic lines generated in this study are available upon request.

RNAi was performed by dsRNA injection (1 µg/µl) into the body cavity as described (Fire et al., 1998). dsRNA was generated from a 318-bp genomic fragment from the beginning of exon 3 of the predicted *ceh-35* coding region.

### Molecular analysis of *unc-39*

The sequences of all oligonucleotides used in this work are available upon request. The *unc-39* rescuing construct was made by PCR amplification of the *ceh-35* locus from genomic DNA (see Fig. 4). Included was 4443 bp of upstream sequence (the *unc-39* promoter), the entire *ceh-35* coding region, and 337 bp of downstream sequence (the *unc-39* 3' region). The full-length *unc-39::gfp* C-terminal tag was constructed by PCR amplification of the promoter and coding region without the stop codon and placement of this fragment upstream of and in frame with *gfp* in a *gfp* expression vector. The *unc-39* 3' region was then inserted downstream of *gfp* in the vector. The *unc-39 promoter::gfp* transgene was constructed by PCR amplification of the *unc-39* promoter (excluding the initiator codon), which was inserted upstream of *gfp*. The *unc-39* 3' region was then placed downstream of *gfp*.

The *unc-39::human Six5 Six/homeodomain* swap construct was made as follows: The sequence of the human *Six5* Six and homeodomains were aligned with *unc-39* (see Fig. 4). The human *Six5* Six and homeodomain sequences (which were designed with optimal *C. elegans* codon usage) were generated by Klenow-mediated extension and subsequent ligation of 11 overlapping oligonucleotides. The contiguous region was then amplified with PCR and subcloned as a *XbaI/BglII* fragment into these unique sites in the *unc-39* genomic rescue construct. Since these sites are in introns, the human sequence was designed to recreate the cloning sites, intron sequences, and splice

junctions found in *unc-39*. A control construct without the Six and homeodomains was created by blunt cloning the *XbaI/BglIII* sites in the genomic rescue plasmid. The sequences of all coding regions amplified by PCR were determined to ensure that no point mutations were introduced by PCR.

The molecular lesions associated with the three *unc-39* alleles *e257*, *ct73*, and *rh72*, were obtained by direct sequencing of PCR fragments spanning the *ceh-35* coding region generated from each mutant.

## Results

### *unc-39* alleles were isolated from diverse genetic screens

The canonical allele of *unc-39*, *e257*, was isolated in a screen for animals with uncoordinated locomotion (Brenner, 1974); *ct73* in a screen for mutations that disrupt neuronal cell migration (Manser and Wood, 1990); and *rh72* in a distinct screen for mutations that affect cell migration (Hedgecock et al., 1987). Some cell morphology and migration defects of *unc-39(ct73)* have been previously described (Hedgecock et al., 1987; Manser and Wood, 1990). Consistent with these findings, we found that all three alleles of *unc-39* display uncoordinated locomotion (Unc), withered tail morphology (Wit), egg-laying defects (Egl), protruding vulvae, and multiple vulvae.

### *unc-39* affects mesodermal specification

*C. elegans* has a small number of mesodermal tissues (Sulston et al., 1983): striated muscles of the body wall; non-striated muscles of the pharynx, uterus, vulva and enteric system; male-tail associated muscles; the somatic component of the gonad (Z1 and Z4, and their descendants); six coelomocytes (nonmuscle mesodermal cells of unknown function); six GLRs; and the two head mesodermal cells. We analyzed the development of multiple mesodermal tissues in *unc-39* mutants and found that *unc-39* affects the development of coelomocytes, the head mesodermal cell, the M mesoblast (which gives rise to multiple mesodermal cell types), vulval muscles, and body wall muscles.

### Coelomocytes

Wild-type animals have six coelomocytes: four are born during embryonic development from the MS.ap and MS.pp lineages (Sulston et al., 1983), and two are born during postembryonic divisions of the M mesoblast cell (Sulston and Horvitz, 1977). M and the four embryonic coelomocytes are born in the anterior of the embryo and migrate posteriorly to their final positions (Fig. 1A). Previous analyses (Manser and Wood, 1990) indicated that *unc-39* mutants were defective in the migrations of the embryonic coelomocytes and the M cell, which often resulted in the localization of all six coelomocytes to the head region (data not shown). We found that in addition to perturbing coelomocyte migration, *unc-39* mutants had defects in coelomo-

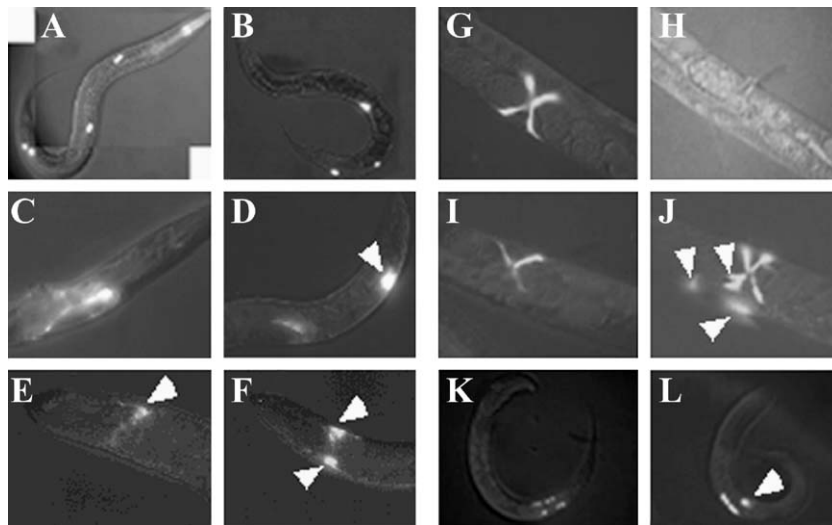


Fig. 1. Mesodermal defects in *unc-39* mutants. Fluorescent micrographs of living animals harboring mesodermal promoter::*gfp* transgenes. Animals in A, C, E, G, and K are wild type and animals in B, D, F, H, I, J, and L are *unc-39(ct73)*. (A, B) Animals harbor a coelomocyte *gfp* reporter (see Materials and methods). (C, D) Animals harbor a *lim-7::gfp* transgene expressed in gonadal sheath cells. An arrow in D points to an additional *lim-7::gfp*-positive cell in the head of *unc-39* mutant animal. (E, F) Animals harboring an *arg-1::gfp* transgene expressed in the hmc, vulval muscles, and enteric muscles. Arrowheads point to hmc expression. The *unc-39* mutant in F has two hmcs. (G–J) Animals harboring an *egl-15::gfp* transgene expressed in the differentiated vm1 vulval muscles. In wild-type animals, the vm1 muscles display a cruciform arrangement (G). The *unc-39* mutant in H lacks all vm1 cells, and the *unc-39* mutant in I lacks some vm1 cells. The *unc-39* mutant in J displays ectopic *egl-15::gfp*-expressing cells (arrowheads). Vulval muscles may be missing and/or ectopic vulval muscles can be seen (arrows in J). (K, L) Animals harboring a *twist::gfp* transgene expressed in M descendants. The wild-type animal in K has eight M progeny, four located dorsally and four located ventrally. The *unc-39* mutant in L has four dorsal progeny, but the ventral M daughter has failed to divide (arrowhead).

cyte specification. We observed coelomocytes using either Nomarski optics or with one of two *gfp* reporters (see Materials and methods). *unc-39(e257* and *ct73)* mutants displayed a variable number of coelomocytes ranging from one to seven (data not shown), and 85% of *ct73* animals and 45% of *e257* animals displayed coelomocyte specification defects (Table 1; Fig. 1B). In *e257* and *ct73*, the specification defects affected both the embryonic and postembryonically derived coelomocytes. In contrast, only the postembryonic coelomocytes were affected in *rh72* mutant animals (56%; Table 1). Thus, *unc-39* acts in coelomocyte specification as well as migration. That some mutants with greater than six coelomocytes were observed suggests that coelomocytes are also ectopically generated in *unc-39* mutants. We have not investigated the origins of the extra coelomocytes.

#### Head mesodermal cell

The right and left head mesodermal cells (hmcR and hmcL) are sisters of the somatic gonad precursors Z1 and Z4, respectively. Whereas Z1 and Z4 migrate posteriorly, hmcR and L migrate dorsally to positions over the pharynx. Each hmc has a unique wiring in the embryonic motor system; however, hmcR undergoes programmed cell death in late embryogenesis (Sulston et al., 1983).

In *unc-39* mutants an extra gonadal mesoblast (designated Z5) was formed in some embryos, for example, 11% of *e257* and 16% of *rh72* embryos (Table 1). Z5 cells could be visualized both with Nomarski optics and with the *lim-7::gfp* reporter transgene (Hobert and Westphal, 2000), which is expressed in Z1 and Z4 and their descendants (Figs. 1C and D). Further experiments (see below) indicated that Z5 was derived from the presumptive hmcR cell. Z1 and Z4 normally give rise to greater than 50 cells that form the somatic gonad (Kimble and Hirsh, 1979). The extent to which Z5 differentiated into a third somatic gonad arm differed between individuals, ranging from a single cell to a complete extra gonad arm that actually attached to the normal somatic gonad and became a functional third arm of the gonad with oocytes and fertilized eggs.

Z5 was always found on the right side of the body wall, usually midway along the Z1 migration path or at the hmcR position over the pharynx, rather than with the other gonadal cells. Possibly, hmcR, which normally undergoes programmed cell death, is sometimes transformed to its

sister fate Z1. However, we observed a late embryonic cell death with the characteristic position and morphology of hmcR, in individuals with a Z5 mesoblast. We suggest instead that hmcR (lineally termed MS.pppaaa) may sometimes divide in *unc-39* mutants, yielding an hmcR-like daughter and a Z1-like daughter (Z5).

In *ced-3* mutants, programmed cell deaths are generally suppressed (Ellis and Horvitz, 1986). In particular, hmcR survived to adulthood in *ced-3* mutants but neither divided nor was transformed in fate (data not shown). In *unc-39; ced-3* double mutants, the frequency of appearance of Z5 was significantly increased [62% in *ced-3; unc-39(e257)* compared to 11% in *unc-39(e257)* alone], consistent with the possibility that MS.pppaaa (hmcR) could be the source of Z5. We speculate that *unc-39* normally suppresses MS.pppaaa division and that, in *unc-39* mutants, this mis-differentiated cell can occasionally undergo programmed cell death, as does the normal hmcR, before it divides.

We also often observed two hmcL-like cells in adult *unc-39* worms (Figs. 1E and F, Table 1). *arg-1::gfp* is a marker for the differentiated hmcL fate (Kostas and Fire, 2002). In *unc-39* mutants, two *arg-1::gfp* positive cells (hmcL-like) are present 10–22% of the time (Table 1). We never observed the two hmcLs and Z5 in the same animal, suggesting that MS.pppaaa (hmcR) is the source of the second adult hmcL and that the Z5 and hmcL fates of MS.pppaaa daughters might be mutually exclusive.

The hmcs are highly branched cells with processes extending around the nerve ring and posteriorly along the dorsal nerve cord. Similar to the defects seen in neuronal branching (see below), we observed defects in hmc branching in *unc-39* mutants. These defects included misguided processes that extended anteriorly, increased numbers of processes extending around the nerve ring, and shortening of the posterior-directed processes (data not shown).

#### M mesoblast

The M mesoblast is a descendant of the MS.ap lineage (Sulston et al., 1983). It is born on the left-hand side of the embryo in the anterior. M migrates posteriorly during embryogenesis, gradually shifting to the right-hand side of the animal where it ultimately resides. M then undergoes postembryonic divisions to give rise to body wall muscles, sex myoblasts (SMs), and two coelomocytes in hermaphrodites (Sulston and Horvitz, 1977). We visualized M and its

Table 1  
Mesodermal defects in *unc-39* mutants

Percentage of animals ( <i>n</i> ) with							
Genotype	<6 ccs	Abnormal "Z5"	Two hmcL cells	hmc branching defects	Missing M cell	Loss of vulval muscles	<95 bwms
+	0 (100)	0 (100)	0 (62)	0 (50)	0 (50)	0 (100)	0 (30)
<i>unc-39(rh72)</i>	56 (72)	16 (100)	10 (50)	28 (50)	21 (102)	40 (162)	71 (42)
<i>unc-39(e257)</i>	45 (200)	11 (100)	21 (100)	27 (100)	4 (260)	21 (444)	71 (39)
<i>unc-39(ct73)</i>	85 (144)	24 (100)	22 (35)	22 (35)	22 (50)	31 (74)	76 (49)

descendants using Nomarski optics and in animals harboring a *twist::gfp* fusion (Harfe et al., 1998b).

Previously, defects in M migration in *unc-39(ct73)* were noted (Manser and Wood, 1990). Using *twist::gfp* as a reporter for the M lineage (Harfe et al., 1998b), we observed that only a very small percentage (less than 2%, data not shown) of animals displayed strong defects in M migration. In contrast, 22% of *ct73* animals did not contain a *twist::gfp*-positive M cell (Table 1); 4% of *e257* and 21% of *rh72* animals likewise lacked a *twist::gfp*-positive M cell.

One result of M misspecification was a loss of the postembryonic M-derived coelomocytes. However, M misspecification is not the sole cause of postembryonic coelomocyte defects, as *unc-39(rh72)*, which only perturbs postembryonic coelomocyte specification, caused a higher proportion of coelomocyte defects (56%) than M specification defects (21%) (Table 1). Thus, M-derived coelomocytes can be misspecified even when M is apparently normal, suggesting that *unc-39* may act at multiple points in the M lineage.

#### Vulval muscles

The M cell also gives rise to the two sex myoblasts (SMs), which divide to produce vulval and uterine muscles (Sulston and Horvitz, 1977). Using *egl-15::gfp* as a reporter for the vm1 vulval muscles, we found that vm1 vulval muscles were often missing or disorganized in *unc-39* mutants (Figs. 1G–I). The frequency of loss of all vulval muscles (Fig. 1H) was consistent with a defect in M specification and also correlated with the percentage of egg-laying-defective animals. We also observed several animals in which one to three of the vm1 muscles were missing (Fig. 1I), suggesting that *unc-39* acts at multiple steps in the M lineage as observed with coelomocyte specification.

Previous studies have shown that the M descendants that normally form the coelomocytes can be transformed to the SM fate (Harfe et al., 1998a). We examined *unc-39* animals for the presence of additional SMs using the *twist::gfp* reporter. Although we did not see an increase in the number of SMs, we often observed additional vulval muscles in *unc-39* animals (Fig. 1J). While these additional cells appeared to be integrated into the body wall musculature, they did not display the striated character of the body wall muscles. Instead, they appeared as smaller elongated cells, similar to the vulval muscles. They also expressed both *arg-1::gfp* and *egl-15::gfp*, indicators of the terminally differentiated vulval muscle fate. These reporters are not expressed by differentiated body wall muscles. While we do not know the origin of these extra vulva muscle-like cells, it is possible that they are misspecified coelomocytes. However, it is also possible that misspecification is occurring in other parts of the M lineage (possibly in body wall muscle specification; see below) leading to the formation of these cells.

#### Body wall muscles

Eighty-one body wall muscles (bwms) are present in the newly hatched L1 larva (Sulston et al., 1983). Eighty result from divisions of the MS, C and D cells, and one is from the AB lineage. The M mesoblast produces 14 additional postembryonic bwms (Sulston and Horvitz, 1977). We analyzed the number of bwms in *unc-39* adults by counting the number of cells that expressed *myo-3::gfp* (Kostas and Fire, 2002) and found that *unc-39* mutants often had fewer than 95 body wall muscles (Table 1), ranging from 81 to 95, suggesting that *unc-39* affects specification of the M-derived body wall muscles. We do not know if *unc-39* mutants have fate transformations of body wall muscles or if the immediate precursors of the body wall muscles are simply not generated.

Together, these data indicate that *unc-39* is involved in the development of multiple types of mesoderm, including the head mesodermal cells, coelomocytes, body wall muscles, and vulval muscles.

#### *unc-39* affects the development of anterior neurons

##### CAN neuron

*unc-39(ct73)* mutants were previously shown to affect CAN cell migration (Manser and Wood, 1990). Here, we show that *unc-39* mutations also affect CAN axon pathfinding. The bilateral CAN neurons, descendants of the AB.alap neuroblast, are born in the anterior of the embryo and migrate posteriorly during embryogenesis to their final lateral positions in the middle of the animal near the vulva (Sulston et al., 1983; White et al., 1986). The anterior CAN axon extends laterally to the head near the amphid ganglion, and the posterior CAN axon extends to the tail near the phasmid ganglion (Fig. 2A). As was observed for *unc-39(ct73)*, the other *unc-39* alleles *e257* and *rh72* also caused defects in the posterior migration of the CAN cells (Fig. 2B, Table 2).

*unc-39* mutants also displayed defects in the pathfinding of the posterior CAN axon (Fig. 2B, Table 2): posterior CAN axons were often misguided and wandered along the body wall outside of their normal lateral route or were prematurely terminated before reaching their normal final positions in the tail. Anterior CAN axons were also affected but at a much lower frequency (data not shown). Furthermore, CAN axons of *unc-39* animals often displayed ectopic axon branching. In some animals, the CAN neurons appeared to be absent. However, we cannot rule out a complete failure of CAN cell migration and posterior axon outgrowth resulting in the CAN cell body residing in the region of the amphid ganglion, obscured by *gfp*-expressing amphid neurons (Figs. 2A and B).

*unc-39* affects both CAN cell migration and axon pathfinding. Previous studies have shown that CAN cell migration and CAN axon pathfinding are under independent control (Lundquist et al., 2001). We analyzed the development of the wild-type CAN neuron using the

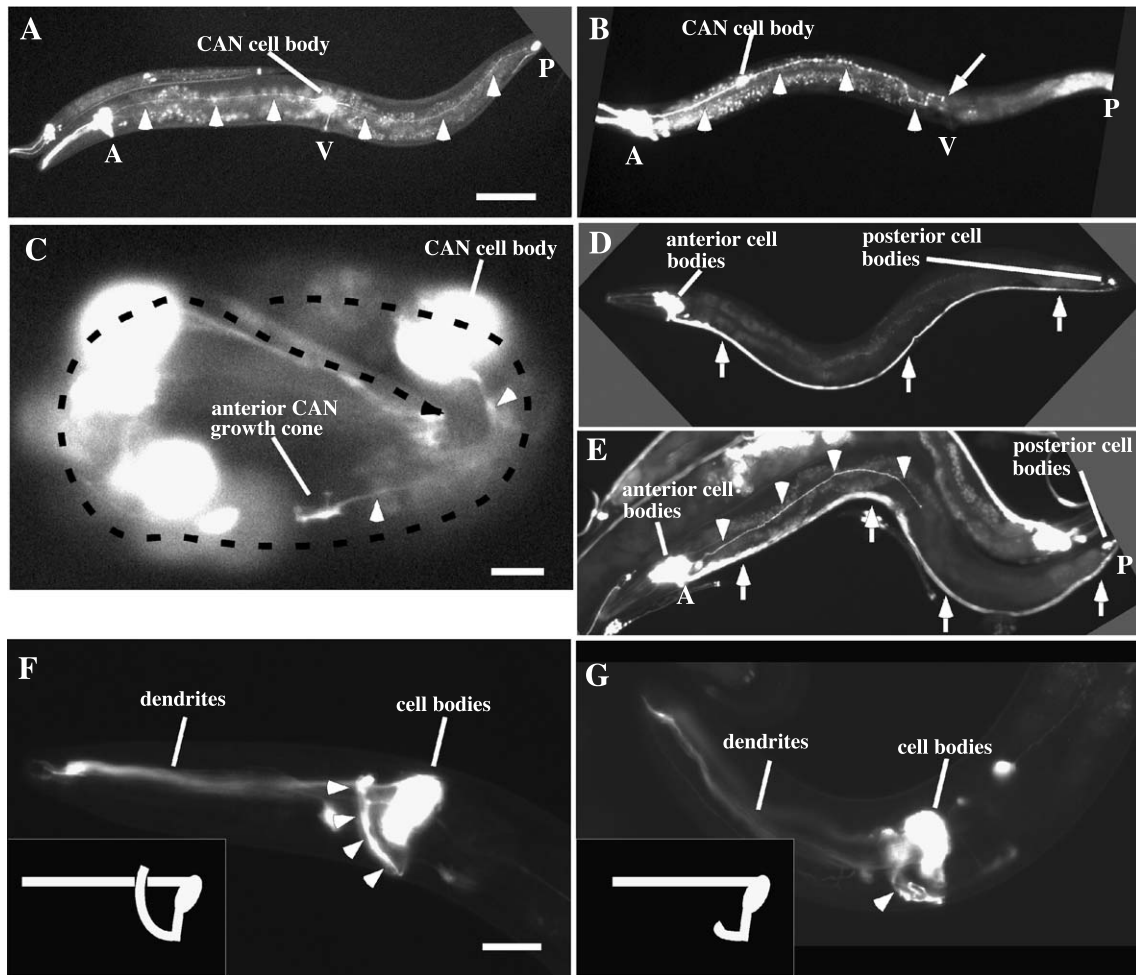


Fig. 2. Neuronal cell migration and axon pathfinding defects in *unc-39* mutants. Fluorescent micrographs of living animals harboring a *ceh-23::gfp* transgene (*lqls27*) expressed in amphid neurons and the CAN neurons (A, B, F, and G), an *unc-39::gfp* transgene expressed in the CAN and other neurons (C), and a *glr-1::gfp* transgene (*zdlis3*) expressed in anterior and posterior interneurons (D and E). Animals in A, C, D, and F are wild type, and animals in B, E, and G are *unc-39* mutants. Anterior is left; dorsal is up in all but C. (A) A CAN neuron cell body in wild type resided laterally at the midbody near the vulva (V). The arrowheads trace the trajectories of the anterior and posterior CAN axons. The amphid (A) and phasmid (P) cell bodies are indicated. (B) An *unc-39(ct73)* mutant, labeled as in A, displayed a CAN cell migration defect (the CAN cell body is located anterior to the vulva) and premature posterior CAN axon termination (arrow). (C) A wild-type 3-fold stage embryo with *unc-39::gfp* expression in the CAN neurons. The dashed line traces the body axis of the embryo folded upon itself in the eggshell; the arrow marks the anterior tip of the animal. The CAN cell body is in its final midbody position; however, a nascent CAN axon (arrowheads), tipped by a growth cone, is in the process of growing anteriorly. (D) A wild-type animal showed no lateral axons from *glr-1*-expressing anterior neurons. Arrows trace the *glr-1*-expressing axons in the ventral nerve cord. (E) An *unc-39(e257)* mutant displayed an ectopic lateral *glr-1*-expressing axon from anterior neurons (arrowheads). A second adult animal is present in the micrograph, dorsal to the labeled animal. (F) The anterior amphid region of a wild-type animal. The amphid axons (arrowheads) extended from the ventral surface dorsally past the dendrites. The inset is a diagram of amphid axon morphology. (G) The amphid axons of an *unc-39(e257)* mutant terminated prematurely before extending dorsally past the dendrites (arrowhead). Scale bar in A = 50 μm for A, B, D, and E; scale bar in C = 5 μm; scale bar in F = 10 μm for F and G.

*unc-39 promotor::gfp* transgene (see below). We found that the CAN cell body first migrates to its final midbody position and then extends anterior and posterior CAN axons (Fig. 2C). These results are consistent with the notion that CAN cell migration and axon pathfinding are independent events and that *unc-39* affects both processes.

#### Interneurons

The cell bodies of anterior *glr-1*-expressing neurons (Maricq et al., 1995), which are mostly interneurons involved in locomotion, normally reside in anterior ganglia

and send axons into the nerve ring or posteriorly down the ventral or dorsal nerve cords (White et al., 1986) (Fig. 2D). Although *unc-39* animals had a normal number of *glr-1*-expressing cells, these cells appeared to differentiate abnormally. In wild-type animals, no processes from these neurons were found along the lateral lengths of the animal. In *unc-39* mutants, abnormal processes were observed that emanated laterally from the anterior cluster of *glr-1*-expressing cell bodies and extended along the body wall (Fig. 2E, Table 2). In most cases, we could not resolve the identity of the individual neurons from which these lateral processes extended, nor could we resolve if these lateral

Table 2  
Neuronal cell migration and axon pathfinding defects in *unc-39* mutants

Genotype	Percentage of animals with defective <sup>a</sup>			
	CAN cell migration <sup>b</sup>	CAN axon pathfinding <sup>c</sup>	Anterior <i>glr-1</i> -expressing axon pathfinding <sup>d</sup>	Amphid axon pathfinding <sup>e</sup>
+	0	0	0	0
<i>unc-39(rh72)</i>	71 ± 9	63 ± 9	52 ± 9	21 ± 6
<i>unc-39(e257)</i>	58 ± 9	25 ± 8	13 ± 6	18 ± 7
<i>unc-39(ct73)</i>	77 ± 7	43 ± 9	30 ± 9	25 ± 6
<i>unc-39(e257); lqEx280 [unc-39(+)]</i>	9 ± 5	1 ± 4	ND	ND
<i>unc-39(RNAi)<sup>f</sup></i>	50 ± 16	38 ± 14	ND	ND

<sup>a</sup> ± 2 SE of the proportion.

<sup>b</sup> Percentage of CAN cell bodies that failed to reach the position of the vulva (see Figs. 2A and B).

<sup>c</sup> Percentage of posterior CAN axons that failed to extend posteriorly to the area of the phasmids (see Figs. 2A and B).

<sup>d</sup> Percentage of animals with lateral *glr-1*-expressing axons from anterior neurons (see Figs. 2D and E).

<sup>e</sup> Percentage of animals with premature amphid axon termination (see Figs. 1F and G).

<sup>f</sup> *unc-39* RNAi led to approximately 62% embryonic and larval arrest, and only those animals that survived to adulthood were scored.

processes were misguided axons or if they were additional, ectopic axons.

#### Amphid sensory neurons

The axons of the sensory amphid neurons were also defective in *unc-39* mutants. The bilateral amphid neuron cell bodies reside in the anteriorly located amphid ganglia (White et al., 1986) (Fig. 2F). Ciliated dendrites extend anteriorly from the cell bodies to the tip of the nose, and amphid axons extend ventrally down the amphid commissure. At the ventral midline, the axons enter the nerve ring and extend dorsally past the dendrites until the two fascicles of the nerve ring meet at the dorsal midline of the pharynx. In *unc-39* mutants, the amphids were disorganized and displayed defects such as misplaced cell bodies and defasciculated dendrites and axons (data not shown). Furthermore, the amphid axons often terminated prematurely near the ventral midline and failed to extend dorsally past the dendrites (Fig. 2G, Table 2). We could detect no gross changes in the number of amphid neurons in *unc-39* mutants (data not shown). Together, these data indicate that *unc-39* affects the development of the CAN neurons, interneurons, and sensory neurons located in or born in the anterior region of the embryo.

Previous studies indicate that *unc-39* mutations do not affect the migrations of other neuronal cell bodies (Manser and Wood, 1990), including the HSN, which is born in the posterior and migrates anteriorly to its final position in the midbody near the vulva. We analyzed the cell positions and axon morphologies of neurons born in midbody and posterior positions, including HSN, PVM, PDE, the phasmid

neurons, PVD, PVQ, PVC, PQR, and the VD and DD motor neurons. We could detect no abnormalities of these midbody and posterior neurons in *unc-39* mutants.

#### *unc-39* affects pharyngeal and buccal development

Each of the three *unc-39* alleles displayed a low-penetrance L1 larval arrest. The animals appeared to be starved, and close examination indicated that many of these animals had defects in pharyngeal development. We observed that the pharyngeal metacarpus was frequently misshapen (Figs. 3A–C) and that the luminal space was uneven (data not shown). In rare animals, we observed that the pharynx was unattached to and retracted from the buccal opening (Figs. 3D and E). Greater than 30% of the viable *unc-39* animals also displayed mild defects in pharyngeal morphology (data not shown). Furthermore, many *unc-39* animals (greater than 60% of all three alleles) displayed a ventrally displaced buccal opening (compare Figs. 3D and F).

#### *unc-39* encodes a molecule similar to Six-family homeodomain transcription factors

Genetic deficiency mapping data placed *unc-39* in a short interval of chromosome V (see Materials and

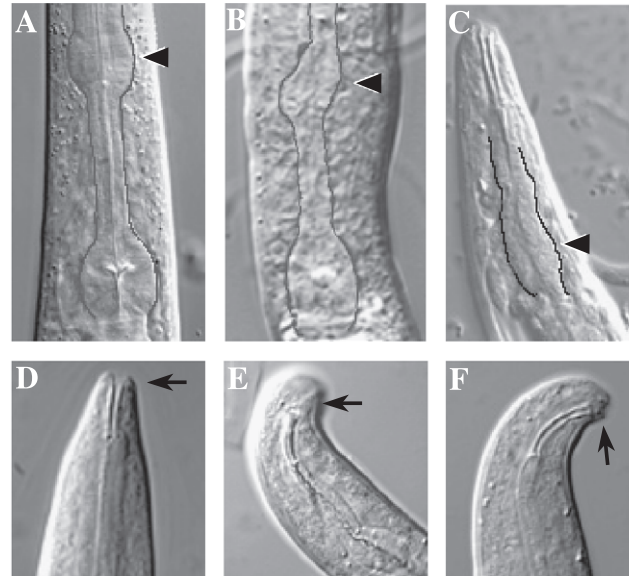


Fig. 3. Pharyngeal and buccal defects in *unc-39* mutants. Nomarski images of (L1) animals showing the pharyngeal bulbs (A–C) and buccal opening (D–F). (A–C) The pharynx has been outlined to emphasize the shape. (A) Wild type. The pharynx is straight and symmetric. (B, C) *unc-39(ct73)*. The pharynx is frequently asymmetric and misshapen. Black arrowheads point to the metacarpus, the region of the pharynx that is most severely affected in mutant animals. (D) Wild-type animal showing that the buccal opening (black arrow) is centered and symmetric. (E) *unc-39(ct73)* mutant showing a pharynx that has retracted from the hypodermis (black arrow). (F) An *unc-39(ct73)* mutant with a ventrally displaced buccal opening (black arrow; ventral is to the right).

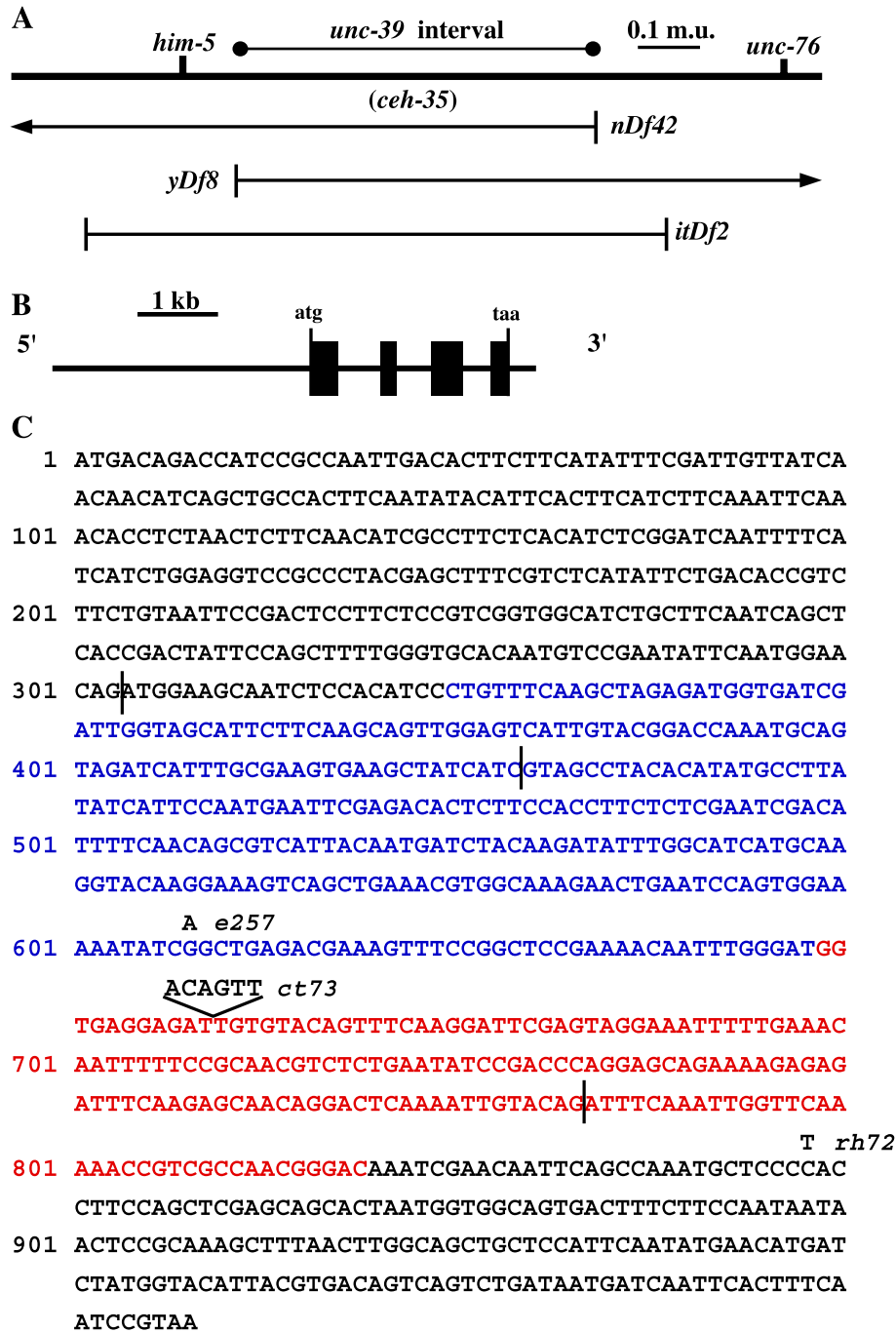


Fig. 4. UNC-39 is similar to *Drosophila* Six4 and human Six5. (A) A genetic map of the *unc-39* region (m.u. = map units). The deficiencies *nDf42*, *yDf8*, and *itDf2* all uncover *unc-39* whereas only *nDf42* and *itDf2* uncover *him-5*. The genetic interval that must contain *unc-39* is indicated by a barbell. This region of the *C. elegans* genome contains the predicted gene *ceh-35*, which can encode a Six-class homeodomain protein. (B) The *ceh-35* genomic region included in the *unc-39* rescuing construct [bases 10759–17539 of cosmid F56A12 (GenBank accession number Z81551)]. The filled boxes represent the exons of the predicted *ceh-35* gene (F56A12.1), beginning with an initiator ATG codon and ending with a TAA stop codon. (C) The nucleotide sequence of the open reading frame encoded by an *unc-39* cDNA (GenBank accession number CB400463; kindly provided by M. Vidal). The blue and red nucleotides represent the coding region of the Six and homeodomains, respectively; vertical bars represent the positions of introns; and the notations above the sequence indicate the nucleotide changes associated with the three *unc-39* alleles. (D) The amino acid sequence of the predicted CEH-35/UNC-39 polypeptide. The Six domain is indicated in blue, the homeodomain is red, and nonconserved regions are black. The numbering represents amino acid residues. The residues that are identical between UNC-39, *Drosophila* Six4 (GenBank accession number AAD39864), and human Six5 (GenBank accession number NP787071) are underlined. The amino acid changes associated with the three *unc-39* mutations are indicated above the sequence: *e257* alters conserved arginine 203 to glutamine in a highly conserved region of the Six domain; *ct73* is a six-base insertion resulting in the insertion of two amino acids in a conserved region of the homeodomain; and *rh72* changes proline 296 to leucine in a nonconserved region of the polypeptide. (E) A comparison of the CEH-35/UNC-39 predicted polypeptide with human Six5 and *Drosophila* Six4. Numbering indicates amino acid residues and the percentage of identical amino acid residues present in the Six domain and homeodomains of the molecules are indicated. The numbers below Dm Six4 are the identities of the regions in DmSix4 to CEH-35/UNC-39.

**D**

1 MTDHPPIDTSSSYFDCYQQHQLPLOYTFTSSSSNSNTSNTSSTSPSHISDQFS

SSGGPPYELSSHILTPSSVIPTSPSPSVASASISSPTIPAFGCTMSEYSME

101 QMEAISTSLFQARDGDRLVAFFKQLESLYGPNVADHLRSEAIIVAYTYAL

YHSNEFETLFHLLSNRHFQQRHYNDLQDIWHHARYKESQLKRGKELNPVE

Q e257 QF ct73

201 KYRLRRKFPAPKTIWDGEEIVYSFKDSSRKFLKQFFRNVSEYPTQEOKRE

rh72 L

ISRATGLKIVQISNWFKNRRQDKSNNSAKCSPSSSSSSTNGGSDFLPII

301 TPQSFNLAAPFNMNIYGTLRDSQSDNDQFTFNP

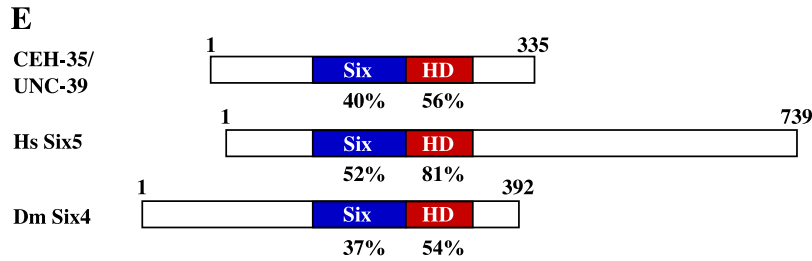


Fig. 4 (continued).

methods and Fig. 4A). This region contains the predicted gene *ceh-35* (The *C. elegans* Genome Sequencing Consortium, 1998) (Figs. 4B and C), which can encode a molecule similar to the Six4/5 family of homeodomain transcription factors (Figs. 4D and E). A cDNA (kindly provided by M. Vidal) confirmed the gene structure and sequence reported here. Six-class homeodomain proteins have been implicated in anterior pattern formation and mesodermal development in other organisms. CEH-35 is similar to *Drosophila* Six4 and human Six5 (Dozier et al., 2001), with highest sequence conservation in the Six and homeodomains and little conservation outside of these regions (Figs. 4D and E).

Three lines of evidence indicated that the *unc-39* mutations affect *ceh-35*. First, we found that a transgene (the *unc-39* rescuing transgene; see Fig. 4B and Materials and methods) containing 4 kb of the *ceh-35* upstream region and the entire *ceh-35* coding region rescued the Unc and Wit phenotypes of *unc-39(ct73)* (data not shown), as well as the coelomocyte defects (Table 3) and the CAN cell migration and axon defects (Table 2) of *unc-39(ct73)*. Second, we found that the *ceh-35* coding regions from the three existing *unc-39* alleles contained mutations (Figs. 4C and D). *unc-39(e257)* had a missense mutation that changed arginine 203 to glutamine in the Six domain, the identical conserved arginine altered in the *Drosophila D-Six4<sup>131</sup>* mutation (Kirby et al., 2001); *unc-39(ct73)* had a six-base insertion resulting in two extra amino acid residues (QF) between residues 220 and 221 in a conserved region of the homeodomain; and *unc-39(rh72)* had a missense mutation that changed proline 296 to leucine in a non-

conserved region of the molecule. Third, RNA-mediated gene interference (RNAi) of *ceh-35* mirrored the *unc-39* phenotype: *ceh-35* RNAi caused 60% larval lethality, and surviving adult animals were Unc, Wit, and displayed CAN cell migration and axon pathfinding defects (Table 1) and coelomocyte specification defects (data not shown). These experiments indicate that *unc-39* mutations affect the *ceh-35* gene. We will refer to this locus as *unc-39*.

No *unc-39* mutation was a deletion or premature stop mutation. Additionally, *unc-39(e257)/Df* animals were sick and slow-growing and had more severe Unc and Wit defects than *unc-39(e257)* homozygotes (data not shown). Furthermore, we found that *unc-39* RNAi led to 60% larval

Table 3  
Rescue of *unc-39* embryonic coelomocyte defects by the human Six5 domain and homeodomain

Number of embryonic coelomocytes	Number of animals				
	0 cc	1 cc	2 cc	3 cc	4 cc
<i>unc-39(ct73)</i>	2	10	29	34	42 (36%)
<i>unc-39(ct73);[unc-39(+)] #1</i>			13		59 (82%)
<i>unc-39(ct73);[unc-39(+)] #2</i>			6	1	19 (73%)
<i>unc-39(ct73);[unc-39::human Six/HD] #1</i>			14	2	24 (60%)
<i>unc-39(ct73);[unc-39::human Six/HD] #2</i>			4	5	75 (89%)
<i>unc-39(ct73);[unc-39::human Six/HD] #3</i>					71 (100%)
<i>unc-39(ct73);[unc-39::Δ Six/HD]</i>		1	11	19	21 (40%)

lethality, more than was seen in any of the three *unc-39* mutants. Together, these observations suggest that no extant mutation is null for *unc-39*. We attempted to obtain

an *unc-39* null allele using two separate F1 non-complementation screens and a PCR-based deletion screen without success.

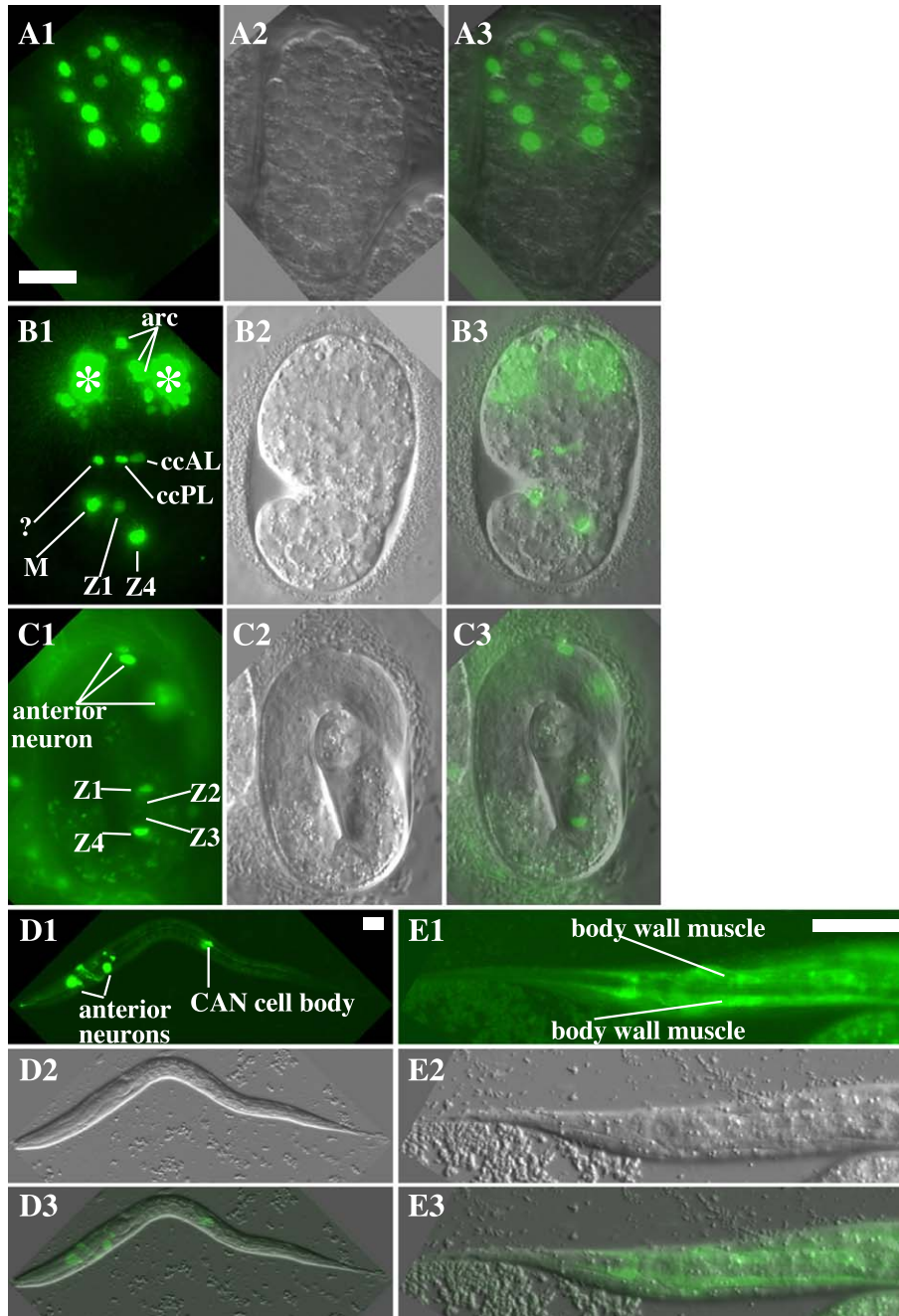


Fig. 5. *unc-39* is expressed in anterior ectoderm and mesoderm. Micrographs of animals harboring a full-length *unc-39::gfp* transgene that rescues *unc-39* mutants (A–C), or an *unc-39 promoter::gfp* transgene (D and E). In A–C, anterior is up. In D and E, anterior is left and dorsal is up. Panels labeled 1 are fluorescence micrographs showing *gfp* expression; panels labeled 2 are the DIC micrographs of the same animals, and panels labeled 3 are overlays of the panels 1 and 2. A1 and B1 are flattened images of deconvolved Z sections through the embryo. (A) At approximately 150 min post-fertilization, UNC-39::GFP was seen to accumulate in the nuclei of 12–16 blast cells in the anterior of the embryo. (B) At approximately 350 min post-fertilization, UNC-39::GFP accumulation was seen in posterior nuclei whose positions were consistent with the M mesoblast, the somatic gonad founders Z1 and Z4, and the embryonic coelomocytes (ccAL and ccPL). A small number of other unidentified nuclei also showed UNC-39::GFP (indicated by a ?). By this time, many anterior nuclei (>50) also showed UNC-39::GFP. The positions of these nuclei are consistent with anterior neurons and neuroblasts. The pharyngeal arcade cells (arc) also showed UNC-39::GFP expression. (C) An 800-min post-fertilization embryo (just before hatching) showed UNC-39::GFP accumulation in a few (<10) anterior nuclei whose positions and morphologies are consistent with anterior neurons, and in Z1 and Z4 (the germ cells Z2 and Z3, which did not express UNC-39::GFP, are indicated). (D) A newly hatched L1 larva displayed *unc-39* promoter activity in anterior neurons in the area of the amphids and in the CAN neurons. (E) A newly hatched L1 larva displayed *unc-39* promoter activity in body wall muscle cells. The scale bars = 10  $\mu$ m (A1 for A–C; D1 for D; and E1 for E).

### *The unc-39 gene is active in anterior ectoderm and mesoderm*

To determine where and when *unc-39* is expressed, we constructed an *unc-39* reporter transgene by fusing 4 kb of the *unc-39* upstream region and the entire *unc-39* coding region in frame to *gfp* (see Materials and methods). This transgene is predicted to produce a full-length UNC-39 molecule with GFP at the C-terminus. This construct rescued the Unc, Wit, and coelomocyte defects of *unc-39(e257 and ct73)* mutants (Table 3 and data not shown), indicating that the transgene can encode a functional UNC-39::GFP molecule.

Expression of the full-length *unc-39::gfp* transgene was dynamic throughout development. UNC-39::GFP was first seen at approximately 100 min post-fertilization when gastrulation and embryonic transcription begin (Figs. 5A1–3). UNC-39::GFP accumulated in the nuclei of 12–16 blast cells in the anterior of the embryo. As embryogenesis proceeded, expression was observed in increasing numbers of anterior nuclei (Figs. 5B1–3) which were most likely the descendants of the earlier-expressing cells, as UNC-39::GFP-expressing pairs of cells in the process of cell division were observed (data not shown). Expression in anterior nuclei continued until about 450 min post-fertilization (2-fold stage), at which time expression became restricted to less than 10 cells and persisted into adulthood (Figs. 5C1–3). The locations and morphologies of some of these cells indicated that they were anterior neuroblasts derived from the AB lineage.

At approximately 240 min post-fertilization, UNC-39::GFP appeared in more posteriorly located nuclei (Figs. 51–3) in addition to the anterior nuclei. Posterior nuclei expressing *unc-39::gfp* included Z1 and Z4 (the founders of the somatic gonad), the M cell, and the embryonic coelomocytes. Occasionally, additional posteriorly located cells that we were unable to unambiguously identify showed UNC-39::GFP expression. Expression in M and the coelomocytes was transient and no longer observed by 600 min post-fertilization, whereas Z1 and Z4 expression persisted into the L1 larval stage (Figs. 5C1–3). Z1, Z4, M, and four coelomocytes are all generated from MS descendants in the anterior region of the embryo and, like the CAN neuron, migrate posteriorly to their final positions (Sulston et al., 1983).

The 4-kb *unc-39* promoter driving the expression of *gfp* alone showed a similar general pattern. However, expression of this construct persisted in a greater number of anterior cells into larval and adult stages (Figs. 5D1–3). These cells had axon and dendrite morphologies characteristic of amphid sensory neurons. Furthermore, the CAN neuron clearly showed *unc-39* promoter expression (Figs. 5D1–3).

In addition to expression in the anteriorly generated mesodermal Z1, Z4, M, and coelomocytes, the *unc-39* promoter was active in all of the body wall muscle cells

of late embryos and L1 larvae (Figs. 5E1–3). Thus, *unc-39::gfp* transgenes were expressed very early in embryogenesis in anterior cells that generate neuronal and mesodermal tissues, and later in body wall muscle cells.

### *The Six domain and homeodomain from human Six5 can substitute for those from UNC-39*

*unc-39* is similar to human *Six5* (Dozier et al., 2001). To determine if UNC-39 and *Six5* are functionally equivalent, we removed the UNC-39 Six domain and homeodomain coding region from the full-length *unc-39::gfp* rescue construct and inserted the Six domain and homeodomain coding region from human *Six5* (see Materials and methods). This *unc-39::human Six/HD* transgene partially rescued the coelomocyte specification, Unc and Wit defects of *unc-39(ct73)* (Table 3 and data not shown). In contrast, a control transgene lacking the Six domain and homeodomain did not rescue the coelomocyte defects of *unc-39(ct73)* (Table 3).

## Discussion

We have cloned and characterized the *unc-39* gene, which encodes a *C. elegans* member of the Six4/5 family and which we showed is a functional homolog of the human myotonic dystrophy-associated homeodomain protein *Six5*. Like other *Six* family members, we have found that *unc-39* affects mesodermal specification and as well as the differentiation of anterior ectoderm, including neuronal migration and axon pathfinding.

### *unc-39 affects mesodermal specification and differentiation*

Although no extant allele of *unc-39* is null, we have analyzed three loss-of-function alleles of *unc-39* and found a variety of defects in mesodermal development. First, we noted that some mesodermal cells were misspecified in *unc-39* mutants. For example, specification defects in the M cell and its descendants led to missing or extra coelomocytes and vulval muscles. Second, though apparently specified normally, we noted that some mesodermal cells displayed defects in differentiation in *unc-39* mutants, such as the failure of coelomocytes to execute their full-length migrations and abnormal hmcR branching.

### *unc-39 affects anterior neuronal differentiation*

In addition to mesodermal defects, we found that *unc-39* mutants displayed defects in the differentiation of anterior neurons, including the CAN neurons, the amphid neurons, and the command interneurons. The CAN neurons often displayed both neuronal cell migration defects and axon pathfinding defects, and the amphid neurons and command interneurons displayed axon pathfinding defects. Unlike in

the mesoderm, we could detect no neuronal specification defects (e.g. missing or extra neurons). Furthermore, we could detect no effect of *unc-39* on midbody or posterior neurons, suggesting that *unc-39* affects the development of anterior neurons only. This idea is strengthened by the observation that *unc-39* was expressed in anterior neuroblasts and neurons but not in posterior neuroblasts or neurons.

#### *UNC-39 controls long-range cell and axon migrations*

The body plan of *C. elegans* is primarily formed by morphogenetic movements rather than by cell migration. Only 14 of the 558 cells in the newly hatched larva have undergone extensive migration, the majority along the A–P axis (Hedgecock et al., 1987; Manser and Wood, 1990; Sulston et al., 1983). We see effects of *unc-39* in many of these migratory cells, including the M mesoblast, coelomocytes, head mesodermal cell, and CAN neurons. All of these cells are derived from anterior blastomeres and execute posterior migrations. Possibly, one role of *unc-39* is to regulate anterior to posterior cell migration.

Interestingly, *unc-39* also affects the fate specification of many of these posteriorward migrating cells. Although *unc-39* could be affecting migration and differentiation independently, it is tempting to speculate that *unc-39* acts at an interface between these processes, integrating migratory cues to influence cell fate determination and terminal differentiation.

#### *UNC-39 is a Six-class homeodomain transcription factor*

We have shown that *unc-39* encodes a Six4/5-class homeodomain transcription factor that is similar to and is a functional homolog of human Six5. Thus, it is likely that UNC-39 regulates the expression of other genes involved in mesodermal and neuronal specification and differentiation. This result along with the observation that many of the cells affected by *unc-39* mutations display *unc-39::gfp* expression suggests that the effects of *unc-39* are cell-autonomous. Based on the model that *unc-39* controls many aspects of neuronal and mesodermal specification and differentiation, we might expect a dynamic expression pattern of *unc-39* in which populations of cells express the gene at different stages of embryonic development. The expression of UNC-39::GFP in the nuclei of dividing and migrating cells of both the mesoderm and anterior ectoderm is consistent with this model. Furthermore, the early expression of *unc-39::gfp* in neuroblasts and mesoderm suggests that *unc-39* might be a key early regulator of neuronal and mesodermal differentiation or that *unc-39* might regulate multiple steps of differentiation (as observed in the M lineage). *unc-39::gfp* is also expressed in cells that show no defects in existing *unc-39* mutants, including the somatic gonad precursors Z1 and Z4 and the embryonically derived body wall muscles.

The *unc-39* mutations *e257* and *ct73* affected the conserved Six domain and homeodomain, respectively. In contrast, *rh72* affected a proline residue in a region of the molecule not conserved in *Drosophila* Six4 or human Six5 but present in UNC-39 from the related nematode species *Caenorhabditis briggsae* (data not shown). Furthermore, *rh72* mutants displayed only postembryonic coelomocyte defects and exhibited a higher loss of vulval muscle cells than did *e257* and *ct73*. Possibly, the *rh72* mutation affects a previously unknown domain of UNC-39 that is not conserved in human Six5 or *Drosophila* Six4.

#### *unc-39 as a model for myotonic dystrophy type I*

Myotonic dystrophy type I is associated with a triplet repeat expansion on human chromosome 19 (Brook et al., 1992; Fu et al., 1993; Harley et al., 1992). The relationship between this expansion and the pathophysiology of the DM1 is not yet understood at a cellular level. One model suggests that the binding of CUG-binding protein (CUG-BP) to the triplet repeat expansion within the 3' untranslated region of DMPK globally alters alternative splicing (Philips et al., 1998). A second model suggests that the expansion creates a silent chromatin domain (Ottens and Tapscott, 1995). To date, three genes within the vicinity of the expansion have been shown to have reduced mRNA levels: the DM associated protein kinase, DMPK (Carango et al., 1993); the downstream gene *Six5* (Klesert et al., 1997; Thornton et al., 1997); and an upstream gene, DM-associated WD-repeat protein, *DMWD* (Alwazzan et al., 1999). We have shown that the *C. elegans unc-39* gene is the homolog of human *Six5*. Several lines of evidence indicate that different genes are likely to be responsible for different subsets of defects in the complex DM1 disease state. *DMPK*(+/-) and (-/-) mice show deficiencies in sodium channel conductance in skeletal myocytes, providing direct evidence that DMPK affects this abnormality in the human disease (Mounsey et al., 2000). Winchester et al. (1999) showed that changes in *Six5* expression in the adult human lens are related to cataract development, one of the most frequent defects in DM1 patients. Similarly, *Six5* knockout mice showed an increased frequency of cataracts (Klesert et al., 2000; Sarkar et al., 2000). In humans, *Six5* expression is detected in the ganglion cells of the eye (Winchester et al., 1999). It is postulated that loss of *Six5* expression might contribute to the reduction in the number of these cells seen in DM1 patients. However, if the situation in DM1 is analogous to *unc-39* animals, the underlying defect in the ganglion cells may also include a failure in ganglion cell migration or axon pathfinding.

The variable expressivity and penetrance of *unc-39* defects are reminiscent of the pleiotropy seen in DM1 patients. For example, *unc-39* mutants display defects in muscle cell specification, neuronal cell migration, and axonal outgrowth, defects which could explain in part the

myotonia (along with *DMPK*) and mental impairment seen in DM1 patients. Neither the mouse nor *Drosophila* models display such similarity to DM1 in terms of the range and variability of defects, suggesting that *C. elegans* may serve as a useful model system for understanding the function of Six5 in DM1. Underscoring the value of *unc-39* as a model for *Six5* function is our observation that the human Six5 Six domain and homeodomain can substitute for these domains of UNC-39 in our *unc-39* genetic rescue assays. These results indicate that some of the DNA binding specificity and some of the protein–protein interactions must have been maintained during the evolution of nematodes and humans. Further studies of *unc-39* will not only elucidate the function of a *Six5* family member in vivo, but also allow identification and characterization of other genes important in *Six5* function, including upstream regulators and downstream targets.

### Acknowledgments

We thank Scott Clark, Wayne Forrester, and Gian Garriga for sharing *C. elegans* strains and reagents; Steve Kostas, Jun Kelly Liu, Chris Weise, Bruce Vogel, and Lisa Timmons for helpful discussions and suggestions; Rachel Brewster and Harry Hochheiser for valuable comments on the manuscript; and Eric Struckhoff and Stephanie Mariano for technical assistance. Some strains used in this work were provided by the *Caenorhabditis* Stock Center, which is funded by the National Center for Research Resources. This work was supported by NIH Grants R01GM37706 to A.Z.F. and R01NS40945 to E.A.L. and an NIH postdoctoral fellowship to J.L.Y.

### References

- Alwazzan, M., Newman, E., Hamshire, M.G., Brook, J.D., 1999. Myotonic dystrophy is associated with a reduced level of RNA from the DMWD allele adjacent to the expanded repeat. *Hum. Mol. Genet.* 8, 1491–1497.
- Boucher, C.A., King, S.K., Carey, N., Krahe, R., Winchester, C.L., Rahman, S., Creavin, T., Meghji, P., Bailey, M.E., Chartier, F.L., et al., 1995. A novel homeodomain-encoding gene is associated with a large CpG island interrupted by the myotonic dystrophy unstable (CTG)n repeat. *Hum. Mol. Genet.* 4, 1919–1925.
- Brenner, S., 1974. The genetics of *Caenorhabditis elegans*. *Genetics* 77, 71–94.
- Brook, J.D., McCurrach, M.E., Harley, H.G., Buckler, A.J., Church, D., Aburatani, H., Hunter, K., Stanton, V.P., Thirion, J.P., Hudson, T., et al., 1992. Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* 69, 385.
- Carango, P., Noble, J.E., Marks, H.G., Funanage, V.L., 1993. Absence of myotonic dystrophy protein kinase (DMPK) mRNA as a result of a triplet repeat expansion in myotonic dystrophy. *Genomics* 18, 340–348.
- Cheyette, B.N., Green, P.J., Martin, K., Garren, H., Hartenstein, V., Zipursky, S.L., 1994. The *Drosophila sine oculis* locus encodes a homeodomain-containing protein required for the development of the entire visual system. *Neuron* 12, 977–996.
- Collet, J., Spike, C.A., Lundquist, E.A., Shaw, J.E., Herman, R.K., 1998. Analysis of *osm-6*, a gene that affects sensory cilium structure and sensory neuron function in *Caenorhabditis elegans*. *Genetics* 148, 187–200.
- Dozier, C., Kagoshima, H., Niklaus, G., Cassata, G., Bürglin, T.R., 2001. The *Caenorhabditis elegans* Six/sine oculis class homeobox gene *ceh-32* is required for head morphogenesis. *Dev. Biol.* 236, 289–303.
- Ellis, H.M., Horvitz, H.R., 1986. Genetic control of programmed cell death in the nematode *C. elegans*. *Cell* 44, 817–829.
- Filippova, G.N., Thienes, C.P., Penn, B.H., Cho, D.H., Hu, Y.J., Moore, J.M., Klesert, T.R., Lobanenkov, V.V., Tapscott, S.J., 2001. CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nat. Genet.* 28, 335–343.
- Fire, A., Xu, S.-Q., Montgomery, M.K., Kostas, S.A., Driver, S.E., Mello, C.C., 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806–811.
- Forrester, W.C., Perens, E., Zallen, J.A., Garriga, G., 1998. Identification of *Caenorhabditis elegans* genes required for neuronal differentiation and migration. *Genetics* 148, 151–165.
- Fu, Y.H., Friedman, D.L., Richards, S., Pearlman, J.A., Gibbs, R.A., Pizzuti, A., Ashizawa, T., Perryman, M.B., Scarlato, G., Fenwick Jr., R.G., et al., 1993. Decreased expression of *myotonin-protein kinase* messenger RNA and protein in adult form of myotonic dystrophy. *Science* 260, 235–238.
- Gallardo, M.E., Lopez-Rios, J., Fernaud-Espinosa, I., Granadino, B., Sanz, R., Ramos, C., Ayuso, C., Sellar, M.J., Brunner, H.G., Bovolenta, P., Rodríguez de Córdoba, S., 1999. Genomic cloning and characterization of the human homeobox gene SIX6 reveals a cluster of SIX genes in chromosome 14 and associates SIX6 hemizygosity with bilateral anophthalmia and pituitary anomalies. *Genomics* 61, 82–91.
- Harfe, B.D., Branda, C.S., Krause, M., Stern, M.J., Fire, A., 1998a. MyoD and the specification of muscle and non-muscle fates during postembryonic development of the *C. elegans* mesoderm. *Development* 125, 2479–2488.
- Harfe, B.D., Vas Gomez, A., Kenyon, C., Liu, J., Krause, M., Fire, A., 1998b. Analysis of a *Caenorhabditis elegans* *Twist* homolog identifies conserved and divergent aspects of mesodermal patterning. *Genes Dev.* 12, 2623–2635.
- Harley, H.G., Rundle, S.A., Reardon, W., Myring, J., Crow, S., Brook, J.D., Harper, P.S., Shaw, D.J., 1992. Unstable DNA sequence in myotonic dystrophy. *Lancet* 339, 1125–1128.
- Harper, P.S., 2001. Myotonic Dystrophy, third ed., vol. 37. Saunders, London.
- Heanue, T.A., Reshef, R., Davis, R.J., Mardon, G., Oliver, G., Tomarev, S., Lassar, A.B., Tabin, C.J., 1999. Synergistic regulation of vertebrate muscle development by *Dach2*, *Eya2*, and *Six1*, homologs of genes required for *Drosophila* eye formation. *Genes Dev.* 13, 3231–3243.
- Hedgecock, E.M., Culotti, J.G., Hall, D.H., Stern, B.D., 1987. Genetics of cell and axon migrations in *Caenorhabditis elegans*. *Development* 100, 365–382.
- Hobert, O., Westphal, H., 2000. Functions of LIM-homeobox genes. *Trends Genet.* 16, 75–83.
- Kawakami, K., Sato, S., Ozaki, H., Ikeda, K., 2000. Six family genes—Structure and function as transcription factors and their roles in development. *BioEssays* 22, 616–626.
- Kimble, J., Hirsh, D., 1979. The postembryonic cell lineages of the hermaphrodite and male gonads in *Caenorhabditis elegans*. *Dev. Biol.* 70, 396–417.
- Kirby, R.J., Hamilton, G.M., Finnegan, D.J., Johnson, K.J., Jarman, A.P., 2001. *Drosophila* homolog of the myotonic dystrophy-associated gene, *SIX5*, is required for muscle and gonad development. *Curr. Biol.* 11, 1044–1049.
- Klesert, T.R., Otten, A.D., Bird, T.D., Tapscott, S.J., 1997. Trinucleotide repeat expansion at the myotonic dystrophy locus reduces expression of DMAHP. *Nat. Genet.* 16, 402–406.
- Klesert, T.R., Cho, D.H., Clark, J.I., Maylie, J., Adelman, J., Snider, L., Yuen, E.C., Soriano, P., Tapscott, S.J., 2000. Mice deficient in *Six5*

- develop cataracts: implications for myotonic dystrophy. *Nat. Genet.* 25, 105–109.
- Kobayashi, M., Toyama, R., Takeda, H., Dawid, I.B., Kawakami, K., 1998. Overexpression of the forebrain-specific homeobox gene *six3* induces rostral forebrain enlargement in zebrafish. *Development* 125, 2973–2982.
- Kostas, S.A., Fire, A., 2002. The T-box factor MLS-1 acts as a molecular switch during the specification of nonstriated muscle in *C. elegans*. *Genes Dev.* 16, 257–269.
- Laclef, C., Hamard, G., Demignon, J., Souil, E., Houbron, C., Maire, P., 2003. Altered myogenesis in *Six1*-deficient mice. *Development* 130, 2239–2252.
- Lagutin, O., Zhu, C.C., Furuta, Y., Rowitch, D.H., McMahon, A.P., Oliver, G., 2001. *Six3* promotes the formation of ectopic optic vesicle-like structures in mouse embryos. *Dev. Dyn.* 221, 342–349.
- Lundquist, E.A., Reddien, P.W., Hartwig, E., Horvitz, H.R., Bargmann, C.I., 2001. Three *C. elegans* Rac proteins and several alternative Rac regulators control axon guidance, cell migration and apoptotic cell phagocytosis. *Development* 128, 4475–4488.
- Manser, J., Wood, W.B., 1990. Mutations affecting embryonic cell migrations in *Caenorhabditis elegans*. *Dev. Genet.* 11, 49–64.
- Maricq, A.V., Peckol, E., Driscoll, M., Bargmann, C.I., 1995. Mechano-sensory signalling in *C. elegans* mediated by the GLR-1 glutamate receptor. *Nature* 378, 78–81.
- Mello, C., Fire, A., 1995. DNA transformation. *Methods Cell Biol.* 48, 451–482.
- Mounsey, J.P., Mistry, D.J., Ai, C.W., Reddy, S., Moorman, J.R., 2000. Skeletal muscle sodium channel gating in mice deficient in myotonic dystrophy protein kinase. *Hum. Mol. Genet.* 9, 2313–2320.
- Otten, A.D., Tapscott, S.J., 1995. Triplet repeat expansion in myotonic dystrophy alters the adjacent chromatin structure. *Proc. Natl. Acad. Sci. U. S. A.* 92, 5465–5469.
- Pasquier, L., Dubourg, C., Blayau, M., Lazaro, L., Le Marec, B., David, V., Odent, S., 2000. A new mutation in the six-domain of *SIX3* gene causes holoprosencephaly. *Eur. J. Hum. Genet.* 8, 797–800.
- Philips, A.V., Timchenko, L.T., Cooper, T.A., 1998. Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science* 280, 737–741.
- Sarkar, P.S., Appukuttan, B., Han, J., Ito, Y., Ai, C., Tsai, W., Chai, Y., Stout, J.T., Reddy, S., 2000. Heterozygous loss of *Six5* in mice is sufficient to cause ocular cataracts. *Nat. Genet.* 25, 110–114.
- Seimiya, M., Gehring, W.J., 2000. The *Drosophila* homeobox gene *optix* is capable of inducing ectopic eyes by an eyeless-independent mechanism. *Development* 127, 1879–1886.
- Serikaku, M.A., O'Tousa, J.E., 1994. *Sine oculis* is a homeobox gene required for *Drosophila* visual system development. *Genetics* 138, 1137–1150.
- Spitz, F., Demignon, J., Porteu, A., Kahn, A., Concordet, J.P., Daegelen, D., Maire, P., 1998. Expression of myogenin during embryogenesis is controlled by *Six/sine oculis* homeoproteins through a conserved MEF3 binding site. *Proc. Natl. Acad. Sci. U. S. A.* 95, 14220–14225.
- Sulston, J.E., Horvitz, H.R., 1977. Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev. Biol.* 56, 110–156.
- Sulston, J.E., Schierenberg, E., White, J.G., Thomson, J.N., 1983. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* 100, 64–119.
- Thornton, C.A., Wymer, J.P., Simmons, Z., McClain, C., Moxley III, R.T., 1997. Expansion of the myotonic dystrophy CTG repeat reduces expression of the flanking DMAHP gene. *Nat. Genet.* 16, 407–409.
- Wakimoto, H., Maguire, C.T., Sherwood, M.C., Vargas, M.M., Sarkar, P.S., Han, J., Reddy, S., Berul, C.I., 2002. Characterization of cardiac conduction system abnormalities in mice with targeted disruption of *Six5* gene. *J. Interv. Card Electrophysiol.* 7, 127–135.
- Wallis, D.E., Roessler, E., Hehr, U., Nanni, L., Wiltshire, T., Richieri-Costa, A., Gillesen-Kaesbach, G., Zackai, E.H., Rommens, J., Muenke, M., 1999. Mutations in the homeodomain of the human *SIX3* gene cause holoprosencephaly. *Nat. Genet.* 22, 196–198.
- White, J.G., Southgate, E., Thomson, J.N., Brenner, S., 1986. The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philos. Trans. R. Soc. London* 314, 1–340.
- Winchester, C.L., Ferrier, R.K., Sermoni, A., Clark, B.J., Johnson, K.J., 1999. Characterization of the expression of DMPK and *SIX5* in the human eye and implications for pathogenesis in myotonic dystrophy. *Hum. Mol. Genet.* 8, 481–492.