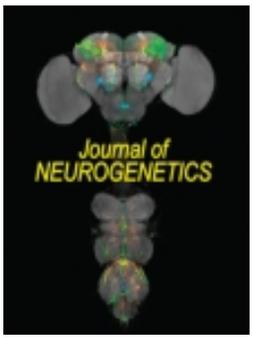



Journal of Neurogenetics

ISSN: 0167-7063 (Print) 1563-5260 (Online) Journal homepage: https://www.tandfonline.com/loi/ineg20

# The Unc-119 Family of Neural Proteins is Functionally Conserved Between Humans, *Drosophila* and *C. Elegans*

**Morris F. Maduro, Michael Gordon, Roger Jacobs & David B. Pilgrim**

**To cite this article:** Morris F. Maduro, Michael Gordon, Roger Jacobs & David B. Pilgrim (2000) The Unc-119 Family of Neural Proteins is Functionally Conserved Between Humans, *Drosophila* and *C. Elegans*, Journal of Neurogenetics, 13:4, 191-212, DOI: 10.3109/01677060009084494

**To link to this article:** https://doi.org/10.3109/01677060009084494

*J. Neurogenetics*, Vol. 13(4), pp. 191–212
Reprints available directly from the publisher
Photocopying permitted by license only



# THE UNC-119 FAMILY OF NEURAL PROTEINS IS FUNCTIONALLY CONSERVED BETWEEN HUMANS, *DROSOPHILA* AND *C. ELEGANS*

MORRIS F. MADURO[a,c], MICHAEL GORDON[b], ROGER JACOBS[b] and DAVID B. PILGRIM[c,*]

[a]*Department of Molecular, Cellular, and Developmental Biology and Neuroscience Research Institute, University of California, Santa Barbara, CA 93106, USA;* [b]*Department of Biology, McMaster University, Hamilton, ON Canada L8S 4K1;* [c]*Department of Biological Sciences, University of Alberta, Edmonton, AB Canada T6G 2E9*

*(Received 1 April 1999; Revised 31 May 1999)*

*C. elegans* animals mutant for the *unc-119* gene exhibit movement, sensory and behavioral abnormalities. Consistent with a nervous system role, *unc-119* reporter genes are expressed throughout the *C. elegans* nervous system. The UNC-119 protein has strong sequence similarity to the predicted protein from a human gene, HRG4/HsUNC-119, whose transcript is abundant in the retina. Using these similarities, we have identified a *Drosphila* homolog, DmUNC-119, which is expressed in the *Drosophila* nervous system. The predicted *C. elegans*, human and *Drosophila* gene products are conserved across two domains. Expression of portions of HRG4/HsUNC-119 or DmUNC-119, directed by the *unc-119* promoter, can fully rescue the *C. elegans unc-119* mutant phenotype. We tested the ability of portions of HRG4/HsUNC-119 to rescue, and found that its function in *C. elegans* requires the conserved carboxyl terminus, while the dissimilar amino terminus is dispensable. UNC-119, HRG4 and DmUNC-119 constitute members of a new class of neural genes whose common function has been maintained through metazoan evolution.

*Keywords:* Nervous system; neural development; retinal gene; novel proteins

* Corresponding author. Tel.: (780)492-2792. Fax: (780)492-9234.
E-mail: dave.pilgrim@ualberta.ca.

## INTRODUCTION

It has long been doctrine that by studying the development of genetically tractable model systems, we will learn fundamental principles which can be applied to vertebrate systems. The development of the nervous system has been an unusually fruitful field of investigation, particularly in *Drosophila*, where the conservation of neural developmental processes at the molecular level has been particularly well demonstrated (Thor, 1995; Bier, 1997). Recently, however, *C. elegans* has made some important contributions to this field. Its small size, nearly invariant cell lineage, ease of genetics and established pattern of neuronal connectivity make *C. elegans* an ideal system in which to study the development of a complete nervous system (Brenner, 1974; Sulston and Horvitz, 1977; White *et al.*, 1986; Wood, 1988). *C. elegans* has one important advantage over *Drosophila*, in that it does not need to move to feed or to mate, and therefore mutations in their nervous or muscle systems that severely compromise the ability to move can be recovered as viable homozygous strains. Such mutations are typically lethal in *Drosophila* (e.g. Van Vactor *et al.*, 1993).

A satisfying outcome of the molecular identification of nervous system genes in the nematode has been that many of these factors are remarkably conserved with vertebrate genes in both structure and function. For example, the genes *unc-5*, *unc-6* and *unc-40*, which are involved in the migration of mesodermal cells and pioneering axons (Hedgecock *et al.*, 1990), have related vertebrate homologs. The UNC-6 protein is structurally related to the netrin family of guidance cues which establish dorsal–ventral gradients, providing cues to commissural interneurons in the developing vertebrate spinal cord (Serafini *et al.*, 1994; Kennedy *et al.*, 1994; Wadsworth *et al.*, 1996). The *unc-40* gene, required primarily for ventral migrations (Hedgecock *et al.*, 1990), encodes a homolog of the DCC cell surface receptor (Chan *et al.*, 1996; Hedrick *et al.*, 1994; Fazeli *et al.*, 1997) which can bind netrin-1 (Keino-Masu *et al.*, 1996). Mutations of *Drosophila* and mouse DCC homologs also result in guidance defects (Kolodziej *et al.*, 1996; Fazeli *et al.*, 1997), implying a similar biological role for DCC and UNC-40.

For cases in which cellular function or molecular interactions of a gene are unknown, a genetically manipulable system such as *C. elegans* can be used to directly assay for conservation of function through the

expression of chimeric transgenes. For example, the novel UNC-76 protein, proposed to have an intracellular role in axon fasciculation, is structurally related to two human proteins, FEZ1 and FEZ2 (fasciculation and elongation protein, zygin/zeta-1; Bloom and Horvitz, 1997). A FEZ1 transgene can partially replace *unc-76* function in *C. elegans*, suggesting that FEZ1 has a similar role in humans (Bloom and Horvitz, 1997). The collection of such conserved neural genes is expanding, adding to the list of factors likely to be important players in neural development and/or function.

We have previously described the behavioral phenotype of *unc-119* null mutants (Maduro and Pilgrim, 1995). Animals exhibit uncoordinated movement, a weak egg laying defect, constitutive (versus regulated) pharyngeal pumping, and an inability to form dauer larvae. The body paralysis appears to result from a motor neuron defect, as muscle ultrastructure is normal and animals respond postsynaptically to an acetylcholine agonist, levamisole, with hypercontraction of body wall muscles (Lewis *et al.*, 1980; Maduro and Pilgrim, 1995). Furthermore, the dauer formation defect is suppressible by mutation of *daf-7*, suggesting that *unc-119* is not essential for entry into dauer *per se* (Maduro and Pilgrim, 1995). The simplest explanation of these results is that *unc-119* is required for correct development or function of the nervous system. This is consistent with the expression of *unc-119*::lacZ reporter transgenes in most, if not all, *C. elegans* neurons (Maduro and Pilgrim, 1995; 1996). Recently, we have shown that a full length fusion of UNC-119 and the Green Fluorescent Protein (GFP) is able to rescue *unc-119* mutants, and the fluorescence is only detected within the neurons (Materi *et al.*, submitted) consistent with lack of a secretory targeting signal on UNC-119. Therefore, since neuronal targeting signals in *C. elegans* have been shown to be expressed from non-neuronal tissues (Ishii *et al.*, 1992; Wadsworth *et al.*, 1996; Colavita *et al.*, 1998) the UNC-119 protein is presumed to act in the neuron as a downstream component of the neuronal targeting or outgrowth machinery.

Shortly after we reported the sequence of *unc-119*, a similar open reading frame (ORF) was identified by Higashide *et al.* (1996) in a screen for retina-specific cDNAs. The transcript of this gene (HRG4, for Human Retinal Gene 4), is apparently enriched in rods and cones (Higashide *et al.*, 1996), where the protein is localized to the synaptic ends of photoreceptors, the ribbon synapse (Higashide *et al.*, 1998).

Although the similarity between UNC-119 and HRG4 has been previously noted (Swanson *et al.*, 1998; Higashide *et al.*, 1998), no functional relatedness between these genes has yet been demonstrated.

In this paper, we describe the similarities between the HRG4 and UNC-119 sequences, and demonstrate the presence of two distinct conserved domains A and B. We use the conservation in domain B to clone and map a homolog, DmUNC-119, from the fruit fly *Drosophila melanogaster*. This protein also exhibits the same bipartite structure as HRG4 and UNC-119. We show that as with *unc-119* in *C. elegans*, the product of DmUNC-119 is expressed in the developing *Drosophila* nervous system. We find that *unc-119*::HRG4 and *unc-119*::DmUNC-119 transgenes can fully rescue *unc-119* null mutations, and with *unc-119*::HRG4, this depends upon the presence of conserved sequences. Our results suggest that UNC-119 is the founding member of a new class of functionally related neural proteins.

## MATERIALS AND METHODS

### Manipulation of DNA and RNA

All cloning was performed according to standard protocols (Sambrook *et al.*, 1989) with enzymes obtained from GibcoBRL unless otherwise noted. All DNA constructs were cloned into pBluescript KS-(Stratagene). The sequences of oligonucleotides are given in Table I.

TABLE I Oligonucleotides used in this work

| *Name* | *Sequence (5′ to 3′) in IUPAC notation* |
|---|---|
| MMA4 | CTACTCGCAGGATCCATAATTTCCCG |
| MMA7 | GCTACAACAGGATCCATGAAGGC |
| MMA14 | CCAGCCCAGAACTGGATCCTCCTGGG |
| MMA19 | CGGTGGATCCGGGATGGCGACGGAGTC |
| MMA20 | TCAGGGGATCCCGCTGTAGGAATAGTC |
| MMA23 | GACGGATCCGIATGATHGAIMGICAY |
| MMA25 | CTGGGATCCACRAARTARAAISWRTC |
| MMA32 | GCCGGATCCAAAATGCCGGGGCCGCTGCAGAG |
| MMA33 | GCCGGATCCAAAATGGTGACTCCCGACGAGGTG |
| MMA34 | GCCGGATCCATATTGCCGCCATCGTAG |

### Molecular Identification of DmUNC-119

A single PCR product encoding part of a *D. melanogaster* homolog was amplified from 100 ng of Oregon R DNA (a gift from Gary Ritzel, University of Alberta) with the degenerate oligonucleotides MMA23 and MMA25 using the following conditions: (95°C, 5 min; 53°C, 90 s; 73°C, 5 min) for one cycle, and (93°C, 75 s; 55°C, 75 s; 73°C, 3 min) for 30 cycles. A single product containing part of HRG4 was also obtained from total human DNA under these conditions (data not shown). A larger genomic fragment of the *Drosophila* gene was obtained by screening a pBluescript KS- mini-library of ~5 kb genomic *Hin*DIII fragments, constructed using a procedure previously described for a size-selected *Sst*I library of *C. elegans* (Maduro and Pilgrim, 1995). Coding regions were deduced by conservation with UNC-119/HRG4 and by prediction of splice sites, which were confirmed by RT-PCR (see below). GenBank accession numbers: HRG4, U40998; UNC-119, U32854; DmUNC-119, AF119102.

### Construction of *unc-119*::HRG4 and *unc-119*::DmUNC-119 Fusions

Novel *Bam*HI sites were generated in pDP#MM016 immediately upstream of the endogenous start (ATG) and stop (TAA) codons using two rounds of oligonucleotide-mediated site-directed mutagenesis (Kunkel, 1985) with the primers MMA4 and MMA7 (Table I). The resultant plasmid was digested with *Bam*HI and the intervening segment containing the entire UNC-119 coding region was replaced with the various HRG4 and DmUNC-119 inserts. Translation is predicted to initiate at the novel ATG of both fragments; the carboxyl terminus is fused to the 3′ noncoding region of *unc-119*. The pDP#MM147 plasmid was obtained by *Pst*I digestion of pDP#MM103 followed by religation, which deletes an in-frame 42-bp fragment from the HRG4 coding region.

### Amplification of HRG4 and DmUNC-119 cDNAs

For HRG4, first-strand cDNA was synthesized from 1 μg of total human retina RNA (a gift from Ian MacDonald, University of Alberta) with primer MMA14. For DmUNC-119, mixed-stage total RNA was prepared with TRIzol reagent (BRL) from adults and larvae of the

*Drosophila* strain *vg/CyO* (a gift from Ross Hodgetts, University of Alberta) using a standard RNA extraction protocol and used for a first-strand reaction with primer MMA34. In both cases, Superscript II Reverse Transcriptase was used with buffers supplied with the BRL 5′RACE kit. PCR was performed on first-strand cDNA using MMA19/20 (for HRG4) or MMA33/34 (for DmUNC-119) using the conditions (95°C, 5 min; 61°C, 45 s; 72°C, 2 min) for one cycle, and (93°C, 1 min; 61°C, 45 s; 73°C, 2 min) for 29 cycles using a Stratagene Robocycler 40 and reagents supplied with *Taq* polymerase (Amersham). The products were digested with *Bam*HI and cloned into pBluescript KS- to generate pDP#MM092 (HRG4) and pDP#MM154 (DmUNC-119). A novel start codon (ATG) was introduced by each upstream primer. The cloning of the HRG4 cDNA was confirmed by dideoxy sequencing (Sanger *et al.*, 1977) using Sequenase 2.0 (Amersham) and cloning of DmUNC-119 was confirmed using Thermo Sequenase (Amersham) by following the instructions provided with each kit. The derivative HRG4 coding regions were obtained by PCR from pDP#MM103 using MMA32/20 for pDP#MM150 and MMA19/25 for pDP#MM148.

### Construction of Transgenic *C. elegans* Strains

Microinjection of the syncytial gonad of young adult hermaphrodite worms was performed as described (Mello *et al.*, 1991). Plasmid DNA was prepared by alkaline lysis (Sambrook *et al.*, 1989) and injected at a concentration of 100 ng/μL. The presence of the transgenes was monitored by the Rol phenotype for the pRF4 marker or for rescue of the *unc-119* phenotype in strains carrying the *unc-119* rescuing clone pDP#MM016B.

Extrachromosomal arrays were integrated as follows. Fifty $P_0$ animals were treated with 1500 rad of $\gamma$ irradiation from a $^{60}Co$ source. Three hundred transgenic $F_1$ animals were picked to self-fertilize individually, and $F_2$ animals were picked from plates showing $> 75\%$ transmission. Homozygosity of the integrants was verified in the $F_3$ prior to backcrossing to N2.

### Phenotypic Assays

Analysis of dauer formation and quantitative analysis of locomotion, pharyngeal pumping, and egg retention were performed as described in Maduro and Pilgrim (1995).

### Localization *in situ* of *Drosophila* Gene Products

Embryo fixation and immunocytochemistry were performed as described (Patel, 1994) employing 22C10 monoclonal (from S. Benzer, further characterized by Brewster and Bodmer, 1996), imaged with an HRP conjugated secondary antibody (Jackson Immunoresearch). The insert from plasmid pDP#MM154 (described above) was used as a probe. Subsequent *in situ* hybridization to RNA probes was performed according to Tautz and Pfeifle (1989) with modifications by D. Mellerick (Mellerick and Nirenberg, 1995).

## RESULTS

### Identification of a *Drosophila* UNC-119 Homolog

The existence of vertebrate homologs of *C. elegans unc-119* (Higashide *et al.*, 1998; Swanson *et al.*, 1998) prompted us to search for a similar gene in *Drosophila*. We used the strong amino acid conservation between the carboxyl halves of HRG4/HsUNC-119 and UNC-119 to construct a set of degenerate oligonucleotides for polymerase chain reaction (PCR) as described in Materials and Methods. A single 444-bp product was obtained upon amplification from genomic DNA of the *D. melanogaster* strain Oregon R. This product was cloned into pBluescript KS-(Stratagene) to make plasmid pDP#MM113.

We sequenced the pDP#MM113 insert and found two short coding regions, separated by an apparent intron, with similarity to HRG4/HsUNC-119 and UNC-119. A *Drosophila* genomic Southern blot probed with this fragment detected a single ~ 5-kbp *Hin*DIII fragment. We cloned this fragment from a *Hin*DIII plasmid mini-library made in pBluescript KS-(pDP#MM121). The sequence of the majority of pDP#MM121 was obtained and used to retrieve expressed sequence tags (ESTs) from the Genbank EST database. Six overlapping ESTs were identified that define a presumptive mRNA of approximately 2.3 kb and spanning five exons (Fig. 1A; Accession numbers AF083305, AA202977, AA264600, AA539927, AA942390 and AA438392). We independently confirmed the splicing of the central portion of the gene by sequencing a partial cDNA obtained by RT-PCR from mixed-stage *Drosophila* embryos (plasmid pDP#MM154). While we cannot exclude alternative transcription, the overlap of the ESTs is consistent with a

single DmUNC-119 message encoding an ORF of 265 amino acids (aa) (Fig. 1B). Based on similarities in amino acid sequence with HRG4/HsUNC-119 and *C. elegans* UNC-119 (see below), we are calling this gene DmUNC-119.

## Mapping of DmUNC-119

The insert from pDP#MM113 was used for *in situ* hybridization to *Drosophila* polytene chromosomes. The DmUNC-119 gene maps to the X chromosome, band 7A1 (E. Woloshyn and D. Nash, unpublished data). While there are no known viable mutations mapped to this region, a FlyBase query (http://flybase.bio.indiana.edu) shows that some lethals map in the vicinity of 7A1 (e.g. lines ESHS24 through ESHS30). If DmUNC-119 has a role in the embryonic nervous system, a lethal phenotype might be expected as a consequence of its mutation; therefore, lethals in the 7A1 region may be candidates for DmUNC-119 disruptions (see Discussion).

## Conservation of UNC-119 Structure

Although DmUNC-119 was identified based on the conservation of amino acids in the carboxyl portions of UNC-119 and HRG4/HsUNC-119, DmUNC-119 shares substantial homology with these proteins in the amino halves as well (Fig. 2A). An alignment of all three proteins

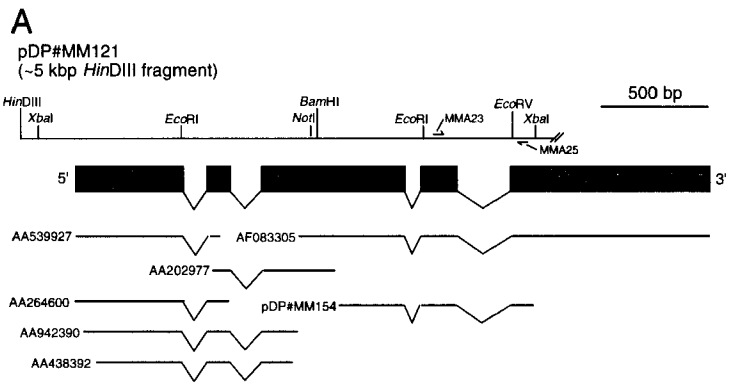



FIGURE 1A

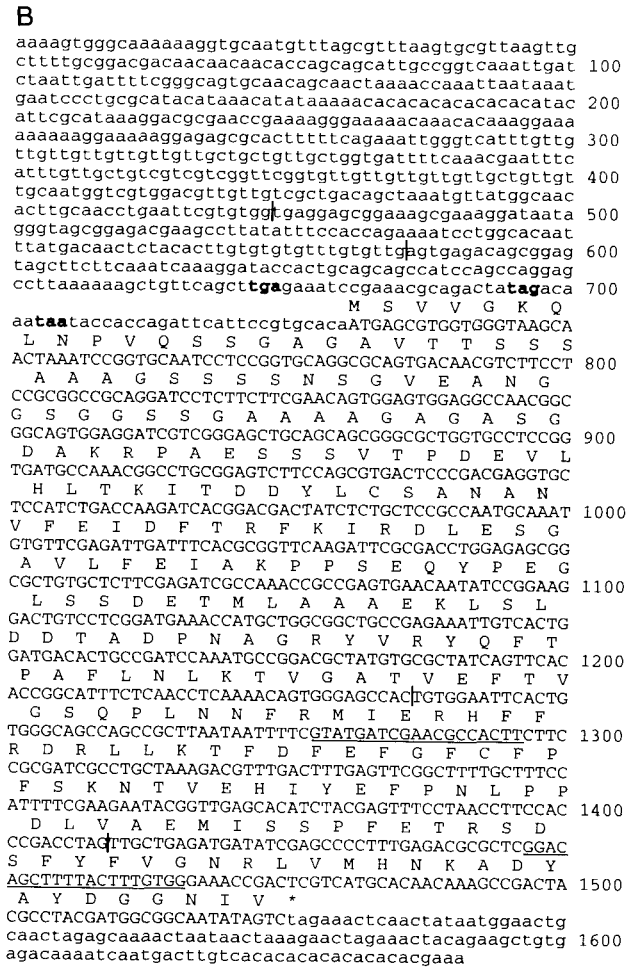



FIGURE 1B

FIGURE 1 (A) Gene structure of DmUNC-119. A restriction map of part of clone pDP#MM121 is shown. The predicted cDNA is indicated as boxes below the map, with darker shading to indicate the DmUNC-119 ORF. The alignment of six *Drosophila* ESTs, and a partial cDNA (pDP#MM154), are shown below the cDNA. The region 3′ to the *Xba*I site in pDP#MM121 has not been determined, but an additional 667 nucleotides are predicted by EST AF083305. (B) Sequence of DmUNC-119 coding region, and flanking regions deduced from ESTs and partial sequence of genomic clone pDP#MM121. The sequence is shown 5′–3′. Positions of introns are indicated with a vertical line (|). The ORF is shown in capital letters, with the predicted amino acid sequence of the DmUNC-119 protein shown in single-letter code above the first base of each codon in the ORF. Upstream stop codons are shown in boldface, while an asterisk (*) indicates the putative DmUNC-119 stop codon. The regions where degenerate primers MMA23 and MMA25 anneal are underlined. The 3′ end of the transcript has not been determined, although EST AF083305 contains an additional 667 nucleotides beyond the last base shown.

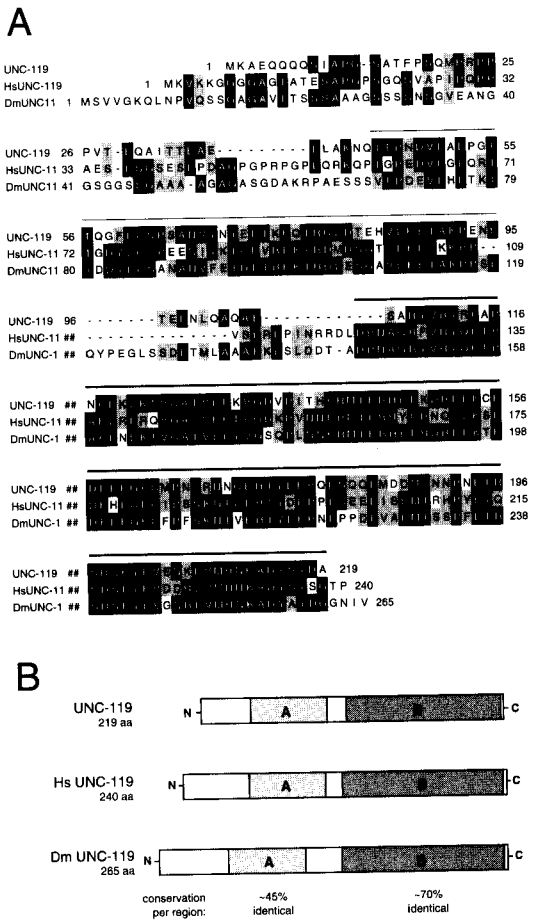



FIGURE 2 Comparison of *C. elegans* UNC-119, human HsUNC-119 and Drosophila DmUNC-119. (A) Amino acid alignment. Identities present in at least two of the three genes are shown as white text in black boxes, and conserved residues are shaded gray. A single line above the alignment indicates the extent of region A, while a double line shows region B. (B) Schematic alignment of the three homologs showing the two regions A and B, and the percentage identities and conserved changes (similarities) with HsUNC-119.

also suggests that, while there are many similarities throughout their coding regions, the homology between UNC-119 family members defines two broad domains which we will refer to as regions A and B (Fig. 2B). Within each region, the degree of conservation is significantly different: the percentage of amino acids identical across all three

homologs is approximately 45% for region A, and almost 70% for region B. Consistent with a bipartite structure, in UNC-119 and HRG4/HsUNC-119 these regions are separated by about 10 aa, while in DmUNC-119 they are almost 25 aa apart; these 'spacer' regions have only limited conservation (Fig. 2A).

## Expression of DmUNC-119 in *Drosophila* Embryos

The expression of *unc-119* throughout the *C. elegans* nervous system (Maduro and Pilgrim, 1995), which implies a role for UNC-119 in many neuron types, contrasts with the apparent photoreceptor-specific expression of HRG4/HsUNC-119 (Higashide *et al.*, 1996; 1998; Swanson *et al.*, 1998). To determine the expression pattern of DmUNC-119, we used the pDP#MM154 partial cDNA as a probe for *in situ* hybridization to *Drosophila* embryos (Fig. 3). The DmUNC-119 transcript appears at

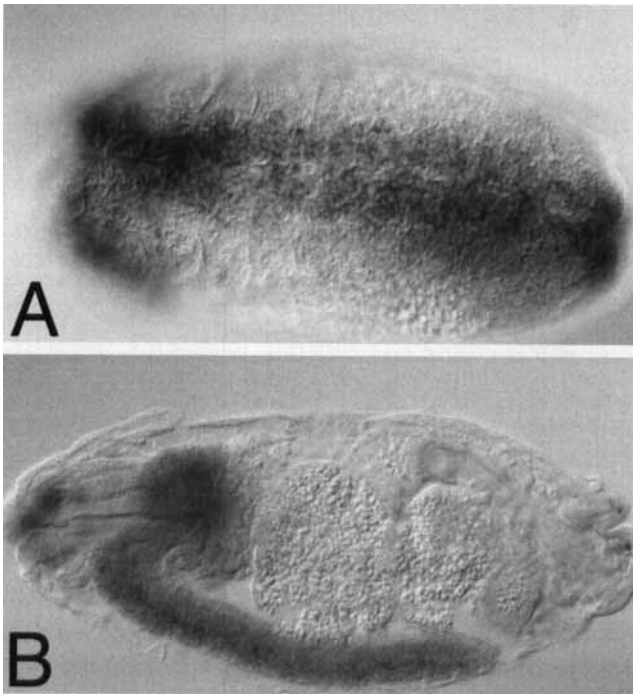



FIGURE 3A and B

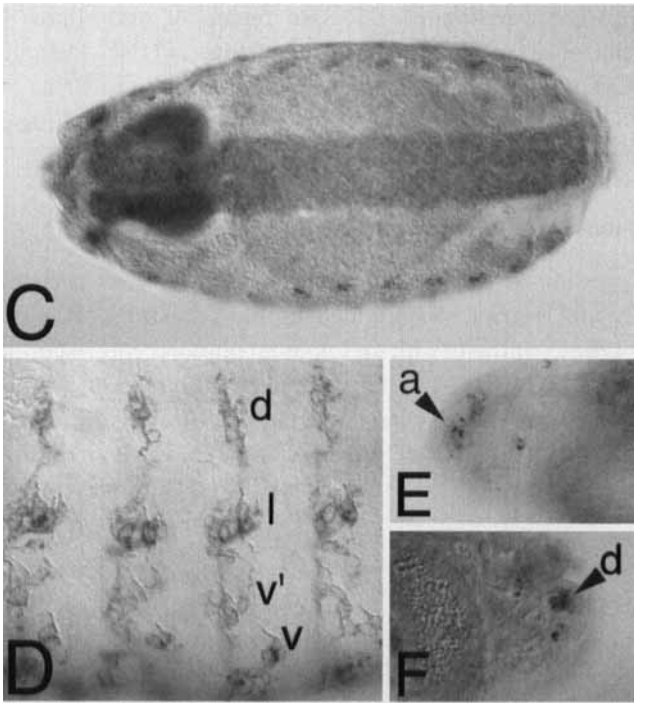



FIGURE 3C–F

FIGURE 3 Distribution of DmUNC-119 transcript in embryos. DmUNC-119 transcript accumulation is first detected in late stage 11 embryos, uniformly through the neuromeres of the central nervous system (A). A uniform and ubiquitous distribution through the nervous system is maintained for the remainder of embryogenesis, as seen in sagittal view at stage 16 (B) and frontal view at stage 13 (C). Neurons of the peripheral nervous system also express DmUNC-119. Lateral ectoderm of a stage 14 embryo double labeled for 22C10 (brown) and DmUNC-119 (blue) demonstrates extensive co-labeling in the dorsal (d), lateral (l) and ventral (v', v) sensory clusters. Neurons of the antenno-maxillary complex (a in panel E) and the dorsal sensory cluster of the telson (d in panel F) also express DmUNC-119. (See Color Plate I at back of the issue.)

late stage 11, after the first neurons are generated from neuroblasts, and as they begin their differentiation. Expression appears to be uniform and ubiquitous in the central nervous system. In the peripheral nervous system, the results are a little clearer, and expression seems to be confined to neurons, although the strength of expression may be variable between neurons. Figure 3 panel F illustrates the terminal segment homolog of the PNS of the other segments. Therefore, while vertebrate

expression of UNC-119 is apparently retina-specific (Higashide *et al.*, 1996; 1998; Swanson *et al.*, 1998), in *Drosophila* and *C. elegans*, expression occurs in a large part of the entire nervous system.

## UNC-119 is Functionally Conserved

The strong structural similarities between UNC-119, DmUNC-119 and HRG4/HsUNC-119 suggest that these proteins might also share the same *in vivo* activity. We tested the ability of these genes to replace UNC-119 function by taking advantage of the known behavioral defects in *C. elegans unc-119* animals (Maduro and Pilgrim, 1995). As the expression of vertebrate UNC-119 is restricted to the retina, we might expect its function to be more diverged from *C. elegans* UNC-119 than the *Drosophila* gene; therefore, ascertainment of rescue by HRG4/HsUNC-119 is likely to be a more discriminating test of conservation. We obtained a partial HRG4/HsUNC-119 cDNA by performing RT-PCR on total human retina RNA (see Materials and Methods). A partial cDNA was used to make plasmid pDP#MM103, an *unc-119*::HRG4/HsUNC-119 fusion gene in which the *C. elegans* UNC-119 coding region has been replaced by HRG4/HsUNC-119. This construct preserves both the 5′ and 3′ flanking regions of the *C. elegans* gene, and is expected to direct expression of the human homolog with the same temporal and spatial pattern as the wild-type *unc-119* gene.

Nematodes homozygous for the *unc-119* null allele *ed3* were made transgenic for pDP#MM103 by gonadal microinjection (Mello *et al.*, 1991). Animals mutant for *unc-119* display uncoordinated locomotion, defective dauer formation, increased egg retention and constitutive pharyngeal pumping (Maduro and Pilgrim, 1995). Among the progeny of *ed3* mutants injected with pDP#MM103 were animals that displayed wild-type locomotion, consistent with rescue of at least part of the *unc-119* phenotype (Fig. 4). We established several independent transmitting lines, and used one of these, *edEx49*, to quantify the rescue of other *unc-119* defects. In summary, the *unc-119*::HRG4/HsUNC-119 transgene was able to rescue the locomotory, dauer formation, egg retention and pharyngeal pumping defects in a manner indistinguishable from a plasmid containing the intact *C. elegans* gene (Fig. 5). We conclude that HRG4/HsUNC-119 is a functionally conserved homolog of *C. elegans* UNC-119.

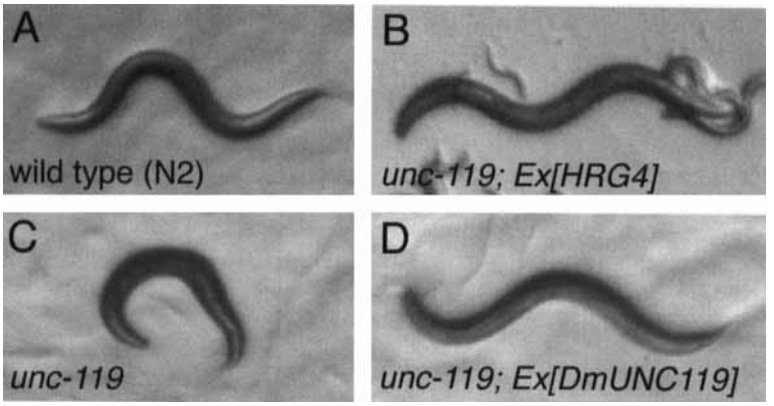
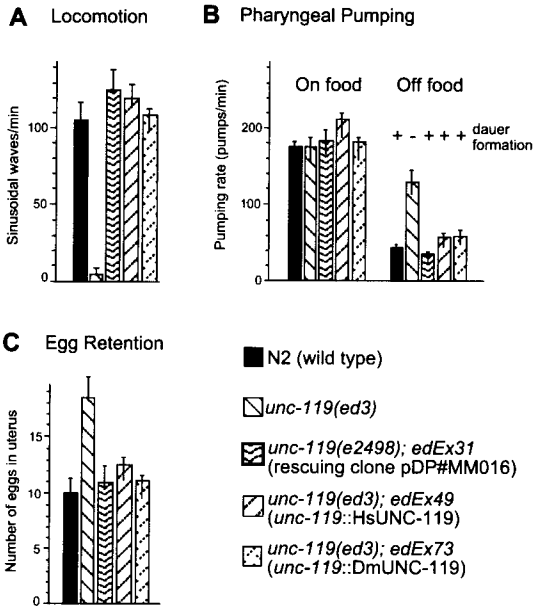



FIGURE 4 Still frames from video images of live animals on agar plates. Anterior is shown to the left, and all images are at the same scale. The adult in (A) is approximately 1 mm long. (A) Wild-type nematode showing typical dorsal–ventral sinusoidal waves propagated during motion. (B) *unc*-119(*e2498*) carrying the extrachromosomal array *edEx51*, which contains the *unc-119*::HRG4 fusion. (C) *unc-119*(*ed9*) mutant displaying ventral coiling and paralysis. (D) *unc-119*(*ed3*) carrying the array *edEx117*, which contains the *unc-119*::DmUNC-119 fusion. Note normal wave as in (A) and (B).

## Function of UNC-119 does not Require the Dissimilar Amino Terminus

The sequence divergence of the amino termini among the three UNC-119 homologs suggests that this region may not be crucial for function. To test the importance of the conserved, versus nonconserved regions, we tested the ability of portions of HRG4/HsUNC-119 to confer rescue of the locomotory and dauer defects (Fig. 5). Removal of the amino terminus of HRG4/HsUNC-119, up to the start of region A, did not affect the ability to rescue *unc-119* (plasmid pDP#MM150). Removal of a small part of the less-conserved region A (HRG4/HsUNC-119 amino acids 56–69) also had no effect, although this transgene still retains the amino terminus (pDP#MM147). In contrast, deletion of the carboxyl terminal, which contains the longest contiguous region of identity among the three homologs (sequence VMHNKADY) completely abolished function (plasmid pDP#MM148). We conclude that functional replacement of UNC-119 by HRG4/HsUNC-119 in *C. elegans* requires the carboxy terminus, but not the structurally divergent amino terminus.




**D** Rescue of *unc-119* by HRG4/HsUNC-119 fragments

| plasmid | portion of HRG4/HsUNC-119 | amino acids | loco-motion | dauer formation |
|---|---|---|---|---|
| pDP#MM103, pDP#MMUGF20 * | | 13-238 | + | + |
| pDP#MMUGF22 * | | 13-55 | – | – |
| pDP#MMUGF21 * | | 13-175 | – | – |
| pDP#MM147 | | 13-55 + 70-238 | + | + |
| pDP#MM148 | | 13-223 | – | – |
| pDP#MM150 | | 55-238 | + | + |
| HsUNC-119 | A | 1-240 | n.d | n.d. |

* fusion to GFP and *unc-54* 3'UTR

FIGURE 5 Ability of HRG4 and DmUNC-119 transgenes to rescue the *C. elegans unc-119* mutant phenotype. (A–C) Rescue of *unc-119* behavioral defects by HRG4/HsUNC-119 and DmUNC-119 transgenes. At least ten animals were scored for each trait. Data collection for panels (A–C) is described in Maduro and Pilgrim (1995). Lines above data bars denote standard error of the mean (SEM). (A) Locomotion measured in liquid (M9 buffer). (B) Pharyngeal pumping. '+' denotes ability to form dauer larvae, and '–' denotes a dauer formation defect. (C) Retention of eggs in adult hermaphrodites in the presence of food. (D) Rescue of locomotory and dauer formation defects by truncated HRG4 transgenes. Data collection is described in the text. The structure of the predicted full-length HRG4 protein is shown at the bottom for comparison; conserved regions A and B are shaded. 'amino acids' indicates the amino acids of HRG4 present in each transgene. Plasmids pDP#MM103, 147, 148 and 150 were injected into *unc-119(e2498)*, the GFP fusions pUGF20–22 were injected into *unc-119(ed3)*, and pDP#MM155 was injected into *unc-119(ed3); him-8(e1489)*. Plasmids pDP#MM147–155 were coinjected with pRF4 (*rol-6*) as transgenic markers.

To test whether the *Drosophila* gene could also confer rescue, we performed a similar experiment for DmUNC-119. The partial cDNA pDP#MM154, which starts at the beginning of region A (DmUNC-119 amino acid number 68) was used to make an *unc-119*::DmUNC-119 fusion gene, which was then injected into *unc-119(ed3)* homozygotes. As with the HsUNC-119 transgene, the fly gene can also confer rescue of the *unc-119* behavioral defects (Fig. 5). Therefore, the UNC-119 family of neural genes has been functionally conserved across metazoan evolution.

## DISCUSSION

The predicted UNC-119 protein has no obvious domains or motifs which give clues as to its molecular role in the cell. We have identified the *unc-119* homolog from the related nematode *C. briggsae*, and shown that it can replace the *C. elegans* gene under the control of its own promoter (Maduro and Pilgrim, 1996). However, at the amino acid level, the two gene products are approximately 90% identical across their lengths, precluding any strong conclusions about which regions of the proteins are likely to be important for function (Maduro and Pilgrim, 1996). Like the two *Caenorhabditis* genes, the vertebrate homologs are strongly related to each other: the human and rat UNC-119 homologs (HRG4 and MRG4, respectively) are 92% identical overall (Higashide *et al.*, 1996; 1998). While the two *Caenorhabditis* genes are expressed throughout the nematode nervous system, including sensory and locomotory neurons, the apparent retina-enhanced expression of the vertebrate homologs (Higashide *et al.*, 1996; Swanson *et al.*, 1998) initially suggested that the nematode and human genes, while structurally similar, may be diverged in function. Here, we have demonstrated the existence of an UNC-119 homolog in *Drosophila*, which is expressed in a large part of the fly nervous system; we have further shown that the human and fly genes are capable of rescuing the mutant phenotypes of *C. elegans unc-119* null mutants, and that this function depends only upon the presence of the conserved regions. Therefore, the UNC-119 homologs constitute a novel, functionally-related class of neural protein.

## Identification of a Neuronally-Expressed *Drosophila* UNC-119 Homolog

By whole-mount *in situ* hybridization, DmUNC-119 is expressed in the embryonic nervous system during development. As several ESTs were obtained from *Drosophila* libraries derived from adult brain, we can conclude that expression continues through adulthood in at least part of the fly nervous system. With respect to neural development, this overlaps the pattern seen for both the nematode and vertebrate UNC-119 homologs. Reporter *unc-119* fusions are expressed in the *C. elegans* nervous system, both embryonically (before neural differentiation), and continuing through adulthood (Maduro and Pilgrim, 1995; 1996). In the rat, a developmental Northern profile has shown expression starting at the onset of photoreceptor cell differentiation, and continuing through adulthood (Higashide *et al.*, 1996). This time course of gene expression strongly suggests a function for UNC-119 during neural development and in fully differentiated neurons. Indeed, there is strong evidence that in *C. elegans*, UNC-119 has a role during axonal outgrowth (W. Materi *et al.*, submitted).

The DmUNC-119 gene was mapped to a single site on the X chromosome (band 7A1); however, there are as yet no known DmUNC-119 mutations. Although *C. elegans unc-119* null mutants are fully viable (Maduro and Pilgrim, 1995), we anticipate that complete loss of *Drosophila* DmUNC-119 function results in an embryonic lethal phenotype. There are several precedents for conserved neural genes with relatively mild null phenotypes in *C. elegans*, and embryonic lethal phenotypes for their homologs in *Drosophila*. For example, homozygous null mutants in *unc-40*, which encodes the nematode homolog of the netrin receptor DCC, are viable (Chan *et al.*, 1996); in contrast, null alleles of the *Drosophila* DCC homolog *frazzled* are homozygous embryonic lethal (Kolodziej *et al.*, 1996). There are several embryonic lethal mutations found in *Drosophila* which map to the 7A1 region, any one of which could hence be a candidate DmUNC-119 mutant.

## Is UNC-119 Strictly Found in Retina?

The widespread neural expression of the invertebrate *unc-119* genes (i.e. *Drosophila* and *Caenorhabditis*) contrasts with the apparent retina-specific expression seen for the human and rat genes (Higashide *et al.*,

1996; Swanson *et al.*, 1998). It is not likely that expression of HRG4/HsUNC-119 is entirely retina-specific. ESTs corresponding to this gene have been found in many tissue preparations other than retina, including those in which contamination by photoreceptors is highly unlikely. One EST (AA101444) was found in a neuronal precursor NT2 cell line (Swanson *et al.*, 1998). In contrast, rhodopsin ESTs are only found in retina, and in preparations in which retina is included. Therefore, it is possible that HRG4/HsUNC-119 is more widely expressed throughout the vertebrate nervous system than the initial data would indicate. It is also possible that expression of vertebrate UNC-119 outside the retina occurs at a much lower abundance. Indeed, Northern analysis in the two *Caenorhabditis* species and in *Drosophila* suggests that the *unc-119* transcripts are of relatively low abundance (M.F. Maduro *et al.*, unpublished observations). Lastly, there may be other (as yet undiscovered) UNC-119 homologs found in other neural tissues.

There are vertebrate retina homologs of other neural *C. elegans* genes. The *tax-2* and *tax-4* gene products are structurally similar to components from cyclic nucleotide-gated channels found in vertebrate photoreceptors (Coburn and Bargmann, 1996; Komatsu *et al.*, 1996), and consistent with their predicted role, are expressed only in *C. elegans* sensory neurons (Coburn and Bargmann, 1996). Expression of *ceh-10*, a homolog of the vertebrate retina-specific homeobox genes Chx10 and Vsx-1, has been observed in a specific set of nematode neurons (Svendsen and McGhee, 1995).

## The UNC-119 Proteins Define a Novel, Functionally-conserved Protein Class Found in Divergent Metazoans

An alignment between the human, worm and fly protein sequences has shown approximately 45% and 70% identity across two extended domains, with regions of low identity at the amino termini and in the 'spacer' between the two domains. The existence of UNC-119 homologs in diverse phyla suggests that there are likely to be homologs in most, if not all, metazoans. For example, there are at least partial UNC-119 genes in the zebrafish *Danio rerio* (A. Manning and D.P., unpublished observations), and in a sipunculate worm (J. McGregor *et al.*, unpublished observations). The UNC-119 proteins define a new class, as they share no significant similarity with other known proteins.

We have shown here that HRG4/HsUNC-119 and DmUNC-119 can fully rescue the behavioral defects seen in *unc-119* null mutants. The function of these homologs depends upon the presence of only the conserved domains predicted by the structural alignment; removal of the most conserved portion, at the carboxyl terminus, abolishes the ability of HRG4/HsUNC-119 to rescue. Therefore, the UNC-119 genes are conserved in function as well as structure. It is likely, then, that these genes evolved from a common ancestral form; indeed, a common intron, occurring in the same position in the more conserved carboxyl domain, has been found in all of the UNC-119 homologs sequenced to date. Hence, the study of UNC-119 homologs may well reveal new insights into fundamentally conserved mechanisms of neural development and function.

## Possible UNC-119/HRG4 Functions

The ability of HRG4/HsUNC-119 and DmUNC-119 to compensate for loss of *unc-119* function in *C. elegans*, and the correlation between the rescuing ability and the presence of the most conserved regions of the proteins, demonstrates that these proteins have similar cellular functions, and that they likely interact with the same target molecules. It is still formally possible that this protein family does not function through protein–protein interactions, and instead has an intrinsic enzymatic or metabolic function, but the lack of any type of obvious structural motif or domain makes this less likely. The sequence conservation and the ability of divergent homologs to rescue suggest that proteins of the UNC-119 class interact with other well-conserved proteins. These factors could be ubiquitous, such as actin or microtubules, or they might be conserved proteins found only in neural cell types, such as UNC-33/CRMP-62 (Goshima *et al.*, 1995). This predicts that there are undiscovered *C. elegans* genes with vertebrate homologs expressed (at least) in the retina, and may reveal as-yet unknown shared mechanisms of neurogenesis.

### *Acknowledgments*

We thank the Caenorhabditis Genetics Center for providing some of the strains used in this work; Andy Fire for providing GFP expression

vectors; Ian Macdonald for the total human retina RNA; Gary Ritzel for the *Drosophila* DNA; Effie Woloshyn and David Nash for *in situ* mapping of DmUNC-119; Angela Manning for contributing unpublished results; Jiangwen Zhu for protocols on integrating extrachromosomal arrays; and Gina Broitman-Maduro for sequencing the DmUNC-119 partial cDNA clone. This work was supported by grants from the Medical Research Council of Canada and the Natural Sciences and Engineering Research Council.

## *References*

Bier, E. (1997) "Anti-neural-inhibition: a conserved mechanism for neural induction," *Cell*, **89**, 681–684.

Bloom, L. and Horvitz, H.R. (1997) "The *Caenorhabditis elegans* gene *unc-76* and its human homologs define a new gene family involved in axonal outgrowth and fasciculation," *Proc. Nat. Acad. Sci. USA*, **94**, 3414–3419.

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

Brewster, R. and Bodmer, R. (1996) "Cell lineage analysis of the *Drosophila* peripheral nervous system," *Dev. Gen.*, **18**, 50–63.

Chan, S.S.-Y., Zheng, H., Su, M.-W., Wilk, R., Killeen, M.T., Hedgecock, E.M. and Culotti, J.G. (1996) "UNC-40, a *C. elegans* homolog of DCC (Deleted in Colorectal Cancer), is required in motile cells responding to UNC-6 netrin cues," *Cell*, **87**, 187–195.

Coburn, C.M. and Bargmann, C.I. (1996) "A putative cyclic nucleotide-gated channel is required for sensory development and function in *C. elegans*," *Neuron*, **17**, 695–706.

Colavita, A., Krishna, S., Zheng, H., Padgett, R.W. and Culotti, J.G. (1998) "Pioneer axon guidance by UNC-129, a *C. elegans* TGF-β," *Science*, **281**, 706–709.

Fazeli, A., Dickinson, S.L., Hermiston, M.L., Tighe, R.V., Steen, R.G., Small, C.G., Stoeckli, E.T., Keino-Masu, K., Masu, M., Rayburn, H., Simons, J., Bronson, R.T., Gordon, J.I., Tessier-Lavigne, M. and Weinberg, R.A. (1997) "Phenotype of mice lacking functional Deleted in colorectal cancer (Dcc) gene," *Nature*, **386**, 796–804.

Goshima, Y., Nakamura, F., Strittmatter, P. and Streittmatter, S.M. (1995) "Collapsin-induced growth cone collapse mediated by an intracellular protein related to UNC-33," *Nature*, **376**, 509–514.

Hedgecock, E.M., Culotti, J.G. and Hall, D.H. (1990) "The *unc-5*, *unc-6*, and *unc-40* genes guide circumferential migrations of pioneer axons and mesodermal cells on the epidermis in *C. elegans*," *Neuron*, **2**, 61–85.

Hedrick, L., Cho, R.R., Fearon, E.R., Wu, T.C., Kinzler, K.W. and Vogelstein, B. (1994) "The DCC gene product in cellular differentiation and colorectal tumorigenesis," *Genes Dev.*, **8**, 1174–1183.

Higashide, T., McLaren, M.J. and Inana, G. (1998) "Localization of HRG4, a photoreceptor protein homologous to Unc-119, in ribbon synapse," *Invest. Ophthalmol. Vis. Sci.*, **39**(5), 690–698.

Higashide, T., Murakami, A., McLaren, M.J. and Inana, G. (1996) "Cloning of the cDNA for a novel photoreceptor protein," *J. Biol. Chem.*, **271**, 1797–1804.

Ishii, N., Wadsworth, W.G., Stern, B.D., Culotti, J.G. and Hedgecock, E.M. (1992) "UNC-6, a laminin-related protein, guides cell and pioneer axon migrations in *C. elegans*," *Neuron*, **9**, 873–881.

Keino-Masu, K., Masu, M., Hinck, L., Leonardo, E.D., Chan, S.S.-Y., Culotti, J.G. and Tessier-Lavigne, M. (1996) "Deleted in Colorectal Cancer (DCC) encodes a netrin receptor," *Cell*, **87**, 175–185.

Kennedy, T.E., Serafini, T., de la Torre, J.R. and Tessier-Lavigne, M. (1994) "Netrins are diffusible chemotropic factors for commissural axons in the embryonic spinal cord," *Cell*, **78**, 425–435.

Kolodziej, P.A., Timpe, L.C., Mitchell, K.J., Fried, S.R., Goodman, C.S., Jan, L.Y. and Jan, Y.N. (1996) "Frazzled encodes a *Drosophila* member of the DCC immunoglobulin subfamily and is required for CNS and motor axon guidance," *Cell*, **87**, 197–204.

Komatsu, H., Mori, I., Rhee, J.-S., Akaike, N. and Ohshima, Y. (1996) "Mutations in a cyclic nucleotide-gated channel lead to abnormal thermosensation and chemosensation in *C. elegans*," *Neuron*, **17**, 707–718.

Kunkel, T.A. (1985) "Rapid and specific site-specific mutagenesis without phenotypic selection," *Proc. Nat. Acad. Sci. USA*, **82**, 488–492.

Lewis, J.A., Wu, C.H., Berg, H. and Levine, J.H. (1980) "The genetics of levamisole resistance in the nematode *Caenorhabditis elegans*," *Genetics*, **95**, 905–928.

Maduro, M. and Pilgrim, D. (1995) "Identification and cloning of unc-119, a gene expressed in the *Caenorhabditis elegans* nervous system," *Genetics*, **141**, 977–988.

Maduro, M. and Pilgrim, D. (1996) "Conservation of function and expression of unc-119 from two *Caenorhabditis* species despite divergence of non-coding DNA," *Gene*, **183**, 77–85.

Mellerick, D.M. and Nirenberg, M. (1995) "Dorsal-ventral patterning genes restrict NK-2 homeobox gene expression to the ventral half of the central nervous system of *Drosophila* embryos," *Dev. Biol.*, **171**, 306–316.

Mello, C.C., Kramer, J.M., Stinchcomb, D. and Ambros, V. (1991) "Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences," *EMBO J.*, **10**, 3959–3970.

Patel, N. (1994) "Imaging neuronal subsets and other cell types in whole-mount *Drosophila* embryos and larvae using antibody probes," *Meth. Cell Biol.*, **44**, 445–505.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) "DNA sequencing with chain-terminating inhibitors," *Proc. Nat. Acad. Sci. USA*, **74**, 5463–5467.

Serafini, T., Kennedy, T.E., Galko, M.J., Mirzayan, C., Jessell, T.M. and Tessier-Lavigne, M. (1994) "The netrins define a family of axon outgrowth-promoting proteins homologous to *C. elegans* UNC-6," *Cell*, **78**, 409–424.

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

Svendsen, P.C. and McGhee, J.D. (1995) "The *C. elegans* neuronally expressed homeobox gene ceh-10 is closely related to genes expressed in the vertebrate eye," *Development*, **121**, 1253–1262.

Swanson, D.A., Chang, J.T., Campochiaro, P.A., Zack, D.J. and Valle, D. (1998) "Mammalian orthologs of *C. elegans unc-119* highly expressed in photoreceptors," *Invest. Ophthalmol. Vis. Sci.*, **39**(11), 2085–2094.

Tautz, D. and Pfeifle, C. (1989) "A non-radioactive *in situ* hybridisation method for the localisation of specific RNAs in Drosophila embryos reveals translational control of the segmentation gene hunchback," *Chromosoma*, **98**, 81–85.

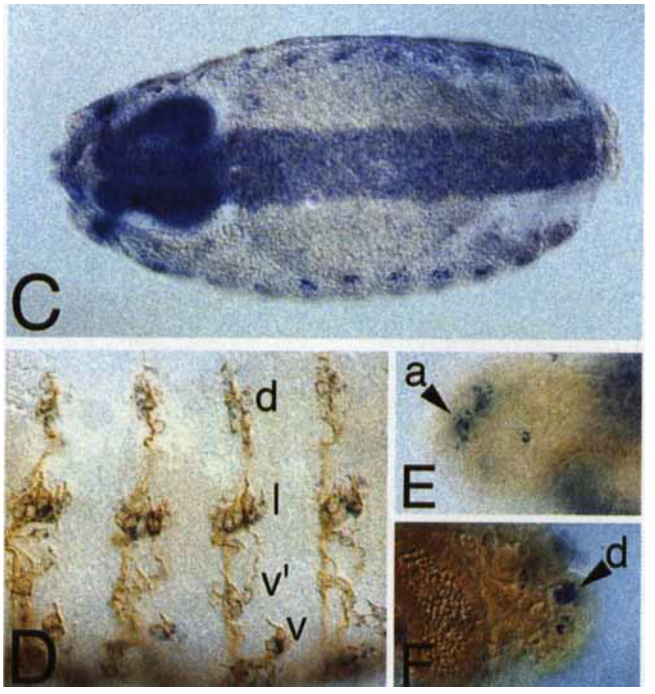
Thor, S. (1995) "The genetics of brain development: conserved programs in flies and mice," *Neuron*, **15**, 975–977.

Van Vactor, D., Sink, H., Fambrough, D., Tsoo, R. and Goodman, C.S. (1993) "Genes that control neuromuscular specificity in Drosophila," *Cell*, **73**, 1137–1153.

Wadsworth, W.G., Bhatt, H. and Hedgecock, E.M. (1996) "Neuroglia and pioneer neurons express UNC-6 to provide global and local netrin cues for guiding migrations in *C. elegans*," *Neuron*, **16**, 35–46.

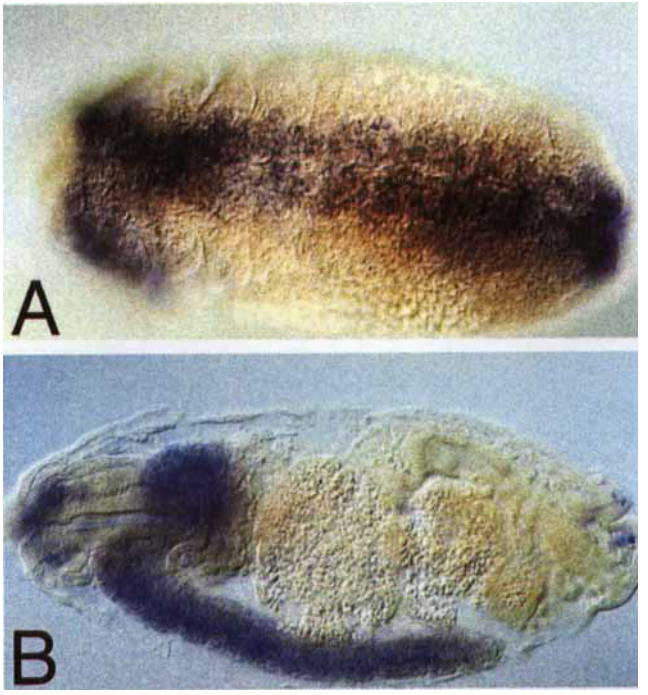
White, J.G., Southgate, E., Thomson, J.N. and Brenner, S. (1986) "The structure of the nervous system of *Caenorhabditis elegans*," *Phil. Trans. R. Soc. Lond. B Biol. Sci.*, **314**, 1–340.

Wood, W.B. (1988) The nematode *Caenorhabditis elegans*, W.B. Wood (Ed.) (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).




**Color Plate 1 (See page 201, Figure 3)**

Distribution of DmUNC-119 transcript in embryos. DmUNC-119 transcript accumulation is first detected in late stage 11 embryos, uniformly through the neuromeres of the central nervous system (A). A uniform and ubiquitous distribution through the nervous system is maintained for the remainder of embryogenesis, as seen in sagittal view at stage 16 (B) and frontal view at stage 13 (C). Neurons of the peripheral nervous system also express DmUNC-119. Lateral ectoderm of a stage 14 embryo double labeled for 22C10 (brown) and DmUNC-119 (blue) demonstrates extensive co-labeling in the dorsal (d), lateral (l) and ventral (v', v) sensory clusters. Neurons of the antennomaxillary complex (a in panel E) and the dorsal sensory cluster of the telson (d in panel F) also express DmUNC-119.




**Color Plate 1 (Continued)**