

Muscle and nerve-specific regulation of a novel NK-2 class homeodomain factor in *Caenorhabditis elegans*

Brian D. Harfe^{1,2} and Andrew Fire^{1,*}

¹Carnegie Institution of Washington, Department of Embryology, 115 W. University Parkway, Baltimore, MD 21210, USA

²The Johns Hopkins University, Graduate Program in Biology, Baltimore, MD 21218, USA

*Author for correspondence (e-mail: fire@mail1.ciwemb.edu)

Accepted 17 November 1997; published on WWW 13 January 1998

SUMMARY

We have identified a new *Caenorhabditis elegans* NK-2 class homeobox gene, designated *ceh-24*. Distinct *cis*-acting elements generate a complex neuronal and mesodermal expression pattern. A promoter-proximal enhancer mediates expression in a single pharyngeal muscle, the donut-shaped m8 cell at the posterior end of the pharynx. A second mesodermal enhancer is active in a set of eight nonstriated vulval muscles used in egg laying. Activation in the egg laying muscles requires an 'NdE-box' consensus motif (CATATG) which is related to, but distinct from, the standard E-box motif bound by the MyoD family of transcriptional activators. Ectodermal expression of *ceh-24* is limited to a subset of sublateral motor neurons in the head of the animal; this activity requires a *cis*-acting activator element that is distinct from the control elements

for pharyngeal and vulval muscle expression. Activation of *ceh-24* in each of the three cell types coincides with the onset of differentiation. Using a set of transposon-induced null mutations, we show that *ceh-24* is not essential for the formation of any of these cells. Although *ceh-24* mutants have no evident defects under laboratory conditions, the pattern of *ceh-24* activity is apparently important for Rhabditid nematodes: the related species *C. briggsae* contains a close homologue of *C. elegans ceh-24* including a highly conserved and functionally equivalent set of *cis*-acting control signals.

Key words: *ceh-24*, Vulval muscle, Pharyngeal muscle, Muscle, Nerve, Motor neuron, NK-2 homeodomain, NdE-box, *Caenorhabditis elegans*, *Caenorhabditis briggsae*

INTRODUCTION

In the past few years, many of the tools of molecular genetics have been used to extend our knowledge of the determination and differentiation of nonskeletal muscles. These muscles differ from skeletal muscles in their structure, in their contractile properties and in lacking expression of the myogenic helix-loop-helix (MyoD) family of transcription factors (Rudnicki and Jaenisch, 1995). A variety of regulatory factors have been proposed to coordinate non-skeletal muscle differentiation. These include members of the NK-2 family of homeodomain proteins (Harvey, 1996). In *Drosophila*, the NK homeodomain factors *tinman* and *bagpipe* are involved in the subdivision of the mesoderm and in the determination of cell fates in the dorsal mesoderm (Azpiazu and Frasch, 1993; Bate, 1993; Bodmer, 1993). The vertebrate gene *Nkx2-5* has been shown to express in cardiogenic progenitor cells and in mature cardiomyocytes (Lints et al., 1993; Komuro and Izumo, 1993). Mice containing a targeted disruption in *Nkx2-5* do not initiate heart looping morphogenesis and die as embryos (Lyons et al., 1995). Several other transcription factors including GATA-4/5/6 (Laverriere et al., 1994), MEF-2 (Yu et al., 1992), MHOx (Cserjesi et al., 1992) and HF-1b (Zhu et al., 1993) have been implicated in vertebrate cardiac muscle

development. Many of these regulatory components are members of structural families with roles in diverse muscle and non-muscle differentiation. In particular, NK homeodomains have been implicated in a wide variety of developmental processes.

The muscles of *C. elegans* can be grouped into three types: striated body wall muscles used in locomotion, single-sarcomere pharyngeal muscles used for feeding and a group of minor muscles responsible for egg laying and defecation (Waterston, 1988). These groups of muscle cells are distinguished by lineage as well as their location and cell morphology. An analysis of regulatory components responsible for myosin heavy chain expression in *C. elegans* pharyngeal muscles led to identification of an NK-2 class homeodomain factor, CEH-22 (Okkema and Fire, 1994). *ceh-22* was identified by the ability of its product to bind the enhancer of the pharyngeal-specific myosin heavy chain gene, *myo-2*. CEH-22 protein is found in a subset of pharyngeal muscles, prior to the detection of *myo-2* product in these cells (Okkema and Fire, 1994).

The diverse roles for NK-2 class homeobox genes in vertebrate and invertebrate mesodermal development led us to search (experimentally and using DNA sequence databases) for additional members of this family in *C. elegans*. Of three

additional NK-2 family members that were found, only one (designated *ceh-24*) exhibits mesodermal activity. In this paper, we present molecular and genetic characterization of *ceh-24*.

MATERIALS AND METHODS

Molecular biology and transgenic strain construction

To identify novel NK homeodomains, *C. elegans* genomic DNA was amplified using degenerate primers AF39 – GCTCTAGAYTNTTYTCNCANGCNCARGT and AF44 – GGAATTCYTTTRTANCKRTGRTTYTGRAACCA (IUPAC notation). Both cDNA and genomic phage libraries were screened with the PCR product. Four independent genomic phage inserts and one cDNA were obtained. The genomic clones were used to map *ceh-24* to chromosome V between *him-5* and *vmp-2* (Coulson et al., 1986; A. Coulson, personal communication). A *C. briggsae* *ceh-24* genomic clone was obtained from screening a Charon 4 genomic library provided by D. Baillie and T. Snutch (Snutch, 1984; Snutch et al., 1988).

Transgenic *C. elegans* were made as described (Mello et al., 1991). The plasmid pRF4 was used as a dominant selectable marker for transgenic animals (Mello et al., 1991).

Whole-mount RNA in situ hybridization

Mixed stage populations of *C. elegans* were processed for in situ hybridization as described by Seydoux and Fire (1995), except that S.T.F. Molecular Biology Grade fixative (Streck Laboratories) was used to fix specimens overnight in place of the 20 minute formaldehyde treatment. Digoxigenin-labeled probe was visualized with alkaline-phosphatase-labeled antibodies. After the addition of coloring reagent, slides were allowed to develop for >5 hours in the dark.

TC1 excision screen

A TC1 insertion in *ceh-24* was obtained from K. van der Linden and R. Plasterk. The insertion did not cause a phenotype. A PCR-based strategy was used to screen for imprecise excisions of the TC1 insertion as described by Plasterk (1995). Four independent deletion alleles of *ceh-24* were recovered.

To address the activity of the *ceh-24* promoter in a *ceh-24* mutant genetic background, we injected construct pBH29.173. This construct produces full *ceh-24* promoter expression in a wild-type background and an identical pattern in the *ceh-24* mutant *cc539*.

Fig. 1. NK-2 homeodomain comparison. Identical sequences are indicated by ‘-’.

The CEH-24 homeodomains in *C. elegans* and *C. briggsae* are identical. Sources for homeodomain sequences are: *C. briggsae* CEH-24 (this paper); TTF-1/Nkx-2.1 (Guazzi et al., 1990); *Dth-2* (Garcia-Fernández et al., 1993); *Nkx2-5/Csx* (Lints et al., 1993); Komuro and Izumo, 1993); *Nkx-2.2*, *Nkx2-3* (Price et al., 1992); XINK-2 (Saha et al., 1993); *Lox-10* (Nardelli-Haeffliger and Shankland, 1993); NK-2/*vnd*, *tin*/NK-4, *bag*/NK-3 (Kim and Nirenberg, 1989); CEH-22 (Okkema and Fire, 1994); EgHbx3 (Oliver et al., 1992); *ceh-27*, *ceh-28* and *ceh-29* are additional related genes that have been sequenced by the ongoing *C. elegans* genome sequencing consortium (Wilson et al., 1994; B. D. Harfe and A. Fire, unpublished results).

CEH-24	RRKRRVLFSSQ	AQVYELERRF	KQAKYLTAPE	REQLANSIRL	TPTQVKIWFQ	NHRYKCKRQE	
<i>C. briggsae</i> CEH-24	-----	-----	-----	-----	-----	-----	100%
Rat TTF-1/Nkx-2.1	-----	-----	--Q--S---	--H--SM-H-	-----	-----M--A	87
Planarian <i>Dth-2</i>	-----I-----	--I-----	--Q--S---	--H--L-N-	-----	-----S-Q	85
Mouse <i>Nkx2-5/Csx</i>	--P-----	-----	--QR--S---	--D--SVLK-	-S-----	-R-----R	80
Mouse <i>Nkx2-3</i>	--P-----	--F-----	--QR--S---	--H--S-LK-	-S-----	-R-----R	80
Mouse <i>Nkx-2.2</i>	K-----K	--T-----	R-QR--S---	--H--SL---	-----	-----M--AR	78
Xenopus X1NK-2	K-----K	--T-----	R-QR--S---	--H--SL---	-----	-----M--AR	78
Leech <i>Lox-10</i>	-----I-----	--I-----	R-Q--S---	--H--TF-G-	-----	-----T-KSK	78
<i>Drosophila vnd</i> /NK-2	K-----TK	--T-----	R-QR--S---	--H--SL---	-----	-----T--AQ	77
<i>C. elegans</i> CEH-22	K-----PK	--T-----	RSQ--S---	--A--MQ---	-----	-----T-KSH	77
<i>C. elegans</i> CEH-27	-----TP	Q--HV----	QINR--SAAD	--N--K--N-	SA-----	-----Q-----	72
<i>C. elegans</i> CEH-28	K--P----T-	H--N---E--	-KQR-V--T-	--E--QCLG-	-A-----	-R-----LA	66
<i>Drosophila tin</i> /NK-4	K--P-----	--L--C--	RLK---GA-	--II-QKLN-	SA-----	-R--S--GD	65
<i>Drosophila bag</i> /NK-3	KKRS AA--H	---F-----	A-QR--SG--	--SEM-K-L-	-E-----	-R--T--KQ	62
Flatworm EgHbx3	QS-----NK	F-ISQ--K--	RKQR----Q-	--QE--HT-G-	-----	--A--M--LF	62
<i>C. elegans</i> CEH-29	--A-TV--D	Q-LQG----	ESQR--ST--	--IE--ALN-	SE----T---	-R-M-H-KVV	53

RESULTS

A novel NK-2 class homeodomain

We designed a degenerate PCR protocol to identify additional members of the NK-2 class homeodomain family in *C. elegans*. The primers used for this protocol [see Materials and Methods] resulted in the identification of genomic sequence for a single novel gene, designated *ceh-24*. The gene is expressed: we obtained a single cDNA from approximately 1×10^6 clones in an embryonic cDNA library. Among characterized NK-2 class members, the CEH-24 homeodomain is most closely related to the vertebrate thyroid transcription factor TTF-1 (Fig. 1). A lower level of identity (77%) is seen with the *C. elegans* NK-2 homeodomain protein CEH-22 (Okkema and Fire, 1994).

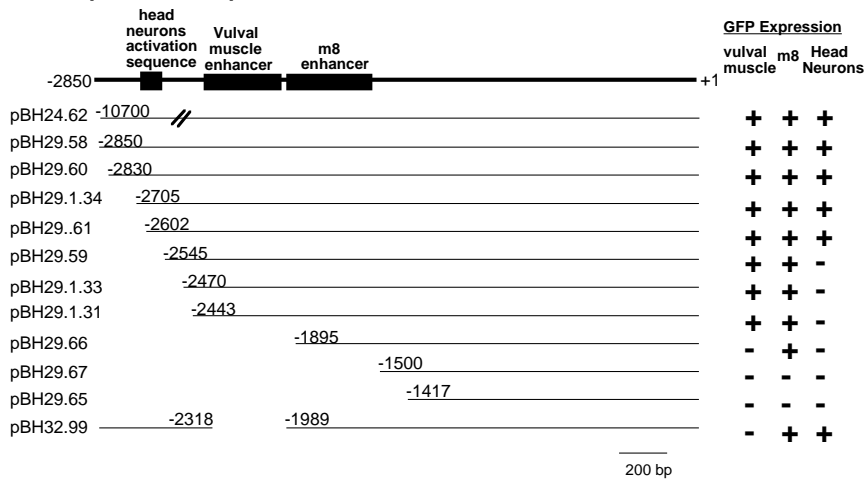
Outside of the homeodomain, CEH-24 lacks the ‘NK-2’ motif that has been found in other members of this family (Price et al., 1992; Lints et al., 1993; Harvey, 1997). Two striking acidic regions in CEH-24 (one just upstream of the homeodomain and another at the C terminus) suggest that the protein might function as a transcriptional activator. The possibility of control at the level of phosphorylation is suggested by an N-terminal region with eight serine residues at 2-3 amino acids intervals.

The *ceh-24* promoter is active in three distinct tissues

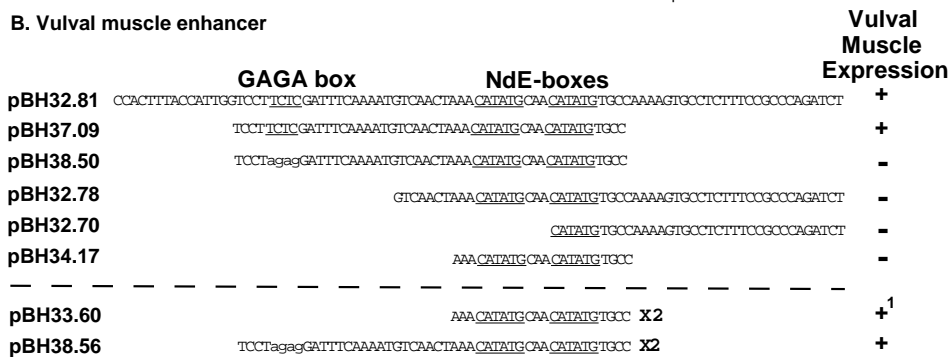
To determine the activity pattern of the *ceh-24* promoter, we constructed a series of *lacZ* and *gfp* translational fusions (Fig. 2). Additional reporter constructs which retained 5', internal, and 3' *ceh-24* sequences were produced by in-frame insertion of *gfp* directly into the *ceh-24* coding region. All of these chimeras have identical expression patterns. The constructs tested for expression cover a region extending 10.7 kb upstream and 0.7 kb downstream of the *ceh-24* coding region.

In adult animals, expression was seen in three distinct areas: in the eight vulval muscles, in 8-10 ventral neurons in the head and in the most posterior pharyngeal muscle cell, m8 (Fig. 3A-C).

Expression in vulval muscle cells was seen from the L4 larval stage onward, starting just before these cells began contractile activity. Vulval muscle cells are born in the late L3 larval stage (Sulston and Horvitz, 1977). Two sex myoblast cells present in

A. Complete *ceh-24* promoter

B. Vulval muscle enhancer



C. Pharyngeal muscle m8 enhancer

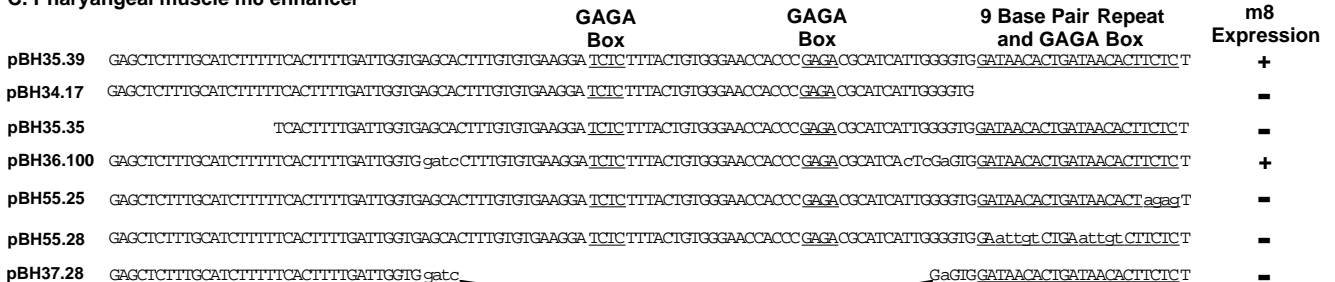


Fig. 2. Characterization of the *ceh-24* promoter and enhancers. (A) Activity of deletion constructs derived from a *ceh-24::gfp* translational fusion. All constructs contained the first 5 exons of *ceh-24*. pBH24.62 was assayed using a *lacZ* reporter. All other constructs were fused in frame to *gfp*. For each construct, GFP or LacZ expression was assayed in at least 20 independent F1 animals, and in at least one heritable line. Expression was scored in the vulval muscles, pharyngeal muscle m8, and in head neurons. Reporter-positive cells are indicated by a '+'; reporter negative cells are indicated by a '-'. Ectopic expression was occasionally seen in gut cells and rarely in anal muscles; this type of ectopic expression has been previously reported to result from juxtaposition of vector sequences (Krause et al. 1994). (B) Delineation of the *ceh-24* vulval muscle enhancer sequence. DNA sequences in the vicinity of the putative vulval muscle enhancer were assayed upstream of a truncated *pes-10* promoter fragment driving *lacZ* (pPD95.18, A.Fire, S. Xu, J. Ahnn and G. Seydoux, personal communication). The truncated *pes-10* promoter produces background expression in the pharynx and gut, but no expression in any other cell type. Ability of individual elements to activate the *pes-10* promoter in vulval muscles is indicated by '+'; lack of expression is indicated by '-'. All constructs were assayed in at least two lines. NdE-boxes and GAGA box are underlined. GAGA box mutations are indicated by lower case letters (pBH38.50 and pBH38.56). pBH33.60 and pBH38.56 each contain two tandem copies of the sequence shown (indicated by the annotation 'X2'). ¹Plasmid pBH33.60 is also expressed in muscles associated with the uterus and intestine, and in the anal depressor muscle. (C) Delineation of an enhancer which is specific to the pharyngeal muscle cell m8. Specific DNA sequences were assayed in front of a truncated *myo-2* promoter (pPD95.62, Okkema et al, 1993; Jantsch-Plunger and Fire, 1994). This promoter by itself does not produce expression in any cell type. Expression was scored in the m8 cell of the pharynx. A 115 bp fragment was sufficient to drive *myo-2* promoter activity in m8 (pBH35.39). In this 115 bp fragment there are three GAGA boxes (underlined) and a 9 base pair repeat containing two GATA sites (underlined). Deletion of the first 18 bp (pBH35.35), an internal fragment containing the two GAGA boxes (pBH37.28), the 9 base pair repeat (pBH55.28), or the third GAGA box (pBH55.25), abolished activity. An internal linker substitution of seven base pairs (lower case letters; pBH36.100) did not affect m8 expression.

the L3 larval stage are the precursors to the eight vulval and eight uterine muscles (Sulston and Horvitz, 1977). No *ceh-24* reporter activity was seen in the sex myoblasts or uterine muscles.

Expression in the pharynx was limited to a single cell, the donut-shaped muscle cell m8. This unique cell forms the rear of the pharynx. m8 expression was first seen in the embryo, beginning approximately 400 minutes after the first cell cleavage. This activity continued throughout the life of the animal.

C. elegans hermaphrodites contain 302 neurons (White et al., 1986; Chalfie and White, 1988). The precise identity of each neuron can, in principle, be determined by the location of its nucleus and the paths of its nerve processes (White et al., 1986). The *ceh-24* reporter construct was active in 8-10 neurons in the ventral head region of *C. elegans*. Using a *ceh-24::gfp* fusion construct (pBH28.101), the 8-10 ventral head neurons were observed to extend nerve processes into the nerve ring and sublateral nerve cords. Based on this observation, six of these neurons have been identified as SMDVL, SMDVR, SIBVL, SIBVR, SIBDL and SIBDR. The four remaining neurons with *ceh-24* promoter activity appear also to belong to the SI/SM class of sublateral neurons.

To localize *ceh-24* RNA transcripts in the embryo, we performed RNA in situ hybridization. *ceh-24* transcripts were first detected approximately 400 minutes after the first cell cleavage, and were seen in ventral head neurons (Fig. 3F-H). This correlates with the first detection of *ceh-24* reporter expression. RNA transcripts could be detected in these cells until hatching. We have not been able to obtain reproducible permeabilization of differentiated pharyngeal or vulval muscles. Difficulty in detecting mRNA in these cells occurs with several unrelated transcripts, likely reflecting a physical barrier to permeabilization of these tissues (G. Seydoux, B. D. Harfe and A. Fire, unpublished results).

Distinct control elements responsible for the *ceh-24* promoter activity pattern

To locate the control elements that are responsible for the observed *ceh-24* promoter activity pattern, we first carried out a unidirectional deletion analysis starting with a *ceh-24::gfp* translational fusion. Activity was analyzed using both an F₁ expression assay and in stable transgenic lines (Okkema et al., 1993).

The unidirectional deletion analysis revealed three distinct DNA elements specifying different aspects of *ceh-24* promoter activity (Fig. 2A). The critical elements reside between -2602 bp and -1500 bp upstream of the start of translation, in the order motor neuron element→vulval muscle element→m8 element→promoter.

Characterization of a vulval muscle enhancer sequence

The deletion analysis indicated a vulval muscle control element positioned between -2443 and -1989 relative to the translational start. To further narrow down this signal, we carried out enhancement assays using a vector based on the minimal promoter segment from *pes-10* (Seydoux and Fire, 1994). The crippled *pes-10* promoter in this vector is not intrinsically active, but can be activated in a variety of tissues by juxtaposition to diverse tissue-specific enhancers (B. D. Harfe, A. Fire, S. Xu and G. Seydoux, unpublished observations). A 452 bp sequence from the vulval enhancer

region of *ceh-24* activated the minimal *pes-10* promoter in all eight vulval muscle cells (Fig. 3D). This fragment was capable of enhancement in either orientation. Further deletions identified a core 48 bp sequence that could activate reporter expression in the eight vulval muscles (pBH37.09, Fig. 2B).

Within the core 48 bp vulval enhancer, there are two repeats of the sequence 'CATATG' and a consensus 'GAGA' site. GAGA sites in other systems have been shown to bind a factor with general ability to aid in enhancer function (Biggin and Tjian, 1988). In constructs with a single copy of the enhancer sequence, the GAGA element is necessary for activity (pBH38.50 and pBH32.78, Fig. 2B). The GAGA site is apparently not the major determinant of tissue specificity, since duplication of an enhancer fragment containing a mutated GAGA sequence can overcome the need for this site (pBH38.56, Fig. 2B).

Assays with duplicated *ceh-24* enhancer segments allowed us to define a 22 bp element which is sufficient to specify expression in vulval muscle. The 22 bp minimal vulval muscle enhancer contains two copies of the sequence 'AACATATG'. We will refer to the sequence 'CATATG' as an 'NdE-box' (named for the restriction enzyme *NdeI*, which recognizes this sequence).

Concatamers of the 22 bp NdE-box segment have a broader tissue specificity than the original *ceh-24* enhancer. In addition to the eight vulval muscles, expression was seen in many of the remaining single sarcomere body muscles: the two intestinal muscles, the anal depressor muscle and the eight uterine muscles (Fig. 3I-L).

Characterization of an enhancer active in posterior pharyngeal muscle m8

The unidirectional deletion analysis suggested a 395 bp region as a candidate for the m8 signal. This region was tested for its ability to activate a minimal (non-active) muscle promoter from *myo-2*. The truncated *myo-2* promoter used in these experiments has no activity by itself but can be activated in a variety of cell types when adjacent to enhancer sequences (Okkema et al., 1993; Jantsch-Plunger and Fire, 1994). A 117 bp fragment from *ceh-24* activated expression in the pharyngeal muscle cell m8 (Fig. 3E). Enhancement by this segment was not seen in any other cell.

Further mutational analysis indicated that the 117 bp m8 enhancer contains three distinct elements that are necessary for enhancer activity:

(1) Three GAGA boxes. m8 enhancer activity is lost following deletion of a 50 bp sequence containing two of these sites, or mutation of the third GAGA box (pBH37.28, pBH55.25, Fig. 2C).

(2) A twice-repeated 9 bp sequence (pBH55.28, Fig. 2C). This sequence contains two 'GATA' motifs.

(3) An AT-rich region at the promoter-distal end of the enhancer (pBH35.35, Fig. 2C).

Null mutations of *ceh-24* have no detectable phenotype

A transposon-based strategy has been used to create targeted deletion mutations in a variety of *C. elegans* genes (Plasterk, 1995). Four *ceh-24* deletion alleles, designated *cc539-cc542*, were isolated using a TC1 insertion in *ceh-24* obtained from K. van der Linden and R. Plasterk. All of these deletions remove the homeobox and the downstream acidic region (Fig. 4A). The largest, *cc539*, deletes the entire CEH-24 coding

region except the first 5 amino acids. Loss of the homeobox sequence in *cc539* and *cc542* was confirmed by probing a genomic Southern blot with a DNA fragment containing either the *ceh-24* homeobox or the entire *ceh-24* gene (Fig. 4B,C).

All four alleles are viable. No phenotypic differences from wild-type were evident. The largest deletion, *ceh-24(cc539)*, was analyzed further for the following properties.

(1) Growth – *ceh-24(cc539)* animals grow at all permissive temperatures at a rate similar to wild-type animals.

(2) Visible phenotypes – *ceh-24(cc539)* animals are not visibly dumpy (Dpy), uncoordinated (Unc) or egg-laying defective (Egl).

(3) Mating – A *ceh-24(cc539)* male stock mated with an efficiency similar to wild type.

(4) Chemotaxis – *ceh-24(cc539)* animals were found to chemotax normally to three volatile odorants (benzaldehyde, diacetyl and pyrazine) and one non-volatile odorant (Cl⁻ ions) (C. Bargmann, personal communication)

(5) Pharyngeal pumping – *ceh-24(cc539)* animals have a normal pharyngeal activity pattern (electropharyngeogram), with effective muscle activity and pumping (L. Avery, personal communication).

(6) VC neuron synapses – the VC neurons were visualized to synapse normally onto the vulval muscles using an *unc-4::gfp* construct (The *unc-4::gfp* strain was a gift from H. Hutter and D. Miller; *unc-4::gfp* is expressed in VC neurons, D. Miller personal communication; Pflugrad et al., 1997).

To address the possibility of autoregulatory phenomena, we introduced a *ceh-24::gfp* reporter construct into *ceh-24(cc539)* animals. The observed expression pattern was identical to that of the same DNA construct in wild-type animals. This experiment confirmed the presence in *ceh-24(cc539)* animals of morphologically normal vulval muscles, sublateral motor neurons and pharyngeal muscle m8. These cells were in their correct positions and their cell shapes appeared to be wild type. The paths of the head neuron processes appeared normal in the

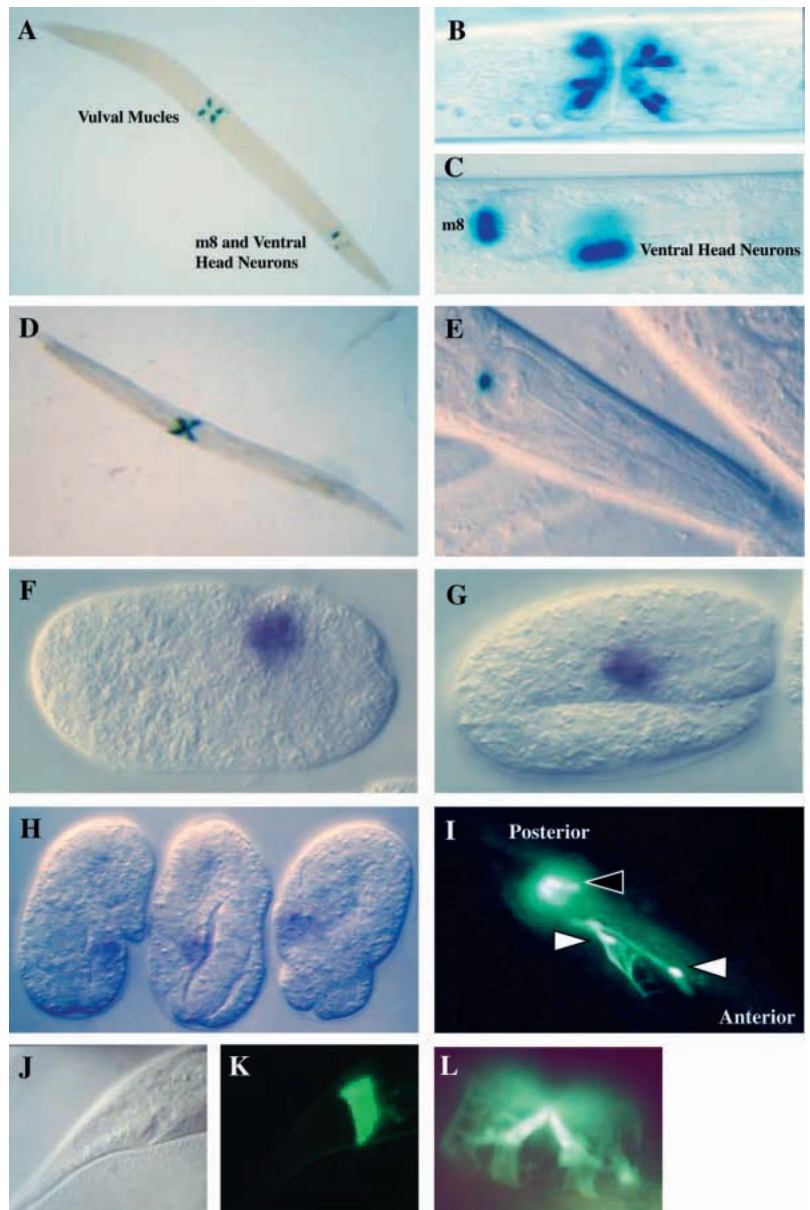


Fig. 3. Micrographs showing *ceh-24* reporter expression and RNA localization. Expression patterns for all constructs were assayed both in transgenic lines and in the first generation following microinjection. Identical expression patterns were observed for all constructs in both heritable lines and F1 assays. (A) *ceh-24::lacZ* expression in an adult animal. Expression is seen in the eight vulval muscles, eight ventral neurons in the head, and in the most posterior pharyngeal muscle, m8. (B) *ceh-24::lacZ* expression in the eight vulval muscles of an adult animal. (C) *ceh-24::lacZ* expression in the adult head. Activity is seen in pharyngeal muscle m8 and in 8-10 ventral head neurons. (D) The *ceh-24* vulval muscle enhancer driving *lacZ* expression from a truncated *pes-10* promoter in the eight vulval muscles. (E) The *ceh-24* m8 enhancer driving *lacZ* expression from a truncated *myo-2* promoter in the pharyngeal muscle m8. (F-H) In situ localization of *ceh-24* RNA. Antisense probes give a pattern with ventral head neurons showing strong expression in comma stage (F), two-fold (G), and three-fold (H) embryos. No signal was seen with a comparable 'sense' probe (not shown). (I-L) GFP expression driven by concatamerized NdE-boxes. GFP activity is both nuclear and cytoplasmic because of the tendency of the small GFP protein to leak out of the nucleus. (I) Anal depressor muscle (black arrowhead) and the two intestinal muscles (white arrowheads) are shown. Fibers of the intestinal muscles can be seen. (J,K) Higher magnification shows cellular outline of a GFP-positive anal depressor muscle. (L) NdE-box driven GFP expression is seen in vulval and uterine muscles. Scale: Adults are 1 μ m long and approximately 100 μ m wide; Embryos are 50 μ m long.

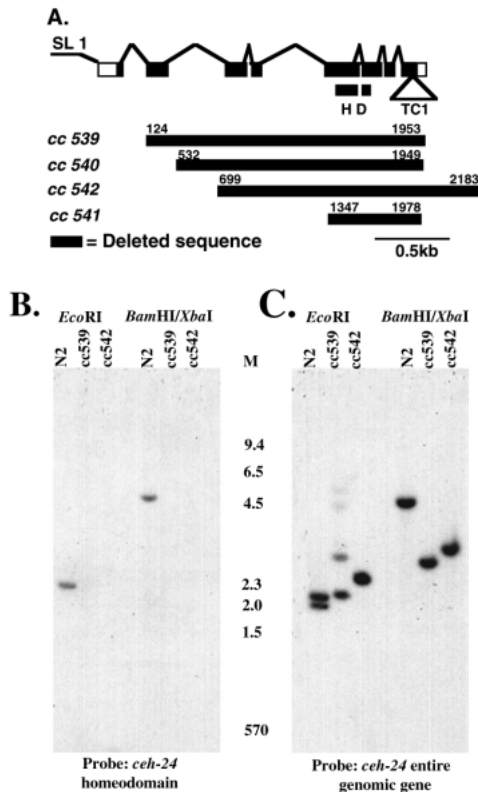


Fig. 4. Generation of null mutations in *che-24*. (A) Four transposon-catalyzed deletions were isolated (*cc539-cc541*; see Materials and Methods). Deletion junctions were amplified by PCR and sequenced. Solid lines indicate deleted DNA. *che-24* genomic organization is shown above the deletions. (B) Southern Blot analysis of deficiency DNA probed with a DNA fragment encoding the CEH-24 homeodomain. No hybridization to the CEH-24 homeodomain probe was detected in either of the two *che-24* deletion mutations. Marker sizes 'M' are in kb. (C) Southern Blot analysis of deficiency DNA probed with the entire *che-24* gene. There are *Eco*RI sites flanking *che-24* and one internal site. Digestion with *Eco*RI thus yields two wild-type bands. The internal *Eco*RI site is absent in both *che-24* deletions, resulting in the detection of just one band. A faint band in *Eco*RI digested *che-24(cc539)* DNA resulted from partial DNA digestion. A *Bam*HI/*Xba*I double digestion produced a wild-type 4.5 kb DNA fragment containing the entire *che-24* gene. A smaller *Bam*HI/*Xba*I band was observed in each *che-24* deletion.

mutant. GFP expression was at approximately the same levels in the mutant as in wild-type animals.

In addition to *che-24*, there is just one known *C. elegans* NK2 family member with mesodermal activity: CEH-22. CEH-22 protein is present in a set of pharyngeal muscle cells that do not show *che-24* promoter activity (Okkema and Fire, 1994). To determine whether CEH-24 was responsible for the lack of CEH-22 activity in m8, we injected a *che-22::lacZ* construct into *che-24(cc539)* animals. For these experiments, we used a *che-22::lacZ* fusion that mimics the native *che-22* expression pattern (pOK29.02; Okkema and Fire, 1994). No difference in *che-22* promoter activity was observed between wild-type and *che-24(cc539)* mutant animals (data not shown).

A mutation in *che-22* has recently been reported (Okkema et al., 1997). *che-22* mutant animals have a defective pharynx and are occasionally sterile. To investigate a possible role for CEH-

A. Protein alignments

C. elegans MSEKETPSPVLDVKKKEKNEETGIDEKSSSEDDCSKRSKVKSNPSKF
 C. briggsae MSEKESPSPOH-KKDEVVDDT-EQETKDSSEDDATKKMKIKEG-SKF

SVNSILSPLESILVRVQQQLLKMAASKSGTPTNAGVPGE--FPYGP
 TMNSILSPLESILARVQQQLLKMAAASQSGFGNGSGGNGANGFPY-PA

RLPGNYFAGFFPGYSGAQPNNWYNGNDPRFAAAAALLPSCIDPVRSAI
 RFATNCFSGPFFPGYNGSQGNWYNGNDARFAAAAALLPCAIDPVRSAI

NHQFSMSMSQRRKRRVLFSAQVYELERRFKQAKYLTAPEREQLAN
 NHQFSMS---QRRKRRVLFSAQVYELERRFKQAKYLTAPEREQLAN

SIRLTPTQVKIWFQNHRYKCKRQEKAKMSGLGHSEGGSSPPFDND
 SIRLTPTQVKIWFQNHRYKCKRQEKAKMSGLGHGGDSSPPFGDD

DDD-KYSIEMDDKDEEEEESEKPVLPKPSGVFGLPYP-PNAAAAA
 DDDKYSIEMDDKDEEEDG-AKPMKLSNGVFLPYPSPANAAAAA

AAAFNFPFAAQGNPAYMWR
 AAAFNFQFGAQQGNPAYFMRW

B. Vulval muscle enhancer

GAGA Box NdE-boxes

C. elegans TCCTTCCTGATTTTC-AAAATGTCACACTAAA CATATGCAA CATATGTGCC
 C. briggsae TCT-TCTCTATGTCTTAAATGTCACACTGAA CATATGCAA CATATGTGTC

C. m8 enhancer

GAGA Box Repeat

C. elegans GAGCTCTTTGCACTCTTTTTCACITTTT-GATTTGGTGAGCACTTTGHTGTAAGGA TCCTCT
 C. briggsae GAGCTCTTTCTCATATTTCTTTCACITTTTGTGATTTGGCGAGCACTTT-TGTGAAGGA TCCTCT

C. elegans TACTGTGGGAACCAACCCGAGAGCAATCATTTGGGGTGA TAACA CTGA TAACA CTCTCT
 C. briggsae TACTGTGGGAACCCCGACGCATATCATTTGGGGAGGC TAACA-A-TAACA CTCTCT

Fig. 5. Comparison between *C. elegans* and *C. briggsae* *che-24*. Conserved sequences are denoted with a '|'. (A) CEH-24 protein alignments. The homeodomains of both proteins are boxed in gray. Within the homeodomain, the proteins are identical; overall they are 79% identical. (B) Vulval enhancer sequences. The NdE-boxes are identical between *C. elegans* and *C. briggsae*; overall the vulval enhancers are 84% identical. The GAGA box, which is required for activity, is also conserved. The two NdE-boxes and single GAGA box are underlined. (C) m8 enhancer sequences. Underlined are two conserved repeated sequences and the lone conserved GAGA box. GenBank accession numbers are AF026056 (*che-24* coding sequence from *C. elegans*) and AF026057 (*che-24* genomic sequence from *C. briggsae*).

22 in *che-24* transcriptional repression, we examined expression of a *che-24::gfp* construct (pBH29.58, Fig. 2A) in *che-22* mutant animals. Homozygous *che-22* mutant animals had normal expression of the *che-24::gfp* construct (data not shown).

The genome of *C. briggsae* contains a very close relative of *C. elegans* *che-24*

C. briggsae is a Rhabditid nematode that contains many genes

of similar function and structure to *C. elegans* (Zucker-Aprison and Blumenthal, 1989; Heschl and Baillie, 1990; Kennedy et al., 1993; Krause et al., 1994). Two overlapping phage inserts were obtained from screening a *C. briggsae* genomic library with a *C. elegans ceh-24* homeobox containing fragment. Within the homeodomain, *C. elegans* CEH-24 and the *C. briggsae* homologue are identical (Fig. 1). Outside the homeodomain, the proteins are 79% identical (Fig. 5A) and five of six intron positions are conserved. Different NK-2 proteins generally do not have high homology outside the homeodomain. The presence of such high homology outside the homeodomain suggests that we have cloned the *C. briggsae ceh-24* homologue and not a more distantly related NK-2 class homeodomain.

Injection of a *C. elegans ceh-24::gfp* reporter construct into *C. briggsae* produced GFP expression in *C. briggsae* head neurons, pharyngeal muscle m8 and the eight vulval muscles (data not shown). This is identical to the expression pattern obtained upon injection of this construct into *C. elegans*. We constructed a *C. briggsae ceh-24* gene tagged with *gfp*. When injected into *C. elegans*, this construct produced expression identical to the *C. elegans ceh-24::gfp* fusion (data not shown).

Sequences upstream of *C. briggsae ceh-24* contain motifs identical to the *C. elegans ceh-24* Nde-boxes and closely related to the m8 enhancer (Fig. 5B,C). Outside the Nde-boxes and m8 enhancer, the upstream sequences from the two genes are not conserved. The *C. briggsae* homologue of the vulval enhancer is apparently a functional homologue; the *C. briggsae* sequence (in either orientation) can drive the *pes-10* minimal promoter in all eight vulval muscles in *C. elegans* (data not shown).

DISCUSSION

A novel NK-2 class homeodomain factor in *C. elegans*

NK-class homeodomain factors have been implicated in a variety of mesodermal and ectodermal differentiation processes in both invertebrate and vertebrate systems (Azpiazu and Frasch, 1993; Lyons et al., 1995; Okkema et al., 1997). Efforts in this laboratory, combined with the ongoing genome sequencing project, have identified four members of the NK-2 homeodomain family in *C. elegans*: CEH-22, CEH-24, CEH-27 and CEH-28 (Okkema and Fire, 1994; Wilson et al., 1994; this work; B. D. Harfe and A. Fire, unpublished data). Although the various activity patterns of NK-class homeodomains in vertebrate and invertebrate development suggest a diversity of regulatory roles, the conserved nature of the family suggests that certain underlying regulatory mechanisms (and possibly specific targets) could be conserved.

In this paper, we describe the characterization of the *C. elegans* NK class homeodomain factor, CEH-24. The CEH-24 homeodomain shows strong homology with the NK-2 subfamily, and in particular is more closely related to vertebrate thyroid transcription factor (TTF-1) than to any of the other known *C. elegans* NK-2 class homeodomains. In our analysis of *ceh-24*, we have focused on tissue activity pattern, on the regulation of expression and on the functional consequences of loss of function.

The *ceh-24* gene is active in specific subsets of neural and muscle tissue

We used a combination of in situ hybridization and reporter gene fusions to analyze the activity pattern of the *ceh-24* gene. The set of cells showing activity, a subset of 8-10 sublateral motor neurons and two types of non-striated muscles was somewhat surprising; these groups of cells had not been previously connected by any morphological or functional criteria.

The sublateral motor neurons had been primarily described in terms of their axonal outgrowth and synaptic connectivity. These cells are unique in that they send processes into the sublateral nerve cords, with synapses to body wall muscle cells (White et al., 1986; D. H. Hall and J. B. Rand, unpublished results). The role played by these cells in *C. elegans* behavior has not been elucidated. (Function has yet to be assigned to a large fraction of the neurons in *C. elegans*; e.g., Avery and Horvitz, 1989, Bargmann and Horvitz, 1991).

By contrast, the vulval and pharyngeal muscles have been assigned clear roles in the control of egg laying and in pumping of food into the intestine, respectively (Waterston, 1988). The proper physiological regulation of these contractile processes are critical to the fitness of the species. Optimal regulation presumably requires both the proper assembly of the contractile tissue and the ability to respond to a plethora of extrinsic signals modulating muscle contraction. These tissues might thus be expected to utilize a variety of regulatory components for specific aspects of cellular function.

ceh-24 and *ceh-22* define distinct identities within the pharyngeal musculature

The *ceh-24* activity pattern in the pharynx is complementary to that of the previously identified *C. elegans* NK-2 homeodomain factor CEH-22. CEH-22 protein is present in the pharyngeal muscles m1, m3, m4, m5 and m7 (Okkema and Fire, 1994), but is not present in the *ceh-24*-positive muscle cell m8. Neither gene is active in muscle cells m2 or m6. The activity pattern for *ceh-22* is unaffected in a *ceh-24* mutant; likewise the *ceh-24* activity pattern is unaffected in a *ceh-22* mutant. Our analysis thus indicates that the two non-overlapping expression patterns result from independent controls and not from an exclusionary mechanism.

GAGA and Nde-box elements combine to specify expression in vulval muscles

In analyzing requirements for the activity pattern of the *ceh-24* promoter, we found an underlying piecemeal organization, with each aspect of the expression pattern mediated by a distinct set of *cis*-acting signals. The vulval muscle enhancer region lies 2104 bp upstream of the mRNA start. The most prominent sequence features are a tandem pair of motifs with the core sequence 'CATATG'. As a mnemonic, we have designated these sequences 'Nde-boxes' (the *Nde*I restriction enzyme recognizes this sequence). Transcription factors of the bHLH family have been found to recognize sequences with a core 'CAnnTG', which would include the Nde-box. It should be noted that the myogenic bHLH family (including *C. elegans* CeMyoD and vertebrate MyoD, myogenin, Myf-5, and MRF4) recognize a distinct sequence, CAGCTG (Blackwell and Weintraub, 1990; Krause et al., 1997). Hence, these known

myogenic proteins would not be expected to readily bind the NdE-box consensus sequence. Consistent with this hypothesis, *C. elegans hhh-1* expression has not been detected in the non-striated vulval muscles (Krause et al., 1990). Instead, it seems likely that activity of the NdE-box element is mediated by a distinct factor, possibly a divergent member of the bHLH family.

In addition to the NdE-boxes, the vulval enhancer contains a 'GAGA' sequence that is required for enhancer function. Similar 'GAGA' sequences in *C. elegans* have been found to be required for activity of the *unc-54* enhancer (Jantsch-Plunger and Fire, 1994), and the *ceh-24* m8 enhancer (this work). These sites could serve as a binding site for a broadly active factor similar to the *Drosophila* GAGA-binding protein *Trithorax-like* (Farkas et al., 1994; Biggin and Tjian, 1988). This class of DNA-binding proteins have been proposed to create nucleosome-free areas of DNA, allowing the binding of additional transcription factors (Tsukiyama et al., 1994). No putative homologue of the *Drosophila* GAGA gene has yet been identified in *C. elegans*.

The ability to bypass the requirements for the GAGA element by duplicating the NdE-box portion of the enhancer indicates that the GAGA site is not the major determinant of vulval muscle activity. Interestingly, the activity of the dimerized 22-mer included additional non-striated muscles outside of the vulva: the intestinal muscles, uterine muscles and anal depressor muscle. This suggests that the GAGA and/or other sequences in the complete 81 bp vulval enhancer may act to limit expression in these additional muscles. The activity of the concatamerized NdE-box sequence in multiple sets of non-striated muscles could reflect either (1) a single factor binding the NdE-boxes and present in most of the minor muscles, or (2) several distinct NdE-binding factors present in different minor muscle subsets.

Is there a role for *ceh-24*?

Several circumstances could explain the lack of a detectable phenotype for a *ceh-24* null mutation.

(1) The function of CEH-24 may be quite subtle, so that a phenotype might not be evident with current techniques. For example, a decrease in a vulval muscle metabolic function or contractile strength would be difficult to quantify.

(2) CEH-24 might be serving a redundant function. We do not yet have any clear candidate for a redundant factor. The only previously characterized member of the *C. elegans* NK-2 class of homeodomains, CEH-22 (Okkema and Fire, 1994), is not present in cells with *ceh-24* activity. The *C. elegans* genome project has recently identified two additional members of the NK-2 family (*ceh-27* and *ceh-28*) and a single member of the closely related brain-specific-homeobox family (Jones and McGinnis, 1993; *ceh-29*). We have begun to characterize these three genes. Preliminary results indicate that none of the three genes are capable of expressing in *ceh-24*-positive cells (B. D. Harfe and A. Fire, unpublished results). In addition, *ceh-27*, *ceh-28*, and *ceh-29* reporter fusions do not have any ectopic expression in a *ceh-24* null mutation (B. D. Harfe and A. Fire, unpublished results). Genes that could compensate for CEH-24, but are not closely related to CEH-24, would not have been identified in this analysis.

(3) CEH-24 may function in a behavioral or developmental program that is not required under laboratory conditions. *C.*

elegans are grown on a plentiful food source under ideal conditions in the laboratory. In the wild, *C. elegans* is subject to many different environmental signals and stresses. Specific behavioral responses requiring CEH-24 might be important for fitness under these non-laboratory conditions.

The presence of a gene closely related to *C. elegans ceh-24* in the rhabditid nematode *C. briggsae* is indicative of an evolutionarily important function for CEH-24. The two genes encode identical homeodomains and yield 79% protein identity over their entire length. *ceh-24* vulval and m8 enhancers are also conserved between *C. elegans* and *C. briggsae*. This conservation is likely to be significant, since the two species are known to have diverged sufficiently to remove any non-selected similarity outside of coding regions (Prasad and Baillie, 1989).

The *C. elegans* and *C. briggsae* control regions showed apparently identical activation patterns when transformed into the heterologous species. This functional equivalence of control regions indicates that the *trans*-acting factors responsible for *ceh-24* activity are conserved both in their binding specificity and activity pattern. Conservation of regulatory properties in the two species argues strongly for a selective advantage conferred by the *ceh-24* expression pattern.

We thank K. van der Linden and R. Plasterk (funded by grant 5 RO1 RR10082-02 of the NIH-NCRR to R. P.) for isolating the Tc1 insertion strain used in this paper. We thank A. Antebi and H. Hutter for help in the identification of neuronal cells, and P. Okkema, M. Montgomery, S. Kostas, J. Hsieh, M. Hsu, H. Rienhoff, and W. Kelly for advice throughout this work. This work was supported by the National Institutes of Health (R01GM37706 to AZF, training grant T32GM07231 to The Johns Hopkins University Department of Biology). Sequences have been submitted to GenBank.

REFERENCES

- Avery, L. and Horvitz, H. R. (1989). Pharyngeal pumping continues after laser killing of the pharyngeal nervous system in *C. elegans*. *Neuron* **3**, 472-485.
- Azpiazu, N. and Frasch, M. (1993). *tinman* and *bagpipe*: two homeobox genes that determine cell fates in the dorsal mesoderm in *Drosophila*. *Genes Dev.* **7**, 1325-1340.
- Bargmann, C. I. and Horvitz, H. R. (1991). Control of larval development by chemosensory neurons in *Caenorhabditis elegans*. *Science* **251**, 1243-1246.
- Bate, M. (1993). The mesoderm and its derivatives. In *The Development of the Drosophila melanogaster*. (ed. M. Bate and A. Martinez-Arias), pp 1013-1090. Plainview, NY: Cold Spring Harbor Lab.
- Biggin, M. D. and Tjian, R. (1988). Transcription factors that activate the *Ultrathorax* promoter in developmentally staged extracts. *Cell* **53**, 699-711.
- Blackwell, T. K. and Weintraub, H. (1990). Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* **250**, 1104-1110.
- Bodmer, R. (1993). The gene *tinman* is required for specification of the heart and visceral muscles in *Drosophila*. *Development* **118**, 719-729.
- Chalfie, M. and White, J. (1988). The nervous system. In *The Nematode Caenorhabditis elegans* (ed. W. B. Wood), pp337-391. Cold Spring Harbor NY: Cold Spring Harbor Laboratory.
- Coulson, A. R., Sulston, J., Brenner, S. and Karn, J. (1986). Toward a physical map of the genome of the nematode *C. elegans*. *Proc. Natl. Acad. Sci. USA* **83**, 7821-7825.
- Cserjesi, P., Lilly, B., Bryson, L., Wang, Y., Sassoon, D. A., and Olson, E. N. (1992). MHox: a mesodermally restricted homeodomain protein that binds an essential site in the muscle creatine kinase enhancer. *Development* **115**, 1087-1101.
- Farkas, G., Gausz, J., Galloni, M., Reuter, G., Gyurkovics, H., Karch, F.

- (1994). The *Trithorax-like* gene encodes the *Drosophila* GAGA factor. *Nature* **371**, 806-808.
- García-Fernández, J., Bagnà, J., and Saló, E.** (1993). Genomic organization and expression of the planarian homeobox genes *Dth-1* and *Dth-2*. *Development* **118**, 241-253.
- Guazzi, S., Price, M., De Felice, M., Damante, G., Mattei, M. -G., and Di Lauro, R.** (1990). Thyroid nuclear factor 1 (TTF-1) contains a homeodomain and displays a novel DNA binding specificity. *EMBO J.* **9**, 3631-3639.
- Harvey, R.** (1996). NK-2 Homeobox genes and heart development. *Dev. Biol.* **178**, 203-216.
- Heschl, M. P. F. and Baillie, D. L.** (1990). Functional elements and domains inferred from sequence comparisons of a heat shock gene in two nematodes. *J. Mol. Evol.* **31**, 3-9.
- Jantsch-Plunger, V. and Fire, A.** (1994). Combinatorial structure of a body muscle-specific transcriptional enhancer in *Caenorhabditis elegans*. *J. Biol. Chem.* **269**, 27021-27028.
- Jones, B. and McGinnis, J. D.** (1993). A new *Drosophila* homeodomain gene, *bsh*, is expressed in a subset of brain cells during embryogenesis. *Development* **117**, 793-806.
- Kennedy, B. P., Aamodt, E. J., Allen, R. L., Chung, M. A., Heschl, M. P. F. and McGhee, J. D.** (1993). The gut esterase gene (*ges-1*) from the nematodes *Caenorhabditis elegans* and *Caenorhabditis briggsae*. *J. Mol. Biol.* **229**, 890-908.
- Kim, Y. and Nirenberg, M.** (1989). *Drosophila* NK-homeobox genes. *Proc. Natn. Acad. Sci. USA* **86**, 7716-7720.
- Komuro, I., and Izumo, S.** (1993). *Cx*: a murine homeobox-containing gene specifically expressed in the developing heart. *Proc. Natn. Acad. Sci. USA* **90**, 8145-8149.
- Krause, M., Fire, A., White Harrison, S., Priess, J., and Weintraub, H.** (1990). *CeMyoD* accumulation defines the body wall muscle cell fate during *C. elegans* embryogenesis. *Cell* **63**, 907-919.
- Krause, M., Harrison, S. W., Xu, S. Q., Chen, L. S. and Fire, A.** (1994). Elements regulating cell- and stage-specific expression of the *C. elegans* myoD family homologue *hhl-1*. *Dev. Biol.* **166**, 133-148.
- Krause, M., Park, Zhang, J., Yuan, J., Harfe, B., Zu, S., Greenwald, L., Cole, M., Paterson, B. and Fire, A.** (1997). A *C. elegans* E/Daughterless bHLH protein marks neuronal but not striated muscle development. *Development* **124**, 2179-2189.
- Laverriere, A. C., MacNeill, C., Mueller, C., Poelmann, R. E., Burch, J. B. E. and Evans, T.** (1994). GATA-4/5/6, a subfamily of three transcription factors transcribed in developing heart and gut. *J. Biol. Chem.* **269**, 23177-23184.
- Lints, T. J., Parsons, L. M., Hartley, L., Lyons, I. and Harvey, R. P.** (1993). *Nkx-2. 5*: a novel murine homeobox gene expressed in early heart progenitor cells and their myogenic descendants. *Development* **119**, 419-431.
- Lyons, I., Parsons, L. M., Hartley, L., Li, R., Andrews, J. E., Robb, R., and Harvey, R. P.** (1995). Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeobox gene *Nkx2-5*. *Genes Dev.* **9**, 1654-1666.
- 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**, 3959-70.
- Nardelli-Haeffiger, D. and Shankland, M.** (1993). *Lox10*, a member of the NK-2 homeobox gene class, is expressed in a segmental pattern in the endoderm and in the cephalic nervous system of the leech *Helobdella*. *Development* **118**, 877-892.
- Okkema, P. G., Harrison, S. W., Plunger, V., Aryana, A. and Fire, A.** (1993). Sequence requirements for myosin gene expression and regulation in *C. elegans*. *Genetics* **135**, 385-404.
- Okkema, P. G. and Fire, A.** (1994). The *Caenorhabditis elegans* NK-2 class homeoprotein CEH-22 is involved in combinatorial activation of gene expression in pharyngeal muscle. *Development* **120**, 2175-2186.
- Okkema, P. G., Ha, E., Haun, C., Chen, W. and Fire, A.** (1997). The *C. elegans* NK-2 homeobox gene *ceh-22* activates pharyngeal muscle gene expression in combination with *pha-1* and is required for normal pharyngeal development. *Development* **124**, 3965-3973.
- Oliver, G., Vispo, M., Mailhos, A., Martínez, C., Sosa-Pineda, B., Fielitz, W. and Ehrlich, R.** (1992). Homeoboxes in flatworms. *Gene* **121**, 337-342.
- Pflugrad, A., Meir, J., Barnes, T. and Miller, D.** (1997). The Groucho-like transcription factor UNC-37 functions with the neural specificity gene *unc-4* to govern motor neuron identity in *C. elegans*. *Development* **124**, 1699-1709.
- Plasterk, R. H. A.** (1995). Reverse genetics: From gene sequence to mutant worm. In *Caenorhabditis elegans: Modern Biological Analysis of an Organism* (ed. H. F. Epstein and D. C. Shakes), pp. 59-80. New York: Academic Press.
- Prasad, S. S. and Baillie, D. L.** (1989). Evolutionary conserved coding sequences in the *dpy-20* - *unc-22* region of *Caenorhabditis elegans*. *Genomics* **5**, 185-198.
- Price, M., Lazzaro, D., Pohl, T., Mattei, M. -G., Rütter, U., Olivo, J. -C., Duboule, D., and Di Lauro, R.** (1992). Regional expression of the homeobox gene *Nkx-2. 2* in the developing mammalian forebrain. *Neuron* **8**, 241-255.
- Rudnicki, M. A. and Jaenisch, R.** (1995). The MyoD family of transcription factors and skeletal myogenesis. *BioEssays* **17**, 203-209.
- Saha, M. S., Michel, R. B., Gulding, K. M., and Grainger, R. M.** (1993). A *Xenopus* homeobox gene defines dorsal-ventral domains in the developing brain. *Development* **118**, 193-202.
- Seydoux, G. and Fire, A.** (1994). Soma-germline asymmetry in the distributions of embryonic RNAs in *Caenorhabditis elegans*. *Development* **120**, 2823-2834.
- Seydoux, G. and Fire, A.** (1995). Whole mount in situ hybridization for the detection of RNA in *Caenorhabditis elegans* embryos. In *Caenorhabditis elegans: Modern Biological Analysis of an Organism* (ed. H. F. Epstein and D. C. Shakes), pp. 323-337. New York: Academic Press.
- Snutch, T. P.** (1984). Molecular and genetic analysis of the heat shock response of *Caenorhabditis elegans*. Ph. D. Thesis, Simon Fraser University, Burnaby, B. C.
- Snutch, T. P., Heschl, M. F. P. and Baillie, D. L.** (1988). The *Caenorhabditis elegans* *hasp70* gene family: A molecular genetic characterization. *Gene* **4**, 241-255.
- Sulston, J. E. and Horvitz, H. R.** (1977). Post-embryonic lineages of the nematode, *Caenorhabditis elegans*. *Dev. Biol.* **56**, 110-156.
- Tsukiyama, T., Becker, P. and Wu, C.** (1994). ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor. *Nature* **367**, 525-532.
- Yu, Y. -T., Breitbart, R. E., Smoot, L. B., Lee, Y., Mahdavi, V., and Nadal-Ginard, B.** (1992). Human myocyte-specific enhancer factor 2 comprises a group of tissue restricted MADS box transcription factors. *Genes Dev.* **6**, 1783-1798.
- Waterston, R. H.** (1988). Muscle. In *The Nematode Caenorhabditis elegans* (ed. W. B. Wood), pp. 281-335. Cold Spring Harbor NY: Cold Spring Harbor Laboratory.
- White, J. G., Southgate, E., Thomson, J. N. and Brenner, S.** (1986). The structure of the nervous system of *Caenorhabditis elegans*. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **314**, 1-340.
- Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J. et al.** (1994). 2. 2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. *Nature* **368**, 32-38.
- Zhu, H., Nguyen, V. T. B., Brown, A. B., Pourhosseini, A., Garcia, A. V., van Bilsen, M., Chien, K. R.** (1993). A novel, tissue-restricted zinc-finger protein (HF-1b) binds to the cardiac regulatory element (HF-1b/MEF-2) in the rat myosin light-chain2 gene. *Mol. Cell. Biol.* **13**, 4432-4444.
- Zucker-Aprison, E. and Blumenthal, T.** (1989). Potential regulatory elements of nematode vitellogenin genes revealed by interspecies sequence comparisons. *J. Mol. Evol.* **28**, 487-496.