Overlapping roles of two Hox genes and the exd ortholog ceh-20 in diversification of the C. elegans postembryonic mesoderm

Jun Liu and Andrew Fire*

Carnegie Institution of Washington, Department of Embryology, 115 West University Parkway, Baltimore, MD 21210, USA
*Author for correspondence (e-mail: fire@ciwemb.edu)

Accepted 20 September; published on WWW 2 November 2000

SUMMARY

Members of the Hox family of homeoproteins and their cofactors play a central role in pattern formation of all germ layers. During postembryonic development of C. elegans, non-gonadal mesoderm arises from a single mesoblast cell M. Starting in the first larval stage, M divides to produce 14 striated muscles, 16 non-striated muscles, and two non-muscle cells (coelomocytes). We investigated the role of the C. elegans Hox cluster and of the exd ortholog ceh-20 in patterning of the postembryonic mesoderm. By examining the M lineage and its differentiation products in different Hox mutant combinations, we found an essential but overlapping role for two of the Hox cluster genes, lin-39 and mab-5, in diversification of the postembryonic mesoderm. This role of the two Hox gene products required the CEH-20 cofactor. One target of these two Hox genes is the C. elegans twist ortholog hlh-8. Using both in vitro and in vivo assays, we demonstrated that twist is a direct target of Hox activation. We present evidence from mutant phenotypes that twist is not the only target for Hox genes in the M lineage: in particular we show that lin-39 mab-5 double mutants exhibit a more severe M lineage defect than the hlh-8 null mutant.

Key words: Hox, mab-5, lin-39, ceh-20, Caenorhabditis elegans, Mesoderm, twist

INTRODUCTION

Anterior-posterior patterns in animals arise from a combined consequence of cellular identities acquired in each of the three germ layers. Members of the Hox family of homeodomain proteins play a central role in this process (see Lewis, 1978; Lawrence and Morata, 1994; Krumlauf, 1994; Biggin and McGinnis, 1997 for review). Each of the three germ layers has a unique pattern of Hox expression, with the eventual pattern of tissues reflecting both autonomous Hox specification in each germ layer and interactions between germ layers (for example, see Bienz, 1994 for review). Two key paradoxes that have arisen in studies of Hox function concern (1) the relatively broad DNA binding specificity exhibited by Hox proteins in vitro, and (2) the ability of Hox proteins with very similar DNA binding properties in vitro to direct distinctive developmental patterns in vivo.

Recent studies have shown that the specificity of Hox factors is augmented in vivo by interaction with a distinctive group of homeodomain cofactors. These cofactors belong to the TALE (three amino acid loop extension) class of homeodomain proteins (Bürglin, 1997; Bürglin, 1998). One group of cofactors is represented by the Drosophila EXD and vertebrate PBX proteins (see Mann, 1995; Wilson and Desplan, 1995; Mann and Chan, 1996; Mann and Affolter, 1998 for review). Heterodimerization with EXD/PBX can greatly increase the specificity of DNA binding by Hox factors. In some cases the nature of the interaction with target genes is also changed, with the PBX/Hox dimers exhibiting a genetic activation function not seen with Hox alone (Pinsonneault et al., 1997; Li et al., 1999). Structural investigation of this regulatory system has recently begun with the analysis of Hox/EXD/DNA complex by X-Ray crystallography (Piper et al., 1999; Passner et al., 1999).

An additional level of regulation of Hox activity involves a second group of TALE class homeodomain proteins; these factors (HTH in Drosophila, MEIS and PREP1 in vertebrates) have been shown to regulate nuclear/cytoplasmic localization of EXD/PBX (Rieckhof et al., 1997; Kurant et al., 1998; Pai et al., 1998; Abu-Shaar et al., 1999; Berthelsen et al., 1999). It has also been shown that HTH/MEIS/PREP1 can form trimeric DNA binding protein complexes with Hox factors and EXD/PBX proteins (Berthelsen et al., 1998; Ryoo et al., 1999; Ferretti et al., 2000).

Although identification of the two families of Hox cofactors helps in explaining the specificity of Hox function in vivo, how Hox/cofactor combinations activate different regulatory networks to specify distinct cellular identity is not well understood. The availability of the entire lineage history of all cells in C. elegans (Sulston et al., 1983; Sulston and Horvitz, 1977) and its powerful genetics provide an approach to this question at a single-cell level.

The major C. elegans Hox complex contains four genes: ceh-13, lin-39, mab-5 and egl-5; these are orthologs of the Drosophila labial, Deformed/Sex combs reduced, Antennapedia/Ubx/Abd-A and Abd-B genes respectively.
(Costa et al., 1988; Schaller et al., 1990; Clark et al., 1993; Wang et al., 1993; Brunschwig et al., 1999). Two additional Abd-B homologs have recently been identified at a distinct chromosomal locus (Ruvkun and Hobert, 1998). Each Hox gene has a distinct expression pattern along the anterior-posterior axis of the animal. For the ectoderm, functions of the Hox factors have been extensively studied (see Bürglin and Ruvkun, 1993; Salser and Kenyon, 1994; Kenyon et al., 1997 for review). C. elegans also contains a single exd/pbx ortholog, ceh-20 (Bürglin, 1992). As with Drosophila exd mutants (Peifer and Wieschaus, 1990; Rauskolb et al., 1995), ceh-20 mutants share similar ectodermal phenotypes with multiple-loss-of-function Hox mutants (E. Chen, M. Robinson and M. Stern, personal communication). The situation in the mesoderm is less clear: although several of the Hox factors are known to be present in developing mesoderm (Wang et al., 1993; Salser, 1995; Ferreira et al., 1999), the roles for Hox and CEH-20 in mesodermal specification remain to be elucidated.

The mesoderm of C. elegans produces a variety of cell types including striated muscle, several types of non-striated muscle, and a limited set of non-muscle cells of diverse function. Cells from each of these classes arise both embryonically and postembryonically. Myogenesis in the embryo produces 81 striated bodywall muscles, 37 non-striated pharyngeal muscles, and four gut-associated enteric muscles and are used for locomotion. The 2 SMs are born in the posterior of L1 larvae, but migrate towards the anterior and reside at the vulval region. At mid L3 stage the SMs divide and give rise to 8 vulval muscles and 8 uterine muscles, which flank the gonad and are involved in egg laying. The coelomocytes are non-muscle mesodermal cells that behave as macrophages with a scavenger-like function (Chitwood and Chitwood, 1974; Fire et al., 1998a). The M lineage provides a valuable microcosm for genetic and experimental manipulations of mesodermal patterning since it is not essential for viability or for the overall body plan of the worm (Sulston and Horvitz, 1977).

Three of the Hox genes are known to be expressed in the M lineage. Although mab-5 expression occurs throughout the M lineage, mab-5 mutants show only a limited set of M lineage defects (Kenyon, 1986; Salser, 1995; Harfe et al., 1998b). lin-39 is expressed in sex myoblasts and their descendants (Wang et al., 1993; Liu and Fire, unpublished), while egl-5 expression has been observed in a subset of posterior muscles, and in the M mesoblast of males. No functional analysis of lin-39 or egl-5 in mesodermal specification has been reported. In this paper, we demonstrate an essential but redundant role for the Hox genes mab-5 and lin-39 in the diversification of the M lineage. We show that this role involves interactions of these factors with the C. elegans ortholog of twist.

MATERIALS AND METHODS

C. elegans strains

Strains were maintained and manipulated under standard conditions as described by Brenner (1974). Analyses were performed at 25°C, unless otherwise noted. The following strains were used in this work: mab-5(e1239) III (Kenyon, 1986), lin-39(n760) III (Clark et al., 1993), egl-5(n945) III (Wang et al., 1993), mab-5(e1239) egl-5(n945) III, lin-39(n760) mab-5(e1239) glf1; him-5 (e1490) III (Wang et al., 1993), lin-39(n760) mab-5(e1239) egl-5(n945) sma-3(e491) mab-5(e1239) egl-5(n945) III (gift from C. Kenyon), ceh-20(n2513) sma-3(e491) unc-32(e189) III (gift from K. Kornfeld).

Strains carrying integrated cell-type-specific reporter transgenes were used to facilitate identification of specific cell fates within the M lineage:

myo-3::gfp–PD4251(ccl154251) I, active in all bodywall muscles and vulval muscles (Fire et al., 1998b).

hls-8::gfp–PD4666(ayv56)x and PD4667(ayls7)IV, active in all undifferentiated cells in the M lineage (Harfe et al., 1998b).

egl-15::gfp–NH2447(ayls2) I, active in adult vm1 muscles (gift from C. Branda and M. Stern).

ndc-box::gfp–PD4655(ccl6555) II, active in eight vulval and eight uterine muscles (Harfe et al., 1998b).

myo-3::secreted gfp (secreted gfp)–GS1919(arls37)I and GS2077(arls39)x, used to visualize coelomocytes (gift from J. Fares and I. Greenwald).

Two integrated transgenic lines used in the heat-shock experiments were CF303(muls9) X for hs::mab-5 (Salser et al., 1993) and CF439(muls23) for hs::lin-39 (linkage group unknown; Hunter and Kenyon, 1995).

Lineage analysis was performed as described by Sulston and Horvitz (1977).

Plasmid constructs

mab-5 promoter constructs

7.5 kb of the mab-5 promoter sequence (~7485 to ~1) was amplified through long range PCR (Boehringer Mannheim) using genomic DNA as template. Primers JKL–192 and JKL–209 were used for the amplification, which resulted in the addition of unique NcoI and NotI sites at the ends of the PCR fragment. This fragment was cloned into PBS/KS+ and used as the mab-5 promoter in the following plasmids:

pJKL443.1: mab-5 promoter::mab-5 cDNA::unc-54 3'TR

pJKL419.5: mab-5 promoter::lin-39 cDNA::unc-54 3'TR

pJKL436.3: mab-5 promoter::egl-5 cDNA::unc-54 3'TR

pJKL439.5: mab-5 promoter::ceh-13 cDNA::unc-54 3'TR

pJKL418.4: mab-5 promoter::hls-8 cDNA::unc-54 3'TR

pJKL420.2: mab-5 promoter::hls-8 cDNA::smg suppressible 3'TR

pJKL481.9: mab-5 promoter::gfp::lacZ::unc-54 3'TR

The unc-54 3’TR is functional in all somatic tissues (Fire et al., 1990); the segment used for these constructs was derived from pPD96.85; the smg suppressible 3’UTR was derived by an extended coding region of let-5188 out of frame (Kelly et al., 1997). The GFP–β-gal fusion in pJKL481 contains a nuclear localization signal (derived from pPD107.94), thus resulting in nuclear localized GFP fluorescence. The cDNAs were used derived from the following plasmids:


Heat-shock constructs

The following constructs were used for ectopic production of the corresponding coding regions upon heat-shock treatment:

** RESULTS**

An essential role for Hox genes *mab-5* and *lin-39* in diversification of the postembryonic mesoderm

We first investigated the roles of Hox genes in patterning of the M lineage by examining the pattern of differentiated cells produced by the lineage in strains carrying different combinations of mutations in these genes.

Using a set of cell fate-specific reporter constructs (Fig. 1, Materials and Methods), we found that each single mutant retained the ability to carry out extensive M lineage diversification. The M lineage in null mutant *egl-5(n945)* or *lin-39(1760)* animals was normal, whereas limited M lineage defects were observed in null *mab-5(e1239)* mutants as reported previously: the M-derived coelomocytes and one or two bodywall muscles transformed to the sex myoblast fate (Harfe et al., 1998b). These results suggested that none of the three individual genes was essential for extensive M lineage diversification.

In contrast to the single mutants, a *lin-39(n1760) mab-5(e1239) egl-5(n945)* triple mutant showed a pronounced and severe defect in the M lineage. These animals produced no postembryonic coelomocytes, none of the body wall muscles normally derived from the M lineage and they lacked differentiated vulval and uterine muscles (Fig. 2). Instead of the 32 cells normally produced by the M lineage, we found a much smaller number of cells. These animals produced no myogenic (expressing *myo-3* and *egl-15* reporter constructs) but lacked the morphology of any normal muscle class.

The *lin-39* and *mab-5* genes, but not *egl-5*, appear to be the key factors in specifying the M lineage. Comparison of the *lin-39(n1760) mab-5(e1239)* double mutant with the *lin-39(n1760) mab-5(e1239) egl-5(n945)* triple mutant showed a similar range of phenotypes. All observed phenotypes were comparable in the two strains, although there was a somewhat higher fraction of animals in the double mutant showing the 1-4 residual myogenic products of the M lineage (Fig. 2). In similar assays for M lineage diversification, we saw no difference between *mab-5(e1239) egl-5(n945)* double mutants and *mab-5(e1239) single mutants (Fig. 2). These results indicate that *egl-5* does not play a critical role in patterning the hermaphrodite M lineage.

To further characterize M lineage products in the *lin-39(n1760) mab-5(e1239)* double mutant, we followed the cells continuously from hatching to early L2 stage by direct observation using Nomarski optics. In the majority of newly hatched mutant L1 larvae, M appeared at the correct position (although a fraction appeared ventralized). Since the final
position of M is a result of a posterior-directed cell migration during embryogenesis (Sulston et al., 1983), this indicates that the mutant M cells expressed aspects of their normal fates needed for migration during embryogenesis. In four of seven lineage mutant animals, M did not divide at all until early L2 stage; in the remaining three animals, M divided once. In wild type animals, M divides to produce 18 cells during this time. The M lineage defect in the double mutant is not due to a requirement for Hox activity in the M lineage. Experiments investigating the requirement for Hox activity in the M lineage. Experiments

**Functional equivalence of Hox factors MAB-5 and LIN-39 in the M lineage**

The synergism of lin-39 and mab-5 mutations suggested a partial redundancy between the two gene products in postembryonic mesoderm development. To ask whether this redundancy resulted from a functional equivalence between the two factors, we performed a series of experiments in which forced expression of one Hox family member was carried out in a genetic background lacking endogenous lin-39 and mab-5 activity. Rescue was assayed by direct analysis of differentiated descendants of the M lineage using specific M-lineage-derived cell type (first sex myoblasts and then sex muscles).
**Fig. 2.** M lineage phenotypes of Hox and ceh-20 mutants. For all panels in this and subsequent figures, lin-39(0), mab-5(0) and egl-5(0) represent lin-39(n1760), mab-5(e1239) and egl-5(n945) respectively, except where noted. All animals, except for the top one in C, are oriented with anterior to the left. (A-C) Vulval muscle phenotypes of (A) wild-type, (B) mab-5(0) and (C) lin-39(0) mab-5(0) animals visualized using egl-15::gfp, which labels vm1 vulval muscles. Note the extra vulval muscles in mab-5(0) animals (arrows in B), and the lack of vulval muscles (top animal in C) or the presence of one highly elongated egl-15::gfp-positive cell (arrow, bottom animal in C) in lin-39(0) mab-5(0) animals. (D,E) Bodywall muscle phenotypes of (D) wild-type and (E) lin-39(0) mab-5(0) animals visualized by myo-3::gfp, which labels all bodywall and vulval muscles. Note the decreased number and increased spacing of muscle nuclei in the double mutants. (F) Summary of M lineage phenotypes. The number of M-derived cells were assayed using cell type-specific GFP markers as described in Materials and Methods. For each genotype, >70 animals were examined for M-derived coelomocytes (secreted gfp). 7-15 animals were counted for M-derived bodywall muscles (myo-3::gfp), >90 animals were assayed for the number of vm1 vulval muscles (egl-15::gfp), and >80 animals were examined for the number of vulval and uterine muscles (NdE-box::gfp). The percentage represents the number of animals in the population examined that exhibited the phenotype indicated. In cases where the percentage is not noted, 100% of the animals exhibited that phenotype. *Although there were no sex muscles in the presumptive vulval region, there were 1-4 highly elongated or irregularly shaped, ‘sex-muscle-like’, cells in the posterior of some lin-39(0) mab-5(0), lin-39(0) mab-5(0) egl-5(0) or ceh-20(n2513) mutants. **The M lineage phenotypes of ceh-20(2513) animals were approximately 90% penetrant, approximately 10% of the animals exhibited partial formation of sex muscles (1-3 vm1s) in the presumptive vulval region. ‘ceh-20 (RNAi)’, progeny from hermaphrodite animals injected with double stranded ceh-20 RNA. After RNAi, all progeny were viable and showed 100% penetrance with respect to all M lineage defects. Moreover, all progeny lacked the ‘sex-muscle-like’ elongated cells which were present in lin-39(0) mab-5(0), lin-39(0) mab-5(0) egl-5(0) or ceh-20(n2513) animals.
in which a short pulse of heat in late embryogenesis was used to transiently produce LIN-39 or MAB-5 resulted in transient activation of the hlh-8 reporter in the L1 stage but not to the later production of vulval muscles expressing egl-15::gfp (data not shown). The ability of LIN-39 and MAB-5 proteins to function in the M lineage appeared specific: forced expression of two other Hox factors (CEH-13 or EGL-5) under the control of the 7 kb mab-5 promoter did not result in rescue of any M lineage defects in the lin-39(n1760) mab-5(e1239) double mutant (Fig. 3). These results suggest that LIN-39 and MAB-5 proteins share specific structural properties and/or activities that allow either protein (in the absence of the other) to direct diversification in the M lineage.

The C. elegans EXD/PBX ortholog CEH-20 acts as a Hox cofactor in M lineage diversification

In both Drosophila and vertebrates, Hox proteins function with a homeodomain protein cofactor, EXD/PBX, to regulate target gene expression (see Mann, 1995; Wilson and Desplan, 1995; Mann and Chan, 1996; Mann and Affolter, 1998 for review). We tested the role of the unique C. elegans exd/pbx ortholog ceh-20 (Bürglin, 1992) in patterning the M lineage using a strong loss-of-function mutation (n2513; E. Chen, M. Robinson and M. Stern, personal communication) and RNA-mediated interference (RNAi; Fire et al., 1998b). Both the n2513 mutation and ceh-20(RNAi) produced M lineage defects that were similar to lin-39(n1760) mab-5(e1239) double mutants. M lineage defects were approximately 90% penetrant in the n2513 mutant and 100% penetrant following ceh-20 RNAi (Fig. 2). Although M was present, all M-derived cells, including the 14 bodywall muscles, the 16 sex muscles and the 2 coelomocytes, were missing in the strongly affected animals.

The similarity in M lineage phenotypes between ceh-20 and
lin-39(n1760) mab-5(e1239) suggested several plausible models. One model that was readily tested was that CEH-20 might activate mab-5 and lin-39 gene expression in the M lineage. A mab-5::gfp-lacZ reporter construct (driven by the 7.5kb mab-5 promoter used in the rescue assays above) was used to examine mab-5 activity, while anti-LIN-39 antibody staining was used to assess LIN-39 localization. We found no change in mab-5 or lin-39 expression in the M lineage in ceh-20(n2513) animals (data not shown). In addition, forced expression of mab-5 or lin-39 using the heatshock promoter failed to rescue M lineage defects or hlh-8::gfp reporter activity in ceh-20(RNAi) animals (data not shown). These results suggest that the major contribution of CEH-20 is not as a regulator of mab-5 and lin-39 expression.

Intriguingly, we found one difference in phenotype between ceh-20(RNAi) and the lin-39(n1760) mab-5(e1239) double-null mutant (Fig. 2). The egl-15::gfp and Nde-box::gfp positive, 'sex muscle-like' cells present in lin-39(n1760) mab-5(e1239) double mutants were absent in ceh-20(RNAi) animals (n>100). Although we cannot rule out the possibility that treatment with dsRNA targeted against ceh-20 interferes with additional genes, there are no genes with sufficient homology to ceh-20 in the nearly complete genome sequence (C. elegans sequencing consortium, 1998) that could serve as common RNAi targets. Hence, the greater severity of the ceh-20(RNAi) phenotype suggests that there might be additional partners for ceh-20 in the M lineage. This is reminiscent of the situation in Drosophila, where EXD has been shown to act as a cofactor for non-Hox homeodomain proteins such as Engrailed (Peifer and Wieschaus, 1990; van Dijk and Murre, 1994).

A critical Hox/CEH-20 target site in the promoter of the C. elegans twist ortholog hlh-8

To further understand the role of mab-5, lin-39 and ceh-20 in mesodermal diversification, we investigated the relationship between these genes and hlh-8, a lineage-specific regulatory factor involved in patterning the M lineage. hlh-8 encodes the C. elegans twist ortholog and is active in undifferentiated cells throughout the M lineage (Harfe et al., 1998b). hlh-8 null mutants have variable defects in the M lineage, including alterations in early cleavage planes within the lineage, incomplete differentiation of sex muscles, variable numbers of M-derived bodywall muscles and lack of expression of two hlh-8 targets, egl-15 and ceh-24 (Corsi et al., 2000). Several previous observations with mab-5 had suggested that hlh-8 might act downstream of Hox function in the M lineage: (a) mab-5 mutants lack hlh-8 reporter expression in the early M lineage (this expression re-appears later in the lineage) and (b) forced expression of mab-5 can activate an hlh-8 reporter in muscle precursors (Harfe et al., 1998b).

We extended the connection between hlh-8 and Hox function by examining hlh-8 reporter activity in various Hox mutant combinations. We found the hlh-8 reporter to be completely off in the M lineage in lin-39(n1760) mab-5(e1239) double mutants (Fig. 4). These mutants retained normal hlh-8 expression in a set of non-muscle cells in the head. Loss of hlh-8 activity throughout the M lineage was similarly observed in the ceh-20(n2513) mutant and following ceh-20 RNAi (data not shown).

To test the hypothesis that lin-39 and mab-5 shared the ability to activate hlh-8 in the lineage, we forced expression of lin-39 using the mab-5 promoter or a heat shock promoter. As with MAB-5 (Harfe et al., 1998b), early embryonic expression of LIN-39 with a heat-shock promoter was sufficient to activate ectopic hlh-8 reporter expression in embryonic muscle precursors (data not shown, see Materials and Methods). Later forced expression of either lin-39 or mab-5, using the heat-shock promoter or the mab-5 promoter, could rescue the loss of hlh-8::gfp expression in lin-39(n1760) mab-5(e1239) mutants (Fig. 3). This rescue appeared to be specific to mab-5 and lin-39; forced expression of egl-5 or ceh-13 using the mab-5 promoter failed to rescue the loss of hlh-8 expression (Fig. 3). These results indicate a necessary and sufficient role in activating hlh-8 expression that can be fulfilled by either lin-39 or mab-5.

The hlh-8 promoter contains four candidate Hox binding sites. A 517 bp fragment of the hlh-8 promoter, which is sufficient for M-lineage specific expression of reporter genes, has four TAAT (or ATTA) sequences resembling core binding sites for Antennapedia type homeodomain proteins (Fig. 5A).
None of these sites matched the reported Hox/EXD consensus site (TGATNNATNN; Mann and Affolter, 1998). However, site 1 (TGAAAAATTA) contains a 3/4 match to the consensus half site (TGAT) for EXD factors (Mann and Affolter, 1998). A set of mutant promoters with clustered alterations in one or more of the Hox sites was created in vitro and tested in the animal using a GFP reporter. As shown in Fig. 5A, mutations in sites 2, 3 and 4 had little or no effect on the activity of the promoter. However, mutations in site 1, including those in the Hox or the CEH-20 half site, or both together, significantly reduced the level of reporter expression in the M lineage. These results suggested that site 1 is critical for *hlh-8* promoter activity in vivo.

A direct interaction between a physiologically critical site in the *hlh-8* promoter and LIN-39/CEH-20

To test if *hlh-8* was a direct target of Hox/CEH-20 dimers, recombinant LIN-39 and CEH-20 proteins were generated and purified from *E. coli*, and in vitro gel mobility shift assays were performed using sequences from the *hlh-8* promoter (as described in Materials and Methods, we were unable to generate full length MAB-5 proteins in *E. coli* or yeast). Oligonucleotides containing site 1 formed a complex with LIN-39 and CEH-20, generating a band which showed distinctly retarded gel mobility (Fig. 5B). Appearance of the putative ternary complex depended on both CEH-20 and LIN-39 proteins. LIN-39 alone produced an apparent binary

<table>
<thead>
<tr>
<th>Constructs (517bp promoter::gfp)</th>
<th>Promoter activity in the M lineage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wildtype</td>
<td>++</td>
</tr>
<tr>
<td>Site 1/Hox mutant</td>
<td>+/-</td>
</tr>
<tr>
<td>Site 1/CEH-20 mutant</td>
<td>+/-</td>
</tr>
<tr>
<td>Site 1/double mutant</td>
<td>+/-</td>
</tr>
<tr>
<td>Site 2 mutant</td>
<td>++</td>
</tr>
<tr>
<td>Site 3 mutant</td>
<td>++</td>
</tr>
<tr>
<td>Site 4 mutant</td>
<td>++</td>
</tr>
<tr>
<td>Sites 2 3 4 triple mutant</td>
<td>++</td>
</tr>
</tbody>
</table>

**A**

**B**

**C**

**D**

Oligonucleotides

<table>
<thead>
<tr>
<th>Oligonucleotides</th>
<th>CEH-20</th>
<th>Hox</th>
<th>Binding to LIN-39</th>
<th>LIN-39/CEH-20</th>
</tr>
</thead>
<tbody>
<tr>
<td>JKL-216</td>
<td>GTTTCGATAGTGGAAAAATTCACCGGGAAGAGAATTGAGACG</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>JKL-279</td>
<td>-----------------ctgct---------------------</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>JKL-281</td>
<td>-----------------tctagc---------------------</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>JKL-283</td>
<td>-----------------ctgct---------------------</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>JKL-285</td>
<td>-----------------ctgct---------------------</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>JKL-287</td>
<td>-----------------ctgct---------------------</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>JKL-289</td>
<td>-----------------ctgct---------------------</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>JKL-291</td>
<td>-----------------ctgct---------------------</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>JKL-293</td>
<td>-----------------ctgct---------------------</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>JKL-295</td>
<td>-----------------ctgct---------------------</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>JKL-297</td>
<td>-----------------ctgct---------------------</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>JKL-299</td>
<td>-----------------ctgct---------------------</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>JKL-218</td>
<td>GTTTCGATAGTGGAAAAATTCACCGGGAAGAGAATTGAGACG</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td><strong>EXD</strong></td>
<td><strong>Hox</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JKL-248</td>
<td>ACCCGCGATGCTAATGACCAGACCAGGAT</td>
<td>+</td>
<td>++</td>
<td></td>
</tr>
<tr>
<td>JKL-250</td>
<td>-----------------ctgct---------------------</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>
complex with site 1 oligonucleotides, while no complex was produce with CEH-20 alone. Analysis of sequences required for formation of the site 1/CEH-20/LIN-39 complex (Fig. 5B) demonstrated involvement of the putative CEH-20 site (TGAA) (Fig. 5B; JKL-295, JKL-293). The putative Hox site (aaATTAA) was also required for formation of the ternary complex, and for the binding of LIN-39 protein to site 1 as a binary complex (Fig. 5B; JKL-287, JKL-289).

Comparison of site 1 with a consensus binding site for HOX/EXD dimers (TGATNNATNN; Mann and Affolter, 1998) showed a single base pair difference in the EXD half site (TGAA instead of TGAT). The difference seems unlikely to reflect divergence in HOX/EXD recognition sequence between nematodes and other systems, since a canonical ANTP/EXD binding site (Knoepfler et al., 1996) formed a ternary complex efficiently with LIN-39 and CEH-20 (Fig. 5C). The greater affinity of LIN-39/CEH-20 to the canonical ANTP/EXD binding site than to site 1 may reflect differences in core sequence or in the four degenerate bases, which could alter the relative spacing or orientation of CEH-20 and Hox sites.

Interestingly, we found a class of mutation in the Site 1 Hox binding site (JKL-218 and possibly JKL-291) that retained Hox binding but lost the ability to form a ternary complex. These mutations may shift the position of the Hox site on the DNA (i.e. creating a new Hox site), or may alter the geometry of the Hox/DNA interaction so as to prevent formation of the ternary complex. In the case of JKL-218, the mutated site had lost the ability to function in vivo as part of the promoter (Fig. 5A). A similar situation was observed with sites 2 and 3 in the hlx-8 promoter; these two sites bound to LIN-39 in vitro but showed no ternary complex formation with LIN-39/CEH-20 (data not shown). Like the JKL-218 mutant of site 1, the natural sites 2 and 3 showed no evident contribution to in vivo activity of the promoter (Fig. 5A). We can make no conclusion concerning the significance of LIN-39:DNA complexes lacking CEH-20 that are formed in vitro: these complexes might fail completely to form in vivo, they might form only transiently, or they might persist but fail to activate gene expression in the M lineage. At least for the hlx-8 promoter, our data support a model in which a ternary complex of LIN-39 with CEH-20 on a defined site (site 1) is critical for activation of gene expression in the M lineage.

Fig. 5. The hlx-8 promoter contains a critical Hox/CEH-20 target site. (A) Wild-type and mutant hlx-8 promoter::gfp constructs and their M lineage promoter activity assayed in vivo. In the diagrams of promoter constructs, only the four putative Hox binding sites are highlighted. Core sequences are in upper case; mutations in lower case. Promoter activity was assayed by examining GFP in at least seven independent transgenic lines: +++, bright M lineage GFP in >95% of animals; ++, faint GFP signals in a low fraction of progeny (3/35 for the CEH-20 half site mutant, 2/100 for the HOX half-site mutant, and 0/100 for the double mutant). Quantitative measurement of fluorescence indicated that the faint signals in the few positive animals from the CEH-20 and HOX half-site mutants were approximately 40-fold lower than signals from the wild-type hlx-8 promoter. (B) Gel mobility shift assays using purified LIN-39 and CEH-20 proteins and dsDNA oligonucleotides (sequences shown in D). 1x and 10x represent 50 ng and 500 ng of purified proteins used respectively. A more rapidly migrating complex (arrowhead) required only LIN-39 protein and sequences in the HOX half site (lanes 14, 18, 26, 30); a more slowly migrating complex (arrow) required CEH-20 and LIN-39 proteins and sequences in the HOX and CEH-20 half sites (lanes 16, 20; 28, 32). (C) Mobility shifts with oligonucleotides with a canonical ANTP/EXD composite site (JKL-248; lanes 1,3,5,7,9,11) and an EXD half site mutant (JKL-250; lanes 2,4,6,8,10,12). A more rapidly migrating complex (arrowhead) was dependent on LIN-39 protein but independent of CEH-20 and the EXD binding site; a more slowly migrating complex (arrow) was dependent on both LIN-39 and CEH-20, and on the EXD half site. Under the conditions of this assay, the ANTP/EXD composite site was apparently more effective than hlx-8 site 1 in complexing with CEH-20/LIN-39, as evidenced by the efficiency of the band shift and the requirement for a lower concentration of CEH-20 (compare lane 4 in B and lane 7 in C). (D) A summary of mobility shift assays. + and – represent presence or absence of a mobility shift (−/+), in some experiments we observed a faint band corresponding to an apparent ternary complex between JKL-291 and LIN-39/CEH-20). Oligonucleotide JKL-216 covered site 1 of the hlx-8 promoter (core sequences underlined). JKL-279 to JKL-299 were derived from JKL-216: mutated residues are shown in lower case with dotted lines indicating un-mutated residues. JKL-248 contained a canonical EXD/EXD site; JKL-250 was a derivative of JKL-248 with the EXD half site mutated. A StyI linker was present on each double stranded oligonucleotide shown in the figure; this linker had no evident effect on binding, since equivalent results were obtained using a wild type site 1 oligonucleotide without the linker (JKL-300: GTTTCGATAGTGAAAATAATTACCCCGGAAA).
ceh-20 mutants are intriguing. These defects did not appear to be a result of homeotic transformation of the fate of M or its descendants. Instead, the mutants exhibit either (1) a loss of all differentiated M-lineage descendants or (2) the precocious production of abnormal mesodermal fates with certain properties of later M lineage products. The precocious appearance of large cells that exhibit SM- and sex muscle-like characteristics suggests that this program might be a default state of M in the absence of Hox function.

The shared role of MAB-5 and LIN-39 in the M lineage appeared to be specific to these two Hox factors. First, forced expression of either lin-39 or mab-5, but not of the neighboring Hox genes ceh-13 and egl-5, was sufficient to activate ectopic expression of M lineage reporters. Second, egl-5 mutants (which are viable either alone or in combination with lin-39 and mab-5 mutants) had no M lineage defects on their own and showed no synergistic effects with lin-39 and mab-5.

Mesodermal roles of Hox and exd genes have also been shown in Drosophila. In the visceral mesoderm, Ubx and abd-A are involved in morphogenesis of the midgut (see Bienz, 1994; Frasch and Nguyen, 1999 for review). In this case, a few targets for Hox genes have been described: Ubx in the visceral mesoderm is directly required with an EXD cofactor for activating expression of the signaling molecule dpp (Capovilla et al., 1994; Chan et al., 1994). In the somatic mesoderm, Ubx and abd-A can each promote the formation of specific sets of muscle precursors (Greig and Akam, 1993; Michelson, 1994). None of the Drosophila Hox mutants or combinations that have been analyzed show as drastic an effect on postembryonic mesoderm as was seen with the lin-39 mab-5 double mutant in C. elegans. This apparent discrepancy may reflect a fundamental difference between the biological systems; alternatively, a more drastic postembryonic requirement for Hox factors in the Drosophila mesoderm might have been missed due to the embryonic lethality of multiple-Hox mutants.

Functional equivalence of mab-5 and lin-39 in the M lineage

Our rescue experiments suggested partially overlapping roles for mab-5 and lin-39 in the M lineage. The modest M-lineage defects seen in mab-5 single mutants, compared with the lack of any M-lineage defects in lin-39 single mutants suggest that under normal circumstances the contribution of mab-5 may be somewhat more substantial at early time points. One conceivable explanation for the ‘either/or’ requirement would involve cross-regulation between Hox genes. In particular, we have tested the possibility that lin-39 expression in the M lineage only occurs in the absence of functional mab-5. This is apparently not the case, as mab-5 mutants show an apparently normal pattern of M lineage staining with antibodies to LIN-39 (date not shown).

Several types of interactions between lin-39 and mab-5 activities in determining cell fate have been reported. In a subset of Pn.aap cells that normally express both lin-39 and mab-5, the lin-39 activity is dominant, preventing mab-5 from functioning in these cells (Salser et al., 1993; Clark et al., 1993). A distinct interaction is seen in male Pn.p cells, where lin-39 and mab-5 are both expressed and act combinatorially to specify a fate that is different from that specified by either alone (Salser et al., 1993; Wang et al., 1993). A third situation (Clandinin et al., 1997; Maloof and Kenyon, 1998) is seen in hermaphrodite vulval precursor cells, for which the loss of Hox (lin-39) activity after specification results in a failure to differentiate; in this lineage, lin-39 and mab-5 activities have the capability to promote distinct and non-overlapping consequences in terms of cell fate. The functional and simultaneous requirement in the M lineage for either mab-5 or lin-39 function represents a further degree of freedom in using these genes to build an organism.

The highly conserved structure of Hox factors is consistent with a view that these genes have evolved by duplication of a single precursor gene (Bürglin, 1994). Under these circumstances, it is not surprising that certain roles for Hox factors would still be maintained as shared (or redundant) between several genes in the cluster (for example, Michelson, 1994; Greig and Akam, 1995; Casares et al., 1996; Favier et al., 1996; Barrow and Capecci, 1999). While the individual genes might have acquired position-specific roles based on their acquisition of intricate patterns of expression, it is certainly conceivable that the entire family (or a large subset) will have maintained a shared role equivalent to that of the ancestral (and unique) Hox factor. While the role of that factor will remain a mystery, the appearance of Hox factors in the developing embryo just prior to the start of differentiation suggests that the ancestral factor could have played a role in developmental timing, perhaps modulating the start of differentiation in a subset of cells.

The C. elegans twist ortholog hlh-8 is a direct and critical target of Hox genes and ceh-20 in the postembryonic M lineage

Our studies of the function of mab-5, lin-39 and ceh-20 in patterning of the postembryonic mesoderm led to the identification of a direct target for these genes, the C. elegans twist ortholog hlh-8. We identified a critical site in the hlh-8 promoter that is a binding site for the LIN-39/CEH-20 protein complex. The similarity between core binding sequences for Drosophila ANTP and DFD proteins in vitro (Ekker et al., 1994), and the functional equivalence of mab-5 and lin-39 in activating hlh-8 expression in the M lineage, strongly suggest that this site is also a binding site for MAB-5/CEH-20.

Although hlh-8 is a target for Hox/CEH-20 function in the M lineage, it is not the only such target. Several indirect observations demonstrate the existence of additional targets. One line of evidence comes from the observation that forced expression of hlh-8 in lin-39(n1760) mab-5(e1239) mutants failed to rescue the M lineage defects. An independent line of evidence comes from a comparison of mutant phenotypes: lin-39(n1760) mab-5(e1239) animals lack both M-derived coelomocytes, the majority of hlh-8(nr2061) mutants (Corsi et al., 2000; this work): (1) While lin-39(n1760) mab-5(e1239) animals lack both M-derived coelomocytes, the majority of hlh-8(nr2061) mutants (76%) contain normal numbers of M-derived coelomocytes. (2) While lin-39(n1760) mab-5(e1239) mutants lack all M-derived bodywall muscle, hlh-8(nr2061) mutants produce variable number of these cells. (3) Sex muscles can be produced in hlh-8(nr2061) mutants, although they are not fully differentiated.

The identity of other Hox targets in the M lineage is not known. We are currently using a genetic approach to identify additional candidates.
We thank QueeLim Ch’ng, Ann Corsi, Jamie Fleener, Brian Harfe, Cynthia Kenyon, Yoji Kohara, Kerry Kornfeld, Mike Krause, Julin Maloof, Siquan Xu for C. elegans strains, cDNA clones and antibodies; Ann Corsi, David Eisenmann, Cynthia Kenyon, Kerry Kornfeld, Steve Kostas, Mike Krause, Steve Salser, Cynthia Wolberqer for helpful discussions and suggestions; Cathy Branda for instruction in lineaging methods; Estella Chen, Matthew Robinson and Michael Stern for communicating unpublished results and for sharing information concerning ceh-20(n2513); and Thomas Bürglin, Ann Corsi, David Eisenmann, Manfred Frasch, Jenny Hsieh, Steve Kostas, Mike Krause, Richard Mann, Susan Parrish, Judy Yanowitz, Mariana Wölfner and two anonymous reviewers for critical comments on the manuscript. Some strains used in this study were obtained from the C. elegans Genetics Center (CGC), which is supported by a grant from the NIH Center for Research Resources. This work is supported by NIH grants (RO1GM37706 to A. Z. F.) and (F32HD08331 to J. L.).

REFERENCES


