Published in final edited form as:
Mech Dev. 2008 ; 125(5-6): 451–461.

Mesodermal expression of the *C. elegans* HMX homolog *mls-2* requires the PBC homolog CEH-20

Yuan Jiang, **Herong Shi**, **Nirav M. Amin**, **Ibrahim Sultan**, and **Jun Liu***
Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853

Abstract

Metazoan development proceeds primarily through the regulated expression of genes encoding transcription factors and components of cell signaling pathways. One way to decipher the complex developmental programs is to assemble the underlying gene regulatory networks by dissecting the *cis*-regulatory modules that direct temporal-spatial expression of developmental genes and identify corresponding *trans*-regulatory factors. Here, we focus on the regulation of a HMX homoebox gene called *mls-2*, which functions at the intersection of a network that regulates cleavage orientation, cell proliferation and fate specification in the *C. elegans* postembryonic mesoderm. In addition to its transient expression in the postembryonic mesodermal lineage, the M lineage, *mls-2* expression is detected in a subset of embryonic cells, in three pairs of head neurons and transiently in the somatic gonad. Through mutational analysis of the *mls-2* promoter, we identified two elements (E1 and E2) involved in regulating the temporal-spatial expression of *mls-2*. In particular, we showed that one of the elements (E1) required for *mls-2* expression in the M lineage contains two critical putative PBC-Hox binding sites that are evolutionarily conserved in *C. briggsae* and *C. remanei*. Furthermore, the *C. elegans* PBC homolog CEH-20 is required for *mls-2* expression in the M lineage. Our data suggests that *mls-2* might be a direct target of CEH-20 in the M lineage and that the regulation of CEH-20 on *mls-2* is likely Hox-independent.

Keywords

C. elegans; mesoderm; M lineage; *mls-2*; *ceh-20*; homeodomain; PBC; HMX; Hox; PBC-Hox binding site; promoter dissection

1. Introduction

Metazoan development proceeds by coordinated regulation of multiple cellular events, including cell proliferation, cell fate specification and differentiation. The underlying mechanism is the regulated expression of genes encoding transcription factors and components of cell signaling pathways. The *C. elegans* postembryonic mesodermal lineage, the M lineage, provides an excellent model system to investigate the coordinated mechanisms that regulate cell patterning, cell proliferation and cell fate determination. The M lineage is derived from a single pluripotent precursor, the M mesoblast. During postembryonic hermaphrodite development, the M cell divides in a characteristic and reproducible pattern to produce six

***Corresponding author:** Jun Liu, 439 Biotechnology Building, Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, Tel: 607-254-8942, Fax: 607-255-6249, E-mail: JL53@cornell.edu.

mesodermal cell types: striated body-wall muscles (BWMs) used for locomotion, coelomocytes (CCs) with a non-essential "scavenging" function, and four types of nonstriated sex muscles (type I and II vulval muscles and type I and II uterine muscles) involved in egg laying (Sulston and Horvitz, 1977).

Genetic and molecular analyses have uncovered a hierarchy of transcription factors that execute specific functions in M lineage development (Kenyon, 1986; Harfe et al, 1998a, 1998b; Corsi et al, 2000; Liu and Fire, 2000; Kostas and Fire, 2002; Jiang et al., 2005; Foehr et al., 2006; Amin et al., 2007). In particular, two Hox proteins (MAB-5 and LIN-39), along with their cofactor CEH-20, are needed to generate the M lineage: the M mesoblast often fails to divide properly in *mab-5 lin-39* double mutants or *ceh-20* single mutants so that no M descendants arise (Liu and Fire, 2000). CEH-20 forms a heterodimeric complex with either MAB-5 or LIN-39 and directly activates the expression of multiple targets including CeTwist/HLH-8, which is required to maintain the patterning and specify correct cell fates of the M lineage (Liu and Fire, 2000; Corsi et al, 2000). In addition, MAB-5 also plays a role in cell fate determination in that *mab-5* mutation causes a fate transformation from M-derived BWMs and CCs to sex myoblasts (Harfe et al., 1998b).

In addition to the two Hox proteins (MAB-5 and LIN-39) and their cofactor CEH-20, another homeodomain protein, MLS-2, also plays critical roles in regulating M lineage proliferation, patterning and cell fate specification. MLS-2 belongs to the NK class of homeodomain proteins and it is present in both the M lineage and a subset of head neurons (Jiang et al., 2005). During the early stage of M lineage development, *mls-2* expression appears to be regulated at both the transcriptional and post-transcriptional level (Jiang et al., 2005). This tight regulation of *mls-2* expression reflects the important roles of MLS-2 in the early M lineage. MLS-2 regulates cell proliferation by regulating the activity of a G1 cyclin, CYE-1, and specifies BWM and CC cell fates through multiple downstream targets, one of which is CeMyoD/HLH-1. MLS-2 also regulates cell cleavage orientation in the M lineage (Jiang et al., 2005). These observations indicate that MLS-2 is functioning at the node of a regulatory network. Thus, it is necessary to understand the molecular mechanisms behind the stringent spatio-temporal control of *mls-2* expression.

In this study, we characterized the *mls-2* promoter region and identified *cis*-elements responsible for *mls-2* expression in the M lineage and in the head neurons respectively. Site-directed mutagenesis reveals the importance of two potential PBC-Hox binding sites in regulating *mls-2* expression in the M lineage. These and additional results indicate that the PBC protein, CEH-20, is likely a direct regulator of *mls-2* expression in the M lineage.

2. Results

2.1. *mls-2* expression is spatially and temporally restricted

We have previously used pYJ59 (*5.5 kb mls-2p::gfp::mls-2::mls-2 3'UTR*), a functional GFP::MLS-2 translational fusion to examine the expression pattern of *mls-2*. pYJ59 rescues the M lineage defects of the *mls-2(cc615)* null mutants, and transgenic animals carrying pYJ59 shows the same expression pattern as MLS-2 antibody staining (Jiang et al., 2005). To further delineate the *cis*-regulatory sequences that are required for proper *mls-2* expression, we generated pYJ55 (*5.5 kb mls-2p::gfp::mls-2 3'UTR*), a transcriptional fusion construct with the *mls-2* promoter and 3' UTR regions, and pYJ51 (*5.5 kb mls-2p::gfp::unc-54 3'UTR*), another transcriptional fusion construct with the *mls-2* 3'UTR replaced by the *unc-54* 3'UTR. Transgenic animals carrying either pYJ55 or pYJ51 showed similar GFP expression pattern to that of transgenic animals carrying the functional GFP::MLS-2 translational fusion pYJ59, as summarized below.

During embryogenesis, GFP expression was detected in a subset of embryonic cells, similar to MLS-2 antibody staining (Jiang et al., 2005; data not shown). During postembryonic development, GFP was observed in several cell types. At the L1 stage, GFP was transiently expressed in the early M lineage (Jiang et al., 2005; Fig. 1A–B). During the L2 and L3 stages, GFP was transiently detected in a group of proliferating cells surrounding the gonad, which are presumably somatic gonad cells from the Z lineage (Fig.1C–D). GFP expression was also detected in three pairs of head neurons throughout postembryonic development (Fig.1A–B, E–H). We have identified two pairs of the neurons expressing *mls-2* as AIM and ASK, based on their location, their axon structures, co-localization with a *ttx-3::rfp* (an AIY specific marker, Wenick and Hobert, 2004) and DiO staining which stains amphid and phasmid neurons in the head (ASI, ADL, ASK, AWB, ASH and ASJ) and tail (PHA and PHB) (Fig.1E–H).

Thus, transgenic animals carrying either pYJ55 or pYJ51 showed a similar GFP expression pattern to that observed in transgenic animals carrying pYJ59 or that via MLS-2 antibody staining (Jiang et al., 2005; this work). Since pYJ51 does not contain any introns or the 3' UTR sequences of *mls-2*, all critical elements required for *mls-2* expression must lie in the 5.5 kb *mls-2* promoter fragment.

2.2. Two *cis*-acting regulatory elements (E1 and E2) are required for the spatiotemporal expression of *mls-2*

To further uncover mechanisms involved in regulating the expression of *mls-2*, we first generated a series of deletion constructs by removing the distal or internal parts of the *mls-2* promoter and tested the consequences of these deletions on *mls-2* promoter activity. The promoter activity of each construct was assayed by both the GFP expression pattern and the ability of GFP tagged MLS-2 to rescue the M lineage defects of *mls-2(cc615)* mutants. The latter assay requires functional levels of transgene expression in the proper temporal and spatial pattern. As summarized in Figure 2, transgenes with pYJ68 (−3825 to −1), pYJ69 (−2926 to −1), pYJ70 (−2469 to −1), pYJ82 (−2308 to −1) or pIS2 (−2221 to −1) showed expression patterns similar to that previously observed for pYJ59 (−5564 to −1). However, further 5' deletions, as in pYJ81 (−2109 to −1), pYJ71 (−1490 to −1), pYJ73 (−661 to −1) and pIS4 (−144 to −1), resulted in the loss of *mls-2* expression in the M lineage (Fig. 2). Three internal deletions, pYJ100 (Δ−639 to −403), pYJ101(Δ−2076 to −144) and pYJ164 (Δ−2077 to −144), did not affect the proper expression pattern of *mls-2* in the M lineage (Fig. 2). Taken together, these results suggest that a 144-bp element, E1, residing between −2.221 kb and −2.078 kb upstream of the *mls-2* initiation codon, is required for *mls-2* expression in the M lineage.

We also identified an element sufficient for *mls-2* expression in the head neurons, and in groups of precursor cells during embryogenesis and the L2–L3 stages. Transgenic animals containing pIS4 with the shortest promoter sequence (143 bp immediately upstream of the ATG) showed nearly normal expression pattern similar to that observed for pYJ59 in embryos as well as in a subset of proliferating cells around the gonad in L2–L3 stage larvae (Fig. 2). The pIS4 transgenic animals also showed normal pattern of neuronal expression, albeit sometimes fainter and mosaic (Fig. 2). No ectopic expression was observed using pIS4. Thus the 143-bp region immediately upstream of the translational start of *mls-2*, the E2 element, is sufficient for *mls-2* expression in the head neurons as well as in groups of precursor cells during embryogenesis and at L2–L3 stages.

Therefore, the deletion analysis identified two distinct *cis*-acting elements (E1 and E2) that regulate the temporal and tissue-specific expression of *mls-2*.

2.3. Comparison of *mls-2* promoter sequences reveals a high degree of sequence conservation among *C. elegans, C. briggsae* and *C. remanei*

To identify conserved regulatory sites in the *mls-2* promoter, we performed a phylogenetic comparison of the genomic sequences of *mls-2* orthologous genes from *C. elegans, C. briggsae* and *C. remanei*. *C. elegans, C. briggsae* and *C. remanei* have diverged between 40 and 100 million years respectively (Rudel and Kimble, 2001; Stein et al., 2003). In the coding region, the *C. elegans* MLS-2 protein shares significant sequence identity with its *C. briggsae* homolog CBG14538 (70% identity at the amino acid level), and its *C. remanei* homolog (77%), suggesting conserved roles of the three homologs. We then generated a construct with 2003bp of the *C. briggsae mls-2* promoter driving a translational *gfp::Ce-mls-2* reporter (pYJ166, the 5' end of the sequence is marked in Fig. 3A). This construct fully rescued the M lineage defects of a null allele of *mls-2*, *cc615*, and directed *gfp* expression that mimicked the endogenous *mls-2* expression pattern (data not shown). This result indicates that functionally important *cis*-regulatory sequences are conserved between the *C. elegans mls-2* and the *C. briggsae mls-2*.

We next compared 3kb of the *mls-2* promoter regions from all three species using Dialign (Morgenstern, 1999) and Dot Plots (Maizel and Lenk, 1981) and observed that the relative positions of the conserved sequence blocks are also well conserved in the *mls-2* promoters from all three species (data not shown). We particularly focused on the two elements described above, E1 (sequences from −2221 to −2078) and E2 (the 143bp sequence immediately upstream of the *mls-2* initiation codon ATG). A high degree of sequence conservation was observed in both elements using the Clustal W method (Thompson et al., 1994) (Fig. 3). We then used the TESS program (http://www.cbil.upenn.edu/tess) to look for known consensus binding sites in both elements. Among the possible transcription factor binding sites found within the E1 elements were two highly conserved sequence motifs, S1 and S2. As shown in Figure 3A, both S1 (TGATTTACGG) and S2 (TCATAAATTC) are highly conserved in *C. briggsae* and *C. remanei*, and both resemble the PBC-Hox heterodimer consensus binding site [5'-TGATNNAT(G/T)(G/A)-3'], with the 5' TGAT half being the PBC binding site and the 3' half being the Hox binding site (Chan and Mann, 1996; Mann and Affolter, 1998, Fig. 4).

2.4. The putative PBC-Hox binding sites (S1 and S2) are essential for *mls-2* expression in the M lineage

We next tested whether the S1 and S2 elements are essential for *mls-2* expression in the M lineage. Towards this goal, we made site-directed point or clustered mutations in S1, S2 or their surrounding regions in the context of the 5.5 kb or the 2.9 kb *mls-2* promoter (the 5.5 kb promoter and the 2.9 kb promoter behaved indistinguishably in all our assays), and tested the ability of these mutated promoters to drive the expression of the translational *gfp::mls-2* reporter in the M lineage, using rescue of the *cc615* M lineage defects as a read-out. The reporter was inferred to be expressed in the M lineage if the construct had the ability to rescue the M lineage defects of *cc615* mutants. The rescuing efficiency of each construct was compared to that of pYJ69, which contains 2.9 kb of the *mls-2* promoter (−2926 to −1) driving a functional translational *gfp::mls-2* fusion (Fig. 4). On average, transgenic lines carrying pYJ69 showed between 10% and 45% rescuing efficiency (45% is shown in Fig. 4). We considered any transgene that gave over 10% rescuing efficiency as being efficiently expressed in the M lineage. All the mutant constructs showed normal expression in the head neurons (data not shown).

As shown in Fig. 4, mutations in the 5' PBC half site within the S1 element (pYJ97) completely abolished the M lineage expression of the transgene, while mutations located one nucleotide upstream of the S1 element (pYJ98) did not significantly affect the M lineage expression of the transgene. These results demonstrate that the putative PBC binding site in S1 is required

for *mls-2* expression in the M lineage. Mutations in the 3' Hox half site within the S1 element (pYJ110 and pYJ108) also resulted in a failure of the transgene to be expressed in the M lineage, indicating that this putative Hox binding site is required for *mls-2* expression in the M lineage. We also mutated the S1 site from TGATTTAC to TGATTTAt (pYJ109), mimicking the perfect PBC-Hox binding site. This construct behaved just like pYJ59, the wild type construct (Fig. 4).

We then tested the importance of the S2 element for *mls-2* activity in the M lineage. This element, TCATAAATTC, also closely matches the PBC-Hox binding site and is conserved in *C. briggsae* and *C. remanei* (Fig. 3A). When base changes were introduced into the 5' PBC half site (pYJ140 and pYJ141) or the 3' Hox half site (pYJ138) of the S2 element, we saw a significant reduction of the rescuing efficiency of the corresponding transgenes, varying from 2% for pYJ138 and pYJ140 to 6% for pYJ141 (n>100). Therefore, we concluded that this putative PBC-Hox binding site, S2, is important, but not essential, for *mls-2* expression in the M lineage.

2.5. The C. elegans PBC homolog CEH-20 specifically regulates *mls-2* expression in the M lineage

There are two PBC homologs in *C. elegans*, *ceh-20* and *ceh-40*, which share redundant functions during embryogenesis (Van Auken K et al., 2002). Previous studies have shown that *ceh-20* plays an important role in M specification and patterning (Liu and Fire, 2000). We found that a weak allele of *ceh-20, ay9* (Takács-Vellai et al., 2007), also causes modest M lineage defects (data not shown). When we examined the M lineage phenotype in a presumably null allele of *ceh-40, gk159*, we found normal M lineage development in this mutant. Furthermore, *ceh-40(gk159)* did not enhance the M lineage defects of *ceh-20(ay9)* mutants (data not shown), suggesting that *ceh-40* does not function in the M lineage. We therefore focused on CEH-20 and tested whether it is required for *mls-2* expression in the M lineage.

We performed antibody staining using anti-MLS-2 antibodies in both *ceh-20(n2513)* and *ceh-20(RNAi)* animals. In over 95% of *n2513* mutants (n>200), MLS-2 was still detectable in M-derived cells and in the head neurons (Fig. 5B). However, the level of MLS-2 expression in the M lineage appeared to be greatly reduced in *n2513* homozygous animals compared to that in *n2513/+* heterozygous animals (Fig. 5A–B). *n2513* is a strong loss of function allele of *ceh-20* and exhibits a 90% penetrance of M lineage defects, while *ceh-20(RNAi)* results in 100% M lineage defects (Liu and Fire, 2000). To test if the presence of MLS-2 in *ceh-20 (n2513)* animals is due to the presence of residual CEH-20 activity, we stained for MLS-2 in *ceh-20(RNAi)* worms and found that while the M mesoblast is still present in *ceh-20(RNAi)* worms, MLS-2 expression was undetectable in the M mesoblast of *ceh-20(RNAi)* worms (n=34) (Fig. 5C). These results suggest that *mls-2* expression in the M lineage is dependent on CEH-20. In *ceh-20(RNAi)* worms, we still observed the embryonic (data not shown) and neuronal expression of *mls-2* (Fig. 5C). Neuronal cells are known to be resistant to RNAi (Kamath et al., 2000 and Timmons et al., 2001). But the presence of MLS-2 in embryonic cells, together with the absence of MLS-2 in the M lineage, of *ceh-20(RNAi)* animals suggests that CEH-20 is specifically required for the M lineage expression of *mls-2*.

2.6. The function of CEH-20 in regulating *mls-2* expression in the M lineage is likely independent of the Hox proteins

A major role of PBC/CEH-20 family members is to act as cofactors of Hox proteins to form PBC-Hox complexes that cooperatively bind to DNA (reviewed by Moens and Selleri, 2006). Both the S1 and S2 elements resemble PBC-Hox binding sites and our mutational studies suggest that the putative Hox half sites are also required for *mls-2* expression in the M lineage (Fig. 3A). Thus we tested whether the Hox genes are required for *mls-2* expression in the M

lineage. We previously showed that MLS-2 is still present in the M lineage in *lin-39(n1760) mab-5(e1239)* double null mutants or *lin-39(n1760) mab-5(e1239) egl-5(n945)* triple null mutants (Jiang et al., 2005, data not shown). Further quantification showed that the MLS-2 protein level in *lin-39(n1760) mab-5(e1239)/+ +* heterozygous mutants or *lin-39(n1760) mab-5 (e1239) egl-5(n945)/+ + +* heterozygous mutants was comparable to that in the homozygous *lin-39(n1760) mab-5(e1239)* double mutants or the homozygous *lin-39(n1760) mab-5(e1239) egl-5(n945)* triple mutants (data not shown). These results suggest that MAB-5, LIN-39 and EGL-5 are not required for *mls-2* expression in the M lineage.

C. elegans has three other Hox genes, *ceh-13, nob-1* and *php-3* (Schaller et al., 1990 and Van Auken et al., 2000). To test whether these genes are involved in regulating *mls-2* expression in the M lineage, we first tested whether any of the three genes is involved in regulating M lineage development, as loss of *mls-2* causes a variety of M lineage defects (Jiang et al., 2005). *ceh-13* alone or *nob-1* and *php-3* together are required for embryonic development (Brunschwig et al., 1999 and Van Auken et al., 2000). To bypass their embryonic requirement, we reduced the activity of each of these genes by RNAi (see Experimental procedures) or by using genetic mutations, *nob-1(ct230)* and *php-3(ok919)* (Van Auken et al., 2000). No M lineage defects were observed in any of the single RNAi or single mutants (data not shown). Since *egl-5, nob-1* and *php-3* encode posterior group Hox proteins and *nob-1* and *php-3* are functionally redundant during embryogenesis (Van Auken et al., 2000), it is also likely that these three posterior group genes share redundant functions in the M lineage. To test this hypothesis, we examined the M lineage phenotype in *nob-1(RNAi) egl-5(n945), php-3(RNAi) egl-5(n945),* and *nob-1(RNAi) php-3(RNAi) egl-5(n945)* mutants. None of the mutants showed any abnormality in the M lineage. With the caveat of RNAi, which may not completely knock out the function of the genes of interest, we concluded that *ceh-13, egl-5, nob-1* and *php-3* are not essential for the M lineage and therefore not responsible for regulating *mls-2* expression in the M lineage.

Taken together, our data suggest that none of the six Hox genes are required for *mls-2* expression in the M lineage. Thus, CEH-20 may regulate *mls-2* expression independently of the Hox factors.

2.7. Additional factors are required for *mls-2* expression in the M lineage

CEH-20 is expressed and functions in multiple cell types in addition to M lineage cells. Thus, additional factors must function together with CEH-20 to restrict the expression of *mls-2* specifically in the M lineage. As an initial step towards identifying these additional factors, we tested whether the E1 element (−2221 to −2078) is sufficient to drive reporter expression in the M lineage. We first inserted the E1 element upstream of the *egl-18* basal promoter (pYJ176) and examined whether it could drive GFP expression in the M lineage (see Materials and Methods). The *egl-18* basal promoter has been previously used to identify potential enhancer elements of other genes (Wagmaister et al, 2006; Zhao et al., 2007). No M lineage GFP expression was detected in transgenic lines carrying pYJ176 (Fig. 6A). Similarly, when the E1 element with the *egl-18* basal promoter was used to drive the expression of GFP-tagged MLS-2 (pJKL805) in the rescue assay, it failed to rescue the M lineage defects of *mls-2(cc615)* mutants (Fig. 6B). Similar constructs (pYJ175 and pJKL804) containing more 5' sequences of the *mls-2* promoter (−2469 to −2078) also failed to direct the expression of GFP (pYJ175, Fig. 6A) or GFP-tagged MLS-2 (pJKL804, Fig. 6B) in the M lineage. Thus, while the E1 element is required for proper expression of *mls-2* in the M lineage, it is not sufficient to act as an M lineage enhancer. Consistent with the above observations, multiple copies of the S1 element (pYJ174 and pJKL809, see Materials and Methods) did not direct M lineage expression of GFP or GFP-tagged MLS-2 (data not shown).

We then tested whether the E1 and E2 elements together can direct reporter expression in the M lineage. We placed both the 144bp E1 element (−2221 to −2078) and the 143bp E2 element (−143 to −1) in tandem upstream of the *egl-18* basal promoter and examined whether it can drive the expression of GFP (pYJ177) or GFP-tagged MLS-2 (pJKL806) in the M lineage. pYJ177 directed weak GFP expression in the M lineage (Fig. 6A) and pJKL806 rescued the M lineage defects of *mls-2(cc615)* mutants (Fig. 6B). This M lineage expression of the reporters requires the presence of both the E1 and the E2 elements, as the E2 element alone is not sufficient to drive the expression of GFP or GFP-tagged MLS-2 in the M lineage (see pJKL811 and pJKL812 in Fig. 6A,B). These results suggest that proper M lineage expression of *mls-2* requires both the E1 and the E2 elements and that together these two elements are sufficient to direct *mls-2* expression in the M lineage. These data also suggest that additional yet-to-be identified factors must be required for *mls-2* expression in the M lineage.

3. Discussion

In this study, we have analyzed the regulatory sequences required for the temporal-spatial expression of *mls-2*, a gene critical for multiple aspects of mesodermal development in *C. elegans*. Our results indicate that proper M lineage expression of *mls-2* requires both the 144bp promoter element (E1) located ~2.1kb upstream the *mls-2* translation initiation site and the 143bp sequence element (E2) located immediately upstream of the *mls-2* translation initiation site. Furthermore, the E2 element is capable of driving *mls-2* expression in neuronal cells. Within the E1 element, there are multiple highly conserved sequences that are required for *mls-2* expression in the M lineage. Two putative PBC-Hox binding sites are found to be close to each other and named as S1 and S2. These two sites are highly conserved in three nematode species analyzed (Fig. 3A). The S1 sequence (TGATTTACGG) is a nearly perfect match to the PBC-Hox heterodimer binding site [TGATNNAT(G/T)(G/A)] (Lu and Kamps, 1996;Chan and Mann, 1996;Knoepfler et al., 1996), while the S2 sequence (TCATAAATTC) is slightly divergent. In canonical PBC-Hox-DNA complexes, PBC contacts the constant TGAT half and Hox recognizes the 3' half (Lu et al., 1994;Knoepfler et al., 1996). Disruption of either the PBC or the Hox core binding site in S1 completely abolished *mls-2* expression in the M lineage. Unlike the S1 site, mutagenesis in the S2 site severely affected but did not completely eliminate *mls-2* expression in the M lineage. These results suggest that while both the S1 and the S2 sites contribute to *mls-2* promoter activity in the M lineage, S2 is not essential for this function.

There are two *C. elegans* PBC homologs, CEH-20 and CEH-40 (Van Auken K et al., 2002). We showed that CEH-20, rather than CEH-40, is required for proper M lineage development (Liu and Fire, 2000; this study), and that CEH-20 is specifically required for *mls-2* expression in the M lineage (Fig. 5). In *ceh-20* mutants, the M mesoblast is present and it still expresses the Hox gene *mab-5* (Liu and Fire, 2000). Thus the lack of *mls-2* expression in the M lineage, together with our findings that the putative PBC binding sites (S1 and S2) in the E1 element are required for the M lineage expression of *mls-2*, led us to propose that CEH-20 may directly bind to the S1 and S2 sites in the E1 element. CEH-20 has been shown to form complexes with Hox proteins MAB-5 or LIN-39 to directly regulate *hlh-8, egl-18* and *egl-1* expression in several tissue types (Liu and Fire, 2000, Koh et al., 2002 and Liu et al., 2006). However, we found that *mab-5, lin-39* and *egl-5* are not required for *mls-2* expression in the M lineage (Jiang et al., 2005 and this study) and that *ceh-13, nob-1* and *php-3* are not required for M lineage development. Thus, CEH-20 might interact with a non-Hox factor(s) with binding sites similar to the Hox binding sites, and together this complex binds to the S1 and S2 sequences. Non-Hox interactors of PBC proteins have been shown in other systems. For example, the *Drosophila* PBC homolog Exd directly interacts with Engrailed, a non-Hox homeodomain protein, *in vitro* (Kobayashi et al., 2003; Peltenburg and Murre, 1996), while the vertebrate PBC homolog Pbx can form a complex with bHLH proteins (such as MyoD) to bind DNA (Knoepfler et al., 1999; Berkes et al., 2004). Recent studies in *C. elegans* demonstrate that in

Q neuroblast cell migration and vulval formation, loss of *ceh-20* activity causes different defects than those seen upon loss of Hox activity (Yang et al., 2005), suggesting that PBC can function independently of Hox. At present, the identity of the CEH-20 cofactor(s) involved in regulating *mls-2* is not known, thus precluding the possibility of gel mobility shift experiments *in vitro*. We attempted to use Chromatin Immunoprecipitation (ChIP) to detect binding of CEH-20 to the S1 and S2 sites in the *mls-2* promoter. However, due to the short time window of early M lineage development and the small number of M-derived cells (1–18) relative to the whole animal (~900 somatic cells), we weren't able to detect strong and consistent signals after ChIP. Further work will be needed to identify other CEH-20 associated factors required for *mls-2* expression in the M lineage.

CEH-20 is broadly expressed in many cell types throughout development (Yang et al., 2005; Liu et al., 2006; Takacs-Vellai K et al., 2006; J.Y. and J. L., unpublished data). Thus other factors, in addition to CEH-20, must function to allow the proper spatiotemporal expression of *mls-2*. The results from our enhancer assays suggested that the E2 element is also required for *mls-2* expression in the M lineage (Fig. 6). Furthermore, we have identified another region in the E1 element, S3 (TGTTGCGAACCAAACC, located between −2152 and −2139), which is required for proper M lineage expression of *mls-2* (data not shown). Further characterizing the E2 element and the S3 site and the identification of the *trans*-acting factors binding to them will help to elucidate the spatio-temporal regulation of *mls-2*.

Experimental procedures

C. elegans strains

Strains were manipulated under standard conditions as described by Brenner (Brenner, 1974). Analyses were performed at 20°C, unless otherwise noted. The alleles used in this work are: LGIII, *ceh-20(n2513)* (Liu and Fire, 2000), *ceh-20(ay9)* (Takács-Vellai et al., 2007), *pha-1 (e2123ts)* (Schnabel and Schnabel, 1990), *lin-39(n1760) mab-5(e1239)/dpy-17(e164) unc-32 (e189)* (Liu and Fire, 2000), *egl-5(n945)* (Wang et al., 1993), *nob-1(ct230), php-3(ok919)* (Van Auken et al., 2000); LG V, *him-5(e1467)* (Hodgkin et al., 1979); LGX, *mls-2(cc615), ceh-40(gk159) (C. elegans* knockout consortium). Other transgenic lines: OH1098: *otIs133 (ttx-3∷rfp)* (Wenick and Hobert, 2004).

Generating transgenic lines

Transgenic lines of *C. elegans* were generated as described by Mello and colleagues (Mello et al., 1991). The plasmid pRF4 (Mello et al., 1991) or the *pha-1* rescuing plasmid pC1 (Granato et al., 1994) were used as markers. The promoter constructs were introduced into wild type and/or *cc615* mutant animals. Expression of the transgene in the M lineage was monitored by assaying GFP∷MLS-2 expression in the M lineage under the fluorescence microscope and the ability of the transgene to rescue the M lineage defects of *cc615* mutants. Neuronal expression was directly monitored under the fluorescence microscope. For enhancer assays, the constructs were introduced into LW1066: *jjIs1066[pJKL705.1(hlh-8p∷mRFP+unc-119(+))]; unc-119 (ed4)*.

Plasmid construction

Three *gfp* reporter constructs have been used to examine *mls-2* expression:

pYJ51: *5.5 kb mls-2p∷gfp∷unc-54 3'UTR*

pYJ55: *5.5 kb mls-2 promoter∷gfp∷mls-2 3' UTR* (Jiang et al., 2005)

pYJ59: *5.5 kb mls-2 promoter∷gfp∷mls-2∷mls-2 3' UTR* (Jiang et al., 2005)

The following constructs were generated from pYJ59 and used in the promoter deletion analysis:

pYJ68 *[3.8 kb of mls-2p (−3825 to −1)∷gfp∷mls-2∷mls-2 3' UTR]*

pYJ69 *[2.9 kb of mls-2p (−2926 to −1)∷gfp∷mls-2∷mls-2 3' UTR]*

pYJ96 *[2.9 kb of mls-2p (−2926 to −1)∷rfp∷mls-2∷mls-2 3' UTR]*

pYJ70 *[2.5 kb of mls-2p (−2469 to −1)∷gfp∷mls-2∷mls-2 3' UTR]*

pYJ82 *[2.3 kb of mls-2p (−2308 to −1)∷gfp∷mls-2∷mls-2 3' UTR]*

pIS2 *[2.2 kb of mls-2p (−2221 to −1)∷gfp∷mls-2∷mls-2 3' UTR]*

pYJ81 *[2.1 kb of mls-2p (−2109 to −1)∷gfp∷mls-2∷mls-2 3' UTR]*

pYJ74 *[1.5 kb of mls-2p (−1490 to −1)∷gfp∷mls-2∷mls-2 3' UTR]*

pYJ73 *[0.6 kb of mls-2p (−661 to −1)∷gfp∷mls-2∷mls-2 3' UTR]*

pIS4 *[144 bp of mls-2p (−144 to −1)∷gfp∷mls-2∷mls-2 3' UTR]*

pYJ100*[−2926bp to −1bp (Δ−639 to −403) of mls-2p∷gfp∷mls-2∷mls-2 3' UTR]*

pYJ101*[−2926bp to −1bp (Δ−2076 to −144) of mls-2p∷gfp∷mls-2∷mls-2 3' UTR]*

pYJ164*[−2201bp to −1bp (Δ−2077 to −144) of mls-2p∷gfp∷mls-2∷mls-2 3' UTR]*

The following constructs were generated for enhancer assays (*egl-18p* in these constructs refers to the *egl-18* basal promoter):

pYJ175 *[392bp of mls-2p (−2469 to −2078)∷egl-18p∷gfp∷unc-54 3' UTR]*

pYJ176 *[144bp of mls-2p (−2221 to −2078)∷egl-18p∷gfp∷unc-54 3' UTR]*

pYJ177 *[144bp of mls-2p (−2221 to −2078)+143bp of mls-2p (−143 to −1) ∷egl-18p∷gfp∷unc-54 3' UTR]*

pJKL812 *[143bp of mls-2p (−143 to −1)∷egl-18p∷gfp∷unc-54 3' UTR]*

pYJ174 *[9X (S3+S1) of mls-2p (−2153 to −2101)∷egl-18p∷gfp∷unc-54 3' UTR]*

pJKL804 *[392bp of mls-2p (−2469 to −2078)∷egl-18p∷gfp∷mls-2∷mls-2 3' UTR]*

pJKL805 *[144bp of mls-2p (−2221 to −2078)∷egl-18p∷gfp∷mls-2∷mls-2 3' UTR]*

pJKL806 *[144bp of mls-2p (−2221 to −2078)+143bp of mls-2p (−143 to −1) ∷egl-18p∷gfp∷mls-2∷mls-2 3' UTR]*

pJKL811 *[143bp of mls-2p (−143 to −1)∷egl-18p∷gfp∷mls-2∷mls-2 3' UTR]*

pJKL809 *[9X (S3+S1) of mls-2p (−2153 to −2101)∷egl-18p∷gfp∷ mls-2∷mls-2 3' UTR]*

The constructs containing small deletions (see above and Fig. 2) or point mutations (listed in Fig. 4 and Fig. 6) were made via the mutagenesis strategy by overlapping PCR using customized primers and pYJ59 or pYJ69 as template (Ho et al., 1989). All mutated products were confirmed by DNA sequencing.

RNA interference

pJKL422.1 (Liu and Fire, 2000) and pYJ133 were used to make dsRNA for *ceh-20* and *php-3* respectively. *yk403d9* and *yk467d4* (gift from Yuji Kohara) were used for making dsRNA against *nob-1*. For *ceh-13* RNAi, primers JKL-601 and JKL-602 were used to amplify the insert in the *ceh-13* RNAi clone III-4C05 from the Ahringer library (Fraser et al., 2000; Kamath et al., 2003). dsRNA was synthesized following the protocol of Fire and colleagues

(1998). dsRNA for *php-3* or *ceh-20* was injected into wild-type gravid adults. Progeny of injected animals were collected at the early L1 stage and examined for M lineage defects or used for immunostaining. Water injected worms were used as controls. For *ceh-13(RNAi)*, *nob-1(RNAi)*, *php-3(RNAi)* or *nob-1(RNAi) php-3(RNAi)*, a soaking protocol modified from Ahringer et al. (2006) was used. Embryos and the subsequent newly hatched L1s were soaked in corresponding dsRNAs or in soaking buffer alone for 24–48 hours at 20°C. The soaked L1 worms were transferred to NGM plates with OP50 and followed throughout development for M lineage phenotypes.

Immunofluorescence staining

For immunostaining using MLS-2 antibodies, early L1 larvae of *ceh-20(n2513)* or *ceh-20 (RNAi)* animals were fixed and stained as described in Hurd and Kemphues (2003). The primary antibodies were preadsorbed rat anti-MLS-2 antibodies (CUMC-R6) (1:400 to 1:1000) (Jiang et al., 2005) and goat anti-GFP antibodies (Rockland Immunochemicals) (1:5000). The secondary antibodies (both from Jackson Immunoresearch Laboratories) were affinity-purified Cy3-conjugated donkey anti-rat (1:400), affinity-purified FITC-conjugated donkey anti-goat (1:100).

DiO staining

For DiO staining, transgenic animals carrying pYJ96 *(mls-2p∷rfp∷mls-2)* were incubated in 2 μg/ml 3,3′-dioctadecyloxacarbocyanine perchlorate (DiO) (from Molecular Probes, Catalog# D-275) staining medium for 2–3 hours at room temperature as described by Perkins et al. (1986) and Herman and Hedgecock (1990). Worms were then washed with M9 buffer three times and transferred to NGM plates with OP50 bacteria. After one hour, worms were examined for DiO and RFP signals using a Leica DMRA2 compound microscope.

Acknowledgements

We thank Ann Corsi and Stephany Meyers for the plasmid containing the *egl-18* basal promoter, Oliver Hobert for the *ttx-3∷rfp* transgenic line, Yuji Kohara for cDNA clones, Alexander Soneru for help with the *ceh-13(RNAi)* experiments and Rachel Fairbank for help generating the *hlh-8p∷mRFP* strain (LW1066). Some strains used in this study were obtained from the *C. elegans* Genetics Center (CGC), which is supported by a grant from the NIH National Center for Research Resources. This work was supported by NIH R01 GM066953 (to J. L.). I. S. was supported by the Summer Research Fellowship from Weill Cornell Medical College in Qatar.

References

Ahringer, J. Reverse genetics. In: The *C. elegans* Research Community. , editor. Wormbook. 2006. http://www.wormbook.org

Amin NM, Hu K, Pruyne D, Terzic D, Bretscher A, Liu J. A Zn-finger/FH2-domain containing protein, FOZI-1, acts redundantly with CeMyoD to specify striated body wall muscle fates in the Caenorhabditis elegans postembryonic mesoderm. Development 2007;134:19–29. [PubMed: 17138663]

Berkes CA, Bergstrom DA, Penn BH, Seaver KJ, Knoepfler PS, Tapscott SJ. Pbx marks genes for activation by MyoD indicating a role for a homeodomain protein in establishing myogenic potential. Mol. Cell 2004;14:465–477. [PubMed: 15149596]

Brenner S. The genetics of Caenorhabditis elegans. Genetics 1974;77:71–94. [PubMed: 4366476]

Brunschwig K, Wittmann C, Schnabel R, Burglin TR, Tobler H, Muller F. Anterior organization of the Caenorhabditis elegans embryo by the labial-like Hox gene ceh-13. Development 1999;126:1537–1546. [PubMed: 10068646]

Chan SK, Mann RS. A structural model for a homeotic protein-extradenticle-DNA complex accounts for the choice of HOX protein in the heterodimer. Proc. Natl. Acad. Sci. U. S. A 1996;93:5223–5228. [PubMed: 8643557]

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 1998;391:806–811. [PubMed: 9486653]

Foehr ML, Lindy AS, Fairbank RC, Amin NM, Xu M, Yanowitz J, Fire AZ, Liu J. An antagonistic role for the C. elegans Schnurri homolog SMA-9 in modulating TGFbeta signaling during mesodermal patterning. Development 2006;133:2887–2896. [PubMed: 16790477]

Fraser AG, Kamath RS, Zipperlen P, Martinez-Campos M, Sohrmann M, Ahringer J. Functional genomic analysis of C. elegans chromosome I by systematic RNA interference. Nature 2000;408:325–330. [PubMed: 11099033]

Granato M, Schnabel H, Schnabel R. pha-1, a selectable marker for gene transfer in C. elegans. Nucleic Acids Res 1994;22:1762–1763. [PubMed: 8202383]

Harfe BD, Branda CS, Krause M, Stern MJ, Fire A. MyoD and the specification of muscle and non-muscle fates during postembryonic development of the C. elegans mesoderm. Development 1998;125:2479–2488. [PubMed: 9609831]

Harfe BD, Vaz Gomes A, Kenyon C, Liu J, Krause M, Fire A. Analysis of a Caenorhabditis elegans Twist homolog identifies conserved and divergent aspects of mesodermal patterning. Genes Dev 1998;12:2623–2635. [PubMed: 9716413]

Herman RK, Hedgecock EM. Limitation of the size of the vulval primordium of Caenorhabditis elegans by lin-15 expression in surrounding hypodermis. Nature 1990;348:169–171. [PubMed: 2234080]

Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 1989;77:51–59. [PubMed: 2744487]

Hurd DD, Kemphues KJ. PAR-1 is required for morphogenesis of the Caenorhabditis elegans vulva. Dev. Biol 2003;253:54–65. [PubMed: 12490197]

Jiang Y, Horner V, Liu J. The HMX homeodomain protein MLS-2 regulates cleavage orientation, cell proliferation and cell fate specification in the C. elegans postembryonic mesoderm. Development 2005;132:4119–4130. [PubMed: 16107479]

Kamath RS, Ahringer J. Genome-wide RNAi screening in Caenorhabditis elegans. Methods 2003;30:313–321. [PubMed: 12828945]

Kamath RS, Martinez-Campos M, Zipperlen P, Fraser AG, Ahringer J. Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in Caenorhabditis elegans. Genome Biol. 2, RESEARCH0002. 2001

Kenyon C. A gene involved in the development of the posterior body region of C. elegans. Cell 1986;46:477–487. [PubMed: 3731276]

Knoepfler PS, Bergstrom DA, Uetsuki T, Dac-Korytko I, Sun YH, Wright WE, Tapscott SJ, Kamps MP. A conserved motif N-terminal to the DNA-binding domains of myogenic bHLH transcription factors mediates cooperative DNA binding with pbx-Meis1/Prep1. Nucleic Acids Res 1999;27:3752–3761. [PubMed: 10471746]

Knoepfler PS, Lu Q, Kamps MP. Pbx-1 Hox heterodimers bind DNA on inseparable half-sites that permit intrinsic DNA binding specificity of the Hox partner at nucleotides 3' to a TAAT motif. Nucleic Acids Res 1996;24:2288–2294. [PubMed: 8710498]

Kobayashi M, Fujioka M, Tolkunova EN, Deka D, Abu-Shaar M, Mann RS, Jaynes JB. Engrailed cooperates with extradenticle and homothorax to repress target genes in Drosophila. Development 2003;130:741–751. [PubMed: 12506004]

Koh K, Peyrot SM, Wood CG, Wagmaister JA, Maduro MF, Eisenmann DM, Rothman JH. Cell fates and fusion in the C. elegans vulval primordium are regulated by the EGL-18 and ELT-6 GATA factors -- apparent direct targets of the LIN-39 Hox protein. Development 2002;129:5171–5180. [PubMed: 12399309]

Kostas SA, Fire A. The T-box factor MLS-1 acts as a molecular switch during specification of nonstriated muscle in C. elegans. Genes Dev 2002;16:257–269. [PubMed: 11799068]

Liu H, Strauss TJ, Potts MB, Cameron S. Direct regulation of egl-1 and of programmed cell death by the Hox protein MAB-5 and by CEH-20, a C. elegans homolog of Pbx1. Development 2006;133:641–650. [PubMed: 16421192]

Liu J, Fire A. Overlapping roles of two Hox genes and the exd ortholog ceh-20 in diversification of the C. elegans postembryonic mesoderm. Development 2000;127:5179–5190. [PubMed: 11060243]

Lu Q, Kamps MP. Structural determinants within Pbx1 that mediate cooperative DNA binding with pentapeptide-containing Hox proteins: proposal for a model of a Pbx1-Hox-DNA complex. Mol. Cell. Biol 1996;16:1632–1640. [PubMed: 8657138]

Lu Q, Wright DD, Kamps MP. Fusion with E2A converts the Pbx1 homeodomain protein into a constitutive transcriptional activator in human leukemias carrying the t(1;19) translocation. Mol. Cell. Biol 1994;14:3938–3948. [PubMed: 7910944]

Maizel JV, Lenk RP. Enhanced graphic matrix analysis of nucleic acid and protein sequences. Proc. Natl. Acad. Sci. U.S.A 1981;78:7665–7669. [PubMed: 6801656]

Mann RS, Affolter M. Hox proteins meet more partners. Curr. Opin. Genet. Dev 1998;8:423–429. [PubMed: 9729718]

Mello CC, Kramer JM, Stinchcomb D, Ambros V. Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. EMBO J 1991;10:3959–3970. [PubMed: 1935914]

Morgenstern B. DIALIGN 2: improvement of the segment-to-segment approach to multiple sequence alignment. Bioinformatics 1999;15:211–218. [PubMed: 10222408]

Peltenburg LT, Murre C. Engrailed and Hox homeodomain proteins contain a related Pbx interaction motif that recognizes a common structure present in Pbx. EMBO J 1996;15:3385–3393. [PubMed: 8698039]

Perkins LA, Hedgecock EM, Thomson JN, Culotti JG. Mutant sensory cilia in the nematode Caenorhabditis elegans. Dev. Biol 1986;117:456–487. [PubMed: 2428682]

Ramji DP, Foka P. CCAAT/enhancer-binding proteins: structure, function and regulation. Biochem. J 2002;365:561–575. [PubMed: 12006103]

Rudel D, Kimble J. Conservation of glp-1 regulation and function in nematodes. Genetics 2001;157:639–654. [PubMed: 11156985]

Schaller D, Wittmann C, Spicher A, Muller F, Tobler H. Cloning and analysis of three new homeobox genes from the nematode Caenorhabditis elegans. Nucleic Acids Res 1990;18:2033–2036. [PubMed: 1970877]

Schnabel R, Schnabel H. Early determination in the C. elegans embryo: a gene, cib-1, required to specify a set of stem-cell-like blastomeres. Development 1990;108:107–119. [PubMed: 2351058]

Stein LD, Bao Z, Blasiar D, Blumenthal T, Brent MR, Chen N, Chinwalla A, Clarke L, Clee C, Coghlan A, et al. The genome sequence of Caenorhabditis briggsae: a platform for comparative genomics. PLoS Biol 2003;1:E45. [PubMed: 14624247]

Sulston JE, Horvitz HR. Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. Dev. Biol 1977;56:110–156. [PubMed: 838129]

Takács-Vellai K, Vellai T, Chen EB, Zhang Y, Guerry F, Stern MJ, Muller F. Transcriptional control of Notch signaling by a HOX and a PBX/EXD protein during vulval development in C. elegans. Dev. Biol 2007;302:661–669. [PubMed: 17084835]

Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–4680. [PubMed: 7984417]

Timmons L, Court DL, Fire A. Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in Caenorhabditis elegans. Gene 2001;263:103–112. [PubMed: 11223248]

Van Auken K, Weaver D, Robertson B, Sundaram M, Saldi T, Edgar L, Elling U, Lee M, Boese Q, Wood WB. Roles of the Homothorax/Meis/Prep homolog UNC-62 and the Exd/Pbx homologs CEH-20 and CEH-40 in C. elegans embryogenesis. Development 2002;129:5255–5268. [PubMed: 12399316]

Van Auken K, Weaver DC, Edgar LG, Wood WB. Caenorhabditis elegans embryonic axial patterning requires two recently discovered posterior-group Hox genes. Proc. Natl. Acad. Sci. U. S. A 2000;97:4499–4503. [PubMed: 10781051]

Wagmaister JA, Miley GR, Morris CA, Gleason JE, Miller LM, Kornfeld K, Eisenmann DM. Identification of *cis*-regulatory elements from the C. elegans Hox gene lin-39 required for embryonic expression and for regulation by the transcription factors LIN-1, LIN-31 and LIN-39. Dev. Biol 2006;297:550–565. [PubMed: 16782085]

Wenick AS, Hobert O. Genomic cis-regulatory architecture and trans-acting regulators of a single interneuron-specific gene battery in C. elegans. Dev. Cell 2004;6:757–770. [PubMed: 15177025]

Yang L, Sym M, Kenyon C. The roles of two C. elegans HOX co-factor orthologs in cell migration and vulva development. Development 2005;132:1413–1428. [PubMed: 15750187]

Zhao J, Wang P, Corsi AK. The C. elegans Twist target gene, arg-1, is regulated by distinct E box promoter elements. Mech Dev 2007;124:377–389. [PubMed: 17369030]

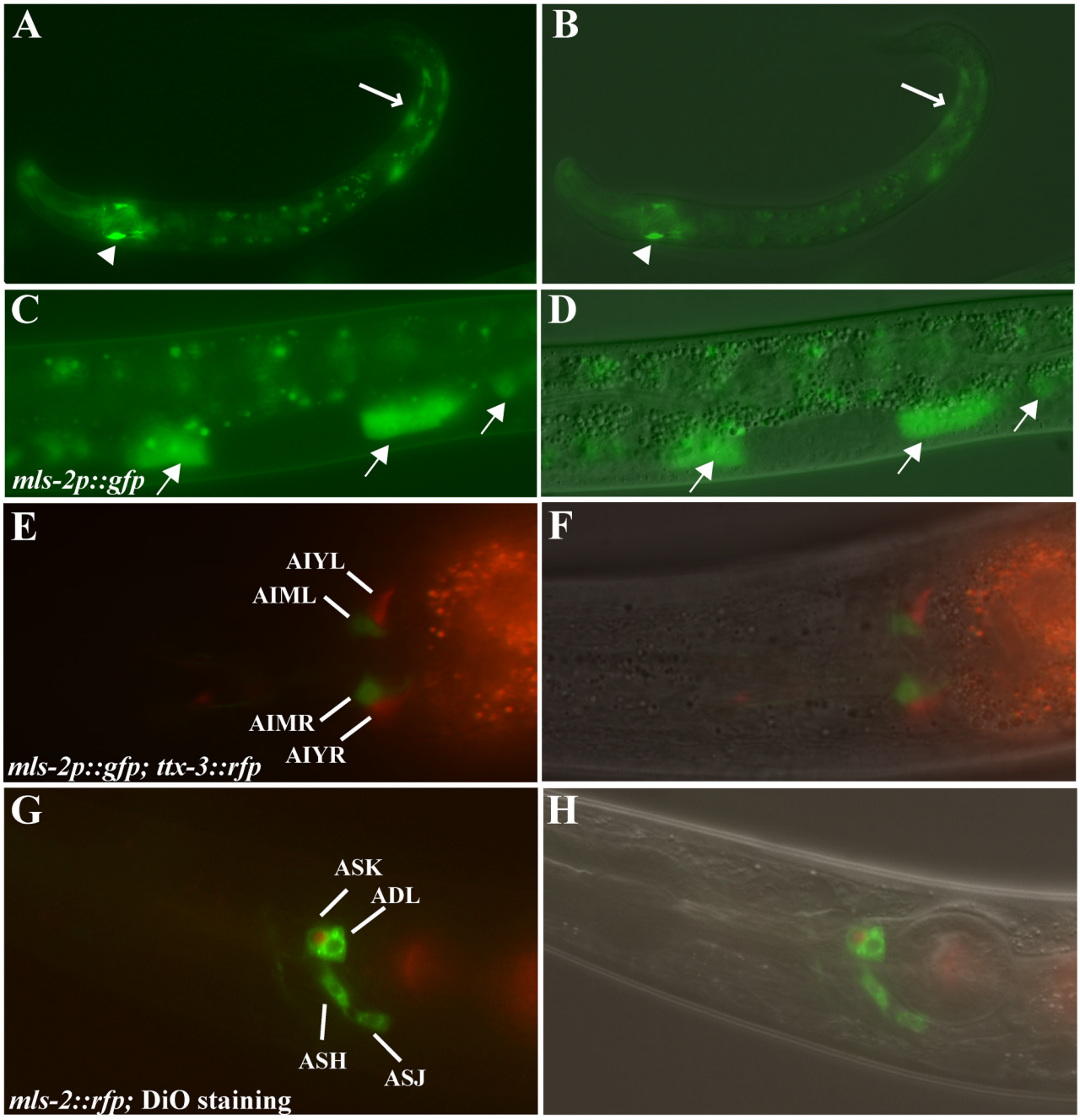

Figure 1. Expression of *mls-2* in the M lineage, the somatic gonad and head neurons

(A, C, E, G) fluorescence images and (B, D, F, H) the corresponding DIC images overlaid with fluorescence images. (A, B) A L1 animal carrying a transcriptional fusion pYJ55 (*mls-2p∷gfp*). *gfp* expression is seen in M descendants (arrows) and head neurons (arrowheads). (C, D) A L3 animal carrying pYJ55. *gfp* expression in cells surrounding the gonad is marked by solid arrows. (E, F) An adult animal carrying pYJ55 (green) and *ttx-3∷rfp* (red) showing *mls-2* expression in the AIM. (G, H) An adult animal carrying pYJ96 (*mls-2p∷rfp∷mls-2*, red) stained with DiO (green), showing *mls-2* expression in ASK.

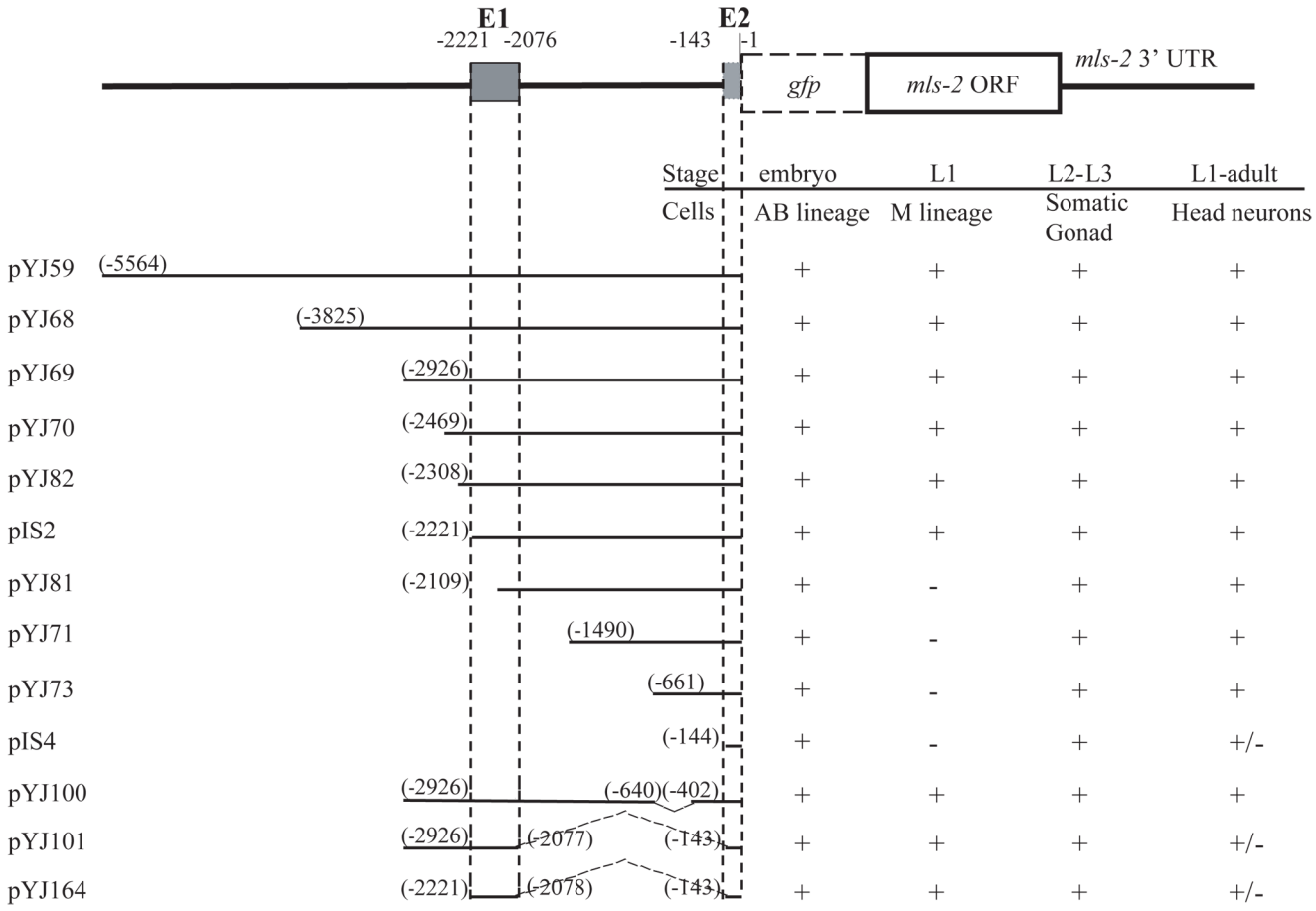

Figure 2. Summary of deletion analysis of the *mls-2* promoter

All of the constructs contain *mls-2* promoters of different sizes (indicated by the numbers in parentheses), the *mls-2* ORF fused with *gfp* and the *mls-2* 3' UTR. The start codon (ATG) is marked as −1. The diagram is not drawn to scale. GFP fluorescence in cells is represented as + (consistent expression), - (no expression or below detection), and +/− (mosaic expression).

Figure 3. Alignment of the *C. elegans* E1 (A) and E2 (B) elements with the corresponding *C. briggsae* and *C. remanei* sequences

Numbers indicate the positions of nucleotides relative to ATG (−1 position, as marked by the red bars in B). S1 and S2 sequences (by dashed blue boxes) contain two putative consensus PBC-Hox binding sequences. The S3 sequence (by solid black box) is also required for *mls-2* promoter activity (data not shown). The red arrow in (A) points to the 5' end point of the *C. briggsae* promoter used in the rescuing experiment.

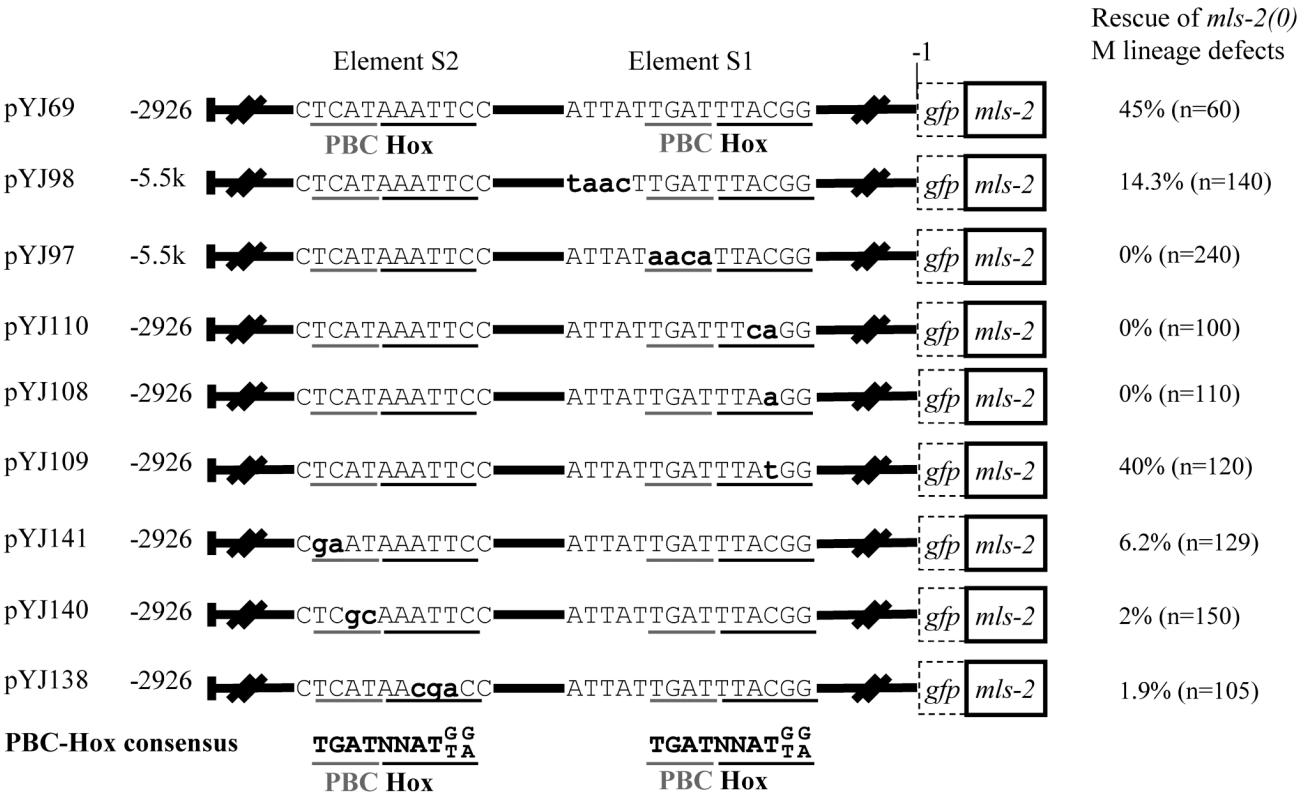

Figure 4. The PBC-Hox binding sites in the S1 and S2 elements are important for proper M lineage expression of *mls-2*

The constructs depicted were tested for their ability to rescue the M lineage defects of *mls-2 (cc615)* mutants in transgenic animals. Mutated sequences are indicated in bold and lower case font. The rescuing percentage was assessed in F1 progeny, with the number of animals examined listed in parentheses. The PBC-Hox consensus binding sites are depicted at the bottom of the diagram.

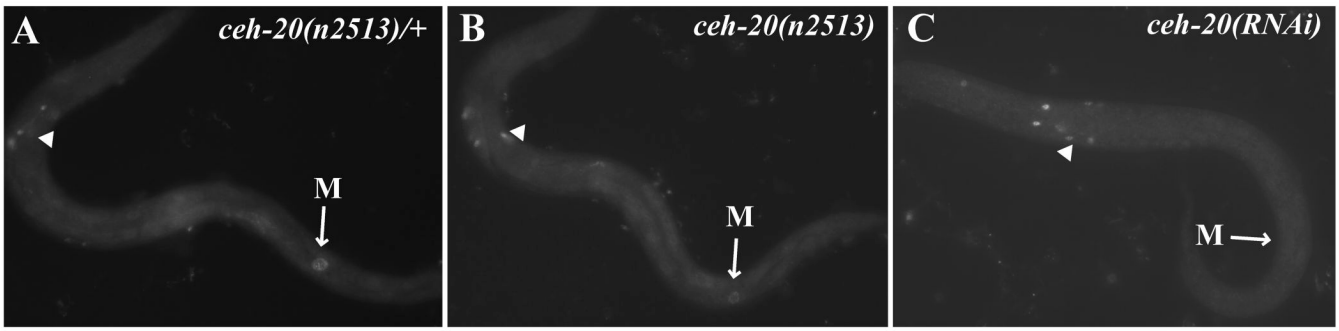

Figure 5. *ceh-20* regulates *mls-2* expression in the M lineage

MLS-2 expression in the M mesoblast (by arrows) and the head neurons (arrowheads) as observed by immunostaining with anti-MLS-2 antibodies in *ceh-20(n2513)*/+ (A), *ceh-20 (n2513)* (B) and *ceh-20(RNAi)* (C) animals. Images in A and B were taken under the same exposure time.

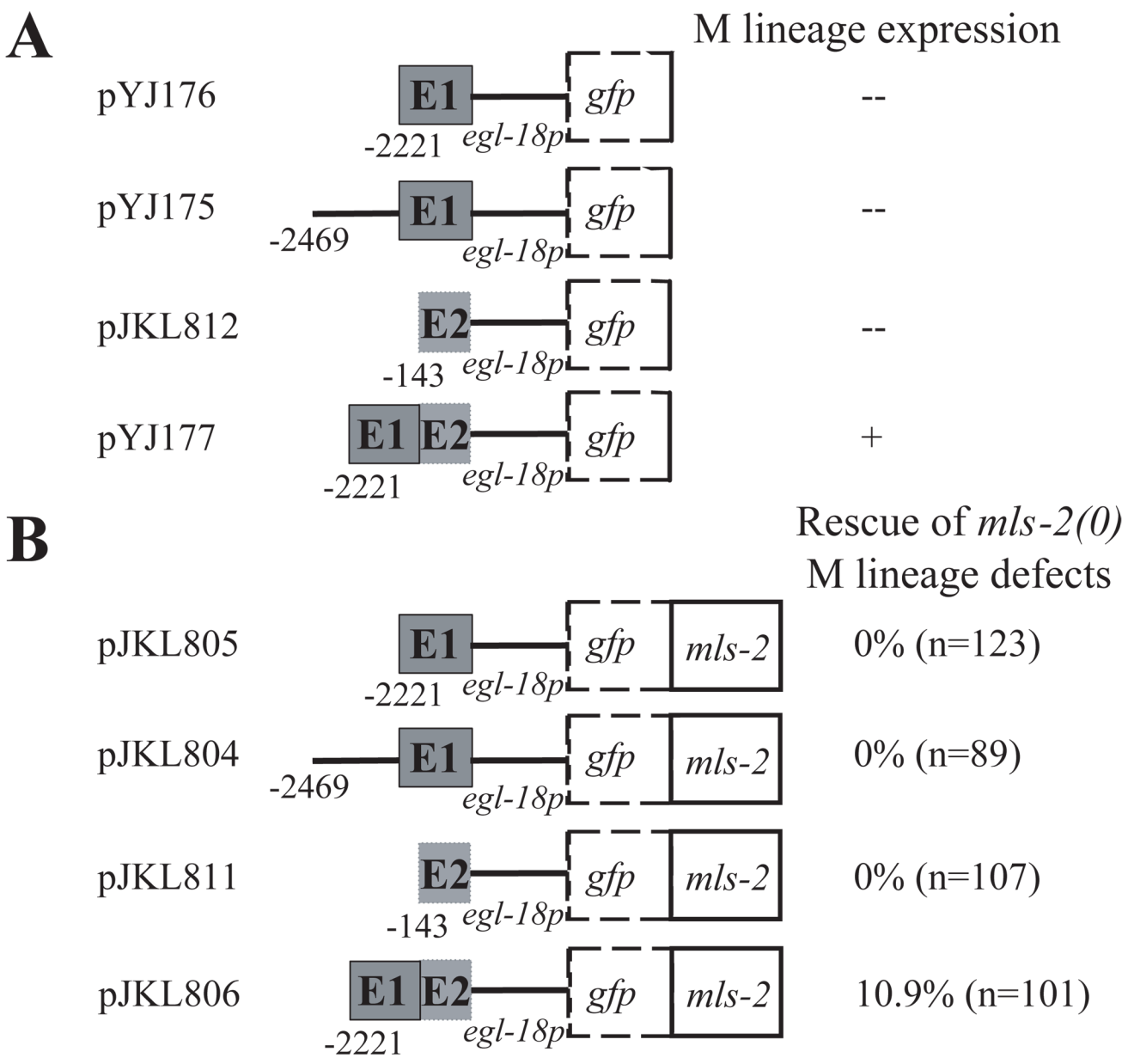

Figure 6. The E2 element of the *mls-2* promoter is required for *mls-2* expression in the M lineage

Enhancer assays showing that the E2 element is required, but not sufficient, for both GFP (A) and GFP::MLS-2 (B) expression in the M lineage. The rescuing percentage was assessed in F1 progeny, with the number of animals examined listed in parentheses.