Development 137, 963-974 (2010) doi:10.1242/dev.044719

The HMX/NKX homeodomain protein MLS-2 specifies the identity of the AWC sensory neuron type via regulation of the *ceh-36* Otx gene in *C. elegans*

Kyuhyung Kim*, Rinho Kim* and Piali Sengupta†

SUMMARY

The differentiated features of postmitotic neurons are dictated by the expression of specific transcription factors. The mechanisms by which the precise spatiotemporal expression patterns of these factors are regulated are poorly understood. In *C. elegans*, the *ceh-36* Otx homeobox gene is expressed in the AWC sensory neurons throughout postembryonic development, and regulates terminal differentiation of this neuronal subtype. Here, we show that the HMX/NKX homeodomain protein MLS-2 regulates *ceh-36* expression specifically in the AWC neurons. Consequently, the AWC neurons fail to express neuron type-specific characteristics in *mls-2* mutants. *mls-2* is expressed transiently in postmitotic AWC neurons, and directly initiates *ceh-36* expression. CEH-36 subsequently interacts with a distinct site in its cis-regulatory sequences to maintain its own expression, and also directly regulates the expression of AWC-specific terminal differentiation genes. We also show that MLS-2 acts in additional neuron types to regulate their development and differentiation. Our analysis describes a transcription factor cascade that defines the unique postmitotic characteristics of a sensory neuron subtype, and provides insights into the spatiotemporal regulatory mechanisms that generate functional diversity in the sensory nervous system.

KEY WORDS: *C. elegans*, HMX/NKX, OTX, Sensory neuron, Transcription factor

INTRODUCTION

Animals recognize and respond to a bewildering array of sensory cues in their environment. Environmental signals are recognized by peripheral sensory neuron types, each of which exhibits a unique functional profile (Bargmann et al., 1993; Malnic et al., 1999; Pichaud et al., 1999). Functional diversity in the sensory nervous system is generated by the expression of specific complements of signaling molecules, such as channels and receptors, as well as molecules that confer neuron-specific morphologies in individual neuron types (Bruhn and Cepko, 1996; Malnic et al., 1999; Etchberger et al., 2007). The precise regulation of complements of neuron-specific genes is therefore essential to ensure fidelity and diversity in sensory responses. The developmental mechanisms by which the differentiated properties of postmitotic sensory neurons are specified are not fully understood.

The sensory nervous system of *C. elegans* provides an excellent model system in which to investigate the developmental cascades that specify cell-specific identities via the initiation and maintenance of gene expression in postmitotic sensory neurons. The bilateral amphid organs of the head comprise 12 sensory neuron pairs, which arise from invariant lineages in the embryo (Sulston et al., 1983). Each of these sensory neuron types has been shown to mediate distinct sensory functions, and to express unique sets of terminal differentiation genes (Troemel et al., 1995; Uchida et al., 2003; Colosimo et al., 2004; Melkman and Sengupta, 2004; Bargmann, 2006; Etchberger et al., 2007). The ability to consistently identify and follow gene expression in individual neuron types, together with the amenability of the *C. elegans* system to forward and reverse genetic analyses at high resolution, have led to the identification of developmental pathways that regulate the expression of postmitotic features of subsets of sensory neurons (Lanjuin and Sengupta, 2004; Hobert, 2008).

A general theme that has arisen from these studies is the concept of a 'terminal selector' transcription factor (TF) that directly or indirectly regulates the expression of all terminal differentiation genes specific to that neuron type, but not that of 'generic' neuronal genes (Hobert, 2008). In the well-studied ASE amphid chemosensory neuron type, it has been shown that the CHE-1 zinc-finger TF directly regulates the expression of the majority of terminal differentiation genes (Chang et al., 2003; Uchida et al., 2003; Etchberger et al., 2007; Etchberger et al., 2009). In *che-1* mutants, the ASE neurons fail to express any ASE-specific characteristics, while retaining pan-neuronal features. Similarly, mutations in the *ceh-36* Otx, *lim-4* LIM-homeobox, *odr-7* nuclear hormone receptor and *ttx-1* Otx genes result in loss of postmitotic characteristics of the AWC, AWB, AWA and AFD amphid sensory neuron types, respectively (Sengupta et al., 1994; Sagasti et al., 1999; Satterlee et al., 2001; Lanjuin et al., 2003). Thus, candidate terminal selector proteins belonging to different TF families specify sensory neuronal identities and generate neuronal diversity in the *C. elegans* sensory system.

A defining feature of these terminal selector TFs is that they are expressed specifically or selectively in a small number of neurons postmitotically, and that their expression is maintained throughout development (Way and Chalfie, 1989; Sagasti et al., 1999; Altun-Gultekin et al., 2001; Sarafi-Reinach et al., 2001; Satterlee et al., 2001). Thus, precise spatiotemporal regulation of expression of these TFs is crucial for correct neuronal differentiation. It is likely that the cis-regulatory regions of these TF genes integrate multiple extrinsic and intrinsic cues to initiate and/or maintain gene

Department of Biology and National Center for Behavioral Genomics, Brandeis University, Waltham, MA 02454, USA.

*These authors contributed equally to this work
†Author for correspondence (sengupta@brandeis.edu)

Accepted 15 January 2010

expression in specific cell types. Analyses of these cues require that the trans-acting factors that direct the expression of terminal selector TFs in specific neuron types be identified and characterized. In the mechanosensory neurons of *C. elegans*, it has been shown that expression of the *mec-3* terminal selector gene is regulated by the POU-domain TF UNC-86 (Way and Chalfie, 1988; Xue et al., 1992; Duggan et al., 1998); *unc-86* expression, in turn, is regulated by multiple lineage-specific upstream factors (Baumeister et al., 1996). Recently, elegant work has shown that members of a Zic and bHLH TF family act together with a transiently acting Wnt/β-catenin pathway to regulate the expression of the terminal selector TFs TTX-3 and CEH-10 in the AIY interneurons, thereby coupling asymmetric cell division to the acquisition of neuronal identity (Bertrand and Hobert, 2009). However, a generalized description of the principles governing spatiotemporal expression of terminal selector proteins requires that additional upstream regulatory mechanisms be identified and analyzed.

Here, we show that the *mls-2* Hmx/Nkx homeobox gene is necessary to drive expression of the candidate terminal selector *ceh-36* Otx homeobox gene in the AWC sensory neurons. In *mls-2* mutants, *ceh-36* expression is specifically abolished in the AWC neurons, and this neuron type fails to express AWC-specific terminal differentiation genes. We find that *mls-2* is expressed transiently in the AWC neurons, and directly initiates *ceh-36* expression. We also demonstrate that following initiation of expression, CEH-36 autoregulates to maintain its own expression via direct binding to a cognate site in its upstream regulatory sequences. Moreover, CEH-36 directly regulates the expression of terminal differentiation genes in the AWC neurons. Our analyses also indicate that MLS-2 acts independently in several lineages to direct the specification of neuronal and non-neuronal cell types. Taken together with previous findings, this work allows us to propose a general scheme for the regulation of terminal selector proteins in chemosensory neuron types, and provides insight into how neuronal diversity is generated via the context-dependent functions of individual TFs.

MATERIALS AND METHODS

C. elegans strains

The wild-type strain used was *C. elegans* variety Bristol, strain N2. Mutant strains were *mls-2(cc615)* (Jun Liu, Cornell University), *mls-2(ns156)* (Shai Shaham, Rockefeller University), *mls-2(tm252)* (Shohei Mitani, National Bioresource Project, Japan), or were obtained from the *Caenorhabditis* Genetics Center (University of Minnesota). Strains carrying stably integrated transgenes were: *ceh-36*p*::gfp* (*oyIs48*), *ceh-36*p*::dsRed* (*otIs151*), *flp-7*p*::gfp* (*ynIs66*), *flp-10*p*::gfp* (*otIs92*), *gcy-8*p*::gfp* (*oyIs17*), *gpa-9*p*::gfp (pkIs586)*, *odr-1*p*::dsRed* (*oyIs44*), *odr-10*p*::gfp* (*kyIs37*), *osm-10*p*::gfp* (*nuIs11*), *sra-6*p*::gfp* (*oyIs14*), *sre-1*p*::gfp* (*otIs24*), *srh-142*p*::dsRed* (*oyIs51*), *srv-33*p*::gfp* (*oyIs47*), *osm-9*p*::gfp* (*oyIs1*), *tbh-1*p*::gfp (nIs107)*, *ttx-3*p*::gfp* (*mgIs18*) and *unc-129*p*::gfp* (*evIs82b*). Double-mutant strains were constructed using standard methods and verified by sequencing or phenotypic analyses.

Isolation and identification of *mls-2(oy88)*

Wild-type or *che-1(p674)* strains carrying stably integrated copies of the *ceh-36*p*::gfp* (*oyIs48*) fusion gene were mutagenized using EMS, and their F2 progeny were examined for mutants exhibiting altered *gfp* expression in the AWC neurons. The *oy88* allele was isolated from a screen of ~14,000 haploid genomes. *oy88* was mapped to LG X, and fine-mapped between the pKP6149 and pKP6106 SNP markers. The gene expression defect in *oy88* mutants was rescued by the C36E6 cosmid in this interval, and *oy88* was confirmed to be an allele of *mls-2* based on genotypic and phenotypic analyses and sequencing. *mls-2(oy88)* was outcrossed at least three times prior to further analyses. This screen also yielded multiple alleles of *lim-4*.

Microscopy

Fluorescence images were captured with a Zeiss Axioplan microscope equipped with a CCD camera (Hamamatsu), 40× (1.30 NA) or 63× (1.25 NA) objectives and OpenLab 4.0 software (Improvision).

Molecular biology

The *mls-2*p*::gfp*, *mls-2*p*::mls-2::mls-2* 3′UTR, and *mls-2*p*::gfp::mls-2::mls-2* 3′UTR constructs were provided by Jun Liu (Jiang et al., 2005; Jiang et al., 2008). The *osm-6*p*::mls-2::mls-2* 3′UTR misexpression construct was created by replacing the *mls-2* promoter with the *osm-6* promoter. Constructs were injected at 10 ng/μl with *unc-122*p*::dsRed* or *odr-1*p*::dsRed* as the co-injection marker.

Fragments of *ceh-36* upstream regulatory sequences were obtained by either restriction digestion or PCR amplification, and inserted upstream of *gfp* or *dsRed* coding sequences in a *C. elegans* expression vector (gift of A. Fire, Stanford University). Mutations in candidate cis-regulatory elements in *ceh-36* or *odr-1* upstream sequences were generated using the QuikChange Site-Directed Mutagenesis Kit (Stratagene). All constructs were injected at 50 ng/μl with *unc-122*p*::dsRed* or *unc-122*p*::gfp* as the co-injection marker.

Identification of CEH-36 and MLS-2 binding sites

Identified binding sites for HMX-3/NKX (Amendt et al., 1999) or CEH-36 (Etchberger et al., 2009) were used in a simple matrix-based motif search program, Possum (http://zlab.bu.edu/~mfrith/possum; threshold set at 5), to identify candidate motifs in *ceh-36* and *odr-1* upstream regulatory sequences. Only one candidate site each for MLS-2 and CEH-36 was identified in ~830 bp of *ceh-36* upstream regulatory sequences. Two candidate CEH-36 binding sites were identified in ~2387 bp of *odr-1* regulatory sequences.

Mosaic analysis

The rescuing cosmid C39E6 and the *odr-1*p::*dsRed* and *unc-122*p*::dsRed* constructs were co-injected into the *che-1(p674); ceh-36::gfp(oyIs48); mls-2(oy88)* strain. Transgenic animals expressing *unc-122*p*::dsRed* were selected and examined under a Zeiss Axioplan microscope using a 40× (1.30 NA) objective. Expression of *odr-1*p*::dsRed* in the AWC neurons indicated the presence of the injected array containing C39E6 in the AWC neurons.

Electrophoretic mobility shift assays (EMSAs)

A full-length *mls-2* cDNA was amplified from an oligo(dT)-primed reverse transcription reaction using total RNA extracted from mixed stage animals, cloned into pGEX4T-1 (gift of J. Ahnn, Hanyang University) and expressed in *E. coli* JM109 cells. GST-MLS-2 fusion protein was purified using the RediPack GST Purification Module following the manufacturer's protocols (GE Healthcare). EMSA was performed using the LightShift Chemiluminescent EMSA Kit (Pierce). GST-MLS-2 fusion protein (100 μg) was incubated with biotin-labeled wild-type or mutant binding site sequences for 30 minutes, followed by electrophoresis in a 5% polyacrylamide gel. For competition assays, excess wild-type or mutant binding site sequences were incubated with protein for 10 minutes prior to addition of labeled wild-type probe. Binding site sequences used were: wild type, 5′-CAAATGGGCG-3′; mutant A, 5′-CGGACGGGCG-3′ (mutated bases are underlined).

RESULTS

The *oy88* mutation affects expression of the *ceh-36* Otx homeobox gene in AWC sensory neurons

The Otx homeobox gene *ceh-36* regulates the expression of terminal differentiation genes in the AWC polymodal sensory neuron type, and is expressed selectively in the AWC and ASE sensory neurons in adults (Lanjuin et al., 2003; Koga and Ohshima, 2004). To identify genes that act upstream of *ceh-36* in the development of the AWC neurons, we mutagenized animals carrying stably integrated copies of a *ceh-36*p*::gfp* fusion gene (p, promoter), and examined their F2 progeny for mutants exhibiting altered *gfp* expression in the

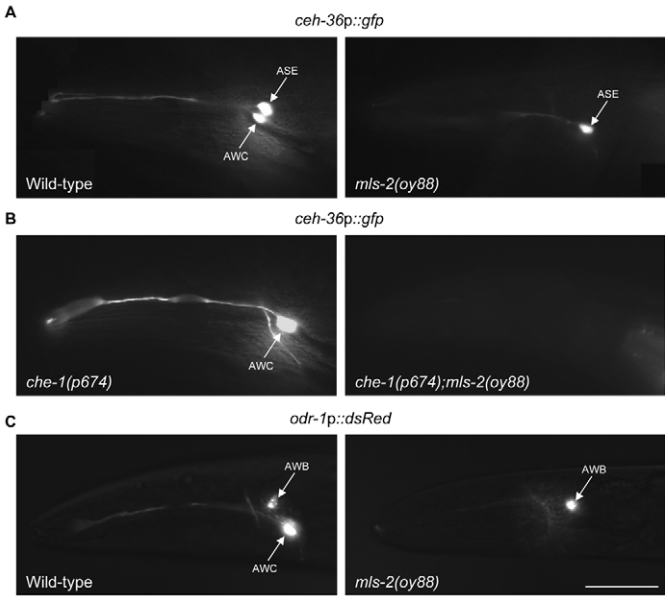


Fig. 1. The *C. elegans oy88* mutation affects gene expression in the AWC neurons. (**A**,**B**) Expression of a *ceh-36*p*::gfp* fusion gene is abolished in the AWC neurons in *oy88* mutants in a wild-type (A) or *che-1(p674)* mutant (B) background. Mutations in the *che-1* zinc-finger transcription factor abolish *ceh-36*p*::gfp* expression in the ASE neurons (Chang et al., 2003). (**C**) Expression of an *odr-1*p*::dsRed* transgene is affected in the AWC neurons in *oy88* mutants. Anterior is to left. Scale bar: 10 μm.

AWC neurons. This screen identified the *oy88* mutant, in which expression of *ceh-36*p*::gfp* was affected in the AWC, but not in the ASE, neurons (Fig. 1A). *gfp* expression was completely abolished in at least one of the bilateral AWC neuron pair in ~50% of *oy88* mutants (Table 1). We noted, however, that ~60% of the subset of AWC neurons that retained marker expression in *oy88* animals nevertheless exhibited multiple morphological defects. These defects included truncated dendrites, axonal branching and abnormal axonal trajectories, as well as altered cell body positions (see Fig. S1 in the supplementary material; data not shown), suggesting that multiple aspects of AWC identity are affected in *oy88* mutants.

To confirm that expression was specifically affected in the AWC, but not ASE, neurons, we examined *ceh-36*p*::gfp* expression in *che-1(p674); oy88* double mutants. *che-1* encodes a zinc-finger TF that regulates *ceh-36* expression in the ASE neurons, and is required for terminal differentiation of the bilateral ASE neuron pair (Chang et al., 2003; Uchida et al., 2003; Etchberger et al., 2007; Etchberger et al., 2009). Expression of *ceh-36*p*::gfp* was abolished in both neuron types in ~50% of *che-1; oy88* mutants (Fig. 1B; Table 1), verifying that the *oy88* mutation affects *ceh-36* expression in the AWC, but not ASE, neurons. Since CEH-36 regulates expression of AWC-specific terminal differentiation genes, such as the *odr-1* receptor guanylyl cyclase (L'Etoile and Bargmann, 2000; Lanjuin et al., 2003), we would expect that the expression of these genes would also be affected in *oy88* mutants. Consistent with this notion, expression of an *odr-1*p*::dsRed* transgene was also affected in the AWC neurons in *oy88* mutants (Fig. 1C; Table 1).

To determine whether *oy88* affects only AWC-specific characteristics, or whether general sensory or pan-neuronal features are also altered, we attempted to examine expression of the *osm-6*p*::gfp* and *rab-3*p*::rfp* fusion genes in the AWC neurons in *oy88* mutants. *osm-6* encodes a component of the intraflagellar transport complex that is required for ciliogenesis in all sensory neurons in *C. elegans* (Perkins et al., 1986; Collet et al., 1998), whereas *rab-3* encodes a small GTPase that is expressed pan-neuronally and is required for synaptic transmission (Nonet et al., 1997). However, owing to mispositioning of the AWC neurons in *oy88* mutants (see Fig. S1 in the supplementary material) (Yoshimura et al., 2008), we could not reliably identify the AWC neuronal cell bodies, and were thus unable to confirm whether expression of the examined markers was altered in *oy88* mutants.

oy88 is an allele of the Hmx/Nkx-type homeobox gene *mls-2*

Mapping with respect to genetic markers and single-nucleotide polymorphisms (see Materials and methods) placed *oy88* in a 600 kb region on LG X. This region contains the Hmx/Nkx-type homeobox gene *mls-2*, in which mutations were initially identified in a screen for genes that regulate postembryonic lineage patterning of the M mesodermal precursor (Jiang et al., 2005). Recently, *mls-2* has also been implicated in the differentiation of the ventral cephalic sheath (CEPshV) glial cells in the head (Yoshimura et al., 2008). It was noted that mutations in *mls-2* affect dendritic extension, axonal morphology, and the trajectories of a subset of sensory neurons,

Table 1. Expression of genes in the AWC neurons is affected in *oy88* mutants

		***gfp* expression in AWC neurons (%)**		
Strain	Marker	Both	One	None
Wild type	*ceh-36*p*::gfp*	100	0	0
mls-2(oy88)	*ceh-36*p*::gfp*	45	43	12
mls-2(oy88); Ex[*C39E6*]*	*ceh-36*p*::gfp*	82	16	2
mls-2(cc615)	*ceh-36*p*::gfp*	42	48	10
mls-2(tm252)	*ceh-36*p*::gfp*	94	3	3
Wild type	*odr-1*p*::dsRed*	100	0	0
mls-2(oy88)	*odr-1*p*::dsRed*	35	44	21
mls-2(cc615)	*odr-1*p*::dsRed*	44	43	13
mls-2(ns156)	*odr-1*p*::dsRed*	43	45	12
mls-2(tm252)	*odr-1*p*::dsRed*	82	16	2
che-1(p674)	*ceh-36*p*::gfp*	100	0	0
che-1(p674); mls-2(oy88)	*ceh-36*p*::gfp*	43	43	14

Expression of stably integrated transgenes was examined in adult animals (*n*>50 for each strain). Animals were grown at 25°C and examined at 400× magnification.
*The *C39E6* cosmid contains *mls-2* genomic sequences, including upstream regulatory sequences.

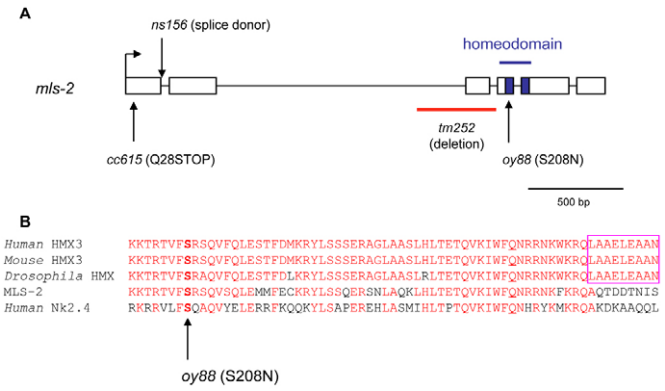


Fig. 2. *oy88* is an allele of the *mls-2* Hmx/Nkx homeobox transcription factor gene. (**A**) Genomic structure of *C. elegans mls-2* showing the molecular nature of the lesions in *mls-2* alleles examined in this work. Exons are indicated by boxes, introns by lines. (**B**) Alignment of the homeodomains of the indicated proteins, including *C. elegans* MLS-2. Residues that are identical in at least two proteins are in red. The residue mutated in *oy88* is indicated. The Q50 residue in the third helix that is crucial for conferring DNA-binding specificity is underlined. The C-terminal HMX motif that is shared among HMX, but not NKX, proteins is boxed.

including the AWC neurons, as well as gene expression in the AWC neurons (Yoshimura et al., 2008), suggesting that *oy88* might be an allele of *mls-2*.

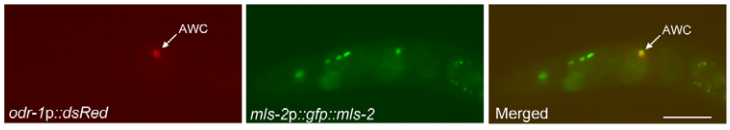
We verified this hypothesis by demonstrating rescue of AWC gene expression defects of *oy88* mutants with a cosmid containing *mls-2* genomic sequences (Table 1). Moreover, the molecular lesion in *oy88* was identified as a missense mutation in a highly conserved residue in the MLS-2 homeodomain (Fig. 2A,B), and animals carrying the previously identified *ns156* (Yoshimura et al., 2008) and *cc615* (Jiang et al., 2005) (Fig. 2A) alleles of *mls-2* also showed defects in *ceh-36*p*::gfp* and *odr-1*p*::dsRed* expression, similar to those observed in *oy88* mutants (Table 1). However, only minor defects were observed in animals carrying the *tm252* semi-dominant gain-of-function allele (Fig. 2A; Table 1). The penetrance of the AWC gene expression phenotype was similar in animals carrying the *oy88* missense and the *cc615* candidate null alleles (Table 1), suggesting that the homeodomain is essential for MLS-2 function in the AWC neurons. These findings indicate that *oy88* is an allele of *mls-2*, and that MLS-2 regulates differentiation of the AWC neurons.

mls-2 is expressed transiently in AWC neurons and functions in this neuron type to regulate their differentiation

Examination of the expression pattern of a rescuing *gfp*-tagged *mls-2* transgene, as well as staining with anti-MLS-2 antibodies, has shown that during embryonic development, MLS-2 is localized to the nuclei of a subset of AB precursor-derived neuronal and non-neuronal cells (Jiang et al., 2005; Yoshimura et al., 2008). However, postembryonic expression of *mls-2* appears to be restricted to three neuronal pairs in the head, the somatic gonad, the M cells and their descendants (Jiang et al., 2005; Jiang et al., 2008). Two of the three MLS-2-expressing neurons have been identified as the ASK chemosensory and the AIM interneurons (Jiang et al., 2008). Consistent with these published observations, we did not detect *mls-2* expression in the AWC neurons in adult animals (data not shown).

Transcription factors that play important roles in the regulation of differentiation of individual neuron types can be expressed transiently following birth of the postmitotic neuron (Finney et al., 1988;

A



B

	AWCL	AWCR	AWCL	AWCR
mls-2 rescuing sequences in :	present	present	absent	absent
Wild-type *ceh-36*p::*gfp* expression	100 %	100 %	44 %	50 %
n	50	51	50	50

C

	AWCL	AWCR	AWCL	AWCR
mls-2 rescuing sequences in :	present	present	absent	absent
Wild-type AWC morphology	99 %	99 %	46 %	48 %
n	300	300	50	50

D

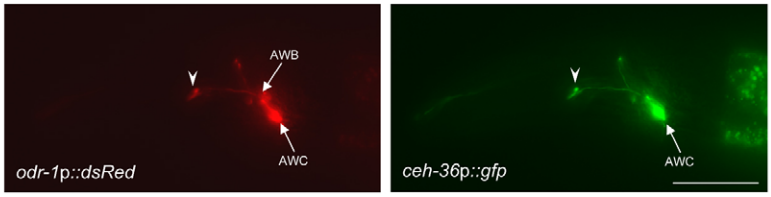


Fig. 3. *mls-2* is expressed transiently in, and acts in, the AWC neurons to regulate gene expression. (**A**) Expression of *odr-1*p*::dsRed* (left) and a *gfp*-tagged *mls-2* gene driven under its own regulatory sequences (middle) in wild-type L1 larvae. Merged image is on the right. The AWC neurons were identified by their position and *odr-1*p*::dsRed* expression. (**B**) Correlation of the presence of *mls-2* rescuing sequences with wild-type *ceh-36*p*::gfp* expression in *che-1(p674); mls-2(oy88)* mutants carrying stably integrated copies of the *ceh-36*p*::gfp* fusion gene. *mls-2* rescuing sequences in the C39E6 cosmid were present on unstable extrachromosomal arrays that also contain *odr-1*p*::dsRed* sequences. The presence of *mls-2* sequences in the AWC neurons was monitored by *odr-1*p*::dsRed* expression. AWCL, left; AWLR, right. (**C**) Correlation of the presence of *mls-2* rescuing sequences with wild-type AWC neuronal morphology in *che-1(p674); mls-2(oy88)* mutants carrying stably integrated copies of the *ceh-36*p*::gfp* fusion gene. The presence of *mls-2* sequences in the AWC neurons was monitored as described in B. (**D**) An AWC neuron in *che-1(p674); mls-2(oy88)* animals containing wild-type *mls-2* sequences as monitored by the expression of *odr-1*p*::dsRed* (left) exhibits a truncated dendrite (arrowhead). The truncated dendrite is also visualized via expression of *ceh-36*p*::gfp* (right). Scale bars: 5 μm in A: 10 μm in D.

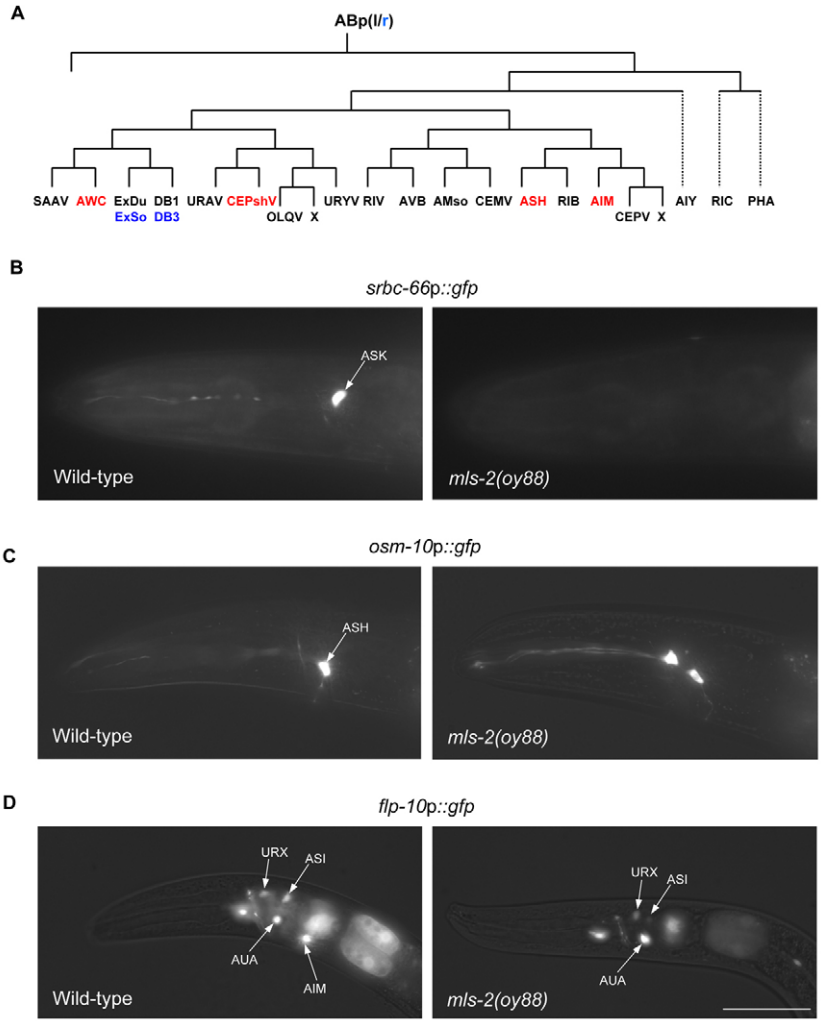


Fig. 4. MLS-2 regulates the differentiation of multiple ABp(l/r)paa-derived cell types. (**A**) Lineage diagram of ABp(l/r)-derived cells (not to temporal scale) (see Sulston et al., 1983). Left/right (l/r) asymmetric cells derived from ABpr are shown in blue. Cells affected in *mls-2* mutants are in red. Dotted lines indicate lineages for which not all cell divisions are shown. (**B**) Expression of the ASK marker *srbc-66*p*::gfp* (Kim et al., 2009) is abolished in *mls-2(oy88)* mutants. (**C**) The ASH sensory neuronal marker *osm-10*p*::gfp* (Hart et al., 1999) is ectopically expressed in *mls-2(oy88)* mutants. (**D**) Expression of *flp-10*p*::gfp* (Kim and Li, 2004) is abolished in the AIM interneurons in *mls-2(oy88)* animals. Scale bar: 10 μm.

Roztocil et al., 1997; Sarafi-Reinach et al., 2001; Lanjuin et al., 2003; Wernet et al., 2006; Sarin et al., 2009). We therefore investigated whether *mls-2* expression could be detected early in development in the postmitotic AWC neurons. Indeed, we observed *mls-2* expression in the AWC neurons in L1 larvae shortly after hatching (Fig. 3A); expression was extinguished by the late L1 larval stage (not shown).

We next performed mosaic analysis to determine the site of action of MLS-2 in the regulation of AWC gene expression. We generated *che-1(p674); mls-2(oy88)* animals carrying stably integrated copies of the *ceh-36*p*::gfp* transgene to assess rescue of the AWC developmental defects. These animals also carried unstable extrachromosomal arrays containing rescuing *mls-2* sequences on a cosmid, as well as the *odr-1*p*::dsRed* transgene as a marker for the presence of the array in the AWC neurons. Since MLS-2 regulates *odr-1* expression, lack of *odr-1*p*::dsRed* expression in the AWC neurons could be due to either absence of the array in the AWC neurons, or to failure of wild-type *mls-2* sequences to rescue the mutant phenotype. However, correlation of the presence of rescuing *mls-2* sequences in the AWC neurons, as assessed by expression of *odr-1*p*::dsRed* from the array, with *ceh-36*p*::gfp* expression in the AWC neurons showed that when the array was present, 100% of the AWC neurons expressed *ceh-36*p*::gfp* (Fig. 3B). These observations suggest that *mls-2* acts in the AWC lineage, and probably cell autonomously in the AWC neurons, to regulate the expression of terminal differentiation genes.

In addition to affecting differentiation of the AWC neurons, mutations in *mls-2* also affect CEPshV glial cell differentiation (Yoshimura et al., 2008). Ablation of the CEPshV glial cells has previously been shown to result in defects in AWC morphology but not in gene expression (Yoshimura et al., 2008) (S. Yoshimura and S. Shaham, personal communication), suggesting that *mls-2* affects AWC morphology non-autonomously, probably via its functions in the CEPshV cells. The CEPshV glial cells are close lineal relatives of the AWC neurons (Sulston et al., 1983), such that distinguishing between these two cell types as the site of action of *mls-2* is challenging. However, we identified rare animals in which the AWC neurons retained altered cell body positions despite containing rescuing *mls-2* sequences (Fig. 3C,D), supporting the notion that MLS-2 acts in a close lineal relative, such as the CEPshV cells, to regulate AWC neuronal morphology.

The AWC neurons do not appear to adopt an alternate fate in *mls-2* mutants

Transcription factors that promote a specific neuronal cell fate can repress the expression of characteristics of another cell type, which may or may not be lineally related (Miller and Niemeyer, 1995; Sagasti et al., 1999; Thaler et al., 1999; Muhr et al., 2001; Esmaeili et al., 2002). To determine whether the AWC neurons in *mls-2* mutants adopt the characteristics of another cell type, we examined the expression of markers for cells that are related by lineage

Table 2. The differentiation of additional, lineally related cell types is affected in *mls-2* mutants

Strain	Marker*	Cells examined	Expression of *gfp* in number of neurons (%)			
			Both	One	None	Other
Wild type	*odr-10*p::*gfp*	AWA	100	0	0	–
mls-2(oy88)	*odr-10*p::*gfp*	AWA	100	0	0	–
Wild type	*odr-1*p::*dsRed*	AWB	100	0	0	–
mls-2(oy88)	*odr-1*p::*dsRed*	AWB	100	0	0	–
mls-2(cc615)	*odr-1*p::*dsRed*	AWB	100	0	0	–
mls-2(ns156)	*odr-1*p::*dsRed*	AWB	100	0	0	–
mls-2(tm252)	*odr-1*p::*dsRed*	AWB	100	0	0	–
Wild type	*gcy-8*p::*gfp*	AFD	100	0	0	–
mls-2(oy88)	*gcy-8*p::*gfp*	AFD	100	0	0	–
Wild type	*ceh-36*p::*gfp*	ASE	100	0	0	–
mls-2(oy88)	*ceh-36*p::*gfp*	ASE	100	0	0	–
mls-2(cc615)	*ceh-36*p::*gfp*	ASE	100	0	0	–
mls-2(tm252)	*ceh-36*p::*gfp*	ASE	100	0	0	–
Wild type	*srh-142*p::*dsRed*	ADF	100	0	0	–
mls-2(oy88)	*srh-142*p::*dsRed*	ADF	100	0	0	–
Wild type	*srv-33*p::*gfp*	ASG	100	0	0	–
mls-2(oy88)	*srv-33*p::*gfp*	ASG	100	0	0	–
Wild type	*osm-10*p::*gfp*	ASH	100	0	0	–
mls-2(oy88)	*osm-10*p::*gfp*	ASH	18	0	0	82 (ectopic)
Wild type	*sra-6*p::*gfp*	ASH	100	0	0	–
mls-2(oy88)	*sra-6*p::*gfp*	ASH	33	0	0	67 (ectopic)
Wild type	*sra-6*p::*gfp*	ASI	100	0	0	–
mls-2(oy88)	*sra-6*p::*gfp*	ASI	100	0	0	–
Wild type	*gpa-9*p::*gfp*	ASJ	100	0	0	–
mls-2(oy88)	*gpa-9*p::*gfp*	ASJ	100	0	0	–
Wild type	*srbc-66*p::*gfp*	ASK	87	11	2	–
mls-2(oy88)	*srbc-66*p::*gfp*	ASK	0	0	100	–
Wild type	*sre-1*p::*gfp*	ADL	100	0	0	–
mls-2(oy88)	*sre-1p*::*gfp*	ADL	100	0	0	–
Wild type	*flp-7*p::*gfp*	SAAV	100	0	0	–
mls-2(oy88)	*flp-7*p::*gfp*	SAAV	100	0	0	–
Wild type	*osm-9*p::*gfp*	OLQV	100	0	0	–
mls-2(oy88)	*osm-9*p::*gfp*	OLQV	100	0	0	–
Wild type	*flp-10*p::*gfp*	AIM	100	0	0	–
mls-2(oy88)	*flp-10*p::*gfp*	AIM	0	0	100	–
Wild type	*ttx-3*p::*gfp*	AIY	100	0	0	–
mls-2(oy88)	*ttx-3*p::*gfp*	AIY	100	0	0	–
Wild type	*tbh-1*p::*gfp*	RIC	100	0	0	
mls-2(oy88)	*tbh-1*p::*gfp*	RIC	100	0	0	
Wild type	*osm-10*p::*gfp*	PHA	100	0	0	–
mls-2(oy88)	*osm-10*p::*gfp*	PHA	100	0	0	–
Wild type	*unc-129*p::*gfp*	DAs, DBs	–	–	–	100†
mls-2(oy88)	*unc-129*p::*gfp*	DAs, DBs	–	–	–	100†

Animals were grown at 25°C and adults were examined at 400× magnification. *n*>50 for each strain.
*The expression of stably integrated transgenes was examined in all cases except for *srbc-66*p::*gfp*, which was present on an extrachromosomal array.
†*unc-129*p::*gfp* is expressed in multiple neuronal and non-neuronal cell types (Colavita et al., 1998). Only expression in the DA and DB motoneurons was quantified.

(Fig. 4A) and/or function to the AWC neurons. We first determined whether the expression of markers specific for other sensory neurons in the head amphid sensory organ was affected in *mls-2* mutants. No effects were observed on the expression of markers for nine of the remaining 11 amphid sensory neurons. However, the expression of a marker for the ASK neurons, which express *mls-2* throughout postembryonic development, was completely abolished (Fig. 4B; Table 2) (Jiang et al., 2008). In addition, markers for the ASH polymodal sensory neuron type were expressed ectopically in at least one additional neuron in ~60-80% of *mls-2(oy88)* mutants (Fig. 4C; Table 2). Six amphid sensory neuronal cells, including the ASH and ASK neurons, fill with lipophilic dyes such as DiO (Perkins et al., 1986; Herman and Hedgecock, 1990). Consistent with the presence of ectopic ASH-like neurons, we observed more than six dye-filled cell pairs in *mls-2(oy88)* mutants (see Fig. S2 in the supplementary material). However, the ASK neurons continued to

dye-fill and exhibited wild-type morphology (see Fig. S2 in the supplementary material), suggesting that mutations in *mls-2* affect a subset of ASK differentiated features.

Since both the ASH and AWC neurons are generated from the ABp(l/r)paa precursors (Sulston et al., 1983) (Fig. 4A), we considered the possibility that the AWC neurons were adopting the fate of the ASH neurons in *mls-2* mutants. To examine this notion, we correlated ectopic expression of the ASH neuronal marker *osm-10*p*::gfp* (Hart et al., 1999) with the AWC neuronal marker *odr-1*p*::dsRed* in individual animals. However, we found no correlation (see Table S1 in the supplementary material), suggesting that the ectopic ASH-like cells in *mls-2(oy88)* mutants do not arise from fate transformation of the AWC neurons.

Additional ABp(l/r)paa-derived cells, the fates of which are affected in *mls-2* mutants, include the CEPshV cells (Yoshimura et al., 2008) (Fig. 4A). To investigate whether the AWC neurons adopted the fate of other cells arising from these precursors, we examined markers for the SAAV interneurons, which are lineal sisters of the AWC neurons, the ventral OLQ sensory neurons, the DB1 and DB3 motoneurons, and the AIM interneurons, which express *mls-2* throughout development. In addition, we examined the expression of markers for the more distantly related AIY, RIC and PHA neurons in *mls-2(oy88)* mutants (Fig. 4A). Although the expression of the majority of the markers examined was unaffected in *mls-2* mutants (Table 2), 100% of *mls-2(oy88)* mutants lacked expression of the AIM interneuronal marker *flp-10*p*::gfp* (Fig. 4D; Table 2) (Kim and Li, 2004). Thus, MLS-2 regulates fate specification of multiple cell types arising from the ABp(l/r)paa precursors (summarized in Fig. 4A). In addition, MLS-2 affects differentiation of the ABa(l/r)pp-derived ASK cell type. Our analysis also suggests that the AWC neurons do not appear to adopt the fate of the examined lineally or functionally related neurons in *mls-2* mutants.

Mutations in *mls-2* affect cell proliferation and cell cleavage orientations in the postembryonic M lineage (Jiang et al., 2005), raising the possibility that alterations in embryonic lineage patterns might similarly give rise to the observed neuronal defects. However, analysis of the ABp(l/r)paa-derived lineages by automated 4D lineage tracking (Bao et al., 2006) did not reveal any obvious defects in the sublineages that give rise to the AWC, ASH or AIM neurons (J. Murray, personal communication) (Yoshimura et al., 2008). These results suggest that the defects observed in the development and differentiation of multiple embryonically derived cell types in *mls-2* mutants might arise as a consequence of independent MLS-2 requirements in individual sublineages.

Spatiotemporal misexpression of *mls-2* is not sufficient to drive ectopic expression of *ceh-36*p*::gfp*

We asked whether misexpression of *mls-2* in other cell types is sufficient to drive *ceh-36*p*::gfp* expression. However, no ectopic *ceh-36*p*::gfp* expression was observed upon misexpression of *mls-2* under the *osm-6* promoter, which drives expression in all ciliated neurons (Collet et al., 1998) (see Fig. S3 in the supplementary material). This result indicates that under the experimental conditions used, MLS-2 is not sufficient to activate *ceh-36* expression in the context of other ciliated neuron types. However, we noted that *ceh-36*p*::gfp* expression in the AWC, but not ASE, neurons was downregulated in 10% of animals (n=80) expressing *osm-6*p*::mls-2* (see Fig. S3 in the supplementary material). Expression of the transiently expressed LIM-homeobox gene *lin-11* under the *osm-6* promoter has similarly been shown to downregulate expression of the putative terminal selector TF gene *odr-7* in the AWA olfactory neurons (Sarafi-Reinach et al., 2001). Since the *osm-6* promoter drives expression through adult stages (Collet et al., 1998), these observations suggest that temporal misregulation of *mls-2* expression might alter AWC-specific gene expression.

We next examined whether CEH-36, once expressed, feeds back directly or indirectly to repress *mls-2* expression in the AWC neurons. However, *mls-2* expression was still extinguished after the L1 stage in the AWC neurons in *ceh-36* mutants [no *ceh-36(ky640)* mutants expressed *mls-2*p*::gfp* in the AWC neurons in older larvae and adults; n>50], suggesting that alternate mechanisms must ensure the downregulation of *mls-2* expression in older animals.

MLS-2 and CEH-36 directly initiate and maintain *ceh-36* expression in the AWC neurons

A model consistent with the transient expression of *mls-2* in the AWC neurons suggests that MLS-2 activates *ceh-36* expression, which subsequently maintains its expression, perhaps via autoregulation. Regulation of expression by MLS-2 and CEH-36 could be direct or indirect. We first confirmed that CEH-36 autoregulates to maintain its own expression by determining that *ceh-36*p*::gfp* expression is not maintained in AWC neurons in *ceh-36* mutants (Fig. 5A).

We next identified the minimal cis-regulatory sequences required to drive *ceh-36* expression in the AWC neurons. The expression of *gfp* driven by different extents of *ceh-36* upstream regulatory sequence was examined in wild-type animals transgenic for promoter*::gfp* fusion constructs. These analyses identified a 33 bp sequence located ~200 bp upstream of the initiator ATG, deletion of which resulted in loss of *gfp* expression in adult AWC neurons (Fig. 5B; compare the expression driven by 238 bp versus 205 bp of upstream regulatory sequence). We identified a motif previously shown to be bound by CEH-36 in vitro (Etchberger et al., 2009) within this 33 bp sequence (Fig. 5B,C), suggesting that CEH-36 might bind to this motif to maintain expression. Consistent with this hypothesis, mutating this motif resulted in loss of *gfp* expression in adults, but not in larvae (Fig. 5B,C). This motif was fully conserved in the regulatory sequences of the *ceh-36* ortholog in the related nematode *C. briggsae* (Fig. 5C), although the motif was located further upstream than in *C. elegans*.

Since sequences lacking the CEH-36 binding site continued to drive expression in the AWC neurons in L1 larvae, we reasoned that these sequences contain a putative MLS-2 binding motif. The sequence of the MLS-2 homeodomain is highly conserved with those of other HMX/NKX homeodomain proteins (Fig. 2B) and includes the Q50 residue in the third helix that is crucial for DNA-binding specificity (Treisman et al., 1989; Gehring et al., 1994). Thus, MLS-2 is expected to bind regulatory motifs recognized by members of the HMX/NKX TF family. We identified a motif previously shown to be recognized by HMX/NKX homeodomain proteins (Okkema and Fire, 1994; Amendt et al., 1999; Wang et al., 2002; Wang et al., 2005) located ~20 bp upstream of the ATG in the *mls-2* genomic sequence (Fig. 5B,C). This motif was also present in the upstream sequence of the *C. briggsae ceh-36* ortholog (Fig. 5C). Mutation of this motif abolished *ceh-36*p*::gfp* expression in larvae (Fig. 5B,C). To determine whether MLS-2 directly bound this motif, we performed electrophoretic mobility shift (EMSA) assays. Bacterially produced MLS-2 protein bound directly to the wild-type sequences, as indicated by the retarded mobility of the complex (Fig. 5D), and was more effectively competed by excess wild-type than mutant sequences (Fig. 5D). Thus, MLS-2 directly binds the identified motif to initiate *ceh-36* expression in the AWC neurons.

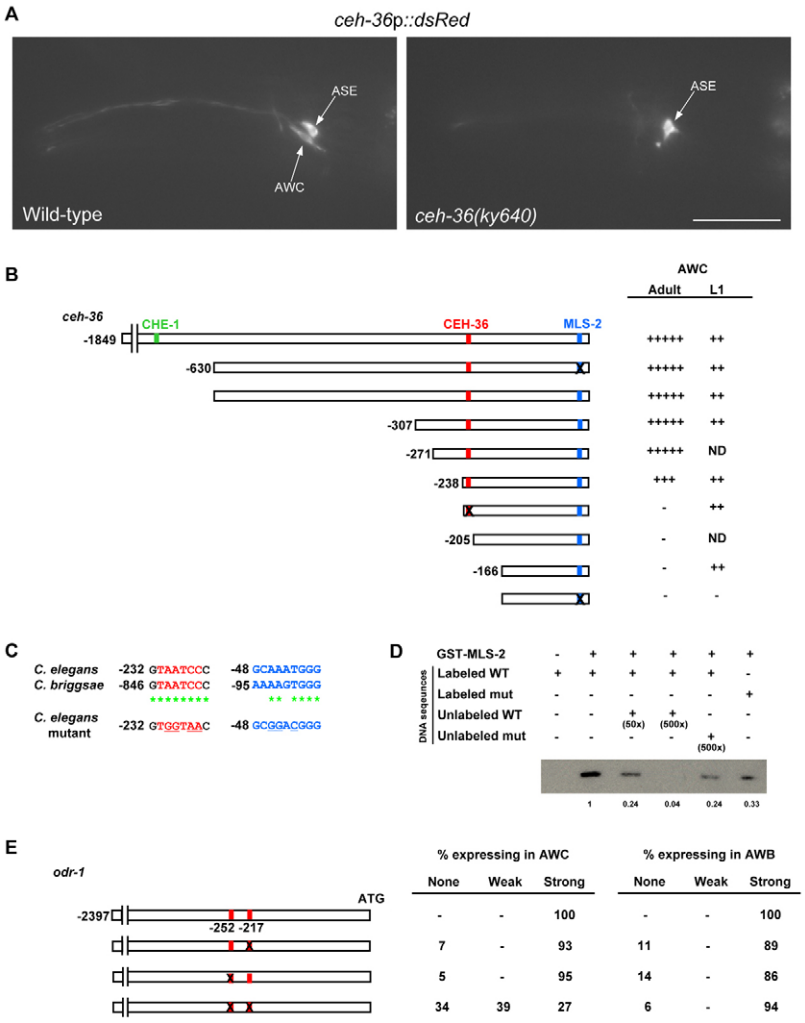


Fig. 5. MLS-2 and CEH-36 interact directly with *ceh-36* upstream regulatory sequences to initiate and maintain *ceh-36* expression, respectively. (**A**) CEH-36 autoregulates to maintain its own expression in the AWC neurons. Forty-seven percent of *ceh-36(ky640)* animals failed to express a *ceh-36p::dsRed* transgene in the AWC neurons. Expression in the ASE neurons was unaffected. *n*=66 animals. Scale bar: 10 μm. (**B**) The extent of *ceh-36* upstream regulatory sequences driving *gfp* expression in the different constructs. The numbering is relative to the initiator ATG. Motifs that interact directly with CEH-36 and MLS-2 are indicated by red and blue boxes, respectively. The CHE-1 binding site is indicated by a green box (Etchberger et al., 2007). Mutated motifs are indicated by an X; the sequences of the mutated motifs are shown in C. The strength of expression is indicated by the number of + symbols (–, no expression detected; ND, not determined). Note that the effects of mutations in the MLS-2 binding motif on transgene-driven gene expression in wild-type AWC neurons can only be observed in the absence of a functional CEH-36 binding site in both L1 larvae and adults. Larval expression driven by sequences containing a functional CEH-36 binding site, but mutated MLS-2 binding site, is likely to be due to the temporal overlap in *mls-2* and *ceh-36* expression in newly hatched animals. At least 50 animals from two independent transgenic lines carrying each construct were examined. (**C**) Conservation of the CEH-36 (red) and MLS-2 (blue) binding sites between *C. elegans* and *C. briggsae*. Asterisks indicate identical nucleotides. Point mutations in the predicted binding sites analyzed (in B and D) are underlined. (**D**) Binding of bacterially expressed GST-MLS-2 fusion protein to the predicted MLS-2 binding site upstream of *ceh-36* as examined by EMSA. Competition was performed with the indicated excess of unlabeled wild-type or mutant MLS-2 binding sequences shown in C. Numbers below each lane in the gel indicate relative band intensities, with the intensity of the band representing GFP-MLS-2 bound to wild-type sequences set at 1. (**E**) Mutation of two predicted CEH-36 binding sites (red boxes; mutation indicated by X) in *odr-1* upstream regulatory sequences affects *odr-1* expression in the AWC neurons. At least 50 animals from two independent transgenic lines carrying each construct were examined.

These results indicate that multiple cis- and trans-regulatory modules direct correct spatiotemporal *ceh-36* expression. These include an MLS-2 binding motif that initiates *ceh-36* expression in the AWC neurons, a CEH-36 binding motif that maintains *ceh-36* expression, and a previously defined CHE-1 binding motif that initiates *ceh-36* expression in the ASE neurons (Chang et al., 2003; Etchberger et al., 2007).

CEH-36 may directly interact with the regulatory sequences of terminal differentiation genes in AWC neurons

We and others have previously shown that *ceh-36* regulates the expression of AWC-specific terminal differentiation genes such as *odr-1* and *tax-2* (Lanjuin et al., 2003; Koga and Ohshima, 2004). Identification of a putative CEH-36 binding motif in the *ceh-36*

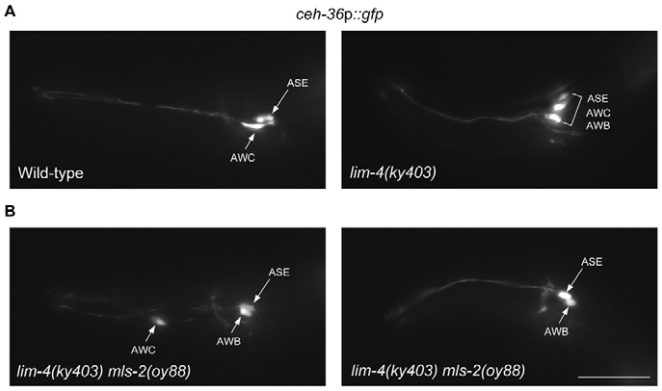


Fig. 6. *mls-2* is not required to drive misexpression of *ceh-36* in the AWB neurons in *lim-4* mutants. (**A**) *ceh-36*p::*gfp* is misexpressed in the AWB neurons in 100% of *lim-4(ky403)* mutants. *n*=50. (**B**) *ceh-36*p::*gfp* expression is retained in the AWB neurons in *lim-4(ky403) mls-2(oy88)* double mutants. *ceh-36*p::*gfp* expression in a mispositioned AWC neuron (left) and in the ASE and AWB neurons (left and right) is shown in *lim-4 mls-2* mutants. All *lim-4 mls-2* mutants retained *ceh-36*p::*gfp* expression in the AWB neurons; *n*=50. Anterior is to left. Scale bar: 10 μm.

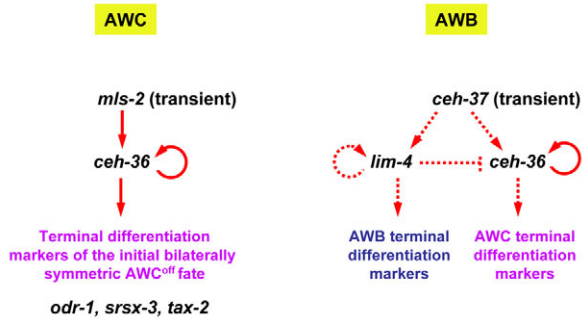


Fig. 7. Models of the developmental regulatory cascades that specify AWC and AWB neuronal identities in *C. elegans*. In the AWC neurons (left), transient expression of the MLS-2 HMX/NKX transcription factor directly initiates expression of the *ceh-36* Otx homeobox gene. CEH-36 maintains expression by autoregulation via direct interaction with its upstream regulatory sequences, and also directly regulates the expression of terminal differentiation genes, including *odr-1*, *srsx-3* and, possibly, *tax-2* (Koga and Ohshima, 2004), to specify the initial, bilaterally symmetric AWC^{OFF} fate. In the AWB neurons (right), transient expression of the CEH-37 OTX homeodomain protein initiates expression of the *lim-4* LIM-homeobox gene. LIM-4 maintains expression by autoregulation and promotes expression of AWB terminal differentiation genes. CEH-37 may also be able to promote expression of *ceh-36* in the AWB neurons, but this expression is normally repressed by LIM-4. In the absence of LIM-4 function, AWB-specific genes are not expressed and the expression of *ceh-36* is derepressed, resulting in the expression of AWC fate. Dotted lines indicate regulation that might be direct or indirect.

regulatory sequences prompted us to investigate whether a similar site was present upstream of *odr-1* and was required for *odr-1* expression in the AWC neurons. Examination of ~2.4 kb of upstream regulatory sequences previously shown to be sufficient to drive *odr-1* expression in the AWC neurons, as well as in other cell types (L'Etoile and Bargmann, 2000), identified two TAATCC motifs in close proximity located ~200-250 bp upstream of the initiator ATG (Fig. 5E). Whereas deletion of each motif did not affect *odr-1* expression in the AWC neurons, mutating both motifs significantly decreased *odr-1*p::*dsRed* expression in the AWC neurons without affecting expression in the AWB neurons (Fig. 5E). Deletion analyses previously also identified a candidate CEH-36 binding site upstream of the left/right asymmetrically expressed *srsx-3* chemoreceptor gene that is required to drive bilateral expression of *srsx-3* in the AWC neurons (Lesch et al., 2009) (see Discussion). These observations suggest that CEH-36 directly regulates the expression of terminal differentiation genes in the AWC neurons.

ceh-36 is misexpressed in the AWB neurons of *lim-4* mutants in an MLS-2-independent manner

The LIM-4 LIM-homeodomain TF regulates terminal differentiation of the AWB olfactory neuron type (Sagasti et al., 1999). In *lim-4* mutants, the AWB neurons lack AWB-specific characteristics, and instead express an AWC-like fate and morphology, suggesting that LIM-4 acts as a binary switch between the AWB and AWC cell fates. LIM-4 could repress the AWC fate via repression of either *mls-2* or *ceh-36* expression in the AWB neurons. Interestingly, we identified several *lim-4* alleles in the genetic screen for mutants exhibiting altered *ceh-36*p::*gfp* expression (see Materials and methods). In these mutants, we observed ectopic expression of *ceh-36*p::*gfp* in neurons identified as AWB neurons by their cell body position (Fig. 6A). Expression of *ceh-36*p::*gfp* in the AWC neurons was unaffected. To determine whether *mls-2* is required for this ectopic expression, we examined *ceh-36*p::*gfp* expression in *lim-4 mls-2* double mutants. We found that whereas, as expected, *ceh-36*p::*gfp* expression in the AWC neurons was affected, *ceh-36* continued to be misexpressed in the AWB neurons in these animals (Fig. 6B), suggesting that in the absence of LIM-4 function, *ceh-36* is misexpressed in the AWB neurons in an *mls-2*-independent manner.

DISCUSSION

We have identified a regulatory TF cascade that specifies the unique identity of the AWC sensory neuron type. Our experiments indicate that the *mls-2* Hmx/Nkx homeobox gene is expressed in the postmitotic AWC neurons but that this expression is extinguished shortly after hatching. Transient *mls-2* expression is sufficient to directly activate *ceh-36* Otx homeobox gene expression in the AWC neurons (Fig. 7). Once expression is initiated, CEH-36 autoregulates to maintain expression via direct interaction with its cis-regulatory sequences (Fig. 7). Together with our previous work (Lanjuin et al., 2003), our data suggest that CEH-36 is both necessary and sufficient to promote the AWC fate via direct regulation of subsets of downstream target genes. These target genes comprise terminal differentiation genes, such as *odr-1* and *srsx-3*, as well as the *tax-2* channel gene (Koga and Ohshima, 2004), but might also include additional downstream TFs that in turn regulate additional subsets of terminal differentiation genes (Hobert, 2008). MLS-2 also appears to regulate the differentiation of additional neuron types and, based on its maintained expression in these neuron types, might act directly as a terminal selector TF in these neurons. HMX/NKX proteins have similarly been shown to play diverse roles in neuronal and, particularly, sensory organ development in both vertebrates and *Drosophila* (Wang and Lufkin, 2005; Alenina et al., 2006; Dessaud et al., 2008).

The bilateral AWC neurons exhibit stochastic left/right asymmetry in the expression of chemoreceptor genes and in sensory functions (Troemel et al., 1999; Wes and Bargmann, 2001; Lesch et al., 2009). Thus, either the left or the right AWC neuron expresses the *str-2* chemoreceptor gene (the AWC^{ON} neuron), whereas the other AWC neuron expresses the *srsx-3* chemoreceptor (the AWC^{OFF} neuron) (Troemel et al., 1999; Lesch et al., 2009). *srsx-3* is initially expressed bilaterally in both AWC neurons (Lesch et al., 2009), indicating that the AWC^{OFF} state is the ground state. Experimental analyses have suggested that CEH-36 activates bilateral *srsx-3* expression via interaction with an OTX binding motif in *srsx-3* regulatory sequences (Lesch et al., 2009). Following this initial bilateral expression, *srsx-3* expression is then repressed in the AWC^{ON} neuron by the NSY-7 atypical homeodomain TF, which binds an overlapping motif, and the expression of which is activated in the AWC^{ON} neuron by neuronal signaling between the left and right postmitotic AWC neurons (Sagasti et al., 2001; Tanaka-Hino et al., 2002; Chuang and Bargmann, 2005; Vanhoven et al., 2006; Chuang et al., 2007; Lesch et al., 2009). Taken together with the results presented in this work, these findings suggest that the MLS-2/CEH-36 regulatory cascade is important for establishing the initial, bilaterally symmetric features of the AWC neurons, whereas additional mechanisms subsequently regulate left/right asymmetric gene expression. This regulatory scheme exhibits similarities to that used in the ASE chemosensory neurons, which also exhibit left/right asymmetry in gene expression and sensory functions (Yu et al., 1997; Pierce-Shimomura et al., 2001; Chang et al., 2003; Johnston and Hobert, 2003; Ortiz et al., 2006; Ortiz et al., 2009). Bilateral expression of genes in the ASE neurons is regulated by the CHE-1 TF, which also acts together with additional factors to restrict asymmetric gene expression to either the left or right ASE neuron (Etchberger et al., 2009). However, whereas CHE-1 has been proposed to act in a feed-forward loop (Alon, 2007) to regulate asymmetric gene expression, it is not yet clear whether CEH-36 is directly required to regulate all asymmetrically expressed genes in the AWC neurons.

The temporal features of the regulatory cascade that is utilized for cell fate specification in the AWC neurons exhibit similarities with cascades previously shown to specify the identities of other chemosensory neuron types in *C. elegans*. The *lin-11* LIM-homeobox, *ceh-37* Otx and *nhr-67* Tailless/TLX genes are expressed transiently in the postmitotic AWA, AWB and ASE amphid chemosensory neurons, respectively (Sarafi-Reinach et al., 2001; Lanjuin et al., 2003; Sarin et al., 2009). Similar to *mls-2* in the AWC neurons, transient expression of these TFs triggers expression of the *odr-7* nuclear hormone receptor, *lim-4* LIM-homeobox and the *che-1* zinc-finger genes in the AWA, AWB and ASE neurons, respectively, which maintain expression via autoregulation and act as terminal selector TFs for these neuron types (Sengupta et al., 1994; Sagasti et al., 1999; Sarafi-Reinach et al., 2001; Lanjuin et al., 2003; Sarin et al., 2009). Thus, similar regulatory principles employing members of different TF families appear to specify the identities of functionally related amphid sensory neuron types.

In this context, it is interesting to address how *ceh-36* expression is initiated in the AWB neurons of *lim-4* mutants to promote the AWC fate. Our experiments indicate that MLS-2 is dispensable for the initiation of *ceh-36* expression, suggesting that alternative mechanisms must operate in the AWB neurons in the absence of LIM-4 function. A hint as to a possible mechanism comes from the observation that loss-of-function mutations in the transiently expressed *ceh-37* Otx gene, which regulates *lim-4* expression in the AWB neurons, does not lead to the adoption of an AWC identity by this neuron type (Lanjuin et al., 2003), and our finding that *ceh-36* contains an OTX homeodomain binding site in its cis-regulatory sequences. We suggest that transient CEH-37 expression is able to activate *ceh-36* expression in the AWB neurons via direct interaction with the OTX homeodomain binding motif in the *ceh-36* cis-regulatory sequences, but that this activation is normally prevented by LIM-4-mediated repression of *ceh-36* in wild-type animals (Fig. 7). Thus, in both the AWC neurons in wild-type animals and in AWB neurons in *lim-4* mutants, transient expression of a TF triggers expression of *ceh-36*, further highlighting the similarities in the regulatory principles of cell specification used by these neuron types.

Temporal regulation of TF expression also plays an important role in postmitotic neuron subtype specification in organisms other than *C. elegans*. In *Drosophila*, transient stochastic expression of the bHLH-PAS-domain TF Spineless in subsets of photoreceptor cells is necessary and sufficient to specify a photoreceptor subtype, and to promote expression of subtype-specific genes such as rhodopsins (Wernet et al., 2006). In spinal cord motoneurons, the expression of specific homeodomain proteins must be temporally regulated in order to specify columnar and functional identities (Tanabe et al., 1998; William et al., 2003; Thaler et al., 2004). The mechanisms by which temporal regulation of TF expression is effected in postmitotic neuron types are not fully understood but are likely to involve both extrinsic and lineage-derived intrinsic cues (e.g. Baumeister et al., 1996; Sockanathan et al., 2003; Bertrand and Hobert, 2009). These cues define the cellular context, which includes the complement of regulators such as other TFs, as well as additional post-transcriptional and post-translational modifiers that are expressed in a given neuron type and which modulate the expression and specific functions of the TF in that cell type. Thus, the cellular context regulates the transient expression of *mls-2* in the AWC neurons and its maintained expression in the ASK sensory and AIM interneurons (Jiang et al., 2008), which in turn define the functions of these neuron types. Identification of the mechanisms that govern regulatory TF expression and functions in specific cell types is important in furthering our understanding of how the distinctive characteristics of differentiated cells are defined, and allows us to more precisely regulate the generation of desired cell types from undifferentiated cells, such as stem cells, in vitro.

Acknowledgements
We are grateful to Mark Alkema, Oliver Hobert, Jun Liu, Shohei Mitani, Shai Shaham and Marian Walhout for reagents, the *Caenorhabditis* Genetics Center for strains, John Murray for lineage analyses, and Shai Shaham and Chris Li for comments on the manuscript. This work was supported by the NIH (R01 GM56223 to P.S.). Deposited in PMC for release after 12 months.

Competing interests statement
The authors declare no competing financial interests.

Supplementary material
Supplementary material for this article is available at
http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.044719/-/DC1

References
Alenina, N., Bashammakh, S. and Bader, M. (2006). Specification and differentiation of serotonergic neurons. *Stem Cell Rev.* **2**, 5-10.
Alon, U. (2007). Network motifs: theory and experimental approaches. *Nat. Rev. Genet.* **8**, 450-461.
Altun-Gultekin, Z., Andachi, Y., Tsalik, E. L., Pilgrim, D., Kohara, Y. and Hobert, O. (2001). A regulatory cascade of three homeobox genes, *ceh-10, ttx-3* and *ceh-23* controls cell fate specification of a defined interneuron class in *C. elegans*. *Development* **128**, 1951-1969.
Amendt, B. A., Sutherland, L. B. and Russo, A. F. (1999). Transcriptional antagonism between *Hmx1* and *Nkx2.5* for a shared DNA-binding site. *J. Biol. Chem.* **274**, 11635-11642.
Bao, Z., Murray, J. I., Boyle, T., Ooi, S. L., Sandel, M. J. and Waterston, R. H. (2006). Automated cell lineage tracing in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* **103**, 2707-2712.

Bargmann, C. I. (2006). Chemosensation in *C. elegans*. In *Wormbook*, pp. 1-29. http://www.wormbook.org.
Bargmann, C. I., Hartwieg, E. and Horvitz, H. R. (1993). Odorant-selective genes and neurons mediate olfaction in *C. elegans*. *Cell* **74**, 515-527.
Baumeister, R., Liu, Y. and Ruvkun, G. (1996). Lineage-specific regulators couple cell lineage asymmetry to the transcription of the *Caenorhabditis* elegans POU gene *unc-86* during neurogenesis. *Genes Dev.* **10**, 1395-1410.
Bertrand, V. and Hobert, O. (2009). Linking asymmetric cell division to the terminal differentiation program of postmitotic neurons in *C. elegans*. *Dev. Cell* **16**, 563-575.
Bruhn, S. L. and Cepko, C. L. (1996). Development of the pattern of photoreceptors in the chick retina. *J. Neurosci.* **16**, 1430-1439.
Chang, S., Johnston, R. J., Jr and Hobert, O. (2003). A transcriptional regulatory cascade that controls left/right asymmetry in chemosensory neurons of *C. elegans*. *Genes Dev.* **17**, 2123-2137.
Chuang, C. F. and Bargmann, C. I. (2005). A Toll-interleukin 1 repeat protein at the synapse specifies asymmetric odorant receptor expression via ASK1 MAPKKK signaling. *Genes Dev.* **19**, 270-281.
Chuang, C. F., Vanhoven, M. K., Fetter, R. D., Verselis, V. K. and Bargmann, C. I. (2007). An innexin-dependent cell network establishes left-right neuronal asymmetry in *C. elegans*. *Cell* **129**, 787-799.
Colavita, A., Krishna, S., Zheng, H., Padgett, R. W. and Culotti, J. G. (1998). Pioneer axon guidance by UNC-129, a *C. elegans* TGF-beta. *Science* **281**, 706-709.
Collet, J., Spike, C. A., Lundquist, E. A., Shaw, J. E. and Herman, R. K. (1998). Analysis of *osm-6*, a gene that affects sensory cilium structure and sensory neuron function in *Caenorhabditis elegans*. *Genetics* **148**, 187-200.
Colosimo, M. E., Brown, A., Mukhopadhyay, S., Gabel, C., Lanjuin, A. E., Samuel, A. D. and Sengupta, P. (2004). Identification of thermosensory and olfactory neuron-specific genes via expression profiling of single neuron types. *Curr. Biol.* **14**, 2245-2251.
Dessaud, E., McMahon, A. P. and Briscoe, J. (2008). Pattern formation in the vertebrate neural tube: a sonic hedgehog morphogen-regulated transcriptional network. *Development* **135**, 2489-2503.
Duggan, A., Ma, C. and Chalfie, M. (1998). Regulation of touch receptor differentiation by the *Caenorhabditis elegans mec-3* and *unc-86* genes. *Development* **125**, 4107-4119.
Esmaeili, B., Ross, J. M., Neades, C., Miller, D. M., 3rd and Ahringer, J. (2002). The *C. elegans* even-skipped homologue, *vab-7*, specifies DB motoneurone identity and axon trajectory. *Development* **129**, 853-862.
Etchberger, J. F., Lorch, A., Sleumer, M. C., Zapf, R., Jones, S. J., Marra, M. A., Holt, R. A., Moerman, D. G. and Hobert, O. (2007). The molecular signature and *cis*-regulatory architecture of a *C. elegans* gustatory neuron. *Genes Dev.* **21**, 1653-1674.
Etchberger, J. F., Flowers, E. B., Poole, R. J., Bashllari, E. and Hobert, O. (2009). *Cis*-regulatory mechanisms of left/right asymmetric neuron-subtype specification in *C. elegans*. *Development* **136**, 147-160.
Finney, M., Ruvkun, G. and Horvitz, H. R. (1988). The *C. elegans* cell lineage and differentiation gene *unc-86* encodes a protein with a homeodomain and extended similarity to transcription factors. *Cell* **55**, 757-769.
Gehring, W. J., Qian, Y. Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Affolter, M., Otting, G. and Wuthrich, K. (1994). Homeodomain-DNA recognition. *Cell* **78**, 211-223.
Hart, A. C., Kass, J., Shapiro, J. E. and Kaplan, J. M. (1999). Distinct signaling pathways mediate touch and osmosensory responses in a polymodal sensory neuron. *J. Neurosci.* **19**, 1952-1958.
Herman, R. K. and Hedgecock, E. M. (1990). Limitation of the size of the vulval primordium of *Caenorhabditis elegans* by *lin-15* expression in surrounding hypodermis. *Nature* **348**, 169-171.
Hobert, O. (2008). Regulatory logic of neuronal diversity: terminal selector genes and selector motifs. *Proc. Natl. Acad. Sci. USA* **105**, 20067-20071.
Jiang, Y., Horner, V. and Liu, J. (2005). The HMX homeodomain protein MLS-2 regulates cleavage orientation, cell proliferation and cell fate specification in the *C. elegans* postembryonic mesoderm. *Development* **132**, 4119-4130.
Jiang, Y., Shi, H., Amin, N. M., Sultan, I. and Liu, J. (2008). Mesodermal expression of the *C. elegans* HMX homolog *mls-2* requires the PBC homolog CEH-20. *Mech. Dev.* **125**, 451-461.
Johnston, R. J., Jr and Hobert, O. (2003). A microRNA controlling left/right neuronal asymmetry in *Caenorhabditis elegans*. *Nature* **18**, 845-849.
Kim, K. and Li, C. (2004). Expression and regulation of an FMRFamide-related neuropeptide gene family in *Caenorhabditis elegans*. *J. Comp. Neurol.* **475**, 540-550.
Kim, K., Sato, K., Shibuya, M., Zeiger, D. M., Butcher, R. A., Ragains, J. R., Clardy, J., Touhara, K. and Sengupta, P. (2009). Two chemoreceptors mediate developmental effects of dauer pheromone in *C. elegans*. *Science* **13**, 994-998.
Koga, M. and Ohshima, Y. (2004). The *C. elegans ceh-36* gene encodes a putative homeodomain transcription factor involved in chemosensory functions of ASE and AWC neurons. *J. Mol. Biol.* **336**, 579-587.
Lanjuin, A. and Sengupta, P. (2004). Specification of chemosensory neuron subtype identities in *Caenorhabditis elegans*. *Curr. Opin. Neurobiol.* **14**, 22-30.
Lanjuin, A., VanHoven, M. K., Bargmann, C. I., Thompson, J. K. and Sengupta, P. (2003). *Otx/otd* homeobox genes specify distinct sensory neuron identities in *C. elegans*. *Dev. Cell* **5**, 621-633.
Lesch, B. J., Gehrke, A. R., Bulyk, M. L. and Bargmann, C. I. (2009). Transcriptional regulation and stabilization of left-right neuronal identity in *C. elegans*. *Genes Dev.* **23**, 345-358.
L'Etoile, N. D. and Bargmann, C. I. (2000). Olfaction and odor discrimination are mediated by the *C. elegans* guanylyl cyclase ODR-1. *Neuron* **25**, 575-586.
Malnic, B., Hirono, J., Sato, T. and Buck, L. B. (1999). Combinatorial receptor codes for odors. *Cell* **96**, 713-723.
Melkman, T. and Sengupta, P. (2004). The worm's sense of smell: development of functional diversity in the chemosensory system of *C. elegans*. *Dev. Biol.* **265**, 302-319.
Miller, D. M., 3rd and Niemeyer, C. J. (1995). Expression of the *unc-4* homeoprotein in *Caenorhabditis elegans* motor neurons specifies presynaptic input. *Development* **121**, 2877-2886.
Muhr, J., Andersson, E., Persson, M., Jessell, T. M. and Ericson, J. (2001). Groucho-mediated transcriptional repression establishes progenitor cell pattern and neuronal fate in the ventral neural tube. *Cell* **104**, 861-873.
Nonet, M. L., Staunton, J. E., Kilgard, M. P., Fergestad, T., Hartwieg, E., Horvitz, H. R., Jorgensen, E. M. and Meyer, B. J. (1997). *Caenorhabditis elegans rab-3* mutant synapses exhibit impaired function and are partially depleted of vesicles. *J. Neurosci.* **17**, 8061-8073.
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.
Ortiz, C. O., Etchberger, J. F., Posy, S. L., Frokjaer-Jensen, C., Lockery, S., Honig, B. and Hobert, O. (2006). Searching for neuronal left/right asymmetry: genomewide analysis of nematode receptor-type guanylyl cyclases. *Genetics* **173**, 131-149.
Ortiz, C. O., Faumont, S., Takayama, J., Ahmed, H. K., Goldsmith, A. D., Pocock, R., McCormick, K. E., Kunimoto, H., Iino, Y., Lockery, S. et al. (2009). Lateralized gustatory behavior of *C. elegans* is controlled by specific receptor-type guanylyl cyclases. *Curr. Biol.* **19**, 996-1004.
Perkins, L. A., Hedgecock, E. M., Thomson, J. N. and Culotti, J. G. (1986). Mutant sensory cilia in the nematode *Caenorhabditis elegans*. *Dev. Biol.* **117**, 456-487.
Pichaud, F., Briscoe, A. and Desplan, C. (1999). Evolution of color vision. *Curr. Opin. Neurobiol.* **9**, 622-627.
Pierce-Shimomura, J. T., Faumont, S., Gaston, M. R., Pearson, B. J. and Lockery, S. R. (2001). The homeobox gene *lim-6* is required for distinct chemosensory representations in *C. elegans*. *Nature* **410**, 694-698.
Roztocil, T., Matter-Sadzinski, L., Alliod, C., Ballivet, M. and Matter, J. M. (1997). NeuroM, a neural helix-loop-helix transcription factor, defines a new transition stage in neurogenesis. *Development* **124**, 3263-3272.
Sagasti, A., Hobert, O., Troemel, E. R., Ruvkun, G. and Bargmann, C. I. (1999). Alternative olfactory neuron fates are specified by the LIM homeobox gene *lim-4*. *Genes Dev.* **13**, 1794-1806.
Sagasti, A., Hisamoto, N., Hyodo, J., Tanaka-Hino, M., Matsumoto, K. and Bargmann, C. I. (2001). The CaMKII UNC-43 activates the MAPKKK NSY-1 to execute a lateral signaling decision required for asymmetric olfactory neuron fates. *Cell* **105**, 221-232.
Sarafi-Reinach, T. R., Melkman, T., Hobert, O. and Sengupta, P. (2001). The *lin-11* LIM homeobox gene specifies olfactory and chemosensory neuron fates in *C. elegans*. *Development* **128**, 3269-3281.
Sarin, S., Antonio, C., Tursun, B. and Hobert, O. (2009). The *C. elegans* Tailless/TLX transcription factor *nhr-67* controls neuronal identity and left/right asymmetric fate diversification. *Development* **136**, 2933-2944.
Satterlee, J. S., Sasakura, H., Kuhara, A., Berkeley, M., Mori, I. and Sengupta, P. (2001). Specification of thermosensory neuron fate in *C. elegans* requires *ttx-1*, a homolog of *otd/Otx*. *Neuron* **31**, 943-956.
Sengupta, P., Colbert, H. A. and Bargmann, C. I. (1994). The *C. elegans* gene *odr-7* encodes an olfactory-specific member of the nuclear receptor superfamily. *Cell* **79**, 971-980.
Sockanathan, S., Perlmann, T. and Jessell, T. M. (2003). Retinoid receptor signaling in postmitotic motor neurons regulates rostrocaudal positional identity and axonal projection pattern. *Neuron* **40**, 97-111.
Sulston, J. E., Schierenberg, E., White, J. G. and Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64-119.
Tanabe, Y., William, C. and Jessell, T. M. (1998). Specification of motor neuron identity by the MNR2 homeodomain protein. *Cell* **95**, 67-80.
Tanaka-Hino, M., Sagasti, A., Hisamoto, N., Kawasaki, M., Nakano, S., Ninomiya-Tsuji, J., Bargmann, C. I. and Matsumoto, K. (2002). SEK-1 MAPKK mediates Ca^{2+} signaling to determine neuronal asymmetric development in *Caenorhabditis elegans*. *EMBO Rep.* **3**, 56-62.
Thaler, J., Harrison, K., Sharma, K., Lettieri, K., Kehrl, J. and Pfaff, S. L. (1999). Active suppression of interneuron programs within developing motor neurons revealed by analysis of homeodomain factor HB9. *Neuron* **23**, 675-687.

Thaler, J. P., Koo, S. J., Kania, A., Lettieri, K., Andrews, S., Cox, C., Jessell, T. M. and Pfaff, S. L. (2004). A postmitotic role for Isl-class LIM homeodomain proteins in the assignment of visceral spinal motor neuron identity. *Neuron* **41**, 337-350.

Treisman, J., Gonczy, P., Vashishtha, M., Harris, E. and Desplan, C. (1989). A single amino acid can determine the DNA binding specificity of homeodomain proteins. *Cell* **59**, 553-562.

Troemel, E. R., Chou, J. H., Dwyer, N. D., Colbert, H. A. and Bargmann, C. I. (1995). Divergent seven transmembrane receptors are candidate chemosensory receptors in C. elegans. *Cell* **83**, 207-218.

Troemel, E. R., Sagasti, A. and Bargmann, C. I. (1999). Lateral signaling mediated by axon contact and calcium entry regulates asymmetric odorant receptor expression in *C. elegans*. *Cell* **99**, 387-398.

Uchida, O., Nakano, H., Koga, M. and Ohshima, Y. (2003). The *C. elegans che-1* gene encodes a zinc finger transcription factor required for specification of the ASE chemosensory neurons. *Development* **130**, 1215-1224.

Vanhoven, M. K., Bauer Huang, S. L., Albin, S. D. and Bargmann, C. I. (2006). The claudin superfamily protein *nsy-4* biases lateral signaling to generate left-right asymmetry in *C. elegans* olfactory neurons. *Neuron* **51**, 291-302.

Wang, L. H., Chmelik, R. and Nirenberg, M. (2002). Sequence-specific DNA binding by the vnd/NK-2 homeodomain of *Drosophila*. *Proc. Natl. Acad. Sci. USA* **99**, 12721-12726.

Wang, L. H., Chmelik, R., Tang, D. and Nirenberg, M. (2005). Identification and analysis of vnd/NK-2 homeodomain binding sites in genomic DNA. *Proc. Natl. Acad. Sci. USA* **102**, 7097-7102.

Wang, W. and Lufkin, T. (2005). *Hmx* homeobox gene function in inner ear and nervous system cell-type specification and development. *Exp. Cell Res.* **306**, 373-379.

Way, J. C. and Chalfie, M. (1988). *mec-3*, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in *C. elegans*. *Cell* **54**, 5-16.

Way, J. C. and Chalfie, M. (1989). The *mec-3* gene of *Caenorhabditis elegans* requires its own product for maintained expression and is expressed in three neuronal cell types. *Genes Dev.* **3**, 1823-1833.

Wernet, M. F., Mazzoni, E. O., Celik, A., Duncan, D. M., Duncan, I. and Desplan, C. (2006). Stochastic spineless expression creates the retinal mosaic for colour vision. *Nature* **440**, 174-180.

Wes, P. D. and Bargmann, C. I. (2001). *C. elegans* odour discrimination requires asymmetric diversity in olfactory neurons. *Nature* **410**, 698-701.

William, C. M., Tanabe, Y. and Jessell, T. M. (2003). Regulation of motor neuron subtype identity by repressor activity of *Mnx* class homeodomain proteins. *Development* **130**, 1523-1536.

Xue, D., Finney, M., Ruvkun, G. and Chalfie, M. (1992). Regulation of the *mec-3* gene by the *C. elegans* homeoproteins UNC-86 and MEC-3. *EMBO J.* **11**, 4969-4979.

Yoshimura, S., Murray, J. I., Lu, Y., Waterston, R. H. and Shaham, S. (2008). *mls-2* and *vab-3* control glia development, *hlh-17*/*Olig* expression and glia-dependent neurite extension in *C. elegans*. *Development* **135**, 2263-2275.

Yu, S., Avery, L., Baude, E. and Garbers, D. A. (1997). Guanylyl cyclase expression in specific sensory neurons: A new family of chemosensory receptors. *Proc. Natl. Acad. Sci. USA* **94**, 3384-3387.