

Hox proteins interact to pattern neuronal subtypes in *Caenorhabditis elegans* males

Andrea K. Kalis ¹, Maria C. Sterrett ², Cecily Armstrong,¹ Amarantha Ballmer,² Kylie Burkstrand,¹ Elizabeth Chilson,¹ Estee Emlen,² Emma Ferrer,² Seanna Loeb,¹ Taylor Olin,¹ Kevin Tran,² Andrew Wheeler,² Jennifer Ross Wolff ^{2,*}

¹Department of Biology, St. Catherine University, St. Paul, MN 55105, USA,

²Department of Biology, Carleton College, Northfield, MN 55057, USA

*Corresponding author: Biology Department, Carleton College, One North College Street, Northfield, MN 55057, USA. Email: jwolff@carleton.edu

Abstract

Hox transcription factors are conserved regulators of neuronal subtype specification on the anteroposterior axis in animals, with disruption of Hox gene expression leading to homeotic transformations of neuronal identities. We have taken advantage of an unusual mutation in the *Caenorhabditis elegans* Hox gene *lin-39*, *lin-39(ccc16)*, which transforms neuronal fates in the *C. elegans* male ventral nerve cord in a manner that depends on a second Hox gene, *mab-5*. We have performed a genetic analysis centered around this homeotic allele of *lin-39* in conjunction with reporters for neuronal target genes and protein interaction assays to explore how LIN-39 and MAB-5 exert both flexibility and specificity in target regulation. We identify *cis*-regulatory modules in neuronal reporters that are both region-specific and Hox-responsive. Using these reporters of neuronal subtype, we also find that the *lin-39(ccc16)* mutation disrupts neuronal fates specifically in the region where *lin-39* and *mab-5* are coexpressed, and that the protein encoded by *lin-39(ccc16)* is active only in the absence of *mab-5*. Moreover, the fates of neurons typical to the region of *lin-39-mab-5* coexpression depend on both Hox genes. Our genetic analysis, along with evidence from Bimolecular Fluorescence Complementation protein interaction assays, supports a model in which LIN-39 and MAB-5 act at an array of *cis*-regulatory modules to cooperatively activate and to individually activate or repress neuronal gene expression, resulting in regionally specific neuronal fates.

Keywords: *Caenorhabditis elegans*; Hox genes; neurogenesis; TALE homeodomain proteins; male-specific neurons; ventral cord neurons; BiFC

Introduction

The specification of neuronal subtypes during development is essential for the organization of animal nervous systems, providing a foundation for integration into a coherent circuitry that mediates behaviors essential to survival and reproduction. In bilaterians, Hox transcription factors are highly conserved effectors of neuronal specification, assigning positional identities on the anteroposterior (AP) axis as well as influencing cell survival and proliferation (reviewed in [Dasen and Jessell 2009](#)). In *Drosophila* and mice, Hox proteins are essential for specification of axial motor neurons that innervate the limbs, with regional Hox gene expression leading to AP specificity of neuron identity and targeting ([Baek et al. 2013](#); [Philippidou and Dasen 2013](#); [Estacio-Gómez and Díaz-Benjumea 2014](#); [Shin et al. 2020](#)). While *Caenorhabditis elegans* axial patterning is less obvious than that evinced in limbed animals, the nervous system is nonetheless patterned on the AP axis, with Hox proteins playing a central role. For example, the posterior Hox protein EGL-5 (AbdB/Hox9-13) specifies posterior subtype identity in touch receptive neurons, the Hox proteins LIN-39 (Scr/Df/Hox4-5) and MAB-5 (Antp/Hox6-8) mediate the direction of migration and final position of neuroblasts on the AP axis, and MAB-5 and EGL-5 coordinate to pattern the finger-like

sensory rays of the male tail ([Salser and Kenyon 1992](#); [Clark et al. 1993](#); [Salser et al. 1993](#); [Chow and Emmons 1994](#); [Zheng et al. 2015](#)).

As in the mouse spinal cord and *Drosophila* ventral nerve cord, neuronal subclass specification in the *C. elegans* ventral nerve cord also depends on the regional activity of Hox proteins. In hermaphrodites, the Hox genes *lin-39*, *mab-5*, and *egl-5* are expressed subclass-specifically in their respective domains and in coordination with the transcription factor UNC-3 help define motor neuron fates by controlling the expression of subclass-specific target genes ([Kratsios et al. 2017](#); [Feng et al. 2020](#)). Whereas many of the neuronal subclasses of the ventral cord are present in both sexes to control locomotion, the ventral cord also includes sex-specific neurons that mediate reproductive behaviors. VC neurons are unique to hermaphrodites, and CA and CP neurons to males, and each class has distinct targets, morphology, and neurotransmitter expression ([Sulston and Horvitz 1977](#); [White et al. 1986](#); [Schinkmann and Li 1992](#); [Duerr et al. 2008](#); [Jarrell et al. 2012](#)). In males, the CPs and their sister cells, the CAs, synapse with a variety of postsynaptic partners including male-specific tail neurons, body wall muscles, and the gonad ([Jarrell et al. 2012](#)). The CA neurons regulate sperm transfer ([Schindelman et al. 2006](#)), whereas

the CP motor neurons are thought to contribute to ventral flexure of the body wall as the male circles the hermaphrodite in search of the vulva (Loer and Kenyon 1993; Liu and Sternberg 1995; Carnell et al. 2005; Whittaker and Sternberg 2009; Serrano-Saiz et al. 2017).

Hox-mediated patterning of the male-specific CPs is particularly striking, with neuronal subclass identities defined by overlapping spatial domains of the Hox proteins LIN-39 and MAB-5 (Fig. 1, a and b). LIN-39 is expressed in an anterior domain that encompasses CPs 1–6 and their precursor cells (P3.aap–P8.aap), and is required for survival of these cells, as well as the division that generates CPs 1–6 (Costa et al. 1988; Clark et al. 1993; Wang et al. 1993; Kenyon et al. 1997; Kalis et al. 2014). MAB-5 is expressed in a more posterior domain that encompasses CPs 5–9 and their precursors (P7.aap–P11.aap), and is required for survival of these cells, as well as the division that generates CPs 7–9 (Costa et al. 1988; Salser et al. 1993; Wang et al. 1993; Kenyon et al. 1997). The fates of the CPs on the AP axis can be divided into 3 zones that correspond to these distinct Hox expression profiles. In anterior CPs 1–4 (zone 1), LIN-39 directs expression of the serotonergic gene *tph-1* and the neuropeptidergic gene *flp-22* (Loer and Kenyon 1993; Sze et al. 2000; Clark and Chiu 2003; Kalis et al. 2014). In posterior CPs 7–9 (zone 3), MAB-5 directs expression of the neuropeptidergic gene *flp-21* (Kalis et al. 2014). In midbody CPs 5–6 (zone 2), LIN-39 and MAB-5 together define CP fates: CPs 5–6 are

characterized by intense expression of *tph-1* and absence of *flp-21* and *flp-22* (Kalis et al. 2014). Here, LIN-39 and MAB-5 appear to act in concert to regulate *tph-1* expression: *lin-39* is required to activate *tph-1::mCherry*, while *mab-5* is required for high levels of expression. In contrast, MAB-5 and LIN-39 act reciprocally to regulate expression of *flp-22* and *flp-21*, which are each repressed in a manner that depends on *mab-5* and *lin-39*, respectively. The interactions of *lin-39* and *mab-5* in zone 2 provide a model for understanding how neuronal fates are defined in situations where 2 Hox genes are coexpressed.

A major outstanding question in our understanding of Hox-mediated specification is how Hox proteins exert both flexibility and specificity in their target regulation. In vitro target selection has demonstrated that Hox proteins can recognize a relatively nonspecific AT-rich binding site that is found at high frequency throughout genomes. Despite this, decades of observations of Hox mutant phenotypes and DNA-binding assays have demonstrated that Hox function in vivo can be quite specific (reviewed in Mann et al. 2009; Merabet and Hudry 2013; Merabet and Mann 2016; De Kumar and Darland 2021). This raises the question of how each Hox protein recognizes targets in its domain of expression and how coexpressed Hox proteins regulate both common and distinct targets in the same cell or tissue. Specificity of Hox proteins for their targets could be influenced by many variables. Some of these include: (1) structural variations in the Hox

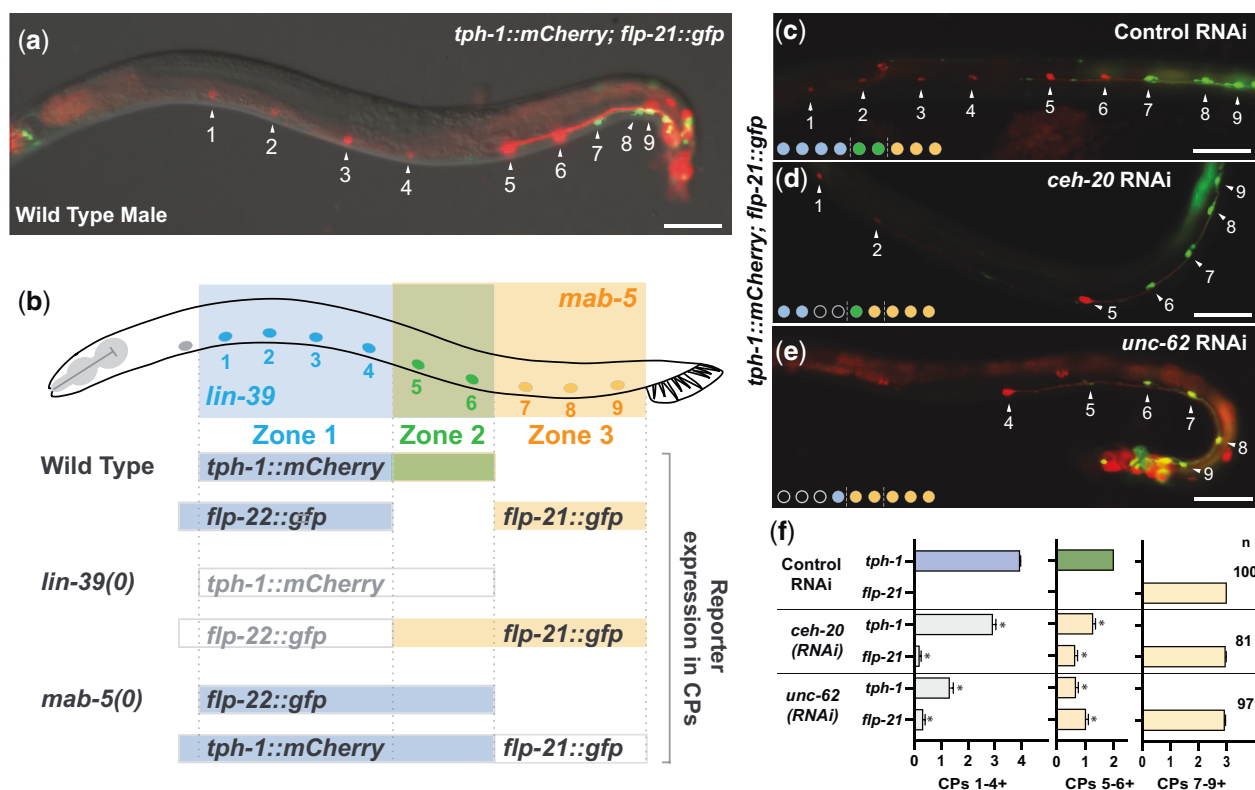


Fig. 1. Hox and TALE transcription factors define 3 anteroposterior zones of CP neuron fate in *C. elegans* males. a) The serotonergic reporter *tph-1::mCherry* is expressed in CPs 1–6 in a wild-type male; *flp-21::gfp* is expressed in CPs 7–9. Expression of *tph-1::mCherry* is relatively less intense in zone 1 CPs 1–4 than in zone 2 CPs 5–6 (see Fig. 3a). b) Schematic showing zone-specific and hox-dependent expression of CP neuron reporters in wild-type, *lin-39* null mutant, and *mab-5* null mutant males. Bar colors indicate zone-typical reporter expression, such that zone 1 = blue, zone 2 = green, zone 3 = yellow. c–f) TALE factors *ceh-20* and *unc-62* are required for normal *tph-1::mCherry* expression in Zones 1 and 2, but not for *flp-21::gfp* expression in zone 3. c) Expression of *tph-1::mCherry* (red) in CPs 1–6 and *flp-21::gfp* (green) in CPs 7–9 in a male treated with control (empty vector) RNAi. d and e) Reduced expression of *tph-1::mCherry* and expanded expression of *flp-21::gfp* in males treated with *ceh-20* RNAi (d) or *unc-62* RNAi (e). c–e) Filled circle colors summarize zone identities of CP neurons. Scale bar = 50 μ m. f) Summary of *tph-1::mCherry* and *flp-21::gfp* expression in CPs in males treated with control, *ceh-20*, and *unc-62* RNAi; CPs+ denotes mean number of fluor-positive CPs/worm in each zone. Bar colors indicate zone identities of CP neurons; error bars indicate standard error of the mean (SEM); number of ventral cords scored is indicated at right; * $P < 0.05$, unpaired t-test, compared to control RNAi.

proteins themselves, in or outside the homeodomain, that influence DNA or cofactor binding; (2) variation in binding site sequence and clustering; and (3) protein–protein interactions among Hox proteins or with cofactors such as the TALE class of homeodomain proteins. For example, in *Drosophila*, interactions between Hox and TALE cofactors of the Pbx and Meis families are known to impart higher specificity to DNA binding both in vitro and in vivo (Slattery et al. 2011a, 2011b). Additionally, amino acid motifs outside the homeodomain have been shown to mediate distinct interactions between Hox proteins and TALE cofactors depending on DNA-binding site context (Dard et al. 2018, 2019; Bridoux et al. 2020). Consequently, any of these variables have the potential to influence one another to form a complex web affecting Hox–target interactions. Thus, Hox proteins, cofactors, and DNA have the potential for a high degree of flexibility in their interactions, which, in turn, could impart a range of target specificity to the repertoire of each Hox protein.

The *C. elegans* CPs constitute a genetically tractable system in which to explore one such web of Hox/cofactor/target interactions in neuronal specification. In the CPs, LIN-39 and MAB-5 exhibit both specificity for their individual targets and flexibility in the sign of their transcriptional regulation, providing an entry point for addressing how the functions of LIN-39 and MAB-5 converge in their region of coexpression. Here, we have used a genetic approach to investigate the Hox system of regulation of subtype specification in CP neurons. We find that coexpression of LIN-39 and MAB-5 is sufficient to confer zone 2 fates in zone 1 CP 1–4 neurons, and have identified cis-regulatory modules (CRMs) that drive zone-specific expression of Hox target genes. We have isolated a homeotic allele of *lin-39* that specifically transforms the fates of zone 2 CPs to fates typical of their more posterior zone 3 counterparts due to failure of LIN-39 and MAB-5 to negotiate interactions that define regionally specific neuronal fates. This mutation disrupts sequences C-terminal to the LIN-39 homeodomain and weakens interactions among LIN-39, MAB-5, and the TALE factor CEH-20. Our work supports a model in which LIN-39 and MAB-5 act both independently and together to regulate target genes that define neuronal subclass identity, in conjunction with TALE cofactors.

Materials and methods

Caenorhabditis elegans strains

Nematodes were cultured as described in Stiernagle (2006), and grown at 15°C, 20°C, or 22°C (for RNAi) on nematode growth media seeded with bacteria (*Escherichia coli* OP51). All strains in which males were scored contained a high incidence of males, *him-8*(e1489), mutation. All mutant alleles and integrated transgenes can be found in the Reagent Table (Table A1).

Identification of *lin-39*(ccc16)

lin-39(ccc16) was isolated in a loss-of-function EMS mutagenesis screen (Kalis et al. 2014), and subsequently mapped to ~1 cM on chromosome III using single nucleotide polymorphisms (Wicks et al. 2001; Davis et al. 2005) and deficiencies, followed by whole-genome sequencing. Whole-genome sequencing of the *lin-39*(ccc16) line also revealed a mutation in the protein-coding region of the closely linked *lin-13* locus, resulting in the amino acid substitution C145Y. Using a line also carrying an *unc-32*(e189) mutation, recombinants between *lin-39*(ccc16) and *unc-32*(e189) were isolated and screened by PCR for the absence of the *lin-13* mutation. A recombinant line with *lin-39*(ccc16) but lacking the *lin-13* and *unc-32* mutations was used for all further analysis,

with the exception of strains double mutant for *lin-39*(ccc16) and *mab-5*(e1239), which retain the *unc-32*(e189) mutation.

Antiserotonin staining

Serotonin antibody staining was performed as in Loer and Rivard (2007) using antiserotonin primary antibody (Sigma-Aldrich #S5545) at 1:100 and AlexaFluor 488 Goat antirabbit IgG (Thermo Fisher Scientific #R37116) secondary antibody at 1:500.

3' RACE

RNA was isolated from either *him-8* or *lin-39*(ccc16); *him-8*; *ccc1s1* worms using the Macherey-Nagel NucleoSpin RNA isolation kit. After the addition of RA1 and BME, worm pellets were lysed 2 times for 1 min each using a micropestle and drill. Splice products were isolated using Takara Bio USA, Inc. SMARTer RACE 5'/3' kit. First-strand 3' RACE-ready cDNA was prepared and 3' RACE was performed using the gene-specific primer GATTACGCCAAGCTTccacagatgcaccagagctacagctcc. RACE products were characterized by gel purification and in-fusion cloning into pRACE vector followed by DNA sequencing.

RNA interference

RNAi of *ceh-20* and *unc-62* in *flp-21::gfp* was carried out as in Kalis et al. (2014). For RNAi of *ceh-20* and *ceh-60* (control) in Bimolecular Fluorescence Complementation (BiFC) extrachromosomal lines, adult transgenic animals were allowed to lay eggs on RNAi plates overnight, grown to young adulthood, heat shocked for 2 h at 33°C, and recovered at 20°C for 5–7 h before imaging. RNAi was performed in triplicate, with 20 worms scored per experiment. *ceh-60* RNAi was used as a negative control.

Cloning of cis-regulatory regions

Segments of regulatory regions in the *flp-22::gfp* reporter plasmid (Kim and Li 2004) were deleted by restriction digest and religation. For all other reporter constructs, PCR products were amplified from genomic DNA and cloned into the *Sal* I and *Sph* I restriction sites in the *pes-10* minimal promoter::gfp plasmid L3135 (Fire Vector Kit, Addgene #1531). All clones were confirmed by sequencing of the inserted or deleted region; 50 ng/μl of each plasmid was injected with 50–100 ng/μl of *str-1::gfp* as a coinjection marker into *him-8* worms. For each injected clone, 2–5 independent lines were examined.

Bimolecular Fluorescence Complementation

The Venus-based BiFC was carried out as designed in Shyu et al. (2008). *lin-39*, *lin-39*(ccc16ex), *lin-39*(ccc16in), *mab-5*, or *ceh-20* were amplified by PCR from cDNA isolated for 3' RACE (above) and cloned 3' to the heat-shock promoter of EcoRI- and XhoI-digested pPV3 (*unc-62* in pCE-BiFC-VN173) or pPV4 (*ceh-20* in pCE-BiFC-VC155) (Van de Walle et al. 2019) using NEBuilder HiFi DNA Assembly Master Mix and transformed into DH5- α competent cells. Homeodomain mutations are a substitution of aa54 of the homeodomain from Asn to Ala (aat → agc) and were induced using NEB Q5 Site-directed Mutagenesis. BiFC constructs were verified by sequencing; primers used for cloning and mutagenesis are listed in the Reagent Table (Table A1). To make transgenic lines, 100 ng/μl of pRF4 (dominant Roller injection marker) and 15 ng/μl of each BiFC plasmid were microinjected into adult hermaphrodite gonads and at least 3 independent lines were established. To induce protein expression, 1–2-day-old adult transgenic worms were heat shocked for 2 h at 33°C and recovered at 20°C for 5–7 h before imaging. For imaging, 30 worms for each extrachromosomal line were imaged on a Leica DM6

fluorescence microscope at 50 ms exposure. The average fluorescence per worm was determined using LasX imaging software, with fluorescence intensity quantified by drawing an ROI around the first 6 in-focus anterior intestinal nuclei of each worm.

CRISPR

For *lin-39* CRISPR, guide sequences were identified using CRISPOR (Concordet and Haeussler 2018). Guide sequences were introduced by NEB Q5 mutagenesis into the pDD162 (Pef3::Cas9 + Empty sgRNA, Addgene 47549), to generate *lin-39*-guide-Cas-9 plasmids, and confirmed by sequencing. *lin-39* repair templates were generated as single-stranded oligonucleotides (Elim Biopharm), with HPLC/PAGE 2-step purification. *lin-39*-guide-Cas-9 plasmids were injected into N2 hermaphrodites at 50 ng/μl, with 50 ng/μl *lin-39* repair template oligo, 50 ng/μl *dpy-10* co-CRISPR plasmid AP568 (Addgene 70047), 20 ng/μl *dpy-10(cn64)* repair template oligo as described in Arribere et al. (2014). *Dpy-10(cn64)/+* Roller hermaphrodites were selected from the F1 progeny of injected worms and allowed to produce progeny at 20°C for 2 days, then screened by PCR followed by RFLP analysis or sequencing to identify individuals heterozygous for the desired *lin-39* mutation. Mutations were subsequently homozygosed, and then crossed into strains containing the *him-8* mutation and CP reporter transgenes. Presence of *lin-39* mutations in CRISPR; *him-8*; reporter strains was confirmed by RFLP analysis and/or sequencing.

Identification of Hox-binding sites using cisBP

Potential LIN-39-, MAB-5-, and CEH-20-binding sites were identified using CisBP's "Scan a DNA region for potential binding sites" tool (Weirauch et al. 2014) with settings as follows: Species: *C. elegans*, Motif Model: PWMs LogOdds, and Threshold 8. Binding sites were labeled as "Hox" if they were predicted to bind LIN-39, MAB-5, or both, and were labeled as "TALE" if they were predicted to bind CEH-20.

Microscopy and photography

Worms were anesthetized in a solution of 0.1% tricaine and 0.01% tetramisole and mounted on 2% agarose pads on glass slides. Images were collected at 20× on a Zeiss Axioimager A1 microscope with epifluorescence using ProgRes Capture software or a Leica Automated DM6 Microscope with epifluorescence using LasX acquisition software.

Fluorescence intensity comparison

For Fig. 3a, using ImageJ, CP neurons were traced and corrected total cell fluorescence (CTCF) calculated for each CP using the formula: $CTCF = \text{Integrated Density} - (\text{area of selected cell} \times \text{mean fluorescence of background readings})$. Mean CTCF of CPs 1-4 was compared with mean CTCF of CPs 5-6 using an unpaired T test.

Statistical analysis

For data quantification, graphs show the mean number of fluorescence-positive CP neurons per worm $\pm$ the standard error of the mean. Statistical analysis was performed using PRISM 9 Software as follows: for data shown in Fig. 5e, an ordinary one-way ANOVA (multiple comparison) with uncorrected Fisher's LSD; for data shown in Fig. 6i, an unpaired nested t-test (two-tailed); for data shown in Fig. 6j and Supplementary Fig. S3, a nested one-way ANOVA (multiple comparison) with uncorrected Fisher's LSD; and for data shown in Figs. 1f, 2i, 3e, 4g, 5e, and 6m,

an unpaired t-test (2-tailed). Differences with $P < 0.05$ were considered statistically significant.

Results

LIN-39, but not MAB-5, requires TALE cofactor activity to specify CP fate

We are interested in understanding mechanisms by which LIN-39 and MAB-5 act separately in zones 1 and 3, and together in zone 2 to define distinct neuronal fates. The TALE homeobox transcription factors CEH-20 (Pbx) and UNC-62 (MEIS) are candidates to interact with LIN-39 and MAB-5 to influence their function, perhaps contributing to their region specificity. We previously investigated the role of *ceh-20* and *unc-62* in specifying zone 1 and 2 neurons, and found that loss of function of either phenocopies *lin-39(lf)*, resulting in reduction of *tph-1* and *flp-22* expression in these zones (Kalis et al. 2014). In this study, we tested the role of TALE cofactors in MAB-5-dependent neuronal specification by using RNAi (*ceh-20* and *unc-62* null mutations are embryonic lethal) in a neuronally sensitized background (*lin-15b nre-1*) to target *ceh-20* and *unc-62* in males expressing *flp-21::gfp* and *tph-1::mCherry*. As previously observed, *ceh-20(RNAi)* or *unc-62(RNAi)* reduces, but does not eliminate, *tph-1::mCherry* expression in CPs 1-6, consistent with a reduction of function of *ceh-20* and *unc-62* by RNAi. In contrast, *ceh-20(RNAi)* or *unc-62(RNAi)* does not disrupt expression of *flp-21::gfp* in CPs 7-9 (Fig. 1, c-f), suggesting that MAB-5 does not require these cofactors to promote this zone 3-specific trait. Moreover, RNAi targeting *ceh-20* and *unc-62* also results in expansion of *flp-21::gfp* into zone 2 CPs 5-6 (Fig. 1, c-f), suggesting that LIN-39 does require CEH-20 and UNC-62 to prevent *flp-21::gfp* expansion into zone 2. Neither CP5 nor CP6 fate is disproportionately affected by *unc-62* or *ceh-20* RNAi (data not shown), suggesting that the disruption is zone-specific rather than CP5- or CP6-specific. Together, these results confirm our previous finding that LIN-39 requires TALE cofactors to specify CP fates, and suggest that MAB-5 directs zone 3-specific fates independently of these TALE cofactors.

Cis-regulatory modules (CRMs) in *tph-1* and *flp-22* drive zone-specific and Hox-responsive expression in CP neurons

The Zone-specific and Hox-responsive expression pattern of *tph-1* and *flp-22* in CP neurons suggests the presence of CRMs that direct this expression. To identify these CRMs in *tph-1* and *flp-22*, we constructed a series of reporters containing putative regulatory regions from these genes, upstream of GFP (Fig. 2). For *tph-1*, we identified 2 elements that can drive expression in zone 2 CPs 5-6 in conjunction with a minimal *pes-10* promoter (Fig. 2a). One of these elements (-585 to -177) drives relatively weak expression only in some transgenic lines, whereas the other (-290 to -91) drives consistent expression in CPs 5-6 in all lines tested. We continued our analysis with an integrated reporter, *tph-1(CP5-6)::gfp(cccls2)*, which contains both elements (-585 to -91) and which is robustly and specifically expressed in zone 2 CPs 5-6. *tph-1(CP5-6)::gfp* is not expressed in CPs 5-6 in *lin-39(n1760)* or *mab-5(e1239)* strong loss-of-function mutants, indicating that both of these Hox proteins must be present for this CRM to be active (Fig. 2, b-d and i). In contrast, expression of the full-length *tph-1::mCherry* reporter requires *lin-39*, but not *mab-5*, for expression in zone 2 (Fig. 1b, Kalis et al. 2014). This suggests the presence of a *lin-39*-responsive CRM that is both active in zone 2 in *mab-5* mutants and located outside the region (-585 to -91) contained in *tph-1(CP5-6)::gfp*. This analysis did not reveal the presence of a

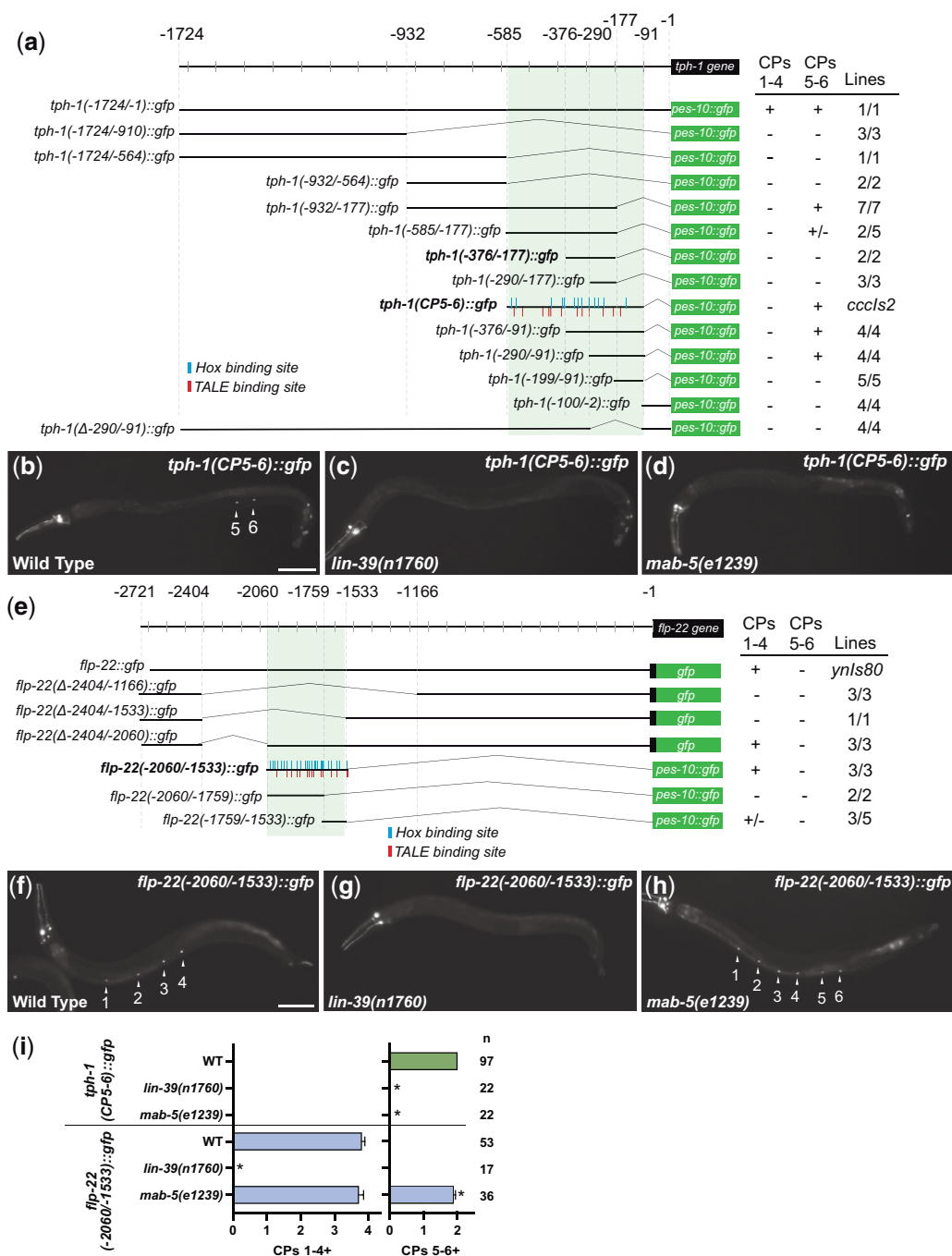


Fig. 2. Cis-regulatory elements in *tph-1* and *flp-22* drive zone-specific and Hox-responsive reporter expression in male CP neurons. a) Identification of *tph-1* CRMs driving CP5-6 (zone 2)-specific expression of GFP in the ventral cord. All lines scored were extrachromosomal arrays, except for the integrated array *cccls2* [*tph-1*(CP5-6)::*gfp*]. Potential Hox- and TALE-binding sites, as identified by cisBP analysis, are indicated for the region encompassed by the *tph-1*(CP5-6)::*gfp* reporter. b-e) The *tph-1*(CP5-6)::*gfp* reporter, which contains upstream sequences from -585 to -91, requires both *lin-39* and *mab-5* for expression in CPs 5-6. b) Expression of *tph-1*(CP5-6)::*gfp* in CPs 5-6 in a wild-type male. c and d) *tph-1*(CP5-6)::*gfp* is not expressed in ventral cord neurons in *lin-39*(e1760) or *mab-5*(e1239) mutants. e) Identification of *flp-22* CRMs driving CP 1-4 (zone 1)-specific expression of GFP in the ventral cord. All lines scored were extrachromosomal arrays, except for the integrated array *ynls80*. Potential Hox- and TALE-binding sites, as identified by cisBP analysis, are indicated for the region encompassed by the *flp-22*(-2,060/-1,533)::*gfp* reporter which contains upstream sequences from -2,060 to -1,533. f-h) The *flp-22*(-2,060/-1,533)::*gfp* reporter requires *lin-39* and *mab-5* for CP1-4-specific expression. (f) Expression of *flp-22*(-2,060/-1,533)::*gfp* in CPs 1-4 in a wild-type male. (g) *flp-22*(-2,060/-1,533)::*gfp* is not expressed in ventral cord neurons in *lin-39*(e1760) mutants. (h) *flp-22*(-2,060/-1,533)::*gfp* is expanded posteriorly, and encompasses CPs 1-6 in *mab-5*(e1239) mutants. i) Summary of *tph-1*(CP5-6)::*gfp* and *flp-22*(-2,060/-1,533)::*gfp* in CPs in wild-type and Hox mutant males. CPs+ denotes mean number of fluor-positive CPs/worm in each zone; error bars indicate standard error of the mean (SEM); number of ventral cords scored is indicated at right; *P < 0.05, unpaired t-test, compared to WT. In a and e) reporter boundaries are indicated in base pairs, relative to the first coding exon of *tph-1* and *flp-22*. + indicates GFP expression in the indicated CP neurons in all transgenic lines examined; ± indicates expression in some lines examined; - indicates no GFP expression in any lines examined. Reporters labeled *pes-10::gfp* included the *pes-10* minimal promoter. Scale bars = 100 μm.

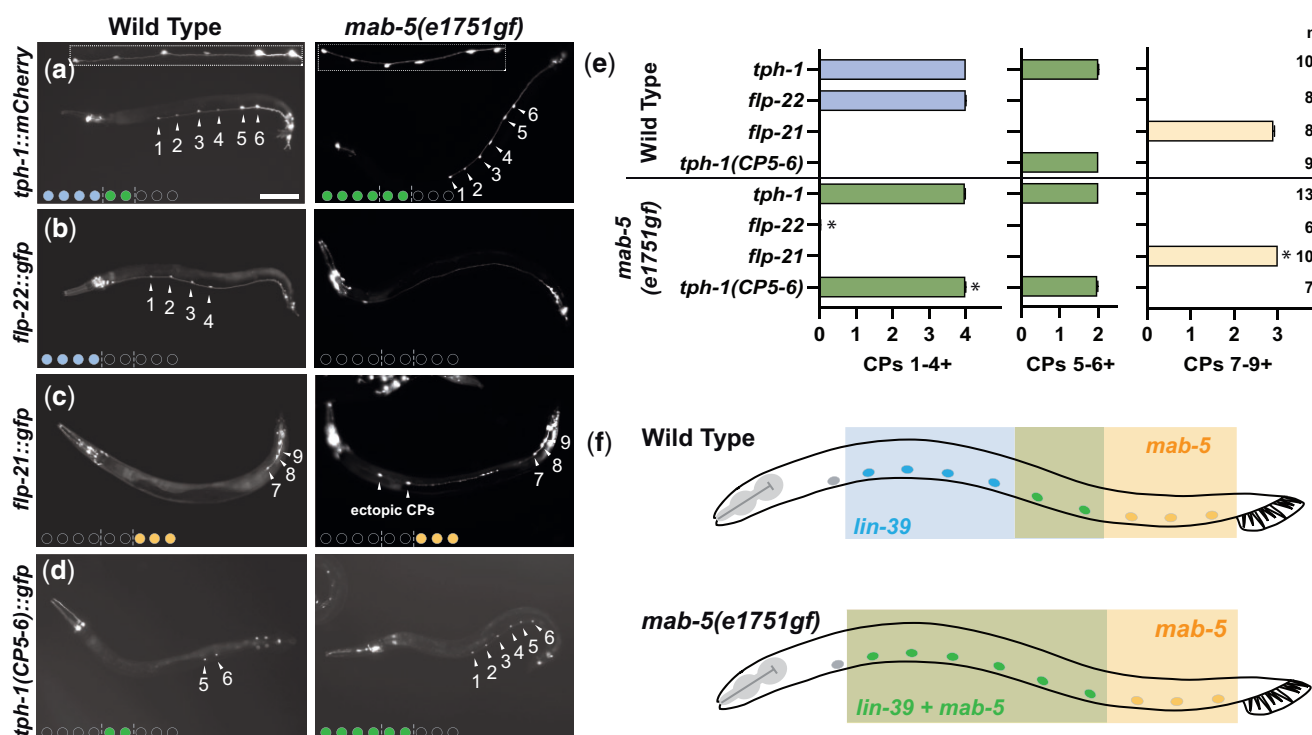


Fig. 3. Anterior expansion of MAB-5 specifies zone 2-like fates in zone 1 CPs. a–d) CP reporter expression in wild-type (left) and *mab-5*(e1751gf) (right) males. Filled circle colors summarize zone identities of CP neurons; scale bar = 100 μ m. a) *tph-1::mCherry* expression in CPs 1–6 in wild-type and *mab-5*(e1751) males. Insets show more intense relative expression in CPs 5–6 than in CPs 1–4 in wild-type ($P < 0.0001$, $n = 15$, unpaired t-test), but equally intense expression throughout CPs 1–6 in *mab-5*(e1751gf) ($P > 0.7$, $n = 17$, unpaired t-test). b) *flp-22::gfp* is expressed in CPs 1–4 in a wild-type male, but is not expressed in CPs in a *mab-5*(e1751gf) male. c) *flp-21::gfp* is expressed in CPs 7–9 in both wild-type and *mab-5*(e1751gf) males. *mab-5*(e1751) males also display expression of *flp-21::gfp* in ectopic anterior CPs. d) *tph-1*(CP5-6)::gfp is expressed in CPs 5–6 in a wild-type male, but expands to CPs 1–6 in a *mab-5*(e1751gf) male. e) Summary of *tph-1::mCherry*, *flp-22::gfp*, *flp-21::gfp*, and *tph-1*(CP5-6)::gfp expression in CPs (CPs+ denotes mean number of fluor-positive CPs/worm in each zone) in wild-type and *mab-5*(e1751gf) males. Bar colors indicate zone identities of CP neurons; error bars indicate standard error of the mean (SEM); number of ventral cords scored is indicated at right; * $P < 0.05$, unpaired t-test, compared to WT. f) Diagram representing CP neuron homeotic transformations in *mab-5*(e1751gf), in which zone 1 neurons (CPs 1–4) assume fates normally reserved for zone 2 neurons (CP 5–6).

zone 1-specific *tph-1* CRM, which may indicate that our constructs disrupt elements required for its activity.

Our analysis of *flp-22* did identify a zone 1-specific CRM in *flp-22*: a 226 base-pair region (–1,759 to –1,533) is sufficient to drive GFP expression in zone 1 CPs 1–4 (Fig. 2e). We used a reporter containing a slightly larger regulatory element (*flp-22*(–2,060/–1,533)::gfp) to assess Hox-responsiveness of the zone 1 CRM in *flp-22*, as this reporter shows more consistent expression. In wild-type males, *flp-22*(–2,060/–1,533)::gfp is expressed zone 1-specifically, and requires *lin-39*, but not *mab-5*. In zone 2, *mab-5* is required to prevent expression of *flp-22*(–2,060/–1,533)::gfp, as expression expands into CPs 5–6 in the *mab-5*(e1239) mutant (Fig. 2, f–i). The identification of these zone 1- and zone 2-specific CRMs suggests that control of expression of CP identity genes is modular, with regulation exerted by Hox proteins in a region-specific manner.

Anterior expansion of MAB-5 causes zone 1 CPs 1–4 to adopt zone 2 CP fates

Because zone 2 fates are dependent on both *lin-39* and *mab-5*, we next asked whether expression of *lin-39* and *mab-5* together in zone 1 CPs 1–4 is sufficient to respecify zone 1 CP neurons to zone 2-like fates. To do so, we made use of the previously described *mab-5*(e1751gf) gain-of-function mutant, in which MAB-5 expression is expanded anteriorly into regions that normally express more anterior Hox proteins, including LIN-39 (Salser et al. 1993). We find this sufficient to transform CPs 1–4 into zone 2-like

neurons (Fig. 3). In wild-type males, *tph-1::mCherry* is normally expressed at low intensity in CPs 1–4, but, in *mab-5*(e1751gf), CPs 1–4 display the intense *tph-1::mCherry* expression that is normally characteristic of only zone 2 CPs 5–6 (Fig. 3, a and e). Furthermore, *mab-5*(e1751gf) CPs 1–4 fail to express *flp-22::gfp* (Fig. 3, b and e) and instead express the normally zone 2-specific *tph-1*(CP5-6)::gfp (Fig. 3, d and e). Taken together, these findings indicate that, in *mab-5*(e1751gf) males, CPs 1–4 resemble CPs 5–6, consistent with LIN-39 and MAB-5 collaborating to promote zone 2 CP fates in cells in which they are both expressed (Fig. 3f). Interestingly, in *mab-5*(e1751gf), *flp-21::gfp* is expressed ectopically in 2 cells anterior to CP1 that would normally undergo programmed cell death or fail to divide in wild-type males, but that survive and divide inappropriately in these mutants (Fig. 3c, Salser et al. 1993). *lin-39* is not expressed in these anterior cells; thus their adoption of a zone 3-like fate (expression of *flp-21* alone) is consistent with a model in which MAB-5 specifies zone 3 fates, acting without LIN-39.

The *lin-39*(ccc16) mutation affects zone 2 CP fates specifically and in a *mab-5*-dependent manner

Our reporter analyses indicate that zone-specific CP gene expression is controlled by modular cis-regulatory elements. Our analysis of an unusual *lin-39* allele, *lin-39*(ccc16), which we recovered in a genetic screen (Kalis et al. 2014), indicates that the specification of zone 2 CP fates is also genetically modular. *lin-39*(ccc16) mutants consistently lack expression of *tph-1* reporters in zone 2

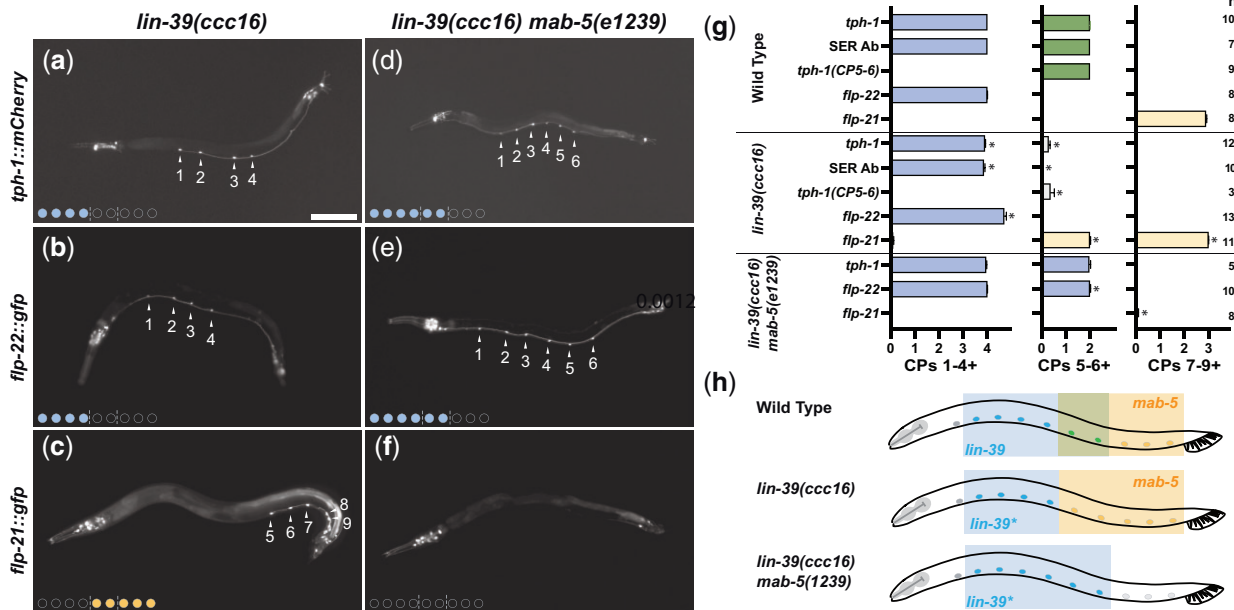


Fig. 4. The *lin-39(ccc16)* mutation affects zone 2 CP fates specifically and in a *mab-5*-dependent manner. a–f) CP reporter expression in *lin-39(ccc16)* and *lin-39(ccc16) mab-5(e1239)* males; refer to Fig. 3 for wild-type expression of CP reporters. *tph-1::mCherry* is expressed in CPs 1–4, but not CPs 5–6 in a *lin-39(ccc16)* male a), but is expressed in CPs 1–6 in a *lin-39(ccc16) mab-5(e1239)* male d). *flp-22::gfp* is expressed in CPs 1–4 in a *lin-39(ccc16)* male b), and in CPs 1–6 in a *lin-39(ccc16) mab-5(e1239)* male e). *flp-21::gfp* is expressed in CPs 5–9 in *lin-39(ccc16)* c), but is not expressed in CPs in a *lin-39(ccc16) mab-5(e1239)* f) male. Filled circle colors summarize zone identities of CP neurons; scale bar = 100 μ M. g) Summary of *tph-1::mCherry*, Serotonin (SER) antibody staining, *tph-1(CP5-6)::gfp*, *flp-22::gfp*, and *flp-21::gfp* expression in CPs (CPs+ denotes mean number of fluor-positive CPs/worm in each zone) in wild-type, *lin-39(ccc16)*, and *lin-39(ccc16) mab-5(e1239)* males. Bar colors indicate zone identities of CP neurons; error bars indicate standard error of the mean (SEM); number of ventral cords scored is indicated at right; * $P < 0.05$, unpaired t-test, compared to WT. h) Diagram representing CP neuron homeotic transformations in *lin-39(ccc16)* and *lin-39(ccc16) mab-5(e1239)* mutants. In *lin-39(ccc16)*, zone 2 neurons (CPs 5–6) assume fates normally reserved for zone 3 CP neurons (7–9). In *lin-39(ccc16) mab-5(e1239)*, the LIN-39(ccc16) protein is able to specify zone 1 fates in zone 2 CPs 5–6 in the absence of *mab-5*. *lin-39** denotes the *ccc16* isoform of *lin-39*.

CPs 5–6, but retain expression of *tph-1* in zone 1 CPs 1–4 (Fig. 4, a and g, refer to Fig. 3 for wild-type expression). Analysis of other markers of CP fate shows that *lin-39(ccc16)* males exhibit typical zone 1 CP fates, including expression of *tph-1::mCherry*, serotonin immunofluorescence, and *flp-22::gfp* in CPs 1–4 (Fig. 4, a, b, and g). In contrast, zone 2 fates are abnormal in *lin-39(ccc16)*, with *tph-1::mCherry*, serotonin immunofluorescence, and *tph-1(CP5-6)::gfp* dramatically reduced, and *flp-21::gfp* ectopically present (Fig. 4, a, c, and g). We hypothesize that the Zone-2-specific loss of *tph-1::mCherry* expression is due to a failure of the protein encoded by *lin-39(ccc16)* to negotiate interactions with MAB-5 that define zone 2 fates. If this is the case, we would expect the abnormal zone 2 fates of CPs 5–6 in *lin-39(ccc16)* to depend on *mab-5*, and this is what we see (Fig. 4, d–g). In *lin-39(ccc16) mab-5(e1239)* double mutants, *tph-1::mCherry* expression is restored in zone 2 CPs 5–6, and *flp-22::gfp* is expanded into zone 2, supporting the idea that LIN-39(ccc16) is able to function in the absence of *mab-5* to specify a zone 1 fate. Taken together, these findings provide strong genetic evidence that the *ccc16* lesion affects a portion of the LIN-39 protein that is required for LIN-39 to interact with MAB-5 to specify CP 5–6 fates (Fig. 4h).

To determine *lin-39(ccc16)*'s function relative to wild-type and other previously described *lin-39(lf)* alleles, we performed complementation analysis (Supplementary Fig. S1). The *lin-39(ccc16)* allele behaves recessively, with *lin-39(ccc16)/+* males producing the wild-type expression pattern of *tph-1::mCherry*. Complementation analysis with previously described *lin-39* alleles revealed that phenotypes of males with *lin-39(ccc16)* in trans to other *lin-39* alleles closely resemble the phenotypes of *lin-39(ccc16)* homozygotes. Taken together, these classical genetic analyses strongly

indicate that the *lin-39(ccc16)* allele produces a protein that is less (hypomorphic) or differently (neomorphic) functional than wild-type LIN-39, but which is not a dominant negative, in that it does not interfere with the wild-type protein.

***lin-39(ccc16)* is able to function in tissues outside the ventral nerve cord**

Given the apparent specificity of the *lin-39(ccc16)* phenotype, we wondered whether tissues outside the ventral cord were abnormal in *lin-39(ccc16)* mutants. In hermaphrodites, *lin-39* is normally required for vulval development; *lin-39(n1760)* mutants are vulvaless (Clark et al. 1993). In contrast, *lin-39(ccc16)* hermaphrodites have a functional vulva; they lay eggs and can be mated with males, although the vulva occasionally appears disorganized. This suggests that the protein encoded by *lin-39(ccc16)* can direct relatively normal vulval development. Because *ccc16* affects zone 2, where MAB-5 and LIN-39 are expressed together, we wondered whether *lin-39(ccc16)* would disrupt development of other tissues with a similar LIN-39/MAB-5 overlap. The male hyp7 syncytium is one such tissue, where LIN-39 and MAB-5 regulate the fusion of the Pn.p cells (sister cells of the Pn.a cells that give rise to ventral cord neurons) into the hyp7 syncytium along the A–P axis (Salser et al. 1993). In wild-type males in late L1, the Pn.p cells in zone 2 fuse to the hypodermis, while Pn.p cells in Zones 1 and 3 remain unfused. This fusion occurs in the presence of both *lin-39* and *mab-5*, or in the absence of both *lin-39* and *mab-5*, but not in the case where only 1 of the 2 is present. We wondered whether *lin-39(ccc16)* would affect Pn.p fusion across the domain of *lin-39* function, or if its phenotype might be specific to zone 2, and used the cell boundary marker *ajm-1::gfp* to assess

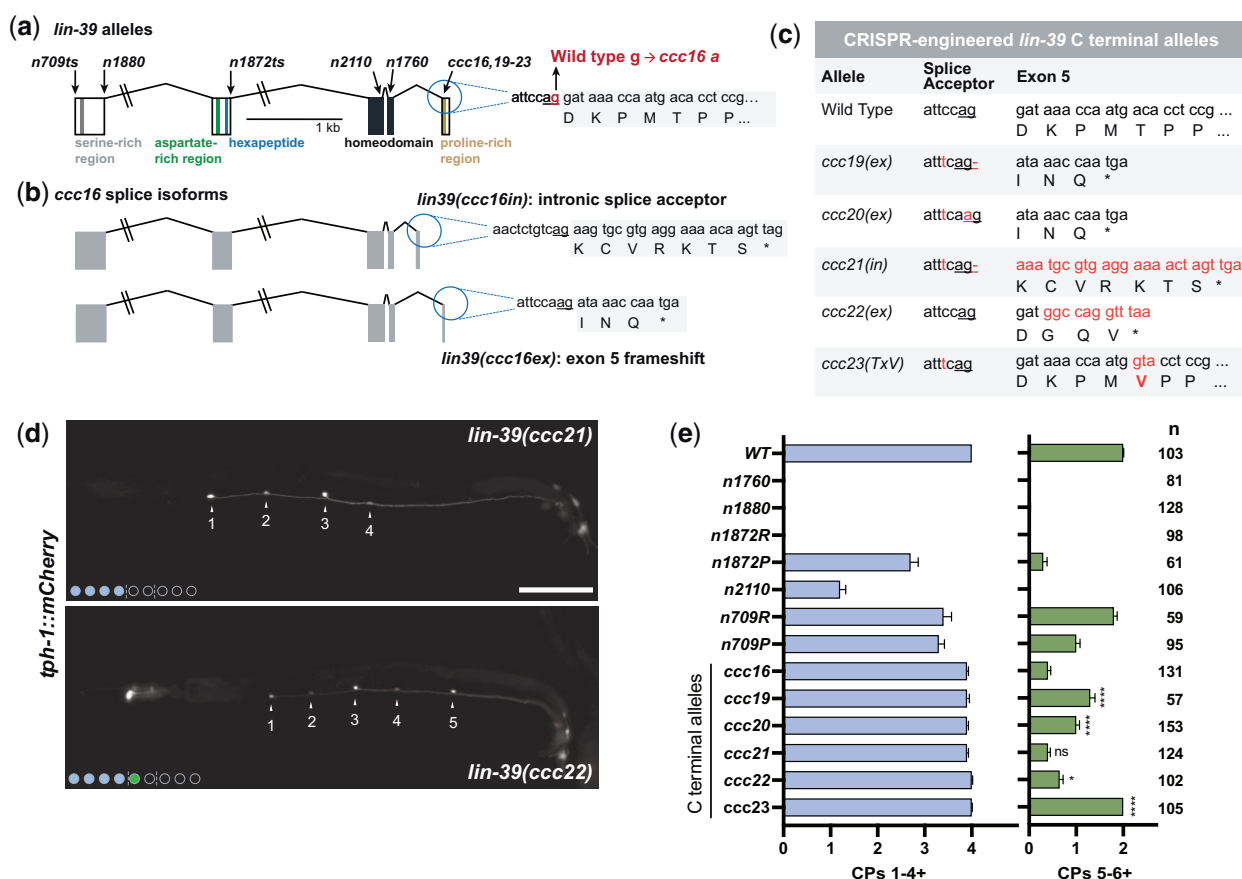


Fig. 5. *lin-39(ccc16)* disrupts sequences C-terminal to the homeodomain. a) Gene structure of *lin-39* indicating protein domains and alleles. Previously identified alleles of *lin-39* are located either before or in the homeodomain: *n709* and *n1872* are splice acceptor and donor mutations, respectively, *n1880* and *n1760* are substitutions that result in a stop codon, and *n2110* a substitution. DNA sequencing indicates that *ccc16* is a g->a exon 5 splice acceptor mutation. b) 3' RACE reveals 2 *ccc16* splice isoforms. Isoform *lin-39(ccc16in)* uses an intronic splice acceptor resulting in 7 amino acids before a stop codon (*). Isoform *lin-39(ccc16ex)* exon 5 splice acceptor results in a frameshift and early stop codon (*). c) CRISPR-engineered changes to *lin-39* exon 5 or splice acceptor. Sequence changes are indicated in red. ex, truncated exon variant; in, intronic splice variant; TxV, valine substitution for threonine in consensus MAPK site. d) Reduced expression of *tph-1::mCherry* in CPs 5–6 in *lin-39(ccc21)* and *lin-39(ccc22)*. Filled circle colors summarize zone identities of CP neurons; scale bar = 100 μ m. e) Summary of *tph-1::mCherry* expression in CPs 1–6 (CPs+ denotes mean number of fluor-positive CPs/worm in each zone) in *lin-39* mutant alleles. *lin-39* temperature-sensitive alleles were scored at both the permissive (P) temperature of 16°C and restrictive (R) temperature of 25°C. Significant differences of *tph-1::mCherry* expression in CPs 5–6 are indicated for comparisons between *lin-39(ccc16)* and other C-terminal alleles (ordinary one-way ANOVA with multiple comparisons, * $P < 0.05$, **** $P < 0.0001$, ns = not significant). Bar colors indicate zone identities of CP neurons; error bars indicate standard error of the mean (SEM); number of ventral cords scored is indicated at right. In e), for CPs 1–4, all means are significantly different from WT at $P < 0.05$ (unpaired t-test) except for *ccc19*, *ccc22*, and *ccc23*; for CPs 5–6, all means are significantly different from WT at $P < 0.05$ except for *ccc23*.

Pn.p fusion. We found that *lin-39(ccc16)* males resemble Wild-Type in zone 1 and occasionally fail to fuse Pn.ps in zone 2 (Supplementary Table S1). Although this weak phenotype is zone 2-specific, this result indicates that the LIN-39(ccc16) protein is largely able to regulate Pn.p fusion in either zone 1 or 2, and suggests that the effects on CP neuron specification are somewhat specific to *lin-39* function in the ventral cord neuron context.

lin-39(ccc16) disrupts sequences C-terminal to the homeodomain

We hypothesize that *lin-39(ccc16)* encodes a protein that is able to function in CPs in the absence of MAB-5, but unable to function in the presence of MAB-5. If this is the case, its molecular properties should shed light on how LIN-39, MAB-5, and their cofactors interact to define neuronal fate. Whole-genome sequencing revealed that *lin-39(ccc16)* disrupts a splice acceptor for the final exon of *lin-39* (Fig. 5a); 3' RACE identified 2 splice products in *lin-39(ccc16)* mutants; both encode proteins that replace the 28 amino acids of exon 5 with alternative peptides. One splice

product, *lin-39(ccc16ex)*, results in a protein in which the last exon encodes the sequence INQ due to a frameshift; the other, *lin-39(ccc16in)*, encodes a protein in which the last exon encodes the sequence KCVRKTS due to an intronic alternative splice site (Fig. 5b). Both of these truncated proteins are disrupted after the homeodomain, which is consistent with our findings that LIN-39(ccc16) remains functional in many cells, including the vulva, Pn.ps, and zone 1 CPs 1–4.

The *lin-39(ccc16)* phenotype indicates that the C-terminus of LIN-39 is important for interactions with MAB-5 that define zone 2 CP neuronal fates. We next asked whether this zone 2-specificity is a consequence of any loss of *lin-39* function, or if it is specific to lesions, such as *lin-39(ccc16)*, that disrupt sequences C-terminal to the homeodomain. To address this question, we examined previously described *lin-39* null and hypomorphic alleles. Males bearing mutations that alter key homeodomain residues or truncate the protein prior to the homeodomain (*n1760*, *n1872R*, *n1880*, *n2110*) lack expression of *tph-1::mCherry* in all or most CPs (Fig. 5e). This is unsurprising, as previous studies (Clark et al.

1993; Salser et al. 1993) have shown that the precursors of CPs 1–4 require *lin-39* to survive and divide, and CPs 1–6 require *lin-39* for expression of *tph-1*. Previously described weak alleles do disproportionately affect CPs 5–6; however, neither of those we examined (*n1872P*, *n709*) appears to be as specific to zone 2 as *lin-39(ccc16)*, in which zone 1 neurons are almost completely normal with respect to *tph-1::mCherry*.

Previous studies identified a consensus MAPK phosphorylation sequence PMTP C-terminal to the homeodomain of *lin-39* (Grandien and Sommer 2001; Wagmaister et al. 2006a). This site is phosphorylated in vitro by Erk2 (Wagmaister et al. 2006a), although transgenes bearing this mutation are able to rescue vulval development in *lin-39(n1760)*, suggesting this site is dispensable for *lin-39*'s function in vulval development (Grandien and Sommer 2001). Both *lin-39(ccc16)* splice variants remove the consensus MAPK target site as well as a putative downstream MAPK interaction domain. Given that *lin-39(ccc16)* mutants do produce a functional vulva, we asked whether a mutation in the putative MAPK phosphorylation site might disrupt LIN-39 function in CP neurons. We used CRISPR to replace the threonine in the PMTP sequence with valine. This mutant, *lin-39(ccc23)*, has wild-type expression of *tph-1::mCherry* in CPs 1–6 (Fig. 5, c and e) suggesting that, as in vulval development, the putative MAPK site is dispensable for LIN-39 function.

We next considered whether loss of the C-terminal exon (as in the *ccc16ex* isoform) could be responsible for the zone 2-specific defects in *lin-39(ccc16)*, or, alternatively, whether the intronic splice (*lin-39ccc16in*) might produce a neomorphic protein that interferes with CP5–6 specification. We used CRISPR to create a series of alleles predicted to encode proteins that mimic each isoform (Fig. 5c). Alleles that, like *lin-39(ccc16ex)*, truncate the C-terminal exon (*ccc19*, *ccc20*, *ccc22*) partially phenocopy *lin-39(ccc16)* in that CPs 1–4 are relatively normally specified; however, more zone 2 CPs 5–6 retain *tph-1::mCherry* expression than in *lin-39(ccc16)* (Fig. 5, d and e). In contrast, the *lin-39(ccc21)* allele, which most closely resembles the product of the intronic splice, phenocopies *lin-39(ccc16)* closely, with normal expression of *tph-1::mCherry* expression in zone 1 CPs 1–4 and nearly complete loss of expression in zone 2 CPs 5–6. This closer resemblance of *lin-39(ccc21)* to *lin-39(ccc16)* supports the idea that a neomorphic function of the C-terminal peptide KCVRKTS contributes to the *lin-39(ccc16)* phenotype. Taken together, our analyses of C-terminal alleles of *lin-39* indicate that the C terminus of LIN-39 is key to coordinating LIN-39/MAB-5 interactions in zone 2 CP specification.

Interactions between LIN-39 and MAB-5 depend on C-terminal sequences that are disrupted in *lin-39(ccc16)*

It is clear that *lin-39* and *mab-5* interact genetically to specify zone 2 CP5–6 fates, and that this interaction is disrupted by the *lin-39(ccc16)* mutation. We next asked whether LIN-39 and MAB-5 are associated with each other and with the TALE cofactor CEH-20 in worms. To test this, we used a Bimolecular Fluorescence Complementation (BiFC) assay in which we introduced pairs of transgenes, each encoding LIN-39, MAB-5, or CEH-20 fused in frame with a portion of the Venus fluorescent protein, under the control of a heat-shock promoter that is strongly active in intestinal nuclei (Hiatt et al. 2008; Shyu et al. 2008). In this assay, close proximity between Hox and/or TALE fusion proteins would bring the 2 halves of the Venus protein together and reconstitute fluorescence activity. We find that pairwise combinations of LIN-39, MAB-5, and CEH-20 reconstitute fluorescence in BiFC, but that

none of these reconstitutes fluorescence when paired with an empty vector (Fig. 6, a, b, e, f, and i; Supplementary Fig. S2). Previous work has established that Hox-TALE complex formation depends on monomer-DNA binding and can be disrupted by a substitution of amino acid 54 of the homeodomain (Hudry et al. 2011; Dard et al. 2018, 2019). Introduction of this homeodomain substitution in MAB-5 reduces interaction between LIN-39 and MAB-5, suggesting that the homeodomain is required (Supplementary Fig. S3). In contrast, we find that homeodomain substitutions in each of LIN-39 or CEH-20 do not significantly reduce their interactions in any pairwise combination. It is possible that only 1 homeodomain is necessary for strong interactions, which is consistent with previously published findings showing that such interactions may only be disrupted when both interacting partners contain the homeodomain substitution (Hudry et al. 2011; Dard et al. 2018, 2019).

The interactions of LIN-39 and MAB-5 with CEH-20 in this assay (Fig. 6, a–i, Supplementary Fig. S2) are not surprising, and are consistent with numerous examples from previously published work (Liu and Fire 2000; Liu et al. 2006; Wagmaister et al. 2006b; Potts et al. 2009). We asked whether proteins encoded by the *lin-39(ccc16ex)* and *ccc16in* isoforms disrupted these interactions, and find that both result in only partial reduction of interaction between LIN-39 and CEH-20, with only 1 of these reductions rising to the level of statistical significance (Fig. 6, a–d and j). This suggests that interaction between LIN-39 and CEH-20 may be less dependent on the C terminus, and may involve other domains of LIN-39. Surprisingly, we find that LIN-39 and MAB-5 are associated with one another in the BiFC assay. Given the genetic interactions between *lin-39* and *mab-5* indicated by our analysis of *lin-39(ccc16)* mutants, we hypothesized that this interaction might depend on the C terminus of LIN-39. Indeed, the LIN-39/MAB-5 BiFC interaction does depend on the C terminus of LIN-39, as it is reduced in constructs encoding the LIN-39(*ccc16ex*) and LIN-39(*ccc16in*) isoforms (Fig. 6, e–h and j). We also asked whether CEH-20 was required for the interaction between LIN-39 and MAB-5, and performed RNAi targeting *ceh-20* in conjunction with BiFC experiments to test this, and find that reduction of function of *ceh-20* by RNAi does reduce LIN-39/MAB-5 interaction (Fig. 6, k–m). Thus, in BiFC assays, we find that LIN-39, MAB-5, and CEH-20 are found together in vivo, and that, consistent with our genetic findings, their interactions are disrupted by the *lin-39(ccc16)* mutation.

Discussion

A model for Hox-mediated specification of CP neuron subtypes

We have taken a genetic approach to investigate the role of Hox transcription factors in patterning of CP neurons in *C. elegans* males. This work sheds light on the question of how Hox proteins can exert specificity and flexibility in target regulation, addressing the highly conserved process of axial patterning of neuronal subtypes in central nervous systems. Our findings support a hypothetical model in which 3 CRMs generate zone-specific neuronal subtype specificity (Fig. 7). At CRM 1 (e.g. in *tph-1* and *flp-22*), a complex containing LIN-39, CEH-20, and UNC-62 is required to activate gene expression, whereas MAB-5 represses expression. CRM 1 is inactive in regions where LIN-39 is not expressed or where CRM 1 is bound by MAB-5. At CRM 2 (e.g. in *tph-1*), both LIN-39 and MAB-5 are required for activation; this CRM is thus normally active only in zone 2, where both are expressed. At CRM 3 (e.g. in *flp-21*) MAB-5 is required to activate gene expression,

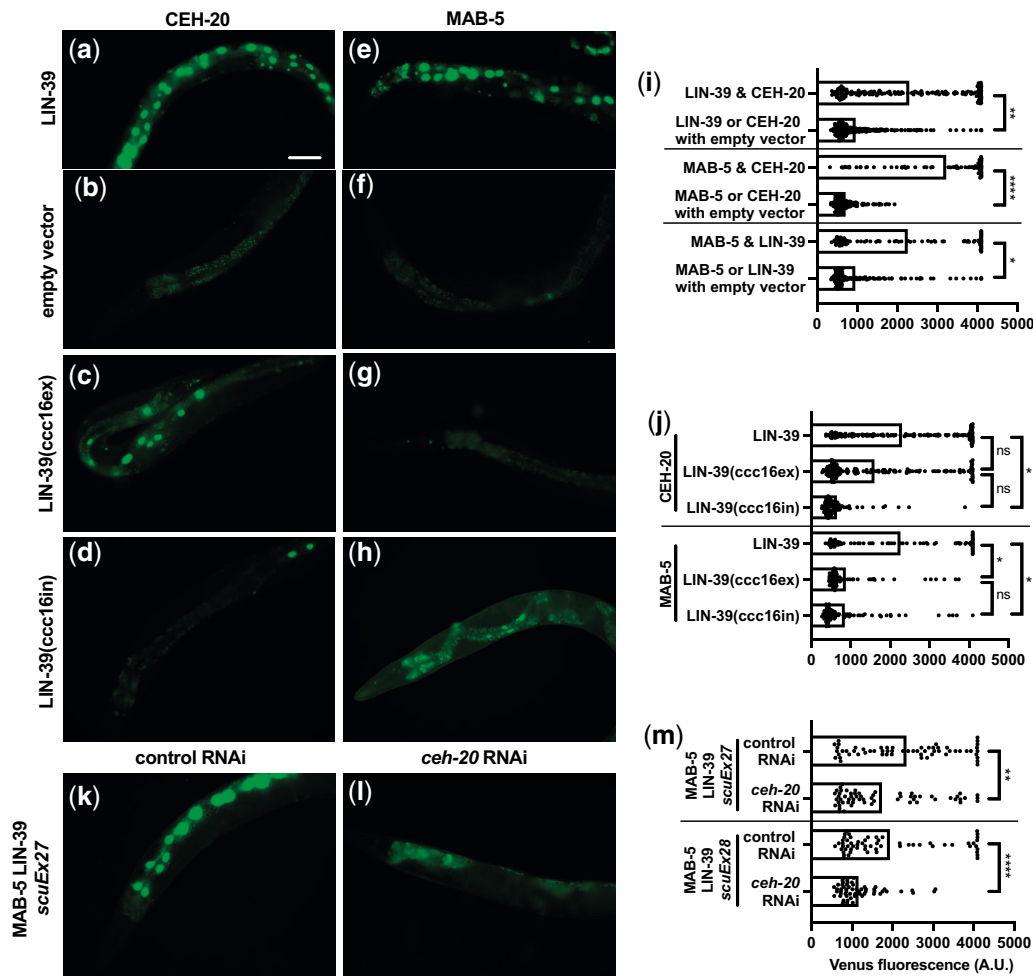


Fig. 6. Bimolecular fluorescence complementation (BiFC) ectopic protein interaction assay reveals interaction between Hox and TALE proteins. LIN-39, LIN-39(ccc16ex), LIN-39(ccc16in), MAB-5, or CEH-20 were fused in frame with the Venus fragments VN173 or VC155, and expressed in transgenic *C. elegans* hermaphrodites via the hsp16.41 heat-shock promoter. a–h) GFP images of worms expressing CEH-20 or MAB-5 with LIN-39, empty vector, LIN-39(ccc16ex), or LIN-39(ccc16in). Ectopic in vivo interaction between LIN-39 and both CEH-20 a) and MAB-5 e) reconstitutes Venus fluorescence prominently in intestinal nuclei. b–d and f–h) No interaction or weaker interaction is observed between CEH-20 or MAB-5 and empty vector or LIN-39 C-terminal mutants. Scale bar = 75µm. i–j and m) Quantification of Venus fluorescence for each pair of proteins. One dot represents the average of 6 intestinal nuclei per worm, bars represent the mean of all worms for a given pair. i and j) For each pairwise combination, 30 worms in each of 3 extrachromosomal lines were examined (see also [Supplementary Fig. 1](#)). i) CEH-20 reconstitutes fluorescence with LIN-39 and MAB-5; additionally, LIN-39 interacts with MAB-5 (nested unpaired t-test). j) LIN-39 C-terminal mutations disrupt LIN-39 and MAB-5 interactions but only partially disrupt LIN-39 and CEH-20 interactions (nested one way ANOVA). k–m) RNAi of *ceh-20* significantly disrupts reconstitution of fluorescence between LIN-39 and MAB-5 in 2 extrachromosomal lines, scuEx27 and scuEx28 ($n = 60$ for each, unpaired t-test). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.0001$, A.U., arbitrary unit, ns, not significant.

whereas LIN-39 represses expression. CRM 3 is inactive in regions where MAB-5 is not expressed or where CRM 3 is bound by the LIN-39 complex.

This model is supported by loss- and gain-of function experiments using both null and homeotic Hox mutations. The homeotic *lin-39(ccc16)* allele encodes a protein that appears to affect interactions between MAB-5 and LIN-39 specifically. Our work supports the hypothesis that LIN-39(ccc16) is functional at CRM 1 when MAB-5 is absent, but that LIN-39(ccc16) is not active at CRM 2 in coordination with MAB-5. Thus, *lin-39(ccc16)* exhibits a homeotic phenotype in which zone 2 neurons assume zone 3 fates. The *mab-5(e1751gf)* mutants likewise display homeotic phenotypes with respect to neurotransmitter expression, in that zone 1 neurons assume zone 2 fates. Furthermore, in *mab-5(e1751gf)*, neurons anterior to zone 1 can assume a zone 3 fate: *flp-21* is ectopically expressed in CPs anterior to CP 1–4 ([Fig. 3c](#)), which survive and divide in these mutants ([Salser et al. 1993](#)).

Previous work has shown that MAB-5 protein is expressed in an anterior-low to posterior-high gradient in the ventral cord ([Salser et al. 1993](#)), raising the possibility that the zone-specific functions of MAB-5 could be a consequence of expression levels. However, we do not think this is the case, given that *mab-5(e1751gf)* is unlikely to recapitulate the typical levels of MAB-5 expression, but nonetheless specifies relatively normal zone 2 and 3 fates as well as ectopic zone 2 fates in zone 1.

Our reporter deletion analysis provides further support for our model, as we identified a 199 base-pair element that likely contains a CRM 2 in *tph-1* and a 226 base-pair element that likely contains a CRM 1 in *flp-22*. However, we were unable to isolate a similar CRM 1 element in *tph-1*, despite strong genetic evidence for its existence. This inability to find a *tph-1* CRM 1 could be because the boundaries of our reporter constructs disrupted key sequences, because CRM 1 activity is distributed over multiple separate sequences, or because CRM 1 overlaps with CRM 2 in *tph-1*.

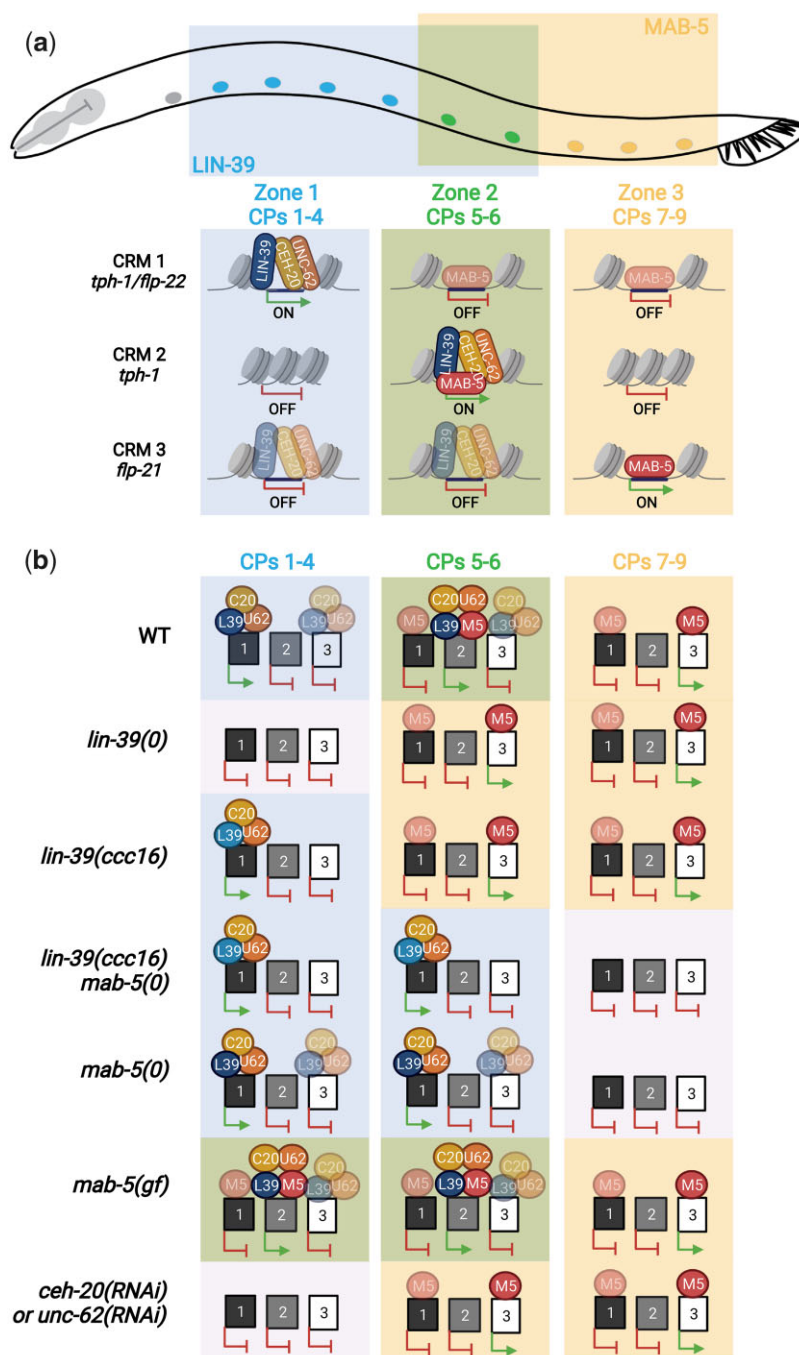


Fig. 7. Model of Hox and TALE regulation of CP neuronal subtype specification along the A-P axis. a) In wild-type males, CRM 1 is active in CPs 1–4, CRM 2 is active in CPs 5–6, and CRM 3 is active in CPs 7–9. b) Summary of regional (CPs 1–4, CPs 5–6, and CPs 7–9) expression at each CRM (1, 2, and 3) in Wild Type, *lin-39(0)*, *lin-39(ccc16)*, *lin-39(ccc16) mab-5(0)*, *mab-5(0)*, *mab-5(gf)*, and *ceh-20(RNAi)* or *unc-62(RNAi)*. a and b) Bar colors indicate zone-typical reporter expression, such that zone 1 = blue, zone 2 = green, zone 3 = yellow, no Zone specified = gray. Green arrows = active expression, red inhibitor = repression, CRM = cis-regulatory module, gf = gain-of-function, 0 = null, dark blue L39 = LIN-39, light blue L39 = LIN-39(ccc16), gold C20 = CEH-20, orange U62 = UNC-62, red M5 = MAB-5. Created with BioRender.com.

Models of Hox specificity: role of trans- and cis-regulation

Both flexibility and specificity of Hox regulatory systems can arise from interactions among *trans*-acting factors, such as Hox and TALE proteins, and from the structure of cis-regulatory elements in target genes. Previous work has established that Hox proteins can regulate gene expression in partnership with cofactors or without TALE cofactors (reviewed in Mann et al. 2009;

Merabet and Hudry 2013; Merabet and Mann 2016; De Kumar and Darland 2021). Additionally, binding sites can exhibit variation in affinity and specificity for their regulatory proteins, ranging from high specificity to a particular Hox protein, to high flexibility, as exemplified by semi-Hox specificity or general specificity for any Hox protein. Moreover, binding sites can influence the nature of the Hox cofactor interactions (Dard et al. 2018, 2019; Bridoux et al. 2020). Hox regulatory complexes can also recruit collaborators

that ultimately determine the sign of transcriptional regulation (Mann et al. 2009). Consistent with these previously proposed models, our work suggests that CP neuronal subtype specificity is brought about by both Hox-TALE interactions and TALE-independent Hox-mediated regulation. Because we see evidence of opposing positive and negative regulation by LIN-39 and MAB-5 at both CRM 1 and CRM 3, these may be examples of semi-Hox specific CRMs. However, we also provide evidence for cooperative Hox-Hox interactions at a single CRM, as LIN-39 and MAB-5 are together both necessary and sufficient to activate expression at CRM 2. One example of such Hox-Hox interaction has been described in *Drosophila*, where dimerization increases the functional specificity of the Hox Protein Scr (Sex combs reduced) (Papadopoulos et al. 2012).

Interactions among Hox and TALE proteins

The results of our BiFC ectopic protein interaction assays lend support to the genetic evidence that the Hox protein LIN-39 acts together with the Pbx-TALE cofactor CEH-20 to activate CRM 1, and repress CRM 3. We find that LIN-39 and CEH-20 can interact (Fig. 6a); our genetic evidence suggests that UNC-62 also participates in this putative complex, which is confirmed by published BiFC experiments demonstrating interaction between CEH-20 and UNC-62 (Van de Walle et al. 2019). The LIN-39/CEH-20 interaction is consistent with previous studies that show this genetic and in vitro interaction on the M lineage target gene *hlh-8* (Liu and Fire 2000), in autoregulation of *lin-39* (Wagmaister et al. 2006b), and in repression of the proapoptotic gene *egl-1* (Potts et al. 2009). In contrast, the *flp-21* CRM 3 is MAB-5-responsive, but does not require UNC-62 or CEH-20. However, MAB-5 and CEH-20 do appear to interact in our ectopic binding assay (Fig. 6a). This could be explained by an interaction among LIN-39, MAB-5, and CEH-20 on CRM 2, or could be explained by interactions between MAB-5 and CEH-20 at other target sites in the genome. For example, MAB-5 and CEH-20 bind in vitro to regulatory regions in the programmed cell death gene, *egl-1* (Liu et al. 2006).

We provide evidence that a complex containing LIN-39, CEH-20, and UNC-62 can both activate CRM 1 and repress CRM 3 in patterning of neuronal cell types. The *ccc16* mutation disrupts LIN-39's ability to repress CRM 3 in *flp-21*, but not its ability to activate CRM 1 in *flp-22*, suggesting that the LIN-39 protein domains required to interact with CEH-20 may be different with respect to the 2 binding site contexts. This is consistent with mouse in vitro and in vivo binding assays that show differing binding sites can change the protein domains with which Hox and PBX TALE proteins interact (Dard et al. 2018, 2019). Flexibility of LIN-39/CEH-20 interactions may likewise depend on context, which may explain why in an ectopic protein interaction assay like BiFC, the LIN-39(*ccc16*) and CEH-20 interaction results in variable levels of fluorescence reconstitution (Fig. 6); when both CRM 1 and CRM 3 are presumably available, LIN-39(*ccc16*) is only able to interact with CEH-20 on CRM 1 but not CRM 3.

Hox-Hox interactions

Since the first descriptions of *lin-39* and *mab-5* in *C. elegans*, it has been clear that they interact genetically in many contexts, and via varying apparent mechanisms (Kenyon 1986; Costa et al. 1988; Salser and Kenyon 1992; Clark et al. 1993; Salser et al. 1993). In the context of Pn.p fusion in late L1, the presence of either LIN-39 or MAB-5 prevents fusion of Pn.p cells with the hypodermis, whereas having neither or both promotes fusion. Alper and Kenyon (2002) proposed that fusion is mediated by LIN-39 and MAB-5 binding together to repress the expression of *ref-2*, which

itself encodes a negative regulator of fusion, but that their individual binding activates *ref-2* and thus prevents fusion. This suggests that LIN-39 and MAB-5 can bring about opposite transcriptional outcomes depending on whether they act alone or together. The *lin-39(ccc16)* mutation, which disrupts the cooperative activating function on LIN-39 and MAB-5 but has little effect on Pn.p fusion in males, suggests that the interactions we propose in CPs are distinct from those proposed for regulation of *ref-2* in Pn.p fusion. In CP neuron targets, LIN-39 and MAB-5 together have an activating effect on expression of CRM 2, but can each either activate or repress at CRM 1 and CRM 3.

Recent studies of ventral cord neuron development in hermaphrodites also suggest a co-regulatory role for LIN-39 and MAB-5, in this case in conjunction with the terminal selector UNC-3 (Feng et al. 2020). Here, LIN-39 and MAB-5 promote subtype-specific fates in cholinergic neurons in the presence of UNC-3. In *unc-3* mutants, LIN-39 no longer regulates cholinergic targets, and instead promotes expression of noncholinergic motor neuron fates in a *mab-5*-independent manner. This study provides an intriguing example of convergence of LIN-39 and MAB-5 on motor neuron subtype specification. However, *unc-3* is not expressed in the CP neurons (Kratsios et al. 2017) and thus other factors that converge on CP fate may remain to be identified. Given the male-specificity of CP neurons, the molecules of the *C. elegans* sex determination pathway are candidates for such roles (Wolff and Zarkower 2008). ChIP experiments in hermaphrodites demonstrate direct binding of both LIN-39 and MAB-5 to at least 1 cholinergic target, *unc-129*, providing strong evidence that coregulation could occur via coincident binding to DNA (Kratsios et al. 2017). Furthermore, global analysis of LIN-39 and MAB-5 via ChIP in hermaphrodites has demonstrated that LIN-39 and MAB-5 share several potential targets; in vivo studies verifying these targets will be needed to determine the nature of these interactions (Niu et al. 2011; Araya et al. 2014).

We find that LIN-39 can interact with MAB-5 in an ectopic protein interaction assay in which both proteins are overexpressed, and that this interaction is disrupted by *lin-39* mutations that affect the C-terminus of the protein. The interaction between LIN-39 and MAB-5 appears to be partially dependent on CEH-20, as RNAi targeting *ceh-20* reduces this interaction in BiFC assays. This is consistent with the observation that LIN-39 and MAB-5 share ChIP targets in hermaphrodites. Whereas global ChIP experiments have made the identification of LIN-39 and MAB-5 targets in hermaphrodites widely available, these community efforts have not yet been completed for males. Our genetic and ectopic protein interaction (BiFC) data support a model in which LIN-39 and MAB-5 interact with TALE factors in varying combinations on CRMs in *tph-1*, *flp-22*, and *flp-21* in males. Given the limitations of BiFC assays, future analyses of LIN-39 and MAB-5 binding in males will shed light on whether this model is supported by in vivo ChIP assays.

Conclusion

This study highlights the continued relevance of the genetic approach to understanding patterning and specification in development. This work has its foundation in an unusual allele identified in a genetic screen for mutations leading to abnormal male neuronal development. This allele, *lin-39(ccc16)*, leads to homeotic transformation of a regionally-specific neuron subtype, and demonstrates the modular nature of Hox activity in the male ventral nerve cord of *C. elegans*. Our model for interactions among the Hox proteins LIN-39 and MAB-5 and their cofactors on

neuronal genes will provide a springboard for experiments that directly test these interactions and for future screens to identify new regulators of neuronal subtype specification. Furthermore, this work highlights the usefulness of non-null alleles, both those generated in screens and by CRISPR, in providing high-resolution gene and protein functional information.

Data availability

The data underlying this article are available in the article and in its online [supplementary material](#).

[Supplemental material](#) is available at GENETICS online.

Acknowledgments

Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). We thank Don Moerman and Stephane Flibotte of the *C. elegans* Knockout Consortium for help with whole-genome sequencing assembly and Chris Li and Liesbet Temmerman for reagents. We thank Shawn Galdeen, Rou-Jia Sung, and Eric Hoopfer for their helpful feedback on the manuscript. We also thank students in the St. Catherine University Fall 2016 and 2018 and Carleton College Winter 2017 and 2018 Genetics courses and the following students for their contributions: Kristina Lodahl, Megan Snyder, Rebecca Luh, and Georgia Schmitt.

Funding

This work was supported by the National Science Foundation Major Research Instrumentation grant 1827514 (St. Catherine University) and the National Science Foundation Research at Undergraduate Institutions grant IOS-1021705 (Carleton College).

Conflicts of interest

None declared.

Literature cited

- Alper S, Kenyon C. The zinc finger protein REF-2 functions with the Hox genes to inhibit cell fusion in the ventral epidermis of *C. elegans*. *Development*. 2002;129(14):3335–3348.
- Araya CL, Kawli T, Kundaje A, Jiang L, Wu B, Vafeados D, Terrell R, Weissdepp P, Gevirtzman L, Mace D, et al. Regulatory analysis of the *C. elegans* genome with spatiotemporal resolution. *Nature*. 2014;512(7515):400–405.
- Arribere JA, Bell RT, Fu BXH, Artilles KL, Hartman PS, Fire AZ. Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in *Caenorhabditis elegans*. *Genetics*. 2014;198(3):837–846.
- Baek M, Enriquez J, Mann RS. Dual role for Hox genes and Hox cofactors in conferring leg motoneuron survival and identity in *Drosophila*. *Development*. 2013;140(9):2027–2038.
- Bridoux L, Zarrineh P, Mallen J, Phuycharoen M, Latorre V, Ladam F, Losa M, Baker SM, Sagerstrom C, Mace KA, et al. HOX paralogs selectively convert binding of ubiquitous transcription factors into tissue-specific patterns of enhancer activation. *PLoS Genet*. 2020;16(12):e1009162.
- Carnell L, Illi J, Hong SW, McIntire SL. The G-protein-coupled serotonin receptor SER-1 regulates egg laying and male mating behaviors in *Caenorhabditis elegans*. *J Neurosci*. 2005;25(46):10671–10681.
- Chow KL, Emmons SW. HOM-C/Hox genes and four interacting loci determine the morphogenetic properties of single cells in the nematode male tail. *Development*. 1994;120(9):2579–2592.
- Clark SG, Chisholm AD, Horvitz HR. Control of cell fates in the central body region of *C. elegans* by the homeobox gene *lin-39*. *Cell*. 1993;74(1):43–55.
- Clark SG, Chiu C. *C. elegans* ZAG-1, a Zn-finger-homeodomain protein, regulates axonal development and neuronal differentiation. *Development*. 2003;130(16):3781–3794.
- Concordet J-P, Haeussler M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res*. 2018;46(W1):W242–W245.
- Costa M, Weir M, Coulson A, Sulston J, Kenyon C. Posterior pattern formation in *C. elegans* involves position-specific expression of a gene containing a homeobox. *Cell*. 1988;55(5):747–756.
- Dard A, Jia Y, Reboulet J, Bleicher F, Lavau C, Merabet S. The human HOXA9 protein uses paralog-specific residues of the homeodomain to interact with TALE-class cofactors. *Sci Rep*. 2019;9(1):5664–5612.
- Dard A, Reboulet J, Jia Y, Bleicher F, Duffraisie M, Vanaker J-M, Forcet C, Merabet S. Human HOX proteins use diverse and context-dependent motifs to interact with TALE class cofactors. *Cell Rep*. 2018;22(11):3058–3071.
- Dasen JS, Jessell TM. Hox networks and the origins of motor neuron diversity. *Curr Top Dev Biol*. 2009;88:169–200.
- Davis MW, Hammarlund M, Harrach T, Hullett P, Olsen S, Jorgensen EM. Rapid single nucleotide polymorphism mapping in *C. elegans*. *BMC Genomics*. 2005;6:118.
- De Kumar B, Darland DC. The Hox protein conundrum: the “specifics” of DNA binding for Hox proteins and their partners. *Dev Biol*. 2021;477:284–292.
- Duerr JS, Han H-P, Fields SD, Rand JB. Identification of major classes of cholinergic neurons in the nematode *Caenorhabditis elegans*. *J Comp Neurol*. 2008;506(3):398–408.
- Estacio-Gómez A, Díaz-Benjumea FJ. Roles of Hox genes in the patterning of the central nervous system of *Drosophila*. *Fly (Austin)*. 2014;8(1):26–32.
- Feng W, Li Y, Dao P, Aburas J, Islam P, Elbaz B, Kolarzyk A, Brown AE, Kratsios P. A terminal selector prevents a Hox transcriptional switch to safeguard motor neuron identity throughout life. *eLife*. 2020;9:e50065.
- Grandien K, Sommer RJ. Functional comparison of the nematode Hox gene *lin-39* in *C. elegans* and *P. pacificus* reveals evolutionary conservation of protein function despite divergence of primary sequences. *Genes Dev*. 2001;15(16):2161–2172.
- Hiatt SM, Shyu YJ, Duren HM, Hu C-D. Bimolecular fluorescence complementation (BiFC) analysis of protein interactions in *Caenorhabditis elegans*. *Methods*. 2008;45(3):185–191.
- Hudry B, Viala S, Graba Y, Merabet S. Visualization of protein interactions in living *Drosophila* embryos by the bimolecular fluorescence complementation assay. *BMC Biol*. 2011;9:5–18.
- Jarrell TA, Wang Y, Bloniarz AE, Brittin CA, Xu M, Thomson JN, Albertson DG, Hall DH, Emmons SW. The connectome of a decision-making neural network. *Science*. 2012;337(6093):437–444.
- Kalis AK, Kissiov DU, Kolenbrander ES, Palchick Z, Raghavan S, Tetreault BJ, Williams E, Loer CM, Wolff JR. Patterning of sexually dimorphic neurogenesis in the *Caenorhabditis elegans* ventral cord by Hox and TALE homeodomain transcription factors. *Dev Dyn*. 2014;243(1):159–171.
- Kenyon C. A gene involved in the development of the posterior body region of *C. elegans*. *Cell*. 1986;46(3):477–487.
- Kenyon CJ, Austin J, Costa M, Cowing DW, Harris JM, Honigberg L, Hunter CP, Maloof JN, Muller-Immerglück MM, Salser SJ, et al.

- The dance of the Hox genes: patterning the anteroposterior body axis of *Caenorhabditis elegans*. *Cold Spring Harb Symp Quant Biol.* 1997;62:293–305.
- Kim K, Li C. Expression and regulation of an FMRFamide-related neuropeptide gene family in *Caenorhabditis elegans*. *J Comp Neurol.* 2004;475(4):540–550.
- Kratsios P, Kerk SY, Catela C, Liang J, Vidal B, Bayer EA, Feng W, De La Cruz ED, Croci L, Consalez GG, et al. An intersectional gene regulatory strategy defines subclass diversity of *C. elegans* motor neurons. *Elife.* 2017;6:e25751.
- 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(4):641–650.
- 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(23):5179–5190.
- Liu KS, Sternberg PW. Sensory regulation of male mating behavior in *Caenorhabditis elegans*. *Neuron.* 1995;14(1):79–89.
- Loer CM, Kenyon CJ. Serotonin-deficient mutants and male mating behavior in the nematode *Caenorhabditis elegans*. *J Neurosci.* 1993;13(12):5407–5417.
- Loer CM, Rivard L. Evolution of neuronal patterning in free-living rhabditid nematodes I: ex-specific serotonin-containing neurons. *J Comp Neurol.* 2007;502(5):736–767.
- Mann RS, Lelli KM, Joshi R. Hox specificity unique roles for cofactors and collaborators. *Curr Top Dev Biol.* 2009;88:63–101.
- Merabet S, Hudry B. Hox transcriptional specificity despite a single class of cofactors: are flexible interaction modes the key? Plasticity in Hox/PBC interaction modes as a common molecular strategy for shaping Hox transcriptional activities. *Bioessays.* 2013;35(2):88–92.
- Merabet S, Mann RS. To be specific or not: the critical relationship between Hox and TALE proteins. *Trends Genet.* 2016;32(6):334–347.
- Niu W, Lu ZJ, Zhong M, Sarov M, Murray JI, Brdlik CM, Janette J, Chen C, Alves P, Preston E, et al. Diverse transcription factor binding features revealed by genome-wide ChIP-seq in *C. elegans*. *Genome Res.* 2011;21(2):245–254.
- Papadopoulos DK, Skouloudaki K, Adachi Y, Samakovlis C, Gehring WJ. Dimer formation via the homeodomain is required for function and specificity of Sex combs reduced in *Drosophila*. *Dev Biol.* 2012;367(1):78–89.
- Philippidou P, Dasen JS. Hox genes: choreographers in neural development, architects of circuit organization. *Neuron.* 2013;80(1):12–34.
- Potts MB, Wang DP, Cameron S. Trithorax, Hox, and TALE-class homeodomain proteins ensure cell survival through repression of the BH3-only gene *egl-1*. *Dev Biol.* 2009;329(2):374–385.
- Salser SJ, Kenyon C. Activation of a *C. elegans* Antennapedia homologue in migrating cells controls their direction of migration. *Nature.* 1992;355(6357):255–258.
- Salser SJ, Loer CM, Kenyon C. Multiple HOM-C gene interactions specify cell fates in the nematode central nervous system. *Genes Dev.* 1993;7(9):1714–1724.
- Schindelman G, Whittaker AJ, Thum JY, Gharib S, Sternberg PW. Initiation of male sperm-transfer behavior in *Caenorhabditis elegans* requires input from the ventral nerve cord. *BMC Biol.* 2006;4:26.
- Schinkmann K, Li C. Localization of FMRFamide-like peptides in *Caenorhabditis elegans*. *J Comp Neurol.* 1992;316(2):251–260.
- Serrano-Saiz E, Pereira L, Gendrel M, Aghayeva U, Bhattacharya A, Howell K, Garcia LR, Hobert O. A neurotransmitter atlas of the *Caenorhabditis elegans* male nervous system reveals sexually dimorphic neurotransmitter usage. *Genetics.* 2017;206(3):1251–1269.
- Shin MM, Catela C, Dasen J. Intrinsic control of neuronal diversity and synaptic specificity in a proprioceptive circuit. *Elife.* 2020;9:e56374.
- Shyu YJ, Hiatt SM, Duren HM, Ellis RE, Kerppola TK, Hu C-D. Visualization of protein interactions in living *Caenorhabditis elegans* using bimolecular fluorescence complementation analysis. *Nat Protoc.* 2008;3(4):588–596.
- Slattery M, Ma L, Négre N, White KP, Mann RS. Genome-wide tissue-specific occupancy of the Hox protein Ultrabithorax and Hox cofactor Homothorax in *Drosophila*. *PLoS One.* 2011a;6(4):e14686.
- Slattery M, Riley T, Liu P, Abe N, Gomez-Alcala P, Dror I, Zhou T, Rohs R, Honig B, Bussemaker HJ, et al. Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins. *Cell.* 2011b;147(6):1270–1282.
- Stiernagle T. Maintenance of *C. elegans*. The *C. elegans* Research Community, editor. *WormBook: the online review of C. elegans biology.* 2006. p. 1–11. doi/10.1895/wormbook.1.101.1, <http://www.wormbook.org>.
- Sulston JE, Horvitz HR. Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev Biol.* 1977;56(1):110–156.
- Sze JY, Victor M, Loer C, Shi Y, Ruvkun G. Food and metabolic signaling defects in a *Caenorhabditis elegans* serotonin-synthesis mutant. *Nature.* 2000;403(6769):560–564.
- Van de Walle P, Geens E, Baggerman G, José Naranjo-Galindo F, Askjaer P, Schoofs L, Temmerman L. CEH-60/PBX regulates vitellogenesis and cuticle permeability through intestinal interaction with UNC-62/MEIS in *Caenorhabditis elegans*. *PLoS Biol.* 2019 17(11):e3000499.
- Wagmaister JA, Gleason JE, Eisenmann DM. Transcriptional upregulation of the *C. elegans* Hox gene *lin-39* during vulval cell fate specification. *Mech Dev.* 2006a;123(2):135–150.
- 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.* 2006b;297(2):550–565.
- Wang BB, Müller-Immergluck MM, Austin J, Robinson NT, Chisholm A, Kenyon C. A homeotic gene cluster patterns the anteroposterior body axis of *C. elegans*. *Cell.* 1993;74(1):29–42.
- Weirauch MT, Yang A, Albu M, Cote AG, Montenegro-Montero A, Drewe P, Najafabadi HS, Lambert SA, Mann I, Cook K, et al. Determination and inference of eukaryotic transcription factor sequence specificity. *Cell.* 2014;158(6):1431–1443.
- White JG, Southgate E, Thomson JN, Brenner S. The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philos Trans R Soc Lond B Biol Sci.* 1986;314(1165):1–340.
- Whittaker AJ, Sternberg PW. Coordination of opposing sex-specific and core muscle groups regulates male tail posture during *Caenorhabditis elegans* male mating behavior. *BMC Biol.* 2009;7:33.
- Wicks SR, Yeh RT, Gish WR, Waterston RH, Plasterk RH. Rapid gene mapping in *Caenorhabditis elegans* using a high density polymorphism map. *Nat Genet.* 2001;28(2):160–164.
- Wolff JR, Zarkower D. Somatic sexual differentiation in *Caenorhabditis elegans*. *Curr Top Dev Biol.* 2008;83:1–39.
- Zheng C, Diaz-Cuadros M, Chalfie M. Hox genes promote neuronal subtype diversification through posterior induction in *Caenorhabditis elegans*. *Neuron.* 2015;88(3):514–527.

Communicating editor: O. Hobert