The homeodomain protein *hmbx-1* maintains asymmetric gene expression in adult *C. elegans* olfactory neurons

Bluma J. Lesch and Cornelia I. Bargmann[1]

Howard Hughes Medical Institute, Laboratory of Neural Circuits and Behavior, The Rockefeller University, New York, New York 10065, USA

Differentiated neurons balance the need to maintain a stable identity with their flexible responses to dynamic environmental inputs. Here we characterize these opposing influences on gene expression in *Caenorhabditis elegans* olfactory neurons. Using transcriptional reporters that are expressed differentially in two olfactory neurons, AWC^{ON} and AWC^{OFF}, we identify mutations that affect the long-term maintenance of appropriate chemoreceptor expression. A newly identified gene from this screen, the conserved transcription factor *hmbx-1*, stabilizes AWC gene expression in adult animals through dosage-sensitive interactions with its transcriptional targets. The late action of *hmbx-1* complements the early role of the transcriptional repressor gene *nsy-7*: Both repress expression of multiple AWC^{OFF} genes in AWC^{ON} neurons, but they act at different developmental stages. Environmental signals are superimposed onto this stable cell identity through at least two different transcriptional pathways that regulate individual chemoreceptor genes: a cGMP pathway regulated by sensory activity, and a *daf-7* (TGF-β)/*daf-3* (SMAD repressor) pathway regulated by specific components of the density-dependent *C. elegans* dauer pheromone.

[*Keywords*: Chemoreceptor, olfactory receptor; homeostasis; neuronal development; olfactory neuron]

Supplemental material is available at http://www.genesdev.org.

Received March 31, 2010; revised version accepted June 25, 2010.

Neurons and other long-lived cells are subject to ongoing modification throughout life, but use transcriptional strategies to maintain stable cell fates (Shirasaki and Pfaff 2002; Ringrose and Paro 2004). In some cells, transcription factors required for the establishment of cell fate remain active throughout life. In the mammalian immune system, the transcription factor Pax5 is required both for initial commitment to the B-cell lineage and for continued expression of B-cell identity: Deletion of Pax5 from committed pro-B cells resulted in their reversion to a multipotential state (Nutt et al. 1999; Mikkola et al. 2002). In other cases, dedicated transcriptional regulators maintain cell identity. For example, the *Caenorhabditis elegans aristaless* homolog *alr-1* is required to maintain, but not establish, functions of sensory glia (Tucker et al. 2005). In addition to cell-specific transcription factors, general chromatin remodeling factors can stabilize cell fates. The Polycomb group (PcG) genes in *Drosophila* maintain a precise Hox gene expression pattern after the disappearance of early developmental regulators (Busturia and Morata 1988; Cao et al. 2002; Muller and Kassis 2006), and the histone acetyltransferase system regulated by BET-1 and MYST maintains cell type-specific gene expression in *C. elegans* neuronal lineages (Shibata et al. 2010).

Sensory neurons are faced with the special challenge of maintaining a stable state while responding to a changing environment. In the senses of taste and smell, heterogeneous populations of sensory neurons express different chemoreceptor genes to detect different environmental chemicals (Buck and Axel 1991; Troemel et al. 1995; Vosshall et al. 1999; Clyne et al. 2000; Etchberger et al. 2007). The expression of chemoreceptor genes is initiated by innate developmental programs. For example, in *C. elegans*, chemoreceptor expression in the two ASE taste neurons is initiated by a general transcriptional regulator for ASE, *che-1*, and refined by a double negative feedback loop that distinguishes right and left ASE fates (Chang et al. 2003, 2004; Johnston and Hobert 2003; Johnston et al. 2005; Etchberger et al. 2007, 2009). *che-1* maintains its own expression in the ASE chemosensory neurons, and also acts on target genes throughout life (Etchberger et al. 2009); this combination of initiation and differentiation roles defines *che-1* as a "terminal selector gene" (Hobert 2008). Analogous genetic studies of other chemosensory cell types—including the AWA, AWB, and AWC

[1]**Corresponding author.**
E-MAIL cori@rockefeller.edu; FAX (212) 327-7243.
Article is online at http://www.genesdev.org/cgi/doi/10.1101/gad.1932610.

olfactory neurons—have generated a sophisticated understanding of the transcription factors that initiate unique neuronal identities (Sengupta et al. 1994; Sagasti et al. 1999; Sarafi-Reinach and Sengupta 2000; Sarafi-Reinach et al. 2001; Colosimo et al. 2003; Lanjuin et al. 2003; Nokes et al. 2009; Kim et al. 2010).

Superimposed on stable chemosensory neuron fates are environmental factors that modify gene expression. The *C. elegans* dauer pheromone, a mixture of compounds containing the sugar ascarylose, represses the expression of chemoreceptor genes in ASH and ASI neurons by regulating intercellular signaling through a TGF-β signaling pathway (Peckol et al. 2001; Nolan et al. 2002; Kim et al. 2009). A salt-inducible kinase affects chemoreceptor expression in AWB olfactory neurons by regulating the transcription factor MEF2 (Lanjuin and Sengupta 2002; van der Linden et al. 2007, 2008). The relationship between these environmental regulators and stable cell fates raises intriguing questions about the relative roles of fixed and variable aspects of neuronal function.

The two *C. elegans* AWC olfactory neurons provide a system in which the acquisition and maintenance of cell fates are distinct. The Otx transcription factor CEH-36 and the HMX/NLX homeodomain protein MLS-2 initiate a general AWC identity, which is subsequently maintained by *ceh-36* (Lanjuin et al. 2003; Kim et al. 2010). In addition to promoting the expression of an AWC-specific transcriptional program, *ceh-36* maintains its own expression, suggesting that *ceh-36* is the terminal selector gene in AWC (Kim et al. 2010). Later in embryogenesis, a stochastic cell fate decision causes the right and left AWC olfactory neurons to take on asymmetric fates, such that one AWC becomes AWC^{ON}, defined as a neuron that expresses the G protein-coupled receptor (GPCR) *str-2* and senses the odor butanone, and the other AWC becomes AWC^{OFF}, which expresses the GPCR *srsx-3* and senses the odor 2,3-pentanedione (Troemel et al. 1999; Wes and Bargmann 2001). The decision to become AWC^{ON} or AWC^{OFF} is made through a signaling pathway that generates the initial asymmetry of chemoreceptor gene expression and also drives asymmetric expression of the transcription factor NSY-7 in AWC^{ON} (Troemel et al. 1999; Lesch et al. 2009). After hatching, the initial signaling pathway becomes inactive, and NSY-7 maintains appropriate chemoreceptor expression in AWC^{ON}. In addition, a cGMP pathway regulated by olfactory signal transduction maintains post-embryonic expression of both *str-2* and *srsx-3* using two receptor-type guanylate cyclases (encoded by *odr-1* and *daf-11*), a cyclic-nucleotide gated cation channel (encoded by *tax-2* and *tax-4*), and a cGMP-responsive protein kinase (encoded by *egl-4*) (Troemel et al. 1999; Lesch et al. 2009).

To better understand the factors responsible for the stability of AWC^{ON} and AWC^{OFF} fates, we performed a screen for mutants that failed to maintain expression of one or both of the asymmetric AWC markers *str-2* and *srsx-3*. Here, we describe the mutants isolated from this screen. We identify the transcription factor *hmbx-1*, a homolog of mammalian HMBOX1, as a regulator of AWC receptor gene expression that acts primarily in adult animals. Using newly identified GPCR genes expressed in AWC neurons, we show that the maintenance of asymmetric receptor gene expression involves at least three repressor pathways: *nsy-7*, an AWC^{ON}-specific cell identity gene; *hmbx-1*, a dosage-sensitive repressor of AWC^{OFF} genes in AWC^{ON} neurons; and *daf-3*, a pheromone-regulated repressor that affects chemoreceptors on a gene-by-gene basis.

Results

Genetic pathways required for maintenance of GPCR expression in AWC

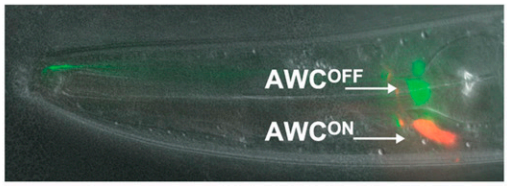
We sought mutants that expressed the AWC^{ON}-specific receptor *str-2* and the AWC^{OFF}-specific receptor *srsx-3* appropriately as early larvae, but failed to maintain expression of one or both of these receptors in adulthood. A strain with an integrated transgene containing *str-2::dsRed2* and *srsx-3::GFP* reporters was mutagenized, and the adult F2 progeny were examined for defects in expression of *str-2, srsx-3*, or both genes (Fig. 1A). After verifying the adult phenotype in subsequent generations, expression of *str-2* and *srsx-3* was evaluated in early larvae, 14 h after hatching (L1 stage), and mutants with a wild-type phenotype at this stage were retained. The screen yielded 19 mutants with defects in the maintenance of AWC markers (Fig. 1B). Genetic mapping and complementation testing indicated that 13 mutations fell in previously characterized genes and pathways, although only a subset had been known to affect AWC gene expression. These genes are described briefly below.

Olfactory transduction Six of the new mutations affect the olfactory cGMP transduction pathway that maintains *str-2* and *srsx-3* expression, including three alleles of *odr-1*, one allele of *daf-11*, one allele of *tax-2*, and one allele of *tax-4* (Fig. 1B; Troemel et al. 1999; Lesch et al. 2009).

Previous studies have revealed both cell-autonomous and nonautonomous effects of sensory signaling proteins on AWC gene expression (Lans and Jansen 2006). Moreover, *tax-4* promotes expression of *daf-7* (Coburn et al. 1998), which acts in ASI neurons to promote *srsx-3* expression in AWC neurons (see below). To ask where *tax-4* acts to regulate *srsx-3* expression, we expressed the *tax-4* cDNA in AWC or ASI in a *tax-4(ky791)* mutant background. AWC-selective expression of TAX-4 rescued the *srsx-3* expression defect of *tax-4(ky791)* mutants, but ASI expression did not (Supplemental Fig. S1). Therefore, *tax-4* acts in AWC to promote *srsx-3* expression.

The screen also yielded a dominant mutation in the Gα subunit *odr-3*, *odr-3(ky879)*, which encodes a G → S missense mutation at position 185. The affected glycine is a conserved residue in the region that changes conformation upon GTP binding; stabilization of the GTP-bound conformation should result in a constitutively active protein (Rens-Domiano and Hamm 1995). The nature of the *odr-3(ky879)* allele suggests that increased olfactory G protein activity disrupts maintenance of

A



B

genotype	str-2::dsRed 2 on ●●	1 on, bright ●○	1 on, faint ○○	2 off ○○	srsx-3::GFP 2 on ●●	1 on, bright ●○	1 on, faint ○○	2 off ○○	n
wild type	0	100	0	0	0	100	0	0	343
Group I: *srsx-3::GFP* defective									
tax-2(ky782) N329K, channel	1	85	4	10	0	0	4	96	210
tax-4(ky791) Q510STOP, after tm domain	1	70	13	16	0	0	10	90	286
daf-7(ky771) Q105STOP	0	97	0	3	0	0	5	95	66
daf-1(ky803) G400E, kinase	0	73	0	27	0	0	0	100	44
unc-3(ky793)	0	98	2	0	0	0	0	100	65
hmbx-1(ky777) H404Y	0	100	0	0	0	0	1	99	199
ky811	0	100	0	0	0	0	7	93	144
Group II: *str-2::dsRed* defective									
nsy-7(ky813)	0	0	0	100	98	0	0	2	53
nsy-7(ky819)	0	0	27	73	100	0	0	0	64
tam-1(ky798) G76S, zinc finger	0	8	24	68	0	95	5	0	248
ky789	0	0	11	89	3	97	0	0	38
ky802	0	0	0	100	0	92	8	0	52
Group III: both markers defective									
odr-1(ky779)	0	0	6	94	0	0	0	100	51
odr-1(ky800)	0	0	9	91	0	1	1	98	89
odr-1(ky807)	0	0	66	34	0	0	2	98	99
daf-11(ky801) D808V, cyclase domain	0	1	36	63	0	21	42	36	184
odr-3(ky879) G185S, GTP binding	0	9	16	74	0	61	23	17	160
ky772	0	2	58	40	2	2	96	0	43
ky773	0	0	12	88	0	0	100	0	52
known alleles									
odr-1(n1936)	0	0	18	82	0	0	9	91	34
daf-11(m47)	0	0	4	96	0	74	4	22	54
tax-2(ks31)	0	94	3	3	0	57	31	12	72
tax-4(p678)	0	90	10	0	0	0	13	87	39
nsy-7(tm3080)	0	0	0	100	95	5	0	0	86
tam-1(cc567)	0	86	12	2	2	98	0	0	43
daf-7(e1372)	0	100	0	0	0	0	0	100	31
daf-1(m40)	0	100	0	0	0	0	0	100	95

Figure 1. Expression phenotypes of mutants defective for *str-2::dsRed* and *srsx-3::GFP* maintenance. (*A*) DIC image of the head of an adult worm overlaid with a fluorescence image of the *str-2::dsRed2* (AWC^{ON}) and *srsx-3::GFP* (AWC^{OFF}) reporters. (*B*) Mutants from the screen. (Group I) *srsx-3* expression affected more strongly than *str-2* expression. (Group II) *str-2* expression affected more strongly than *srsx-3* expression. (Group III) Both markers affected. Other alleles of genes identified in the screen are shown for reference. Bright and faint fluorescence were scored qualitatively; in other figures, these two categories are combined into a single "1 AWC" class.

GPCR expression, perhaps by reducing cGMP levels (Chalasani et al. 2007). Loss-of-function mutations in *odr-3* have little effect on *str-2* expression, although *str-2* expression is reduced when *odr-3(lf)* alleles are combined with mutations in other Gα subunits (Lans and Jansen 2006).

Transcriptional regulation Two alleles of *nsy-7* isolated in the screen had defective maintenance of *str-2* expression, accompanied by bilateral *srsx-3* expression—the same phenotype observed in previously characterized *nsy-7* alleles (Lesch et al. 2009).

Another mutation affected the *tam-1* gene, which encodes a transcriptional regulator that inhibits silencing of repetitive transgenes (Hsieh et al. 1999). A null allele of *tam-1* also affects *str-2* expression, but had a weaker defect than the new missense allele from the screen (Fig. 1B). The stronger phenotype of the missense allele could result from an altered function of the mutant protein, or from modifying effects of background mutations. *tam-1*

was not known previously to affect gene expression in AWC, but diminished expression of the AWC reporter genes is consistent with the general reduction in transgene expression reported in *tam-1* mutants.

DAF-7/TGFβ signaling Several mutations from this screen primarily affected *srsx-3* and not *str-2* expression. These included an allele of *daf-7* (the TGFβ ligand that regulates the dauer developmental decision and chemoreceptor gene expression in ASH and ASI neurons) and an allele of *daf-1* (a TGFβ type I receptor) (Figs. 1B, 2A,B; Georgi et al. 1990; Ren et al. 1996; Schackwitz et al. 1996; Peckol et al. 2001; Nolan et al. 2002). A mutation in *unc-3*, a transcription factor required for expression of *daf-7* in the ASI neurons, was also isolated in the screen, and had a similar phenotype (Prasad et al. 1998; Kim et al. 2005). Existing *daf-4* (the type II TGFβ receptor), *daf-8* (a Smad-like protein), and *daf-14* (a Smad-like protein) mutants displayed similar adult phenotypes to *daf-7* and *daf-1* (Fig. 2C). These results indicate that TGFβ signaling maintains *srsx-3* expression in AWC^{OFF}.

daf-3, which encodes a co-Smad that binds DNA, and *daf-5*, which encodes a proline-rich transcriptional regulator, act downstream from and antagonistically to *daf-1* to promote dauer formation (Patterson et al. 1997; da Graca et al. 2004). Expression of *str-2* and *srsx-3* reporters was normal in *daf-3* and *daf-5* single mutants, and in *daf-3; daf-1* or *daf-5; daf-1* double mutants, recapitulating the regulatory relationships seen in dauer formation (Fig. 2B,C). Expression of DAF-1 in AWC rescued the *srsx-3* expression defect of *daf-1(m40)* mutants, suggesting that TGF-β signals directly to AWC to maintain receptor gene expression (Fig. 2C).

daf-7 expression is inhibited by the dauer pheromone, a mixture of structurally related chemicals termed ascarosides (Jeong et al. 2005; Butcher et al. 2007, 2008). To ask whether *srsx-3* expression responded acutely to pheromones, we exposed adult worms to the ascarosides C3, C6, and C9 for 4 h, and monitored expression of *srsx-3* using a destabilized GFP protein that has a half-life of ~1 h in *C. elegans* (Gaudet and Mango 2002; Frand et al. 2005). An equal mixture of C3, C6, and C9 suppressed *srsx-3* expression in a dose-dependent manner (Fig. 2D). The ascarosides C3 and C6 each suppressed *srsx-3* expression about as well as the mixture, while C9 was less effective. All effects of ascarosides were blocked in *daf-3* mutants, suggesting that C3 and C6 pheromones regulate *srsx-3* expression through the TGFβ pathway (Fig. 2D).

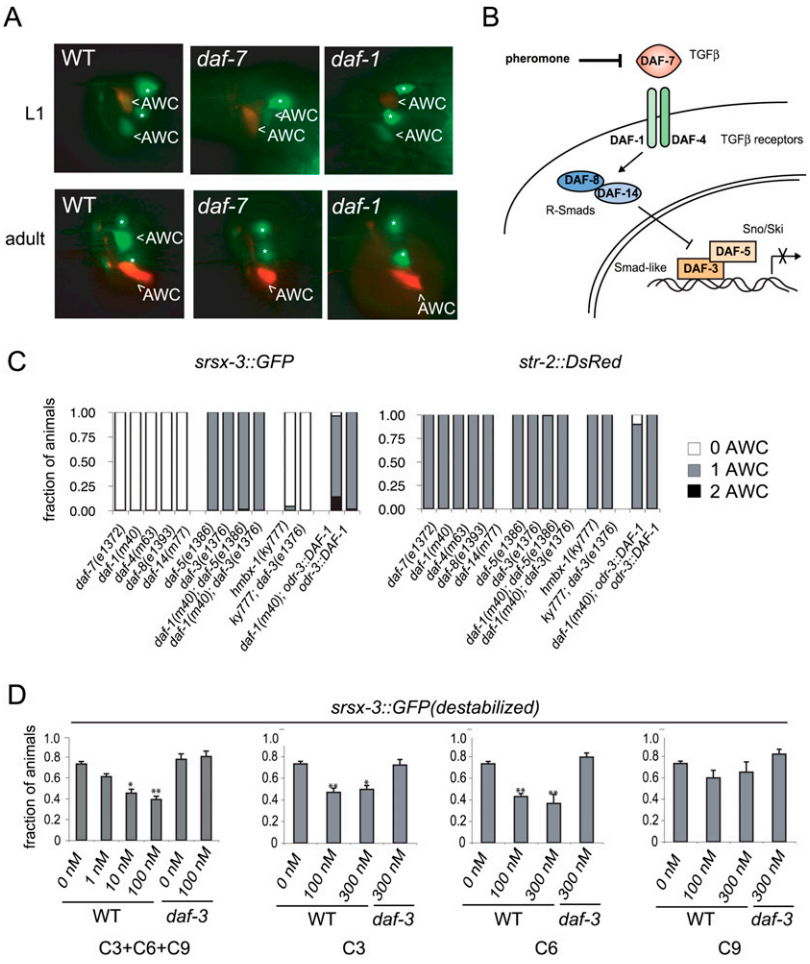


Figure 2. TGFβ and dauer pheromone signals regulate *srsx-3* expression. (*A*) Images of L1 larvae and adults in wild-type, *daf-7(e1372)*, and *daf-1(m40)* animals. AWC neurons are labeled, and asterisks mark AWB neurons. (*B*) Schematic illustration of the *daf-7* TGFβ pathway. (*C*) *srsx-3* and *str-2* expression phenotypes for *daf-7*/TGF-β pathway mutants and rescue of *daf-1(m40)* under the AWC-selective promoter *odr-3*. Also shown is the phenotype of the *hmbx-1(ky777); daf-3(e1376)* double mutant. $n > 25$ for all genotypes. (*D*) Regulation of *srsx-3* by a mixture of C3, C6, and C9 ascarosides (*left*) and by each ascaroside alone (two *middle* and *right*). Graphs show the fraction of adults expressing a destabilized GFP in AWC under the control of the *srsx-3* promoter. (*) $P < 0.01$; (**) $P < 0.001$ compared with no pheromone control (one-way ANOVA with Dunnett's post-test).

A missense allele of the transcription factor hmbx-1 *suppresses* srsx-3 *expression*

In the mutant *ky777*, *srsx-3* expression in AWC^{OFF} was lost after L1, but *str-2* expression in AWC^{ON} was retained (Figs. 1, 3A). Although this phenotype resembled that of *daf-7*/TGFβ mutants, *ky777* mutants were normal for dauer formation, and the *srsx-3* expression defect in these mutants was not suppressed by *daf-3* (Fig. 2C). Therefore, *ky777* appeared to act separately from the TGFβ pathway. *ky777* was mapped to an interval of ~440 kb on the left arm of chromosome I, but no rescue was observed after injection of fosmids or PCR products covering the interval. Therefore, a Solexa-Illumina whole-genome sequencing approach was used to identify mutations in the interval. Unique alignments of *ky777* sequence reads to the *C. elegans* genome accounted for 84% of sequence in the interval, with an average coverage depth of 7.9×. The gigaBayes program identified 19 high-probability point mutations and three single-base indels between *ky777* and the reference genome. Two missense mutations and one silent mutation were present in coding exons, one mutation was in a 3′ untranslated region (UTR), and one mutation was in a 5′UTR. PCR and conventional sequencing determined that one of the coding mutations and both UTR mutations were present both in the mutant strain and in the original, unmutagenized strain, but that the remaining coding mutation was present in the *ky777* mutant but not in the original strain. This mutation represents a C → T transition in the predicted gene *F54A5.1* that results in a missense H → Y mutation at the C-terminal end of the protein (Fig. 3B, red box), and was considered the most likely candidate for *ky777*. Transgenic introduction of a wild-type genomic copy of *F54A5.1* partially restored *srsx-3* expression in *ky777* (Fig. 4C; see below), supporting further analysis of this gene.

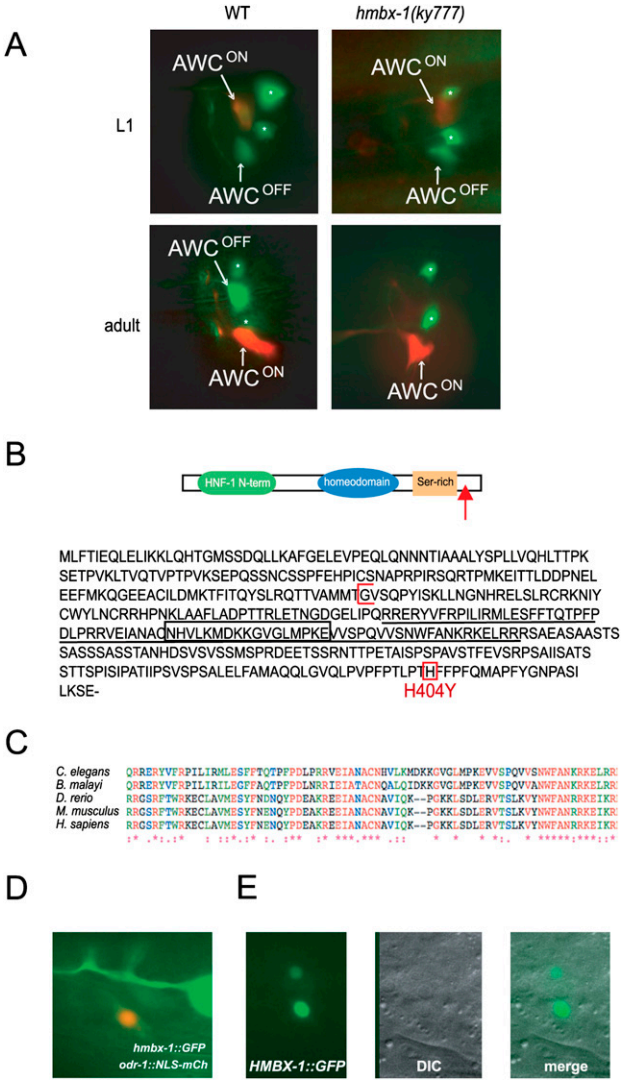


Figure 3. *ky777* is an allele of the homeodomain transcription factor *hmbx-1*. (*A*) Fluorescence images of early larval and adult wild-type and *ky777* mutant animals showing late loss of *srsx-3::GFP* from AWC^{OFF}. AWC neurons are labeled, and asterisks mark AWB neurons. (*B*, *top*) Domain structure and amino acid sequence of HMBX-1. Red arrow indicates the location of the *ky777* mutation in the domain structure diagram. (*Bottom*) The *ky777* mutation is boxed in red, the beginning of the *tm1274* deletion is indicated with a red bracket, the homeodomain is underlined, and a 17-amino-acid insertion in the homeodomain is boxed in black. (*C*) Conservation of the HMBX-1 homeodomain. (*D*) Expression of *hmbx-1(promoter)::GFP* in AWC. AWC is marked by a nuclear-localized mCherry reporter under the control of the *odr-1* promoter. Also visible are parts of the anterior process and cell body of the FLP neuron. (*E*) Nuclear localization of HMBX-1::GFP fusion protein in AWC. GFP is restricted to nuclei, visualized by DIC microscopy.

F54A5.1 encodes a predicted conserved homeodomain transcription factor that contains an HNF-1 N-terminal-like domain and a serine-rich region at its C terminus; it is closely related to the mammalian gene HMBOX1 (Fig. 3C). The homeodomain of *F54A5.1* and its homologs includes an unusual 17-amino-acid insertion not present in other homeodomains (Fig. 3B, black box); the entire protein is highly conserved in multiple species, including zebrafish, mice, and humans (Fig. 3C). Because of this high degree of conservation with HMBOX1 genes, we named *F54A5.1 hmbx-1*.

A transgene in which GFP was expressed under the control of 7 kb of the *hmbx-1* upstream region drove expression in both AWCs; in the chemosensory neurons ASI, AFD, ASH, and URX; in the mechanosensory neurons ALM, PLM, PVD, and FLP; in a few additional head and tail neurons; and in the seam cells of the hypodermis (Fig. 3D; data not shown). The expression of *hmbx-1* in the AWCs supports the hypothesis that it regulates *srsx-3* expression in these neurons. An HMBX-1 cDNA tagged with GFP localized to the nucleus of AWC, consistent with its predicted function as a transcription factor (Fig. 3E). A GFP-tagged HMBX-1(H404Y) protein corresponding to the *ky777* missense mutant also localized to the nucleus.

ky777 *is an altered-function allele of* hmbx-1

A deletion allele of *hmbx-1*, *tm1274*, was kindly provided by the National BioResource Project in Japan. The

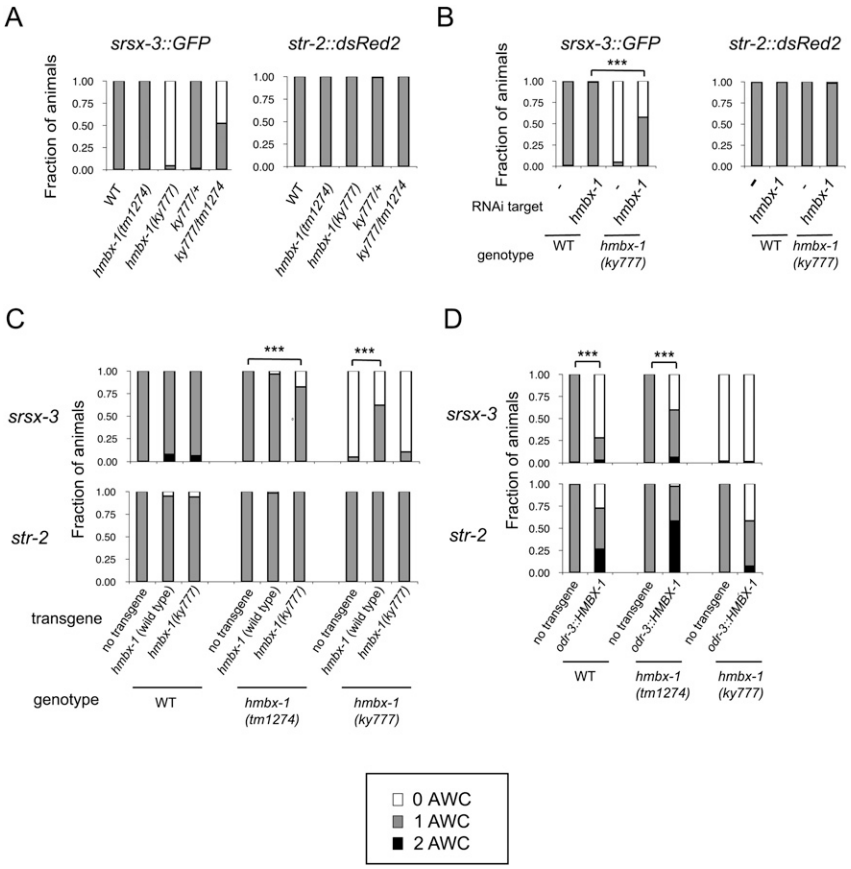


Figure 4. *ky777* is an altered-function allele of *hmbx-1*. (*A*) *srsx-3* and *str-2* expression phenotypes of wild-type, *hmbx-1(tm1274)*, and *hmbx-1(ky777)* homozygotes, and *ky777*/+ and *ky777/tm1274* heterozygotes. $n > 50$ for all genotypes. (*B*) RNAi against *hmbx-1* in wild-type and *ky777* backgrounds. All strains contained the RNAi-sensitizing *eri-1(mg366)* and *lin-15B(n744)* mutations. (***) Different at $P < 0.001$ (Fisher's exact test). $n > 40$ for all conditions. (*C*) Moderate overexpression of *hmbx-1* under endogenous regulatory elements. Expression of *str-2* and *srsx-3* in wild-type, *ky777*, or *tm1274* mutant animals expressing a genomic fragment covering wild-type or *ky777* mutant genomic coding sequence with 7 kb of upstream sequence. (***) Different from transgene-negative control at $P < 0.001$ (χ^2 test). $n > 30$ for all conditions. (*D*) Overexpression of *hmbx-1* from a strong AWC promoter. Expression of *str-2* and *srsx-3* in wild-type, *hmbx-1(tm1274)*, and *hmbx-1(ky777)* animals overexpressing a wild-type HMBX-1 cDNA under the *odr-3* promoter. (***) Different from transgene-negative control at $P < 0.001$ (χ^2 test). $n > 60$ for all genotypes.

deletion eliminates the HMBX-1 homeodomain, and results in a frameshift and early stop codon, and is therefore likely to be a null allele. The *hmbx-1(tm1274)* mutant was healthy and fertile, and had mild defects in chemotaxis to odors sensed by AWC neurons (Supplemental Fig. S2). Surprisingly, *srsx-3* expression was normal in the *hmbx-1(tm1274)* mutants (Fig. 4A). The different phenotypes of the *hmbx-1(ky777)* and *hmbx-1(tm1274)* mutants indicate that *ky777* is not a null allele of *hmbx-1*.

A set of genetic experiments suggested that the *ky777* allele results in altered, dosage-sensitive activity of the *hmbx-1* gene. *hmbx-1(ky777)*/+ animals were largely normal, indicating that *ky777* is recessive to the wild-type allele (Fig. 4A). *hmbx-1(ky777)/hmbx-1(tm1274null)* animals had an intermediate phenotype compared with either starting strain, a result suggesting that *tm1274* eliminates the wild-type gene activity that suppresses *ky777*, and supporting the hypothesis that the two mutations affect the same gene (Fig. 4A). Reducing *hmbx-1* expression using RNAi in wild-type animals had little effect on *srsx-3* expression, but RNAi against *hmbx-1* in *hmbx-1(ky777)* mutants restored *srsx-3* expression to many animals (Fig. 4B). These results suggest that RNAi is reducing an altered *hmbx-1* activity to generate an *hmbx-1*-null (wild-type-like) phenotype.

A moderate increase in *hmbx-1* activity was attained by injecting wild-type *hmbx-1* and *hmbx-1(ky777)* genomic DNAs into wild-type, *hmbx-1(ky777)*, and *hmbx-1(tm1274)* animals. In a wild-type background, neither the wild-type nor the mutant *hmbx-1* gene had a significant effect on *srsx-3* expression, in agreement with the observation that *hmbx-1(ky777)* is a recessive allele (Fig. 4C). In a null background, expression of the *hmbx-1(ky777)* mutant DNA, but not wild-type *hmbx-1*, partly repressed *srsx-3* expression, confirming that *hmbx-1(ky777)* represses *srsx-3* under conditions in which the wild-type *hmbx-1* gene does not. Finally, in a *hmbx-1(ky777)* mutant background, wild-type *hmbx-1* partly restored *srsx-3* expression, indicating that it antagonized *hmbx-1(ky777)* (Fig. 4C).

The results described above indicate that wild-type *hmbx-1* antagonizes *hmbx-1(ky777)*, and are consistent with two additional possibilities: (1) On its own, the *hmbx-1(ky777)* allele has a high or unregulated level of *hmbx-1* activity, or (2) the *hmbx-1(ky777)* allele has an abnormal activity unrelated to normal *hmbx-1* function. To distinguish between these alternatives, wild-type *hmbx-1* cDNA was expressed at high levels in the AWC neurons using the *odr-3* promoter, and the effects were examined in wild-type, *hmbx-1(tm1274)*, and *hmbx-1(ky777)* animals. High-copy *odr-3::hmbx-1* transgenes repressed *srsx-3* in wild-type and *tm1274* backgrounds, like the recessive *ky777* mutant (Fig. 4D). These results suggest that *hmbx-1(ky777)* mutants resemble animals with increased *hmbx-1* repressor activity in AWC. However, a few complications suggest that the effects of

hmbx-1 overexpression may not be entirely straightforward. First, *str-2* expression was sometimes misregulated in *odr-3::hmbx-1* animals, but not in *hmbx-1(ky777)* animals (Fig. 4D). Second, expression of *hmbx-1* under the *osm-3* promoter, which drives expression in 26 chemosensory neurons but not in AWC (Tabish et al. 1995), resulted in ectopic expression of *srsx-3* in AWC^{ON} in some animals (Supplemental Fig. S3). *hmbx-1* may therefore have cell-nonautonomous as well as cell-autonomous effects on *srsx*-3.

Bearing these potential complications in mind, the genetic results suggest that *hmbx-1(ky777)* has unregulated or increased *hmbx-1* activity that inappropriately represses *srsx-3* activity in AWC^{OFF}. A GFP-tagged HMBX-1(ky777) protein was expressed at similar levels to a tagged wild-type protein, suggesting that there were no major effects on protein stability. It is possible that the mutation affects an autoregulatory activity of *hmbx-1*, as expression of an *hmbx-1::GFP* transcriptional reporter was reduced in the AWC neurons of *hmbx-1(ky777)* mutants (Supplemental Fig. S3).

Single-copy srsx-3::GFP *transgenes are regulated by* hmbx-1 *and its predicted binding site*

During the gene dosage studies, we were struck by the variation in the mutant phenotype, depending on whether the copy number of *hmbx-1* was low (endogenous *hmbx-1*), intermediate (genomic *hmbx-1* DNA injection), or high (*odr-3::hmbx-1 cDNA* injection). In all of these experiments, *str-2::dsRed2* and *srsx-3::GFP* reporter genes were present at high-copy number in the *kyIs408* transgene, which could distort their interactions with a dosage-sensitive *hmbx-1* transcription factor. Therefore, to provide a more natural context for examining gene regulation effects, we generated single-copy *srsx-3::GFP* transgene reporters at a defined site on chromosome II using the Mos Single-Copy Insertion (MosSCI) technique (Frokjaer-Jensen et al. 2008).

In wild-type animals, the single-copy *srsx-3::GFP* reporter recapitulated the expression pattern of a high-copy *srsx-3::GFP* array, albeit with a weaker GFP signal (Fig. 5A). Young larvae expressed GFP in both AWC neurons, but expression was restricted to a single AWC^{OFF} neuron in adults (Fig. 5A; data not shown). Expression of the single-copy *srsx-3::GFP* transgene was lost in *hmbx-1(ky777)* mutants, and genetic interactions with *hmbx-1(ky777)* mutations were similar to those with a high-copy *srsx-3::GFP* array (Fig. 5A,B; Supplemental Fig. S4). Remarkably, the single-copy *srsx-3::GFP* reporter uncovered an opposite phenotype for the *hmbx-1(tm1274)*-null allele: The *srsx-3::GFP* reporter was misexpressed in both AWCs in a fraction of *hmbx-1(tm1274)* adults (Fig. 5A,B). Thus, in an *hmbx-1*-null mutant, *srsx-3* is derepressed in AWC^{ON}, whereas, in the *hmbx-1(ky777)* mutant, *srsx-3* is inappropriately repressed in AWC^{OFF}. These straightforward results with single-copy transgenes support and extend the conclusions from high-copy arrays. They suggest that *hmbx-1* represses *srsx-3* expression in AWC^{ON} neurons in the adult stage, that its effect is partly redundant with other repressors, as it is only partially penetrant, and that *hmbx-1(ky777)* is a recessive gain-of-function allele of *hmbx-1* that inappropriately represses *srsx-3* in AWC^{OFF} neurons.

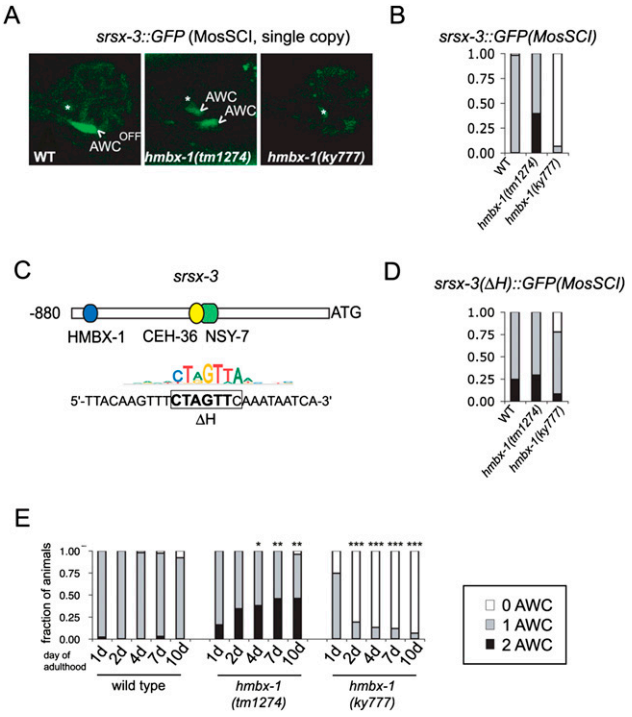


Figure 5. Regulation of single-copy *srsx-3::GFP* lines by *hmbx-1*. (*A*) Confocal images of the single-copy *srsx-3::GFP* reporter in wild-type, *hmbx-1(tm1274)*, and *hmbx-1(ky777)* animals. (Arrowheads) AWC; (asterisks) AWB. (*B*) Expression of singly integrated *srsx-3::GFP* in wild-type and mutant adults. $n > 40$ for all conditions. (*C*, *top*) A diagram of the *srsx-3* promoter with the positions of predicted transcription factor-binding sites. (*Bottom*) Binding site for HMBOX1, the mouse homolog of HMBX-1, shown *above* the sequence in the *srsx-3* promoter that was deleted in the *srsx-3(ΔH)::GFP* reporter. (*D*) Phenotypes of wild-type, *hmbx-1(tm1274)*, and *hmbx-1(ky777)* animals expressing the singly integrated *srsx-3(ΔH)::GFP* reporter. $n > 40$ for all genotypes. (*E*) Expression of singly integrated *srsx-3::GFP* in young (1-d-old; 12 h after the L4 stage) and older (2-, 4-, 7-, or 10-d-old) adults in wild-type and mutant backgrounds. For *hmbx-1(tm1274)*, one asterisk (*) indicates difference from young (1-d-old) adults at $P < 0.05$ and two asteriaks (**) indicate difference from young adults at $P < 0.01$ (Fisher's exact test). For *hmbx-1(ky777)*, three asterisks (***) indicate difference from young adults at $P < 0.001$.

The binding site of the mouse homolog of HMBX-1, HMBOX1, has been identified using an in vitro binding assay (Berger et al. 2008). A similar site is present in the *srsx-3* promoter, suggesting a potential site for regulation by HMBX-1 (Fig. 5C). The significance of this site was examined by deleting it from the *srsx-3* promoter [*srsx-3(ΔH)*] and introducing a single-copy insertion of the *srsx-3(ΔH)::GFP* sequence into the same MosSCI site used for the wild-type *srsx-3::GFP* reporter. Unlike the wild-type reporter, the mutated reporter was expressed in both AWCs in a fraction of wild-type animals, suggesting that

the sequence normally represses *srsx-3* in AWC^{ON} (Fig. 5D). The expression of single-copy *srsx-3* reporters with and without the predicted HMBX-1-binding site was then compared in wild-type, *hmbx-1(tm1274)*, and *hmbx-1(ky777)* backgrounds. The single-copy *srsx-3* reporter lacking the binding site behaved identically in wild-type and *hmbx-1(tm1274)* mutants (Fig. 5D), as predicted if HMBX-1 regulates *srsx-3* expression by binding to this site. Deletion of the predicted binding site partly suppressed the effects of the *hmbx-1(ky777)* mutation, suggesting that the altered-function protein also interacts with this site (Fig. 5D). Thus, the use of single-copy *srsx-3::GFP* reporters identified a loss-of-function phenotype for *hmbx-1*, and a likely site for HMBX-1 binding and regulation in vivo.

In agreement with the idea that *hmbx-1* is involved in long-term maintenance of AWC gene expression, the effects of both *hmbx-1(tm1274)* and *hmbx-1(ky777)* mutants were greater in older adults than in young adults (Fig. 5E). For example, only 16% of young *hmbx-1(tm1274)* adults (12 h after the L4 stage) expressed *srsx-3::GFP* in both AWC neurons, but this fraction reached ~45% by 1 wk of adulthood.

Multiple repressors interact to regulate asymmetric AWC-specific genes

Because *str-2* is the only known AWC^{ON}-specific gene and *srsx-3* is the only known AWC^{OFF}-specific gene, it was not clear whether transcription factors such as *hmbx-1*, *nsy-7*, and *daf-3* regulate individual receptor genes or the entire asymmetric AWC identity. Therefore, we sought and characterized additional markers that distinguished the AWC^{ON} and AWC^{OFF} neurons. The *C. elegans* gene expression project at the University of British Columbia has reported expression of several dozen predicted chemoreceptor genes in head sensory neurons (Dupuy et al. 2007); by examining 20 of these strains, we found that the *srt-26* and *srt-28* reporter genes were expressed strongly in a single AWC neuron, and the *srt-29* gene was expressed weakly in a single AWC. These three genes were coexpressed with *str-2* but not *srsx-3*, indicating that their expression was specific to AWC^{ON}.

AWC^{OFF}-specific reporters were sought using the binding site for the transcription factor NSY-7, which was defined previously by direct DNA-binding experiments (Lesch et al. 2009). Fourteen predicted chemoreceptor genes in the *C. elegans* genome contain the CCTTAAC NSY-7-binding sequence within 300 base pairs (bp) of the coding start site. Fluorescent reporters for these 14 genes were generated by fusing 2 kb upstream of the start site to GFP; two of these 14 genes, *srt-45* and *srt-47*, were expressed strongly in a single AWC neuron and weakly in an additional pair of head neurons. In both cases, expression was present in AWC^{OFF} but not AWC^{ON}, based on coexpression with the *srsx-3::mCherry* reporter (Fig. 6A). It is interesting that five genes of the *srt* family are expressed in AWC neurons, but four other tested *srt* genes were expressed in different neurons, so it is not a universal pattern (data not shown).

The newly identified AWC^{OFF} markers *srt-45* and *srt-47* and the AWC^{ON} markers *srt-26* and *srt-28* were examined in mutants that affect *srsx-3* and *str-2* expression. In *nsy-7*

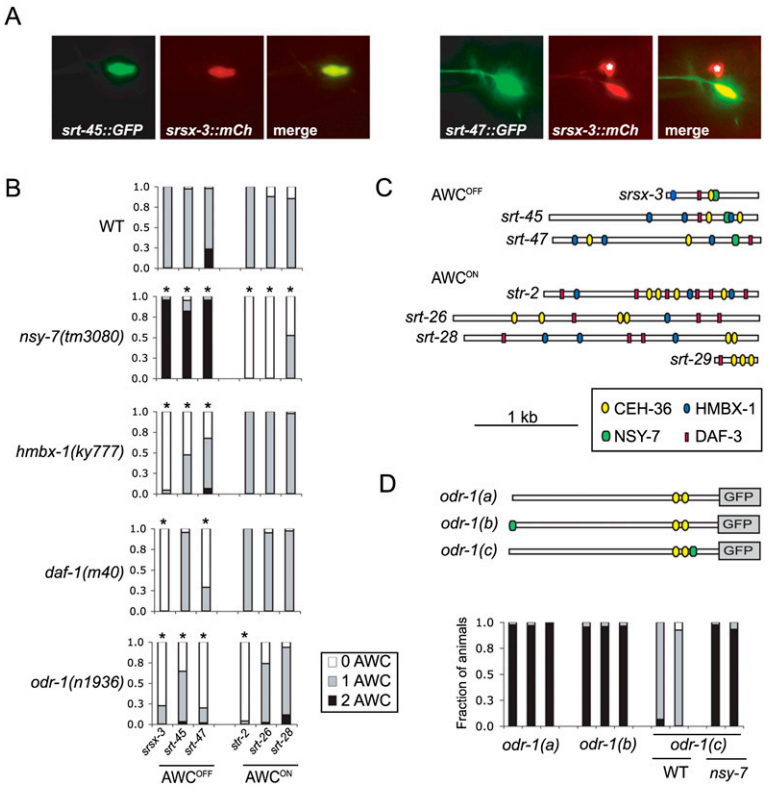


Figure 6. The NSY-7-binding site predicts AWC^{OFF}-specific expression, and is sufficient for repression in AWC^{ON}. (*A*) *srt-45* and *srt-47* are coexpressed with the AWC^{OFF} reporter *srsx-3::mCherry*, which also labels AWB (asterisks). (*B*) Expression of the AWC^{OFF} markers (*srsx-3*, *srt-45*, and *srt-47*) and the AWC^{ON} markers (*str-2*, *srt-26*, and *srt-28*) in wild-type and mutant backgrounds. $n > 30$ for all conditions; (*) difference from wild-type at $P < 0.01$ (Fisher's exact test). (*C*) Schematic of AWC^{ON} and AWC^{OFF} promoters showing predicted binding sites for the transcription factors CEH-36, HMBX-1, NSY-7, and DAF-3. (*D*, *top*) Schematic of *odr-1::GFP* constructs inserted as single copies. [*odr-1(a)*] Wild-type 2.4-kb *odr-1* promoter fragment; [*odr-1(b)*] NSY-7-binding site added to the 5′ end of the *odr-1* promoter; [*odr-1(c)*] NSY-7-binding site added 200 bp upstream of the ATG in the *odr-1* promoter. For *odr-1(c)*, the NSY-7-binding site is inserted close to, but does not disrupt, two CEH-36-binding sites involved in driving *odr-1* expression (Kim et al. 2010). (*Bottom*) Expression of integrated single-copy *odr-1* plasmids in wild-type and *nsy-7(tm3080)* backgrounds. Multiple bars within the same condition represent independently integrated lines. $n > 25$ for all genotypes.

mutants, all three AWC^{OFF} markers were expressed in both AWCs, and all three AWC^{ON} markers were reduced or absent (Fig. 6B). As predicted by earlier studies, this result suggests that, in *nsy-7* mutants, AWC^{ON} is transformed into AWC^{OFF}.

hmbx-1(ky777) regulated AWC^{OFF} chemoreceptor expression in a cell-wide manner, reducing expression of all three AWC^{OFF}-specific markers, but sparing the three AWC^{ON} markers (Fig. 6B). The symmetric AWC signaling genes *odr-3* and *odr-1* were expressed normally in AWC^{OFF} neurons in *hmbx-1(ky777)* mutants, and AWC^{OFF} morphology appeared normal, indicating that the AWC^{OFF} neurons lost their asymmetric identity, but did not die or degenerate (Supplemental Fig. S5; data not shown).

Both *daf-1* TGFβ mutants and *odr-1* cGMP olfactory transduction mutants showed a distinct, gene-specific pattern of regulation of AWC^{OFF} and AWC^{ON} chemoreceptors. In *daf-1* mutants, expression of the AWC^{OFF} markers *srsx-3* and *srt-47* was reduced, but the AWC^{OFF} marker *srt-45* and all AWC^{ON} markers were expressed at wild-type levels (Fig. 6B). *odr-1* mutants were defective for expression of all three AWC^{OFF} markers and the AWC^{ON} marker *str-2*, but expression of the AWC^{ON} markers *srt-26* and *srt-28* was largely intact (Fig. 6B). Thus, the TGFβ and cGMP-dependent pathways regulate subsets of chemoreceptor genes in both AWC^{OFF} and AWC^{ON} neurons.

The relationships between the different transcriptional pathways were probed by examining double mutants, using the multicopy *str-2* and *srsx-3* reporters and the *srsx-3* single-copy reporter. In general, double mutants recapitulated single mutant phenotypes in simple patterns, suggesting that *hmbx-1*, *nsy-7*, and *daf-3* act independently of each other (Supplemental Fig. S6). *odr-1* defects were partly suppressed by *hmbx-1(tm1274)*, suggesting that these two pathways are additive and independent; *hmbx-1(ky777)* and *nsy-7* effects were also additive.

The seven asymmetric AWC promoters had many potential binding sites for AWC-regulating proteins (Fig. 6C). All seven promoter fragments had predicted binding sites for CEH-36, the AWC terminal selector gene, and potential binding sites for DAF-3; six of seven promoters had matches to the consensus binding site for HMBX-1. NSY-7 sites were only present in the three AWC^{OFF} promoters. In previous studies, analysis of the *srsx-3* promoter suggested that NSY-7 binding was necessary for asymmetric expression (Lesch et al. 2009). The identification of additional transcriptional repressors raised the question of whether NSY-7 binding was also sufficient for asymmetric expression. To address this question, a single copy of the CCTTAAC sequence was inserted into an *odr-1::GFP* reporter, which is ordinarily expressed in both AWC neurons (and also in AWB neurons). Single-copy *odr-1::GFP* reporters with and without CCTTAAC sequences were inserted at a defined site on chromosome II using the MosSCI method. The wild-type single-copy *odr-1::GFP* fusion was expressed in both AWC neurons, but an *odr-1* plasmid with a CCTTAAC site 200 bp upstream of the ATG was expressed asymmetrically in a single AWC neuron (Fig. 6D). Coexpression of this transgene with *srsx-3* but not *str-2* reporters indicated that the CCTTAAC site repressed expression in AWC^{ON}. When crossed into a *nsy-7* mutant, the modified transgene was again expressed in both AWC neurons (Fig. 6D). A more distal insertion of the NSY-7 site was expressed bilaterally. These results support the hypothesis that NSY-7 is a transcriptional repressor, and demonstrate that a single promoter-proximal NSY-7-binding site is sufficient to repress gene expression in AWC^{ON}.

Discussion

Multiple environmental and cell-intrinsic influences converge at the transcriptional level to regulate chemoreceptors in AWC neurons. Together with the symmetric AWC terminal selector gene *ceh-36* (Kim et al. 2010), at least four different systems for transcriptional regulation contribute to adult AWC^{OFF}-specific expression of the chemoreceptor gene *srsx-3*. First, the previously identified transcriptional repressor NSY-7, which is expressed preferentially in AWC^{ON}, can repress *srsx-3* expression by direct binding to a consensus site. Second, cGMP signaling promotes *srsx-3* expression via a cGMP-dependent protein kinase, a cell-autonomous cGMP-gated channel, and unknown transcriptional regulators (Lesch et al. 2009). Third, density-dependent dauer pheromones repress expression of the secreted TGFβ homolog DAF-7, which otherwise acts continuously to maintain *srsx-3* expression. The TGFβ pathway regulates multiple chemoreceptor genes; in agreement with this observation, binding sites for SMAD transcription factors such as DAF-3 are among the most common sequences found upstream of chemoreceptor start sites (McCarroll et al. 2005). Finally, the conserved transcription factor *hmbx-1* represses *srsx-3* expression preferentially in adult AWC^{ON} neurons, although it is expressed in both AWC neurons and can act in AWC^{OFF} when bearing the *ky777* point mutation or when overexpressed.

Genetic analysis of *hmbx-1* suggests that it is involved in long-term maintenance of a specific cell identity, not the developmental establishment of that identity or the regulation of specific genes. The original *hmbx-1(ky777)* mutation is a recessive gain-of-function allele with effects that resemble those of *hmbx-1* overexpression. Because of its nature and its dosage sensitivity, it was identified only by whole-genome sequencing. Although recessive gain-of-function alleles are relatively rare, they have been described in genes encoding potassium channels and their regulators, tyrosine kinase receptors, and the ETS domain transcription factor LIN-1 (Klingler et al. 1988; Jacobs et al. 1998; Perez de la Cruz et al. 2003). LIN-1-recessive gain-of-function mutations disrupt a negative regulatory domain; it is possible that *ky777* does the same to HMBX-1.

The analysis of a single-copy insertion of the *srsx-3::GFP* reporter uncovered a phenotype for the *hmbx-1*-null mutant that was not evident with a multicopy reporter, demonstrating that *hmbx-1* normally represses

AWC^{OFF} chemoreceptor genes in AWC^{ON} neurons. Single-copy reporter genes also identified a predicted HMBX-1-binding site required for repression of *srsx-3* by *hmbx-1*. Multicopy integrated and extrachromosomal arrays are commonly and successfully used to analyze gene expression in *C. elegans*. However, high-copy transgenes are subject to repeat-induced silencing in the germline and, to a lesser extent, somatic tissues (Hsieh and Fire 2000), and this leads to altered genetic requirements for their expression. For example, specific genes including *tam-1*, a gene isolated in our screen, are required for efficient expression of high-copy transgenes, but not of the corresponding endogenous genes (Hsieh et al. 1999). Moreover, high-copy transgenes have a high propensity to form heterochromatin, which is not observed with low-copy transgenes of the same sequence (Meister et al. 2010). We suggest that the heterochromatic state of the high-copy transgene may bypass the normal requirement for *hmbx-1* and perhaps other repressors that maintain gene silencing. The ability to introduce single-copy transgenes into defined genomic locations by MosSCI represents a significant advance for controlling copy number and genomic context effects on gene expression (Frokjaer-Jensen et al. 2008).

hmbx-1 has effects that are temporally distinct from those of *nsy-7*, a transcription factor that acts in AWC^{ON} beginning in the L1 larval stage. The relative importance of wild-type *hmbx-1* is greater in older compared with younger adults, and, likewise, the elevated repressive activity of the *hmbx-1(ky777)* mutant increases in adults over time. It may be that NSY-7 loses activity in older animals, or that other changes in gene expression alter the relative importance of the two transcription factors as the animal ages.

The results described here point to a significant role for single transcription factor-binding sites in chemoreceptor expression. The presence of a single binding site for NSY-7 is sufficient to repress expression of an *odr-1* promoter ordinarily expressed in both AWC neurons. Likewise, deletion or mutation of the HMBX-1-binding site in the *srsx-3* promoter permitted misexpression of *srsx-3* in AWC^{ON}. The relative ease with which sensory receptor expression can be altered suggests genetic malleability and potentially evolutionary flexibility in the specificity of GPCR expression in chemosensory neurons (Jovelin 2009). Particularly in *C. elegans*, where ectopic expression of a gene in a single additional neuron can dramatically change behaviors, altered expression due to new mutations might have significant consequences for the animal's behavior and fitness (Troemel et al. 1997).

Despite the potent effects of single binding sites, many additional candidate binding sites were identified in each asymmetric AWC gene. Functional studies of asymmetrically expressed promoters in ASE neurons have uncovered multiple repressor and activator elements in each promoter (Etchberger et al. 2009). In ASE, *che-1*-binding sites with different affinities contribute to left–right asymmetry of gene expression as well as general ASE identity (Etchberger et al. 2009); it may be informative that the AWC^{ON}-specific promoters contain more predicted CEH-36-binding sites than the AWC^{OFF}-specific promoters.

hmbx-1 is expressed in *C. elegans* sensory neurons, and its human and mouse homologs also display high levels of nervous system expression (Chen et al. 2006). The human homolog exhibits repressor activity in vitro, and associates with telomeric sequences in several human cell lines (Chen et al. 2006; Dejardin and Kingston 2009). Telomeres are zones of strong transcriptional repression, suggesting a possible repressor function for HMBOX-1 at these locations (Gottschling et al. 1990). The sequence conservation of HMBOX-1 genes between worm and mammalian homologs, their conserved repressor activity, the apparently conserved binding site specificity, and their neuronal expression pattern all hint at possible conserved neuronal functions. With this in mind, it will be interesting to determine the identities of *hmbx-1* target genes, particularly those that are important in older adult animals, and to explore the association of human HMBOX1 with telomeric repeats during normal growth and aging.

Materials and methods

Genetics and strains

C. elegans strains were cultured using standard methods (Brenner 1974). All strains were grown at 20°C unless otherwise specified. Mutants were isolated by direct inspection of GFP and dsRed fluorescence following mutagenesis of the strain CX7894 *kyIs408 [srsx-3::GFP;str-2::dsRed2;elt-2::GFP] II* (Lesch et al. 2009). Mutagenesis was performed with ethyl methane sulfonate (EMS) using standard protocols (Brenner 1974).

Among the new mutations, *tax-4* and *tax-2* alleles differed in phenotype from *tax-4* and *tax-2* mutants characterized previously using the *kyIs140 [str-2::GFP] I* transgene (Troemel et al. 1999; Lans and Jansen 2006). Whereas expression of *str-2::GFP* from *kyIs140* was lost in most *tax-2* and *tax-4* mutants, the effect of *tax-2* and *tax-4* on *str-2::DsRed* from *kyIs408* was milder. Analysis of the canonical *tax-4(p678)*-null mutant indicated that the difference was due to the transgene and not to the nature of the mutant alleles (Fig. 1B). *kyIs140* fluorescence is much dimmer than that of *kyIs408*, so it may be more sensitive to small reductions in expression level.

A strain list appears as Supplemental Table S1.

Molecular biology

Standard molecular biology techniques were used. The *gfp-pest* reporter was made by overlap extension PCR between the GFP template from the pSM-GFP expression vector and the PEST domain template from a commercially available destabilized GFP vector (d1EGFP-N1, Clontech). The product was then cloned into MCS2 of pSM with the *srsx-3* promoter in MCS1 to create the *srsx-3::gfp-pest* vector.

A PCR fragment containing the *hmbx-1* genomic coding region and 7 kb of upstream sequence was amplified from N2 lysate using the primers 5′-CATATCACATCCTCCTGTTTTC AAG-3′ and 5′-CCAACAAAATTTCAGGGCAAGC-3′. HMBX-1 cDNAs were obtained by PCR from a *C. elegans* poly(A) cDNA library using primers at the beginning and end of the predicted ORF (5′-ATGCTATTCACAATTGAGCAACTG-3′ and 5′-TCA CTCAGACTTCAAAATAGAAG-3′), and were cloned into

MCS2 of a modified pSM expression vector (containing a NotI site in MCS2 instead of MCS1) using the restriction sites SalI and NotI.

The *hmbx-1(7kb)::GFP*, *srt-45::GFP*, and *srt-47::GFP* reporters were created by PCR fusion of the promoter, amplified from N2 lysate, with GFP template amplified from pPD95.75. Reactions were carried out as described in Hobert (2002) with the following primers: *hmbx-1(7kb)*: A, 5′-TGCCGATTGCCAAAA ACTTCC-3′, and A*, 5′-CATATCACATCCTCCTGTTTTCA AG-3′, and B, 5′-AGTCGACCTGCAGGCATGCAAGCTGTGT AAAATAAAATTTACTACGAGGTT-3′; *srt-45*: A, 5′-CTTTTA AGTATTGTCACTCCGGC-3′, and A*, 5′-TTGGACATGTAAT CATGCATG-3′, and B, 5′-AGTCGACCTGCAGGCATGCAAG CTTGCGGAATCTGAAGTTTTTCG-3′; *srt-47*: A, 5′-GCGTA AAACACAGCAGAAAA-3′, and A*, 5′-ATCTTTAAAGGTGC ATTTTATTTGG-3′, and B, 5′-AGTCGACCTGCAGGCATGC AAGCTTTGGTGATAAAGAGAGGTTATAG-3′.

To create the *srt-26::GFP*, *srt-28::GFP*, and *srt-29::GFP* reporters, the *srt-26*, *srt-28*, and *srt-29* promoter sequences were amplified by PCR and cloned into MCS1 of the pSM-GFP expression vector using the FseI and AscI restriction sites.

For MosSCI experiments, NSY-7-binding sites were added to the *odr-1::GFP* expression vector, and the *srsx-3::GFP* plasmid was modified to generate the *srsx-3(ΔH)::GFP* reporter by site-directed mutagenesis using the the QuikChange II Site-Directed Mutagenesis Kit (Stratagene). *odr-1::GFP* or *srsx-3::GFP* sequences were cut out of pSM using the FseI and SpeI restriction sites, and were cloned into a pCFJ151 MosSCI insertion vector (Frokjaer-Jensen et al. 2008) that had been modified to include a FseI site in the MCS.

Whole-genome sequencing

Whole-genome sequencing of the *ky777* mutant was performed by the Rockefeller Genomics Resource Center using Solexa-Illumina Genome Analyzer technology. SNP mapping was used to determine a genetic interval of ~440 kb for the mutation. Purified genomic DNA was prepared from 1000–2000 worms, and was used for library preparation. fastq sequences were aligned to the ce6 genome assembly at the University of California at Santa Cruz using the Mosaik alignment program (http://bioinformatics.bc.edu/marthlab/Mosaik#Current_Documentation), and single-base changes, insertions, or deletions were predicted using GigaBayes polymorphism detection software (http://bioinformatics.bc.edu/marthlab/GigaBayes). High-probability single-base changes that were predicted to fall within an exon, a 3′UTR, or a 5′UTR were checked by PCR and conventional (Sanger) sequencing.

Sequence alignment

The *hmbx-1* homeodomain sequence was aligned with HMBOX1 homeodomain sequences using ClustalW at Pôle BioInformatique Lyonnais (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_clustalw.html). UniProt ID numbers are as follows: for *C. elegans*, Q9TYT0; for *Brugia malayi*, A8QCW9; for *Danio rerio*, Q4V904; for *Mus musculus*, Q8BJA3; for *Homo sapiens*, A8K3R8.

RNAi

RNAi was performed by injection of dsRNA. All assays were performed in an *eri-1(mg366);lin-15B(n744)*-sensitized background (Sieburth et al. 2005). For injection, a dsDNA template corresponding to an exon of the target ORF was amplified from N2 lysate with a T7 sequence (TAATACGACTCACTATAGGG AGA) added at the 5′ ends. The following gene-specific sequences were used for these primers: *nsy-7* (exon 2), 5′-GTTGCG AAAGGATATTCAGATG-3′ and 5′-CTTAGCAAACAAGTTG GTGAGT-3′; *hmbx-1* (exon 3), 5′-ACAAGCTCCCGTAACAC AAC-3′ and 5′-TCACTCAGACTTCAAAATAGAAGCC-3′.

Transcription was performed using the T7 RiboMAX Express RNAi System (Promega) according to instructions, and the unpurified reaction mix was injected into the body cavity, gut, or gonad of adult hermaphrodites. F1 progeny from eggs laid at least 24 h after the injection were scored after 3–4 d for *str-2* and *srsx-3* expression. Control experiments established the normal RNAi response of *hmbx-1(ky777)* animals: RNAi against *nsy-7* in *ky777* mutants resulted in a 100% loss of *str-2* expression (44 out of 44 $RNAi^+$ vs. zero out of 90 $RNAi^-$ animals; $P < 0.0001$ [Fisher's exact test]), consistent with the phenotype of *hmbx-1(ky777);nsy-7(tm3080)* double mutants.

MosSCI integrations

Mos single-copy integrants were generated using the direct insertion protocol described in Frokjaer-Jensen et al. (2008). Thirty to 50 EG4322 *ttTi5605*; *unc-119(ed3)* worms were injected with *rab-3::mCherry*, *myo-2::mCherry*, *myo-3::mCherry*, pJL43.1 (a vector containing the Mos1 transposase under the control of the germline promoter *glh-2*), and a vector containing the specific promoter::GFP sequence to be inserted flanked by sequences homologous to the insertion site. Animals that were rescued for the *unc-119* phenotype (array-positive) were allowed to starve out twice, and then *unc-119* rescued animals that lacked the three mCherry coinjection markers (integrant-positive, array-negative) were cloned out from separate plates to found independent integrated lines. These lines were outcrossed twice to wild-type animals, and the presence of the intact insertion was verified by PCR and sequencing.

Microscopy

For all microscopy, live animals were immobilized on an agarose pad containing 5 mM NaN_3.

Fluorescence microscopy was carried out on a Zeiss Axioplan2 imaging system with a Hamamatsu Photonics C2400 CCD camera, or a Zeiss Axio Imager.Z1 with ApoTome with a Zeiss AxioCam MRm CCD camera. Most animals were scored under a 20× or 40× Plan-Neofluar objective, where "bright" and "faint" fluorescence were scored qualitatively, and photographs were taken under a 40× Plan-Neofluar or 63× Plan-Apochromat objective. Confocal microscopy (Fig. 5A) was done under a 40×/1.2 W C-Apochromat water immersion objective on a Zeiss LSM 510 confocal imaging system using the Zeiss LSM 510 version 3.2 confocal software.

Developmental timing

To evaluate marker expression in the L1 larval stage, larvae were staged by hatch-off. Late embryos were picked to an NGM plate seeded with the *Escherichia coli* strain OP50. After 30 min, just-hatched L1s were transferred to a fresh plate and grown for 14 h at 20°C (for L1s) or ~70 h at 20°C (for adults). To compare young and old adults, 25 L4 animals were picked per plate and grown for 12 h (1 d) at 20°C, 36 h (2 d) at 20°C, 84 h (4 d) at 20°C, 156 h (7 d) at 20°C, or 228 h (10 d) at 20°C before scoring. Animals were transferred to new plates every 24 h to prevent crowding and starvation.

Pheromone assays

Ascarosides C3, C6, and C9 (generously provided by R. Butcher and J. Clardy, Harvard Medical School) were added to liquid agar at the concentrations indicated. For negative controls, the same volume of solvent (ethanol) was added to the agar. Ten milliliters of agar was poured into 6-cm culture dishes and allowed to cool. Plates were then seeded with 100 μL of OP50 bacteria and dried in a hood for 1.5–2 h. These plates were either used immediately or stored overnight at 4°C. Twenty *srsx-3::gfp-pest* array-positive young adults (older than L4, with no eggs yet visible in the gonad) were picked to each plate and incubated for 4 h at 25°C. Animals were scored for presence or absence of GFP in AWC neurons under a 40× objective.

Acknowledgments

We thank Shohei Mitani for the *hmbx-1(tm1274)* strain, Rebecca Butcher and Jon Clardy for ascaroside pheromones, and Mike Chiorazzi and Diana McKeage for contributing to the mutant characterization. This work was supported by NIDCD grant DC004089. B.J.L. was supported by MSTP grant GM07739, and by an individual NRSA (F30MH084482). C.I.B is an Investigator of the Howard Hughes Medical Institute.

References

Berger MF, Badis G, Gehrke AR, Talukder S, Philippakis AA, Pena-Castillo L, Alleyne TM, Mnaimneh S, Botvinnik OB, Chan ET, et al. 2008. Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. *Cell* **133:** 1266–1276.

Brenner S. 1974. The genetics of *Caenorhabditis elegans*. *Genetics* **77:** 71–94.

Buck L, Axel R. 1991. A novel multigene family may encode odorant receptors: A molecular basis for odor recognition. *Cell* **65:** 175–187.

Busturia A, Morata G. 1988. Ectopic expression of homeotic genes caused by the elimination of the Polycomb gene in *Drosophila* imaginal epidermis. *Development* **104:** 713–720.

Butcher RA, Fujita M, Schroeder FC, Clardy J. 2007. Small-molecule pheromones that control dauer development in *Caenorhabditis elegans*. *Nat Chem Biol* **3:** 420–422.

Butcher RA, Ragains JR, Kim E, Clardy J. 2008. A potent dauer pheromone component in *Caenorhabditis elegans* that acts synergistically with other components. *Proc Natl Acad Sci* **105:** 14288–14292.

Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y. 2002. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298:** 1039–1043.

Chalasani SH, Chronis N, Tsunozaki M, Gray JM, Ramot D, Goodman MB, Bargmann CI. 2007. Dissecting a circuit for olfactory behaviour in *Caenorhabditis elegans*. *Nature* **450:** 63–70.

Chang S, Johnston RJ Jr, Hobert O. 2003. A transcriptional regulatory cascade that controls left/right asymmetry in chemosensory neurons of *C. elegans*. *Genes Dev* **17:** 2123–2137.

Chang S, Johnston RJ Jr, Frokjaer-Jensen C, Lockery S, Hobert O. 2004. MicroRNAs act sequentially and asymmetrically to control chemosensory laterality in the nematode. *Nature* **430:** 785–789.

Chen S, Saiyin H, Zeng X, Xi J, Liu X, Li X, Yu L. 2006. Isolation and functional analysis of human HMBOX1, a homeobox containing protein with transcriptional repressor activity. *Cytogenet Genome Res* **114:** 131–136.

Clyne PJ, Warr CG, Carlson JR. 2000. Candidate taste receptors in *Drosophila*. *Science* **287:** 1830–1834.

Coburn CM, Mori I, Ohshima Y, Bargmann CI. 1998. A cyclic nucleotide-gated channel inhibits sensory axon outgrowth in larval and adult *C. elegans*: A distinct pathway for maintenance of sensory axon structure. *Development* **125:** 249–258.

Colosimo ME, Tran S, Sengupta P. 2003. The divergent orphan nuclear receptor ODR-7 regulates olfactory neuron gene expression via multiple mechanisms in *Caenorhabditis elegans*. *Genetics* **165:** 1779–1791.

da Graca LS, Zimmerman KK, Mitchell MC, Kozhan-Gorodetska M, Sekiewicz K, Morales Y, Patterson GI. 2004. DAF-5 is a Ski oncoprotein homolog that functions in a neuronal TGF β pathway to regulate *C. elegans* dauer development. *Development* **131:** 435–446.

Dejardin J, Kingston RE. 2009. Purification of proteins associated with specific genomic loci. *Cell* **136:** 175–186.

Dupuy D, Bertin N, Hidalgo CA, Venkatesan K, Tu D, Lee D, Rosenberg J, Svrzikapa N, Blanc A, Carnec A, et al. 2007. Genome-scale analysis of in vivo spatiotemporal promoter activity in *Caenorhabditis elegans*. *Nat Biotechnol* **25:** 663–668.

Etchberger JF, Lorch A, Sleumer MC, Zapf R, Jones SJ, Marra MA, Holt RA, Moerman DG, Hobert O. 2007. The molecular signature and *cis*-regulatory architecture of a *C. elegans* gustatory neuron. *Genes Dev* **21:** 1653–1674.

Etchberger JF, Flowers EB, Poole RJ, Bashllari E, Hobert O. 2009. Cis-regulatory mechanisms of left/right asymmetric neuron-subtype specification in *C. elegans*. *Development* **136:** 147–160.

Frand AR, Russel S, Ruvkun G. 2005. Functional genomic analysis of *C. elegans* molting. *PLoS Biol* **3:** e312. doi: 10.1371/journal.pbio.0030312.

Frokjaer-Jensen C, Davis MW, Hopkins CE, Newman BJ, Thummel JM, Olesen SP, Grunnet M, Jorgensen EM. 2008. Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nat Genet* **40:** 1375–1383.

Gaudet J, Mango SE. 2002. Regulation of organogenesis by the *Caenorhabditis elegans* FoxA protein PHA-4. *Science* **295:** 821–825.

Georgi LL, Albert PS, Riddle DL. 1990. *daf-1*, a *C. elegans* gene controlling dauer larva development, encodes a novel receptor protein kinase. *Cell* **61:** 635–645.

Gottschling DE, Aparicio OM, Billington BL, Zakian VA. 1990. Position effect at *S. cerevisiae* telomeres: Reversible repression of Pol II transcription. *Cell* **63:** 751–762.

Hobert O. 2002. PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic *C. elegans*. *Biotechniques* **32:** 728–730.

Hobert O. 2008. Regulatory logic of neuronal diversity: Terminal selector genes and selector motifs. *Proc Natl Acad Sci* **105:** 20067–20071.

Hsieh J, Fire A. 2000. Recognition and silencing of repeated DNA. *Annu Rev Genet* **34:** 187–204.

Hsieh J, Liu J, Kostas SA, Chang C, Sternberg PW, Fire A. 1999. The RING finger/B-box factor TAM-1 and a retinoblastoma-like protein LIN-35 modulate context-dependent gene silencing in *Caenorhabditis elegans*. *Genes Dev* **13:** 2958–2970.

Jacobs D, Beitel GJ, Clark SG, Horvitz HR, Kornfeld K. 1998. Gain-of-function mutations in the *Caenorhabditis elegans lin-1* ETS gene identify a C-terminal regulatory domain phosphorylated by ERK MAP kinase. *Genetics* **149:** 1809–1822.

Jeong PY, Jung M, Yim YH, Kim H, Park M, Hong E, Lee W, Kim YH, Kim K, Paik YK. 2005. Chemical structure and

biological activity of the *Caenorhabditis elegans* dauer-inducing pheromone. *Nature* **433:** 541–545.
Johnston RJ, Hobert O. 2003. A microRNA controlling left/right neuronal asymmetry in *Caenorhabditis elegans*. *Nature* **426:** 845–849.
Johnston RJ Jr, Chang S, Etchberger JF, Ortiz CO, Hobert O. 2005. MicroRNAs acting in a double-negative feedback loop to control a neuronal cell fate decision. *Proc Natl Acad Sci* **102:** 12449–12454.
Jovelin R. 2009. Rapid sequence evolution of transcription factors controlling neuron differentiation in *Caenorhabditis*. *Mol Biol Evol* **26:** 2373–2386.
Kim K, Colosimo ME, Yeung H, Sengupta P. 2005. The UNC-3 Olf/EBF protein represses alternate neuronal programs to specify chemosensory neuron identity. *Dev Biol* **286:** 136–148.
Kim K, Sato K, Shibuya M, Zeiger DM, Butcher RA, Ragains JR, Clardy J, Touhara K, Sengupta P. 2009. Two chemoreceptors mediate developmental effects of dauer pheromone in *C. elegans*. *Science* **326:** 994–998.
Kim K, Kim R, Sengupta P. 2010. 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*. *Development* **137:** 963–974.
Klingler M, Erdelyi M, Szabad J, Nusslein-Volhard C. 1988. Function of *torso* in determining the terminal anlagen of the *Drosophila* embryo. *Nature* **335:** 275–277.
Lanjuin A, Sengupta P. 2002. Regulation of chemosensory receptor expression and sensory signaling by the KIN-29 Ser/Thr kinase. *Neuron* **33:** 369–381.
Lanjuin A, VanHoven MK, Bargmann CI, Thompson JK, Sengupta P. 2003. Otx/otd homeobox genes specify distinct sensory neuron identities in *C. elegans*. *Dev Cell* **5:** 621–633.
Lans H, Jansen G. 2006. Noncell- and cell-autonomous G-protein-signaling converges with Ca2+/mitogen-activated protein kinase signaling to regulate *str-2* receptor gene expression in *Caenorhabditis elegans*. *Genetics* **173:** 1287–1299.
Lesch BJ, Gehrke AR, Bulyk ML, Bargmann CI. 2009. Transcriptional regulation and stabilization of left-right neuronal identity in *C. elegans*. *Genes Dev* **23:** 345–358.
McCarroll SA, Li H, Bargmann CI. 2005. Identification of transcriptional regulatory elements in chemosensory receptor genes by probabilistic segmentation. *Curr Biol* **15:** 347–352.
Meister P, Towbin BD, Pike BL, Ponti A, Gasser SM. 2010. The spatial dynamics of tissue-specific promoters during *C. elegans* development. *Genes Dev* **24:** 766–782.
Mikkola I, Heavey B, Horcher M, Busslinger M. 2002. Reversion of B cell commitment upon loss of Pax5 expression. *Science* **297:** 110–113.
Muller J, Kassis JA. 2006. Polycomb response elements and targeting of Polycomb group proteins in *Drosophila*. *Curr Opin Genet Dev* **16:** 476–484.
Nokes EB, Van Der Linden AM, Winslow C, Mukhopadhyay S, Ma K, Sengupta P. 2009. Cis-regulatory mechanisms of gene expression in an olfactory neuron type in *Caenorhabditis elegans*. *Dev Dyn* **238:** 3080–3092.
Nolan KM, Sarafi-Reinach TR, Horne JG, Saffer AM, Sengupta P. 2002. The DAF-7 TGF-β signaling pathway regulates chemosensory receptor gene expression in *C. elegans*. *Genes Dev* **16:** 3061–3073.
Nutt SL, Heavey B, Rolink AG, Busslinger M. 1999. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* **401:** 556–562.
Patterson GI, Koweek A, Wong A, Liu Y, Ruvkun G. 1997. The DAF-3 Smad protein antagonizes TGF-β-related receptor signaling in the *Caenorhabditis elegans* dauer pathway. *Genes Dev* **11:** 2679–2690.
Peckol EL, Troemel ER, Bargmann CI. 2001. Sensory experience and sensory activity regulate chemosensory receptor gene expression in *Caenorhabditis elegans*. *Proc Natl Acad Sci* **98:** 11032–11038.
Perez de la Cruz I, Levin JZ, Cummins C, Anderson P, Horvitz HR. 2003. *sup-9*, *sup-10*, and *unc-93* may encode components of a two-pore K^+ channel that coordinates muscle contraction in *Caenorhabditis elegans*. *J Neurosci* **23:** 9133–9145.
Prasad BC, Ye B, Zackhary R, Schrader K, Seydoux G, Reed RR. 1998. *unc-3*, a gene required for axonal guidance in *Caenorhabditis elegans*, encodes a member of the O/E family of transcription factors. *Development* **125:** 1561–1568.
Ren P, Lim CS, Johnsen R, Albert PS, Pilgrim D, Riddle DL. 1996. Control of *C. elegans* larval development by neuronal expression of a TGF-β homolog. *Science* **274:** 1389–1391.
Rens-Domiano S, Hamm HE. 1995. Structural and functional relationships of heterotrimeric G-proteins. *FASEB J* **9:** 1059–1066.
Ringrose L, Paro R. 2004. Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu Rev Genet* **38:** 413–443.
Sagasti A, Hobert O, Troemel ER, Ruvkun G, Bargmann CI. 1999. Alternative olfactory neuron fates are specified by the LIM homeobox gene *lim-4*. *Genes Dev* **13:** 1794–1806.
Sarafi-Reinach TR, Sengupta P. 2000. The forkhead domain gene *unc-130* generates chemosensory neuron diversity in *C. elegans*. *Genes Dev* **14:** 2472–2485.
Sarafi-Reinach TR, Melkman T, Hobert O, Sengupta P. 2001. The *lin-11* LIM homeobox gene specifies olfactory and chemosensory neuron fates in *C. elegans*. *Development* **128:** 3269–3281.
Schackwitz WS, Inoue T, Thomas JH. 1996. Chemosensory neurons function in parallel to mediate a pheromone response in *C. elegans*. *Neuron* **17:** 719–728.
Sengupta P, Colbert HA, Bargmann CI. 1994. The *C. elegans* gene *odr-7* encodes an olfactory-specific member of the nuclear receptor superfamily. *Cell* **79:** 971–980.
Shibata Y, Takeshita H, Sasakawa N, Sawa H. 2010. Double bromodomain protein BET-1 and MYST HATs establish and maintain stable cell fates in *C. elegans*. *Development* **137:** 1045–1053.
Shirasaki R, Pfaff SL. 2002. Transcriptional codes and the control of neuronal identity. *Annu Rev Neurosci* **25:** 251–281.
Sieburth D, Ch'ng Q, Dybbs M, Tavazoie M, Kennedy S, Wang D, Dupuy D, Rual JF, Hill DE, Vidal M, et al. 2005. Systematic analysis of genes required for synapse structure and function. *Nature* **436:** 510–517.
Tabish M, Siddiqui ZK, Nishikawa K, Siddiqui SS. 1995. Exclusive expression of *C. elegans osm-3* kinesin gene in chemosensory neurons open to the external environment. *J Mol Biol* **247:** 377–389.
Troemel ER, Chou JH, Dwyer ND, Colbert HA, Bargmann CI. 1995. Divergent seven transmembrane receptors are candidate chemosensory receptors in *C. elegans*. *Cell* **83:** 207–218.
Troemel ER, Kimmel BE, Bargmann CI. 1997. Reprogramming chemotaxis responses: Sensory neurons define olfactory preferences in *C. elegans*. *Cell* **91:** 161–169.
Troemel ER, Sagasti A, Bargmann CI. 1999. Lateral signaling mediated by axon contact and calcium entry regulates asymmetric odorant receptor expression in *C. elegans*. *Cell* **99:** 387–398.

Tucker M, Sieber M, Morphew M, Han M. 2005. The *Caenorhabditis elegans aristaless* orthologue, *alr-1*, is required for maintaining the functional and structural integrity of the amphid sensory organs. *Mol Biol Cell* **16:** 4695–4704.

van der Linden AM, Nolan KM, Sengupta P. 2007. KIN-29 SIK regulates chemoreceptor gene expression via an MEF2 transcription factor and a class II HDAC. *EMBO J* **26:** 358–370.

van der Linden A, Wiener S, You YJ, Kim K, Avery L, Sengupta P. 2008. The EGL-4 PKG Acts with the KIN-29 SIK and PKA to regulate chemoreceptor gene expression and sensory behaviors in *Caenorhabditis elegans*. *Genetics* **180:** 1475–1491.

Vosshall LB, Amrein H, Morozov PS, Rzhetsky A, Axel R. 1999. A spatial map of olfactory receptor expression in the *Drosophila* antenna. *Cell* **96:** 725–736.

Wes PD, Bargmann CI. 2001. *C. elegans* odour discrimination requires asymmetric diversity in olfactory neurons. *Nature* **410:** 698–701.