© 2014. Published by The Company of Biologists Ltd | Development (2014) 141, 422-435 doi:10.1242/dev.099721

RESEARCH ARTICLE

The LIM and POU homeobox genes *ttx-3* and *unc-86* act as terminal selectors in distinct cholinergic and serotonergic neuron types

Feifan Zhang[1], Abhishek Bhattacharya[1], Jessica C. Nelson[2], Namiko Abe[1], Patricia Gordon[1], Carla Lloret-Fernandez[3], Miren Maicas[3], Nuria Flames[1,3], Richard S. Mann[1], Daniel A. Colón-Ramos[2] and Oliver Hobert[1,4,*]

ABSTRACT

Transcription factors that drive neuron type-specific terminal differentiation programs in the developing nervous system are often expressed in several distinct neuronal cell types, but to what extent they have similar or distinct activities in individual neuronal cell types is generally not well explored. We investigate this problem using, as a starting point, the *C. elegans* LIM homeodomain transcription factor *ttx-3*, which acts as a terminal selector to drive the terminal differentiation program of the cholinergic AIY interneuron class. Using a panel of different terminal differentiation markers, including neurotransmitter synthesizing enzymes, neurotransmitter receptors and neuropeptides, we show that *ttx-3* also controls the terminal differentiation program of two additional, distinct neuron types, namely the cholinergic AIA interneurons and the serotonergic NSM neurons. We show that the type of differentiation program that is controlled by *ttx-3* in different neuron types is specified by a distinct set of collaborating transcription factors. One of the collaborating transcription factors is the POU homeobox gene *unc-86*, which collaborates with *ttx-3* to determine the identity of the serotonergic NSM neurons. *unc-86* in turn operates independently of *ttx-3* in the anterior ganglion where it collaborates with the ARID-type transcription factor *cfi-1* to determine the cholinergic identity of the IL2 sensory and URA motor neurons. In conclusion, transcription factors operate as terminal selectors in distinct combinations in different neuron types, defining neuron type-specific identity features.

KEY WORDS: *Caenorhabditis elegans*, Homeobox, Neuron differentiation

INTRODUCTION

The development of the nervous system is a multistep process that employs a series of sequentially acting regulatory factors that successively restrict and determine cellular fates. During the process of terminal differentiation, individual neuron types acquire specific, hard-wired features that are maintained by the neuron type throughout the life of the animal. A number of transcription factors have been identified that initiate and maintain specific terminal differentiation programs in the developing nervous system (Hobert, 2011). For example, in mouse, the Nurr1 (Nr4a2) transcription factor initiates and maintains the terminal differentiation program of dopaminergic neurons in the midbrain (Smidt and Burbach, 2009), whereas the Pet1 transcription factor initiates and maintains the terminal differentiation program of serotonergic neurons (Liu et al., 2010). However, few neuronal transcription factors are expressed exclusively in only one specific neuronal cell type (Gray et al., 2004; Lein et al., 2007). For example, in addition to being expressed in midbrain dopaminergic neurons, Nurr1 is expressed in other non-dopaminergic neuronal cell types in which its function is not well understood, such as the adult olfactory bulb, specific cortical areas and the hippocampus (Zetterström et al., 1996). The expression of a given transcription factor in distinct neuronal populations poses the fundamental question of whether there are underlying common themes in the activity of the transcription factor in distinct neuronal cell types.

We have undertaken a systematic, in-depth comparison of the activity of two transcription factors in the development of several distinct neuronal cell types in the nematode *C. elegans*, examining whether there are indeed conceptual similarities in the activities of a given transcription factor in distinct neuron types. We used, as a starting point, a member of the LIM homeobox gene family, an ancient family of neuronal patterning genes that display complex expression patterns in the nervous system of many different species, from invertebrates to vertebrates (Hobert and Westphal, 2000; Simmons et al., 2012; Srivastava et al., 2010). One unifying theme is their expression in terminally differentiating neurons (Hobert and Ruvkun, 1998; Moreno et al., 2005). We focus here on the *ttx-3* LIM homeobox gene, which is the sole *C. elegans* member of the Lhx2/9 subclass of LIM homeobox genes. In vertebrates, Lhx2 is expressed in multiple neuronal cell types and is required for the differentiation of olfactory sensory neurons (Hirota and Mombaerts, 2004; Kolterud et al., 2004), the specification of cortical neuron fate (Mangale et al., 2008) and the differentiation of thalamic neurons (Peukert et al., 2011). Whether there is a common theme in the function of Lhx2 in these distinct neuronal cell types is not known.

The *C. elegans* Lhx2/9 ortholog *ttx-3* is exclusively expressed in a small number of neurons in distinct head ganglia (Altun-Gultekin et al., 2001). *ttx-3* null animals display broad differentiation defects in the cholinergic AIY interneuron class. AIY interneurons of *ttx-3* null mutants are generated and still express pan-neuronal features, but fail to express scores of terminal identity markers that define the functional properties of AIY, including genes required to synthesize and package acetylcholine, genes encoding neuropeptide receptors, various types of ion channels and many others (Altun-Gultekin et al., 2001; Hobert et al., 1997; Wenick and Hobert, 2004). TTX-3 exerts this control through direct binding to a *cis*-regulatory motif shared by all of its target genes. *ttx-3* expression is turned on in the

[1]Department of Biochemistry and Molecular Biophysics, Columbia University Medical Center, New York, NY 10032, USA. [2]Department of Cell Biology, Yale University School of Medicine, New Haven, CT 06520, USA. [3]Instituto de Biomedicina de Valencia, IBV-CSIC, 46010 Valencia, Spain. [4]Howard Hughes Medical Institute, Columbia University Medical Center, New York, NY 10032, USA.

*Author for correspondence (or38@columbia.edu)

Received 31 May 2013; Accepted 11 October 2013

Development (2014) doi:10.1242/dev.099721

neuroblast that generates AIY and its expression is maintained throughout the life of the neuron through an autoregulatory feedback loop (Bertrand and Hobert, 2009) to ensure persistent expression of its target genes. A number of transcription factors have been described in the *C. elegans* nervous system that display similar broad-ranging effects on the terminal differentiation programs executed by the neurons in which they are expressed. These transcription factors have been called 'terminal selectors' (Hobert, 2008; Hobert, 2011). It is still an open question how broadly the terminal selector concept applies throughout the nervous system; that is, how common it is that many distinct and functionally unrelated identity features of a specific neuron type are directly co-regulated by a transcription factor or a combination of transcription factors.

Here, we investigate the role of *ttx-3* in two additional neuron classes in which it is normally expressed, namely the cholinergic AIA interneuron class and the serotonergic NSM neuron class. We find in all three neuron classes that there is a common theme of *ttx-3* function in that it is broadly required to induce many distinct and functionally unrelated terminal identity features of the respective neuron class. Yet the downstream targets of *ttx-3* in these neuron classes are distinct and are determined by the cooperation of *ttx-3* with a distinct set of transcription factors in different neuron classes. One of these factors is the POU homeobox gene *unc-86*, which is required together with *ttx-3* to control the identity of the serotonergic NSM neurons. *unc-86* in turn cooperates with the ARID-type transcription factor *cfi-1* to control many terminal identity features of the cholinergic IL2 sensory and URA motor neurons. Our studies therefore provide further support for the terminal selector concept and show that, in combination with other regulatory factors, one factor can serve as terminal selector in distinct neuronal cell types regulating distinct neuronal differentiation programs.

RESULTS

Expression pattern of *ttx-3* in the *C. elegans* larval and adult nervous system

A *ttx-3* reporter gene that contains the *ttx-3* locus together with a few kilobases upstream but no downstream sequences (*ttx-3*^{promA}*::gfp*; Fig. 1) was previously shown to be continuously expressed in five distinct neuronal cell types: the cholinergic AIY and AIA interneuron classes, the ASI and ADL chemosensory neuron classes and a previously uncharacterized neuronal pair in the pharyngeal nervous system (Altun-Gultekin et al., 2001). Transient expression was observed in the AIN and SMDD neurons at embryonic stages (Bertrand and Hobert, 2009). A fosmid reporter construct, which contains more than 30 kb surrounding the *ttx-3* locus and which

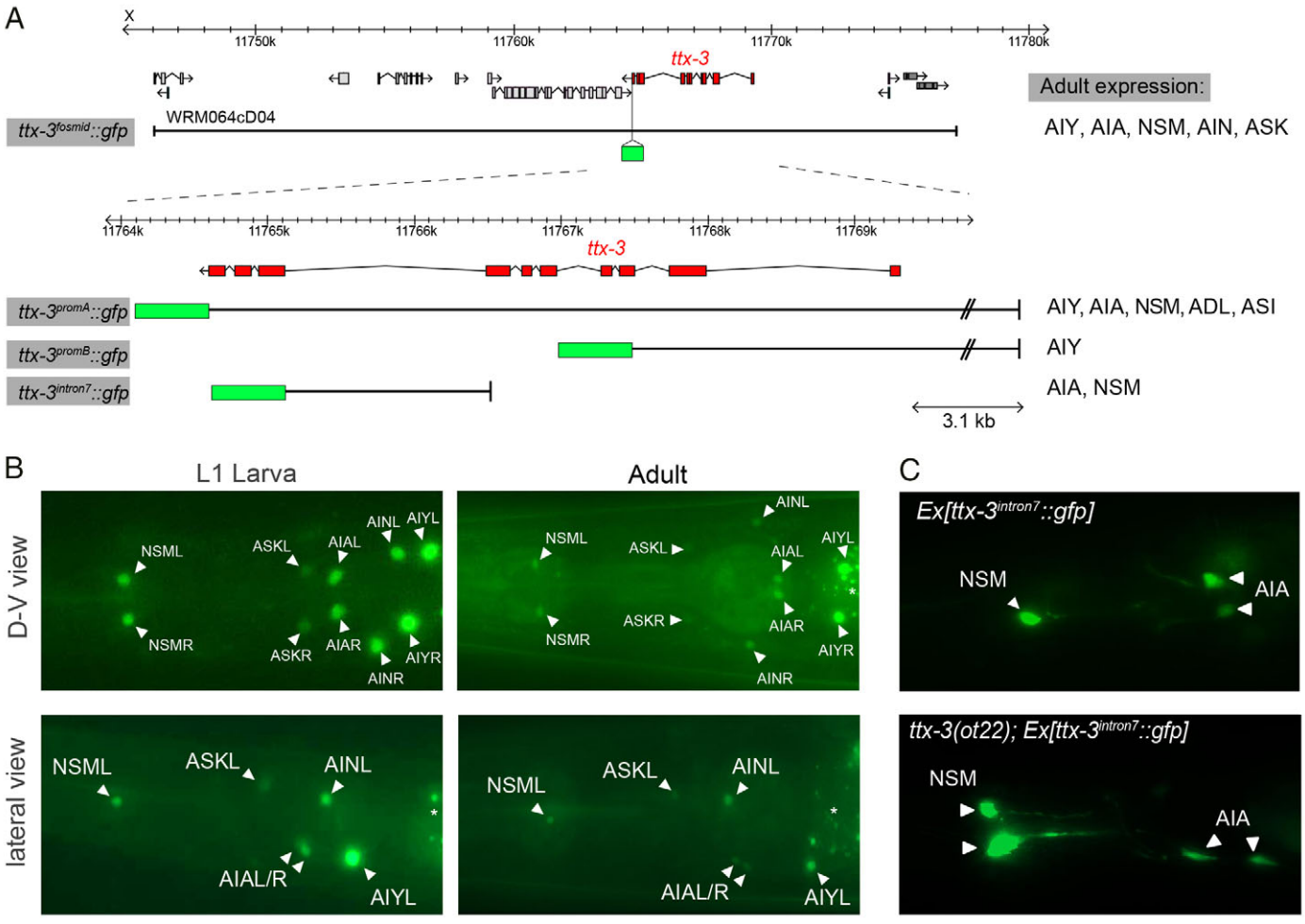


Fig. 1. Expression pattern of the *C. elegans ttx-3* LIM homeobox gene. (A) *ttx-3* expression constructs and summary of neuronal expression pattern. The *promA::gfp* and *promB::gfp* constructs were described previously (Altun-Gultekin et al., 2001; Wenick and Hobert, 2004) and are shown here for comparison only. (B) *ttx-3* fosmid expression (*wgIs68*) in first larval stage animals and in adult animals. D-V, dorsoventral. White asterisks indicate gut autofluorescence. (C) The seventh intron of the *ttx-3* locus contains *cis*-regulatory elements driving reporter gene expression in AIA and NSM neurons. These regulatory elements do not depend on *ttx-3*. Expression is shown in adult animals.

Development (2014) doi:10.1242/dev.099721

rescues the AIY differentiation defect of *ttx-3* mutant animals, mirrors the expression of the smaller, locus-restricted reporter construct in the AIY, the AIA, the AIN and the pharyngeal neuron class (Fig. 1). Based on position, morphology and colabeling with the NSM marker *mgl-1::mCherry*, we identified the pharyngeal neurons that express *ttx-3* as the NSM neuron pair. The NSM neurons are serotonergic, neurosecretory cells that are thought to be involved in sensing food (Albertson and Thomson, 1976; Harris et al., 2011; Horvitz et al., 1982).

There are also notable differences in the expression pattern of the fosmid reporter and the smaller reporters. First, expression in the AIN neurons is maintained throughout development with the fosmid reporter, whereas it is restricted to embryos with smaller reporters (Bertrand and Hobert, 2009). Second, the expression in amphid sensory neurons is markedly different. In larval and adult animals, the fosmid reporter is expressed in the ASK neuron class, whereas the smaller reporters are expressed in the ADL and ASI sensory neurons (Fig. 1).

Previous studies have shown that *ttx-3* expression in the AIY interneuron pair is controlled by a distal initiator element ~1 kb upstream of the *ttx-3* locus and a maintenance element in the second intron of the *ttx-3* locus (Bertrand and Hobert, 2009; Wenick and Hobert, 2004). We find that the expression of *ttx-3* in the NSM and AIA is controlled via regulatory elements present in the seventh intron of the *ttx-3* locus (Fig. 1). As mentioned above, *ttx-3* expression is maintained throughout the life of the AIA and NSM neurons, but maintained expression of a *ttx-3* reporter gene construct (*ttx-3*^{intron7}*::gfp*; Fig. 1A) in the AIA and NSM neuron types does not require *ttx-3* gene activity (Fig. 1C).

ttx-3 controls the differentiation program of AIA interneurons

We focused our analysis of *ttx-3* mutants on the cholinergic AIA interneurons and the serotonergic NSM neurons, which both continuously express *ttx-3* throughout their lifetime. We have previously reported that expression of the marker of cholinergic identity, *unc-17* (vesicular ACh transporter), as well as the expression of an orphan G protein-coupled receptor (GPCR), *sra-11*, is reduced in the AIA neurons of *ttx-3* mutants (Altun-Gultekin et al., 2001). We extended this analysis by examining the expression of seven additional markers of terminal AIA fate: the choline reuptake transporter encoded by *cho-1*; the metabotropic glutamate receptor *mgl-1*; the ionotropic glutamate receptor *glr-2*; the neuropeptides *flp-2* and *ins-1*; the receptor tyrosine kinase *scd-2*; and the receptor guanylyl cyclase *gcy-28d*. Each of these markers is expressed in terminally differentiated AIA interneurons and several of them have previously been implicated in AIA interneuron function (Shinkai et al., 2011; Tomioka et al., 2006). The expression of each of these seven markers is affected in the AIA neurons of *ttx-3* mutants (Fig. 2). Their expression in other neuron types is unaffected in *ttx-3* mutants, with the exception of two markers that are also downregulated in NSM neurons (*mgl-1*, *scd-2*, as described below). *ttx-3* is likely to act cell-autonomously since the AIA differentiation defects are rescued in transgenic *ttx-3* mutant animals that express *ttx-3* cDNA under control of the *ins-1* promoter (supplementary material Table S1).

AIA neurons remain present in the *ttx-3* null mutant, as assessed by the weak but recognizable expression of some terminal differentiation genes (Fig. 2). However, their normally unipolar neurite morphology appears disrupted; ectopic branches can be observed to emanate from the cell body and the main neurite appears blebbed in *ttx-3* mutants (supplementary material Fig. S2). The AIY interneurons, which have a unipolar axon morphology similar to that of AIA interneurons in wild-type animals, display similar morphological defects in *ttx-3* mutants (Hobert et al., 1997). The expression of terminal identity markers that label several distinct neuron types that are lineally related to AIA is not altered (data not shown) (Altun-Gultekin et al., 2001), suggesting that the AIA neuron pair might remain in an undifferentiated state, rather than switching to an alternate fate. Based on a more extensive cell fate marker analysis, a similar conclusion was previously drawn about the fate of the AIY neuron class in *ttx-3* mutants (Altun-Gultekin et al., 2001). Taken together, our fate marker and morphological analyses indicate that *ttx-3* broadly affects the AIA terminal differentiation program. These effects are comparable to the previously described broad effects that loss of *ttx-3* has on the terminal differentiation of AIY interneurons.

A shared *cis*-regulatory signature of AIA-expressed terminal identity features

On a mechanistic level, *ttx-3* operates in a distinct manner in the AIA versus AIY neurons since it operates with distinct co-factors and through distinct *cis*-regulatory elements. The co-factor of *ttx-3* in AIY, the *ceh-10* homeobox gene (Altun-Gultekin et al., 2001), is not expressed in AIA neurons, and AIA neurons display no differentiation defects in *ceh-10* null mutants (two markers tested). Moreover, the *cis*-regulatory motifs through which *ttx-3* acts to control AIY versus AIA identity are distinct. In the AIY neurons, *ttx-3* acts on its many target genes through a *cis*-regulatory motif, termed the 'AIY motif', that provides a cooperative binding site for a TTX-3–CEH-10 heterodimer (Wenick and Hobert, 2004). Mutation of the AIY motif in a locus that is expressed in AIY and AIA neurons, the cholinergic *cho-1* locus, results in a severe reduction in expression in the AIY interneurons but not in the AIA interneurons (Fig. 3A).

In the AIA neurons, by contrast, *ttx-3* acts through a distinct *cis*-regulatory signature, which we deciphered through a mutational analysis of the *cis*-regulatory control regions of three AIA-expressed, *ttx-3*-dependent terminal differentiation genes: *mgl-1*, *ins-1* and *cho-1*. We generated transgenic animals that express nested, shorter versions of these three reporters and identified a 259 bp element in the *cho-1* promoter, a 74 bp element in the *mgl-1* promoter and a 68 bp element in the *ins-1* promoter that are sufficient to direct *gfp* expression to AIA neurons (Fig. 3A-C). Examining these elements for common patterns, we noted that all these elements contain a shared and phylogenetically conserved G(A/G)ATC motif (Fig. 3D). Mutating this motif in the context of any of the three promoters resulted in a reduction of AIA expression of the respective reporter (Fig. 3A-C). In the case of *mgl-1*, two G(A/G)ATC motifs are present in the minimal promoter; mutation of either causes an intermediate reduction in reporter gene expression, and mutation of both motifs results in complete loss of expression (Fig. 3A-C).

Since G(A/G)ATC does not match the consensus binding site for a LIM homeodomain transcription factor such as TTX-3, we also examined the minimal reporters for the presence of conserved TAAT motifs, which comprises the core consensus site for LIM homeodomain transcription factors (Berger et al., 2008). We indeed found several TAAT motifs in the three *cis*-regulatory modules and for each of them we identified a TAAT motif that, when mutated, affected reporter gene expression *in vivo* (Fig. 3A-C). These TAAT motifs can be assembled into a larger sequence matrix, TAATTNGA (Fig. 3D). In two cases, mutation of the TAATTNGA alone affected reporter gene expression, whereas in the third case (*cho-1*) a

Development (2014) doi:10.1242/dev.099721

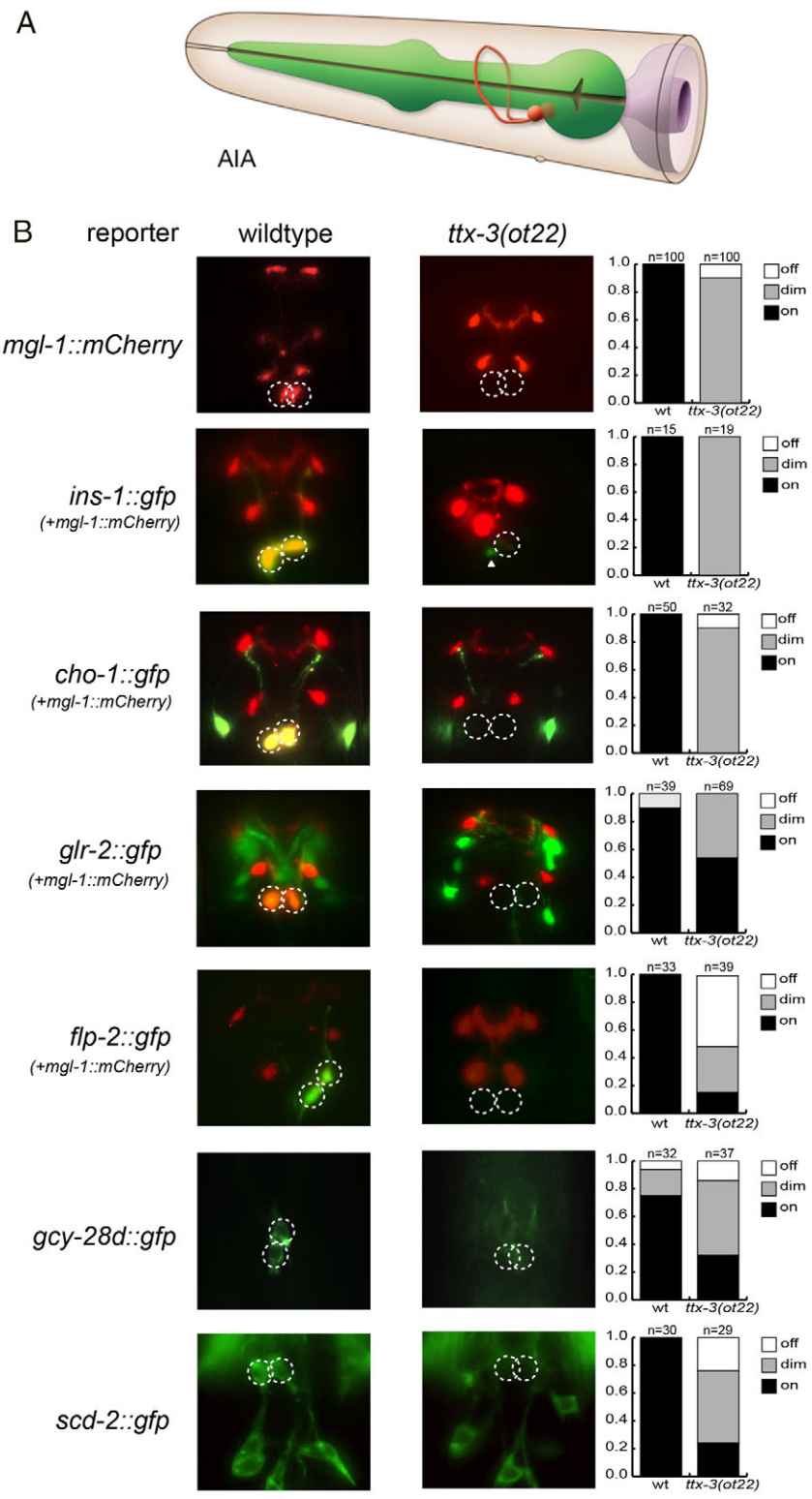


Fig. 2. *ttx-3* affects the terminal differentiation of AIA neurons. (A) Schematic representation of the AIA interneuron pair [reproduced with permission (Altun et al., 2002-2013)]. (B) The expression of terminal differentiation markers of AIA identity is affected in *ttx-3* mutants. Reporter gene arrays were crossed into *ttx-3(ot22)* null mutants. Positions of AIA neurons are outlined (dashed circles). The fraction of animals that show the indicated phenotype is presented in the bar charts. Transgenic arrays are: *otIs317* for *mgl-1*, *otIs326* for *ins-1*, *otIs379* for *cho-1*, *otEx4687* for *glr-2* and *otEx5056* for *flp-2* (see Materials and methods for more detail on the arrays; the *Ex[gcy-28d::gfp]* and *Ex[scd-2::gfp]* arrays were kindly provided by Takeshi Ishihara). Anterior is up in all panels.

complete loss of expression can only be observed upon simultaneous mutation of both the GAATC motif and the TAAT-containing motif (Fig. 3A). The residual AIA expression of a *cho-1* reporter construct in which the GAATC motif is mutated, but the TAAT motif is left intact, is abolished in *ttx-3* mutants (data not shown), consistent with *ttx-3* operating through the TAAT motif.

We examined whether the TAATTNGA motif is indeed a TTX-3 binding site using gel shift assays with bacterially produced TTX-3 protein and probes derived from the *mgl-1* and *cho-1* locus. We found that TTX-3 is able to bind these sites *in vitro* (Fig. 3E). Deletion of the TAAT site that is required for reporter gene expression *in vivo* resulted in the loss of TTX-3 binding *in vitro* (Fig. 3E).

The combination of G(A/G)ATC and TAAT motifs might define a *cis*-regulatory signature that is generally required for gene expression in AIA neurons, since we found a combination of these two motifs to be present in the *cis*-regulatory control regions of the

Development (2014) doi:10.1242/dev.099721

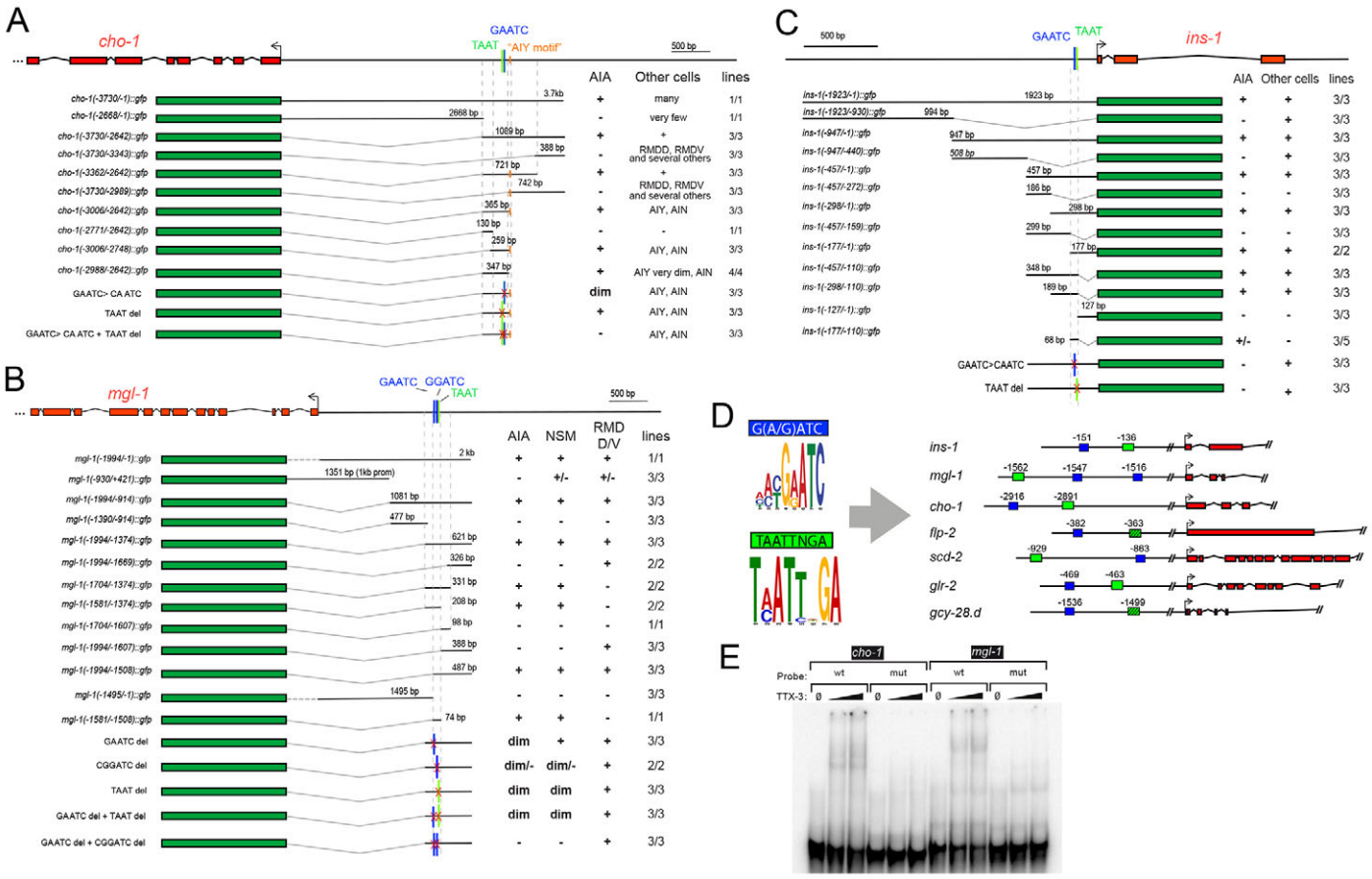


Fig. 3. Co-regulation of AIA-expressed genes by two *cis*-regulatory motifs. (A-C) Mutational dissection of the *cis*-regulatory elements of three AIA-expressed terminal identity markers. (D) Position weight matrix of the two motifs required for AIA expression, based on the motifs from *ins-1*, *cho-1* and *mgl-1* and orthologs in other nematode species. Perfect (filled box) and imperfect (stippled box) matches to the two *cis*-regulatory motifs [blue, G(A/G)ATC; green, TAATTNGA] in other AIA terminal identity markers are shown on the right. (E) TTX-3 binds to *cho-1* and *mgl-1* regulatory elements containing the HD (TAAT) motif. Deletion of the HD motif abolishes binding. EMSA was performed with 250 nM and 750 nM TTX-3.

other four *ttx-3*-dependent terminal AIA markers (Fig. 3D). Taken together, these data show that AIA identity features are co-regulated by a shared *cis*-regulatory signature that is controlled by TTX-3 and an as yet unknown co-factor.

ttx-3 controls the terminal differentiation of serotonergic NSM neurons

We next analyzed the effect of loss of *ttx-3* on the terminal differentiation program of the serotonergic NSM neurons, a neuron type that has not previously been examined in *ttx-3* mutants. Many terminal identity markers of NSM have been described, including the battery of genes that are required to synthesize, package and reuptake serotonin: *tph-1/TPH* (tryptophan hydroxylase), *cat-4/GTPCH* (GTP cyclohydrolase), *cat-1/VMAT* (vesicular monoamine transporter), *bas-1/AAAD* (aromatic amino acid decarboxylase) and *mod-5/SERT* (serotonin reuptake transporter) (Fig. 4A) (Jafari et al., 2011; Ranganathan et al., 2001; Sze et al., 2002). Previous expression analysis of a vesicular glutamate transporter, *eat-4*, suggested that NSM might use the neurotransmitter glutamate (Lee et al., 1999). However, a fosmid-based *eat-4* reporter does not show expression in NSM neurons (Serrano-Saiz et al., 2013) (supplementary material Fig. S1A).

To broaden the spectrum of available terminal markers, we analyzed the expression of other *C. elegans* orthologs of enzymes involved in monoaminergic transmitter metabolism (Fig. 4A) and identified another NSM-expressed terminal marker, *ptps-1* (Fig. 4C; supplementary material Fig. S3). In addition to examining these serotonin (5HT)-related markers, we also examined the expression of three metabotropic neurotransmitter receptors (*mgl-1*, *mgl-3*, *dop-3*), three neuropeptides (*nlp-13*, *flp-4*, *nlp-3*), a glycoprotein hormone alpha subunit (*flr-2*) and a receptor tyrosine kinase (*scd-2*). All of these genes are expressed throughout the life of the NSM neurons. As mentioned above, *scd-2* and *mgl-1* are also expressed in AIA neurons, where their expression is affected by *ttx-3*. We find that the expression of five of these 14 NSM terminal identity markers is either partially or completely eliminated in the NSM neurons of *ttx-3* null mutants (Fig. 4C, Fig. 5, Table 1). *ttx-3* is likely to act cell-autonomously since we can rescue the NSM differentiation defects by driving *ttx-3* cDNA under the control of a *cat-1* promoter fragment, which is expressed in a subset of monoaminergic neurons of *C. elegans* (supplementary material Table S1).

The POU homeobox gene *unc-86* also controls NSM identity

We recently reported that the effects of the loss of a terminal selector type transcription factor in dopaminergic neurons can be partially compensated for by other, co-expressed terminal selectors (Doitsidou et al., 2013). Therefore, we considered the possibility that

Development (2014) doi:10.1242/dev.099721

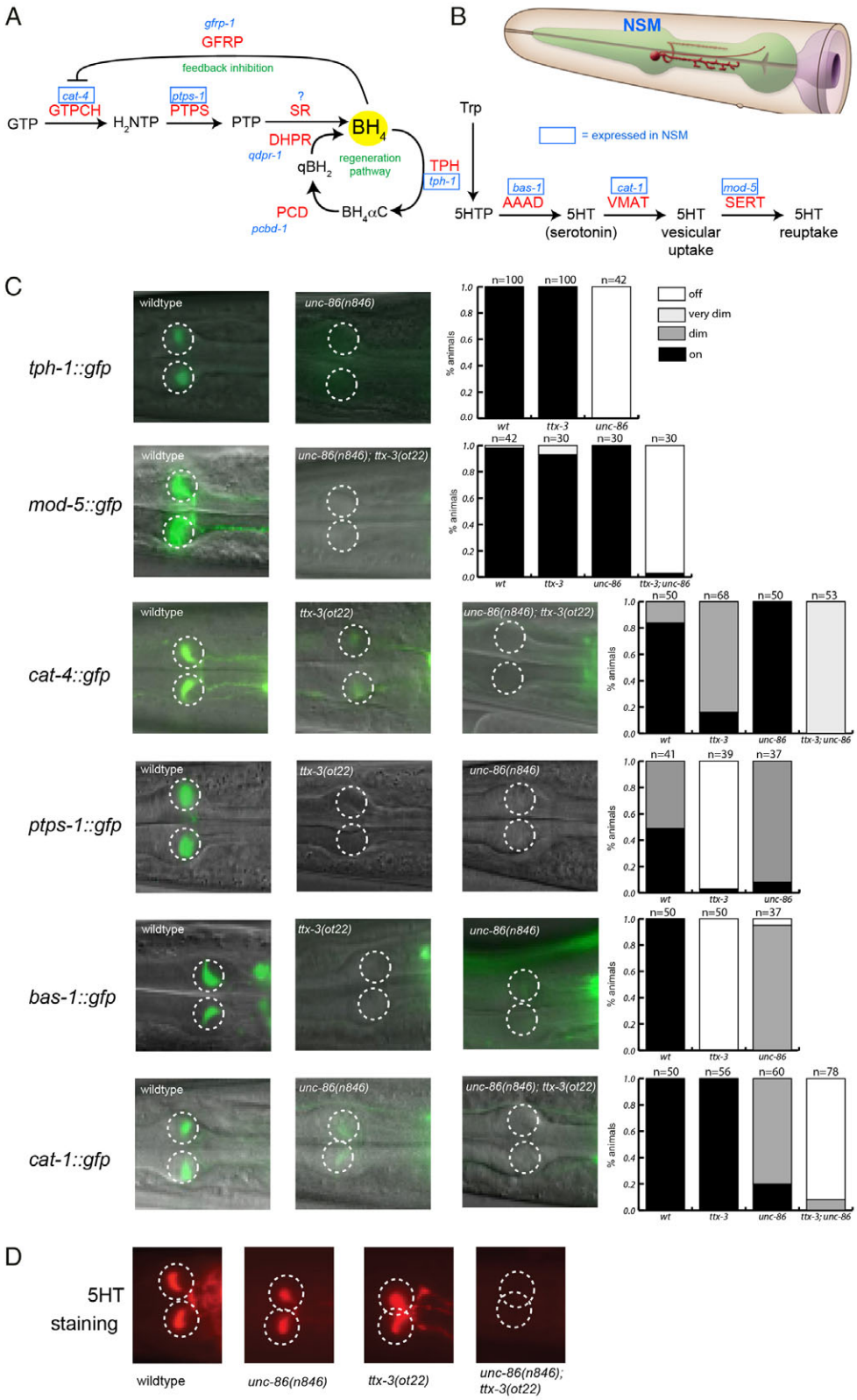


Fig. 4. The effect of *unc-86* and *ttx-3* on the serotonergic identity of NSM neurons. (A) The 5HT pathway including tetrahydrobiopterin biosynthesis genes (Deneris and Wyler, 2012). '?' indicates that a unique homolog of SR could not be identified in the worm genome. (B) Schematic representation of the NSM interneuron pair [reproduced with permission (Altun et al., 2002-2013)]. (C) The expression of serotonergic identity features of NSM (dashed circles) is affected in *unc-86(n846)*, *ttx-3(ot22)* or *unc-86(n846); ttx-3(ot22)* double-null mutants. Reporter gene arrays were crossed into the respective mutant backgrounds. Transgenic arrays are: *zdIs13* for *tph-1*; *otEx4781* for *mod-5*; *otIs225* for *cat-4*; *otEx5280* for *ptps-1*; *otIs226* for *bas-1*; and *otIs224* for *cat-1* (see Materials and methods for more detail on the arrays). Images are only shown for mutant genotypes with effects on reporter expression. (D) Serotonin antibody staining. Thirty animals were scored for each genotype. In the double mutant, no animal showed staining in NSM (circled), whereas in the other genotypes all animals showed staining.

Development (2014) doi:10.1242/dev.099721

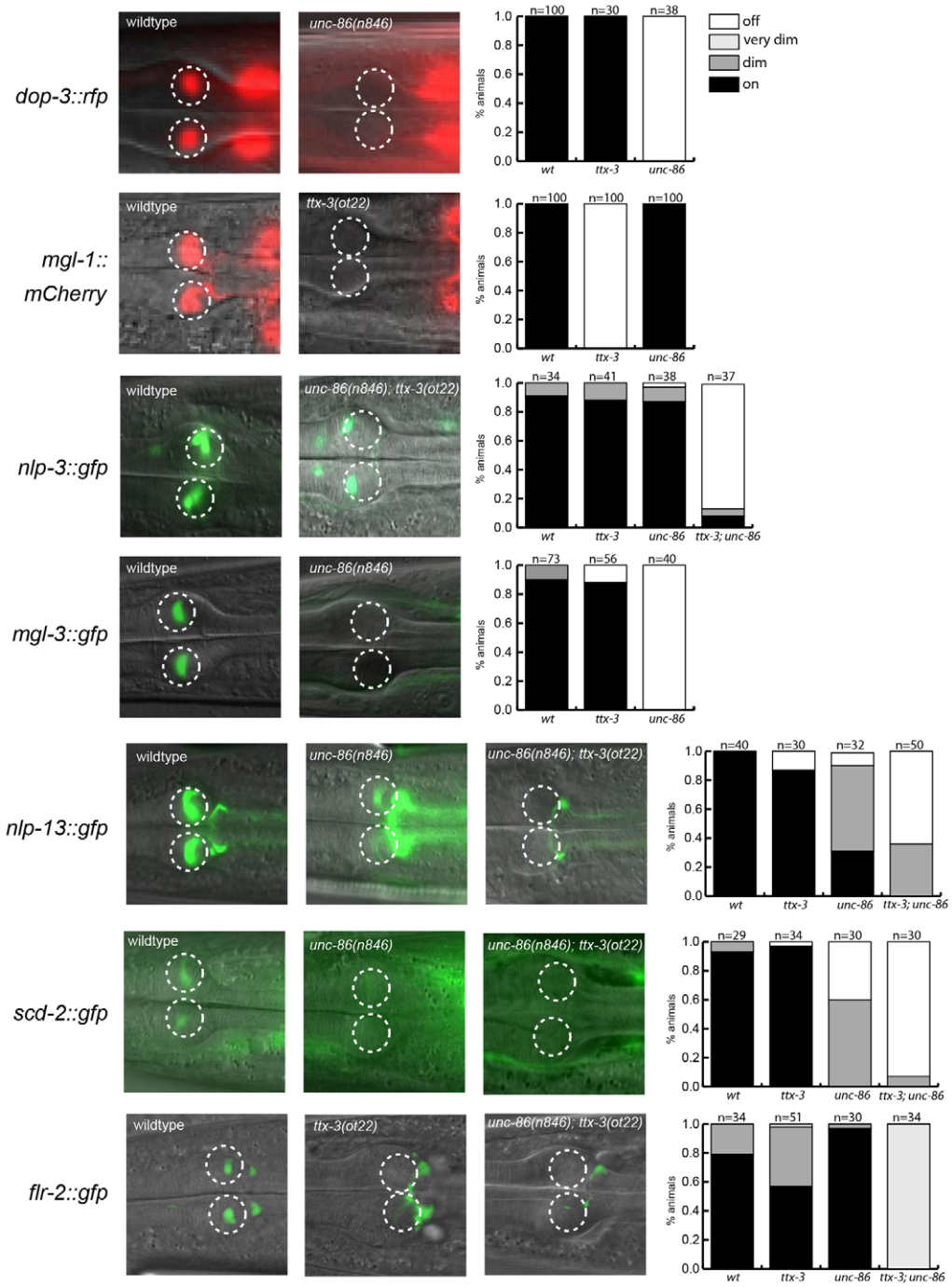


Fig. 5. The effect of *unc-86* and *ttx-3* on other identity features of NSM neurons. The expression of other identity features of NSM is also affected in *unc-86(n846)*, *ttx-3(ot22)* or *unc-86(n846); ttx-3(ot22)* double-null mutants. Reporter gene arrays were crossed into the respective mutant backgrounds. Transgenic arrays are: *vsIs33* for *dop-3*; *otIs317* for *mgl-1*; *otEx5163* for *nlp-3*; *otEx5364* for *mgl-3*; *otEx5163* for *nlp-13*; *otEx5055* for *scd-2*; and *otEx5363* for *flr-2* (see Materials and methods for more detail on the arrays). Micrographs are only shown for mutant genotypes with effects on reporter expression. Dashed circles indicate the position of NSM neurons. See also Table 1.

the lack of an impact of *ttx-3* loss on nine out of 14 NSM markers could be due to the activity of compensatory terminal selector type transcription factors. We sought to identify such a factor, focusing on two homeodomain transcription factors previously shown to be expressed in NSM, namely the empty spiracles homolog *ceh-2* and the POU homeobox gene *unc-86* (Aspöck et al., 2003; Finney and Ruvkun, 1990). We observed no NSM differentiation defects in *ceh-2* null animals (data not shown), but we observed striking NSM differentiation defects in *unc-86* mutants. Loss of *unc-86* was previously shown to affect the expression of *tph-1* and *cat-1* in NSM neurons, but without effect on 5HT antibody staining (Sze et al., 2002). Other differentiation features of NSM neurons had not previously been examined in *unc-86* mutants. Upon examining the expression of all 14 markers of NSM fate in *unc-86* null mutants, we found that the expression of eight is partially or completely eliminated (Figs 4, 5, Table 1).

To examine whether *unc-86* directly affects the expression of these terminal identity features, we analyzed the *cis*-regulatory control regions of four of them: *tph-1*, *bas-1*, *cat-1* and *cat-4*. Through mutational analysis, we defined small (~200 bp) elements that still yielded expression in the NSM neurons (Fig. 6) and, within each of these elements, we identified predicted POU homeodomain binding sites (Rhee et al., 1998). We introduced mutations into these sites in the context of two loci (*tph-1* and *bas-1*) and found that these mutations resulted in a loss of reporter gene expression *in vivo* (Fig. 6A,B). Gel shift analysis further confirmed that these POU homeodomain sites indeed bind bacterially produced UNC-86 protein *in vitro* (Fig. 6E).

Development (2014) doi:10.1242/dev.099721

Table 1. Summary of the effects of *ttx-3* and *unc-86* null mutants on terminal NSM identity markers

Identity feature	Function	*ttx-3(–)*	*unc-86(–)*	*unc-86(–); ttx-3(–)*	Interaction
cat-1	5HT pathway	wt	dim	off	Synergism
cat-4	5HT pathway	dim	wt	very dim	Synergism
mod-5	5HT pathway	wt	wt	off	Synergism
nlp-13	Neuropeptide	wt	dim	off	Synergism
nlp-3	Neuropeptide	wt	wt	off	Synergism
flr-2	Transmembrane	wt	wt	dimmer	Synergism
scd-2	Kinase	wt	dim	off	Synergism
flp-4	Neuropeptide	dim	wt	stronger expression	Synergism
5HT antibody staining	Neurotransmitter	wt	wt	off	Synergism
bas-1	5HT pathway	off	dim	n.d.	?
ptps-1	5HT pathway	off	dim	n.d.	?
tph-1	5HT pathway	wt	off	n.d.	?
mgl-3	GPCR	wt	off	n.d.	?
dop-3	GPCR	wt	off	n.d.	?
mgl-1	GPCR	off	wt	n.d.	?

gfp reporter (or antibody staining): wt, as bright as in wild-type animals; dim, dimmer than in wild type; off, no expression observed.
n.d., not determined because single mutant already shows completely penetrant loss of expression.
Gray shading indicates presence of defect.

unc-86 cooperates with *ttx-3* to control NSM identity

We noted that terminal markers of NSM identity that were severely affected in *unc-86* mutants tended to be those that were weakly or unaffected in *ttx-3* mutants; vice versa, markers unaffected in *ttx-3* mutants tended to be affected in *unc-86* mutants (Table 1). Even though this observation might simply indicate that *unc-86* and *ttx-3* act completely independently of one another, we considered the possibility that *unc-86* and *ttx-3* might collaboratively control NSM identity but that their relative importance may be distinct for different target genes. To investigate this possibility, we examined *unc-86; ttx-3* double-null mutants and found that markers that are either partially or unaffected in *ttx-3* and *unc-86* single mutants are more strongly affected in the double mutant (Figs 4, 5, Table 1). This also holds for 5HT antibody staining, which is not affected in either single mutant but completely abrogated in the *ttx-3*; *unc-86* double mutant (Fig. 4D), probably owing to the combined effect that both genes have on the expression of the 5HT reuptake transporter *mod-5*. As summarized in Table 1, nine of the 15 tested identity features (14 reporter genes and 5HT antibody staining) are affected by both *ttx-3* and *unc-86*, with effects either visible in both single mutants, or as a non-additive, synergistic effect revealed in the double mutant. As described below, there are also synergistic effects of *ttx-3* and *unc-86* on NSM morphology. In six of the 15 cases, either *ttx-3* or *unc-86* already has completely penetrant effects (Table 1). Taken together, these data argue that *unc-86* and *ttx-3* jointly control terminal NSM differentiation. The mechanistic basis of the cooperation is unclear at present because we have so far not been able to identify functional TTX-3 binding sites in terminal NSM identity marker genes.

To further examine potential interactions of *unc-86* and *ttx-3*, we investigated whether they affect each others expression. We find that continuous expression of *unc-86* in NSM neurons depends on *unc-86* itself [autoregulation of *unc-86* was also previously noted (Baumeister et al., 1996)], but not on *ttx-3* (supplementary material Fig. S1B,C). Vice versa, *ttx-3* expression in NSM neurons is not affected in *unc-86* or in *unc-86; ttx-3* mutants (data not shown).

unc-86 and *ttx-3* affect axonal arborization and presynaptic specializations

Apart from affecting the expression of terminal identity markers, loss of *unc-86* and *ttx-3* also results in specific effects on the morphology of the NSM neurons. During embryonic stages, these neurons normally extend a neurite posteriorly toward the nerve ring, which then bifurcates to form a ventral and a dorsal neurite (Axäng et al., 2008) (Fig. 4B). We observe that in *unc-86(n846)* mutants the NSM somas are correctly positioned but there are significant defects in outgrowth of the ventral neurite (61% of animals show outgrowth defects; n=31). By contrast, in *ttx-3(ot22)* mutants, the primary defect observed in ventral neurite outgrowth is the formation of aberrant bifurcations (41% of animals show such defects; n=39). Both *ttx-3* and *unc-86* single mutants also show defects in dorsal axon termination [25% (n=32) of *ttx-3* mutants and 29% (n=34) of *unc-86* mutants].

In early larval stages, ventral NSM neurites begin to extend elaborate arbor structures onto the nerve ring target field (Axäng et al., 2008). These axon arborizations require the NSM-expressed netrin receptor UNC-40/DCC, which is tightly localized to puncta within the main shaft of the NSM neurite and at the tips of axon arbors (Nelson and Colón-Ramos, 2013). These arbor structures persist into the adult stage and contain presynaptic sites, as assessed with a *rab-3* marker (Nelson and Colón-Ramos, 2013). In *ttx-3* mutants, these ultrastructural features are unaffected, but *unc-86* mutants display a highly penetrant defect in axon arborization (Fig. 7A,C). Furthermore, *unc-86* mutants display defects in the dynamic regulation of UNC-40 localization (Fig. 7B,D). In wild-type animals, UNC-40::GFP is diffusely distributed at the L1 stage and becomes localized to bright puncta in the NSM neurite and at the tips of axon arbors as axons are arborizing at the L4 stage. By the adult stage, UNC-40::GFP again becomes diffusely distributed. However, a significant fraction of *unc-86* mutant NSMs retain a juvenile-like pattern of UNC-40 localization during the adult stage, in which UNC-40::GFP remains localized to bright puncta (Fig. 7B).

We observe synergistic morphological defects in *unc-86(n846); ttx-3(ot22)* double mutants. Ventral neurites never reach the middle of the pharyngeal isthmus and are often truncated immediately following the guidance decision to turn posteriorly (100% premature ventral neurite termination, n=18; Fig. 7E). Furthermore, neurites contain large anterior swellings not seen in wild-type animals (33% contain additional anterior swellings, n=18; Fig. 7E). The morphological appearance of NSM neurites in *unc-86; ttx-3* mutants is reminiscent of the normal morphology of M3 neurons (Albertson and Thomson, 1976), which are lineally related to NSM (Sulston et al., 1983). M3 neurons are glutamatergic (Lee et al., 1999) and we indeed find that in *unc-86* mutants the vesicular glutamate

Development (2014) doi:10.1242/dev.099721

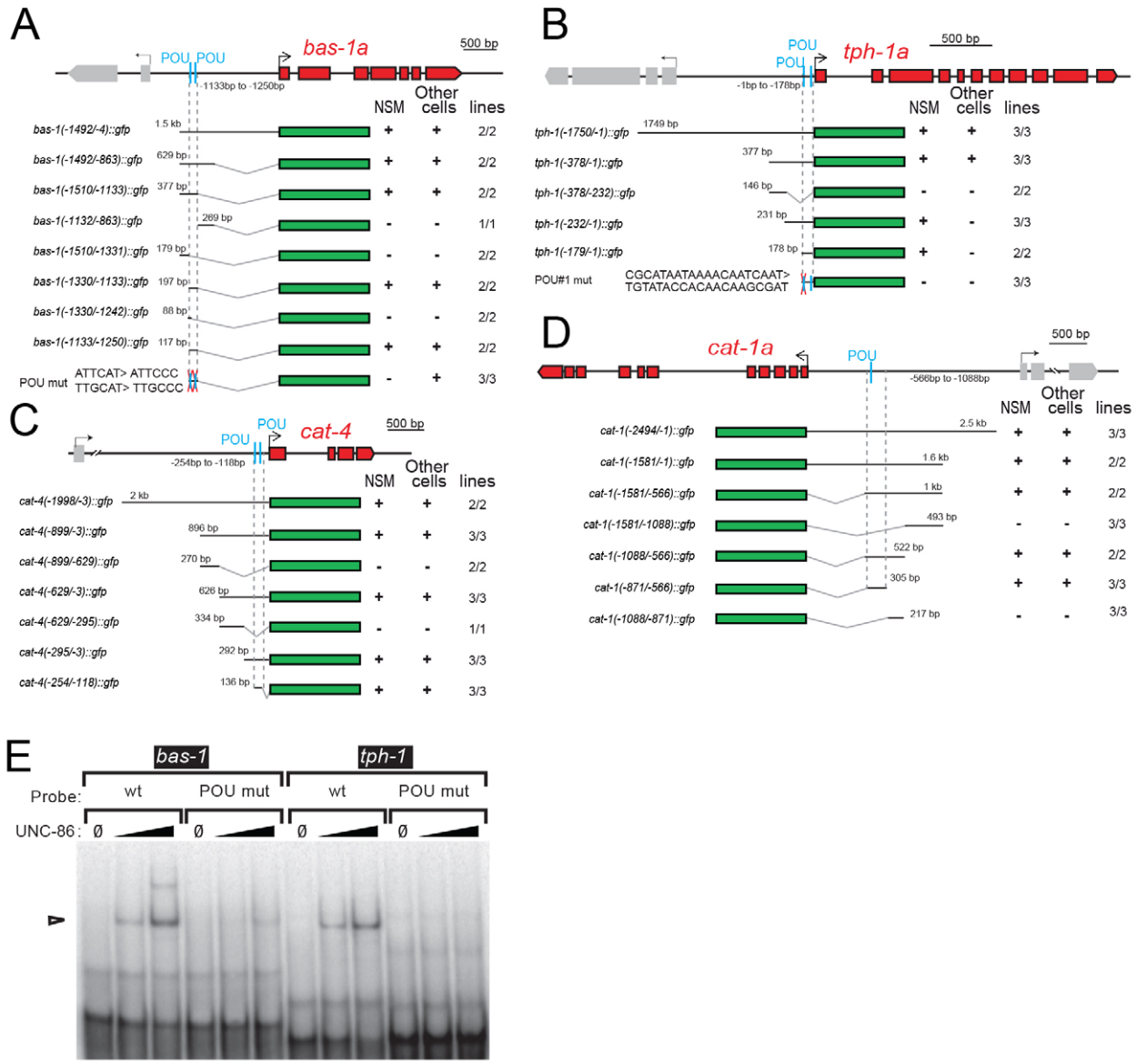


Fig. 6. *Cis*-regulatory analysis of NSM identity specification. (A-D) Dissection of the *cis*-regulatory elements of four NSM-expressed serotonin pathway genes. All minimal *cis*-regulatory elements contain predicted POU sites. (E) EMSAs with UNC-86 protein on *bas-1* and *tph-1* regulatory elements. Mutated POU sites are those that also disrupt reporter gene activity when deleted from the gene contexts of *bas-1* and *tph-1* (A). EMSA was performed with 10 nM and 30 nM UNC-86. Arrowhead indicates UNC86-bound DNA probe.

transporter *eat-4* is ectopically expressed in NSMs (supplementary material Fig. S1A).

unc-86 controls terminal differentiation of the cholinergic IL2, URA and URB sensory, motor and interneurons

Apart from our description of *unc-86* terminal selector function in the serotonergic NSM neurons, *unc-86* had previously been described to broadly affect the terminal differentiation program of other serotonergic (Sze et al., 2002) as well as glutamatergic (Duggan et al., 1998; Serrano-Saiz et al., 2013) neurons. We asked whether *unc-86* might affect the terminal differentiation program of neurons that use yet another neurotransmitter system. We turned to the six IL2 sensory neurons that are involved in nictation behavior (Lee et al., 2012). The IL2 neurons express *unc-86* throughout their lifetime and have been inferred to be cholinergic (Lee et al., 2012). We corroborated the cholinergic identity of the IL2s by finding that reporter fusions to the *unc-17/cha-1* locus and to the choline reuptake transporter *cho-1* are expressed in IL2 neurons (Fig. 8A). The expression of these two key markers of cholinergic identity is eliminated in *unc-86* mutants (Fig. 8A). Expression of the nicotinic acetylcholine receptor subunit *des-2* is also lost in the IL2 neurons of *unc-86* mutants (Treinin et al., 1998).

In addition to these cholinergic markers, we examined the expression of other genes previously shown to be expressed in IL2 neurons, namely the *unc-5* netrin receptor, the guanylyl cyclase *gcy-19*, the kinesin *klp-6* and the Notch ligand *lag-2* (which is expressed in IL2 neurons at the dauer stage) (Leung-Hagesteijn et al., 1992; Ortiz et al., 2006; Ouellet et al., 2008; Peden and Barr, 2005). The expression of all of these terminal markers of IL2 identity is eliminated in IL2 neurons of *unc-86* mutants (Fig. 8A). IL2 neurons also fail to take up dye in *unc-86* mutants (Tong and Bürglin, 2010), suggesting morphological defects. The IL2 neurons are nevertheless

Development (2014) doi:10.1242/dev.099721

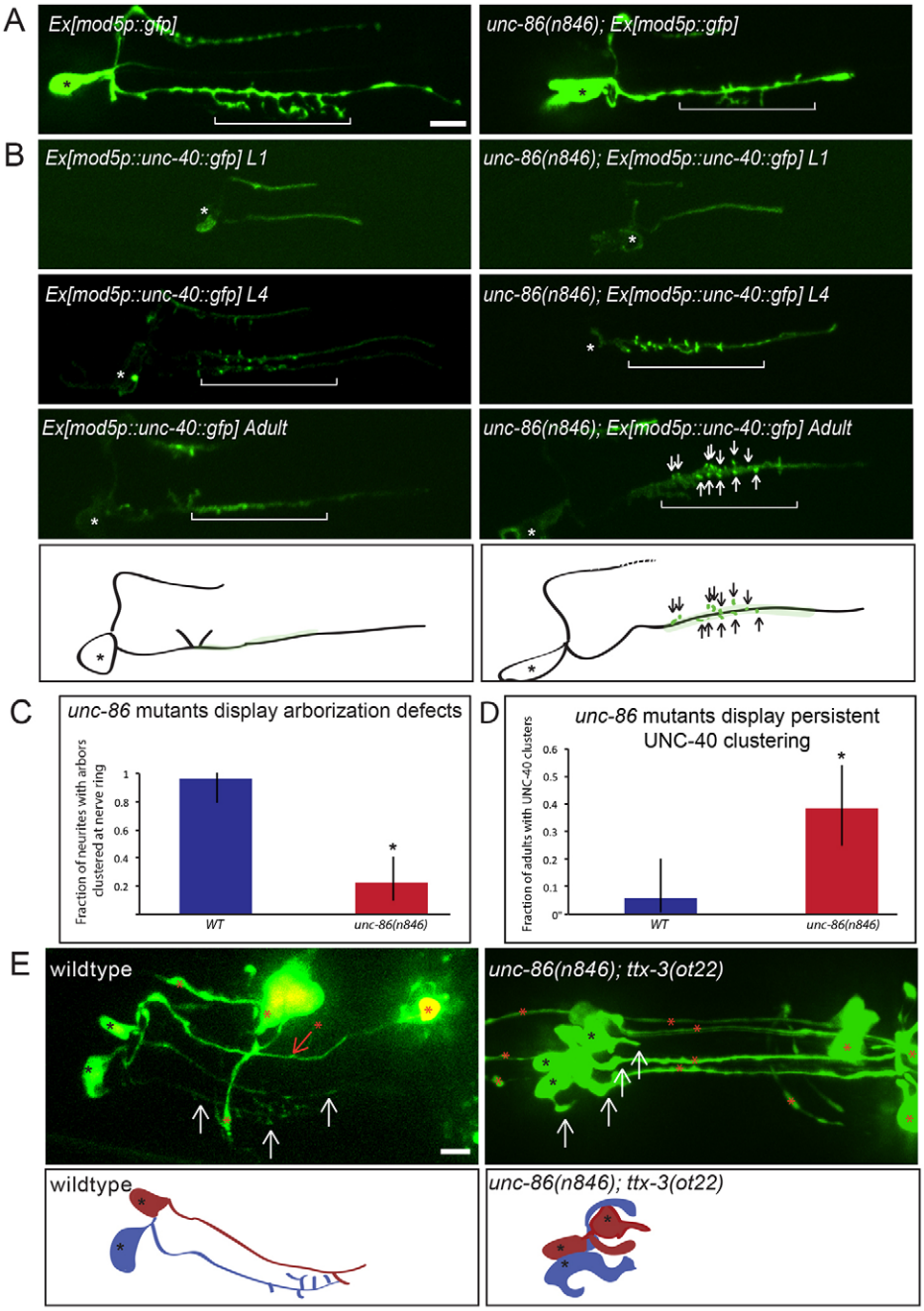


Fig. 7. *unc-86* and *ttx-3* affect NSM morphology. (A) *unc-86(n846)* mutant adults display shorter ventral neurites and fewer and shorter axon arbors. *mod-5p::gfp* (*olaEx1446*) is used to visualize NSM morphology. (B) UNC-40::GFP localization (transgene: *olaEx1448*) remains in a juvenile state in *unc-86* mutant animals. White arrows indicate UNC-40::GFP puncta. (C) Quantification of the *unc-86(n846)* arborization phenotype. Displayed is the fraction of animals with clusters of arbors in the nerve ring region in wild-type and *unc-86(n846)* animals. The difference between wild type and *unc-86* is significant (**P*<0.0001). (D) Quantification of the *unc-86(n846)* UNC-40::GFP localization phenotype. The fraction of animals with multiple, bright UNC-40::GFP puncta in the nerve ring region is displayed. The difference between wild type and *unc-86* is significant (**P*=0.0017). Error bars indicate 95% confidence intervals. (E) *unc-86(n846); ttx-3(ot22)* double mutants display numerous NSM morphology defects, as visualized with *flp-4::gfp (olaEx1485*). In all images, anterior is to the left and ventral down. Asterisks indicate cell bodies (A,B) or additional cell-body-like swellings (E). Brackets denote the nerve ring terminal field where arbors form. White arrows indicate NSM neurites, red arrows and asterisks denote other non-NSM structures. Fisher's *t*-test was used for statistical analysis. Scale bars: 5 µm.

generated in *unc-86* mutants, as assessed by intact expression of the pan-neuronal marker *rab-3* and the pan-sensory marker *osm-6* (50 animals were scored for each marker).

unc-86 is expressed in two additional cholinergic neuron classes in the anterior ganglion besides the IL2 sensory neurons, namely the URA motoneurons [which are synaptically connected to the IL2 neurons (White et al., 1986)] and the URB interneurons. We found that cholinergic identity was also strongly affected in both URA and URB neurons of *unc-86(n846)* loss-of-function mutants (supplementary material Fig. S4).

unc-86 cooperates with the ARID transcription factor *cfi-1* to control IL2 and URA identity

Since none of the previously known co-factors of *unc-86* [*mec-3* for touch neurons (Duggan et al., 1998) and *ttx-3* for NSM neurons (this paper)] is expressed in IL2, URA or URB neurons, *unc-86* is likely to act with another co-factor in IL2 neurons. *cfi-1* is an ARID transcription factor previously shown to be co-expressed with *unc-86* specifically in IL2 and URA neurons (Shaham and Bargmann, 2002). Loss of *cfi-1* results in ectopic expression of identity markers for the CEM neuron in IL2 and URA neurons (Shaham and Bargmann, 2002), which prompted us to investigate whether *cfi-1* might also positively control their cholinergic identity. We find that the cholinergic identity of both IL2 and URA neurons is affected in *cfi-1(ky651)* loss-of-function mutants, albeit not as strongly as in *unc-86* null mutants (Fig. 8A; supplementary material Fig. S4). To investigate whether *unc-86* and *cfi-1* genetically interact, we examined non-additive synergistic interactions of the two genes using a hypomorphic *unc-86* allele, *n848*. Animals carrying this allele show mild IL2 and URA differentiation defects, but in

Development (2014) doi:10.1242/dev.099721

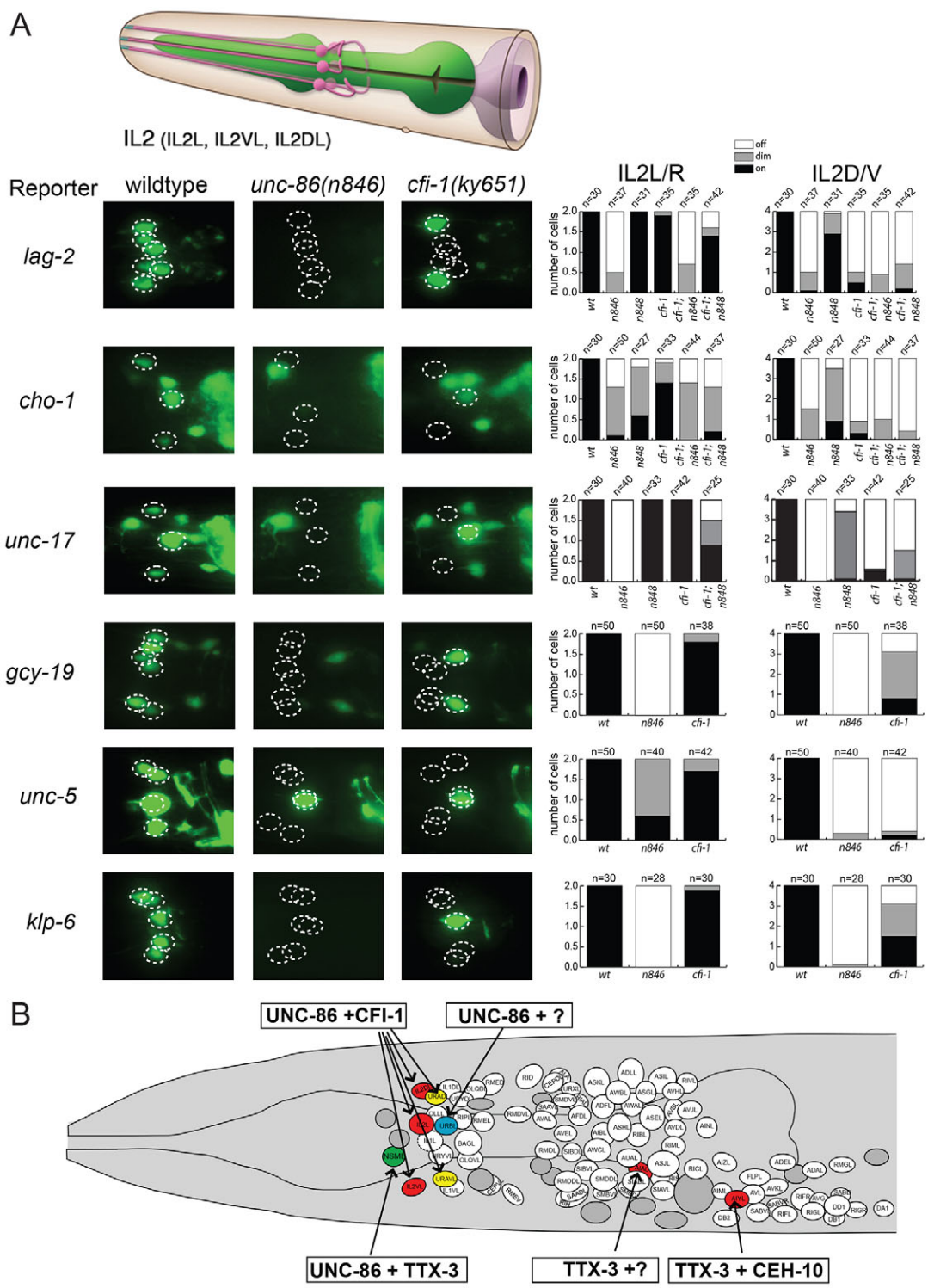


Fig. 8. *unc-86* and *cfi-1* control cholinergic IL2 neuron identity. (A) Animals are late L4 or young adults, with the exception of the *lag-2::gfp* transgenic animals which are dauers. The differential importance of *cfi-1* in the dorsal IL2DL/R and ventral IL2VL/R neurons versus the lateral IL2L/R neurons mirrors morphological differences of the ventral versus lateral neurons, with the lateral neurons having a distinct spectrum of synaptic partners (White et al., 1986). See also supplementary material Fig. S4. IL2 schematic reproduced with permission (Altun et al., 2002-2013). (B) Summary of terminal selector combinatorial codes in head ganglia of *C. elegans*. Colors refer to neurotransmitter identities: green, serotonergic; red, cholinergic; yellow, glutamatergic. Support or blast cells are in gray.

combination with the *cfi-1(ky651)* mutant allele there are strong synergistic, i.e. non-additive, defects in IL2 and URA differentiation (Fig. 8A; supplementary material Fig. S4). We conclude that *unc-86* and *cfi-1* cooperate to control IL2 and URA identity.

DISCUSSION

Two main conclusions can be drawn from the data presented in this paper. First, our data provide general support for the terminal selector concept. Second, our data show that a given transcription

Development (2014) doi:10.1242/dev.099721

factor can operate as a selector of terminal neuron identity in distinct neuronal cell types and that this is achieved through cooperation with distinct co-factors (summarized in Fig. 8B). In other words, individual neuronal cell types use distinct combinatorial codes of terminal selectors, and individual components of the code are reused in distinct combinations in different cell types.

The terminal selector concept was initially proposed based on a relatively small number of *C. elegans* transcription factor mutant phenotypes (Hobert, 2008). In each of these mutant backgrounds, a neuronal cell is born and expresses pan-neuronal features but fails to adopt neuron type-specific identity features. Importantly, terminal differentiation is very broadly affected in terminal selector mutants, such that not only functionally linked features (such as enzymes and transporter in a neurotransmitter synthesis/transport pathway), but also seemingly completely independent differentiation features that have no obvious biochemical connection (e.g. sensory receptors, neuropeptides and ionotropic neurotransmitter receptors) fail to be expressed. That the removal of an individual transcription factor results in such broad defects could not necessarily be assumed since transcriptomic approaches generally show that individual cell types expresses several dozen transcription factors (e.g. Etchberger et al., 2007). This could be interpreted to mean that the identity features of a neuron are regulated in a piecemeal manner, rather than being 'mastered' by a single transcription factor or a small combination thereof (Hobert, 2011). Two major questions raised by the terminal selector concept were how broadly it applies to different cell types in the *C. elegans* nervous system and how it applies to transcription factors expressed in distinct neuron types.

Here, we have shown that the terminal differentiation programs of very distinct neuron types – a cholinergic interneuron (AIA), a serotonergic sensory/motor neuron (NSM) and cholinergic sensory and motor neuron classes (IL2 and URA) – are controlled by distinct combinatorial codes of transcription factors. These factors regulate many distinct identity features of these distinct neuron types, ranging from neuropeptides to neurotransmitter synthesis pathway genes to neurotransmitter receptors and other signaling molecules.

In the case of the cholinergic AIA interneuron, we found that the expression of every tested terminal differentiation marker is affected in *ttx-3* mutants. Since the available AIA marker collection essentially represents a random snapshot of terminal markers that characterize AIA identity, one might extrapolate the regulatory impact of *ttx-3* on each one of these genes to the many hundreds, if not thousands, of genes that are expressed in AIAs, such that *ttx-3* is likely to affect a very large number of them. The estimated very broad effect of *ttx-3* on AIA identity is consistent with what we observed for the cholinergic AIY interneuron, in which *ttx-3* mutation also affects the expression of all known identity features (Altun-Gultekin et al., 2001; Wenick and Hobert, 2004). Even though both neuron types have similar morphologies, are cholinergic, and are directly postsynaptic to various sensory neurons, AIY and AIA have different functions (Hobert et al., 1997; Shinkai et al., 2011; Tomioka et al., 2006), connect to a different spectrum of synaptic partners (White et al., 1986) and express distinct gene batteries. Yet, in both cases, *ttx-3* very broadly affects the differentiation of each neuron type.

The distinct target gene specificities of *ttx-3* in AIA and AIY neurons can be explained by neuron type-specific co-factors and by *ttx-3* acting through distinct *cis*-regulatory motifs. AIY-expressed genes display a characteristic *cis*-regulatory signature that is recognized by a combination of the TTX-3 and CEH-10 homeodomain proteins (Wenick and Hobert, 2004). As we have shown here, AIA-expressed genes share a distinct *cis*-regulatory signature that is composed of two separate motifs located in close proximity, one a TTX-3 binding site and the other a binding site for a presumptive TTX-3 co-factor. This is analogous to the situation in the AIY interneuron class, in which TTX-3 and CEH-10 operate through a bipartite motif (the 'AIY motif') composed of a TTX-3 and a CEH-10 binding site (Wenick and Hobert, 2004). Genes that are expressed in both AIY and AIA neurons (e.g. *cho-1*) contain a modular assembly of both the AIY and AIA *cis*-regulatory signature.

Similar to the *ttx-3*-dependent control of the central cholinergic interneurons AIY and AIA, the mouse LIM homeobox gene *Lhx7* is required for the terminal differentiation of cholinergic striatal interneurons (Lopes et al., 2012). As with other terminal selector transcription factors, *Lhx7* function appears to be continuously required to maintain cholinergic identity. Co-factors that operate together with *Lhx7* are currently not known. *Lhx7* is expressed in many other neurons in the CNS. It will be interesting to determine whether *Lhx7* also operates as a terminal selector in these other neuron types.

ttx-3 activity is not restricted to cholinergic neurons. We find that *ttx-3* is also a key regulator of serotonergic neuron identity. The activity of *ttx-3* in the serotonergic NSM neuron class is, however, distinct from that of AIA and AIY. Whereas the expression of several NSM-expressed effector genes is completely eliminated in *ttx-3* mutants, the expression of some effector genes is only partially affected or not affected at all. In cases in which only partial or no effect was observed, joint removal of another homeobox gene, *unc-86*, resulted in much stronger or complete loss of effector gene expression. Vice versa, the expression of effector genes that are unaffected in expression in *unc-86* mutants is lost in either *ttx-3* mutants or in the *ttx-3; unc-86* double mutant. Taken together, elimination of both of the POU/LIM homeobox genes *unc-86* and *ttx-3* has profound effects on NSM identity, paralleling the profound effect that another POU/LIM homeobox combination (*unc-86* and *mec-3*) has on touch neuron differentiation (Duggan et al., 1998). How *unc-86* and *ttx-3* interact to control NSM differentiation is currently unclear. Both genes are continuously expressed in NSM neurons, but do not regulate the expression of each other. Based on the synergistic nature of the effect of joint *ttx-3* and *unc-86* removal on the expression of some target genes (no or limited effect in single mutants, complete loss in double mutant), we propose that both transcription factors act jointly on common target gene promoters. For some target genes, the loss of one regulatory factor can be completely or partly compensated for by the other regulatory factor; in other cases, such compensation is not possible. *unc-86* and *ttx-3* might therefore not always act in a strict cooperative sense, but rather act independently on target gene promoters. There is already a notable precedent for such a mechanism, as we recently found that a combination of three different transcription factors controls dopaminergic neuron identity. For some target genes, individual transcription factor mutants display very limited effects, but double mutants strongly affect target gene expression (Doitsidou et al., 2013). In the case of NSM, we cannot however rule out the possibility that some genes are exclusively regulated by *unc-86* whereas others are exclusively regulated by *ttx-3*.

Apart from demonstrating *ttx-3* terminal selector function in distinct neuron types, we have also shown here that the POU homeobox gene *unc-86* can similarly act as a terminal selector in distinct neuron types. A role of *unc-86* in the differentiation of serotonergic and glutamatergic touch neurons has been described previously (Desai et al., 1988; Duggan et al., 1998; Sze et al., 2002; Serrano-Saiz et al., 2013). We show here that *unc-86* also controls the terminal differentiation programs of three distinct cholinergic

Development (2014) doi:10.1242/dev.099721

neuron types. Two of these cholinergic neuron types are synaptically connected and form a simple sensory-to-motor circuit (White et al., 1986). The role of *unc-86* in controlling cholinergic IL2 sensory neuron specification is reminiscent of, and might even be homologous to, the function of the POU homeobox gene *acj6* in controlling expression of the cholinergic gene locus in *Drosophila* olfactory neurons (Lee and Salvaterra, 2002). The ARID-type transcription factor *cfi-1* cooperates with *unc-86* to control the cholinergic identity of IL2 and URA neurons. Although neuronal differentiation functions have been reported for the *cfi-1* homolog *dead ringer* (*retained* – FlyBase) in *Drosophila* (Ditch et al., 2005), the functions of vertebrate orthologs (*Arid3* genes) in the nervous system remain to be explored.

MATERIALS AND METHODS

Strains and transgenes

For a list of strains and transgenes and notes on their generation see supplementary material Table S2.

Serotonin antibody staining

Young adult animals were fixed in 4% paraformaldehyde overnight and then treated with 5% β-mercaptoethanol overnight followed by 1000 units/ml collagenase (Sigma-Aldrich) treatment. Rabbit anti-serotonin whole serum (Sigma-Aldrich, S5545) was used at 1:100 dilution. Worms were then washed and incubated with Alexa Fluor 555 donkey anti-rabbit IgG (1:1000; Life Technologies, A-31571).

Cis-regulatory analysis

DNA sequences were subcloned into pPD95.75 expression vector (Addgene). For some smaller constructs, PCR products were directly amplified from subcloned constructs that have the same 3′ end of the promoter sequences. DNAs for injection were PCR amplified to eliminate vector backbone, gel purified and then injected as complex arrays (10 ng/µl) with digested *rol-6(d)* (3 ng/µl) as injection marker, or plasmid mix was directly injected [50 ng/µl together with 100 ng/µl *rol-6(d)*].

Gel shift analysis

Full-length *unc-86* cDNA was cloned into the pET-21b His tag expression vector (EMD Millipore) and transformed into BL21(DE3) pLysS bacteria (Novagen). Protein expression was induced using 1 mM IPTG for 4 hours at 37°C and batch purified using Ni-NTA resin (Qiagen) under denaturing conditions as described (Wenick and Hobert, 2004). TTX-3 was purified and electrophoretic mobility shift assays (EMSAs) were performed as described (Wenick and Hobert, 2004). Probe sequences are listed in supplementary material Table S3.

Acknowledgements

We thank Q. Chen and B. Alarcon for expert assistance in generating transgenic strains, V. Reinke for providing the *ttx-3* fosmid reporter, E. Serrano for *eat-4* reporters, members of the worm community for providing reporter genes and members of the O.H. laboratory for comments on the manuscript.

Competing interests

The authors declare no competing financial interests.

Author contributions

F.Z. and O.H. initiated the study. A.B. and P.G. performed the analysis of the IL2 neurons; A.B. performed analysis of the URA and URB neurons; C.L.-F., M.M. and N.F. undertook the mutational analysis of the serotonergic pathway promoters; J.C.N. and D.A.C.-R. performed and supervised the morphological analysis of the NSM neurons; N.A. and R.S.M. performed and supervised the gel-shift analyses. F.Z. performed all other experiments. O.H. wrote the paper.

Funding

This work was funded by the National Institutes of Health [R01NS039996-05 and R01NS050266-03 to O.H.; R01NS076558 to D.A.C.-R.; and R01NS070644 to R.S.M.]; a March of Dimes Foundation Grant (to D.A.C.-R.); the Spanish Government [SAF2011-26273 to N.F.]; a Marie Curie Career Integration Grant (to N.F.); a VAL i+d Fellowship from Generalitat Valenciana (to C.L.-F.); and a European Research Council Starting Grant (to N.F.). N.F. is a National Alliance for Research on Schizophrenia and Depression (NARSAD) Young Investigator. O.H. is an Investigator of the Howard Hughes Medical Institute. Deposited in PMC for release after 6 months.

Supplementary material

Supplementary material available online at
http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.099721/-/DC1

References

Albertson, D. G. and Thomson, J. N. (1976). The pharynx of Caenorhabditis elegans. *Philos. Trans. R. Soc. B* **275**, 299-325.

Altun, Z. F., Herndon, L. A., Crocker, C., Lints, R. and Hall, D. H. (**eds**) (2002-2013). WormAtlas, http://www.wormatlas.org.

Altun-Gultekin, Z., Andachi, Y., Tsalik, E. L., Pilgrim, D., Kohara, Y. and Hobert, O. (2001). A regulatory cascade of three homeobox genes, ceh-10, ttx-3 and ceh-23, controls cell fate specification of a defined interneuron class in C. elegans. *Development* **128**, 1951-1969.

Aspöck, G., Ruvkun, G. and Bürglin, T. R. (2003). The Caenorhabditis elegans ems class homeobox gene ceh-2 is required for M3 pharynx motoneuron function. *Development* **130**, 3369-3378.

Axäng, C., Rauthan, M., Hall, D. H. and Pilon, M. (2008). Developmental genetics of the C. elegans pharyngeal neurons NSML and NSMR. *BMC Dev. Biol.* **8**, 38.

Baumeister, R., Liu, Y. and Ruvkun, G. (1996). Lineage-specific regulators couple cell lineage asymmetry to the transcription of the Caenorhabditis elegans POU gene unc-86 during neurogenesis. *Genes Dev.* **10**, 1395-1410.

Berger, M. F., Badis, G., Gehrke, A. R., Talukder, S., Philippakis, A. A., Peña-Castillo, L., Alleyne, T. M., Mnaimneh, S., Botvinnik, O. B., Chan, E. T. et al. (2008). Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. *Cell* **133**, 1266-1276.

Bertrand, V. and Hobert, O. (2009). Linking asymmetric cell division to the terminal differentiation program of postmitotic neurons in C. elegans. *Dev. Cell* **16**, 563-575.

Clark, S. G. and Chiu, C. (2003). C. elegans ZAG-1, a Zn-finger-homeodomain protein, regulates axonal development and neuronal differentiation. *Development* **130**, 3781-3794.

Deneris, E. S. and Wyler, S. C. (2012). Serotonergic transcriptional networks and potential importance to mental health. *Nat. Neurosci.* **15**, 519-527.

Desai, C., Garriga, G., McIntire, S. L. and Horvitz, H. R. (1988). A genetic pathway for the development of the Caenorhabditis elegans HSN motor neurons. *Nature* **336**, 638-646.

Ditch, L. M., Shirangi, T., Pitman, J. L., Latham, K. L., Finley, K. D., Edeen, P. T., Taylor, B. J. and McKeown, M. (2005). Drosophila retained/dead ringer is necessary for neuronal pathfinding, female receptivity and repression of fruitless independent male courtship behaviors. *Development* **132**, 155-164.

Doitsidou, M., Flames, N., Topalidou, I., Abe, N., Felton, T., Remesal, L., Popovitchenko, T., Mann, R., Chalfie, M. and Hobert, O. (2013). A combinatorial regulatory signature controls terminal differentiation of the dopaminergic nervous system in C. elegans. *Genes Dev.* **27**, 1391-1405.

Duggan, A., Ma, C. and Chalfie, M. (1998). Regulation of touch receptor differentiation by the Caenorhabditis elegans mec-3 and unc-86 genes. *Development* **125**, 4107-4119.

Etchberger, J. F., Lorch, A., Sleumer, M. C., Zapf, R., Jones, S. J., Marra, M. A., Holt, R. A., Moerman, D. G. and Hobert, O. (2007). The molecular signature and cis-regulatory architecture of a C. elegans gustatory neuron. *Genes Dev.* **21**, 1653-1674.

Finney, M. and Ruvkun, G. (1990). The unc-86 gene product couples cell lineage and cell identity in C. elegans. *Cell* **63**, 895-905.

Flames, N. and Hobert, O. (2009). Gene regulatory logic of dopamine neuron differentiation. *Nature* **458**, 885-889.

Gray, P. A., Fu, H., Luo, P., Zhao, Q., Yu, J., Ferrari, A., Tenzen, T., Yuk, D. I., Tsung, E. F., Cai, Z. et al. (2004). Mouse brain organization revealed through direct genome-scale TF expression analysis. *Science* **306**, 2255-2257.

Greer, E. R., Perez, C. L., Van Gilst, M. R., Lee, B. H. and Ashrafi, K. (2008). Neural and molecular dissection of a C. elegans sensory circuit that regulates fat and feeding. *Cell Metab.* **8**, 118-131.

Harris, G., Korchnak, A., Summers, P., Hapiak, V., Law, W. J., Stein, A. M., Komuniecki, P. and Komuniecki, R. (2011). Dissecting the serotonergic food signal stimulating sensory-mediated aversive behavior in C. elegans. *PLoS ONE* **6**, e21897.

Hirota, J. and Mombaerts, P. (2004). The LIM-homeodomain protein Lhx2 is required for complete development of mouse olfactory sensory neurons. *Proc. Natl. Acad. Sci. USA* **101**, 8751-8755.

Hobert, O. (2008). Regulatory logic of neuronal diversity: terminal selector genes and selector motifs. *Proc. Natl. Acad. Sci. USA* **105**, 20067-20071.

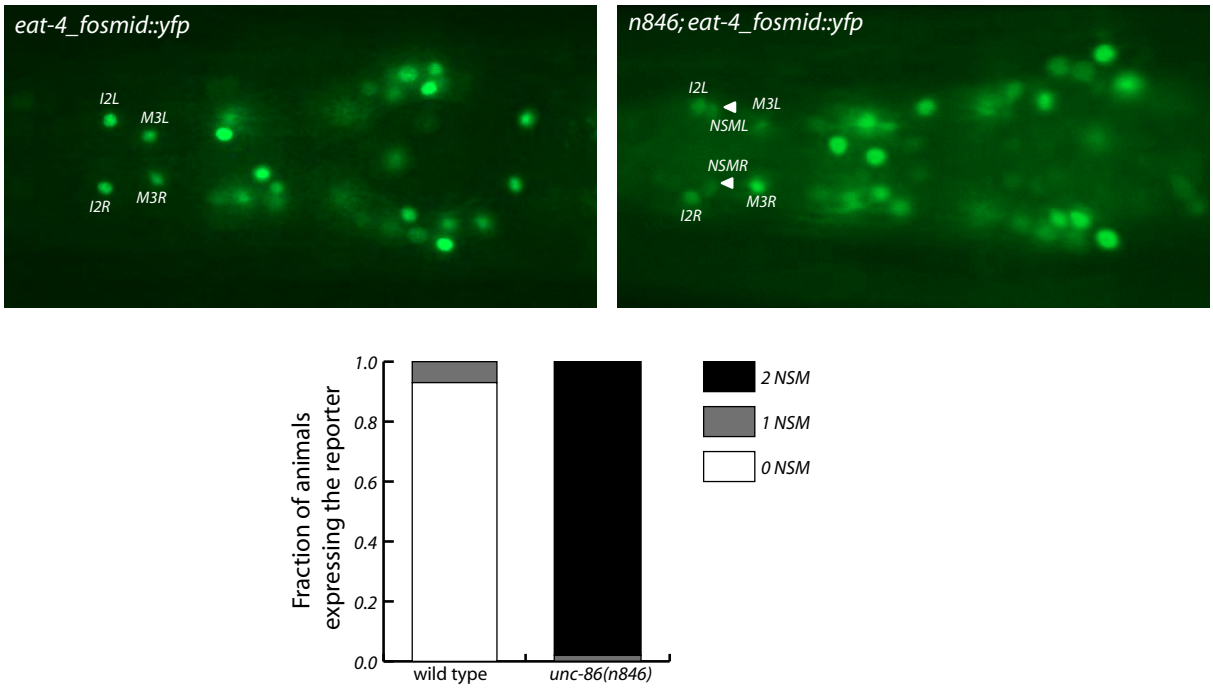
Hobert, O. (2011). Regulation of terminal differentiation programs in the nervous system. *Annu. Rev. Cell Dev. Biol.* **27**, 681-696.

Hobert, O. and Ruvkun, G. (1998). A common theme for LIM homeobox gene function across phylogeny? *Biol. Bull.* **195**, 377-380.

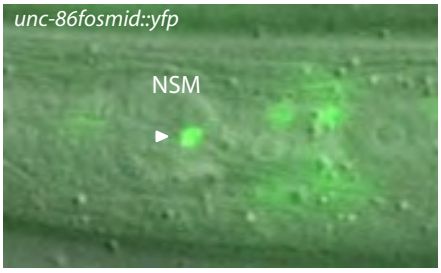
Hobert, O. and Westphal, H. (2000). Functions of LIM-homeobox genes. *Trends Genet.* **16**, 75-83.

Hobert, O., Mori, I., Yamashita, Y., Honda, H., Ohshima, Y., Liu, Y. and Ruvkun, G. (1997). Regulation of interneuron function in the C. elegans thermoregulatory pathway by the ttx-3 LIM homeobox gene. *Neuron* **19**, 345-357.
Horvitz, H. R., Chalfie, M., Trent, C., Sulston, J. E. and Evans, P. D. (1982). Serotonin and octopamine in the nematode Caenorhabditis elegans. *Science* **216**, 1012-1014.
Jafari, G., Xie, Y., Kullyev, A., Liang, B. and Sze, J. Y. (2011). Regulation of extrasynaptic 5-HT by serotonin reuptake transporter function in 5-HT-absorbing neurons underscores adaptation behavior in Caenorhabditis elegans. *J. Neurosci.* **31**, 8948-8957.
Kolterud, A., Alenius, M., Carlsson, L. and Bohm, S. (2004). The Lim homeobox gene Lhx2 is required for olfactory sensory neuron identity. *Development* **131**, 5319-5326.
Lee, M. H. and Salvaterra, P. M. (2002). Abnormal chemosensory jump 6 is a positive transcriptional regulator of the cholinergic gene locus in Drosophila olfactory neurons. *J. Neurosci.* **22**, 5291-5299.
Lee, R. Y., Sawin, E. R., Chalfie, M., Horvitz, H. R. and Avery, L. (1999). EAT-4, a homolog of a mammalian sodium-dependent inorganic phosphate cotransporter, is necessary for glutamatergic neurotransmission in caenorhabditis elegans. *J. Neurosci.* **19**, 159-167.
Lee, H., Choi, M. K., Lee, D., Kim, H. S., Hwang, H., Kim, H., Park, S., Paik, Y. K. and Lee, J. (2012). Nictation, a dispersal behavior of the nematode Caenorhabditis elegans, is regulated by IL2 neurons. *Nat. Neurosci.* **15**, 107-112.
Lein, E. S., Hawrylycz, M. J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A. F., Boguski, M. S., Brockway, K. S., Byrnes, E. J. et al. (2007). Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168-176.
Leung-Hagesteijn, C., Spence, A. M., Stern, B. D., Zhou, Y., Su, M. W., Hedgecock, E. M. and Culotti, J. G. (1992). UNC-5, a transmembrane protein with immunoglobulin and thrombospondin type 1 domains, guides cell and pioneer axon migrations in C. elegans. *Cell* **71**, 289-299.
Liu, C., Maejima, T., Wyler, S. C., Casadesus, G., Herlitze, S. and Deneris, E. S. (2010). Pet-1 is required across different stages of life to regulate serotonergic function. *Nat. Neurosci.* **13**, 1190-1198.
Lopes, R., Verhey van Wijk, N., Neves, G. and Pachnis, V. (2012). Transcription factor LIM homeobox 7 (Lhx7) maintains subtype identity of cholinergic interneurons in the mammalian striatum. *Proc. Natl. Acad. Sci. USA* **109**, 3119-3124.
Mangale, V. S., Hirokawa, K. E., Satyaki, P. R., Gokulchandran, N., Chikbire, S., Subramanian, L., Shetty, A. S., Martynoga, B., Paul, J., Mai, M. V. et al. (2008). Lhx2 selector activity specifies cortical identity and suppresses hippocampal organizer fate. *Science* **319**, 304-309.
Moreno, N., Bachy, I., Rétaux, S. and González, A. (2005). LIM-homeodomain genes as territory markers in the brainstem of adult and developing Xenopus laevis. *J. Comp. Neurol.* **485**, 240-254.
Nathoo, A. N., Moeller, R. A., Westlund, B. A. and Hart, A. C. (2001). Identification of neuropeptide-like protein gene families in Caenorhabditis elegans and other species. *Proc. Natl. Acad. Sci. USA* **98**, 14000-14005.
Nelson, J. C. and Colón-Ramos, D. A. (2013). Serotonergic neurosecretory synapse targeting is controlled by netrin-releasing guidepost neurons in Caenorhabditis elegans. *J. Neurosci.* **33**, 1366-1376.
Ortiz, C. O., Etchberger, J. F., Posy, S. L., Frøkjaer-Jensen, C., Lockery, S., Honig, B. and Hobert, O. (2006). Searching for neuronal left/right asymmetry: genomewide analysis of nematode receptor-type guanylyl cyclases. *Genetics* **173**, 131-149.
Ouellet, J., Li, S. and Roy, R. (2008). Notch signalling is required for both dauer maintenance and recovery in C. elegans. *Development* **135**, 2583-2592.
Peden, E. M. and Barr, M. M. (2005). The KLP-6 kinesin is required for male mating behaviors and polycystin localization in Caenorhabditis elegans. *Curr. Biol.* **15**, 394-404.
Peukert, D., Weber, S., Lumsden, A. and Scholpp, S. (2011). Lhx2 and Lhx9 determine neuronal differentiation and compartition in the caudal forebrain by regulating Wnt signaling. *PLoS Biol.* **9**, e1001218.
Ranganathan, R., Sawin, E. R., Trent, C. and Horvitz, H. R. (2001). Mutations in the Caenorhabditis elegans serotonin reuptake transporter MOD-5 reveal serotonin-dependent and -independent activities of fluoxetine. *J. Neurosci.* **21**, 5871-5884.
Rhee, J. M., Gruber, C. A., Brodie, T. B., Trieu, M. and Turner, E. E. (1998). Highly cooperative homodimerization is a conserved property of neural POU proteins. *J. Biol. Chem.* **273**, 34196-34205.
Röhrig, S. (2000). *Modulation of UNC-86 Activity During Caenorhabditis elegans Neurogenesis*. München, Germany: Herbert Utz Verlag.
Serrano-Saiz, E., Poole, R. J., Felton, T., Zhang, F., De La Cruz, E. and Hobert, O. (2013). Modular control of glutamatergic neuronal identity in C. elegans by distinct homeodomain proteins. *Cell* **155**, 659-673.
Shaham, S. and Bargmann, C. I. (2002). Control of neuronal subtype identity by the C. elegans ARID protein CFI-1. *Genes Dev.* **16**, 972-983.
Shinkai, Y., Yamamoto, Y., Fujiwara, M., Tabata, T., Murayama, T., Hirotsu, T., Ikeda, D. D., Tsunozaki, M., Iino, Y., Bargmann, C. I. et al. (2011). Behavioral choice between conflicting alternatives is regulated by a receptor guanylyl cyclase, GCY-28, and a receptor tyrosine kinase, SCD-2, in AIA interneurons of Caenorhabditis elegans. *J. Neurosci.* **31**, 3007-3015.
Simmons, D. K., Pang, K. and Martindale, M. Q. (2012). Lim homeobox genes in the Ctenophore Mnemiopsis leidyi: the evolution of neural cell type specification. *Evodevo* **3**, 2.
Smidt, M. P. and Burbach, J. P. (2009). Terminal differentiation of mesodiencephalic dopaminergic neurons: the role of Nurr1 and Pitx3. *Adv. Exp. Med. Biol.* **651**, 47-57.
Srivastava, M., Larroux, C., Lu, D. R., Mohanty, K., Chapman, J., Degnan, B. M. and Rokhsar, D. S. (2010). Early evolution of the LIM homeobox gene family. *BMC Biol.* **8**, 4.
Sulston, J. E., Schierenberg, E., White, J. G. and Thomson, J. N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. *Dev. Biol.* **100**, 64-119.
Sze, J. Y., Zhang, S., Li, J. and Ruvkun, G. (2002). The C. elegans POU-domain transcription factor UNC-86 regulates the tph-1 tryptophan hydroxylase gene and neurite outgrowth in specific serotonergic neurons. *Development* **129**, 3901-3911.
Tomioka, M., Adachi, T., Suzuki, H., Kunitomo, H., Schafer, W. R. and Iino, Y. (2006). The insulin/PI 3-kinase pathway regulates salt chemotaxis learning in Caenorhabditis elegans. *Neuron* **51**, 613-625.
Tong, Y. G. and Bürglin, T. R. (2010). Conditions for dye-filling of sensory neurons in Caenorhabditis elegans. *J. Neurosci. Methods* **188**, 58-61.
Treinin, M., Gillo, B., Liebman, L. and Chalfie, M. (1998). Two functionally dependent acetylcholine subunits are encoded in a single Caenorhabditis elegans operon. *Proc. Natl. Acad. Sci. USA* **95**, 15492-15495.
Tursun, B., Cochella, L., Carrera, I. and Hobert, O. (2009). A toolkit and robust pipeline for the generation of fosmid-based reporter genes in C. elegans. *PLoS ONE* **4**, e4625.
Wenick, A. S. and Hobert, O. (2004). Genomic cis-regulatory architecture and trans-acting regulators of a single interneuron-specific gene battery in C. elegans. *Dev. Cell* **6**, 757-770.
White, J. G., Southgate, E., Thomson, J. N. and Brenner, S. (1986). The structure of the nervous system of the nematode Caenorhabditis elegans. *Philos. Trans. R. Soc. B* **314**, 1-340.
Zetterström, R. H., Williams, R., Perlmann, T. and Olson, L. (1996). Cellular expression of the immediate early transcription factors Nurr1 and NGFI-B suggests a gene regulatory role in several brain regions including the nigrostriatal dopamine system. *Brain Res. Mol. Brain Res.* **41**, 111-120.

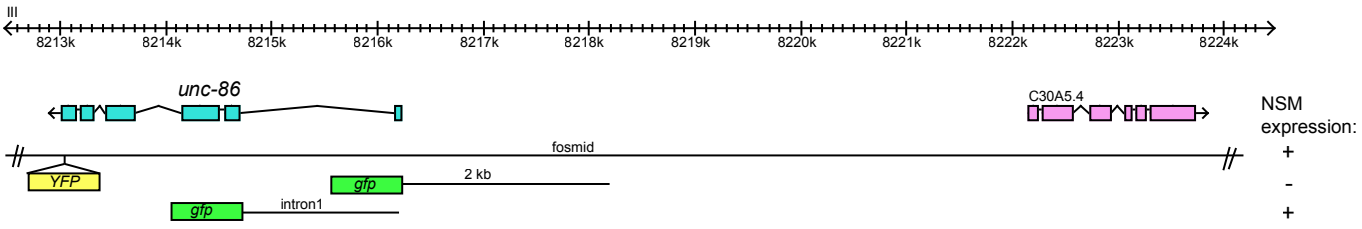
A

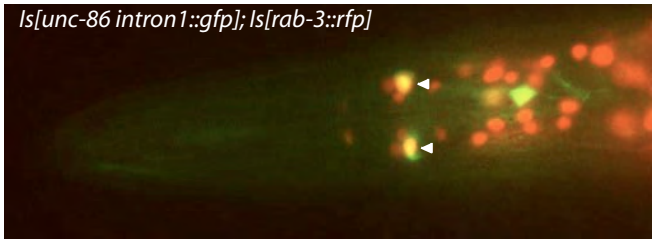


B



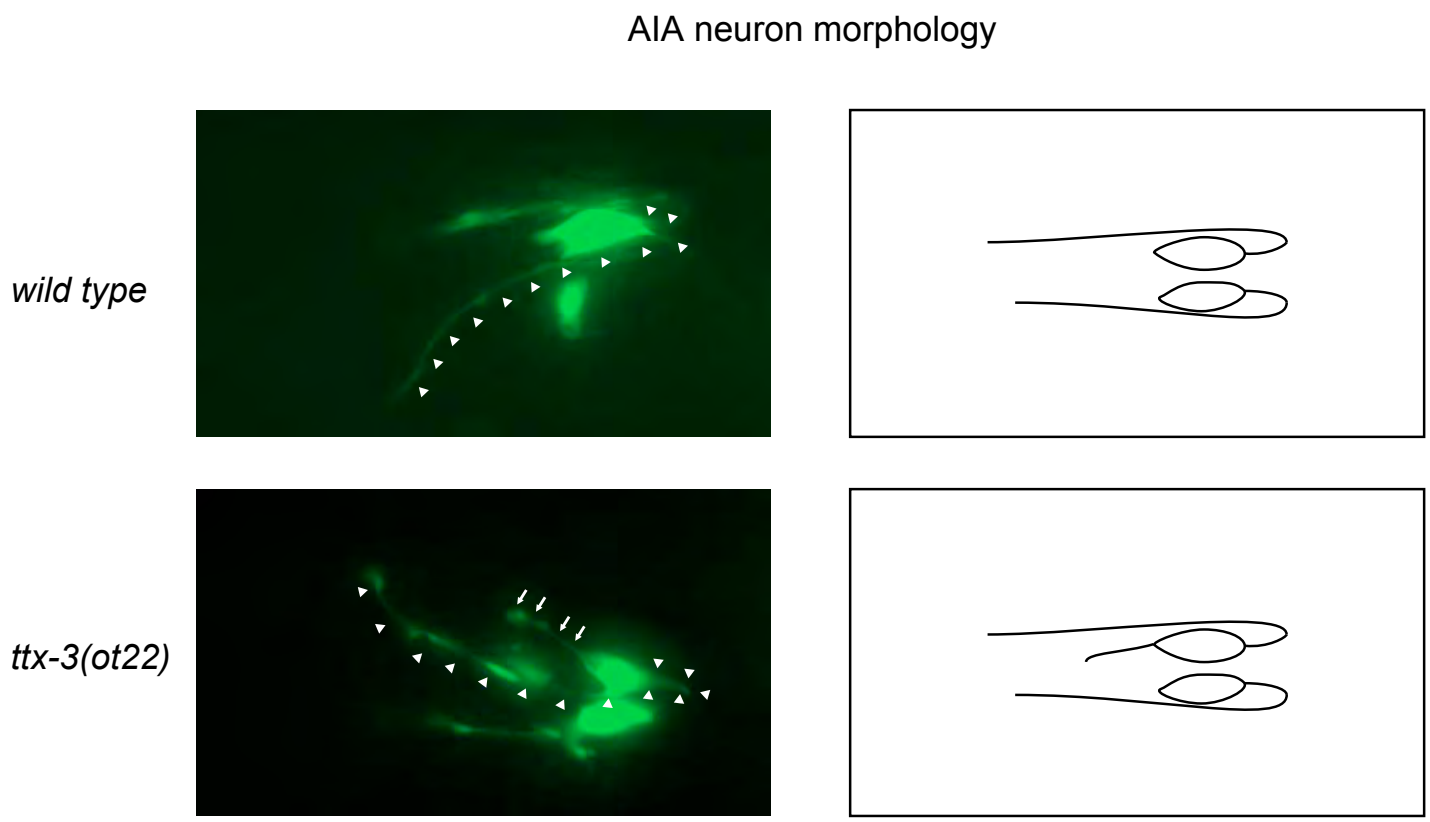
C





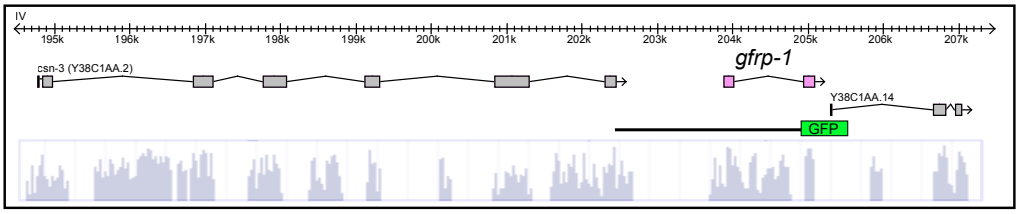
	% NSM on	n
Is[unc-86 intron1::gfp]	93%	28
n846; Is[unc-86 intron1::gfp]	0%	56

Suppl. Figure 1: Analysis of the NSM neurons.
A: A fosmid reporter of the *eat-4* locus was kindly provided by E. Serrano and will be published elsewhere.
B: A *unc-86* fosmid reporter construct is expressed in the NSM neurons of adult animals.
C: A *cis*-regulatory element from the *unc-86* locus drives expression in NSM and this expression depends on *unc-86*.

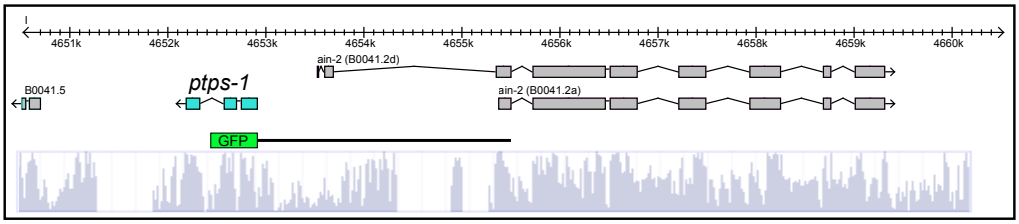


Suppl. Figure 2: AIA morphology in wildtype and *ttx-3(ot22)* mutant animals.
AIA morphology was visualized with *ins-1::gfp (otIs326*). Left panels shows AIA morphology schematically. White triangles in the *gfp* images indicate one of the main axons (normal axon). White arrows indicate the ectopic branching from the cell body in *ttx-3* mutants. Note also the blebbing of the main axon.

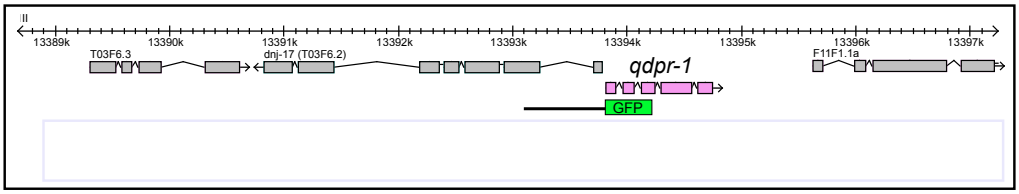
Expression pattern:



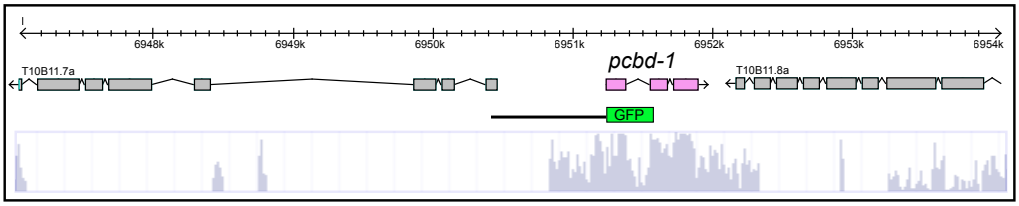
not NSM,
not HSN
a few very dim cells in the head



NSM
HSN
VC4 and VC5
no ADF
more VNC neurons
tail neuron

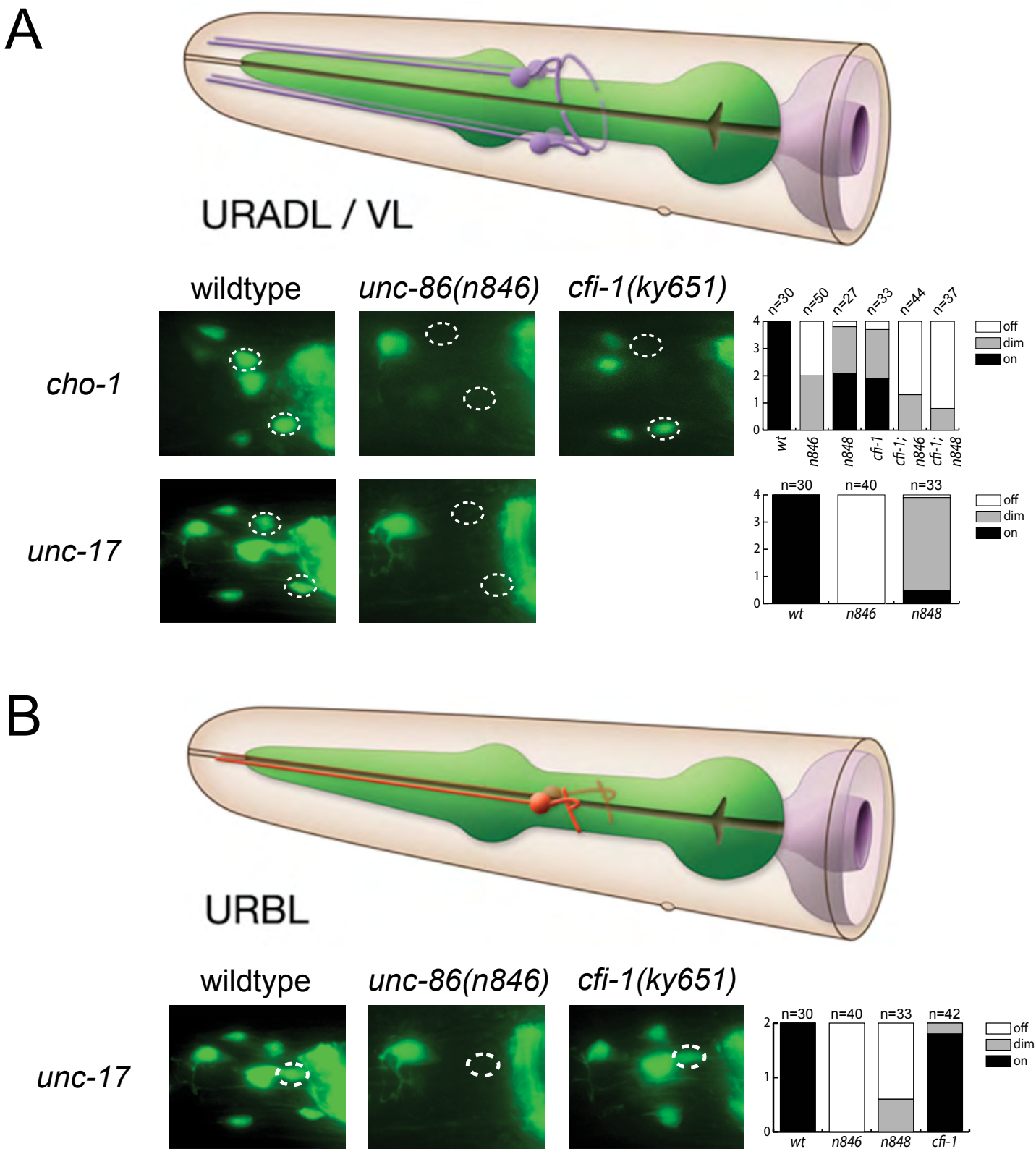


not NSM,
not HSN
several dim cells in the head (ADF?)
tail and head hypodermal cells



not NSM,
not HSN
not ADF
a couple very faint cells in head
no obvious neuronal expresison
tail hypodermal cells
pharyngeal muscle expression

Suppl. Figure 3: BH4 pathway reporters.
Reporter genes and overview of expression pattern of BH4 pathway genes. Expression patterns were observed with multiple lines.



Suppl. Figure 4: *unc-86* controls the identity of the cholinergic URA and URB neurons.
A: *unc-86* and *cfi-1* affect URA identity. Lateral views (anterior to left) are shown. Bar graphs indicate average number of cells expressing *gfp* in the four URA neurons.
B: *unc-86*, but not *cfi-1* affects URB identity as assessed with two different *unc-86* allele, *n846* and *n848*. Lateral views (anterior to left) are shown. Bar graphs indicate average number of cells expressing *gfp* in the two URB neurons.

Table S1. Rescue of NSM and AIA differentiation defects of *ttx-3* mutant animals

	% animals expressing *mgl-1::mcherry* in NSM	% animals expressing *mgl-1::mcherry* in AIA	*n*
wild type	100	100	>100
ttx-3(ot22)	0	0	>100
ttx-3(ot22); Ex[cat-1^{prom}::ttx-3cDNA] line1	92	0	36
ttx-3(ot22); Ex[cat-1^{prom}::ttx-3cDNA] line2	89	0	37
ttx-3(ot22); Ex[ins-1^{prom}::ttx-3cDNA] line1	0	79	38
ttx-3(ot22); Ex[ins-1^{prom}::ttx-3cDNA] line2	0	58	36
ttx-3(ot22); Ex[ins-1^{prom}::ttx-3cDNA] line3	0	79	38

Injection marker: *rol-6(d)*. Note that the *ins-1* promoter is still weakly expressed in *ttx-3* mutants and can hence be used to drive *ttx-3* in *ttx-3* mutants.

Table S2. Strains and transgenes used in this study

Strain/Array	Notes
ttx-3(ot22)	Premature stop before the homeobox (Altun-Gultekin et al., 2001).
unc-86(n846)	A likely null allele resulting in protein loss (Röhrig, 2000). The molecular identity of this strong allele had not been previously described. We sequenced this alleles and it to harbor a G>A splice acceptor site mutation at end of the second intron of the C30H5.7a transcript in the middle of the POU domain (aatacttcagGCGG to aatacttcaaGCGG)
unc-86(n848)	The molecular nature of this temperature-sensitive allele is a GT to AT splice donor site mutation in intron 4 (Röhrig, 2000)
cfi-1(ky651)	A splice acceptor site mutation before the DNA binding domain (Shaham and Bargmann, 2002)
otIs224	*Is[cat-1::gfp]* (Flames and Hobert, 2009)
otIs225	*Is[cat-4::gfp]* (Flames and Hobert, 2009)
otIs226	*Is[bas-1::gfp]* (Flames and Hobert, 2009)
zdIs13	*Is[tph-1::gfp]* (Clark and Chiu, 2003)
wgIs68	*Is[ttx-3fosmid::EGFP-FLAG, unc-119(+)*]. Kindly provided by Valerie Reinke and the ModEncode consortium. Based on fosmid WRM064cD04. The tag was TY1 EGFP 3xFLAG and was added at the C-terminus.
otIs337	*Is[unc-86 fosmid* ^{WRM0612cF07}*::NLS::YFP::H2B; ttx-3::mCherry].* The *unc-86* fosmid reporter was generated bacterial recombineered as previously described (Tursun et al., 2009), fusing an SL2::NLS::YFP::H2B reporter cassette at the C-terminus of *unc-86* in fosmid WRM0612cF07.
vsIs33	*Is[dop-3::dsRed].* Kindly provided by Michael Koelle
otIs317	*Is[mgl-1*$^{long\ prom}$*::mcherry, pha-1].* DNA kindly provided by Kaveh Ashrafi (Greer et al., 2008)
otIs341	*Is[mgl-1*$^{short\ prom}$*::gfp]* -1994 to -1374 bp upstream of ATG
otIs379	*Is[cho-1*^{AIAprom}*::gfp; rol-6(d)]* -3006 to -2642 bp upstream of ATG
otIs326	*Is[ins-1::gfp; rol-6(d)]* -289 bp upstream of ATG
otEx4687	*Ex[glr-2::gfp; rol-6(d)]* -1798 bp upstream of ATG
otEx4886	*Ex[ttx-3 intron7::gfp; rol-6(d)]* GGAAG+intron7+CGTCTACCGATGAAGATG cloned into pPD95.75
otEx5056	*Ex[flp-2::gfp; rol-6(d)]* -2002 bp upstream of ATG
otEx4781	*Ex[mod-5*$^{NSM\ prom::gfp}$*; elt-2::gfp].* First intron of mod-5 cloned into pPD95.75(CACCAGCAGCTGCAAG+ intron1+ CTGAACTCTCC driving GFP)
otEx5280	*Ex[ptps-1::gfp; rol-6(d)]* -2600 bp upstream of ATG
otEx5163	*Ex[nlp-3::gfp; rol-6(d)].* DNA kindly provided by Hart lab (Nathoo et al., 2001)
o*tEx5364*	*Ex[mgl-3::gfp; rol-6(d)].*

	DNA kindly provided by Ashrafi lab (Greer et al., 2008)
otEx5163	*[nlp-13::gfp; rol-6(d)]* -1967 bp upstream of ATG
otEx5055	*[scd-2::gfp; rol-6(d)]* -2045 bp upstream of ATG
otEx5363	*[flr-2::gfp; rol-6(d)]* DNA kindly provided by Takeshi Ishihara
otEx4917	*Ex[unc-86 intron1::gfp; rol-6(d)]* GACGACAACCGCTTCAAAAATGCAACCT+intron1+TTCAACAAC AGTTTATTTGGATCATTCGATGACCC cloned into pPD95.75
otEx4969, otEx4970	2 independent lines of *Ex[cat-1^{prom14}::ttx-3; rol-6(d)]*
otEx5073, otEx5074, otEx5075	3 independent lines of *Ex[ins-1^{457bp_prom}::ttx-3; rol-6(d)] (*-457 bp upstream of ATG)
Ex[gcy-28.d::gfp]	the complete genotype of this array is *Ex[gcy-28.dp::gcy-28.d::GFP, AIA-specific ins-1p::SNB-1::mRFP,rol-6(+)]* (transgene kindly provided by Takeshi Ishihara)
Ex[scd-2::gfp]	*Ex[scd-2::gfp]*: the complete genotype of this array is *Ex[scd-2p::scd-2::GFP, AIA-specific ins-1p::mRFP, lin-44::gfp]* (transgene kindly provided by Takeshi Ishihara
olaEx1446	*Ex* [mod-5p::egfp (2ng/ul)/unc-122p::gfp (20ng/ul)]
olaEx1485	*Ex[*flp-4p::egfp (30ng/ul)/unc-122p::DSRED (20ng/ul)
nuIs9	*Is[unc-5::gfp]* (transgene kindly provided by Josh Kaplan)
otEx2310	*Ex[gcy-19::gfp; unc-122::gfp]* (Ortiz et al., 2006)
lqIs3	*Is[osm-6::gfp]* (transgene kindly provided by Erik Lundquist)
qIs56	*Is[lag-2::gfp]* (transgene kindly provided by Judith Kimble)
vsIs48	*Is[unc-17::gfp]* (transgene kindly provided by Michael Koelle)
otIs323	*Is[cho-1_fosmid::gfp; elt-2::dsRed]* (transgene kindly provided by Paschalis Kratsios)

Table S3. Probe sequences for gel shift analysis

Probe	Sequence
cho-1 wt:	5’tacacacacatcgaaatatgaatcttctcttaaaaagaaggttgtccaattagtttcccctattcaGCTTTCGTTCGTCGCCT
cho-1 TAAT del	5’tacacacacatcgaaatatgaatcttctcttaaaaagaaggttgtccagtttcccctattcaGCTTTCGTTCGTCGCCT
mgl-1 wt	5’gtttccatactcatagtgctcattagaatagcacggatcgtgtttcgcctctcgccttgttaaccgaatctgccGCTTTCGTTCGTCGCCT
mgl-1 TAAT del	5’gtttccatactcatagtgctcgaatagcacggatcgtgtttcgcctctcgccttgttaaccgaatctgccGCTTTCGTTCGTCGCCT
bas-1 wt	5’cccaacaccacattattcatgtatttcctccaaaccactgaaccatctcattctcaaaccagtttctatccgtttgtttgcattcaattaaatttttGCTTTCGTTCGTCGCCT
bas-1 HD mut	5’cccaacaccacgttattcatgtatttcctccaaaccactgaaccatctcattctcaaaccagtttctatccgtttgtttgcattcagttgaattttt GCTTTCGTTCGTCGCCT
bas-1 POU mut	5’cccaacaccacattattcccgtatttcctccaaaccactgaaccatctcattctcaaaccagtttctatccgtttgtttgccctcaattaaatttttt GCTTTCGTTCGTCGCCT
tph-1 wt	5’tctttgtttgcgcataataaaacaatcaatcaacacagcaaagacccctctcaacctcatttcatgattttctttGCTTTCGTTCGTCGCCT
tph-1 HD mut	5’tctttgtttgcgcatagtaaaacaatcaatcaacacagcaaagacccctctcaacctcatttcatgattttctttGCTTTCGTTCGTCGCCT
tph-1 POU mut	5’tctttgtttgTgTataCCaCaacaaGcGatcaacacagcaaagacccctctcaacctcatttcCCgattttctttGCTTTCGTTCGTCGCCT