2010

# Mechanisms for Maintaining Cell Identity in C.elegans Olfactory Neurons

Bluma J. Lesch


# MECHANISMS FOR MAINTAINING CELL IDENTITY IN *C. ELEGANS* OLFACTORY NEURONS

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Bluma J. Lesch

June 2010

# MECHANISMS FOR MAINTAINING CELL IDENTITY IN *C. ELEGANS* OLFACTORY NEURONS

Bluma J. Lesch, Ph.D.
The Rockefeller University 2010

Maintenance of cell identity is a complex process that depends on developmentally determined transcriptional states and on environmental input. In neurons, which are both highly differentiated and highly sensitive to external stimuli, maintenance of identity is especially challenging. In this thesis, I describe the isolation and characterization of several genes involved in maintaining the identities of two olfactory neuron subtypes, $AWC^{ON}$ and $AWC^{OFF}$, in the nematode *Caenorhabditis elegans*.

The $AWC^{ON}$ and $AWC^{OFF}$ identities are specified by a stochastic decision during embryogenesis, but several of the genes involved in this embryonic decision are subsequently downregulated. Additional mechanisms must therefore act to maintain the $AWC^{ON}$ and $AWC^{OFF}$ cell fates. I cloned and characterized a transcription factor, *nsy-7*, that was required to maintain expression of the $AWC^{ON}$ marker *str-2*. *nsy-7* mutants also misexpressed the $AWC^{OFF}$ marker *srsx-3* in both AWCs. The chemotaxis phenotype of these mutants indicated that their defects in *str-2* and *srsx-3* expression corresponded to a more general change in cell identity from $AWC^{ON}$ to $AWC^{OFF}$. *nsy-7* expression was restricted to $AWC^{ON}$ by the signaling pathway that controls the initial fate decision. I collaborated with the Bulyk lab to discover a specific seven-nucleotide DNA sequence bound by the NSY-7 protein. This sequence was present in the *srsx-3* promoter and required for NSY-7-mediated repression

of *srsx-3* in AWC$^{ON}$. Using this sequence, I was also able to predict new transcriptional targets of NSY-7.

In a screen to identify additional genes affecting maintenance of *str-2* and *srsx-3* expression, I isolated nineteen mutants, including the uncharacterized conserved transcription factor *hmbx-1* and components of the TGFβ pathway. I showed that TGFβ signaling is required continuously in adults to maintain *srsx-3* expression and depends on dauer pheromone, an external sensory cue. I then identified new AWC$^{ON}$- and AWC$^{OFF}$-specific markers and examined the extent to which changes in expression of *str-2* or *srsx-3* correlate with larger-scale changes in gene expression in the two AWC neurons. Together, my results indicate that several interlocking genetic pathways combine to maintain the AWC$^{ON}$ and AWC$^{OFF}$ cell identities, including several factors not previously known to be involved in this process.

To my parents,

who I love and admire very much

# Acknowledgments

I am extraordinarily grateful to my advisor, Cori Bargmann, for her patience and guidance during the last four years. She is always available and never intrusive, and sets a high standard for the practice of science that has inspired me and will continue to do so.

My committee members, David Allis, Ali Brivanlou, Michael Young, and Jennifer Zallen, have been extremely supportive. Their advice has allowed me to approach my work from new directions, and I regret that I have not had the chance to follow up on more of their suggestions. I would also like to thank Fred Cross, who was a valuable member of my committee during my first year and whose genetics seminars were fun, interesting, and extremely useful for my graduate education. Mary Beth Hatten was an excellent rotation advisor and continued to be involved and supportive after I had started my thesis project. I am also extremely grateful to Claude Desplan for agreeing to serve on my examination committee, and for very useful comments and questions about *nsy-7*.

Andrew Gehrke and Martha Bulyk made our collaboration on the protein binding microarray experiments easy and enjoyable.

Scott Dewell at the Rockefeller Genomics Resource Center was indispensable in making sense of the Solexa whole-genome sequencing data. Alison North and Shiva Bhuvanendran at the Rockefeller BioImaging Resource Center were endless sources of knowledge about confocal microscopy.

Francesca Spagnoli provided friendly conversations and advice along with protocols, and generously lent me equipment and reagents.

Christian Woods is the stable core at the center of daily life in the Bargmann lab. His good nature and infinite patience make administrative hassles miraculously disappear on a regular basis. Hernan Jaramillo is also wonderfully good-natured, patient, and capable, and I greatly appreciate all the work that went in to freezing the hundreds of strains I have given him. Holly Hunnicutt tears through paperwork with seeming effortlessness, and her energy and humor were an integral part of my daily experience.

Mike Chiorazzi and Diana McKeage were responsible for many enjoyable conversations and for hours of screening, mapping, and picking worms. Makoto Tsunozaki, Sarah Bauer Huang, Sreekanth Chalasani, Patrick McGrath, Tapan Maniar, and Jennifer Garrison were reliable sources of advice, rambling conversation, and general lab knowledge. Andres Bendesky, Sreekanth Chalasani, Evan Macosko, Evan Feinberg, Greg Lee, Andrew Gordus, and Yasunori Saheki made coming in to the lab a daily adventure, producing music and conversation that was more amusing, though less harmonious, than most commercially available entertainment.

I would also like to thank Yael Levin and Keira Cohen for their support at various points throughout graduate school. Chelly has been, as always, wonderfully patient and sympathetic, and continues to make me think in new and interesting ways. Charlie is an interested audience and a friend as much as a sibling. My parents have been and continue to be inspiring, encouraging, and always willing to listen.

# Table of Contents

# List of Figures

# List of Tables

# Chapter 1

## Introduction

Much work has been done to elucidate the mechanisms required for specification of neuronal subtype identity. Known developmental networks require elaborate intercellular signaling and finely-tuned transcriptional responses, and result in a highly complex, multifunctional nervous system. The intricacies of this developmental system, however, are wasted unless these decisions can be maintained, and maintenance of this complex, multilayered system is not trivial. Neurons are highly differentiated, long-lived cells. In most cases, neuronal identity is specified early in life and maintained until death amid changing cellular and environmental contexts (Muotri and Gage 2006; Miller and Kaplan 2007). Only occasionally is the combination of factors initially responsible for the cell's differentiation retained. Indeed, external stimuli are frequently required to sustain the state initiated by a transient combination of developmental regulators (Spitzer 2006). Thus, the mechanisms involved in sustaining neuronal identity include multiple cell-intrinsic and extrinsic factors that operate over a long period of time, and these mechanisms must allow the neuron to retain its core characteristics while also endowing it with the flexibility to respond to novel stimuli. The goal of this thesis was to identify the factors involved in maintaining neuronal identity in two specific neuronal cell types, and to determine how these factors carry out their function.

**Transcription factors involved in cell fate choice may be continuously expressed**

In many cases, information about the final differentiated state of the neuron is determined early in development and passed down through several cell divisions to the postmitotic differentiated neuron. The generation of multiple lineages from neural precursor cells is frequently timing-dependent, so that the identity or quantity of the factors inherited by an early lineage differs from that inherited by later-born cells derived from the same precursor (Decembrini et al. 2006; Zhu et al. 2006; Maurange et al. 2008). This information must then be retained in subsequent generations, so that the daughter cell maintains the identity it has inherited. Hence, the source of cell fate information is temporally sensitive, and a daughter cell's suppression of the continuing temporal progression of the mother cell is essential. In the *Drosophila* neuroblast, the decision to exit the cell cycle is specifically dependent on the conserved homeodomain gene Prospero, whose mammalian homologs are also required for appropriate differentiation of neurons in the forebrain and retina (Wigle et al. 1999; Dyer et al. 2003; Choksi et al. 2006). Ganglion mother cells (GMCs) inherit Prospero, after which they divide only once to produce terminally differentiating neurons or glia of many different types. A GMC defective for Prospero reverts to the neuroblast state (Choksi et al. 2006). The transition to a generalized differentiated GMC state is separate from the decision about what the final differentiated fate will be, and is controlled by a shared transcription factor in multiple neural cell types. With this decision, the GMC begins to transition from a state where it accepts and responds to differentiation cues to a final state where it must reinforce the decisions it has already made.

After this transition, some cell types may sustain expression of the transcription factors that defined their terminal identity. Frequently, these transcription factors are autoregulatory, ensuring their own continued transcription as well as that of their target genes. In *C. elegans*, the homeodomain transcription factor MEC-3 is required for differentiation of the mechanosensory neurons; it continues to be expressed in those neurons, and this continued expression is dependent on its regulation of its own promoter (Xue et al. 1992). Likewise, the ASE chemosensory neurons in *C. elegans* require the transcription factor CHE-1 for both their initial differentiation and later maintenance (Etchberger et al. 2009). A similar mechanism maintains the commitment of pro-B cells in the mammalian immune system, where the transcription factor Pax5 is required both for the original commitment decision and for retention of the committed state. Pro-B cells that lose Pax5 expression revert to a multipotential hematopoetic progenitor state (Nutt et al. 1999; Rolink et al. 1999; Mikkola et al. 2002).

**Dedicated mechanisms maintain cell fate in the absence of differentiation cues**

Frequently, differentiation of a particular cell type is determined not by a single transcription factor whose ongoing expression can sustain the identity of the cell, but by a combination of factors whose expression may change over time. For example, the generation of motor neuron subtypes in the developing spinal cord operates by a combinatorial mechanism, as does control of body segment identity in *Drosophila* (Shirasaki and Pfaff 2002; Peel et al. 2005). When this is the case, specific mechanisms must stabilize the transcriptional state of the differentiated cell as the signals responsible for its differentiation change.

Several general mechanisms are known to operate in this context. These systems frequently involve chromatin modification and remodeling of nucleosomes surrounding target genes to stabilize their active or inactive transcriptional states. The archetypal examples of this type of mechanism are the Polycomb group (PcG) and Trithorax group (TrxG) complexes, which were initially identified for their role in stabilization of homeotic gene expression in *Drosophila* development (Lewis 1978; Busturia and Morata 1988). Since then, they have been found to operate in multiple contexts to maintain developmentally-induced transcriptional states, most frequently in relation to HOX gene-induced differentiation. Consistent with this role, transcription factors are the most strongly conserved class of PcG targets (Schuettengruber et al. 2007). PcG complexes generally act as repressors of target gene expression, while TrxG complexes antagonize PcG to permit expression. Both PcG and TrxG complexes modify chromatin marks at or near sites of regulation: some members of the complex methylate lysines on histone H3, and others carry out the nucleosome remodeling associated with these histone marks. It is assumed, but not fully demonstrated, that these chromatin alterations mediate the repressive or activating functions of PcG and TrxG complexes (Ringrose and Paro 2004; Schwartz and Pirrotta 2007).

Another general mechanism for mediating long-term repression is methylation of cytosines at CpG dinucleotides in the promoter regions of target genes. DNA methylation is essential in mammals for stable, large-scale gene repression in X-chromosome inactivation and genomic imprinting (Robertson 2005; Bienvenu and Chelly 2006; Ideraabdullah et al. 2008). In addition, some specific methylated sites are bound by methyl-binding domain (MBD) proteins,

which in turn recruit histone deacetylases, histone methyltransferases, and nucleosome remodeling machinery to induce an inactive chromatin state (Nan et al. 1998; Ballestar and Wolffe 2001; Fuks et al. 2003; Bienvenu and Chelly 2006). In general, MBD proteins recognize methylated CpGs but are otherwise not sequence-selective; the mechanism by which MBD proteins select target genes may involve sequence-specific DNA binding factors, but is not fully understood (Ballestar and Wolffe 2001; Roloff et al. 2003). The founding member of the MBD family, MeCP2, is a transcriptional repressor that is highly expressed in the mammalian brain, and defects in the MeCP2 gene are associated with specific neurodevelopmental defects. In humans, MeCP2 mutations lead to Rett Syndrome, in which apparently normal girls progressively lose neural function starting between six and eighteen months after birth (Dragich et al. 2000; Bienvenu and Chelly 2006). Mice lacking MeCP2 likewise show postnatal reversals of neural development (Chen et al. 2001). These late deficits, involving loss of neural functions that had already been acquired, indicate a role for MeCP2 in maintenance rather than establishment of these functions (Kishi and Macklis 2004).

Interestingly, it is becoming evident that neither PcG binding nor DNA methylation is as stable as was once assumed. Both of these systems are thought to mediate a long-standing transcriptional state at target genes, and therefore were presumed to stably associate with and inactivate their targets. However, PcG complex proteins have been shown to exchange rapidly at their target binding sites on a time scale of minutes (Ficz et al. 2005). Moreover, PcG-mediated repression can be overcome by high levels of morphogen signaling: imaginal disc identity in *Drosophila* is ordinarily maintained by PcG-based

repression, but Decapentaplegic/TGFβ and wingless can induce transdetermination in imaginal discs (Maves and Schubiger 1998; Klebes et al. 2005). Similarly, methylation patterns in mammalian neurons, once thought to be static, can be altered by membrane depolarization. In cultured neurons, membrane depolarization alters the distribution of DNA methylation and also the distribution of the methyl-binding transcriptional regulator, MeCP2, associated with changes in chromatin state (Chen et al. 2003; Martinowich et al. 2003). MeCP2 itself, once thought to function universally as a repressor of transcription, has recently been shown to directly activate a large number of genes (Chahrour et al. 2008). Thus, transcriptional regulators traditionally thought to mediate unchanging transcriptional states may be actually be malleable and, in the case of DNA methylation, responsive to acute neural stimuli.

**External inputs modify neuronal fate**

In many neurons, the transcriptional state induced at the time of differentiation is not sufficient to sustain the identity and function of the cell. Instead, activity-dependent mechanisms reinforce the identity of the neuron and its role in the network in which it functions. The activity required may be spontaneous, or may be induced by external stimulation from sensory or synaptic input. Activity-dependent mechanisms are also required for the maturation of adult-born neurons (Spitzer 2006). Most neurons continue to receive input throughout their lives and to respond by altering their transcriptional state, their sensitivity to stimulation, and the strength and structure of their synaptic connections (Greer and Greenberg 2008). Specific

mechanisms must be invoked that allow them to sustain their initial identity throughout this time.

Neuronal activity is integral to the developmental program of neuronal subtypes in multiple species. Even before synapse formation, calcium transients can be detected in neuronal precursors. Spontaneous calcium transients have been detected in *Xenopus*, chick, mouse, and rat embryonic neurons and are believed to regulate intrinsic excitability and neurotransmitter expression in these cells (Gu et al. 1994; Owens and Kriegstein 1998; Carey and Matsumoto 1999; Ashworth and Bolsover 2002). In *C. elegans*, an embryonic gap junction network is thought to transmit calcium waves between neurons to determine the specification of neuronal subtypes (Chuang et al. 2007). Global calcium increases in developing mouse neurons also activate a transcriptional program that regulates dendritic growth (Redmond et al. 2002). In general, these calcium signals interact with developmentally-determined transcription factors to influence the identity of the neuron and its relationship with neighboring cells (Spitzer 2006). Activity therefore plays a major role in determination of neural characteristics during development.

After synapses have formed, spontaneous, synchronized waves of activity are necessary for the generation of coordinated circuits in mammalian neural networks (Ben-Ari et al. 1989; Garaschuk et al. 2000; Stellwagen and Shatz 2002; Corlew et al. 2004; Torborg et al. 2005). Here, too, activity interacts with the transcriptional program already present in any given neuron. In postnatal mice, visual experience dramatically alters transcription in the visual cortex. However, these effects are strongly age-dependent, indicating that the effects of activity are closely tied to the developmental state of the neurons involved (Majdan and

Shatz 2006). It is therefore difficult to fully separate developmental, genetically-programmed effects on the state of a neuron from activity-dependent inputs. Instead, activity and developmental programming interact in the immature neuron to coordinate the final stages of its differentiation.

**Adult neurons modify transcription in response to activity**

As the organism matures, activity continues to regulate neuronal transcription. Calcium influx through L-type voltage-sensitive calcium channels, NMDA receptors, and in some cases AMPA receptors initiates signaling cascades that result in activation of transcriptional regulators such as CREB, MEF2, MeCP2, and NPAS4 (Sheng et al. 1991; Deisseroth et al. 1996; Chen et al. 2003; Flavell et al. 2006; Lin et al. 2008). These transcription factors activate or repress transcription of target genes that modulate neural excitability, synapse formation or elimination, and the balance of excitatory and inhibitory influences on the cell's synaptic partners (Greer and Greenberg 2008). Thus there is an extensive transcriptional response to neural activity in the adult as well as the developing nervous system.

One function of adult activity-dependent transcription is to facilitate learning and memory. In one mouse model of neurodegenerative disease, memory impairment associated with loss of neurons was reduced in an enriched environment, and this reduced impairment was associated with rapid remodeling of chromatin in remaining neurons. Histone acetylation was especially important in this context, as the addition of histone deacetylase (HDAC) inhibitors further increased memory recovery (Fischer et al. 2007). This result implies that reactivation of repressed genes allows surviving neurons to

replace those lost in a complex network. At the level of the network, these cells take on a new role and are able to restore the function of the system as a whole.

Ordinarily, however, the transcriptional changes induced by neural activity are transient. For example, transcription from the *Bdnf* promoter IV, a major target of activity-dependent transcription that is regulated by MeCP2 and CREB, returns to baseline within hours of activation (Greer and Greenberg 2008). Most known functions of MeCP2, CREB, MEF2, and other transcriptional regulators are thought to involve transient transcriptional changes in response to activity, not stabilization of previously established transcriptional states. Stable baseline levels of many genes regulated by activity, including BDNF, are required for proper neural development. The mechanisms that stabilize transcription of activity-responsive genes are therefore important to understand in greater detail.

**Establishment and maintenance of neuronal structure also require separate mechanisms**

A neuron's structure and morphology constitutes a major aspect of its identity. Neuronal morphology in adult animals is also poised between the competing influences of activity-dependent remodeling and maintenance of structural features. Specific cell-intrinsic and extrinsic mechanisms are required to maintain dendritic branching, axon integrity, neuronal polarity, and individual synaptic connections. As is the case for maintenance of gene expression, the molecular mechanisms responsible for structural maintenance differ from those required for its establishment. In *C. elegans*, maintenance mechanisms involve a set of transmembrane molecules that interact with the

extracellular matrix or with neighboring cells to resist displacement of neurons or neuronal processes by mechanical stress (Benard and Hobert 2009). Maintenance of neuron structure and position represents a slightly different problem from maintenance of the cell's transcriptional state, but it is interesting that, in both cases, the mechanisms controlling maintenance and establishment are genetically separable.

**Chemosensory receptor expression is related to olfactory neuron identity**

In flies and mammals, the decision to express a particular olfactory receptor (OR) plays a key role in the specification of an olfactory neuron's identity. In mammals, each olfactory neuron expresses one allele of a single OR gene. The choice is stochastic, and the mechanism that selects one specific gene from the ~1300 encoded in the genome and prevents expression of others is an area of active research (Buck and Axel 1991; Chess et al. 1994; Serizawa et al. 2000; Fuss and Ray 2009). It is likely that exclusive expression involves negative feedback from the receptor protein itself, as expression of a nonfunctional OR coding region permits expression of a second OR gene (Serizawa et al. 2003; Lewcock and Reed 2004). In mice, the choice of receptor is also partly regional or lineage-dependent, as different classes of ORs tend to be expressed in different zones within the olfactory epithelium (Ressler et al. 1993; Vassar et al. 1993; Miyamichi et al. 2005; Bozza et al. 2009). Supporting the contribution of regional lineage influences, mutations in closely-spaced O/E (Olf1/Early B-cell factor) and homeodomain-type regulatory motifs in OR promoters have been shown to alter the zone of expression for subsets of ORs, and the homeodomain proteins *Lhx2* and *Emx2* are required for expression of some but not all OR classes (Hoppe

et al. 2003; Hoppe et al. 2006; Michaloski et al. 2006; Hirota et al. 2007; McIntyre et al. 2008). However, OR specification is also flexible: immature olfactory neurons appear to be capable of switching from expression of one OR to another within the same zone. Following this flexible period, expression of a single OR is fixed in a given neuron (Shykind et al. 2004). Stochastic choice of receptor expression is therefore biased by cell-intrinsic transcription and stabilized late during the olfactory neuron's differentiation; once made, the choice is stable and essential for the continued function of the neuron.

Most *Drosophila* chemosensory neurons also express a single cell-type specific chemoreceptor, along with the universal co-receptor OR83b (Vosshall et al. 2000; Larsson et al. 2004; Neuhaus et al. 2005). However, a few classes of neuron express two receptor genes along with OR83b, and selection of the OR gene expressed is not believed to involve negative feedback. Unlike mice, flies do not repress transgenic *Or*-promoter constructs in the presence of the endogenous OR protein, and neurons expressing a nonfunctional OR gene do not respond to odor, suggesting that they do not select an alternative OR locus (Dobritsa et al. 2003; Fishilevich and Vosshall 2005; Goldman et al. 2005). Consistent with this lack of negative feedback, the decision to express a particular receptor in *Drosophila* is strongly lineage-based. The neurons in a given functional class of sensory sensillum always express the same OR genes (Dobritsa et al. 2003; Elmore et al. 2003; Hallem et al. 2004; Couto et al. 2005; Fishilevich and Vosshall 2005; Goldman et al. 2005). Multiple cis-regulatory elements and several transcription factors control expression of subsets of *Drosophila* OR genes, implying that there may be a combinatorial code for

sensillum-specific OR gene expression (Ray et al. 2007; Ray et al. 2008; Tichy et al. 2008).

**Chemoreceptor expression in *C. elegans***

In contrast to the fly and mammalian systems, the 32 chemosensory neurons in *C. elegans* express multiple rhodopsin-like GPCRs and regulate expression of many of these genes in a dynamic fashion; expression of a single receptor gene is therefore less closely tied to the identity of the neuron than it is in flies and mammals (Troemel et al. 1995; Peckol et al. 2001; Nolan et al. 2002; van der Linden et al. 2007; van der Linden et al. 2008). The expression of individual chemoreceptors in olfactory neurons can be altered by external cues and depends on the activity of genes required for odor sensation in these neurons. For example, the TAX-2/TAX-4 cGMP-gated cation channel is required for expression of the receptor gene *srd-1* in the ASI chemosensory neurons, and the guanylate cyclases ODR-1 and DAF-11 are required for expression of *str-1* in AWB and for maintenance of *str-2* expression in the $AWC^{ON}$ neuron (Troemel et al. 1999; Peckol et al. 2001; van der Linden et al. 2008). ODR-1 and DAF-11 are thought to be regulated by GPCR activity in response to odor; production of cGMP from the guanylate cyclases opens the TAX-2/TAX-4 channel, allowing calcium influx (Bargmann et al. 1993; Coburn and Bargmann 1996; Komatsu et al. 1996; Roayaie et al. 1998; L'Etoile and Bargmann 2000; Fujiwara et al. 2002). Activity-dependent regulation by ODR-1 and the TAX-2/TAX-4 channel is also involved in remodeling of the sensory cilia in the AWB neuron, possibly altering the cell's ability to response to odor stimulus (Mukhopadhyay et al. 2008).

A TGFβ signal encoded by the *daf-7* gene also alters chemoreceptor expression in a stimulus-dependent manner. DAF-7 is required to prevent entry into the dauer diapause state, and its expression is inhibited by dauer pheromone, a mixture of structurally-related chemicals constitutively released by *C. elegans* that induces larvae to enter the dauer state when population density is high (Butcher et al. 2007; Butcher et al. 2008; Butcher et al. 2009). Concentrations of pheromone too low to trigger dauer entry can alter chemosensory gene expression in the ASI and ASH neurons, and this activity is dependent on the *daf-7*/TGFβ signaling pathway (Nolan et al. 2002). EGL-4, a cGMP-dependent kinase, acts upstream of the transcriptional regulator *daf-3* in the TGFβ pathway and downstream of the *odr-1* and *daf-11* guanylate cyclases, linking these two pathways (Daniels et al. 2000; L'Etoile et al. 2002; van der Linden et al. 2008). *egl-4* also acts in parallel to the *kin-29* Ser/Thr kinase to regulate transcription in response to sensory input via the histone deacetylase *hda-4* and the transcription factor *mef-2* (Lanjuin and Sengupta 2002; van der Linden et al. 2007; van der Linden et al. 2008). However, cGMP signaling, TGFβ signaling, and *kin-29* activity regulate only a subset of receptors within a single neuron type, and expression of the same receptor can respond differently to these signals in different classes of neurons (Peckol et al. 2001; Nolan et al. 2002; van der Linden et al. 2007). Because multiple receptors are expressed in each neuron, altered expression of a subset of these genes can leave the chemosensory function of the neuron intact. The expression of chemosensory genes in *C. elegans* is therefore partially decoupled from the function of the neurons themselves.

The information specified during neural subtype differentiation can be encoded in a stable fashion even when it is not reflected in the expression of marker genes. For example, the two morphologically symmetric gustatory neurons ASEL and ASER are functionally distinct and express different subsets of genes in adults (Chang et al. 2003). ASEL preferentially senses sodium and is activated by increases in stimulus concentration, while ASER preferentially senses chloride and potassium and is activated by decreases in stimulus concentration (Pierce-Shimomura et al. 2001; Suzuki et al. 2008). The ASEL or ASER fates are specified by signals present at the four-cell stage of embryogenesis; each ASE neuron descends from a different precursor present at this stage (Poole and Hobert 2006). The decision to become ASEL or ASER is executed by a regulatory loop involving transcription factors and microRNAs, and results in the asymmetric expression of a subset of genes in ASEL or ASER (Johnston et al. 2005). The transcription factor CHE-1 plays multiple roles in this process, acting both as a determinant of general ASE fate and as a regulator of asymmetric ASEL and ASER expression (Etchberger et al. 2009). Interestingly, however, several genes that are restricted to a single ASE neuron in adults are expressed in both ASEs in early larvae, including genes that participate in the execution of the asymmetric fate decision, such as the miRNA *lsy-6* (Johnston et al. 2005; Poole and Hobert 2006). Therefore, the precursors of ASEL and ASER retain a memory of their fate decision through several rounds of cell division, and even after embryogenesis they do not fully express the transcriptional programs associated with that decision. During this interval, a mechanism separate from those known to regulate ASE asymmetry must ensure the persistence of the original developmental decision.

Like the ASE neurons, the two AWC chemosensory neurons are morphologically symmetric but take on asymmetric cell fates that differ in the receptors they express and the odors to which they respond. $AWC^{ON}$ expresses the GPCR *str-2* and responds to the odor butanone, while $AWC^{OFF}$ expresses the GPCR *srsx-3* and responds to the odor 2,3-pentanedione (Troemel et al. 1999; Wes and Bargmann 2001; Bauer Huang et al. 2007). Unlike the ASE neurons, however, the AWC decision is stochastic, so that the $AWC^{ON}$ and $AWC^{OFF}$ identities are not linked to the right or left AWC neuron. Also unlike the ASE neurons, the asymmetric fate decision is made late in embryogenesis. This decision is mediated by communication between the two AWCs involving the innexin NSY-5 and the claudin-like protein NSY-4, and transduced by calcium signaling through calcium channel subunits UNC-2 and UNC-36 and a kinase cascade involving the CaMKII UNC-43, the MAPKKK NSY-1, and the MAPKK SEK-1 (Sagasti et al. 2001; Tanaka-Hino et al. 2002; Chuang and Bargmann 2005; Vanhoven et al. 2006; Bauer Huang et al. 2007; Chuang et al. 2007). Several of the genes involved in this decision are downregulated after hatching. Thus, during early larval stages the ASE neurons must complete the execution of a programmed decision made early in development, while the AWC neurons must maintain a decision made stochastically during late embryogenesis.

Underlying cell fate determines the set of receptors initially expressed by a given neuron, but external factors can extensively rearrange expression of GPCR receptor genes in the *C. elegans* chemosensory neurons. While the presence of multiple receptors in a single neuron reduces the effect of changes in expression of a single receptor gene, some receptors can play a significant role in the function of the neuron (Troemel et al. 1997). Furthermore, it is possible that

some of the signals that modulate GPCR expression may also alter the expression of additional genes specific to the cell. The interaction between the activity-dependent signals that modify receptor expression and the genes that regulate cell fate is therefore interesting to examine. cGMP, TGFβ, and *kin-29* signaling affects different receptors differently, but to what extent do these influences depend on the identity of the cell in which they act? The factors that respond to activity-dependent signaling and the factors that regulate cell fate will necessarily converge on transcriptional control of genes specific to the cell in question. It will be interesting to determine the interactions between sensory stimuli and cell fate regulators in the control and maintenance of cell-specific transcription in chemosensory neurons.

# Chapter 2

## Transcriptional regulation and stabilization of left-right neuronal identity in *C. elegans*


**Summary**

At discrete points in development, transient signals are transformed into long-lasting cell fates. For example, the asymmetric identities of two *C. elegans* olfactory neurons called AWC$^{ON}$ and AWC$^{OFF}$ are specified by a purely embryonic signaling pathway, but maintained throughout the life of an animal. Here we show that the DNA-binding protein NSY-7 acts to convert a transient, partially differentiated state into a stable AWC$^{ON}$ identity. Expression of an AWC$^{ON}$ marker is initiated in *nsy-7* loss-of-function mutants, but subsequently lost, so that most adult animals have two AWC$^{OFF}$ neurons and no AWC$^{ON}$ neurons. *nsy-7* encodes a protein with distant similarity to a homeodomain that is expressed in AWC$^{ON}$; it is an early transcriptional target of the embryonic signaling pathway that specifies AWC$^{ON}$ and AWC$^{OFF}$, and its expression anticipates future AWC asymmetry. The NSY-7 protein binds a specific optimal DNA sequence that was identified through a complete biochemical survey of 8-mer DNA sequences. This sequence is present in the promoter of a known AWC$^{OFF}$ marker and essential for its asymmetric expression. An 11-bp sequence within the AWC$^{OFF}$ promoter has two activities: one region activates expression in both AWCs, and the overlapping NSY-7-binding site inhibits expression in AWC$^{ON}$. Our results suggest that NSY-7 responds to transient embryonic signaling by repressing AWC$^{OFF}$ genes in AWC$^{ON}$, thus acting as a transcriptional selector for a randomly specified neuronal identity.

**Introduction**

Most neurons are born in embryogenesis but maintain their morphologies, physiological properties, and patterns of gene expression throughout life. The genetic pathways required to maintain neural identity can be distinct from those required for its establishment, suggesting that maintenance is an active genetic process. For example, MeCP2, a transcriptional regulator defective in the autism-like disorder Rett syndrome, acts in mature neurons to maintain dendritic complexity and synaptic function. Because of this late action of MeCP2, children with Rett syndrome are normal at birth and become symptomatic only months or years later, often losing language and motor skills that they acquired early in life (Dragich et al. 2000; Chen et al. 2001; Kishi and Macklis 2004; Guy et al. 2007).

In the nematode *Caenorhabditis elegans,* distinct genetic mechanisms control the initial specification and subsequent maintenance of a pair of olfactory neuron identities called $AWC^{ON}$ and $AWC^{OFF}$. $AWC^{ON}$ and $AWC^{OFF}$ neurons have similar morphologies but different functions and patterns of gene expression (Troemel et al. 1999). $AWC^{ON}$ senses the odor butanone, expresses the G protein-coupled receptor gene *str-2,* and can promote either attractive or repulsive behaviors based on modulatory inputs from a guanylate cyclase (Troemel et al. 1999; Wes and Bargmann 2001; Tsunozaki et al. 2008). The contralateral neuron, $AWC^{OFF}$, senses the odor 2,3-pentanedione, expresses the G protein-coupled receptor gene *srsx-3,* and has only been observed to mediate attractive behaviors (Wes and Bargmann 2001; Bauer Huang et al. 2007). Each animal has one $AWC^{ON}$ and one $AWC^{OFF}$ neuron, but these identities are randomly assumed by the right or left AWC neuron in each individual (Troemel

et al. 1999). This element of randomness is unusual in *C. elegans*, which is otherwise notable for the predictability of its cell fate decisions (Sulston and Horvitz 1977). The initial $AWC^{ON}$/$AWC^{OFF}$ decision depends on a signaling pathway in which a claudin-like transmembrane protein, NSY-4, and an embryonic gap junction network generated by the innexin NSY-5 induce one AWC to exit the $AWC^{OFF}$ default identity and become $AWC^{ON}$ (Vanhoven et al. 2006; Chuang et al. 2007). The induced $AWC^{ON}$ neuron then provides feedback to the contralateral neuron to stabilize the $AWC^{OFF}$ identity. Genetic analysis indicates that NSY-5/NSY-4 signaling acts through the calcium channel UNC-2/UNC-36 and the membrane protein OLRN-1, and regulates a kinase cascade that includes UNC-43 (CaMKII), NSY-1 (ASK1/MAPKKK), SEK-1 (SEK/MAPKK), and TIR-1 (an adaptor protein). High activity of the calcium/kinase cascade is associated with the future $AWC^{OFF}$ neuron, and low activity with the future $AWC^{ON}$ (Sagasti et al. 2001; Tanaka-Hino et al. 2002; Chuang and Bargmann 2005; Bauer Huang et al. 2007).

The NSY-5 gap junction network observed in electron micrographs disappears after hatching, and the downstream signaling pathway apparently becomes inactive as well (Chuang and Bargmann 2005; Chuang et al. 2007). AWC left-right asymmetry is stable throughout the lifetime of the animal, so these transient signaling events must be captured and stabilized by molecules that act after embryogenesis. Indeed, postembryonic expression of AWC markers requires a different set of genes than those that act in the embryo. Several genes that affect olfactory signal transduction are required to maintain the $AWC^{ON}$ marker *str-2::GFP* after the first larval stage (Troemel et al. 1999). These include *odr-1* and *daf-11*, which encode receptor guanylate cyclase

homologs that are localized to the sensory cilia and thought to produce cGMP during chemosensation (Birnby et al. 2000; L'Etoile et al. 2000). Mutations in genes encoding olfactory Ga proteins also boost *str-2* expression during larval stages (Lans et al. 2004; Lans and Jansen 2006). The genetic requirement for olfactory signal transduction molecules suggests that ongoing sensory activity maintains AWC gene expression.

The contrast between embryonic signaling mutants and postembryonic maintenance mutants defines a period of interest near hatching when the AWC identity is crystallized. Here we show that *nsy-7*, which encodes a divergent transcriptional regulator, coordinates the transition from establishment to maintenance of the $AWC^{ON}$ identity. *nsy-7* mutants exhibit evidence of left-right asymmetry during the first larval stage, but lose this asymmetry and markers of $AWC^{ON}$ function by adulthood. We find that *nsy-7* expression is an early target of the developmental signaling pathway that induces $AWC^{ON}$, and provide evidence that it directly represses $AWC^{OFF}$ genes to ensure a stable $AWC^{ON}$ identity.

**Results**

**Left-right AWC asymmetry emerges in the L1 larval stage**

To visualize the $AWC^{ON}$ and $AWC^{OFF}$ neurons simultaneously, we generated a transgenic strain, *kyIs408*, with integrated *str-2::dsRed2* ($AWC^{ON}$) and *srsx-3::GFP* ($AWC^{OFF}$) reporter genes. In adult animals, *srsx-3::GFP* and *str-2::dsRed2* were expressed in a mutually exclusive pattern in contralateral AWC neurons (Figure 2.1A). We used this strain to characterize the appearance of left-right asymmetry during development.

In newly hatched animals, *srsx-3::GFP* was expressed in both AWCs (Figure 2.1B). The two AWCs differed approximately 5-fold in their level of *srsx-3* expression, suggesting that left-right asymmetry had been initiated by this time (Figure 2.1C). Over the course of the first larval stage (14 h), expression of *srsx-3* increased two- to three-fold in both AWC neurons. Between the L2 larval stage and the adult stage, *srsx-3* expression fell in $AWC^{ON}$ (defined as the AWC with low *srsx-3* expression), and remained high in $AWC^{OFF}$ (the AWC with high *srsx-3* expression).

As previously observed with the marker *str-2::GFP* (Troemel et al., 1999), *str-2::dsRed* was not expressed in the AWCs at hatching. Beginning about five hours after hatching (mid-L1), faint *str-2* expression could be detected in one AWC neuron, which was invariably the AWC with weaker expression of *srsx-3*. At the L1/L2 transition, more than 95% of animals expressed *str-2*, always in the AWC with weaker *srsx-3* expression (Figure 2.1B-D). If either $AWC^{ON}$ or $AWC^{OFF}$ is killed at this stage, the surviving AWC does not change its pattern of gene expression, suggesting that the $AWC^{ON}$ and $AWC^{OFF}$ identities have been irreversibly determined (Troemel et al. 1999). Between the L2 larval stage and the adult stage, *str-2* expression increased in the $AWC^{ON}$ neuron.

These results establish the timing of three events in AWC differentiation: repression of *srsx-3* in $AWC^{ON}$, which begins by hatching; induction of *str-2* in $AWC^{ON}$ during the mid- to late L1 stage; and the increased expression of *srsx-3* and *str-2* in the appropriate AWCs between L1 and adult. As expected from previous studies (Troemel et al. 1999), AWCL and AWCR were equally likely to enter the $str\text{-}2^{+}srsx\text{-}3^{-}$ $AWC^{ON}$ state or the $str\text{-}2^{-}srsx\text{-}3^{+}$ $AWC^{OFF}$ state.

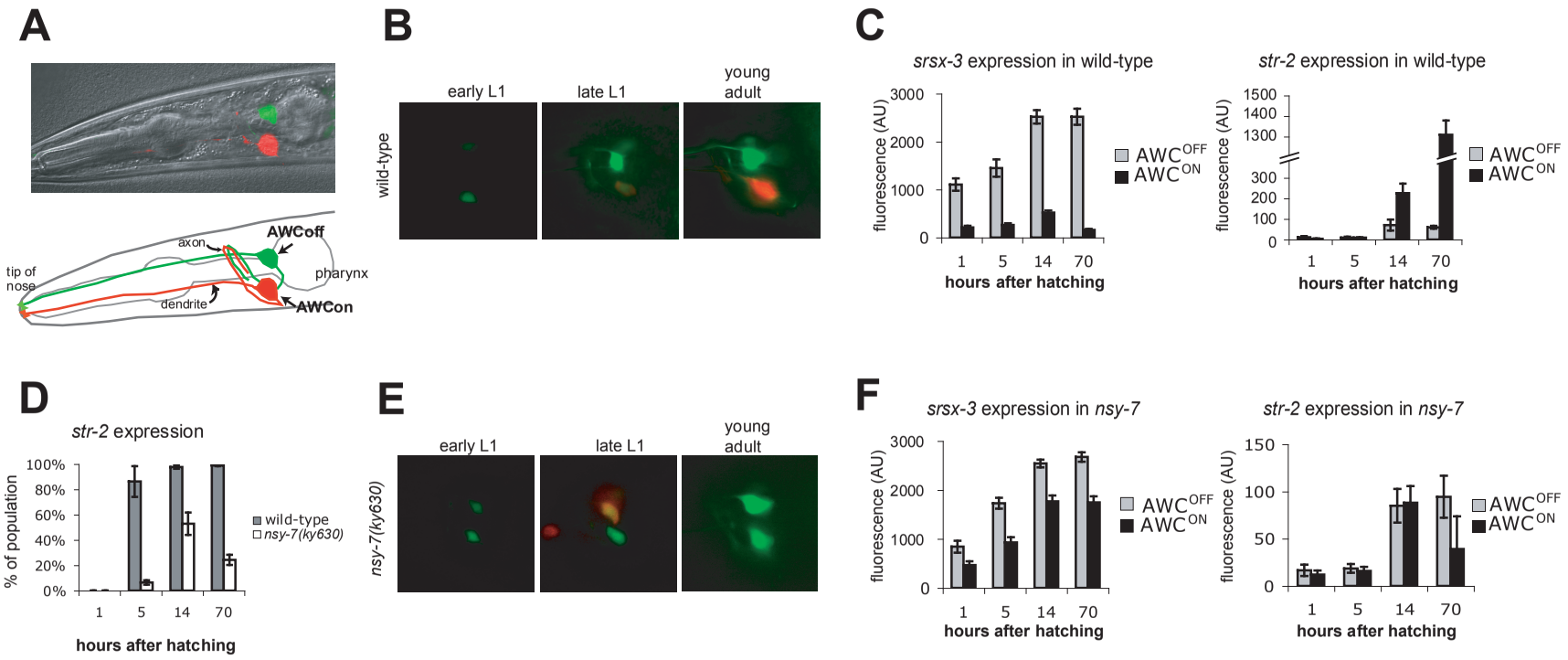



**Figure 2.1. *nsy-7(ky630)* mutants are defective for maintenance of asymmetric identity in AWC.** A, Adult animal with the integrated array *kyIs408*. A fluorescence image of AWC$^{ON}$, expressing *str-2::dsRed2*, and AWC$^{OFF}$, expressing *srsx-3::GFP*, is overlaid on a DIC image. B, AWC neurons in wild-type animals in early L1 (1 hour after hatching), late L1 (14h after hatching), or in the young adult. Red fluorescence, *str-2::dsRed2*; green fluorescence, *srsx-3::GFP*. C, Quantitation of *srsx-3::GFP* and *str-2::dsRed2* fluorescence (n= 25-38 animals). In each animal, the cell with higher *srsx-3::GFP* expression was defined as AWC$^{OFF}$. D, Fraction of *nsy-7(ky630)* animals expressing *str-2::dsRed2* at 1, 5, 14, and 70 hours after hatching. 1-14 hours are L1 stage, 70 hours is adult. E, AWC neurons in *nsy-7(ky630)* animals, as in B. F, Quantitation of *srsx-3::GFP* and *str-2::dsRed2* fluorescence in *nsy-7(ky630)* mutants, as in C (n = 33-45 animals).

### *nsy-7(ky630)* initiates but fails to maintain AWC$^{ON}$ asymmetry

*nsy-7(ky630)* was isolated from a screen for mutants that fail to express *str-2* in either AWC neuron (Vanhoven et al. 2006). As adults, these mutants had a typical 2-AWC$^{OFF}$ phenotype in which both AWC neurons expressed *srsx-3*. Like olfactory transduction mutants, but unlike most other mutants from this screen, *nsy-7(ky630)* larvae exhibited evidence of AWC left-right asymmetry during the L1 stage, when many animals expressed *str-2* in one of the two AWC neurons (Figure 2.1D). The fraction of animals expressing *str-2* decreased later in development (Figure 2.1D).

To define the differences between wild type and *nsy-7* more precisely, the levels of *str-2::dsRed* and *srsx-3::GFP* expression were compared quantitatively at different developmental stages. Throughout development, both AWCs in *nsy-7* mutants expressed *srsx-3* at similar levels resembling the levels in normal AWC$^{OFF}$ neurons (Figure 2.1E, F). At the L1/L2 transition, half of the *nsy-7* animals also expressed *str-2* in one AWC neuron (Figure 2.1D, E), the signature of an AWC$^{ON}$-like identity. *str-2* expression was always restricted to one of the two AWCs; this cell was equally likely to be the cell with slightly higher or slightly lower *srsx-3::GFP* expression (Figure 2.1F). The fraction of animals expressing *str-2::dsRed* fell after the L1 stage, as did the level of *str-2::dsRed* expression within AWC (Figure 2.1F). Some adult animals expressed *str-2::dsRed* at low levels, but because of the low levels, the fraction scored as positive varied between experiments. All AWC neurons in adults expressed *srsx-3*. These results suggest that *nsy-7(ky630)* mutants initiate elements of an asymmetric AWC$^{ON}$-like identity, but cannot repress the AWC$^{OFF}$ marker or maintain the AWC$^{ON}$ marker after the L1 stage.

If *nsy-7* mutants can not stabilize the postembryonic AWC$^{ON}$ identity, AWC$^{ON}$ markers should be lost regardless of the activity of the upstream signaling pathway. This prediction was tested by making double mutants between *nsy-7* and genes in the signaling pathway for AWC asymmetry. Null mutants in the calcium channel genes *unc-2* and *unc-36* or in the downstream MAPKKK gene *nsy-1* have a 2-AWC$^{ON}$ phenotype due to disruptions in signaling between the embryonic AWC neurons (Troemel et al. 1999). *nsy-7; unc-2* mutants, *nsy-7; unc-36* mutants, and *nsy-7; nsy-1* mutants all resembled *nsy-7*: in the double mutants, both AWCs expressed *srsx-3* and neither expressed *str-2*, a 2-AWC$^{OFF}$ phenotype (Table 2.1). These results are consistent with an essential function for *nsy-7* downstream of *unc-2*, *unc-36*, and *nsy-1*, and suggest that *nsy-7* may maintain the AWC$^{ON}$ identity and suppress the AWC$^{OFF}$ identity.

**Olfactory cGMP transduction pathways affect both AWC$^{ON}$ and AWC$^{OFF}$**

The time course of *str-2* expression in *nsy-7* mutants resembles that of *odr-1* olfactory transduction mutants, which initiate but fail to maintain asymmetric *str-2* expression in AWC$^{ON}$ (Troemel et al. 1999). *odr-1* encodes a guanylate cyclase homolog involved in olfactory transduction, whose late role may represent an activity-dependent input into AWC gene expression (L'Etoile and Bargmann 2000). Unlike *nsy-7* mutants, however, we found that *odr-1* mutants were also defective in maintenance of the AWC$^{OFF}$ marker *srsx-3*, so that *srsx-3* expression was present in young *odr-1* larvae but absent in *odr-1* adults (Table 2.1 and data not shown). These results suggest defects in both AWC$^{ON}$ and AWC$^{OFF}$ in olfactory transduction mutants.

## Table 1. *nsy-7* double mutants

| Percentage of adults expressing *str-2::dsRed2* in 0, 1, or 2 AWCs | | | | |
|---|---|---|---|---|
| | 0 AWC | 1 AWC | 2 AWC | n |
| *kyIs408* | 0 | 100 | 0 | 265 |
| *kyIs405* | 0 | 100 | 0 | 109 |
| | | | | |
| *nsy-7 (ky630)* | 61 | 39 | 0 | 187 |
| *unc-36 (e251)* | 0 | 0 | 100 | 42 |
| *unc-2(lj1)* | 6 | 23 | 71 | 31 |
| *nsy-1(ky542);kyIs405* | 0 | 0 | 100 | 48 |
| *odr-1 (n1936)* | 94 | 6 | 0 | 106 |
| *egl-4(ks60)* | 91 | 9 | 0 | 54 |
| | | | | |
| *unc-36;nsy-7* | 100 | 0 | 0 | 33 |
| *unc-2;nsy-7* | 100 | 0 | 0 | 68 |
| *nsy-1;nsy-7;kyIs405* | 100 | 0 | 0 | 46 |
| *odr-1;nsy-7* | 99 | 1 | 0 | 83 |
| *egl-4;nsy-7* | 100 | 0 | 0 | 76 |

| Percentage of adults expressing *srsx-3::GFP* in 0, 1, or 2 AWCs | | | | |
|---|---|---|---|---|
| | 0 AWC | 1 AWC | 2 AWC | n |
| *kyIs408* | 0 | 100 | 0 | 265 |
| *kyIs405* | 0 | 100 | 0 | 109 |
| | | | | |
| *nsy-7 (ky630)* | 0 | 2 | 98 | 187 |
| *unc-36 (e251)* | 100 | 0 | 0 | 42 |
| *unc-2(lj1)* | 68 | 26 | 6 | 31 |
| *nsy-1(ky542);kyIs405* | 100 | 0 | 0 | 48 |
| *odr-1(n1936)* | 96 | 4 | 0 | 106 |
| *egl-4(ks60)* | 100 | 0 | 0 | 54 |
| | | | | |
| *unc-36;nsy-7* | 0 | 0 | 100 | 33 |
| *unc-2;nsy-7* | 0 | 0 | 100 | 68 |
| *nsy-7;nsy-1;kyIs405* | 0 | 0 | 100 | 46 |
| *odr-1;nsy-7* | 100 | 0 | 0 | 83 |
| *egl-4;nsy-7* | 100 | 0 | 0 | 76 |

| Percentage of adults expressing *str-2::dsRed2* in 0, 1, or 2 AWCs | | | | |
|---|---|---|---|---|
| | 0 AWC | 1 AWC | 2 AWC | n |
| N2 | 30 | 70 | 0 | 175 |
| *nsy-1(ky542)* | 36 | 29 | 35 | 121 |
| *nsy-5(ky634)* | 99 | 1 | 0 | 68 |

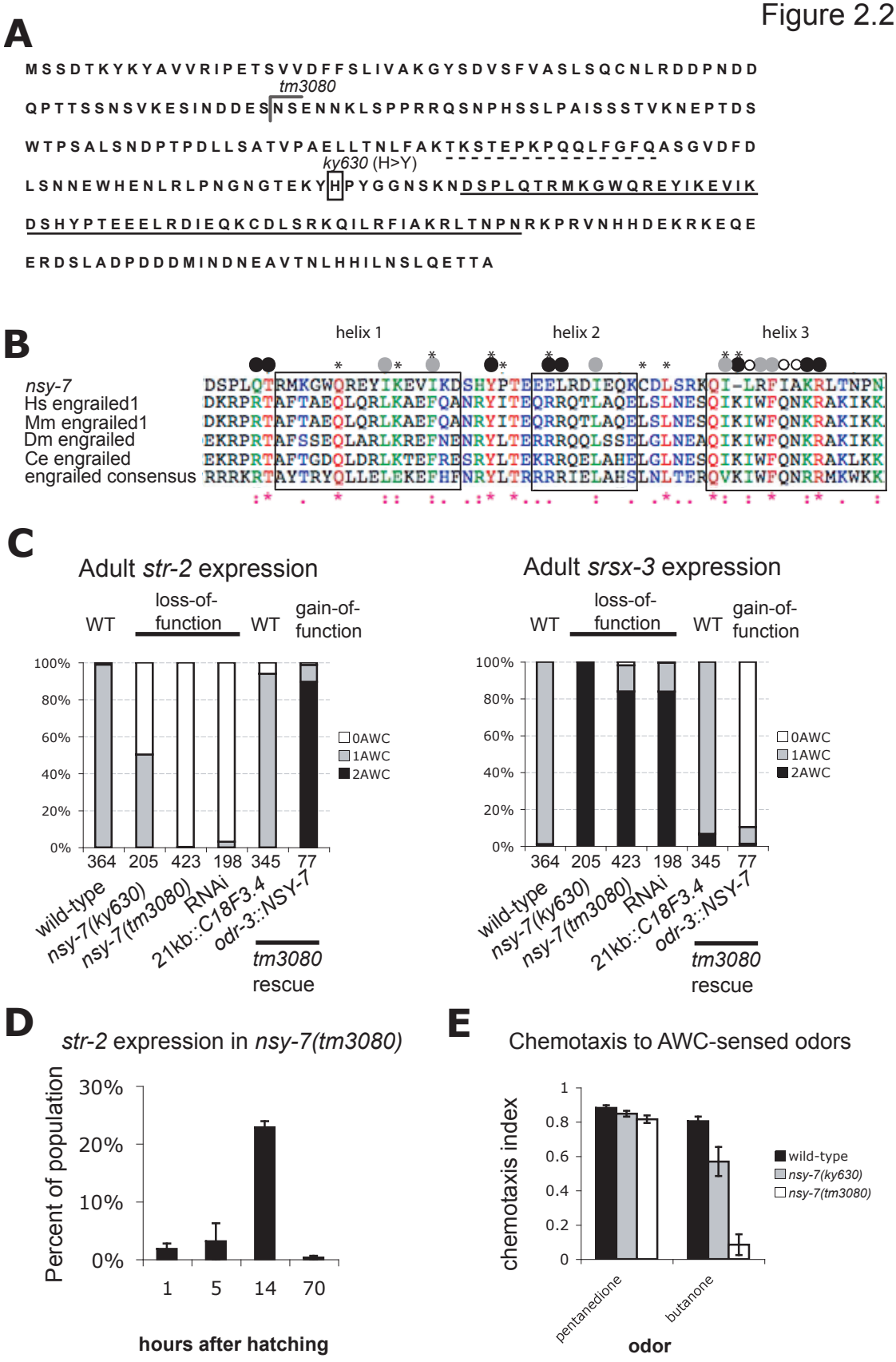
A possible target of cGMP produced by ODR-1 is EGL-4, a cGMP-dependent kinase homolog that is required for normal olfaction and olfactory plasticity in AWC, and for expression of some chemoreceptor genes in AWB neurons (Daniels et al. 2000; L'Etoile et al. 2002; van der Linden et al. 2008). Like *odr-1* mutants, *egl-4* null mutants expressed *str-2* and *srsx-3* in the AWC neurons of L1 animals, but not in adults (Table 2.1 and data not shown). Thus *odr-1* and *egl-4* are required to maintain and increase expression of both $AWC^{ON}$ and $AWC^{OFF}$ genes between the L1 stage and the adult stage, in contrast with *nsy-7*, which is required specifically for appropriate gene induction and repression in $AWC^{ON}$. *nsy-7; odr-1* and *nsy-7; egl-4* double mutants resembled the *odr-1* or *egl-4* single mutants: adults expressed neither *str-2* nor *srsx-3* (Table 2.1).

**_nsy-7_ encodes a homeodomain-like protein that acts in AWC**

*nsy-7(ky630)* was mapped using standard methods to the predicted ORF C18F3.4, which encodes an uncharacterized protein (Figure 2.2A). Although it had no significant homologs outside nematodes by BLAST searches, low-stringency searches and motif searches revealed a region with distant similarity to a homeodomain (Figure 2.2B) (Gehring et al. 1994; Draganescu and Tullius 1998; Fraenkel et al. 1998; Sato et al. 2004; Chi 2005). The *ky630* mutation was associated with a C→T transition in the third exon of the predicted gene, resulting in a missense H→Y mutation immediately before the homeodomain-like region (Figure 2.2A). *nsy-7(ky630)* is recessive and RNAi against C18F3.4 generated a *nsy-7*-like phenotype, with two *srsx-3*-expressing $AWC^{OFF}$ cells and

**Figure 2.2. *nsy-7* encodes a protein with distant similarity to homeodomains.** A, Sequence of the long isoform of NSY-7. The *ky630* and *tm3080* mutations are marked, and the predicted homeodomain-like region is underlined. Dotted line indicates residues absent in the short NSY-7 isoform. B, Alignment of the NSY-7 homeodomain-like region with engrailed-family homeodomains. Engrailed residues that bind specific bases (open circles) or the DNA backbone (black circles) and residues that form the hydrophobic core (grey circles) are marked. Asterisks above the sequences mark residues important for homeodomain function (Gehring et al. 1994; Draganescu and Tullius 1998; Fraenkel et al. 1998; Sato et al. 2004; Chi 2005). C, Expression of *srsx-3::GFP* and *str-2::dsRed* in AWC neurons of wild type, *nsy-7* mutants, C18F3.4-RNAi animals, and *nsy-7(tm3080)* rescued with a genomic clone or an *odr-3::nsy-7* cDNA clone. Numbers of animals scored are indicated. D, Percentage of *nsy-7(tm3080)* animals expressing *str-2::dsRed2* at 1, 5, 14, and 70 hours after hatching. 1-14 hours are L1 stage, 70 hours is young adult. E, Chemotaxis to the $AWC^{ON}$-specific odor 2-butanone and the $AWC^{OFF}$-specific odor 2,3-pentanedione.

Figure 2.2


A
MSSDTKYKYAVVRIPETSVVDFFSLIVAKGYSDVSFVASLSQCNLRDDPNDD
tm3080
QPTTSSNSVKESINDDESNSENNKLSPPRRQSNPHSSLPAISSSTVKNEPTDS
WTPSALSNDPTPDLLSATVPAELLTNLFAKTKSTEPKPQQLFGFQASGVDFD
ky630 (H>Y)
LSNNEWHENLRLPNGNGTEKYHPYGGNSKNDSPLQTRMKGWQREYIKEVIK
DSHYPTEEELRDIEQKCDLSRKQILRFIAKRLTNPNRKPRVNHHDEKRKEQE
ERDSLADPDDDMINDNEAVTNLHHILNSLQETTA
B
helix 1
helix 2
helix 3
nsy-7
Hs engrailed1
Mm engrailed1
Dm engrailed
Ce engrailed
engrailed consensus
C
Adult str-2 expression
WT
loss-of-function
WT
gain-of-function
100%
80%
60%
40%
20%
0%
364
205
423
198
345
77
0AWC
1AWC
2AWC
wild-type
nsy-7(ky630)
nsy-7(tm3080)
RNAi
21kb::C18F3.4
odr-3::NSY-7
tm3080 rescue
Adult srsx-3 expression
D
str-2 expression in nsy-7(tm3080)
Percent of population
30%
20%
10%
0%
1
5
14
70
hours after hatching
E
Chemotaxis to AWC-sensed odors
chemotaxis index
1
0.8
0.6
0.4
0.2
0
wild-type
nsy-7(ky630)
nsy-7(tm3080)
pentanedione
butanone
odor

no *str-2*-expressing AWC$^{ON}$ cells (Figure 2.2C). These results identify *nsy-7(ky630)* as a likely reduction-of-function mutation in C18F3.4.

A deletion mutation in C18F3.4, *tm3080*, was kindly provided by the National Bioresource Project in Japan. *tm3080* deletes parts of the second and third exons and then generates a frameshift that leads to early termination. Therefore, *tm3080* probably represents a null mutation in C18F3.4. The *nsy-7(tm3080)* mutant phenotype was similar to but stronger than the *nsy-7(ky630)* phenotype: almost all adult animals had two *srsx-3*-expressing AWC$^{OFF}$ cells and no *str-2*-expressing AWC$^{ON}$ cells (Figure 2.2C). A fraction of L1 larvae expressed *str-2* in one AWC neuron, suggesting initiation of the AWC$^{ON}$ identity, but this expression was unstable (Figure 2.2D). *nsy-7(ky630)* and *nsy-7(tm3080)* mutants had no obvious defects in growth rate, brood size, morphology, or coordination.

The functions of AWC$^{ON}$ and AWC$^{OFF}$ were examined by testing chemotaxis to butanone and 2,3-pentanedione, odors sensed by AWC$^{ON}$ and AWC$^{OFF}$, respectively (Wes and Bargmann 2001). *nsy-7(ky630)* animals had a mild defect and *nsy-7(tm3080)* had a severe defect in butanone chemotaxis, and both mutants were normal for chemotaxis to 2,3-pentanedione (Figure 2.2E). The milder butanone chemotaxis defect in *nsy-7(ky630)* correlated with its weaker effect on adult *str-2* expression (Figure 2.2C). Thus *nsy-7* is required for multiple properties of AWC$^{ON}$, including sensation of butanone, expression of *str-2*, and repression of *srsx-3*.

When expressed in both AWC neurons of *nsy-7(tm3080)* mutants under the AWC-selective *odr-3* promoter, C18F3.4 cDNAs caused an overexpression phenotype opposite to the *nsy-7* mutant phenotype. Both AWCs in these transgenic animals expressed *str-2* and not *srsx-3*, exhibiting the properties of

AWC$^{ON}$ cells (Figure 2.2C). These results suggest that C18F3.4 functions in AWC to regulate expression of *str-2* and *srsx-3* and confer the AWC$^{ON}$ identity.

**The AWC asymmetry pathway regulates *nsy-7* transcription**

Regulatory sequences for *nsy-7* were identified by characterizing the genomic regions required for rescue of *nsy-7(tm3080)* mutants. Partial rescue was obtained with the C18F3.4 genomic coding region and 9 kb of upstream sequence, and rescue was improved by increasing the amount of upstream sequence (Figure 2.3). Near-complete rescue was achieved using the genomic coding region with 21 kb upstream of the start site (Figure 2.2C).

When fused to a GFP reporter, the 21-kb region upstream of *nsy-7* drove expression in numerous cell types including gut, the amphid sheath glial cells, and head and tail neurons including AWC, ASE, and ASH (Figure 2.4A and data not shown). Its broad expression pattern suggests that *nsy-7* may have uncharacterized functions that are unrelated to AWC asymmetry. In adult animals, *nsy-7::GFP* was asymmetrically expressed in one AWC neuron (Figure 2.4A). The *nsy-7*-expressing neuron was identified by crossing a line expressing *nsy-7::GFP* with a line expressing an *srsx-3::mCherry* (AWC$^{OFF}$) reporter or a *str-2::dsRed2* (AWC$^{ON}$) reporter. *nsy-7* was consistently expressed in the same cell as *str-2* (145/158 animals), and contralateral to the cell expressing *srsx-3* (99/99 animals). Therefore *nsy-7* appears to be expressed in AWC$^{ON}$ but not AWC$^{OFF}$.

The developmental time course of *nsy-7::GFP* expression was examined in a transgenic strain expressing *nsy-7::GFP* and a marker for both AWC neurons, *odr-1::mCherry*. At the 3-fold embryonic stage, when the signaling pathway initiates AWC asymmetry, most animals expressed *nsy-7::GFP* asymmetrically in

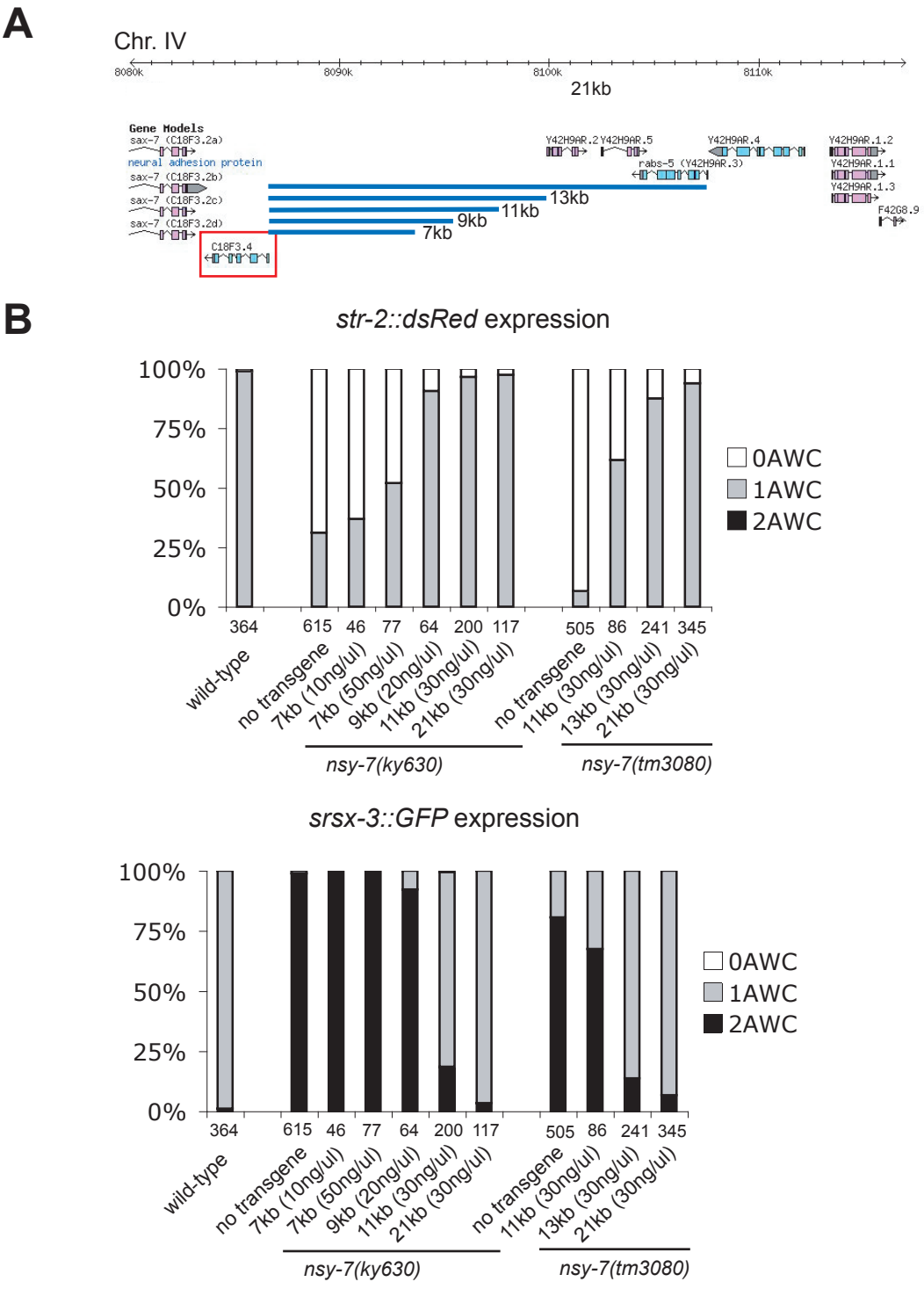



**Figure 2.3. Mapping and rescue of *nsy-7*.** A, C18F3.4 (*nsy-7*) and its genomic context. C18F3.4 is boxed, and promoter regions used for rescue are marked. B, Rescue of *nsy-7(ky630)* and *nsy-7(tm3080)* mutants with *nsy-7* genomic sequence expressed under different promoter regions. Numbers of animals scored are indicated.

one AWC neuron, which was equally likely to be AWCL or AWCR (Figure 2.4B). At this stage, a subset of embryos expressed *nsy-7::GFP* bilaterally in AWC; bilateral AWC expression fell beginning at the L1 larval stage, and was not observed in adults. These results suggest that *nsy-7::GFP* expression is an early marker for AWC asymmetry. To determine whether the apparent asymmetry was regulated by the embryonic signaling pathway, *nsy-7::GFP* expression was examined in signaling mutants with 2-AWC$^{OFF}$ or 2-AWC$^{ON}$ phenotypes. In 2-AWC$^{OFF}$ *nsy-5(ky634)* mutants, adult *nsy-7::GFP* expression was lost (Table 2.1). In 2-AWC$^{ON}$ *nsy-1* mutants, 35% of adults expressed *nsy-7::GFP* in both AWC neurons (Table 2.1). The regulation of *nsy-7::GFP* by the embryonic AWC asymmetry genes suggests that the *nsy-7* promoter is a transcriptional target of the AWC signaling pathway.

*nsy-7::GFP* was expressed bilaterally in the ASE neurons, another pair of chemosensory neurons that exhibits asymmetry. The asymmetric identities of the ASE neurons can be followed using the reporters *gcy-5::GFP*, expressed in ASER, and *gcy-7::GFP*, expressed in ASEL (Yu et al. 1997). Both *gcy-5::GFP* and *gcy-7::GFP* were appropriately and asymmetrically expressed in *nsy-7* mutants, indicating that ASE asymmetry does not require *nsy-7* (data not shown).

**NSY-7 is a nuclear protein that binds the consensus DNA sequence CCTTAAC**

NSY-7 has distant similarity to homeodomain proteins, suggesting that it might act as a transcription factor. Consistent with this possibility, a C-terminally-tagged NSY-7::GFP fusion protein was localized to the nucleus of AWC (Figure 2.4C). The fusion protein was biologically active, as it induced a gain-of-function 2-AWC$^{ON}$ phenotype when expressed in both AWC neurons in a wild-type

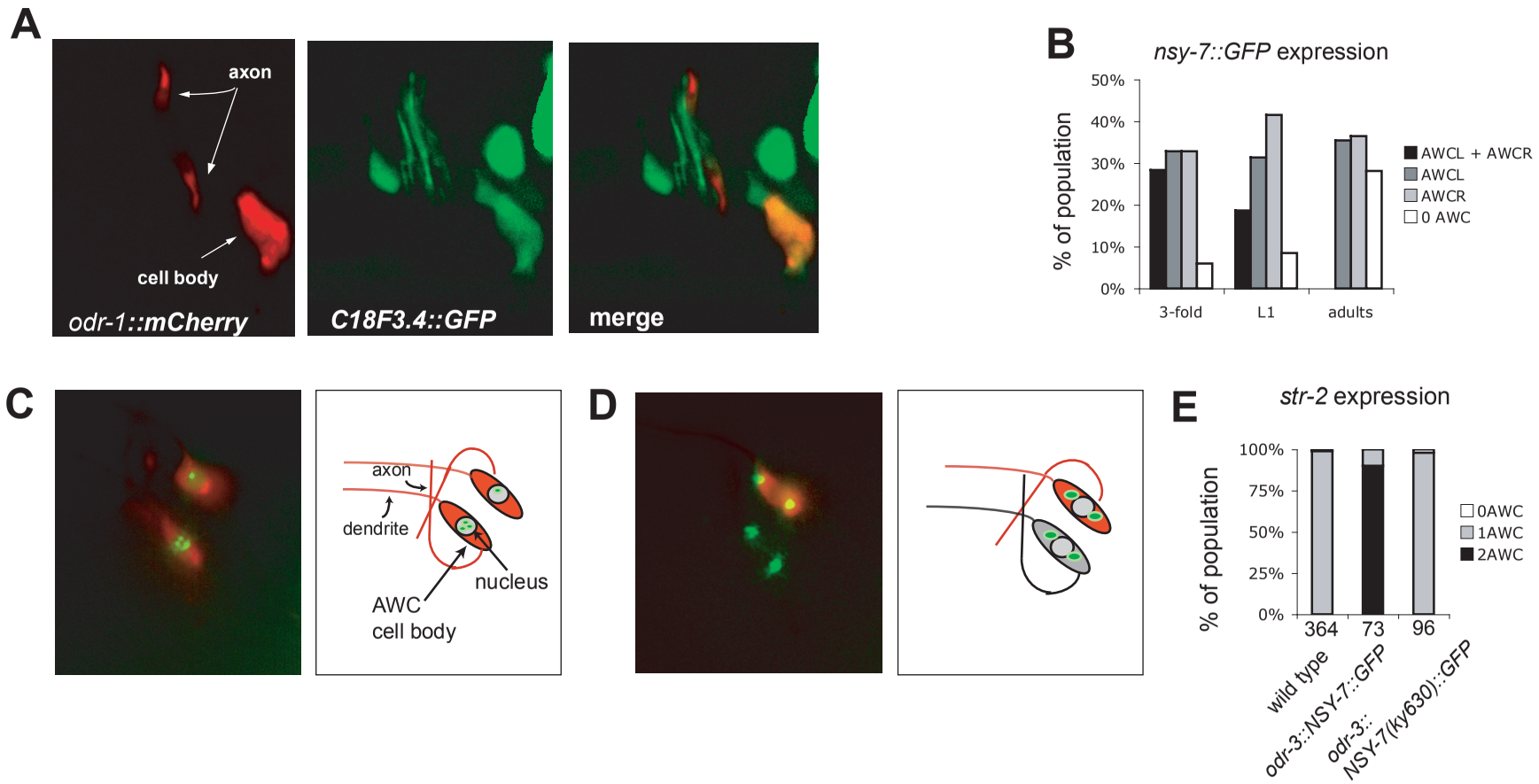



**Figure 2.4. Expression of *nsy-7::GFP* and nuclear localization of NSY-7.** A, Expression of a 21 kb *nsy-7::GFP* promoter fusion (green) and the AWC marker *odr-1::mCherry* (red). The AWC cell body and AWC axons in the nerve ring are marked. B, *nsy-7::GFP* expression across development. C, Nuclear localization of a NSY-7::GFP protein expressed in both AWCs under the *odr-3* promoter. *str-2::dsRed* marks AWC$^{ON}$; note gain-of function 2-AWC$^{ON}$ phenotype. D, Cytoplasmic localization of NSY-7::GFP with the *ky630* mutation (*odr-3* promoter, *str-2::dsRed* marker); note wild-type asymmetric expression of *str-2::dsRed*. E, Quantification of *str-2::dsRed* expression in wild type, or wild type overexpressing *odr-3::nsy-7::GFP* with or without the *ky630* mutation. Numbers of animals scored are indicated.

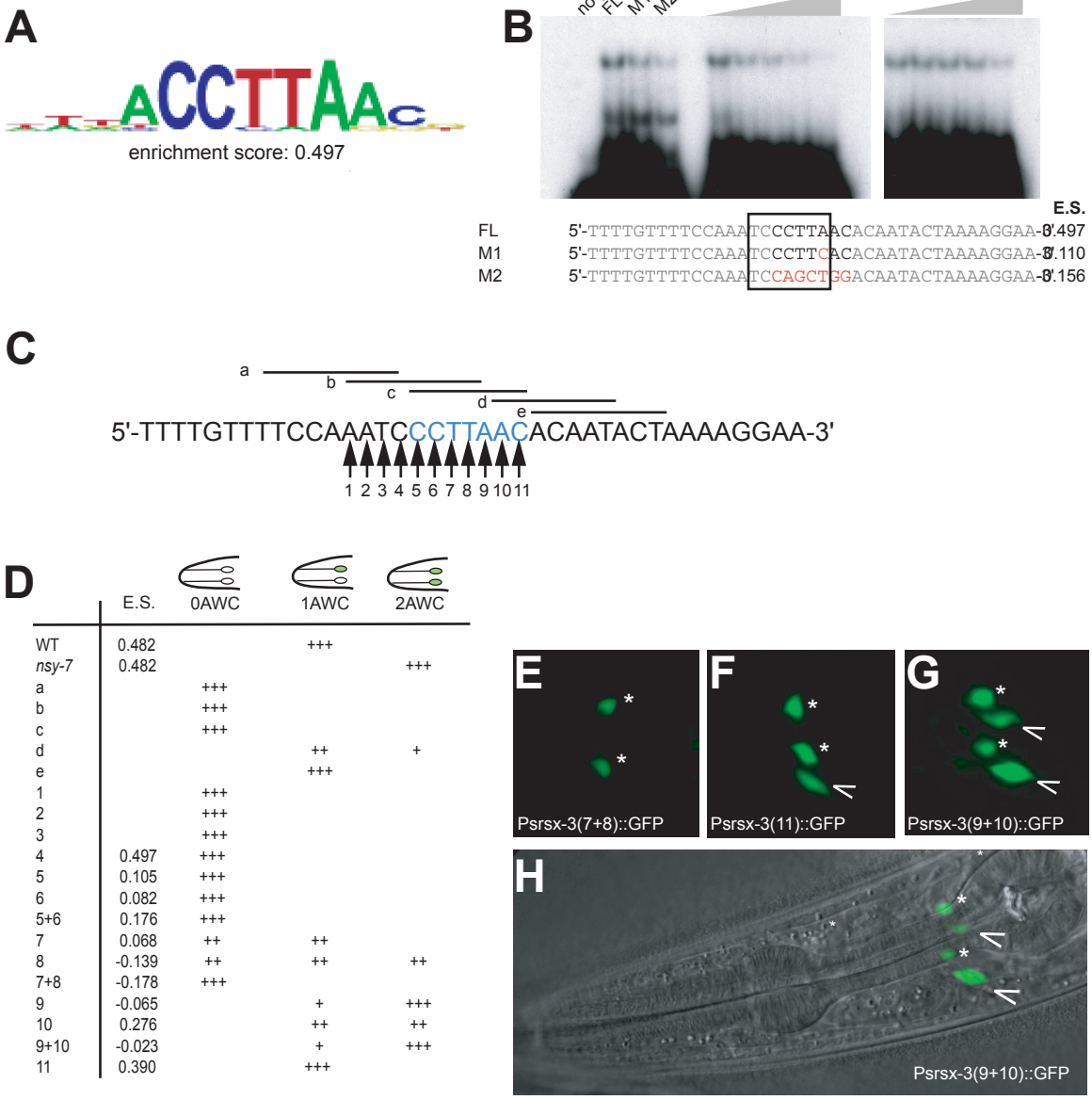
background (Figure 2.4E). A NSY-7::GFP fusion protein carrying the missense mutation H179Y encoded by the *ky630* mutation failed to localize to the nucleus (Figure 2.4D), and also failed to induce a 2-AWC$^{ON}$ phenotype when expressed in both AWCs (Figure 2.4E). These results suggest that NSY-7 nuclear localization may be required for its activity.

To ask whether NSY-7 can bind to DNA and to identify potential target sequences, we made use of universal protein-binding microarrays (Berger et al. 2006; Berger et al. 2008). The DNA microarrays consisted of 41,944 60-mer oligonucleotide elements that collectively represent all possible 8-mer DNA sequences, each repeated approximately 32 times. Full-length GST-tagged NSY-7 protein was produced in *E. coli*, purified, and applied to the DNA microarray; NSY-7 binding at each DNA spot was detected and quantified using anti-GST antibody. NSY-7 binding preferences over all 8-mers were calculated and used to determine a position weight matrix (PWM) and corresponding sequence logo, as well as an 'enrichment score' between -0.5 and 0.5, reflecting the preference of the protein for any given 8-nucleotide sequence. NSY-7 bound the specific recognition site ACCTTAAC with an enrichment score of 0.497 (Figure 2.5A; this data is freely available on the UniPROBE database (Newburger and Bulyk 2008)); an enrichment score greater than 0.45 is indicative of non-random binding (Berger et al. 2008). These results demonstrate that NSY-7 is a sequence-specific DNA binding protein.

The promoter regions of genes that are expressed in AWC neurons were examined for the presence of the ACCTTAAC binding site identified in the protein-binding microarrays. None of these promoters contained the complete 8-mer, but a perfect match for the sequence CCTTAAC was found in the *srsx-3*

**Figure 2.5. NSY-7 binds to the sequence CCTTAAC.** A, Consensus NSY-7 binding sequence determined by protein binding microarray experiments. The enrichment score correlates with the binding affinity of NSY-7 for the sequence, and is measured on a scale of -0.5 to 0.5. B, In vitro binding of 6His-tagged NSY-7 to the predicted consensus sequence. FL, full-length consensus sequence in a 39-nucleotide context from the *srsx-3* promoter. Consensus binding site is boxed. M1 and M2, mutations predicted to decrease the enrichment score (E.S.), shown to the right of the sequences. All lanes except the first contain 500 ng of 6His:NSY-7. Lanes 5-9 contain labeled FL probe and increasing amounts of cold FL competitor; lanes 10-14, labeled FL probe + cold M2. Lower band in lanes 2-4 probably represents a degradation product of the recombinant protein. C, Mutations made in the *srsx-3* promoter and D, resulting expression patterns. Enrichment scores are shown for each point mutation. (+++) indicates 80-100% of population; (++) 30-80%; (+)10-30%; blank <1%. E, GFP expression under an *srsx-3* promoter containing mutations 7 and 8. F, GFP expression under an *srsx-3* promoter containing mutation 11. G, GFP expression under an *srsx-3* promoter containing mutations 9 and 10. H, DIC image of worm shown in (G). In E-H, asterisks indicate AWB and arrowheads indicate AWC.

Figure 2.5




| | E.S. | 0AWC | 1AWC | 2AWC |
|---|---|---|---|---|
| WT | 0.482 | | +++ | |
| *nsy-7* | 0.482 | | | +++ |
| a | | +++ | | |
| b | | +++ | | |
| c | | +++ | | |
| d | | | ++ | + |
| e | | | +++ | |
| 1 | | +++ | | |
| 2 | | +++ | | |
| 3 | | +++ | | |
| 4 | 0.497 | +++ | | |
| 5 | 0.105 | +++ | | |
| 6 | 0.082 | +++ | | |
| 5+6 | 0.176 | +++ | | |
| 7 | 0.068 | ++ | ++ | |
| 8 | -0.139 | ++ | ++ | ++ |
| 7+8 | -0.178 | +++ | | |
| 9 | -0.065 | | + | +++ |
| 10 | 0.276 | | ++ | ++ |
| 9+10 | -0.023 | | + | +++ |
| 11 | 0.390 | | +++ | |

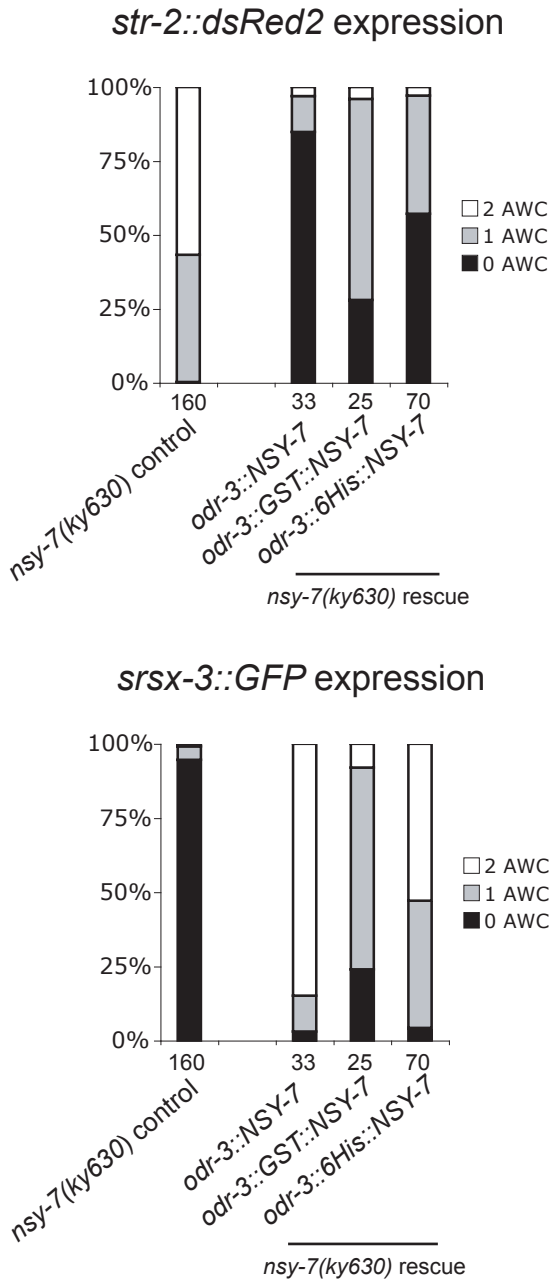



**Figure 2.6. Activity of GST::NSY-7 and 6His::NSY-7 proteins.** Expression of *str-2* and *srsx-3* in *nsy-7(ky630)* strains with bilateral AWC expression of NSY-7, GST:NSY-7 (used for protein binding microarray experiments), and 6His:NSY-7 (used for EMSAs) under an *odr-3* promoter. Controls are sibling *nsy-7(ky630)* animals lacking the unstable *odr-3::nsy-7* transgene; data represents the average of transgene-negative animals from each of the three lines. Numbers of animals scored are indicated.

promoter, 425 base pairs upstream of the predicted start site. Binding of NSY-7 to the preferred sequence from the *srsx-3* promoter was verified by electromobility shift assays (EMSAs) using a $^{32}$P-labelled double-stranded oligonucleotide with the sequence CCTTAAC flanked by 16 nucleotides on each side. A full-length, bacterially-produced His-tagged NSY-7 protein bound the labeled probe, and this interaction was competed away by unlabelled probe (Figure 2.5B, 2.6). Mutations in the predicted binding sequence reduced binding of NSY-7 to labeled probe and reduced the ability of unlabelled probe to compete NSY-7 protein away from wild-type labeled probe (M1 and M2, Figure 2.5B). NSY-7 also weakly bound the M2 mutant probe, suggesting some sequence non-specific binding activity. We conclude that NSY-7 protein can selectively bind the sequence CCTTAAC.

**The NSY-7 binding sequence CCTTAAC is required for asymmetric *srsx-3* expression**

The importance of the putative NSY-7 binding sequence in the *srsx-3* promoter was assessed by mutating an *srsx-3::GFP* reporter gene. Deletion of the CCTTAAC sequence eliminated all expression of *srsx-3::GFP* in AWC [Figure 2.5C(b,c), 2.5D(b,c)]. Further mutational analysis of this element and surrounding sequences in the *srsx-3* promoter was performed by deleting short sequence blocks in the binding site and surrounding sequence (Figure 2.5C, a-e) and by mutating individual nucleotides to residues that produced the lowest enrichment score in the protein-binding microarray experiment (Figure 2.5C, 1-11). Expression of these mutated elements in transgenic *C. elegans* demonstrated that the CCTTAAC sequence was embedded in a larger site with the ability to

either repress or activate AWC *srsx-3* expression, AATCCCTTAAC (Figure 2.5C-H). Mutation or deletion of the four bases preceding the NSY-7 binding site, or the first three bases of the site itself (deletions a-c, point mutations 1-7), resulted in loss of all *srsx-3* expression in AWC. Mutation of the following three bases of the NSY-7 binding site produced a phenotype resembling that of the *nsy-7* mutant, in which both AWCs expressed *srsx-3* (point mutations 8-10). Mutation of the last base of the site or bases beyond it had no effect. Thus, the region near the NSY-7 binding site is a compound element associated with both activating and repressing activities. The simplest explanation for these results is that the compound element binds NSY-7, which represses expression in $AWC^{ON}$, and also binds a second activating factor that is present in both $AWC^{OFF}$ and $AWC^{ON}$ (see discussion).

The extended activator/repressor sequence defined by mutagenesis of the *srsx-3* promoter, AATCCCTTAAC, occurs only 22 times in the *C. elegans* genome. To ask whether this sequence was associated with asymmetric AWC expression in other cases, we examined one additional gene in which this sequence appears 1.5 kb upstream of the transcriptional start site, *hlh-11*. *hlh-11* encodes one of ~35 helix-loop-helix transcription factors in the *C. elegans* genome (Okkema and Krause 2005); it was examined purely as a reporter gene, as the deletion mutant *hlh-11(ok2944)* has no effect on *str-2* or *srsx-3* expression in AWC (n=82). An *hlh-11::GFP* reporter with 3 kb of upstream sequence was expressed in one AWC neuron in 65% of animals, and in both AWC neurons in 30% of animals (Figure 2.7A, B); it was also expressed in a variety of additional cell types. The AWC neuron expressing *hlh-11* was significantly more likely to be $AWC^{OFF}$ than $AWC^{ON}$ (74% coexpression of *hlh-11::GFP* with *srsx-3::mCherry*, n=54). Expression

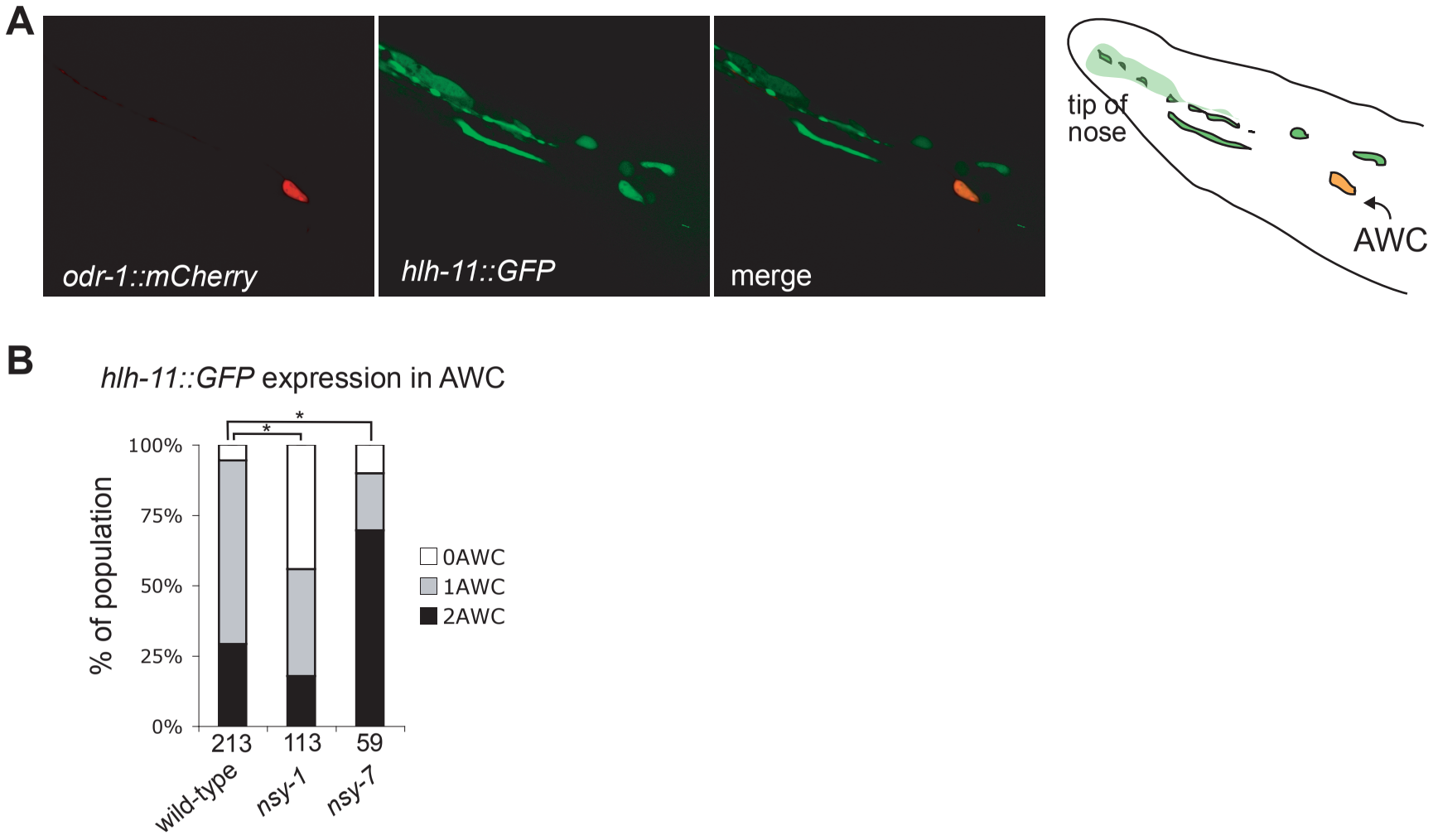



**Figure 2.7. *hlh-11* is regulated by *nsy-7*.** A, Expression of an *hlh-11::GFP* reporter with 3 kb of sequence upstream of the start site, showing colocalization with the AWC marker *odr-1::mCherry*. B, *hlh-11::GFP* expression in wild-type, *nsy-1(ky542)* and *nsy-7(tm3080)* mutant backgrounds. Asterisk indicates results different at P<0.001 (Chi-square test). Numbers of animals scored are indicated.

of *hlh-11::GFP* in AWC was less frequent in *nsy-1* (2-AWC$^{ON}$) mutants, and more frequent in *nsy-7* mutants, suggesting that *nsy-7* represses *hlh-11::GFP* in AWC$^{ON}$ (Figure 2.7B). These results suggest that the 11-nucleotide motif can predict AWC$^{OFF}$–biased expression in AWC neurons.

**Discussion**

The AWC neurons acquire asymmetric identities early in development, but require three sets of genes to maintain these properties into adulthood: (1) *nsy-7,* which is specifically required to maintain AWC$^{ON}$ properties (2) *odr-1* and *egl-4,* which are required to maintain markers for both AWC$^{ON}$ and AWC$^{OFF}$, and (3) *odr-3* and other G proteins, which increase *str-2* expression during larval stages (Troemel et al. 1999; Lans et al. 2004; Lans and Jansen 2006). *nsy-7* mutants initiate *str-2* expression but do not maintain AWC$^{ON}$ identity, as the adult animals do not express *str-2* or sense butanone. A *nsy-7::GFP* reporter is an early marker of the AWC$^{ON}$ neuron, suggesting that asymmetric expression of *nsy-7* is a transcriptional readout of the embryonic signaling pathway. The sequence-specific DNA-binding properties of NSY-7 suggest that it binds directly to promoters of genes expressed in AWC$^{OFF}$ to repress their expression in AWC$^{ON}$.

A complex signaling pathway involving a multicellular gap junction network, claudins, calcium channels, and protein kinases acts in embryogenesis to distinguish the AWC$^{ON}$ and AWC$^{OFF}$ cells (Troemel et al. 1999; Sagasti et al. 2001; Tanaka-Hino et al. 2002; Chuang and Bargmann 2005; Vanhoven et al. 2006; Bauer Huang et al. 2007; Chuang et al. 2007). Previous genetic studies and cell ablations suggest that the AWC$^{OFF}$ identity is a default, or ground state, and AWC$^{ON}$ is an induced state (Troemel et al. 1999). The developmental analysis of

asymmetric marker genes presented here provides a dynamic view of AWC$^{ON}$ induction (Figure 2.8). The bilateral early expression of the AWC$^{OFF}$ marker *srsx-3* is consistent with an AWC$^{OFF}$-like ground state. The first detectable marker of AWC$^{ON}$ identity is *nsy-7::GFP*, whose asymmetric expression beginning in the embryo overlaps with and anticipates the repression of the AWC$^{OFF}$ markers *srsx-3* and *hlh-11* in the L1 stage. Repression of *srsx-3* is followed by asymmetric expression of the AWC$^{ON}$ marker *str-2* in late L1 larvae.

The induction of *str-2* expression in AWC$^{ON}$ is not fully explained by the action of *nsy-7*, suggesting the existence of additional transcriptional regulators for AWC asymmetry. *str-2* expression is undetectable in null mutants for signaling genes such as *nsy-5*, but is detectable in L1 larvae of *nsy-7* null mutants, suggesting that another target of the signaling pathway initiates *str-2* expression in AWC$^{ON}$. *nsy-7; odr-1* double mutants also express *str-2* in the L1 stage, so the olfactory transduction pathway involving *odr-1* and *egl-4* is unlikely to represent this missing target. Moreover, *nsy-7* acts genetically to maintain *str-2* expression in AWC$^{ON}$, but there is no predicted NSY-7 binding site in the *str-2* promoter. As *str-2* activation in AWC$^{ON}$ is delayed compared to *srsx-3* repression, we speculate that *nsy-7* activates AWC$^{ON}$ markers indirectly, for example by repressing a transcriptional repressor. Further genetic screens may identify additional transcriptional effectors of AWC asymmetry.

NSY-7 binds to a specific sequence, CCTTAAC, that is present in the promoter region of the *srsx-3* and *hlh-11* target genes that are repressed in AWC$^{ON}$. Therefore, it is likely to act directly as a repressor of AWC$^{OFF}$ markers. This DNA binding specificity was determined by protein binding microarrays, which established the DNA binding ability of this non-conserved protein,

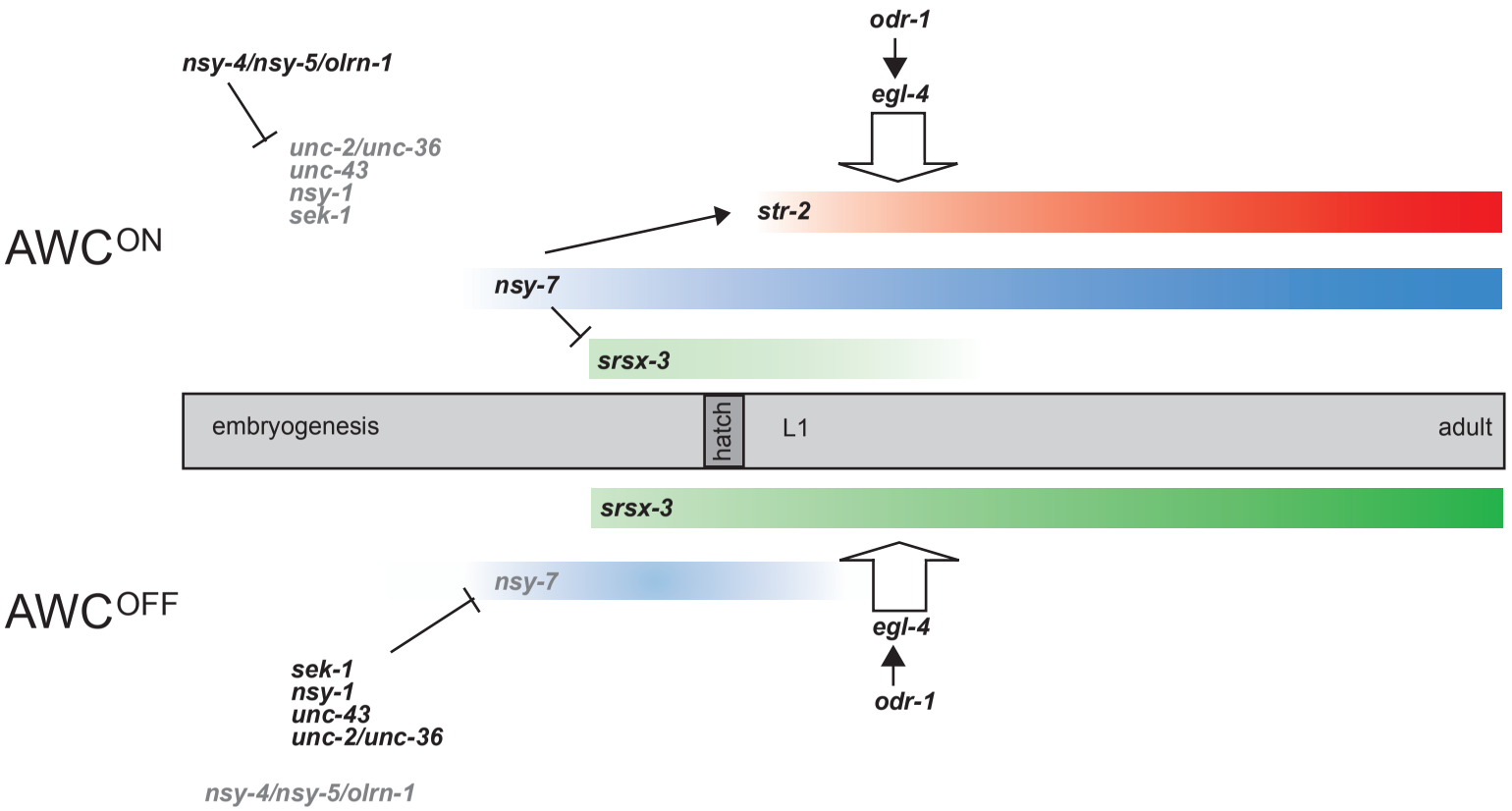



**Figure 2.8. Developmental stabilization of stochastic AWC asymmetry.** Gene expression patterns in AWC$^{ON}$ and AWC$^{OFF}$ across development are shown, along with inferred roles of the initial signaling cascade, the *nsy-7* transcriptional regulator, and the *odr-1*/*egl-4* activity-dependent pathway.

confirmed its sequence-specificity, and identified an optimal site in a single experiment. The DNA-binding specificities of 168 mammalian homeodomain proteins have been established using this technique, and the NSY-7 binding site was distinct from any known site (Berger et al. 2006; Berger et al. 2008). This result is not unexpected, given the divergence of the potential DNA-binding domain of NSY-7 from canonical homeodomains. The I/V47, Q/K50, and N51 residues that recognize the core TAAT in the homeodomain-binding motif are not present in NSY-7. However, the hydrophobic residues that stabilize the homeodomain protein core are conserved or retain hydrophobic character in NSY-7, and several residues that interact with the DNA backbone are present (Draganescu and Tullius 1998; Fraenkel et al. 1998; Chi 2005). The NSY-7 sequence is therefore consistent with a homeodomain-like structure that recognizes a noncanonical DNA sequence.

The NSY-7-binding motif falls within a larger element responsible for asymmetric *srsx-3* expression in $AWC^{OFF}$, AATCCCTTAAC. This 11-bp element includes sequences that promote symmetric AWC expression, which overlap partly but not completely with the NSY-7-binding site. The separation of activator and repressor functions by point mutations suggests that the motif binds at least two different proteins. We propose that the compound element binds the asymmetric repressor protein NSY-7 in $AWC^{ON}$, and an activator protein that is present in both AWC neurons. A candidate for the activator protein is CEH-36, a homeodomain protein that is expressed symmetrically in AWC neurons and required for AWC fates (Lanjuin et al. 2003; Koga and Ohshima 2004). *ceh-36* mutants fail to express both *str-2* or *srsx-3* in AWC, and the consensus binding site for Otx-class K50 homeodomains such as *ceh-36* is

TAATCC, which is similar to the AATCCC sequence that activates AWC expression of the *srsx-3* promoter (Affolter et al. 1990; Schier and Gehring 1993; Lanjuin et al. 2003; Chaney et al. 2005; Berger et al. 2008). NSY-7 might compete with CEH-36 or another activator for binding to the compound 11-bp element, specifically repressing expression in $AWC^{ON}$. This proposed mechanism resembles the function of the *Drosophila* Brinker protein in DPP signaling. DPP receptors regulate the activity of MAD/SMAD transcriptional activators, and regulate the expression of Brinker, a poorly-conserved repressor protein with a divergent homeodomain that competes with MAD proteins at MAD-binding DNA elements in target genes (Campbell and Tomlinson 1999; Jazwinska et al. 1999; Rushlow et al. 2001; Saller and Bienz 2001). An alternative model is that NSY-7 and the proposed activator protein can bind the compound element simultaneously, but that NSY-7 recruits transcriptional corepressors and converts activation into repression in $AWC^{ON}$.

Maintenance of AWC cell identity involves both asymmetric transcriptional regulation by *nsy-7*, which is controlled by a developmental signaling pathway, and symmetric transcriptional regulation by activity-dependent *odr-1*/*egl-4* cGMP signaling. Environmental odor cues cause the cGMP-dependent kinase EGL-4 to translocate to the nucleus, where it might interact with NSY-7 and other transcriptional regulators (L'Etoile et al. 2002). The convergence of developmental and activity-dependent regulators of neuronal function is of potential neurological and psychiatric interest. Disorders such as autism, schizophrenia, affective disorders, and epilepsy have genetic and developmental components, but are also influenced by experience, and manifest themselves in childhood, adolescence, or even adulthood (Charney 2004;

Keshavan 2004; Chubb et al. 2008; Silva 2008; Wigle and Eisenstat 2008). Analysis of neuronal maintenance pathways like those used in $AWC^{ON}$ may provide insight into some of these late-onset disruptions of neuronal function.

# Chapter 3

## Refining the mechanistic basis of NSY-7 activity in AWC

**Summary**

To understand the mechanism of action of the DNA-binding protein NSY-7, we wished to better characterize the timing of NSY-7 function and the nature of NSY-7 interaction partners. We found that adult AWC neurons are competent to respond to the transcriptional influence of NSY-7 to activate *str-2*, but that this response occurs only after a significant delay. One possible downstream effector of *nsy-7* function that may contribute to this delay is *hlh-11*, a transcriptional target of *nsy-7* that encodes a predicted DNA-binding protein. Overexpression of *hlh-11* in AWC has a significant effect on expression of both *str-2* and *srsx-3*, although the absence of a loss-of-function phenotype suggests that it may overlap functionally with other transcription factors. To gain a better understanding of NSY-7 function within the cell, we visualized the behavior of GFP-tagged NSY-7 in the presence of an extrachromosomal array carrying a known binding target, the *srsx-3* promoter, finding that NSY-7::GFP forms distinct patterns within the nucleus. Additional factors that might mediate NSY-7 function were sought by knocking down select components of the chromatin remodeling machinery using RNAi, but little effect was observed on *str-2* and *srsx-3* expression. Finally, experiments inspired by the observation that the NSY-7 DNA-binding domain contains weak similarity to methyl-binding domains demonstrated that NSY-7 is capable of binding to methylated DNA in addition to its specific target sequence CCTTAAC.

## Introduction

*nsy-7* encodes a DNA-binding protein that is an output of the embryonic pathway that establishes the asymmetric AWC$^{ON}$ and AWC$^{OFF}$ cell fates. It acts in AWC$^{ON}$ to repress the AWC$^{OFF}$ marker *srsx-3* and promote expression of the AWC$^{ON}$ marker, *str-2*. NSY-7 binds directly to the *srsx-3* promoter to repress transcription in AWC$^{OFF}$ (Figure 2.5). However, how NSY-7 achieves this repressive activity remains unclear. Furthermore, the mechanism by which *nsy-7* promotes *str-2* expression in AWC$^{ON}$ is entirely unknown: NSY-7 may have the ability to activate as well as to repress expression of direct targets, or it may act indirectly on *str-2* and other AWC$^{ON}$-specific genes by repressing another transcriptional repressor. In this chapter, I describe experiments that test hypotheses about the mechanisms by which NSY-7 controls expression of its target genes.

Transcription factors that, like NSY-7, promote the long-term stabilization of a cellular identity generally act in one of two ways. Some are continuously required to promote expression of target genes, while others act transiently during development to induce a stable transcriptional state. In the former case, a transcription factor may remain active at the promoters of its target genes to sustain their repression or activation. This kind of mechanism has been shown to operate in several contexts in *C. elegans*: the transcription factor CHE-1 is required continuously to maintain the identities of the ASE gustatory neurons, and likewise the homeodomain transcription factor MEC-3 is required to maintain its own expression and that of the genes required for determining

mechanosensory neuron fate (Xue et al. 1992; Etchberger et al. 2009). It is possible, therefore, that NSY-7 acts continuously to repress *srsx-3* and promote *str-2* expression.

Alternatively, cell type-specific transcription factors may recruit transcriptional repressor or activator complexes, as occurs in the maintenance of Hox gene effects by the Polycomb and Trithorax group complexes. In this situation, transient developmental activity of a particular transcription factor, such as *Drosophila Ubx,* is translated into a long-term effect by remodeling of the surrounding chromatin toward an inactive (promoted by the Polycomb complex) or active (promoted by the Trithorax complex) state (Chan et al. 1994; Pirrotta et al. 1995; Ringrose and Paro 2004; Schuettengruber et al. 2007). The Polycomb group (PcG) complex mediates repression primarily by inducing methylation of nearby histone H3 proteins at lysines K9 and K27, whereas other repressive complexes, such as the NuRD (Nucleosome Remodeling and Deacetylation) and Sin3 complexes induce histone H3 methylation and also promote histone deacetylation (Ahringer 2000; Schuettengruber et al. 2007; McDonel et al. 2009). PcG and NuRD complexes have both been shown to have repressive effects in *C. elegans*, so members of these complexes represent candidate mediators of NSY-7 activity (Unhavaithaya et al. 2002; Ross and Zarkower 2003; Zhang et al. 2003; Cui et al. 2006; Zhang et al. 2006; Kiefer et al. 2007).

Another mechanism of transcriptional stabilization is modification of DNA by methylation of cytosine nucleotides. Methyl-CpG binding domain (MBD) proteins are well-known mediators of inactive chromatin states in vertebrates. While different MBD proteins bind different specific cofactors, they almost universally recruit histone deacetylases and ATP-dependent nucleosome

remodeling enzymes, which together act to make a given region of DNA inaccessible to the transcriptional machinery (Ballestar and Wolffe 2001; Fan and Hutnick 2005). Weak similarity between the NSY-7 DNA-binding domain and methyl-CpG binding domains, together with the analogous repressive functions of NSY-7 and the MBD proteins, led us to investigate the ability of NSY-7 to act as a methyl-binding factor.

In the studies presented here, we begin to shed light on the details of NSY-7 activity in regulation of the $AWC^{ON}$ cell identity. In most cases, however, our results are more suggestive than conclusive. While we now have intriguing clues about how NSY-7 functions, the details of its mechanism of action remain largely unknown.

## Results and Discussion

### I. Induction of NSY-7 expression by heat shock rescues *str-2* expression after a 24-hour delay

#### Results

The temporal requirement for NSY-7 in regulation of *str-2* and *srsx-3* expression was examined by inducing expression of NSY-7 by heat shock in a *nsy-7(tm3080)* null mutant background. Heat shock expression of NSY-7 was induced at different developmental stages, the worms were returned to physiological temperatures, and *str-2* and *srsx-3* expression were scored after the animals had reached adulthood (Figure 3.1A). Because the *hsp-16.2* promoter is

expressed in both AWCs, NSY-7 should be overexpressed in both AWCs and in other cells after induction, rather than regaining its normal asymmetric AWC$^{ON}$ expression pattern (Stringham et al. 1992).

Embryonic heat shock caused the majority of *hsp16.2::NSY-7* array-bearing animals to arrest development as embryos or L1s. This toxicity may result from the ectopic expression of NSY-7 in tissues in which it is not normally expressed. Unfortunately, for this reason it was not possible to evaluate the effect of NSY-7 on *srsx-3* or *str-2* expression during embryogenesis.

When NSY-7 expression was induced during larval stages, there was very little effect on adult *srsx-3* expression, suggesting but not proving an embryonic requirement for NSY-7 in *srsx-3* repression. By contrast, *str-2* expression was somewhat sensitive to postembryonic NSY-7 expression, as *str-2* was weakly induced in adults when heat shock was performed during the first larval stage (Figure 3.1A). Induction of NSY-7 in L4s or young adults had little effect on *str-2* expression following two hours of recovery at 20°C, but a 24-hour recovery period rescued *str-2* expression to a similar extent as did heat shock during the L1 larval stage (Figure 3.1A). This result implies that the developmental stage of NSY-7 expression may not be as important for the induction of *str-2* as the recovery time after NSY-7 induction.

To further examine this phenomenon, NSY-7 was induced in young adults that were then allowed to recover at 20°C for various periods of time before scoring (Figure 3.1B). These experiments revealed that NSY-7 induction had a profound but delayed effect on *str-2* expression. *str-2* first became apparent at ~24 hours following heat shock, and nearly 100% of animals expressed *str-2* in one or both AWCs by 36 hours after heat shock (Figure 3.1B). *str-2* expression

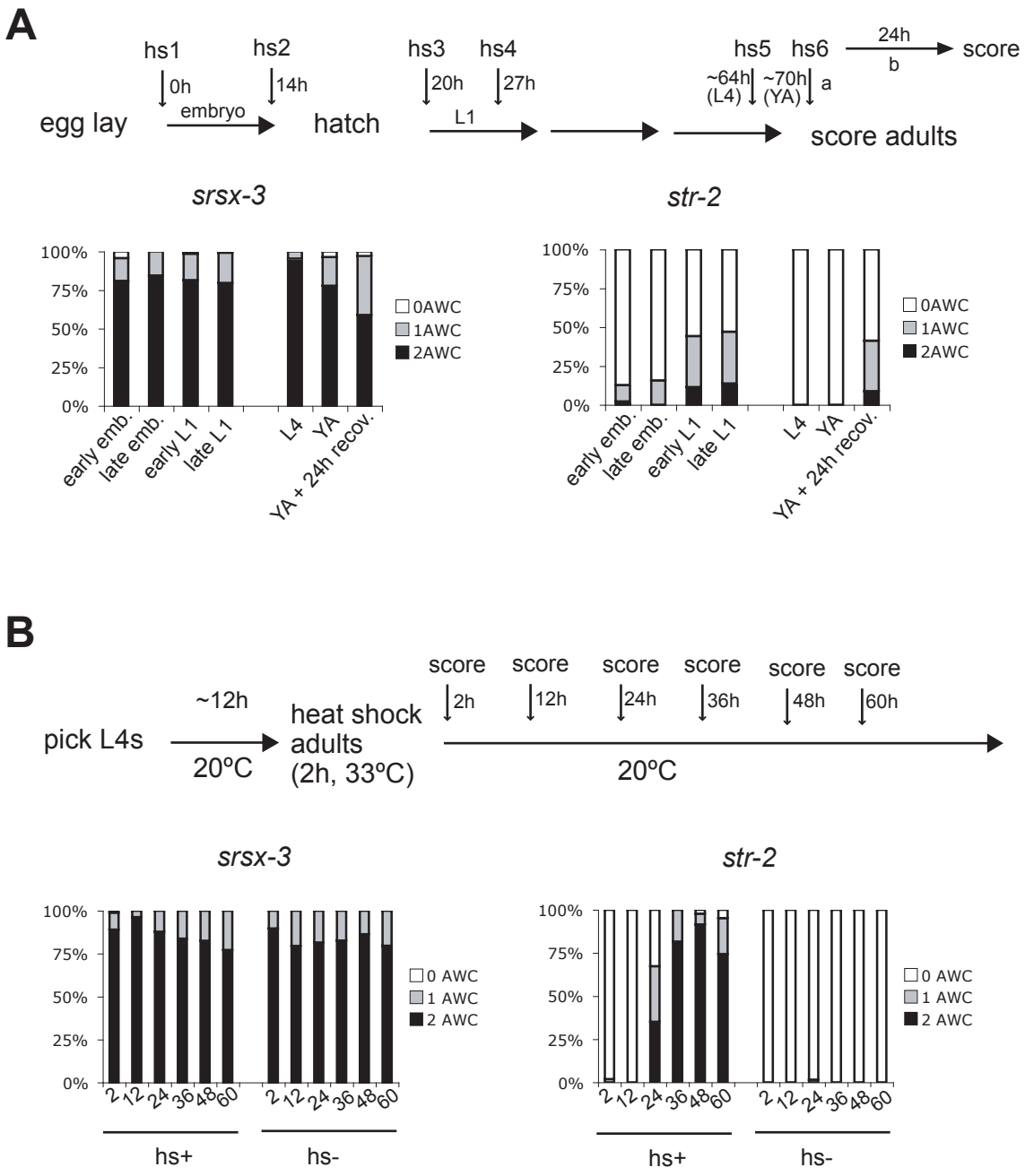



**Figure 3.1. NSY-7 heat shock protocols and results.** A, Diagram of developmental heat shock protocol (top) and *str-2* and *srsx-3* expression (bottom). *str-2* and *srsx-3* were scored ~70 hours after hatching, when worms had reached adulthood. The final heat shock (hs6) was done at 70 hours and scored 2 hours or 24 hours later. B, Diagram of adult heat shock and scoring protocol (top) and *str-2* and *srsx-3* expression (bottom). Young adult worms were heat shocked and scored after the marked intervals. hs+, array-bearing animals after heat shock. hs-, array-bearing animals that did not undergo heat shock. Array-negative animals that underwent heat shock looked similar to the hs- condition (data not shown).

was maintained until at least 60 hours after induction. As observed in the developmental heat shocks, NSY-7 induction had little effect on *srsx-3* expression at any time point.

The 24-hour delay to *str-2* expression is only partially explained by the slow maturation time of the dsRed2 fluorophore used as a reporter: detection time by microscopy for dsRed2 is reported to be twelve to sixteen hours (Clontech documentation), but expression was not evident in AWC for an additional twelve hours after this interval.

**Discussion**

The ability of the AWCs to respond to NSY-7 expression during adulthood by upregulating *str-2* expression indicates that an irreversible decision has not yet been made regarding transcription at the *str-2* promoter. Rather, the adult neurons are still capable of altering their GPCR expression repertoire and potentially their overall fate following the appearance of an inducing factor, NSY-7. However, it is significant that these experiments were conducted in a *nsy-7* mutant background, in which the identities of the AWC neurons in adults may be more receptive to NSY-7 expression than they are in wild-type worms. It will be interesting to perform the same set of experiments in a wild-type background, to determine whether *str-2* can be induced in $AWC^{OFF}$ when development has progressed normally. Another question is whether this effect is specific to *str-2* or whether it represents a more general effect on cell fate. To address this question, it will be important to examine the response of additional $AWC^{ON}$ and $AWC^{OFF}$-specific markers to heat shock induction.

It is enticing to make the connection between the 24-36 hour time delay between heat shock induction of NSY-7 and expression of *str-2*, and the approximate time interval in wild-type larvae between hatching and robust *str-2* expression in $AWC^{ON}$. It is unclear exactly when in development *nsy-7* expression normally begins; however, it is present in the embryo at least by the three-fold stage (Figure 2.4B). The onset of stable *str-2* expression in wild-type larvae occurs around the L1-L2 transition, about 14-15 hours after hatching. This time interval (at least 15-20 hours between the beginning of *nsy-7* expression and the onset of *str-2* expression) may therefore be slightly shorter than that observed in our heat shock experiments, but both the 15-20 hour and 24-36 hour intervals are significantly longer than would be expected for direct transcriptional activation of *str-2* by NSY-7. This interval may represent a delay in translation of the NSY-7 or STR-2 transcripts, or it may be due to the expression of intermediate transcriptional regulators in the regulation of *str-2* expression. The latter possibility is supported by the apparent absence of NSY-7 binding sites in the *str-2* promoter, and by the fact that NSY-7 has known transcriptional repressor activity at the *srsx-3* promoter and at a modified *odr-1* promoter (Figures 2.5, 4.8), but has an activating effect on *str-2* expression.

The lethality of *hsp16.2::NSY-7* after embryonic heat shock precludes an accurate picture of when NSY-7 is required to stabilize $AWC^{ON}$ and $AWC^{OFF}$ identities. No effect of postembryonic NSY-7 induction on *srsx-3* expression was observed under any condition. This negative result may be due to perdurance of the GFP fluorophore or to variable expression of NSY-7 after heat shock induction. The timing of NSY-7 action can be better addressed by eliminating rather than inducing NSY-7 expression during development. For example, a

loxP-containing NSY-7 transgene could be expressed in AWC to rescue the *nsy-7* mutant, and the Cre recombinase could then be induced at various time points, with the expectation that *srsx-3* expression should be induced upon loss of NSY-7. This experiment would reveal a time window in which NSY-7 is required, bypassing the potential GFP perdurance and expression level problems associated with heat shock rescue of *srsx-3*. Alternatively, RNAi could be performed against NSY-7 at different developmental stages; this approach has been successful for targeting the transcription factor CHE-1 in the ASE gustatory neurons postembryonically (Etchberger et al. 2009) and would allow the experiment to be conducted in the context of wild-type levels of NSY-7, rather than in a rescued mutant background.

## II. Regulation of transcription factor expression by NSY-7

**Results**

Transcription factors involved in maintenance of cell fate are frequently autoregulatory, especially if they are required continuously to maintain the differentiated state. To ask whether NSY-7 is required to maintain its own asymmetric expression in AWC, a *nsy-7::GFP* reporter was examined in a *nsy-7(tm3080)* null background. The promoter contained a predicted NSY-7 binding site approximately 20kb upstream of the ATG. *nsy-7::GFP* expression in *nsy-7* mutants was occasionally de-repressed in AWC$^{OFF}$, but the overall expression pattern was asymmetric, as it is in wild-type animals (Figure 3.2A). Thus, no clear autoregulatory effect was observed.

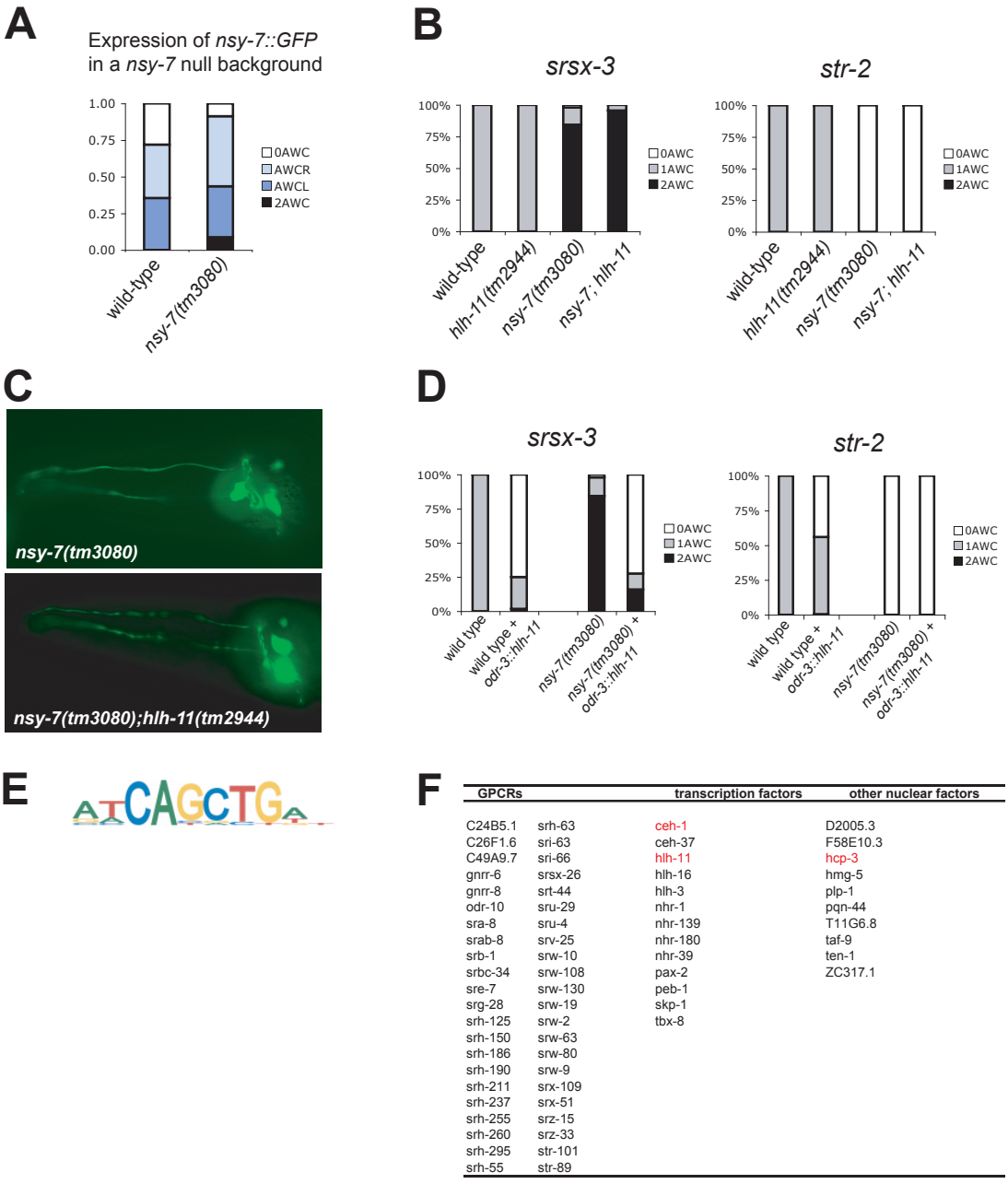
## Table 1: Selected NSY-7 target promoters*

| GPCRs | transcription factors |
|---|---|
| C53C7.1 | attf-1 |
| glb-33 | ceh-1 |
| R11F4.2 | ceh-17 |
| sprr-2 | daf-16 |
| sra-20 | egl-38 |
| sra-26 | egrh-3 |
| srbc-43 | hcp-3 |
| sre-34 | his-42 |
| srg-1 | hlh-11 |
| srh-195 | ldb-1 |
| srh-277 | mut-14 |
| sri-33 | nhr-103 |
| sri-7 | nhr-19 |
| srt-24 | nhr-190 |
| srt-45 | pqn-67 |
| srt-47 | sox-2 |
| srx-38 | |
| Y48C3A.11 | |

*see Methods
red: shared with HLH-11

The delay between NSY-7 expression and its effects on *str-2*, combined with the fact that no close match to the NSY-7 binding site is found in the *str-2* promoter, implies that additional transcription factors may act downstream of NSY-7 to regulate target genes such as *str-2*. Consistent with this idea, several genes containing the NSY-7 binding site CCTTAAC within 300 base pairs of their transcriptional start sites are known or predicted transcription factors (Table 3.1) (see Methods for identification of targets).

*hlh-11*, a helix-loop-helix transcription factor, contains a NSY-7 binding site ~2kb upstream of its transcriptional start site and its expression is repressed by NSY-7 in $AWC^{ON}$ (Figure 2.7B). *hlh-11* is therefore a candidate downstream transcriptional effector of NSY-7. The function of *hlh-11* was evaluated by examining the phenotype of a *hlh-11(tm2944)* deletion mutant, which is predicted to be a null allele. The *hlh-11(tm2944)* mutant had a wild-type phenotype for both *str-2* and *srsx-3* expression (Figure 3.2B). A *nsy-7(tm3080);hlh-11(tm2944)* double mutant had a *nsy-7*-like phenotype, indicating that the *hlh-11* mutant was unable to suppress *nsy-7* defects in *str-2* and *srsx-3* expression (Figure 3.2B). Interestingly, *hlh-11(tm2944)* did suppress a low-penetrance morphological defect in the *nsy-7(tm3080)* mutant: one AWC cell was sometimes displaced anteriorly in *nsy-7(tm3080)*, but the *nsy-7(tm3080);hlh-11(tm2944)* double mutant was always morphologically wild type (Figure 3.2C). It is possible that rescue of this defect was due to loss of an unrelated mutation during the cross; however, the *nsy-7(tm3080)* mutant had been previously outcrossed four times, and rescue of this defect was not evident in other double mutants containing the *nsy-7(tm3080)* mutation.




| GPCRs | | transcription factors | other nuclear factors |
|---|---|---|---|
| C24B5.1 | srh-63 | ceh-1 | D2005.3 |
| C26F1.6 | sri-63 | ceh-37 | F58E10.3 |
| C49A9.7 | sri-66 | hlh-11 | hcp-3 |
| gnrr-6 | srsx-26 | hlh-16 | hmg-5 |
| gnrr-8 | srt-44 | hlh-3 | plp-1 |
| odr-10 | sru-29 | nhr-1 | pqn-44 |
| sra-8 | sru-4 | nhr-139 | T11G6.8 |
| srab-8 | srv-25 | nhr-180 | taf-9 |
| srb-1 | srw-10 | nhr-39 | ten-1 |
| srbc-34 | srw-108 | pax-2 | ZC317.1 |
| sre-7 | srw-130 | peb-1 | |
| srg-28 | srw-19 | skp-1 | |
| srh-125 | srw-2 | tbx-8 | |
| srh-150 | srw-63 | | |
| srh-186 | srw-80 | | |
| srh-190 | srw-9 | | |
| srh-211 | srx-109 | | |
| srh-237 | srx-51 | | |
| srh-255 | srz-15 | | |
| srh-260 | srz-33 | | |
| srh-295 | str-101 | | |
| srh-55 | str-89 | | |

**Figure 3.2. *nsy-7* autoregulation and *hlh-11* function.** A, Expression of a *nsy-7* promoter driving GFP in *nsy-7(tm3080)* mutants. B, *str-2* and *srsx-3* expression in *hlh-11(tm2944)* and *hlh-11(tm2944);nsy-7(tm3080)* mutants. C, AWC morphology defect in *nsy-7(tm3080)* animals and wild-type AWC morphology in *nsy-7(3080); hlh-11(tm2944)* animals. Arrowheads point to AWC cell body. D, *srsx-3* and *str-2* phenotypes in animals expressing the *hlh-11* genomic sequence under the AWC-selective *odr-3* promoter. E, Binding site for HLH-11 predicted using protein-binding microarrays (Grove et al. 2009). F, List of GPCRs and transcriptional regulators containing the predicted HLH-11 binding site within 500 base pairs upstream of the coding start site (Grove et al. 2009). Red, predicted target shared with NSY-7.

The effects of overexpressing *hlh-11* were examined by injecting the *hlh-11* genomic sequence under the control of the AWC-selective *odr-3* promoter; the genomic sequence was used instead of a cDNA because *hlh-11* encodes at least ten predicted splice forms. Overexpression of *hlh-11* had an effect on both *srsx-3* and *str-2* expression, strongly repressing *srsx-3* and partially repressing *str-2* (Figure 3.2D). Overexpression of the *hlh-11* genomic sequence in a *nsy-7(tm3080)* mutant background had an equally strong effect on *srsx-3* expression, repressing it in both AWCs, but had no effect on *str-2* expression (which was already absent in the *nsy-7(tm3080)* background) (Figure 3.2D). *hlh-11* is therefore not required for *str-2* and *srsx-3* expression in AWC, but nevertheless has a strong activity when overexpressed that may indicate a repressive function in wild-type animals.

**Discussion**

NSY-7 was not required to maintain its own asymmetric expression in AWC. This result argues that NSY-7 must recruit additional transcription factors to maintain the asymmetric expression of its direct targets, and suggests that some of its long-term effects may be mediated by downstream transcription factors such as HLH-11.

HLH-11 is a strong candidate for a downstream effector of NSY-7. It is a transcriptional target of NSY-7, which represses it in $AWC^{ON}$ and therefore biases its expression toward $AWC^{OFF}$. An *hlh-11* deletion mutant had no effect on *srsx-3* and *str-2* expression, either by itself or in the context of a *nsy-7* deletion mutant. However, overexpression of *hlh-11* in AWC had a strong repressive effect on expression of both markers. Some wild-type worms express *hlh-11* in both

AWCs, while the remainder express *hlh-11* in AWC$^{OFF}$ only (Figure 2.7); it is possible, therefore, that the AWC overexpression phenotype of *hlh-11* corresponds to a natural function of the protein in one or both AWCs. If so, the wild-type phenotype of the mutant indicates that *hlh-11* is likely to act redundantly with other regulators.

Recently, Grove et al. published a study of the binding sites, dimerization partners, and predicted transcriptional targets of all *C. elegans* HLH genes, including *hlh-11* (Grove et al. 2009). They found that HLH-11 is likely to act as a homodimer, based on its ability to bind DNA when expressed in the absence of any other HLH proteins, and identified a target DNA sequence (Figure 3.2E). The predicted targets of *hlh-11*, defined as those genes containing a binding site within the 500 base pairs upstream of their start sites, are enriched for rhodopsin-like GPCRs (Figure 3.2F). This enrichment suggests that *hlh-11* may represent an intermediary between *nsy-7* and regulation of GPCR expression.

The predicted transcriptional targets of HLH-11 include several transcriptional regulators, one of which is HLH-11 itself, implying an autoregulatory role for the protein. NSY-7 is also predicted to bind the promoters of several transcription factors, and indeed at least two transcription factor promoters are predicted targets of both NSY-7 and HLH-11: *ceh-1* and *hlh-11* itself. The transcriptional regulatory network controlling the identities of the two AWCs may therefore be complex and highly interconnected. Examination of the roles played by some of the predicted transcriptional targets of NSY-7 and HLH-11 will be important to better define this network. Additionally, it will be useful to see whether HLH-11 overexpression can exert its strong repressive effect on additional GPCRs expressed asymmetrically in the AWCs.

## III. NSY-7::GFP forms subnuclear structures that may be cell type-specific

**Results**

Localization of a GFP-tagged NSY-7 protein was examined in AWC and intestinal nuclei. NSY-7 is normally expressed in both of these tissues; it has an established regulatory function in AWC, and the size and visibility of intestinal nuclei made the intestine an appealing tissue for comparison. Using a 'spot analysis', NSY-7::GFP was expressed under a cell-specific promoter, along with a separate array containing a Gal4 binding domain (Gal4BD)::mCherry expression vector with the same cell-specific promoter, the Gal4 UAS, and a NSY-7 target, the *srsx-3* promoter (modified from (Dawes et al. 1999; Chu et al. 2002). Ideally, the mCherry-tagged Gal4BD should bind to the Gal4 UAS in the second array, thereby marking the location of that array within the nucleus. The localization of NSY-7::GFP relative to the array can then be examined; if it binds to its target, *srsx-3*, in this context, it should colocalize with the Gal4BD::mCherry spot (Figure 3.3A). At least two independent NSY-7::GFP arrays and at least two independent Gal4BD::mCherry arrays were tested in different combinations for each cell type, and similar effects were observed for different combinations of arrays in each cell type. Therefore, the differences in appearance described below are likely to be due either to specific effects of the *elt-2* and *odr-3* promoters, or to cell type-specific factors.

NSY-7::GFP formed distinct subnuclear structures both in the presence and absence of the Gal4 array. These structures appeared different in the AWCs

than in the intestinal nuclei. In the intestine, NSY-7::GFP formed clumps at the nuclear periphery, whereas in AWC it formed tight, dot-like clusters at various locations in the nucleus (Figure 3.3B-E). In both cases, clusters became more defined at older developmental stages, becoming most prominent in adults (Figure 3.3D,E).

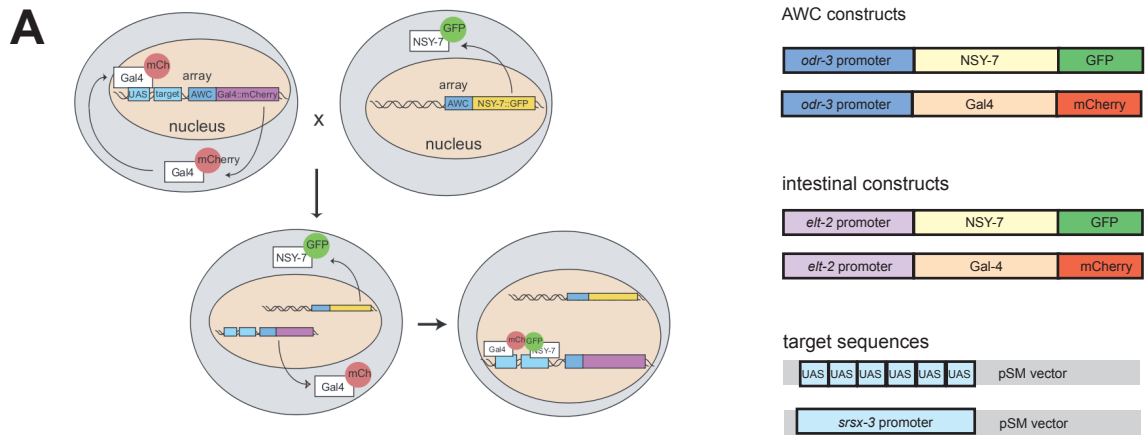
The appearance of the Gal4BD::mCherry protein and its relationship to NSY-7::GFP also differed between cell types. In the intestine, Gal4BD::mCherry usually formed tight, localized spots at the nuclear periphery, while in AWC it often invaded a greater proportion of the total nuclear area (Figure 3.3B-E). Unlike NSY-7::GFP, Gal4BD::mCherry localization did not become significantly more defined depending on the age of the animal.

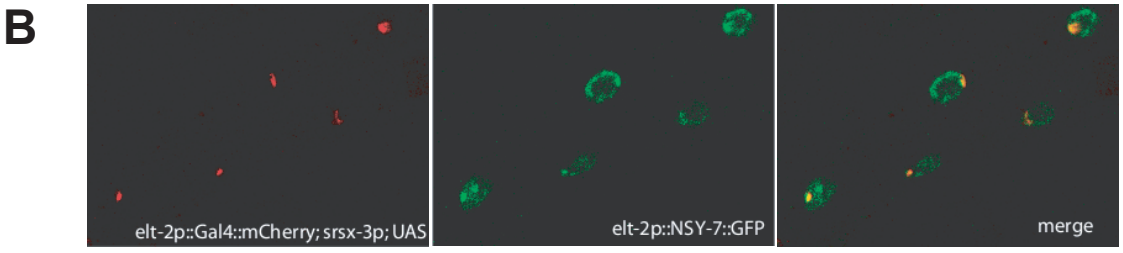
In the intestine, NSY-7::GFP frequently covered most of the area occupied by Gal4BD::mCherry in addition to regions of the nucleus in which Gal4BD::mCherry was not found (Figure 3.3B). In AWC, NSY-7::GFP sometimes partly colocalized with Gal4BD::mCherry in animals of intermediate age (L2-L3) (Figure 3.3C), but in older animals was frequently found interwoven with but not overlapping the Gal4::mCherry cluster (Figure 3.3D), and sometimes occupied an entirely separate part of the nucleus (Figure 3.3E).

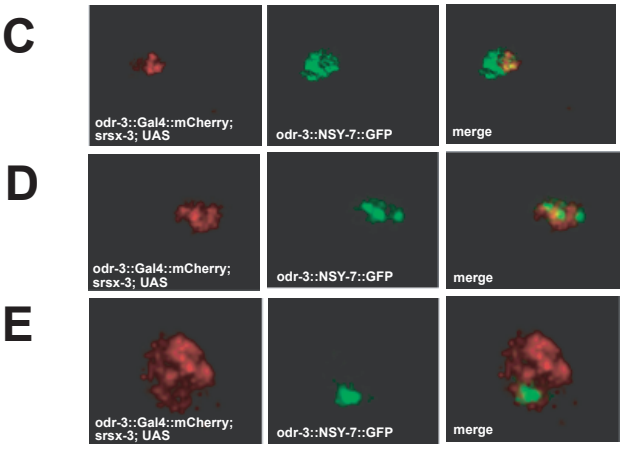
We conclude that both the extrachromosomal arrays and NSY-7::GFP exhibit cell-type specific localization. However, the behavior of NSY-7::GFP and Gal4BD::mCherry in these experiments, while interesting, was difficult to interpret. The NSY-7::GFP-expressing arrays contained cell-specific promoters: *odr-3* in the AWCs and *elt-2* in the intestine. Either or both of these promoters might recruit NSY-7, altering the arrangement of NSY-7::GFP in the nucleus. Furthermore, multi-copy arrays may be heterochromatic, increasing the

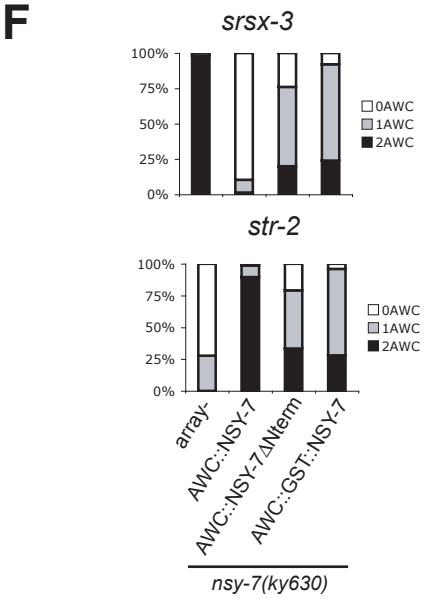
**Figure 3.3. Visualization of NSY-7::GFP using spot analysis.** A, left, Diagram of the spot analysis technique and arrays used for expression in intestinal cells or AWC. Right, vectors used for expression of NSY-7::GFP and Gal4::mCherry in AWC under the *odr-3* promoter (top) or in intestinal cells under the *elt-2* promoter (middle), and for expression of the *srsx-3* and Gal4 UAS target sequences (bottom). B, Expression of Gal-4::mCherry and NSY-7::GFP in intestinal nuclei under the *elt-2* promoter. C-E, Expression of Gal-4::mCherry and NSY-7::GFP in AWC in three different animals under the *odr-3* promoter, illustrating range of observed patterns. C, L3 animal. D,E, adult animals. F, Overexpression in AWC of wild-type NSY-7, NSY-7 lacking the first seven amino acids after the methionine (NSY-7DNterm), and GST::NSY-7 in a nsy-7(ky630) mutant background.

Figure 3.3


A
mCh
Gal4
array
UAS
target
AWC
Gal4::mCherry
nucleus
mCherry
Gal4
x
GFP
NSY-7
array
AWC
NSY-7::GFP
nucleus
GFP
NSY-7
mCh
Gal4
AWC constructs
odr-3 promoter
NSY-7
GFP
odr-3 promoter
Gal4
mCherry
intestinal constructs
elt-2 promoter
NSY-7
GFP
elt-2 promoter
Gal-4
mCherry
target sequences
UAS
pSM vector
srsx-3 promoter
pSM vector


B
elt-2p::Gal4::mCherry; srsx-3p; UAS
elt-2p::NSY-7::GFP
merge


C
odr-3::Gal4::mCherry; srsx-3; UAS
odr-3::NSY-7::GFP
merge
D
odr-3::Gal4::mCherry; srsx-3; UAS
odr-3::NSY-7::GFP
merge
E
odr-3::Gal4::mCherry; srsx-3; UAS
odr-3::NSY-7::GFP
merge


F
srsx-3
100%
75%
50%
25%
0%
0AWC
1AWC
2AWC
str-2
100%
75%
50%
25%
0%
0AWC
1AWC
2AWC
array-
AWC::NSY-7
AWC::NSY-7ΔNterm
AWC::GST::NSY-7
nsy-7(ky630)

proportion of DNA in the nucleus that is in an inactive state and possibly altering the interaction of NSY-7 with the array. Thus, in addition to the unknown effects of the *odr-3* and *elt-2* promoters, the presence of any array may change the transcriptional milieu of the nucleus and may significantly perturb nuclear function.

It is possible that the formations adopted by these NSY-7::GFP clusters were artefacts caused by aggregation of an overexpressed protein. However, the C-terminally tagged NSY-7 used in this experiment is known to be functional, based its the overexpression phenotype in the presence of a *str-2::dsRed2* reporter. When the NSY-7 protein was tagged with GFP at the N-terminus rather than the C-terminus, nuclear GFP appeared smooth and regular, without the clumpy appearance evident with the C-terminally tagged protein. Interference with the N-terminal end of NSY-7, either by deleting the seven most N-terminal amino acids or by tagging the N-terminus with a GST protein, results in reduced protein function: NSY-7(DNterm) and GST::NSY-7 constructs were less capable of inducing an overexpression phenotype in a *nsy-7(ky630)* mutant background than the wild-type protein (Figure 3.3F). However, we did not evaluate whether the GFP::NSY-7 protein exhibited the same reduced overexpression phenotype as NSY-7(DNterm) and GST::NSY-7. Thus, some activity of the NSY-7 N-terminus may contribute to protein function and clustering of the protein, but this possibility should be more thoroughly tested.

**Discussion**

NSY-7::GFP forms subnuclear structures that may be related to its function. However, the details of these preliminary experiments make it difficult

to draw specific conclusions about the significance and implications of these structures. Most problematic is the presence of the NSY-7::GFP-expressing array within the nucleus, which may interfere with or alter NSY-7 function. A cleaner way to examine NSY-7 subnuclear localization would be by antibody staining, which would eliminate the problems associated with expression of the transgenic array. Antibody staining would also assay NSY-7 localization at endogenous levels and without the GFP tag, reducing possible aggregation artifacts due to the GFP tag or to overexpression.

To separate the possible effects of cell type-specific factors in the AWC and intestinal cells from the effects of the promoters used to drive NSY-7::GFP and Gal4BD::mCherry, it would be useful to repeat these experiments using alternative promoters in each cell type. One possibility is to drive AWC expression of NSY-7::GFP and Gal4::mCherry using an *srsx-3* promoter whose NSY-7 binding site has been removed; removal of this binding site greatly reduces the interaction of the promoter with NSY-7 (Figure 2.5). If similar results are observed using multiple promoters in the same cell type, it is more likely that the array behavior is due to the action of cell-type-specific factors. It is also possible to physically separate cell-specific promoters from target promoters by expressing NSY-7::GFP and Gal4BD::mCherry from one array, and *srsx-3* and the Gal4 UAS on a different array. Separating cell-specific promoters in this manner may help to define the specificity of Gal4BD::mCherry and NSY-7::GFP binding.

Although more remains to be done to demonstrate conclusively that NSY-7 localizes to specific regions within the nucleus, the possibility that it does so encourages speculation about the possible function of such clusters. Clustering may localize the protein to a nuclear compartment with a specific transcriptional

function. PcG complexes are distributed in speckles in the nucleus, and association of PcG target elements enhances silencing by the complex (Bantignies et al. 2003). Subnuclear 'granules' consisting of clustered genomic loci have also been reported in which clustered loci are activated rather than repressed; in at least one case, this clustering requires the histone demethylase LSD1 (Hu et al. 2008). If the NSY-7::GFP nuclear clusters observed are real, it will be interesting to identify any additional proteins and genomic loci that are also found in these structures. Preliminary experiments hint that the NSY-7 N-terminal domain might be involved in cluster formation; determining the function of this domain may therefore also be important in elucidating the mechanism by which NSY-7 exerts its transcriptional influence.

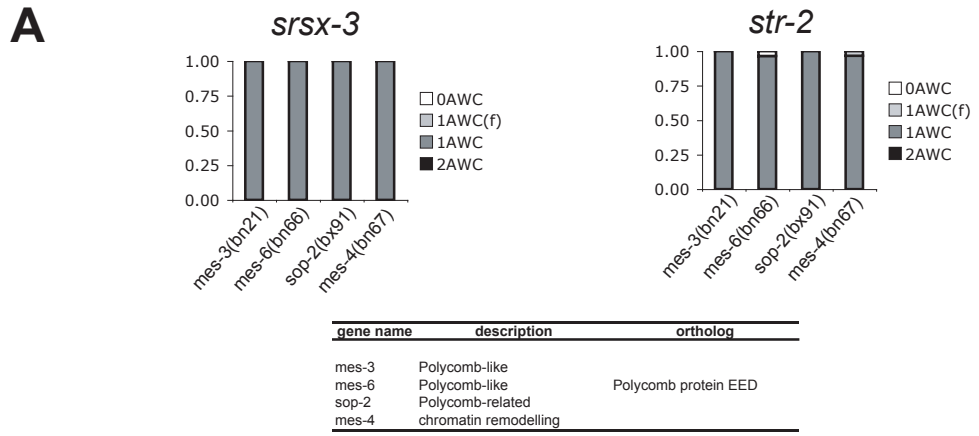
## IV. Known Polycomb and NuRD complex genes are not required for maintenance of the $AWC^{ON}$ and $AWC^{OFF}$ identities

**Results**

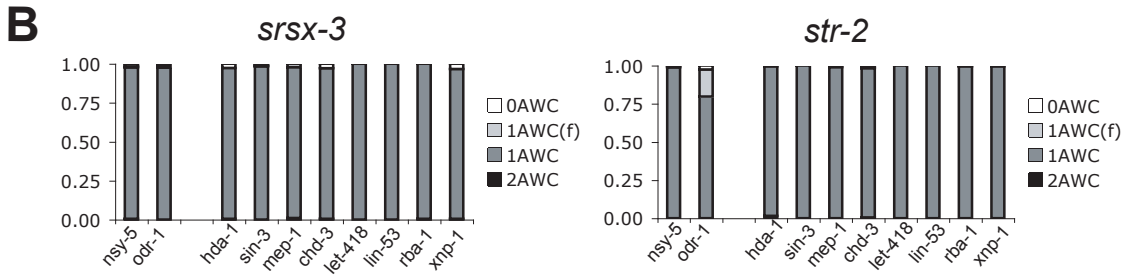
Chromatin remodeling factors can act to convert a transient transcriptional effect into long-term repression or activation of target genes. The pathway that establishes AWC asymmetry acts late in the embryo and is maintained throughout life, prompting an evaluation of the AWC phenotypes of worms defective for components of the chromatin remodeling machinery. Because NSY-7 acts as a repressor, the Polycomb and NuRD (Nucleosome Remodeling and Deacetylase) complexes, both known to exert repressive effects on their target genes, were strong candidates. Other genes thought to be involved in chromatin remodeling and modification were also examined.

The AWC markers *str-2* and *srsx-3* were examined in mutants for the Polycomb and Polycomb-related genes *mes-3*, *mes-6*, and *sop-2*, as well as *mes-4*, which interacts with the *mes-3* and *mes-6* gene products and is thought to be involved in chromatin remodeling, but is not similar to any known Polycomb genes (Xu et al. 2001; Unhavaithaya et al. 2002; Ross and Zarkower 2003; Zhang et al. 2003; Bender et al. 2006). Available alleles for all of these genes are missense mutations that are not null but are thought to be loss-of-function alleles. *mes-3*, *mes-4*, and *mes-6* have a maternal-effect sterility phenotype, so animals scored were the homozygous progeny of homozygous mutants to account for any possible maternal effect. All of these mutants had a wild-type phenotype for both *str-2* and *srsx-3* expression (Figure 3.4A).

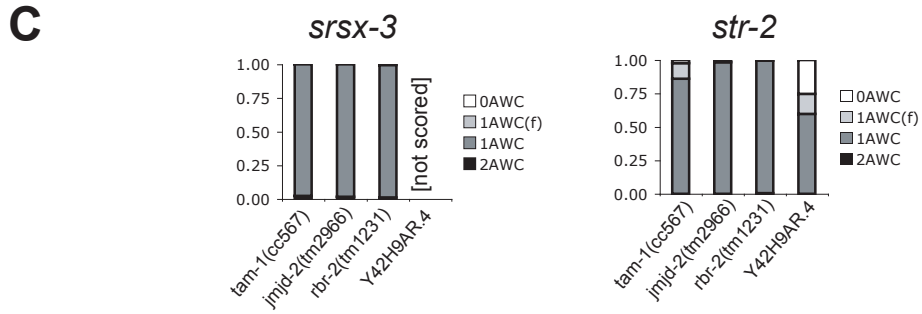
Many NuRD complex genes are required for viability. Therefore, these genes were evaluated by feeding RNAi for the genes to L1 larvae. This approach bypasses the lethality of early loss of function of these genes, but might still reveal defects in *str-2* and *srsx-3* expression if the RNAi targets are required continuously to stabilize $AWC^{ON}$- and $AWC^{OFF}$-specific gene expression. Because *C. elegans* neurons are insensitive to RNAi, all experiments were performed in an *eri-1;lin-15B* sensitized mutant background. If the NuRD complex acts with NSY-7 to repress its targets, knockdown should produce a *nsy-7*-like phenotype, with *srsx-3* expressed in both cells and *str-2* expression reduced or absent. However, little or no effect was seen following knockdown of NuRD complex components. Likewise, two genes orthologous to NuRD complex components in other species but not known to participate in the *C. elegans* NuRD complex, *chd-3*/CHD3 and *rba-1*/RbAp46, had little or no effect on *str-2* and *srsx-3* expression (Figure 3.4B). Knockdown of *xnp-1*, an ATP-dependent SNF2 DNA helicase that interacts with




| gene name | description | ortholog |
|---|---|---|
| mes-3 | Polycomb-like | |
| mes-6 | Polycomb-like | Polycomb protein EED |
| sop-2 | Polycomb-related | |
| mes-4 | chromatin remodelling | |




| gene name | description | ortholog |
|---|---|---|
| hda-1 | histone deacetylase, NuRD complex | HDAC1 |
| sin-3 | histone deacetylase subunit, NuRD complex | Sin3A |
| mep-1 | zinc finger protein, NuRD complex | - |
| chd-3 | SNF2 nucleosome remodelling, NuRD complex | CHD3 |
| let-418 | SNF2 nucleosome remodelling, NuRD complex | Mi2 |
| lin-53 | nucleosome remodelling, NuRD complex | RbAp48 |
| rba-1 | nucleosome remodelling, NuRD complex | RbAp46 |
| xnp-1 | SNF2 family DNA helicase | - |




| gene name | description | ortholog |
|---|---|---|
| tam-1 | chromatin remodelling? | |
| jmjd-2 | histone demethylase | jmjC-domain protein 3A |
| rbr-2 | H3K4 demethylase | JARID1A-D |
| Y42H9AR.4 | similar to micronuclear linker histone | |

**Figure 3.4. Loss-of-function phenotypes of chromatin remodeling factors.** A, Polycomb mutants. B, RNAi against NuRD complex components and related proteins. C, Additional genes evaluated by analysis of mutant or RNAi phenotype.

NuRD complex members (Cardoso et al. 2005), also resulted in a largely wild-type phenotype (Figure 3.4B).

For RNAi experiments, *nsy-5*, which acts only during embryogenesis to control the $AWC^{ON}/AWC^{OFF}$ decision, was included as a negative control, and *odr-1*, which is involved in maintenance of expression beginning in the L1 stage, served as a positive control. RNAi against *odr-1* had only a weak effect (reduced *srsx-3* in ~5% and reduced *str-2* in ~20% of worms). The *odr-1* RNAi phenotype, while weak, was stronger than that observed for any other RNAi target evaluated alone. However, the low efficiency of *odr-1* RNAi even in an RNAi-sensitive background makes it difficult to interpret the negative results obtained from these RNAi experiments.

Three additional mutants involved in chromatin remodeling in *C. elegans* were also examined: *tam-1*, *jmjd-2*, and *rbr-2*. The RING finger/B-box gene *tam-1* enhances expression of repetitive DNA, and may act by modifying the surrounding chromatin structure; *tam-1(cc567)* worms exhibited a mild reduction in *str-2* expression in ~15% of animals (Figure 3.4C). *tam-1(cc567)* encodes an early translational stop and is likely to be a null allele (Hsieh et al. 1999).

*jmjd-2* and *rbr-2* are related jumonji domain-containing histone demethylases. *jmjd-2* contains a zinc finger/PHD DNA-binding domain predicted to bind a sequence similar to the AATCC activator sequence in the *srsx-3* promoter (Figure 2.5). Since *jmjd-2* demethylates the H3K9 repressive methylation mark, it is predicted to have an activating effect and therefore would be a candidate for an *srsx-3* activator (Whetstine et al. 2006). However, *jmjd-2* and *rbr-2* deletion mutants displayed no defect in *str-2* or *srsx-3* expression (Figure 3.4C).

Last, RNAi against genes surrounding *nsy-7* revealed one, *Y42H9AR.4*, that exhibited a *nsy-7*-like phenotype for *str-2* expression in adults following RNAi by injection: ~25% of animals did not express *str-2*, and expression levels were reduced in an additional ~15% of worms. *srsx-3* expression was not evaluated in this experiment (Figure 3.4C). The *Y42H9AR.4* locus is separated from *nsy-7* by three genes (~25kb) and encodes a poorly-conserved protein with multiple nuclear localization sites and some similarity to a *Tetrahymena* micronuclear linker histone (Wu et al. 1994). *Y42H9AR.4* might represent a candidate nuclear factor that can interact with NSY-7 to promote maintenance of *str-2* expression.

**Discussion**

NSY-7 may recruit additional proteins to mediate its repressive effects at target promoters. Given the involvement of NSY-7 in maintaining long-term transcriptional effects, chromatin remodeling factors constitute likely candidates for these cofactors. However, no substantial role was uncovered for any well-characterized chromatin factors in maintaining expression of *str-2* or *srsx-3*.

This mutant and RNAi analysis of a few chromatin factors represents a far-from-exhaustive evaluation of the involvement of chromatin modification in the $AWC^{ON}$/$AWC^{OFF}$ decision. Further exploration of potential interactions between NSY-7 and chromatin remodeling factors, or the role of chromatin state in the establishment and maintenance of $AWC^{ON}$ and $AWC^{OFF}$ identities, would be interesting. A preliminary assay might involve staining *str-2* or *srsx-3* array-bearing animals with antibodies against specific active and inactive histone modifications; it is possible that these experiments would detect gene-specific

differences in the two AWCs, or even differences in overall chromatin state between $AWC^{ON}$ and $AWC^{OFF}$. If specific histone methylation or acetylation marks differ between the two AWCs, the enzymes responsible for creating or destroying these marks might represent likely candidates for NSY-7 interactors.

The strongest candidate interactor for *nsy-7* to emerge from the experiments reported here is *Y42H9AR.4*. Further work on the function of this gene must be done to fully elucidate its role in $AWC^{ON}$ and $AWC^{OFF}$ identities. To begin, it will be interesting to test the *srsx-3* phenotype of *Y42H9AR.4* knockdown animals.

Identification of additional interacting factors might be accomplished using either genetic or biochemical approaches. Specifically, a genetic screen for suppressors of *nsy-7* would identify downstream genes, potentially including transcriptional regulators. Yeast two-hybrid assays using NSY-7 as bait would help to identify physical interactors, as would co-immunoprecipitation experiments, using antibodies against NSY-7::GFP or GST::NSY-7 fusion proteins. However, the latter two approaches would likely identify proteins that interact with NSY-7 in other tissues where *nsy-7* is expressed, including the gut, a much larger tissue than AWC. This effect would reduce the probability of identifying AWC-specific interactions.

## V. NSY-7 binds to methylated DNA in a context-dependent manner

**Results**

During searches for NSY-7-related proteins in other species, we noted that the NSY-7 DNA-binding domain displays weak similarity to the DNA-binding

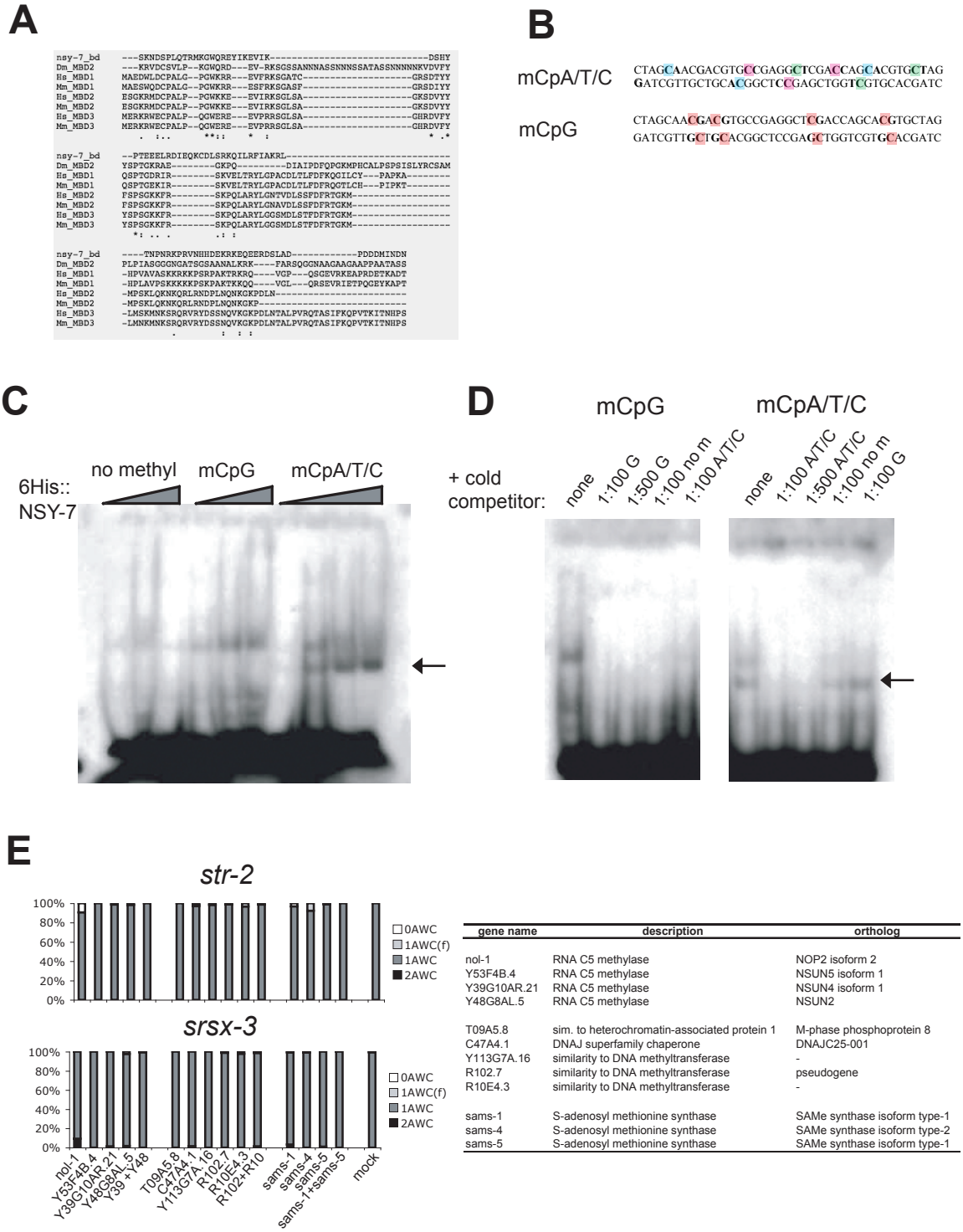
domains of methyl-cytosine binding proteins (Figure 3.5A). The worm genome is not believed to contain methylated DNA (Hendrich and Tweedie 2003). However, the same was thought of *Drosophila* until 2000, when it was found that *Drosophila* embryos exhibit low levels of methylation, representing ~0.4% of total DNA, and these levels drop off drastically in larvae and adults (Gowher et al. 2000; Lyko et al. 2000). Moreover, DNA methylation in *Drosophila* is unusual in that it occurs primarily asymmetrically at CpT and CpA dinucleotides, rather than symmetrically at CpG dinucleotides as is the rule in mammals (Lyko et al. 2000).

EMSA assays were used to evaluate NSY-7 binding to a DNA fragment containing CpG methylation and also to a fragment containing a mixture of CpT, CpA, and CpC methylated sites. The nucleotide sequences of both of these double-stranded fragments were identical: they differed only in their methylation sites (Figure 3.5B). The nucleotide sequence did not contain the CCTTAAC 7-mer previously shown to bind NSY-7 in a sequence-specific manner.

NSY-7 was capable of binding to the CpT/A/C methylated fragment, but not to the CpG methylated fragment (Figure 3.5C, arrow; upper band is nonspecific). Binding was dose-dependent, and could be efficiently competed away by cold CpT/A/C methylated oligo but not by an unmethylated fragment of the same sequence or by the CpG methylated fragment (Figure 3.5C). We then attempted to separate the CpT, CpA, and CpC binding specificities by performing EMSAs using oligos containing only one of these three methylation sequences; however, NSY-7 did not bind to any of these fragments (data not shown). In order to include the same total number of methylation sites in each of

**Figure 3.5. NSY-7 can bind to methylated DNA.** A, Alignment of NSY-7 homeodomain with the methyl-cytosine binding domains of Drosophila MBD2 and mouse and human MBD1, MBD2, and MBD3 proteins. B, Methylated DNA fragments used as probes in EMSAs. Methylated residues are colored. C, EMSAs for NSY-7 binding to unmethylated DNA probe, probe methylated at CpG sites, and probe methylated at a mixture of CpA, CpT, and CpC sites. D, Competition experiments for NSY-7 binding. Left, NSY-7 with labeled CpG methylated probe, competed with 1:100 or 1:500 cold CpG methylated probe, or with 1:100 cold unmethylated or CpA/T/C methylated probe. Right, NSY-7 with labeled CpA/T/C methylated probe, competed with 1:100 or 1:500 cold CpA/T/C methylated probe or with 1:100 cold unmethylated or CpG methylated probe. E, RNAi against candidate C5 methyltransferase genes and against S-adenosyl methionine synthase (*sams*) genes.

Figure 3.5




| gene name | description | ortholog |
|---|---|---|
| nol-1 | RNA C5 methylase | NOP2 isoform 2 |
| Y53F4B.4 | RNA C5 methylase | NSUN5 isoform 1 |
| Y39G10AR.21 | RNA C5 methylase | NSUN4 isoform 1 |
| Y48G8AL.5 | RNA C5 methylase | NSUN2 |
| T09A5.8 | sim. to heterochromatin-associated protein 1 | M-phase phosphoprotein 8 |
| C47A4.1 | DNAJ superfamily chaperone | DNAJC25-001 |
| Y113G7A.16 | similarity to DNA methyltransferase | - |
| R102.7 | similarity to DNA methyltransferase | pseudogene |
| R10E4.3 | similarity to DNA methyltransferase | - |
| sams-1 | S-adenosyl methionine synthase | SAMe synthase isoform type-1 |
| sams-4 | S-adenosyl methionine synthase | SAMe synthase isoform type-2 |
| sams-5 | S-adenosyl methionine synthase | SAMe synthase isoform type-1 |

this second set of fragments, the DNA sequence of the target was altered; it is possible, therefore, that NSY-7 requires some specific sequence surrounding the methylation site in order to bind.

The *C. elegans* genome is not predicted to contain a classical DNA methyltransferase, but has several genes that might encode atypical DNA methyltransferases, including the SUN family of RNA C5 methyltransferases and several genes found by BLAST searches to have weak similarity to known DNA methyltransferases. Mutants of most of these genes are lethal or do not exist, and RNAi feeding clones are also not available. Therefore, expression of these genes was knocked down by injecting double-stranded RNA into an RNAi-sensitized strain expressing the *str-2* and *srsx-3* reporters . None of the knockdown experiments had a strong effect on *str-2* or *srsx-3* expression, although knockdown of *nol-1*, a SUN family RNA methylase, resulted in a very weak *nsy-7*-like defect (Figure 3.5D). We also knocked down three S-adenosyl methionine synthase (*sams*) genes; S-adensoyl methionine (SAMe) is the methyl donor in DNA methylation (Bestor 2000). The three *sams* genes tested were lethal following RNAi, but a few worms that did reach adulthood sometimes also exhibited a *nsy-7*-like phenotype. Overall, the high level of lethality following knockdown of many of these targets made it difficult to fully evaluate the effects of these genes on *srsx-3* and *str-2* expression.

**Discussion**

It is intriguing that NSY-7 seems to be capable of binding both a specific nucleotide sequence, CCTTAAC, and a random sequence containing CpT, CpA,

and CpC methylation. While our data is highly preliminary, it is interesting to speculate about the possible implications of these results should they prove true.

Given that *C. elegans* is not reported to methylate its DNA, the ability to bind methylated sequence may reflect an activity of the NSY-7 DNA-binding domain that is never utilized *in vivo*. A highly speculative alternative explanation is that there could be a structural relationship between *bona fide* DNA methyl-binding proteins and divergent homeodomains such as NSY-7. The structure for human MBD1 bound to DNA has been solved by NMR, and residues important for CpG binding were determined (Ohki et al. 2001). Four of these are retained in NSY-7, and two are not present in NSY-7 but are also not conserved in other MBD proteins. The remaining two are arginines in MBD1 and lysines in NSY-7; one of these positions was shown to tolerate lysine substitutions, and the other has not been evaluated (Ohki et al. 2001). Methyl-binding proteins frequently mediate transcriptional repression, a biological activity similar to that of NSY-7 (Ballestar and Wolffe 2001). It is possible that a sequence-specific DNA-binding protein such as NSY-7 might be co-opted to mediate a methylation-based system of transcriptional repression under special conditions where DNA methylation arises. Another possibility is that nematodes lost a methylation system present in ancestral species, and that NSY-7 represents a remnant of that system that has adopted a new, sequence-specific function. Consistent with this hypothesis, the genome of the related nematode *Pristionchus pacificus* contains a DNA methyltransferase 2 (DNMT2)-like gene, and the absence of a homolog in *C. elegans* has been proposed to be the result of recent gene loss (Gutierrez and Sommer 2004).

Unlike most known methyl-binding proteins, NSY-7 appeared to exhibit some sequence specificity in combination with methyl binding, interacting with one set of methylated DNA fragments used in our EMSAs (Figure 3.5) but not another. MBD1 can bind to unmethylated DNA sequence and affect transcription at unmethylated or hypomethylated promoters, and it has recently been shown that efficient binding of MeCP2 to methylated sites in mammals is dependent on an enrichment of A's and T's adjacent to the methylated cytosines (Ballestar and Wolffe 2001; Klose et al. 2005). A similar context-dependent effect could therefore alter NSY-7 binding. This characteristic may again point to an interesting and possibly evolutionarily relevant intermediate between sequence-specific transcription factors and more promiscuous factors such as methyl-binding proteins.

In light of the apparent ability of NSY-7 to interact with methylated DNA, it might be interesting to revisit the question of whether methylated DNA is in fact present in the *C. elegans* genome. If it is present at specific developmental stages and in small quantities, as is true in *Drosophila*, earlier attempts at detection might have missed the relatively weak signal. Bisulfite sequencing of candidate DNA regions at various developmental time points might identify a methylation site. Alternatively, the antibodies for methylated cytosines that are currently available might be used to detect low levels of methylation in a pure DNA preparation (Marhold et al. 2004; Zilberman and Henikoff 2007). L1 larvae arrested by starvation can live for weeks or months, as can dauer larvae; these states require long-term stabilization of a transcriptional program, including repression of multiple genes expressed under more favorable conditions, and might be interesting developmental stages in which to look for low levels of

DNA methylation (Liu et al. 2004; Baugh and Sternberg 2006; Fisher and Lithgow 2006).

The absence of any predicted DNA methyltransferase ortholog in the *C. elegans* genome is discouraging. There are, however, alternative pathways to methylcytosine synthesis. For example, *Arabidopsis* expresses the CMT3, DRM1, and DRM2 methyltransferases, which are not related by sequence to animal cytosine methyltransferases (Cao and Jacobsen 2002a; Cao and Jacobsen 2002b; Ponger and Li 2005). It is also possible to produce 5-methylcytosine by amination of thymine, although this pathway has not been demonstrated *in vivo* (Friedkin and Wood 1956; Holliday and Pugh 1975).

## Conclusions

As shown in Chapter 1, the role of NSY-7 in regulating *srsx-3* expression is straightforward: it binds directly to the *srsx-3* promoter to repress expression of that gene. Here we found that NSY-7 induces *str-2* only after a significant time delay, and that this delay apparently relies more on the absolute time since the appearance of NSY-7 than on the developmental stage at which it appears. The delay suggests the involvement of additional transcription factors acting either together with or downstream of *nsy-7*. One of these may be *hlh-11*, a transcriptional target of *nsy-7* that has a repressive effect on both *str-2* and *srsx-3* when overexpressed in a *nsy-7* mutant background. Another could be *Y42H9AR.4*, an uncharacterized protein with weak similarity to histone H1 whose knockdown by RNAi leads to defects in *str-2* expression.

NSY-7::GFP forms punctate structures in the nucleus of AWC that are suggestive of the formation of subnuclear domains. However, several factors make the significance of these structures difficult to interpret. NSY-7::GFP may be highly overexpressed, leading to aggregation of the protein. It is also possible that the presence of the *odr-3* or *elt-2* promoters used to drive expression of NSY-7::GFP from extrachromosomal arrays interferes with NSY-7 function and induces formation of the punctate structures.

The evidence presented here suggests that the NSY-7 mechanism of action is complex and not yet fully understood. It is likely that additional transcriptional regulators interact with NSY-7 in the nucleus to promote or modify its transcriptional activity, and NSY-7 itself regulates the transcription of *hlh-11* and possibly other transcription factors. The identities of both of these groups of genes should be uncovered to better understand the action of NSY-7 and the nature of the $AWC^{ON}$ neuron.

# Chapter 4

## Cell-specific factors modify maintenance of neuronal identity by environmental input


**Summary**

Neuronal cell fate maintenance involves both cell-intrinsic processes and environment-dependent control of gene expression. To better define the interplay between these two processes and to identify genes involved, we performed a screen for mutants defective for maintenance of expression of cell-specific markers in the two *C. elegans* olfactory neurons, $AWC^{ON}$ and $AWC^{OFF}$. Our screen uncovered three groups of genes: those affecting cGMP-dependent olfactory signal transduction; members of a TGFβ signaling pathway; and several transcriptional regulators, including a previously uncharacterized HNF homeodomain protein, *hmbx-1*. Using several new $AWC^{ON}$- or $AWC^{OFF}$-specific GPCR markers, we examined the effects of mutations on suites of markers within the same cell. We found that the TGFβ and cGMP pathways affect individual promoters within a neuron, whereas *hmbx-1* alters the sensitivity of $AWC^{OFF}$ but not $AWC^{ON}$ to TGFβ and cGMP signaling. Moreover, the *srsx-3* promoter responds differently to regulatory inputs when it is expressed in $AWC^{ON}$ *versus* $AWC^{OFF}$. Based on our results, we propose that environmental signals modulate the state of a neuron by altering its receptor repertoire, and this modulation is overlaid on separate sets of signals that control the responsiveness of the neuron as a whole.

**Introduction**

The early development of the nervous system reliably produces specific neuronal subtypes and synaptic connections. However, most animals modify their nervous system after birth to better tune it to the environment. Individual synapses may be established, eliminated, strengthened or weakened; new cells may be produced and integrated into the preexisting nervous system; and differentiated neurons may alter their existing role, either by modulating their level of responsiveness or by developing the ability to respond to new stimuli. An example of a developmentally hard-wired switch occurs in terminally-differentiated photoreceptor neurons that express the rhodopsin Rh5 and respond to blue light in the *Drosophila* larva, and which are retained during metamorphosis and become part of the adult eye. During this time, however, they stop expressing Rh5 and begin expressing a different rhodopsin, Rh6, resulting in a switch in sensitivity from blue to green light (Sprecher and Desplan 2008). An example of an environmental switch occurs in mammalian cortical neurons, where calcium influx induced by neural activity activates multiple transcription factors including NPAS4. NPAS4 regulates the balance of inhibitory and excitatory synapses in excitatory neurons, thus altering the nature of the neuron and its role within a network (Lin et al. 2008; Zhang et al. 2009).

If terminally differentiated neurons can undergo large-scale transcriptional changes, it is also important to ask how stable aspects of neuronal identities are retained. Following differentiation, expression of many developmental regulatory genes ceases, but cellular identity is stable in their absence (Shirasaki and Pfaff 2002; Ringrose and Paro 2004). What buffers core cellular identity against transcriptional changes induced by environmental

inputs? First, a developmental switch may 'lock in' the cell fate decisions that have been made. Such switches are usually presumed to involve chromatin modifications. The Polycomb group (PcG) proteins act in this manner to maintain Hox gene expression; PcG proteins can promote an inactive chromatin state by nucleosome remodeling and recruitment of histone methyltransferase and histone deacetylase factors, thereby stabilizing the transcriptional state of their target genes (Busturia and Morata 1988; Kennison and Tamkun 1988; Cao et al. 2002; Muller and Kassis 2006; Schuettengruber et al. 2007).

Second, transcription factors required for maintenance of subtype-specific genes may be expressed continuously, and continuously required for retention of cell identity. These transcription factors typically maintain their own expression in addition to that of their target genes. In the mammalian immune system, the transcription factor Pax5 is required for initial commitment to the B cell lineage, and continued expression of Pax5 is also required to maintain that commitment: deletion of Pax5 from committed pro-B cells resulted in their reversion to a multipotential state capable of differentiating into mature macrophages (Nutt et al. 1999; Rolink et al. 1999; Mikkola et al. 2002). Likewise, in the *C. elegans* nervous system, the transcription factor CHE-1 maintains its own expression in the ASE chemosensory neurons and is required both for the initial specification and later maintenance of these neurons (Etchberger et al. 2009).

Third, external inputs such as sensory or synaptic stimulation may be required for maintenance of a specific aspect of neural function or identity. Visual input and electrical activity establish ocular dominance columns in the visual cortex, and in addition specific genes in cortical neurons are continously regulated by the presence or absence of light (Majdan and Shatz 2006). In

*Drosophila*, the Tv neurons require a retrograde bone morphogenic protein (BMP) signal from synaptic targets to maintain expression of a cell type-specific peptide, FMRFa (Eade and Allan 2009).

In the *C. elegans* AWC olfactory neurons, factors required for stabilization of cell fate differ from those required for its initiation. The right and left AWC olfactory neurons take on asymmetric fates in a stochastic manner, such that one AWC becomes $AWC^{ON}$ and one becomes $AWC^{OFF}$(Troemel et al. 1999). $AWC^{ON}$ expresses the GPCR *str-2* and senses the odor butanone, while $AWC^{OFF}$ expresses the GPCR *srsx-3* and senses the odor 2,3-pentanedione (Troemel et al. 1999; Wes and Bargmann 2001). Early in development, both AWCs express the $AWC^{OFF}$ marker *srsx-3*; an embryonic signaling pathway converts one AWC into a *str-2*-expressing $AWC^{ON}$ cell while stabilizing the $AWC^{OFF}$ state in the opposite cell (Troemel et al. 1999; Sagasti et al. 2001; Tanaka-Hino et al. 2002; Chuang and Bargmann 2005; Vanhoven et al. 2006; Bauer Huang et al. 2007; Chuang et al. 2007; Lesch et al. 2009). The transcription factor NSY-7 acts as an output of this pathway to repress *srsx-3* expression in $AWC^{ON}$, and is also required for maintenance of *str-2* expression in $AWC^{ON}$ (Lesch et al. 2009). After hatching, the signaling pathway responsible for odor detection maintains both *str-2* and *srsx-3* expression. This pathway is regulated by G-protein signaling, which is thought to regulate cGMP levels; targets of cGMP in AWC include a cyclic-nucleotide-gated cation channel and a cGMP-responsive protein kinase (Bargmann et al. 1993; Coburn and Bargmann 1996; Komatsu et al. 1996; Roayaie et al. 1998; L'Etoile and Bargmann 2000; Fujiwara et al. 2002; L'Etoile et al. 2002). Thus, odor detection by GPCRs inhibits cGMP signaling, calcium entry, and activation of the kinase. Several olfactory Gα proteins have been shown to modulate *str-2*

expression during early larval stages, and the guanylate cyclases ODR-1 and DAF-11 are required to maintain expression of both *str-2* and *srsx-3*, as is the cGMP-dependent kinase EGL-4 (Troemel et al. 1999; Lans et al. 2004; Lans and Jansen 2006; Lesch et al. 2009). Because these genes are required for odor sensation, olfactory input may play a role in maintenance of GPCR expression during larval stages.

Here, we describe a mutant screen for genes that maintain expression of *srsx-3* and *str-2*. We find that these genes fall into three categories: olfactory signaling, TGFβ signaling, and transcriptional regulation. First, we confirm that the olfactory signaling pathway is essential for maintenance of GPCR expression in the AWCs. Second, we find that TGFβ signaling induced by *daf-7* acts in $AWC^{OFF}$ to maintain *srsx-3* expression, and that signaling through this pathway is responsible for ongoing regulation of expression in adults in response to environmental pheromone signals. Third, we identify a missense mutation in the conserved transcription factor HMBX-1 that highlights a role for this gene in maintenance of GPCR expression in $AWC^{OFF}$. We identify five new $AWC^{ON}$- and $AWC^{OFF}$-specific GPCRs and demonstrate that, while *hmbx-1* affects all known $AWC^{OFF}$-specific markers and no $AWC^{ON}$-specific markers, cGMP signaling affects expression of a subset of markers in either cell and TGFβ signaling regulates a subset of markers in $AWC^{OFF}$ only. A single promoter, *srsx-3*, is regulated differently by *hmbx-1* in $AWC^{ON}$ or $AWC^{OFF}$. The convergence of multiple pathways regulates the receptors expressed by a sensory cell in a given environmental context, an expression signature that depends both on the transcriptional state of the cell as a whole and on the effects of external stimuli.

## Results

### Nineteen mutants defective in maintenance of GPCR expression

To isolate genes involved in maintenance of GPCR expression in AWC, we conducted a screen for 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 receptors in adulthood. Worms carrying stably integrated *str-2::dsRed2* and *srsx-3::GFP* reporter constructs were mutagenized and their adult F2 progeny examined for defects in expression of *str-2, srsx-3*, or both. After verifying the adult phenotype in subsequent generations, we evaluated expression of *str-2* and *srsx-3* in early larvae, 14 hours after hatching (L1 stage), and kept mutants with a wild-type phenotype at this stage. We obtained nineteen mutants with maintenance defects (Figure 4.1A). Genetic mapping and complementation tests demonstrated that these mutants included three alleles of *odr-1*, one of *daf-11*, one of *tax-2*, and one of *tax-4*. These genes are involved in the olfactory signal transduction pathway and have been previously reported to be involved in maintenance of *str-2* and *srsx-3* expression (Troemel et al. 1999; Lesch et al. 2009).

We also isolated two alleles of *nsy-7*, one allele of *tam-1*, and one allele of *ceh-36*. *nsy-7* encodes a transcription factor that affects maintenance of *str-2* and is required for restriction of *srsx-3* to AWC$^{OFF}$ (Lesch et al. 2009). *tam-1* is a transcriptional regulator that inhibits silencing of repetitive transgenes and regulates expression of some endogenous genes during development (Hsieh et al. 1999; Kiefer et al. 2007); this gene was not previously known to affect gene

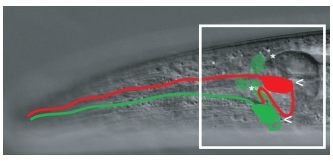
**Figure 4.1. A screen for AWC maintenance mutants.** A, Summary. Adult mutant phenotypes are shown as pie charts for *srsx-3* (green) and *str-2* (red) expression (grey, no expression; bright green/red, expressed in one AWC; dark green/red, expressed in both AWCs). The gene affected and coding change are listed when known. Phenotypic data for known alleles is shown at the right side of the chart. n>30 for each mutant. B, Three mutants specifically defective for maintenance of *srsx-3* expression in $AWC^{OFF}$. Left, DIC image of the head of a worm overlaid with a diagram of the $AWC^{ON}$ (red) and $AWC^{OFF}$ (green) neurons. Arrowheads, AWC; asterisks, AWB. Right, adult and early larval phenotypes of wild-type, *daf-7(e1372)*, *daf-1(m40)*, and *ky777* worms.

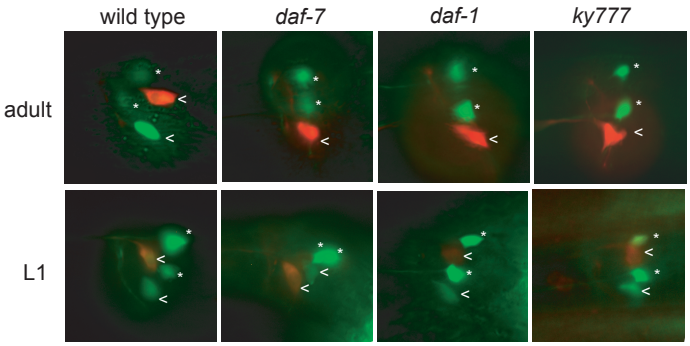
Figure 4.1

**A**

| allele | adult phenotype srsx-3 | adult phenotype str-2 | gene | coding change | domain | characterized allele | characterized allele srsx-3 | characterized allele str-2 |
|---|---|---|---|---|---|---|---|---|
| wild type | | | | | | | | |
| *ky779* | | | *odr-1* | ? | ? | *odr-1(n1936)* | | |
| *ky800* | | | *odr-1* | ? | ? | | | |
| *ky807* | | | *odr-1* | ? | ? | | | |
| *ky801* | | | *daf-11* | D808V | cyclase domain | *daf-11(m47)* | | |
| *ky782* | | | *tax-2* | N329K | channel | *tax-2(ks31)* | | |
| *ky791* | | | *tax-4* | Q510STOP | after tm domain | *tax-4(p678)* | | |
| *ky879* | | | *odr-3(gf)* | G185S | stabilization of GTP binding | | | |
| *ky813* | | | *nsy-7* | ? | ? | *nsy-7(tm3080)* | | |
| *ky819* | | | *nsy-7* | ? | ? | | | |
| *ky798* | | | *tam-1* | G76S | zinc finger | *tam-1(cc567)* | | |
| *ky777* | | | *F54A5.1* | H404Y | ? | | | |
| *ky771* | | | *daf-7* | Q105STOP | - | *daf-7(e1372)* | | |
| *ky803* | | | *daf-1* | G400E | kinase | *daf-1(m40)* | | |
| *ky793* | | | *ceh-36* | A91T | homeodomain | *ceh-36(ky646)* | | |
| *ky772* | | | *?* | ? | ? | *?* | | |
| *ky773* | | | *?* | ? | ? | *?* | | |
| *ky789* | | | *?* | ? | ? | *?* | | |
| *ky802* | | | *?* | ? | ? | *?* | | |
| *ky811* | | | *?* | ? | ? | *?* | | |

**B**





expression in AWC. *ceh-36* is an Otx-family homeodomain transcription factor that is required to establish general AWC cell fate (Lanjuin et al. 2003; Koga and Ohshima 2004). Interestingly, the *ceh-36(ky793)* allele isolated from our screen strongly and specifically affects *srsx-3* but not *str-2* expression. Because *ceh-36(ky793)* encodes a missense mutation in the CEH-36 DNA binding domain, it is possible that this mutation specifically affects binding to the *srsx-3* promoter.

In addition, we isolated a dominant mutation in the Gα subunit *odr-3*, *odr-3(ky879)*, that encodes a G → S missense mutation at position 185. The affected glycine is highly conserved, and this region is involved in the conformational change that occurs upon GTP binding; stabilization of the GTP-bound conformation might be expected to result in a constitutively active protein (Rens-Domiano and Hamm 1995). Loss-of-function mutations in *odr-3* alone have little effect on *str-2* expression, although *str-2* expression is lost in adults when *odr-3(lf)* alleles are combined with mutations in other Gα subunits (Lans and Jansen 2006). In contrast, the *odr-3(ky879)* allele alone has a strong effect on *str-2* as well as a weak effect on *srsx-3* expression, reinforcing the role of this olfactory signal transduction component in the maintenance of GPCR expression.

**Members of the TGFβ pathway affect maintenance of *srsx-3* expression**

Our screen identified several mutations that primarily alter *srsx-3* and not *str-2* expression. *tax-4(ky791)* and *tax-2(ky782)* share this phenotype, along with *ceh-36(ky793)* and three additional mutant alleles. *srsx-3* is ordinarily expressed in AWC$^{OFF}$ and in both AWB neurons. The *ky771, ky803,* and *ky777* alleles, which represent three separate complementation groups, each preserved expression of

*srsx-3* in AWB, but lost expression in AWC$^{OFF}$ between larval and adult stages (Figure 4.1B). *ky771* and *ky803* also exhibited a strongly dauer-constitutive phenotype; dauer larvae are an alternative larval stage usually induced under conditions of starvation, crowding, or other stress. Consistent with this phenotype, *ky771* mapped to *daf-7*, which encodes a TGF-β ligand required to prevent dauer entry, while *ky803* mapped to *daf-1*, which encodes the type I TGF-β receptor that acts downstream of *daf-7* (Georgi et al. 1990; Ren et al. 1996; Schackwitz et al. 1996).

**DAF-7/ TGFβ signals through a canonical pathway to maintain *srsx-3* expression**

To further evaluate the signaling pathway acting downstream of *daf-7*/ TGFβ in maintenance of *srsx-3* expression, we examined the *str-2* and *srsx-3* expression phenotypes of mutants with defective DAF-7 signaling in dauer formation. We found that *daf-4* (the sole *C. elegans* 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*. In each case, *str-2* expression was normal, whereas *srsx-3* expression was absent in 100% of adult animals (Figure 4.2). *daf-3*, a co-Smad that binds DNA, and *daf-5*, a proline-rich protein thought to be involved in transcriptional regulation, act downstream of *daf-7*, *daf-1*, *daf-4*, *daf-8*, and *daf-14* to promote dauer formation; they are antagonized by upstream components of the pathway (Estevez et al. 1993; Patterson et al. 1997; Inoue and Thomas 2000b; Inoue and Thomas 2000a; da Graca et al. 2004). *str-2* and *srsx-3* expression patterns in *daf-3*, *daf-5*, and *daf-3;daf-1* or *daf-5;daf-1* mutants were

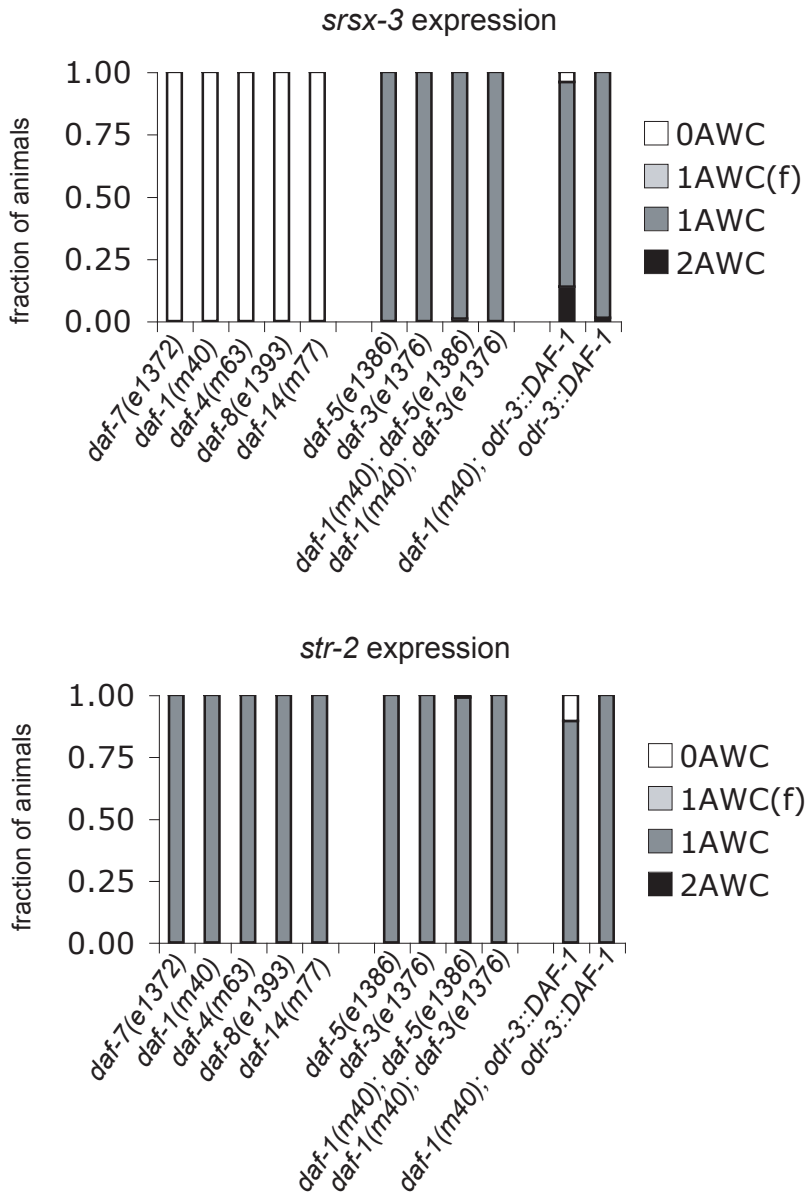



**Figure 4.2. TGFβ signaling regulates *srsx-3*.** *srsx-3* and *str-2* expression phenotypes for *daf-7*/TGFβ pathway mutants and rescue of *daf-1(m40)* under the AWC-selective promoter *odr-3*. In the legend, 1AWC(f) refers to expression in one AWC that is qualitatively fainter than wild-type; this convention is used in all subsequent figures.

normal (Figure 4.2). We conclude that DAF-7/ TGFβ signals through the canonical TGFβ pathway to regulate *srsx-3* expression.

To ask whether the effects of DAF-7 TGFβ signaling in $AWC^{OFF}$ were cell autonomous, we expressed the *daf-1* cDNA under the AWC-selective *odr-3* promoter in a *daf-1(m40)* mutant background. Expression of DAF-1 in AWC rescued the *srsx-3* expression defect almost completely; 6% of animals expressed *srsx-3* in both AWCs, suggesting a possible overexpression defect. These experiments suggest that DAF-7/ TGFβ signaling in AWC is required to maintain expression of the GPCR *srsx-3* in $AWC^{OFF}$.

TAX-4, the cGMP-gated cation channel a subunit, promotes expression of *daf-7* in the ASI chemosensory neuron (Chang et al. 2006). *tax-4* mutants have an *srsx-3* phenotype similar to that of the TGFβ pathway mutants, where *srsx-3* expression is lost while *str-2* expression remains intact (Figure 4.1A). To ask whether TAX-4 regulates *srsx-3* expression though control of *daf-7* expression or whether it acts cell autonomously in AWC, 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 (Figure 4.3). Therefore, TAX-4 acts in AWC to promote *srsx-3* expression, in a process that is separate from its effects on *daf-7*/TGFβ expression in ASI.

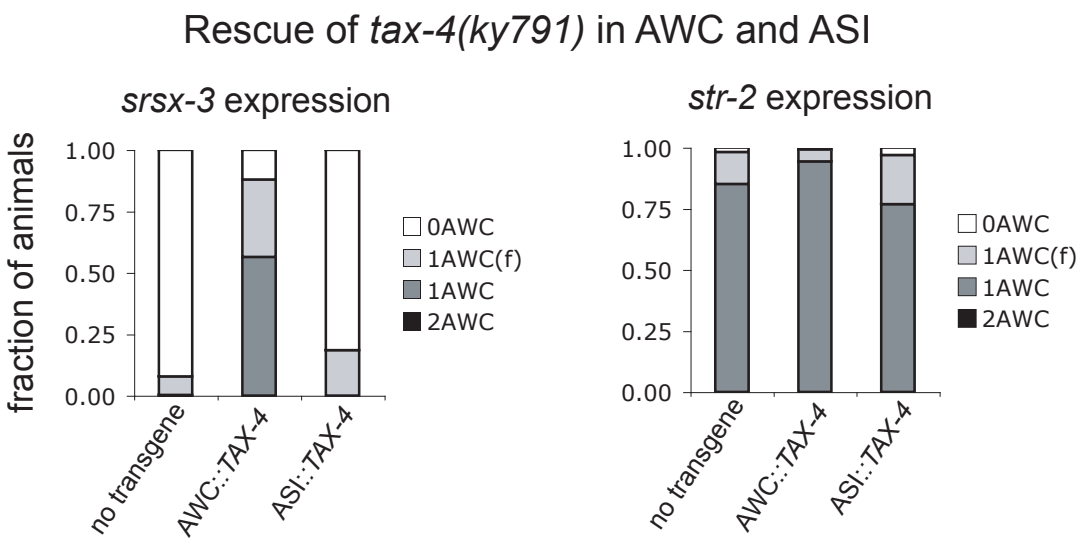



**Figure 4.3. *tax-4* acts cell-autonomously in AWC.** Rescue of *tax-4(ky791)* by expression of a TAX-4 cDNA in AWC using the *odr-3* promoter, or in ASI using the *str-3* promoter.

**Dauer pheromone components C3 and C6 regulate *srsx-3* via DAF-7/ TGFβ signaling**

DAF-7/ TGFβ signaling plays a significant role in early larval development, when it is required to prevent entry into the dauer stage (Ren et al. 1996; Schackwitz et al. 1996). However, *srsx-3* expression in TGFβ pathway mutants was normal during early larval stages; defects were observed only in late larvae and adults. Therefore, we asked whether TGFβ signaling regulates gene expression in adult animals.

*daf-7* expression is inhibited by the density-dependent dauer pheromone, a mixture of structurally related chemicals, termed ascarosides, that are constitutively produced by *C. elegans* (Butcher et al. 2007; Butcher et al. 2008; Butcher et al. 2009). If DAF-7 is continuously required in adult animals to maintain *srsx-3* expression, pheromone stimulation and the ensuing downregulation of *daf-7* should result in decreased *srsx-3* expression in adult animals. To test this hypothesis, we exposed adult worms to the dauer pheromone component ascarosides C3, C6, and C9 alone or in combination and monitored expression of *srsx-3*. The GFP protein was destabilized by adding a PEST domain to its C-terminus (Gaudet and Mango 2002); similar GFP-PEST constructs have been shown to have a half-life of approximately one hour in worms (Frand et al. 2005). Wild-type adult worms expressing a *srsx-3::gfp-pest* reporter gene were exposed to pheromone for four hours at low worm densities in the presence of food. A mixture of equal parts C3, C6, and C9 suppressed *srsx-3* expression in a dose-dependent manner, reducing the fraction of animals expressing GFP by ~50% (from 0.73 at 0nM to 0.39 at 100nM, $p<0.001$ [one-way

ANOVA with Dunnett's post-test]) (Figure 4.4A). *daf-3(e1376)* loss-of-function mutants did not suppress *srsx-3* expression in response to pheromone, indicating the involvement of the TGFβ signaling pathway.

The pheromone components C3, C6, and C9 were then tested individually, at 100nM (the concentration of each ascaroside in the 100nM C3/C6/C9 mixture) and 300nM (the total concentration of pheromone in the 100nM C3/C6/C9 mixture). C3 and C6 each suppressed *srsx-3* expression to a similar extent as the mixture (0.73 vs. 0.47 for 100nM C3, $p<0.001$; 0.73 vs. 0.43 for C6, $p<0.001$ [one-way ANOVA with Dunnett's post-test]) (Figure 4.4B,C), but C9 had no significant effect (0.73 vs. 0.60 for 100nM C9, n.s.) (Figure 4.4D). The *daf-3(e1376)* mutation suppressed the effects of both C3 and C6 on *srsx-3* expression (Figure 4.4B,C).

We conclude from these experiments that the DAF-7/ TGFβ signaling pathway maintains and dynamically regulates expression of *srsx-3* in adult worms, and that signaling through this pathway is sensitive to external cues such as the ascarosides C3 and C6.

**_ky777_ is a dominant-negative mutation in the transcription factor _hmbx-1_**

The *ky777* adult phenotype was similar to that of the TGFβ pathway mutants, in that *srsx-3* expression was lost but *str-2* expression was retained (Figure 4.1A,B). However, *ky777* mutants did not display a dauer-constitutive phenotype, and the *srsx-3* expression defect in these mutants was not suppressed by *daf-3* (Figure 4.7A). Therefore, *ky777* appears to represent a separate mechanism for maintenance of *srsx-3* expression. *ky777* mapped to an interval of

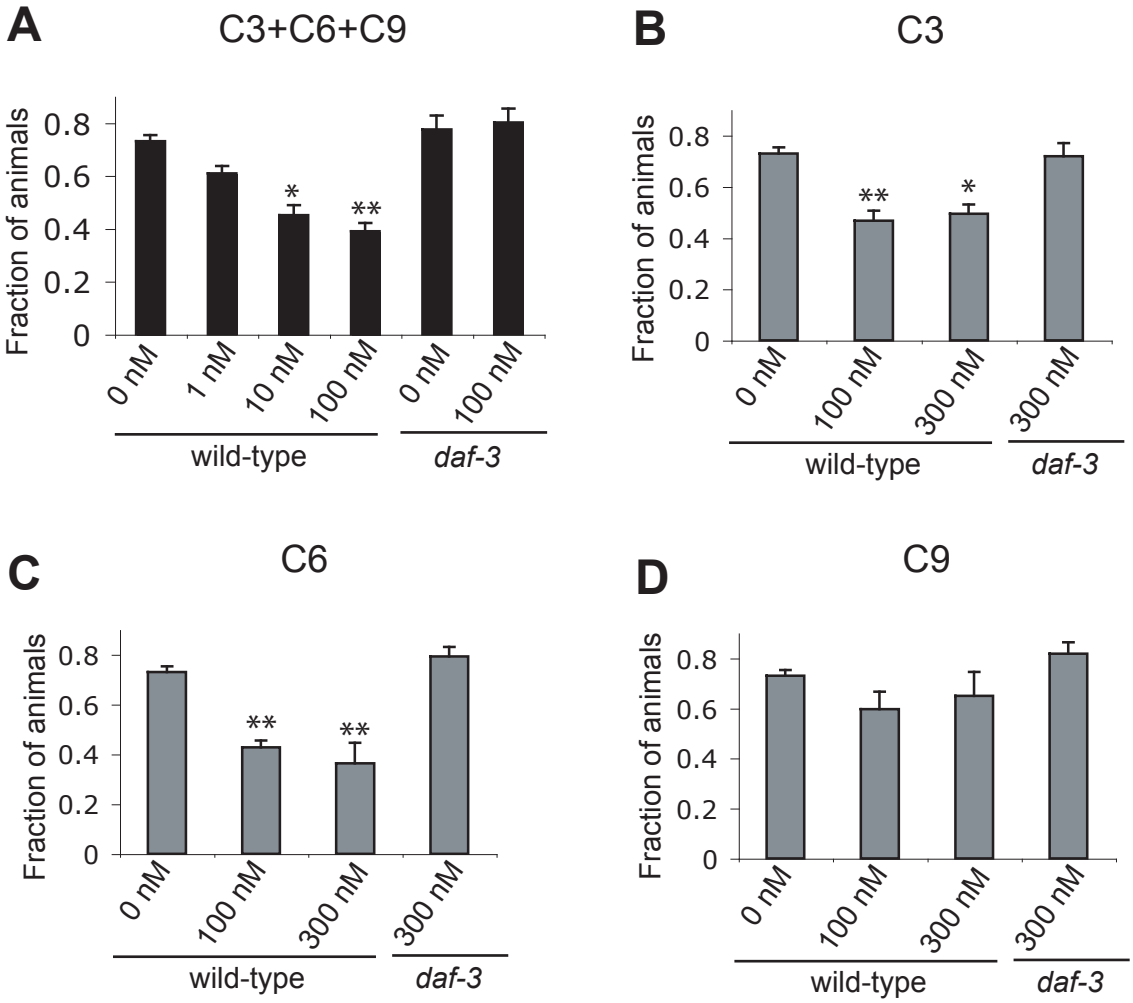


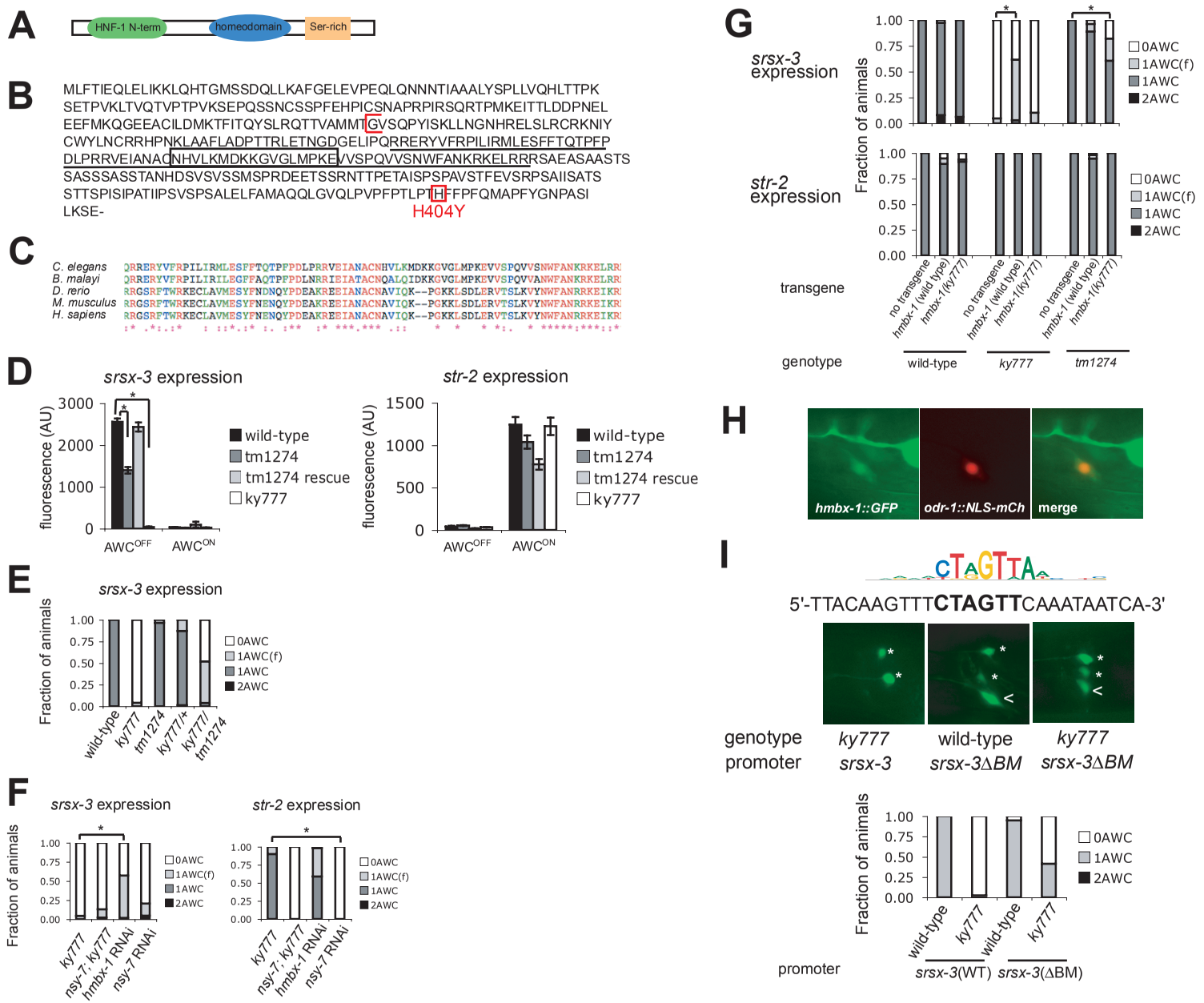

**Figure 4.4. Regulation of *srsx-3* by pheromone.** Panels show fraction of adults expressing a destabilized GFP under the control of the *srsx-3* promoter. A, Equimolar concentrations of C3, C6, and C9. Concentrations refer to the concentration of each ascaroside in the mixture. B, Ascaroside C3. C, Ascaroside C6. D, Ascaroside C9. [*] indicates $p<0.01$ and [**] indicates $p<0.001$ compared to no pheromone control (one-way ANOVA with Dunnett's post-test).

~440 kilobases on the left arm of chromosome I, but no rescue was observed using 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 ce6 genome assembly at UCSC accounted for 84% of sequence in the interval, with an average coverage depth of 7.9x. The gigaBayes program identified 19 high-probability point mutations and three single-base indels between *ky777* and the reference genome. Of these, three were in exons, including one silent mutation, one was in a 3'UTR, and one was in a 5'UTR. These five single-base changes were checked by PCR and conventional sequencing. One exon mutation and both UTR mutations were found to be present as predicted but were also present in the original, unmutagenized strain. The region surrounding the silent mutation is GC-rich; it failed to amplify and could not be verified. The remaining 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 mutation at the C-terminal end of the protein (Figure 4.5B). We concluded that this mutation was the most likely candidate for *ky777*.

*F54A5.1* encodes a predicted conserved homeodomain transcription factor that also contains an HNF-1 N-terminal-like domain and a serine-rich region at its C-terminus, and is related to mammalian HMBOX1 (Figure 4.5A). The homeodomain of *F54A5.1* and its homologs includes an unusual 17-amino acid insertion not present in other homeodomains (Figure 4.5B, black box); the entire protein is highly conserved in multiple species, including zebrafish, mouse, and human, with the most highly conserved region corresponding to the homeodomain (Figure 4.5C). Because of this high degree of conservation with

**Figure 4.5. *ky777* is a recessive gain-of-function allele of the conserved homeodomain transcription factor *hmbx-1*.** A, Domain structure of HMBX-1. B, Amino acid sequence of HMBX-1. The *ky777* mutation is boxed in red, the beginning of the *tm1274* deletion is indicated with a red bracket, the homeodomain is underlined, and an unusual 17-amino acid insertion is boxed in black. C, Conservation of the HMBX-1 homeodomain. D, Quantitation of *srsx-3::GFP* and *str-2::dsRed2* fluorescence in wild-type, *tm1274*, and *ky777* mutants. [*] indicates difference from wild-type at p<0.0001 (two-tailed t-test). E, *srsx-3* expression phenotypes of *ky777*/+ heterozygotes and *ky777*/*tm1274* trans-heterozygotes. F, RNAi against *hmbx-1* in a *ky777* background partially restores *srsx-3* expression. RNAi against *nsy-7* in a *ky777* background was performed as a control. The *ky777;nsy-7(tm3080)* double mutant is shown for comparison with *nsy-7* RNAi. All strains except *ky777;nsy-7* contained the RNAi-sensitizing *eri-1(mg366)* and *lin-15B(n744)* mutations. [*] indicates difference from RNAi-negative control at p<0.0001 (Fisher's Exact test). G, 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 7kb of upstream sequence. [*] indicates difference from transgene-negative control at p<0.0001 (Chi-square test). H, Expression of *hmbx-1::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. I, (top), Binding site for HMBOX1, the mouse homolog of HMBX-1, shown above the corresponding site in the context of the *srsx-3* promoter. (middle), GFP expressed under the control of the full-length *srsx-3* promoter or an *srsx-3* promoter lacking the HMBOX1 binding site (*srsx-3*Δ*BM*) in wild-type and *ky777* backgrounds. (bottom), fraction of wild-type or *ky777* animals expressing GFP in AWC under the control of full-length or mutant *srsx-3* promoters.

Figure 4.5


A
HNF-1 N-term
homeodomain
Ser-rich
B
MLFTIEQLELIKKLQHTGMSSDQLLKAFGELEVPEQLQNNNTIAAALYSPLLVQHLTTPK
SETPVKLTVQTVPTPVKSEPQSSNCSSPFEHPICSNAPRPIRSQRTPMKEITTLDDPNEL
EEFMKQGEEACILDMKTFITQYSLRQTTVAMMTGVSQPYISKLLNGNHRELSLRCRKNIY
CWYLNCRRHPNKLAAFLADPTTRLETNGDGELIPQRRERYVFRPILIRMLESFFTQTPFP
DLPRRVEIANACNHVLKMDKKGVGLMPKEVVSPQVVSNWFANKRKELRRRSAEASAASTS
SASSSASSTANHDSVSVSSMSPRDEETSSRNTTPETAISPSPAVSTFEVSRPSAIISATS
STTSPISIPATIIPSVSPSALELFAMAQQLGVQLPVPFPTLPTHFFPFQMAPFYGNPASI
LKSE-
H404Y
C
C. elegans
B. malayi
D. rerio
M. musculus
H. sapiens
D
srsx-3 expression
str-2 expression
fluorescence (AU)
wild-type
tm1274
tm1274 rescue
ky777
AWCOFF
AWCON
E
srsx-3 expression
Fraction of animals
0AWC
1AWC(f)
1AWC
2AWC
wild-type
ky777
tm1274
ky777/+
ky777/ tm1274
F
srsx-3 expression
str-2 expression
ky777
nsy-7; ky777
hmbx-1 RNAi
nsy-7 RNAi
G
srsx-3 expression
str-2 expression
transgene
no transgene
hmbx-1 (wild type)
hmbx-1(ky777)
genotype
wild-type
ky777
tm1274
H
hmbx-1::GFP
odr-1::NLS-mCh
merge
I
5'-TTACAAGTTTCTAGTTCAAATAATCA-3'
genotype
promoter
ky777
srsx-3
wild-type
srsx-3ΔBM
srsx-3(WT)
srsx-3(ΔBM)

HMBOX1 genes, we named *F54A5.1 hmbx-1*. The putative *ky777* mutation results in a H→Y missense mutation at the C-terminal end of the HMBX-1 protein (Figure 4.5B, red box). The affected histidine is not highly conserved, and the functional implications of this mutation are not clear.

A deletion allele of *hmbx-1*, *tm1274*, was kindly provided by the National BioResource Project in Japan. The *srsx-3* expression phenotype of the *hmbx-1(tm1274)* mutant was much less severe than that of the *hmbx-1(ky777)* mutant: whereas *srsx-3* expression was entirely absent in >95% of *ky777* mutant adults, expression was present in all *tm1274* animals but reduced in intensity compared to wild-type (Figure 4.5D). The reduced expression levels in the *tm1274* background were rescued by a fosmid including the entire gene and 15kb of upstream region. Neither the *ky777* nor the *tm1274* mutant had significant chemotaxis defects to the $AWC^{OFF}$-specific odor 2,3-pentanedione, so the effects of *hmbx-1* may be quantitative or specific to a subset of GPCRs in $AWC^{OFF}$.

In *ky777*/*tm1274* trans-heterozygotes, 48% of animals did not express GFP, although *ky777* was fully recessive to a wild-type allele of *hmbx-1* (Figure 4.5E). These results suggest that *ky777* may represent a recessive allele that interferes with *hmbx-1* and another unknown target. To further test this hypothesis, *hmbx-1* expression was reduced using RNAi in a *ky777* background. Knockdown of *hmbx-1(ky777)* restored *srsx-3* expression in 57% of *hmbx-1(ky777)* animals (32/56 RNAi+ vs. 4/90 RNAi-, $p<0.0001$ [Fisher's exact test]) (Figure 4.5F), supporting the hypothesis that *ky777* is an altered-function allele of *hmbx-1*. 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

(44/44 RNAi+ vs. 0/90 RNAi- animals, $p<0.0001$ [Fisher's exact test]), consistent with the phenotype of *hmbx-1(ky777);nsy-7(tm3080)* double mutants.

Since the *ky777* allele is recessive to the wild-type *hmbx-1* allele (Figure 4.5E), the wild-type allele must outcompete the mutant allele when both are expressed. Therefore, overexpressing the wild-type *hmbx-1* gene in a *ky777* background might rescue the mutant phenotype. Indeed, a DNA fragment containing the *hmbx-1* genomic coding sequence with 7kb of upstream region partially restored *srsx-3* expression in the *ky777* mutant, but the same fragment bearing the *ky777* C→T mutation did not (Figure 4.5G). Conversely, expression of a *ky777* mutant DNA fragment in the *hmbx-1(tm1274)* null mutant reduced *srsx-3* expression in 20-40% of animals, but a wild-type fragment did not, consistent with enhanced repressive activity of *ky777* (Figure 4.5G).

To determine the site of action of *hmbx-1*, GFP was expressed under the control of 7kb of upstream region. This transgene drove expression in both AWCs (Figure 4.5H), 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. The expression of *hmbx-1* in the AWCs supports the hypothesis that it regulates *srsx-3* expression.

We next cloned the *hmbx-1* cDNA and expressed it in both AWCs in wild-type, *ky777*, and *tm1274* mutant backgrounds. AWC expression of the wild-type cDNA repressed *srsx-3* in wild-type and *tm1274* backgrounds, similar to the phenotype of the *ky777* mutant (Figure 4.6A). This suggests that the wild-type protein has some repressive activity in addition to its role as an enhancer of *srsx-3* expression. Expression of *hmbx-1* in AWC had a mixed effect on *str-2*

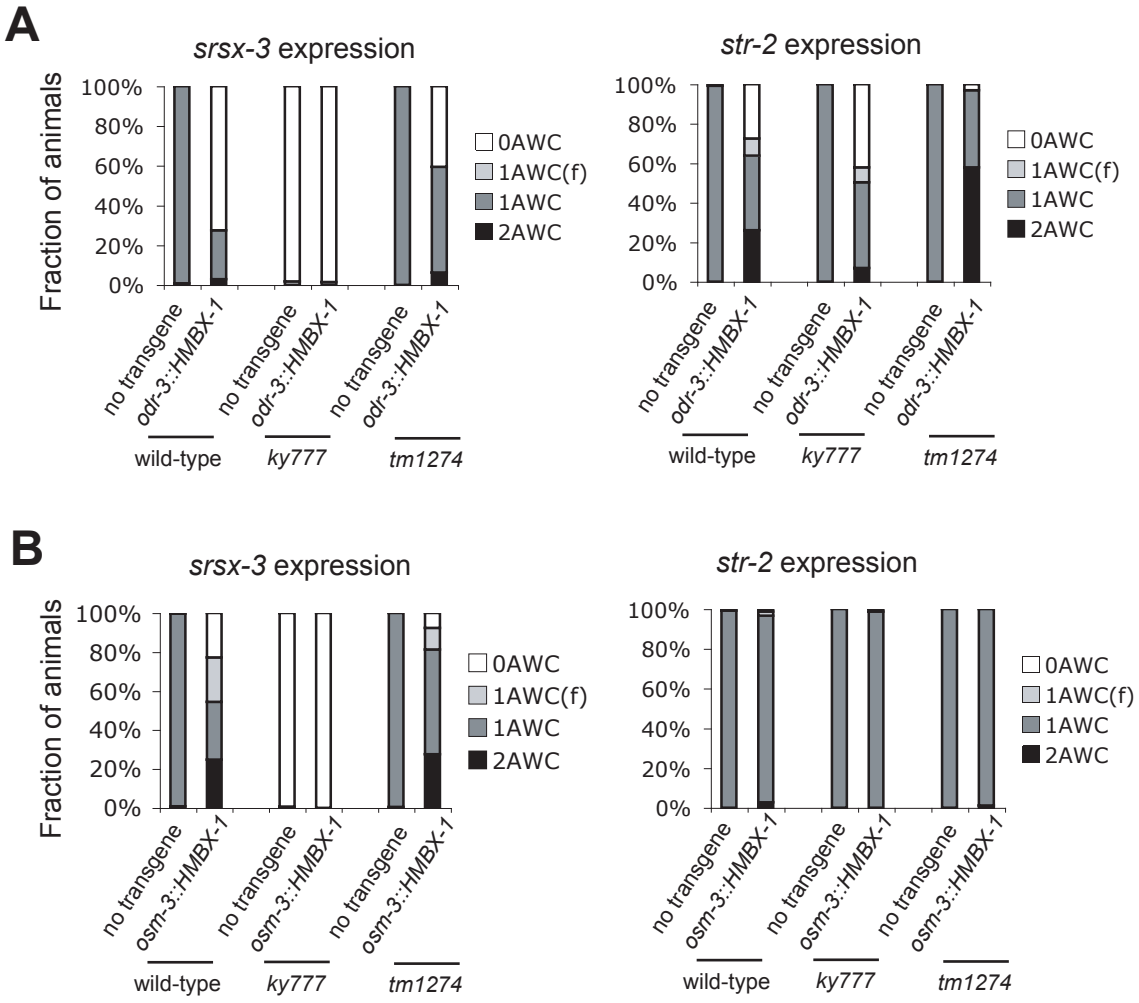



**Figure 4.6. Effects of *hmbx-1* cDNA expression.** A, Expression of *str-2* and *srsx-3* in wild-type, *ky777*, and *tm1274* animals overexpressing a wild-type HMBX-1 cDNA under the control of the AWC-selective *odr-3* promoter. B, Expression of *str-2* and *srsx-3* in wild-type, *ky777*, and *tm1274* animals overexpressing a wild-type HMBX-1 cDNA under the control of the *osm-3* promoter, which drives expression in 26 chemosensory neurons not including AWC.

expression, promoting expression in both AWCs in some worms and repressing it in others (Figure 4.6A). Intriguingly, expression of *hmbx-1* under the *osm-3* promoter, which drives expression in 26 chemosensory neurons, including the amphid neurons ADF, ADL, ASE, ASG, ASH, ASI, ASJ, and ASK, but not in AWC (Tabish et al. 1995), had a weak effect on *srsx-3* but not *str-2* expression, and this effect was opposite that observed with AWC expression. That is, ~25% of worms expressing *hmbx-1* under the *osm-3* promoter expressed *srsx-3* in both rather than one AWC (Figure 4.6B). Enhancement of *srsx-3* expression by *osm-3::hmbx-1* was not observed in the *ky777* mutant. *hmbx-1* may therefore have cell nonautonomous as well as cell autonomous effects, and these two influences antagonize each other in control of *srsx-3*.

**A predicted HMBX-1 binding site is required for expression of the *ky777* mutant phenotype**

Although nothing is known about the HMBX-1 DNA binding site, Berger *et al* predicted the binding site of its mouse homolog, HMBOX1, using an *in vitro* binding assay (Berger et al. 2008). A site similar to the one predicted in mouse is present in the *srsx-3* promoter (Figure 4.5I, top). If *hmbx-1(ky777)* represses *srsx-3* expression by binding to this sequence, then eliminating the sequence should relieve *hmbx-1(ky777)*-mediated repression of *srsx-3*. Indeed, deleting this sequence partially restored expression of *srsx-3* in AWC$^{OFF}$ in a *ky777* background, but had no effect on expression in a wild-type background (Figure 4.5I). The incomplete effect of the binding site deletion may indicate that there are other undiscovered binding sites elsewhere in the *srsx-3* promoter;

alternatively, other *hmbx-1* targets in AWC or cell nonautonomous effects of *hmbx-1* may account for the remaining repression in the *ky777* mutant.

### *daf-3* and *hmbx-1* respond differently to cGMP-based signaling

Null mutants for two transcriptional regulators implicated in maintenance of *srsx-3* expression, *hmbx-1* and the co-Smad *daf-3*, have qualitatively wild-type or subtle phenotypes for expression of AWC markers. To examine the interactions between maintenance pathways, we made double mutants between these mutants and upstream genes in the TGFβ or cGMP signaling pathways.

The relationship between *hmbx-1* and *daf-3* was examined in *hmbx-1(tm1274);daf-3* and *hmbx-1(ky777);daf-3* double mutants. *tm1274;daf-3* double mutants had a qualitatively wild-type phenotype, like each single mutant. *ky777;daf-3* double mutants, like *ky777* single mutants, failed to express *srsx-3* (Figure 4.7A). As *hmbx-1(ky777)* does not require *daf-3* to repress *srsx-3* expression, it is predicted to act downstream of or parallel to *daf-3*. *hmbx-1(tm1274);daf-1* double mutants had a *daf-1*-like phenotype, indicating that *hmbx-1* does not act downstream of the TGFβ signal (Figure 4.7C). Together, these results suggest that *hmbx-1* and *daf-3* are parallel regulators of *srsx-3*.

Double mutants between *daf-3* and the cGMP mutants gave unexpectedly complex results. *daf-3* fully suppressed the *srsx-3* expression defect, but not the *str-2* defect, in *egl-4* (cGMP-dependent kinase) mutants (Figure 4.7B). However, *daf-3* was unable to suppress the *srsx-3* or *str-2* expression defects in *odr-1* (guanylate cyclase) mutants (Figure 4.7B). These results imply that (1) in AWC$^{OFF}$, *odr-1* acts through a pathway that does not require *daf-3* to regulate *srsx-*

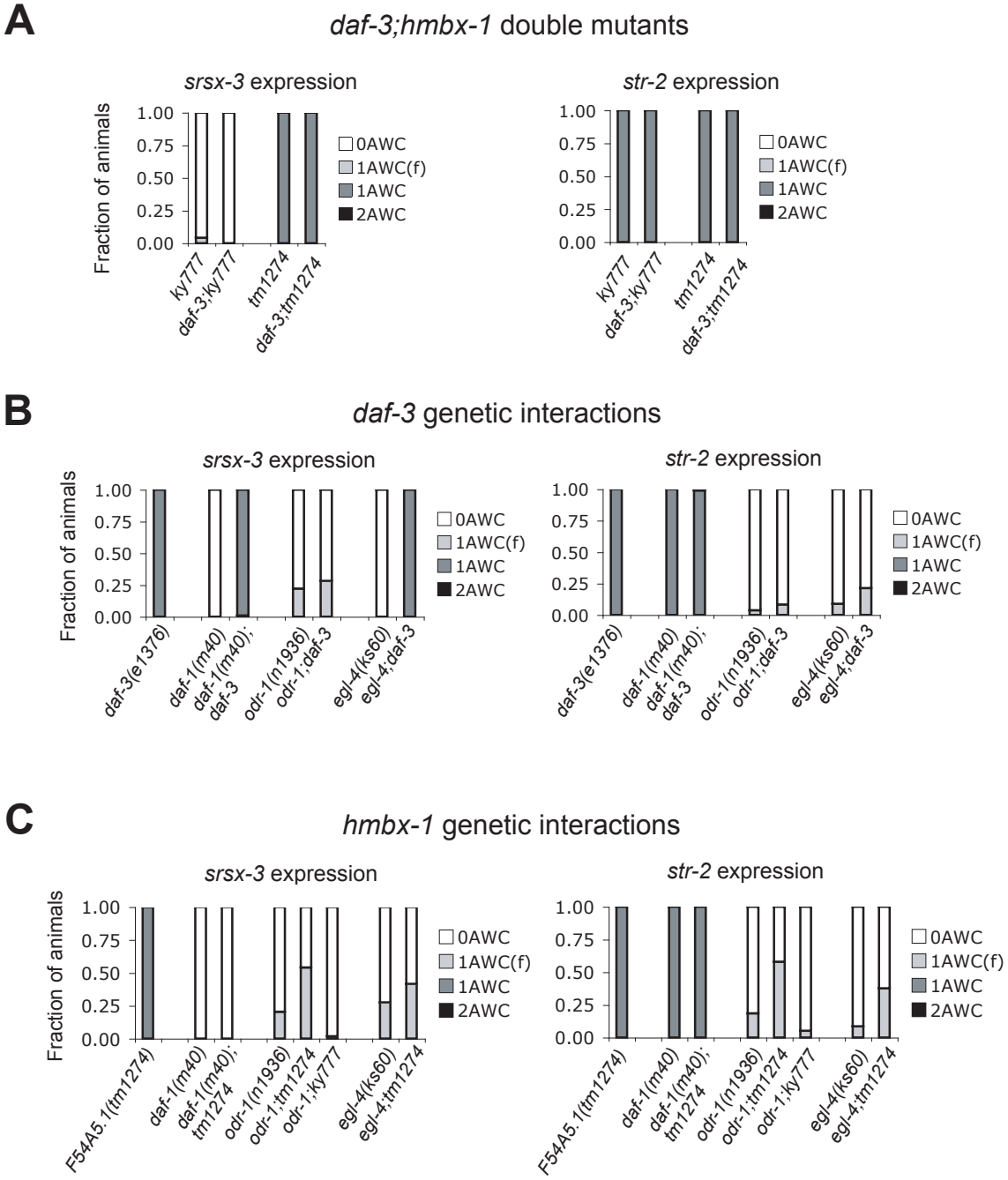



**Figure 4.7. The TGFβ and *hmbx-1* pathways are distinct.** A, Expression of *str-2* and *srsx-3* in *daf-3(e1376);hmbx-1(ky777)* and *daf-3(e1376);hmbx-1(tm1274)* double mutants. B, Expression of *str-2* and *srsx-3* in double mutants between *daf-3(e1376)* and *daf-1(m40)*, *odr-1(n1936)*, and *egl-4(ks60)*. C, Expression of *str-2* and *srsx-3* in double mutants between *hmbx-1(tm1274)* and *daf-1*, *odr-1*, and *egl-4*.

*3* expression, while *egl-4* acts separately and requires *daf-3*, and (2) in $AWC^{ON}$, *odr-1* and *egl-4* regulate *str-2* expression in a pathway that does not involve *daf-3*.

Unlike *daf-3*, *hmbx-1(tm1274)* partially suppressed the *srsx-3* and *str-2* expression defects of the *odr-1* mutant (Figure 4.7C). *tm1274* also partially suppressed the *str-2* but not the *srsx-3* expression defects of *egl-4* mutants. These results suggest that *hmbx-1* could act downstream of *odr-1* and *egl-4*, although these two genes must have additional downstream targets. Furthermore, *odr-1* and *egl-4* were again genetically distinct in their effects on *srsx-3* and *str-2*, and in their dependence on *hmbx-1*. The activity of these two transcriptional regulators therefore depends on the promoter examined and on the cell in which that promoter is expressed.

**An analysis of the NSY-7 binding site reveals additional $AWC^{OFF}$-specific GPCRs**

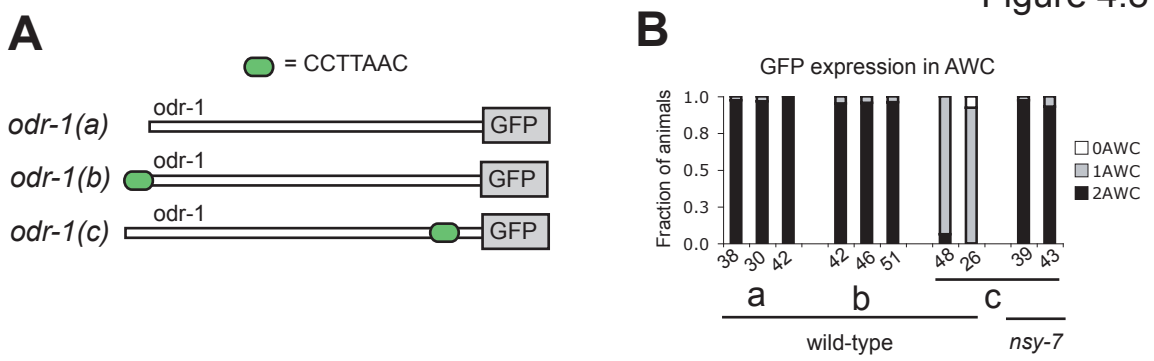
With only one $AWC^{ON}$- and one $AWC^{OFF}$-specific marker, it was not possible to determine whether *hmbx-1*, the TGFβ pathway, and the cGMP pathway affected only *srsx-3* and/or *str-2*, or whether they acted more generally in $AWC^{OFF}$ and $AWC^{ON}$. Therefore, we sought additional genes with expression specific to one of the two AWCs. GPCR expression was examined in approximately twenty of the reporter strains available from the *C. elegans* gene expression database at the University of British Columbia, uncovering three GPCRs, *srt-26*, *srt-28*, and *srt-29*, that were expressed in a single AWC. These three genes are located adjacent to one another on chromosome V, and all were coexpressed with *str-2* but not *srsx-3*, indicating expression specific to $AWC^{ON}$.

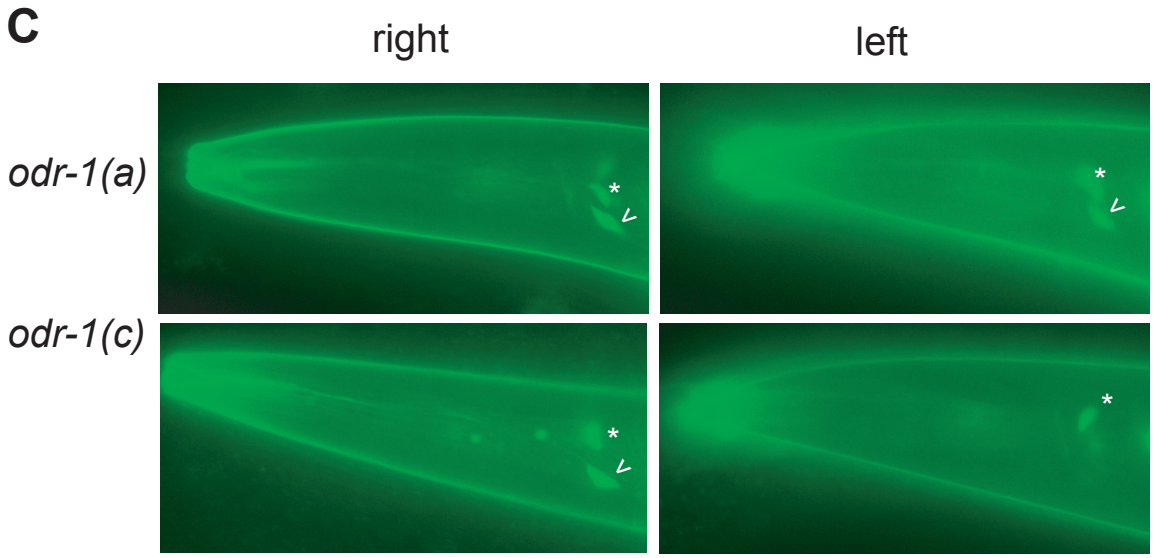
AWC$^{OFF}$-specific reporters were sought using the defined binding site of the transcription factor NSY-7. NSY-7 binds the site CCTTAAC in the *srsx-3* promoter, and acts to repress that promoter in AWC$^{ON}$. If the CCTTAAC sequence is sufficient for AWC$^{ON}$ repression by NSY-7, then insertion of the sequence into a promoter normally expressed in both AWCs should result in AWC$^{OFF}$-specific expression. This idea was tested by inserting a single copy of the CCTTAAC sequence into an *odr-1::GFP* reporter, which is ordinarily expressed in both AWBs and both AWCs. The site was inserted either ~2.4 kb upstream of the ATG (Figure 4.8A, *odr-1(b)*) or 200bp upstream of the ATG (Figure 4.8A, *odr-1(c)*).

Expression of these modified *odr-1::GFP* reporters from extrachromosomal arrays was variable (Figure 4.9), perhaps because of their high copy numbers. Therefore, we made single copy insertions of each reporter at a common site in the genome, using the Mos Single Copy Insertion (MosSCI) technique developed by Frøkjær-Jensen *et al* (Frokjaer-Jensen et al. 2008). A single copy of the wild-type *odr-1* promoter (*odr-1(a)*) and the plasmid containing a -2.4kb distal NSY-7 binding site (*odr-1(b)*) still drove GFP expression in both AWCs, but the *odr-1(c)* plasmid with a -200bp proximal NSY-7 site consistently drove expression in a single AWC (Figure 4.8B,C). A *nsy-7* mutation restored expression to both AWCs, demonstrating NSY-7-dependent repression in on AWC. The GFP-expressing AWC invariably co-expressed *srsx-3* and not *str-2*, indicating that the modified *odr-1* promoter was expressed in AWC$^{OFF}$ and not AWC$^{ON}$ (Figure 4.8D). These results indicate that promoter-proximal NSY-7 binding to its target site is sufficient to repress transcription of a target gene in AWC$^{ON}$.

**Figure 4.8. The NSY-7 binding site predicts AWC$^{OFF}$-specific expression.** A, Schematic of *odr-1(a)*, wild type 2.4-kilobase *odr-1* promoter fragment; *odr-1(b)*, NSY-7 binding site added to the 5' end of the *odr-1* promoter; and *odr-1(c)*, NSY-7 binding site added 200 base pairs upstream of the ATG in the odr-1 promoter. B, Expression of integrated single-copy *odr-1* plasmids in wild-type and *nsy-7(tm3080)* backgrounds. Numbers on the x-axis indicate number of animals scored. Multiple bars within the same condition represent independently integrated lines. C, Images of integrated constructs *odr-1(a)* and *odr-1(c)*, showing *odr-1(a)* expressed in AWB and AWC bilaterally and *odr-1(c)* expressed in AWB bilaterally and AWC unilaterally. Arrowheads, AWC; asterisks, AWB. D, Coexpression of *odr-1(c)::GFP* with *str-2::dsRed2* or *srsx-3::mCherry*. E, *srt-45* and *srt-47* promoters each contain the NSY-7 binding site within 300 bp of the ATG, and are coexpressed with the AWC$^{OFF}$-specific reporter *srsx-3::mCherry*.

Figure 4.8

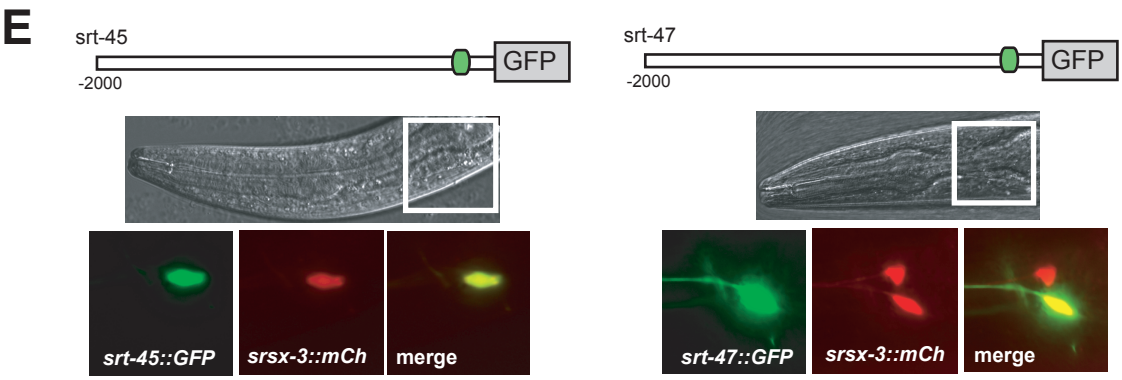



D

| line # | % coexpressing GFP with srsx-3 | n |
|---|---|---|
| 1 | 100 | 25 |
| 2 | 100 | 30 |

| line # | % coexpressing GFP with str-2 | n |
|---|---|---|
| 1 | 0 | 36 |
| 2 | 0 | 30 |



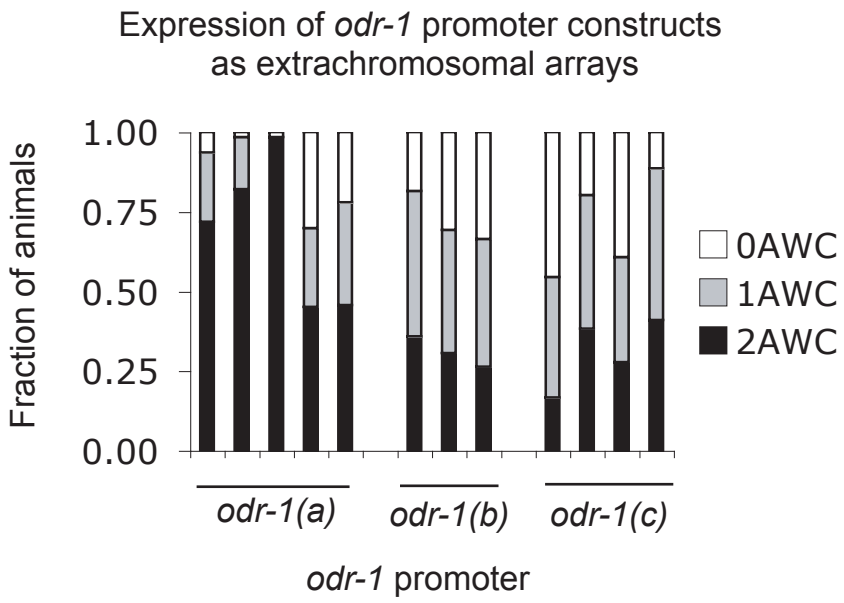



**Figure 4.9. *odr-1(a)*, *odr-1(b)*, and *odr-1(c)* plasmids from Figure 4.8 expressed as extrachromosomal arrays.** Different bars within each promoter group represent independent lines expressing the indicated reporter.

Based on these results, we searched the genome for GPCRs containing the CCTTAAC sequence within 300 bp of the coding start site. Fourteen GPCRs with promoters that fit these criteria were further examined by placing GFP under the control of 2kb of upstream region. Reporter genes for two GPCRs, *srt-45* (on chromosome V) and *srt-47* (on chromosome IV), were expressed in AWC$^{OFF}$ but not AWC$^{ON}$, based on coexpression with the *srsx-3::mCherry* reporter (Figure 4.8E). In *nsy-7* mutants, *srt-45* and *srt-47* were expressed in both AWCs, consistent with the hypothesis that the proximal NSY-7 binding sites confer AWC$^{OFF}$-specific expression (Figure 4.8B).

In combination with previous work, these experiments uncovered three AWC$^{OFF}$ GPCR markers, *srsx-3*, *srt-45*, and *srt-47*, and four AWC$^{ON}$ GPCR markers, *str-2*, *srt-26*, *srt-28*, and *srt-29*.

**AWC$^{ON}$ and AWC$^{OFF}$-specific GPCRs are regulated in tandem by *hmbx-1* and individually by TGFβ and cGMP signaling**

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 affecting maintenance of *srsx-3* and *str-2* expression. Correlated changes in expression within the sets of AWC$^{ON}$ and AWC$^{OFF}$ markers might indicate generalized effects in AWC$^{OFF}$ or AWC$^{ON}$ cell identities, while independent changes would indicate regulation of expression at the level of the gene rather than of the cell. In *nsy-7* mutants, all AWC$^{OFF}$ markers were expressed in both AWCs, and all AWC$^{ON}$ markers were reduced or absent (Figure 4.10B). Thus, *nsy-7* regulates expression on a cell-wide level, affecting both AWC$^{ON}$ and AWC$^{OFF}$ markers.

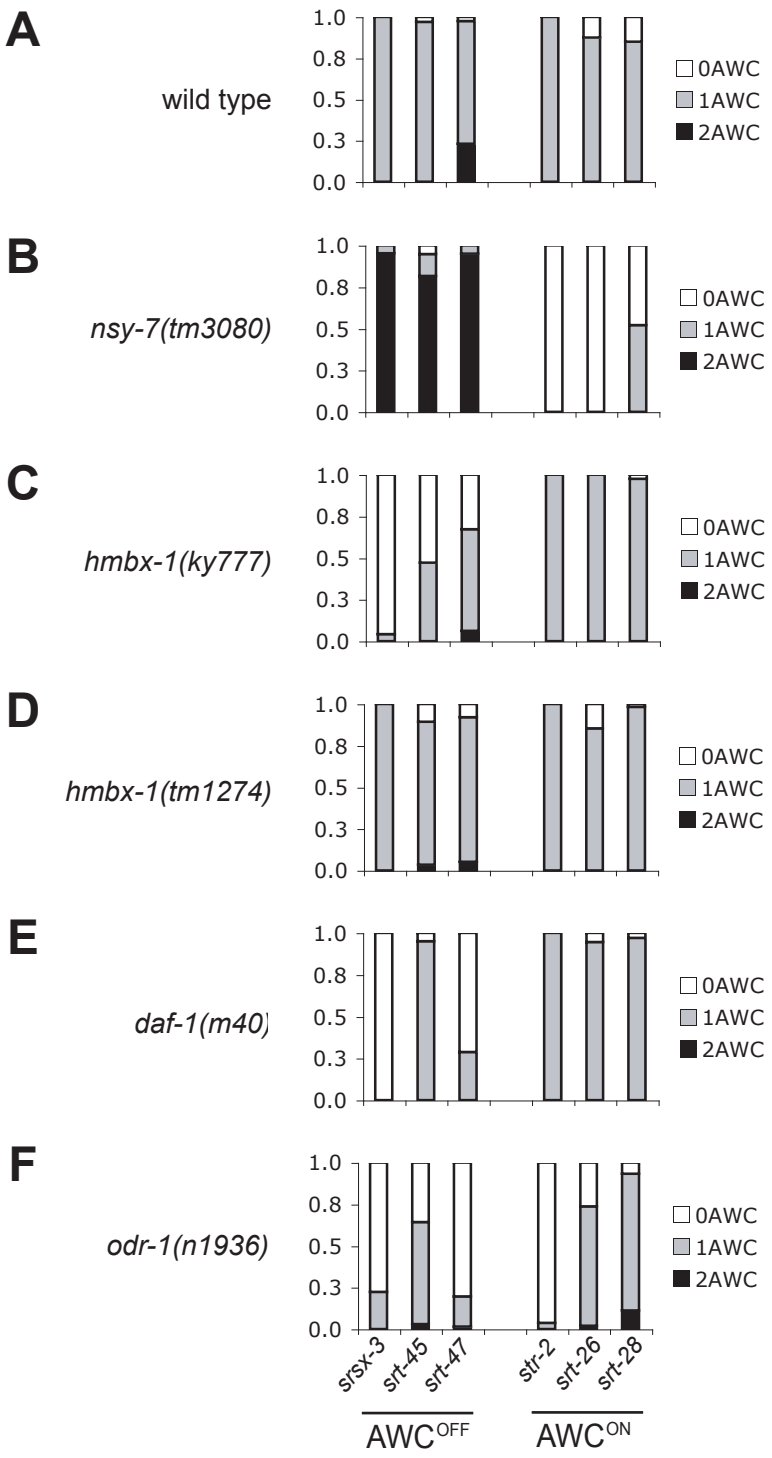



**Figure 4.10. Expression of the new AWC$^{OFF}$ markers *srt-45* and *srt-47* and the new AWC$^{ON}$ markers *srt-26* and *srt-28* in wild-type and mutant backgrounds.** *srsx-3* and *str-2* expression phenotypes are shown for comparison. A, wild-type. B, *nsy-7(tm3080)*. C, *hmbx-1(ky777)*. D, *hmbx-1(tm1274)*. E, *daf-1(m40)*. F, *odr-1(n1936)*.

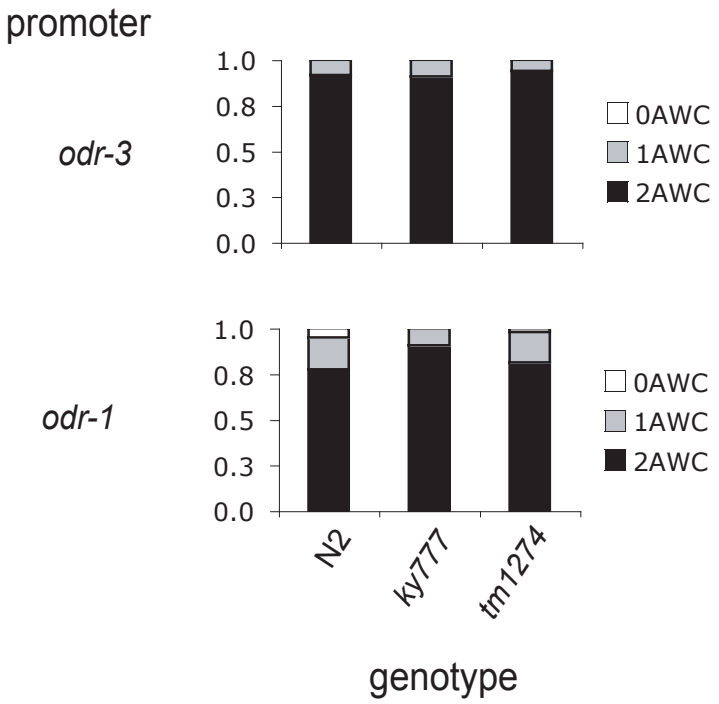



**Figure 4.11. Expression of *odr-3::GFP* and *odr-1::GFP* reporters in wild type, *hmbx-1(ky777)*, and *hmbx-1(tm1274)* animals.**

*hmbx-1* also regulated expression in a cell-wide manner, but only affected $AWC^{OFF}$. *hmbx-1(ky777)* reduced expression of all three $AWC^{OFF}$ markers, but none of the $AWC^{ON}$ markers (Figure 4.10C). *hmbx-1(tm1274)* had little qualitative effect on any markers examined (Figure 4.10D). *hmbx-1(ky777)* did not alter expression of the AWC signaling genes *odr-3* and *odr-1* in $AWC^{OFF}$, indicating that its cell-specific effect on GPCR expression is not due to generalized transcriptional repression in $AWC^{OFF}$ (Figure 4.11).

In *daf-1* mutants, expression of the $AWC^{OFF}$ markers *srsx-3* and *srt-47* was reduced, but *srt-45* expression was not (Figure 4.10E). All $AWC^{ON}$ markers were expressed at wild-type levels. These mixed effects on $AWC^{OFF}$ markers imply that *daf-1* and the TGFβ pathway alter expression at the level of individual receptor genes, and the apparent insensitivity of $AWC^{ON}$ markers suggests that the effects of the TGFβ pathway are restricted to $AWC^{OFF}$.

*odr-1* mutants were strongly defective for expression of the $AWC^{OFF}$ markers *srsx-3* and *srt-47*, but not *srt-45*, and the $AWC^{ON}$ marker *str-2* but not *srt-26* or *srt-28* (Figure 4.10F). Like TGFβ signaling, therefore, *odr-1* cGMP signaling affects specific receptors rather than suites of cell-specific GPCRs, but unlike the TGFβ pathway, the cGMP pathway affects both $AWC^{ON}$ and $AWC^{OFF}$.

### *hmbx-1* regulates *srsx-3* differently in $AWC^{ON}$ compared to $AWC^{OFF}$

To better separate cell-specific from promoter-specific effects, we examined expression of a single promoter, *srsx-3*, in the context of $AWC^{ON}$ and $AWC^{OFF}$. The *srsx-3(N7bs11)* promoter fragment includes two point mutations in the NSY-7 binding site, eliminating repression by NSY-7 in $AWC^{ON}$ (Lesch et al.

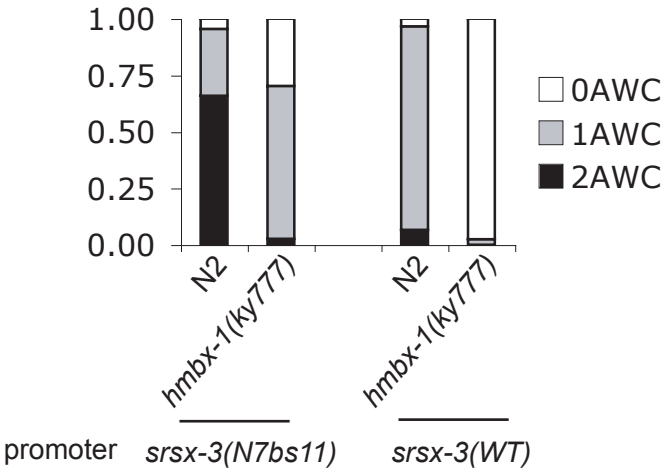



**Figure 4.12. Expression of wild-type *srsx-3::GFP* and mutant *srsx-3(N7bs11)::GFP* reporters in wild type and *ky777* animals.** The *srsx-3(N7bs11)* promoter contains two AG transitions in the NSY-7 binding site, allowing ectopic expression of srsx-3 in $AWC^{ON}$.

2009). Thus, this promoter is expressed in AWC$^{ON}$ as well as AWC$^{OFF}$, but does not affect either AWC cell fate. *hmbx-1(ky777)* mutants expressed the *srsx-3(N7bs11)* promoter predominantly in one AWC neuron (Figure 4.12). Experiments are in progress to determine whether, the single cell expressing the *srsx-3(N7bs11)* promoter is AWC$^{ON}$, and to determine the expression of the *srsx-3(N7bs11)::GFP* reporter in *odr-1* and *daf-1* backgrounds.

**Discussion**

Environmental stimuli and the underlying identity of the cell cooperate to determine the set of receptors expressed by the AWCs. The cGMP-dependent pathway that transduces odor signals in AWC regulates a subset of asymmetrically-expressed GPCRs in both AWCs, while the *daf-7*/TGFβ pathway maintains expression of an overlapping but not identical set of GPCRs, is restricted to AWC$^{OFF}$ in its effects, and responds acutely to external pheromone signals. The previously uncharacterized transcription factor *hmbx-1* also plays a role in maintenance of AWC$^{OFF}$-specific receptor genes. *hmbx-1*, TGFβ, and the cGMP pathway interact so that each receptor is regulated by a specific combination of inputs. Expression of the *srsx-3* promoter responds to different maintenance signals in AWC$^{ON}$ or AWC$^{OFF}$. Thus, the developmental program associated with a given cell such as AWC$^{ON}$ or AWC$^{OFF}$ determines what receptors can be expressed by that cell, while environmental factors alter the particular set of receptors it expresses, and therefore the stimuli to which it can respond, at any given time (Figure 4.13). The combination of two AWC cells, at least two types of environmental stimulus, and multiple receptors may allow the

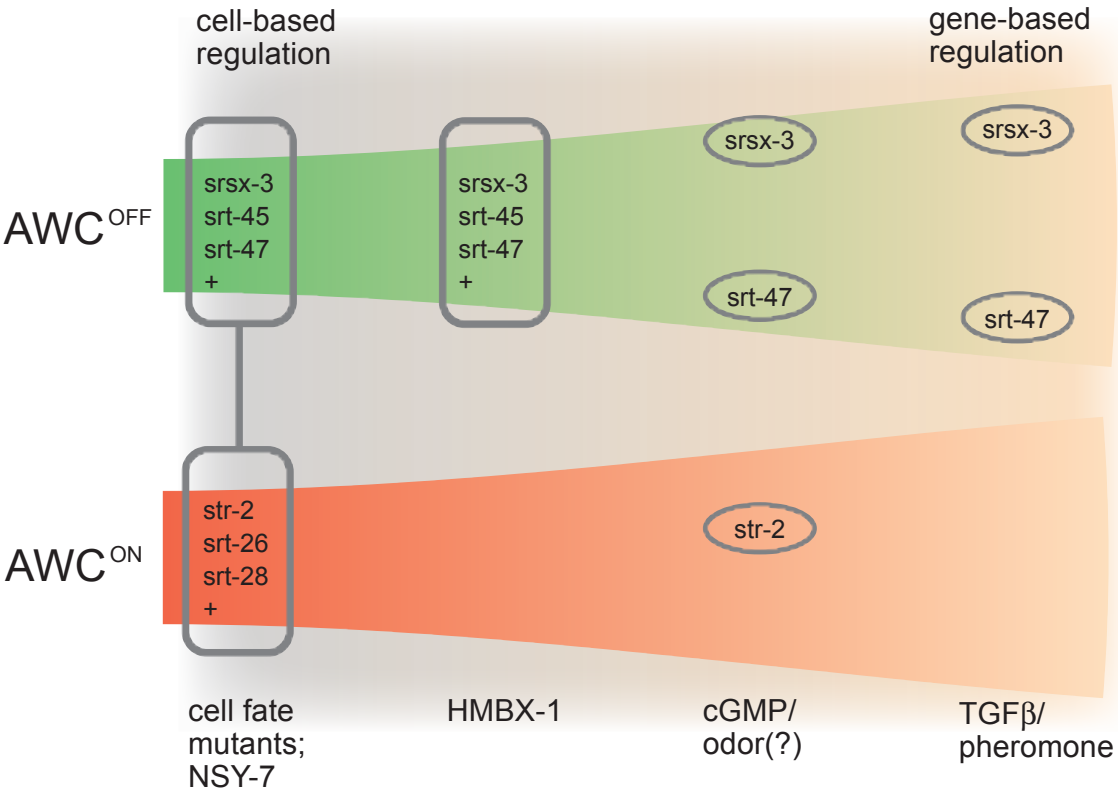



**Figure 4.13. Cell identity and environmental factors influence GPCR expression.** Cell identity determines the set of receptors expressed by a given neuron (regulated by NSY-7) and sensitivity of that neuron to environmental influences (regulated by HMBX-1). The effects of environmentally-modulated signals such as cGMP and TGFβ depend on the receptor gene and on the cell in which the receptor gene is expressed.

animal to tune the responsiveness of its sensory system to its specific surroundings.

Many of the maintenance-defective mutants from our screen affect the olfactory signal transduction pathway. However, different genes within the same pathway have stronger or weaker effects on *str-2* and *srsx-3*. Some, such as the guanylate cyclase *daf-11* and the G-protein α subunit *odr-3*, appear to have a greater effect on *str-2*, while others, such as the cGMP-gated channel subunits *tax-2* and *tax-4*, preferentially affect *srsx-3*. The *tax-4(ky791)* expression defect is rescued in AWC, suggesting cell-autonomous action.

TGFβ signaling targets receptor genes that overlap with those regulated by the cGMP pathway, but acts through different downstream effectors. cGMP and TGFβ downstream signaling converges on the cGMP-dependent kinase *egl-4*, which requires *daf-3* to regulate expression of *srsx-3*. *egl-4* also requires *daf-3* to mediate many additional effects in *C. elegans*, including dauer formation, body size, egg laying, and chemosensation, and the cGMP-dependent kinase PKG enhances BMP signaling in mammalian cells via direct interactions with the BMP receptors and downstream R-Smads (Daniels et al. 2000; Schwappacher et al. 2009). *egl-4* likewise represents a point of intersection between the cGMP and TGFβ pathways in *srsx-3* regulation in $AWC^{OFF}$. By contrast, *daf-3* is not required for regulation of *str-2* by *egl-4*, or for *odr-1* regulation of either *str-2* or *srsx-3*.

The divergence between regulation of *str-2* and regulation of *srsx-3* by both the cGMP and TGFβ pathways could represent a difference between the two promoters, or a more fundamental difference between $AWC^{ON}$ and $AWC^{OFF}$.

To address this question, we identified additional asymmetrically-expressed GPCRs in AWC$^{ON}$ and AWC$^{OFF}$. Analysis of these markers uncovered differences in promoters expressed in the same cell; in AWC$^{OFF}$, for example, *srt-47* is strongly regulated by TGFβ whereas *srt-45* is not. However, *hmbx-1* mutants affect expression of all three AWC$^{OFF}$ receptors, but no AWC$^{ON}$ receptors, indicating that HMBX-1 may have a more general function in AWC$^{OFF}$. A mutant *srsx-3* promoter that is expressed in both AWC$^{ON}$ and AWC$^{OFF}$ in wild-type animals is expressed predominantly in one of the two AWCs in *hmbx-1(ky777)* mutants; since *hmbx-1(ky777)* strongly represses *srsx-3* expression in AWC$^{OFF}$, the cell that retains expression in this context is more likely to be AWC$^{ON}$, and therefore the effects of *hmbx-1* may be AWC$^{OFF}$-specific. Experiments are underway to examine expression of the mutant *srsx-3(N7bs11)* reporter in cGMP and TGFβ pathway mutants. The *srsx-3* promoter mutation specifically alters a NSY-7 binding site and is unlikely to affect general cell fate in AWC$^{ON}$ or AWC$^{OFF}$; we will therefore be able to compare the response of a single promoter to cGMP, TGFβ, and *hmbx-1* in two different cells, and thereby evaluate the extent to which the cell's fate influences its responsiveness to each of these factors.

*hmbx-1* is expressed predominantly in sensory neurons. Its human and mouse HMBOX1 homologs also display high levels of nervous system expression, albeit in the hypothalamus, hippocampus, and cerebral cortex. The strong sequence conservation of this gene, along with its conserved functional role as a repressor, its apparently conserved binding site specificity, and its neuronal expression pattern in both worm and mammal, hints at possible

conserved neuronal functions (Chen et al. 2006). The recessive gain-of-function allele *ky777* may be a useful tool for untangling its regulatory role.

Our results imply that single transcription factors and single binding sites can play a surprisingly large role in GPCR expression in *C. elegans* sensory neurons. For example, two mutations (one amino acid change in *hmbx-1* and two nucleotide changes in the *srsx-3* promoter) are sufficient to switch GPCR expression from one AWC to the other. 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 the *C. elegans* chemosensory neurons.

The presence of a single binding site for the transcriptional repressor, NSY-7, is sufficient to repress expression of a transgene when it is present as a stably-integrated single copy, although not when it is part of a multi-copy array. Many promoters contain several redundant transcription factor binding sites and therefore will be relatively insensitive to such effects. However, some promoters may contain only one crucial binding site, and the specific expression pattern of these promoters can then be easily altered by mutation. Particularly in *C. elegans*, where ectopic expression of a gene in a single additional neuron can drastically change developmental pathways and behavioral states, altered expression due to natural mutation might have significant consequences (Troemel et al. 1997). Moreover, transcription factors controlling neuronal development in *C. elegans* are thought to be evolving more rapidly than structural genes, particularly in chemosensory neurons (Jovelin 2009). Thus changes in transcription factors and transcription factor binding sites have the potential to drive rapid adaptive changes in chemosensory neurons.

The results presented here indicate that fundamental cell identity determines the set of receptors expressed by a given cell, and also the extent to which that receptor repertoire can be modified by environmental inputs. These environmental inputs determine the combination of receptors that is actually expressed at a given time, and hence the nature of the stimulus perceived by the animal. For the AWC neurons, such environmental influences include dauer pheromone, transduced by a TGFβ signal, and odor, transduced by cGMP signaling. The conserved transcription factor *hmbx-1* alters the sensitivity of the $AWC^{OFF}$ neuron to these influences. It will be interesting to determine the extent to which environmentally-induced changes in receptor expression, described here, correspond to changes in odor detection and chemotaxis, a behavioral response that depends directly on the set of receptors expressed.

# Chapter 5

## Discussion and future directions

This work begins to elucidate the mechanisms by which $AWC^{ON}$ and $AWC^{OFF}$ identities are maintained. However, the picture that is emerging is far from complete. Several transcriptional regulators are involved in maintenance, but more remains to be discovered about the molecular mechanisms by which these factors operate to maintain the $AWC^{ON}$ and $AWC^{OFF}$ fates. Additional transcription factors will probably contribute to the regulation of *str-2* and *srsx-3*, and regulation of other genes specific to each of the two AWCs is likely to be equally complex. Furthermore, some preliminary evidence suggests that other neurons, such as the ASI sensory neurons, participate in maintenance of GPCR expression in $AWC^{OFF}$; future studies should refine the identity of these cells, their exact role in regulating expression in $AWC^{OFF}$, and the means by which they accomplish this function. A final question involves the relationship between expression of the *str-2* and *srsx-3* markers and the fundamental identities of the $AWC^{ON}$ and $AWC^{OFF}$ neurons; some factors evidently regulate expression of these markers without universally affecting neuronal identity. It will be interesting to learn whether there is any correlation between changes in receptor expression and changes in chemotaxis behavior, especially in response to environmental inputs. A more thorough understanding of this connection will further reveal the nature of identity maintenance in the AWC neurons.

**Transcription factors in AWC$^{ON}$ and AWC$^{OFF}$**

These studies revealed several transcriptional regulators that participate in the establishment and maintenance of the AWC$^{ON}$ and AWC$^{OFF}$ cell fates (Table 5.1). NSY-7 is a divergent homeodomain protein that controls AWC$^{ON}$ identity in several ways. It is required to repress AWC$^{OFF}$ markers in AWC$^{ON}$; it acts to maintain the expression of *str-2* and possibly other AWC$^{ON}$ markers; and it regulates expression of at least one downstream transcription factor, HLH-11. NSY-7 does not appear to be required to maintain its own expression pattern, arguing that other factors are involved to promote its asymmetric expression in AWC$^{ON}$. NSY-7 is a sequence-specific DNA binding factor and might be able to recruit a more general machinery to mediate long-term repression of its targets. Understanding the factors that mediate NSY-7 function will be important for understanding identity maintenance in the AWC neurons.

*hlh-11* represents a second transcription factor that is asymmetrically expressed, in this case in AWC$^{OFF}$, but its precise function is unclear. Based on its overexpression phenotype, it appears to have repressive activity (Figure 2.2D) and it is predicted to bind to the promoters of several GPCRs and other transcription factors (Grove et al. 2009). However, it is not required for *str-2* and *srsx-3* expression and may act redundantly in AWC.

HMBX-1, another homeodomain protein, is expressed in both AWC neurons but appears to act predominantly in AWC$^{OFF}$. A missense allele of HMBX-1 affects expression of all known AWC$^{OFF}$ markers but does not seem to alter AWC$^{ON}$ marker expression, and the mechanisms that limit its function to AWC$^{OFF}$ have not been examined. Understanding these mechanisms may be one

## Table 5.1. Transcription factors involved in maintaining the AWC$^{ON}$ and AWC$^{OFF}$ fates.

| transcription factor | binding site | binding site evidence | mutant phenotype | AWC overexpression phenotype (*odr-3p*) | expression pattern |
|---|---|---|---|---|---|
| CEH-36 | **TAATCC** | EMSA (Etchberger et al. 2009) | *ky640, ky646*: no *str-2* or *srsx-3* expression in AWC<br>*ky793*: *srsx-3* defective, *str-2* wild-type | mixed phenotype | AWC (both), ASE[1] |
| NSY-7 |  | PBM & EMSA (Lesch et al. 2009) | *tm3080*: *srsx-3* in both AWCs, no *str-2*<br>*ky630*: *srsx-3* in both AWCs, faint *str-2* in some worms | represses *srsx-3* and drives *str-2* expression in both AWCs | AWC$^{ON}$; ASE, ASH, other head neurons, amphid sheath cells, intestine |
| HLH-11 |  | PBM (Grove et al. 2009) | *tm2944*: wild type | represses both *srsx-3* and *str-2* expression | AWC (both; bias toward AWC$^{OFF}$), intestine, other cells |
| HMBX-1 |  (mouse) | PBM (Berger et al. 2008) | *ky777*: no *srsx-3* expression in AWC<br>*tm1274*: reduced *srsx-3* expression in AWC | represses *srsx-3* expression; both enhances and represses *str-2* expression | AWC (both), ASI, ASH, AFD, URX, PVD, FLP, ALM, PLM, seam cells, tail neurons |
| DAF-3 | **GTCTC** | EMSA (Thatcher et al. 1999) | suppression of *srsx-3* expression defect in *daf-7*/TGF-β pathway mutants | n.d. | head neurons, ventral nerve cord, tail neurons, intestine, variable hypodermal expression[2] |

[1]Lanjuin et al. 2003; Koga and Oshima 2004
[2]Patterson et al 1997

gateway into understanding functional differences between $AWC^{ON}$ and $AWC^{OFF}$.

CEH-36, a fourth transcription factor, was previously known to be required for establishment of general AWC identity, but the *ceh-36(ky793)* allele isolated from my screen reveals a specific role in maintenance of *srsx-3* expression. Based on this phenotype, *ceh-36* appears to act in a similar manner to *che-1* in the ASE neurons: it has a general role in establishment of bilateral AWC cell fate, and subsequently assumes a role in maintenance of asymmetric AWC identity (Etchberger et al. 2009). *ceh-36* therefore represents an example of a transcription factor involved in early cell fate decisions whose ongoing expression in the cell helps to maintain that cell's identity, although its role in identity establishment differs from its role in maintenance. Because *ky793* encodes a missense mutation in the CEH-36 binding site, an interesting possibility is that it competes with NSY-7 for binding to the *srsx-3* promoter. If so, CEH-36 may represent the 'activator' predicted in Chapter 1 to bind a site overlapping the NSY-7 site in this promoter.

Interestingly, while null *ceh-36* mutants lose expression of both *srsx-3* and *str-2*, they mis-express the $AWC^{OFF}$ markers *srt-45* and *srt-47* in both AWCs (data not shown). A possible explanation for this complex pattern is that unlike *srsx-3*, *srt-45* and *srt-47* may be regulated directly by NSY-7 but not by CEH-36. In this scenario, failure to express *ceh-36* leads to loss of *nsy-7* expression in $AWC^{ON}$, de-repressing *srt-45* and *srt-47* expression in that cell. If this is the case, the role of *ceh-36* in the AWC neurons is more complicated than previously thought: it controls general AWC fate, it maintains expression of at least one gene in $AWC^{OFF}$, and it indirectly restricts two others to $AWC^{OFF}$ through its effect on

*nsy-7* in AWC$^{ON}$. The expression phenotypes of *srt-45* and *srt-47* have not been examined in a *ceh-36(ky793)* background. Assessment of *nsy-7*, *srt-45*, and *srt-47* expression in *ceh-36(ky793)* mutants will help to test this hypothesis, and to determine whether the protein encoded by the *ceh-36(ky793)* allele is specifically defective for binding to *srsx-3* or is also defective for binding to other possible targets.

The Smad-like DNA-binding protein DAF-3 is also involved in maintenance of gene expression in the AWCs, in this case specifically affecting expression of the AWC$^{OFF}$ promoters *srsx-3* and *srt-47*. DAF-3 antagonizes DAF-7/TGFβ signaling by repressing target genes. However, an *srsx-3* promoter lacking the single predicted *daf-3* binding site is still repressed in a *daf-1* mutant background (data not shown), in a *daf-3*-dependent manner. Either there are other undiscovered DAF-3 binding sites in the *srsx-3* promoter, or DAF-3 interferes with *srsx-3* expression in a more complex indirect way. DAF-3 represents a point of intersection between two otherwise separate signals controlling maintenance of GPCR expression in the AWCs, TGFβ signaling and cGMP signaling. *daf-3* mutants suppress both *daf-1* and *egl-4* (PKG) defects in *srsx-3* expression, although *daf-3* does not suppress the *odr-1* (guanylate cyclase) phenotype. This result highlights the requirement for additional downstream effectors of *odr-1* function, since *odr-1* cannot act solely through *egl-4* to regulate *srsx-3*.

In addition to the five sequence-specific DNA binding factors described, two other factors that have not been shown to bind directly to DNA participate in regulation of *str-2* and/or *srsx-3* expression. First, DAF-5, a divergent Sno/Ski

protein that acts as a cofactor for DAF-3, is required along with DAF-3 to repress *srsx-3* downstream of the TGFβ signal. Second, the *Y42H9AR.4* gene was implicated in regulation of *str-2* by RNAi experiments; the predicted product of this gene is basic and contains some similarity to histone-like proteins, and therefore might act as a transcriptional regulator. However, we have not characterized the nature of this protein or its function in AWC.

Although I have implicated several transcription factors in maintaining the $AWC^{ON}$ and $AWC^{OFF}$ identities, it is unlikely that these represent the full complement of transcriptional regulators required to carry out and maintain the AWC asymmetry decision. It is notable that most of the transcription factors identified act as repressors; CEH-36 may constitute an exception to this pattern. Additional transcriptional activators, particularly in $AWC^{ON}$, may help to fully explain the regulation of asymmetric markers in the AWCs. The predicted transcriptional targets of NSY-7 and HLH-11 may represent some of these missing factors. In particular, *ceh-1*, whose promoter contains predicted binding sites for both NSY-7 and HLH-11, might be a promising candidate to examine further.

**New markers of AWC asymmetry**

I identified three new GPCRs expressed in $AWC^{ON}$ but not $AWC^{OFF}$, and two new GPCRs expressed in $AWC^{OFF}$ but not $AWC^{ON}$. The availability of seven asymmetric AWC chemoreceptor genes, in addition to the two asymmetrically-expressed transcription factors *nsy-7* and *hlh-11*, is a significant improvement over the two that had previously been identified, and allowed me to ask more sophisticated questions about the nature of the $AWC^{ON}$ and $AWC^{OFF}$ identities

and the genes involved in their regulation. Interestingly, all five new $AWC^{ON}$ or $AWC^{OFF}$-specific receptors are members of the *srt* family of serpentine GPCRs. This correlation may reflect some functional significance for this family in AWC. One intriguing possibility is that some members of this family might encode the receptors for odors known to be sensed by the AWCs, including benzaldehyde and isoamyl alcohol, sensed by both neurons, butanone, sensed by $AWC^{ON}$, and 2,3-pentanedione, sensed by $AWC^{OFF}$ (Bargmann et al. 1993; Wes and Bargmann 2001). *srt-26*, *srt-28*, and *srt-29* are therefore candidate receptors for the odor butanone, and *srt-45* and *srt-47* are candidate receptors for 2,3-pentanedione. In addition to these five genes, the family includes 69 additional receptors; the expression and function of these genes in AWC may be interesting to assess.

**Missing components of the maintenance pathways**

The *daf-7*/TGFβ pathway is well understood, and appears to act similarly to maintain *srsx-3* expression as it does in other contexts in *C. elegans*. However, the means by which cGMP produced by *odr-1* and *daf-11* acts to communicate with the nucleus and regulate gene expression is unclear. *egl-4*, the cGMP-dependent kinase, has a similar mutant phenotype to *odr-1*; however, while *egl-4* may mediate cGMP-dependent control of *str-2* signaling, *egl-4* and *odr-1* must act independently to regulate *srsx-3* expression, as *egl-4* requires *daf-3* while *odr-1* does not. *odr-1* regulation of *srsx-3* may depend on the TAX-2/TAX-4 cGMP-gated cation channel. *tax-4* acts cell-autonomously in AWC, and both *tax-2* and *tax-4* more strongly affect *srsx-3* compared to *str-2* expression. However, we do not know what factors might transduce the signal produced by the TAX-2/TAX-4 channel to the nucleus. To discover factors that act downstream of *odr-1* to

regulate gene expression, it would be interesting to perform a screen for suppressors of *odr-1*. Mutants that suppress the *odr-1* defect for *str-2* and/or *srsx-3* expression would be easy to identify, and a screen for these mutants might identify genes that, like *daf-3* and *daf-5*, have a wild-type phenotype by themselves but can be acutely inhibited by an upstream signal such as cGMP.

We are also missing information about the upstream signals responsible for regulation of the *hmbx-1* transcription factor. Analysis of double mutant phenotypes indicates that *hmbx-1* may act downstream of *odr-1* and *egl-4*, but *hmbx-1* deletion mutants only weakly suppress the *odr-1* and *egl-4* phenotypes. *odr-1* and *egl-4* therefore require additional cGMP-responsive transcriptional regulators, and it is likely that *hmbx-1* responds to alternative upstream signals. To begin to better understand the role of *hmbx-1*, it will be interesting to look at the interaction of this gene with *nsy-7* and *ceh-36*, the two other homeodomain transcription factors known to be involved in AWC cell fate. For example, the expression pattern of *hmbx-1* in *nsy-7* and *ceh-36* mutants will help to determine whether *hmbx-1* plays an early role in determination of AWC identity, or whether its expression in AWC requires *ceh-36* or *nsy-7*. Conversely, the expression pattern of *nsy-7* in a *hmbx-1* mutant background can be analyzed to determine whether *hmbx-1* helps to specify *nsy-7* expression.

**Cell nonautonomous effects on expression in AWC$^{ON}$ and AWC$^{OFF}$**

Another emerging pattern from this work is the influence of other neurons on maintenance of marker expression in AWC$^{ON}$ and AWC$^{OFF}$. The foremost example of these effects is the requirement for TGFβ signaling in maintenance of *srsx-3* and *srt-47* expression. DAF-7/TGFβ is released from the ASI sensory

neurons and acts on many tissues to repress the transcriptional changes that control dauer formation (Ren et al. 1996; Schackwitz et al. 1996). Expression of *daf-7* in ASI is thought to be regulated by detection of pheromone by the ASK neurons (Kim and Sengupta, in press). As a result, regulation of GPCR expression in AWC$^{OFF}$ in response to pheromone depends on the relay of information through three cells, from ASK to ASI and from ASI to AWC. Because ASK does not synapse onto ASI, information transfer from ASK to ASI may occur through secreted peptides or may require an additional intermediate cell such as AIA, which is postsynaptic to ASK and forms gap junctions with ASI (White 1986). Maintenance of receptor expression in AWC$^{OFF}$ therefore extends beyond cell-intrinsic factors, requiring environmental input (pheromone) as well as the activity of several other neurons in the network.

HMBX-1, which has a cell-autonomous function in regulation of the *srsx-3* promoter in AWC$^{OFF}$ (Figure 4.5I), also appears to be capable of acting in other neurons to influence *srsx-3* expression. *hmbx-1* is normally expressed in several head neurons in addition to AWC, and when overexpressed under the *osm-3* promoter, which is not expressed in AWC, *hmbx-1* had a strong effect on *srsx-3* expression in both AWC neurons. Under these conditions, *srsx-3* was expressed in both AWCs in both wild-type and *hmbx-1(tm1274)* null backgrounds (Figure 4.6B). Expression of *hmbx-1* and *osm-3* overlaps in ASI and ASH, and this effect may therefore be related to an *hmbx-1* function in those neurons. The potential role of ASI is intriguing, since ASI is also the source of the TGFβ signal. A *daf-7::GFP* reporter was expressed normally in a *hmbx-1(ky777)* background (data not shown), so lack of DAF-7 is not responsible for the *srsx-3* defect of *hmbx-1(ky777)* mutants. However, it is possible to speculate that another diffusible signal may

be released from ASI and received by AWC. ASI also forms synapses onto AWC, and *hmbx-1* activity in ASI may promote signaling through this route. In either case, a single transcription factor appears to be capable of coordinating expression the *srsx-3* GPCR from multiple cells, and is capable of acting on AWC indirectly when expressed in neighboring neurons.

**Maintenance of core identity**

I focused much of my effort on examining expression of two GPCRs, *str-2* and *srsx-3*, and used expression of these markers to approximate the overall cell identities of $AWC^{ON}$ and $AWC^{OFF}$. However, the distinction between expression of an individual gene and regulation of overall cell identity is crucial to understanding the fundamental processes I wished to study. The discovery of additional GPCRs specific to each of the two cells allowed me to approach the question of cell identity more closely by looking at correlated changes of multiple markers within a cell, but this analysis is still limited. Suites of GPCRs may be regulated together in the absence of any other changes in the nature of the cell. Once major question, then, is whether altered GPCR expression is correlated with changes in neuronal function. For example, does loss of a set of subtype-specific GPCRs in adults correspond to loss of the animal's ability to respond to odors specifically sensed by that neuron? Several of the mutants described here, including *odr-1*, *daf-11*, and *nsy-7*, are defective for chemotaxis to these odors as adults, but we have not yet tested whether they are also defective as L1 larvae, when they still express *str-2* and *srsx-3* appropriately.

Another approach to this problem is to look for factors required for maintenance of non-GPCR genes specific to $AWC^{ON}$ or $AWC^{OFF}$. Experiments to

profile the differences in mRNA expression between wild-type and *nsy-7(tm3080)* mutant animals are underway, and will help to highlight differences in the transcriptional profiles of $AWC^{ON}$ and $AWC^{OFF}$ (Blencowe et al. 2009). Alternatively, some genes may be expressed in both AWCs, but have $AWC^{ON}$- or $AWC^{OFF}$-specific functions. For example, *gcy-28* alters the nature of the $AWC^{ON}$ but not $AWC^{OFF}$ response to odor; loss of *gcy-28* switches $AWC^{ON}$ from a mediator of attractive behaviors to a mediator of repulsion (Tsunozaki et al. 2008). Factors required to maintain the expression or function of *gcy-28* might therefore play a role in maintaining the nature and identity of the $AWC^{ON}$ neuron.

**Relationship between environmental inputs and changes in transcriptional state**

To further examine the relationship between maintenance of GPCR expression and neuronal function, it will be important to determine whether the environmental inputs that modulate GPCR expression can also alter the behaviors associated with corresponding cell types. For example, pheromone stimulation downregulates *srsx-3* expression in $AWC^{OFF}$; it may be interesting to see whether it also reduces chemotaxis efficiency to 2,3-pentanedione. If detection by ASK is responsible for the effects of pheromone on *srsx-3* expression in AWC, then activation of ASK might also be expected to downregulate *srsx-3* expression and possibly to reduce 2,3-pentanedione chemotaxis. Similarly, if odor inhibits the generation of cGMP from the ODR-1 and DAF-11 guanylate cyclases, then odor exposure may also reduce expression of *str-2*, *srsx-3*, and other cGMP-responsive receptors. Consistent with this possibility, exposure to

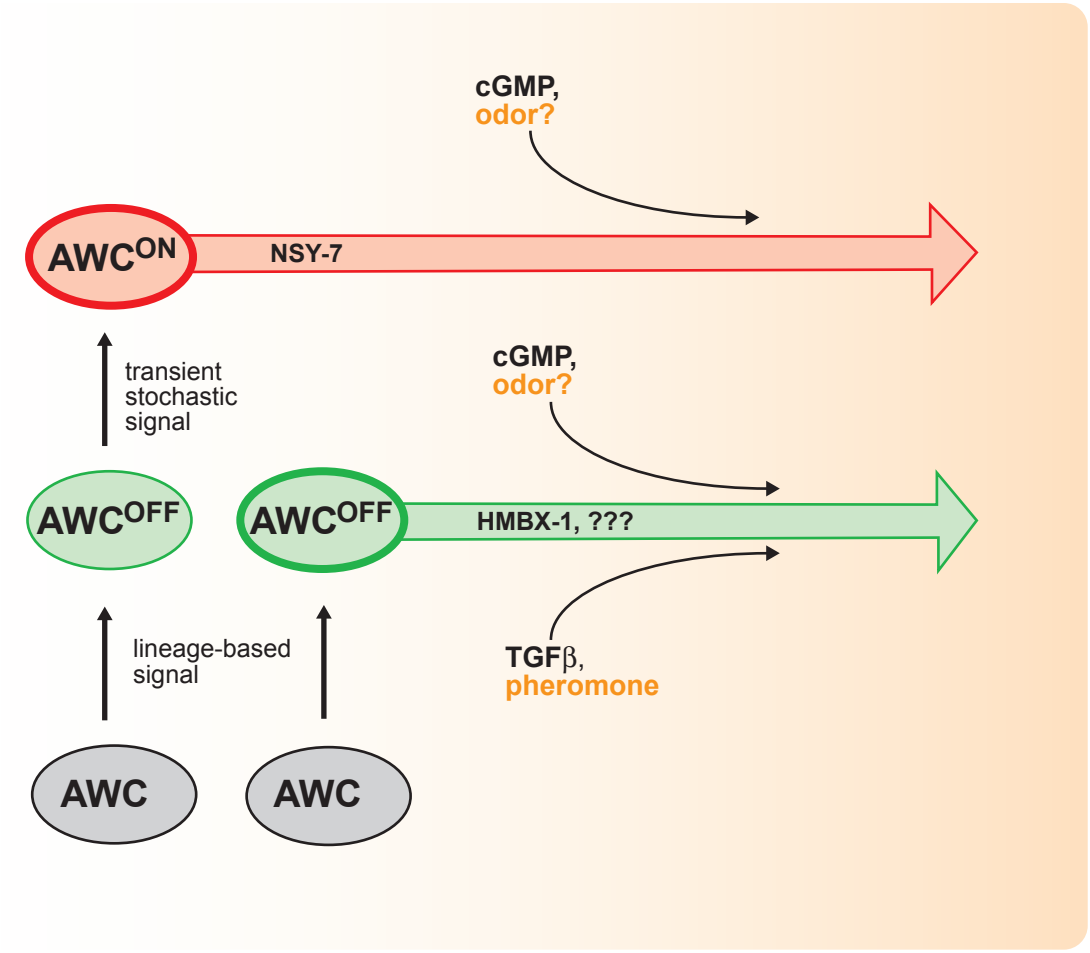



**Figure 5.1. Maintenance of the AWC$^{ON}$ and AWC$^{OFF}$ identities.** The AWC$^{ON}$ state is induced by a transient signal during embryogenesis and maintained by the transcription factor NSY-7. The AWC$^{OFF}$ (default) state is maintained by a combination of external and internal signals, including the transcription factor HMBX-1, a TGFβ signaling pathway that responds to dauer pheromone stimulation, and a cGMP-dependent pathway that responds to stimulation by odor.

high levels of odor can lead to adaptation and interfere with chemotaxis; however, exposure to low concentrations of odor in the presence of food can also enhance chemotaxis (L'Etoile et al. 2002; Torayama et al. 2007). It may be interesting to ask whether these established behavioral effects of odor pre-exposure correlate with differences in receptor expression, indicating that pre-exposure paradigms drive transcriptional changes in the cell.

**Summary**

This thesis has uncovered a variety of factors involved in maintenance of GPCR expression in $AWC^{ON}$ and $AWC^{OFF}$ (Figure 5.1). Some of the genes involved are transcriptional regulators, and appear to affect fundamental cell identity as well as GPCR expression. It is interesting that multiple transcription factors seem to be required to maintain asymmetric AWC cell fate. The involvement of multiple interlocking sets of transcriptional regulators might itself serve a stabilizing function, buffering a cell's overall transcriptional state against any single external influence and allowing more precise modulation of expression levels in response to stimuli. Indeed, the effects of a single transcription factor on target gene expression can be drastically altered by a few well-placed mutations, and therefore the presence of multiple transcriptional influences may help to prevent most such changes from exerting a major effect on the overall state of the neuron. The specific means by which these transcription factors stabilize cell identity remains to be determined, as does the nature of their interaction with activity-dependent influences such as cGMP and TGFβ. It is clear, however, that maintenance of the $AWC^{ON}$ and $AWC^{OFF}$ identities requires a system separate from that responsible for specifying the

initial fate decision, and the studies have provided some insight into the identity and function of the factors involved in these maintenance-specific mechanisms.

# Methods

### Genetics & Strains

*C. elegans* strains were cultured using standard methods (Brenner 1974). All strains were grown at 20°C unless otherwise specified.

### Genetic Mapping

*nsy-7(ky630)* was isolated from a screen for 2-AWC$^{OFF}$ mutants (Vanhoven et al. 2006). In three-factor mapping experiments, *ky630* mapped to chromosome IV between *unc-5(e53)* (IV:1.78) and *dpy-20(e1282ts)* (IV:5.22). Trans-heterozygotes of *ky630* with the deletions *stDf7*, *eDf19*, and *sDf2* each displayed a wild-type *str-2* expression phenotype, suggesting that the mutation fell in the interval IV:1.78-2.43 or IV:3.50-3.60. We confirmed that the mutation fell in the latter region by SNP mapping: the strain CX8147 *nsy-7(ky630) dpy-20(e1282ts) IV; kyIs140 I* was crossed to the polymorphic strain CB4856, and lines displaying either Nsy non Dpy or Dpy non Nsy phenotypes were evaluated at known polymorphic loci. The *nsy-7* mutation fell between the SNPs CE4-150 (IV:3.59) and C07G1:1158 (IV:3.65). Sequencing of the ORF C18F3.4 in this interval revealed a G→A mutation at position 30769 of cosmid C18F3.

*odr-1(ky779)*, *odr-1(ky800)*, and *odr-1(ky807)* were identified by their failure to complement *odr-1(n1936)*. Likewise, *nsy-7(ky813)* and *nsy-7(ky819)* were identified by failure to complement *nsy-7(tm3080)*. The nucleotide change for these alleles is unknown.

*daf-11(ky801)* does not confer the dauer-constitutive phenotype associated with most *daf-11* mutants. *ky801* was mapped using CB4856 SNPs (Davis et al.

2005) to a 2.8Mb interval between pkP5118 and pkP5068 at the center of chromosome V, and then identified by its failure to complement *daf-11(m47)* for the *srsx-3* expression phenotype. Sequencing revealed an A→T transversion in the coding sequence.

*ky791* was preliminarily mapped to chromosome III. *ky791* mutants were then crossed to the polymorphic strain CB4856 for SNP mapping. After selecting F2s that were homozygous for *ky791* based on defective *srsx-3* expression, it was observed that 0/100 recombinant lines displayed the bordering and clumping behavior associated with the CB4856 *npr-1* allele. Because *tax-4* is on chromosome III and is known to suppress the *npr-1* phenotype, the genomic locus was sequenced and a C→T mutation identified. Based on the similar *srsx-3* and *str-2* expression phenotype of the *ky782* mutation and its location on chromosome I, we then sequenced the *tax-2* locus in this mutant and identified a T→A transversion.

*ky879* was mapped to chromosome V, then crossed to a *unc-42(e270) sma-1(e30)* double mutant for three-point mapping and found lie to the right of *sma-1* (V:3.54). SNP mapping indicated that it lay to the left of a SNP at V:6 (Davis et al. 2005). This interval contained the *odr-3* locus; sequencing of this candidate gene revealed the corresponding G→A transition.

*tam-1(ky798)* was preliminarily mapped chromosome V, and three-point crosses to *unc-46(e177) dpy-11(e224)* and *unc-62(e644) dpy-11(e224)* indicated that it lay to the left of *unc-62* (V: -5.22) and to right of *unc-46* (V:-2.49). Partial rescue was observed with the cosmid C02E7, which covers the 3' end of *tam-1*, and

sequencing of the *tam-1* locus revealed a G→A nucleotide change in the coding region.

*ky771* and *ky803* were mapped using CB4856 SNPs to chromosomes III and IV, respectively, and candidate genes were identified based on the strong dauer constitutive phenotype of both alleles. Sequencing identified the *daf-7(ky771)* C→T and *daf-1(ky803)* G→A transitions.

*ky793* and *ky777* were mapped to intervals of <500kb using CB4856 SNPs, and mutations were identified using Solexa-Illumina whole-genome sequencing (see below).

**Whole-genome sequencing**

Whole-genome sequencing of the *ky777* and *ky793* mutants was done in the Rockefeller Genomics Resource Center using the Solexa-Illumina Genome Analyzer technology. For each of these mutants, SNP mapping was used to determine a genetic interval of <500kb for the mutation. Purified genomic DNA was prepared from 1000-2000 worms and used for library preparation. For *ky793*, fastq sequences were aligned to the WormBase reference sequence WS199 and single-base changes were predicted using the MAQ alignment program (http://maq.sourceforge.net). For *ky777*, fastq sequences were aligned to the ce6 genome assembly at UCSC 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). For both *ky793* and *ky777*,

high-probability single-base changes that were predicted to fall within an exon, 3'UTR, or 5'UTR were checked by PCR and conventional (Sanger) sequencing. The original *ky793* strain contained two linked coding mutations within the interval predicted by SNP mapping (now called *odr-1(ky992)* and *ceh-36(ky793)*), and recombinant strains containing each of these mutations separately were created in order to determine the *ky793* causative mutation.

**Sequence Alignment**

The *nsy-7* homeodomain sequence was aligned with engrailed homeodomain sequences using ClustalW at PBIL (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_clustalw.html). GenBank accession numbers are:

*H. sapiens* engrailed1: NM_001426; *M. musculus* engrailed1: NM_010133; *D. melanogaster* engrailed isoform A: NM_078976; *C. elegans* engrailed (*ceh-16*): NM_066497. The engrailed consensus sequence is from Gehring et al., 1994.

The *nsy-7* homeodomain sequence was aligned with methyl-binding domain sequences using MAFFT (BLOSUM62 matrix; opening gap penalty 1.53; extending gap penalty 0.00; strategy E-INS-i) at MyHits (http://myhits.isb-sib.ch/cgi-bin/mafft).

UniProt ID numbers are:

*D. melanogaster* MBD2A: CG8208; *H. sapiens* MBD1: Q9UIS9; *M. musculus* MBD1: Q9Z2E2; *H. sapiens* MBD2: Q9UBB5; *M. musculus* MBD2: Q9Z2E1; *H. sapiens* MBD3: O95983; *M. musculus* MBD3: Q9Z2D8.

**Molecular Biology**

Standard molecular biology techniques were used. The 21 kb *nsy-7* promoter was PCR amplified from wild-type (N2) lysate and cloned into the pSM-GFP expression vector between the NotI and AscI sites in MCS1 (primers: 5′-ttaagaacaacacacatgacctac-3′ and 5′-tctgaaaaaaagtttagatgtttga-3′). For rescue, the same promoter was inserted into MCS1 of the pSM expression vector, and the *nsy-7* genomic sequence was inserted into MCS2 between NheI and KpnI sites (primers: 5′-ttccgaatcaaacatctaaactt-3′ and 5′-cgaagaaaccaaacacttttttca-3′).

NSY-7 cDNAs were obtained by PCR from a *C. elegans* cDNA library using primers at the beginning and end of the predicted open reading frame (5′-atgtcttcggatacaaaatacaaa-3′ and 5′-tcaagccgtcgtctcttgca-3′). Sequencing of cDNA fragments made from a circularized *C. elegans* RNA library revealed that some NSY-7 sequences included an SL1 spliced leader sequence at the predicted 5′ end, confirming that the predicted 5′ exon is (at least in some cases) the true starting exon. Two classes of cDNAs for C18F3.4 were identified by RT-PCR analysis, differing in the presence or absence of 15 amino acids at the beginning of the third exon. Both cDNAs resulted in a 2-AWC$^{ON}$ phenotype when expressed in both AWC neurons under the *odr-3* promoter (Figure 2.2C and data not shown); only the longer isoform was examined in subsequent experiments.

The NSY-7::GFP fusion plasmid was constructed by insertion of the NSY-7 cDNA into pSM-GFP between the NheI and KpnI restriction sites in MCS2, in frame with the GFP coding sequence. The GFP::NSY-7 fusion plasmid was created by amplification of a GFP:NSY-7 fragment by PCR using overlap extension, with pSM-GFP and the NSY-7 cDNA as starting templates. This

fragment was cloned into MCS2 of pSM between the NheI and KpnI restriction sites.

To create the GST::NSY-7 bacterial expression vector for protein production, the full-length NSY-7 cDNA was cloned into pGEX 4T-3 using SalI and NotI sites. For the 6His:NSY-7 bacterial expression vector, the same NSY-7 cDNA was cloned into pSV271 (Pokala and Handel 2004) using NarI and XhoI sites. For expression in *C. elegans*, the GST:NSY-7 sequence was cloned into MCS2 of a modified pSM vector (containing a NotI site in MCS2 instead of MCS1) between the NheI and NotI sites, with a SalI site between the GST and NSY-7 sequences. 6His:NSY-7 was PCR-amplified from the bacterial expression vector to include NheI and KpnI sites, and inserted into MCS2 of pSM using these sites.

Mutation of the *srsx-3::GFP* plasmid (Bauer Huang et al. 2007), was carried out by site-directed mutagenesis using the QuikChange II Site-Directed Mutagenesis Kit (Stratagene, 200524).

The *gfp-pest* reporter was made by overlap extension PCR between GFP template from the pSM-GFP expression vector and 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.

The *hlh-11* promoter sequence was PCR-amplified from N2 lysate (primers: 5′-tcgtttctctcgtttcctcc-3′ and 5′-tctggattacctgaaactttacaa-3′) and cloned into MCS1 of pSM-GFP between the FseI and AscI sites. The *hlh-11* genomic sequence was amplified from N2 lysate and cloned into MCS2 between the NheI

and KpnI sites (primers: 5′-aagtttcaggtaatccagacg-3′ and 5′-taaatattggtactcttgtctcagt-3′).

A PCR fragment containing the *hmbx-1* genomic coding region and 7kb of upstream sequence was amplified from N2 lysate using the primers 5′-catatcacatcctcctgttttcaag-3′ and 5′-ccaacaaaatttcagggcaagc-3′. HMBX-1 cDNAs were obtained by PCR from a *C. elegans* cDNA library using primers at the beginning and end of the predicted open reading frame (5′-atgctattcacaattgagcaactg -3′ and 5′-tcactcagacttcaaaatagaag -3′) and 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.

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 the pSM-GFP expression vector using the FseI and AscI restriction sites.

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′-tgccgattgccaaaaacttcc-3′; A*, 5′-catatcacatcctcctgttttcaag-3′; B, 5′-agtcgacctgcaggcatgcaagctgtgtaaaataaaatttactacgaggtt-3′

*srt-45*: A, 5′-cttttaagtattgtcactccggc-3′; A*, 5′-ttggacatgtaatcatgcatg-3′; B, 5′-agtcgacctgcaggcatgcaagcttgcggaatctgaagtttttcg-3′

*srt-47*: A, 5′-gcgtaaaacacagcagaaaa-3′; A*, 5′-atctttaaaggtgcattttatttgg-3′; B, 5′- agtcgacctgcaggcatgcaagctttggtgataaagagaggttatag-3′

For spot analysis, the UAS vector was created by cloning six tandem repeats of the Gal4 UAS sequence into MCS1 of the pSM expression vector between the HindIII and NotI sites. The Gal4 DNA binding domain was cloned into the pSM-mCherry expression vector using the MCS2 NheI and KpnI restriction sites. A nuclear localization sequence was added at the 5′ end by cloning into the MCS2 EcoNI and NheI sites.

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

**RNAi**

RNAi was performed in one of two ways: by feeding and by injection of double-stranded RNAi. All assays were performed in an *eri-1(mg366;lin-15B(n744)* sensitized background (Sieburth et al. 2005).

For injection, double-stranded DNA template corresponding to an exon of the target ORF was amplified from N2 lysate with a T7 sequence (TAATACGACTCACTATAGGGAGA) added at the 5′ ends. The following gene-specific sequences were used in these primers:

*nsy-7* (exon 2): 5′-gttgcgaaaggatattcagatg-3′; 5′-cttagcaaacaagttggtgagt-3′

*odr-1*: 5′-ctttgactgaccccaccaat-3′; 5′-ggtttgtgccaaatgttctg-3′

*xnp-1*: 5′-agcctctggctgtgaagaaa-3′; 5′-gattcgaaccagcgaattgt-3′

*Y42H9AR.4*: 5′-ttagctcgtcgccacgtta-3′; 5′-tgcattctcctcgtcaagtg-3′

*nol-1*: 5′-agatcatctgccaatcgagaa-3′; 5′-gagcgagttggctctaatgg-3′

*Y53F4B.4*: 5′-cctaaagcacgctctgaagc-3′; 5′-ctcttgacccgttcaaaaaca-3′

*Y39G10AR.21*: 5′-aatccctcaacttcgaccaa-3′; 5′-ctcgctatcctctggcacat-3′

*Y48G8AL.5*: 5′-cgctgctacatgctaatcca-3′; 5′-cctttgcctcttgctcagtc-3′

*T09A5.8*: 5′-gttccaaaagtcccaacgaa-3′; 5′-cgttgggataactgaagatgc-3′

*C47A4.1*: 5′-tccatcaactgtccgaaaatc-3′; 5′-gcaaccctataagaaagagcaga-3′

*Y113G7A.16*: 5′-gaaatggatcggaatggaaa-3′; 5′-acgaaatagttcccgcgttt-3′

*R102.7*: 5′-aacacatgtcacctcaaacga-3′; 5′-ttttgagcaaaacctccattg-3′

*R10E4.3*: 5′-atagctggaggcacagttcg-3′; 5′-agcgcctcgcttaacaatta-3′

*sams-1*: 5′-gaccatgccattgactctca-3′; 5′-gcgtaggaaagttggacgag-3′

*sams-4*: 5′-agctgtaattcctgccaatctc-3′; 5′-tgataatctttcccggtctca-3′

*sams-5*: 5′-catgcaatgttcttgttgctc-3′; 5′-atctggcgaatgttgggtag-3′

*hmbx-1* (exon 3): 5′-acaagctcccgtaacacaac-3′; 5′-tcactcagacttcaaaatagaagcc-3′

Transcription was performed using the T7 RiboMAX Express RNAi System (Promega, P1700) according to instructions. 8ul of unpurified PCR reaction was used in a reaction volume of 20ul. The transcription reaction was diluted 4x in RNAse-free water and 2ul of each reaction was run on an agarose gel for quantification. The unpurified reaction mix was injected into the body cavity, gut, or gonad of adult hermaphrodites. Injected animals were transferred to fresh plates after 24 hours of recovery and transferred again after 48 hours; only F1s from the second and third sets of plates were scored. F1 progeny were scored after 3-4 days for *str-2* and *srsx-3* expression.

For feeding experiments, plates containing 1.5mM IPTG, 50 mg/ml carbenicillin, and 5 mg/ml tetracycline were poured and allowed to dry for at least 4 days at room temperature. A 1ml culture of the *E. coli* HT115 RNAi

feeding strain corresponding to the target gene was grown for 8-10h in LB with 25mg/ml carbenicillin, and plates were then seeded with 100 ml of this culture. Worms were synchronized at the L1 stage by bleaching adults, washing the eggs several times to remove bleach and OP50, and allowing the eggs to hatch in M9 buffer overnight. 50-100 L1 larvae were added to the RNAi plates, grown at 20°C, and scored as adults.

**MosSCI integrations**

Mos single-copy integrants were generated using the direct insertion protocol described in (Frokjaer-Jensen et al. 2008). 30-50 *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.

**NSY-7 target prediction**

Transcriptional targets of NSY-7 were predicted by searching the genome for genes with a CCTTAAC NSY-7 binding site located within 300bp upstream of their predicted transcriptional start sites. The list in Table 3.1 also includes *hcp-3*

and *hlh-11*, which are transcribed in opposite directions and share the same regulatory sequence. A NSY-7 binding site is located ~1.6 kb upstream of *hlh-11* and ~4kb upstream of *hcp-3*. Although this site is greater than 300bp from both genes, both *hlh-11* and *hcp-3* are expressed preferentially in $AWC^{OFF}$ and *hlh-11* is regulated by *nsy-7*. Both were therefore included in the list of possible NSY-7 targets.

**Protein Production and Purification**

GST:NSY-7 and 6His:NSY-7 were produced by expression and purification in *E. coli* strain BL21(DE3). Transformed cells were grown overnight in LB medium + 50ug/ml carbenicillin (for GST::NSY-7) or 50ug/ml kanamycin (for 6His::NSY-7). Overnight cultures were diluted 1:100 in fresh LB + antibiotic; diluted cultures were grown at 37°C to $OD_{600}$ ~0.5, induced with IPTG at a final concentration of 1mM, grown for an additional 4 hours, and pelleted by spinning at ~2500xg for 15 minutes at 4°C. Pellets were washed and resuspended in ice-cold PBS, and then lysed by sonication. For 6His:NSY-7, the pellet was incubated on ice for 30 minutes in lysis buffer (50mM $NaH_2PO_4$, 300mM NaCl, 10mM imidazole, pH 8.0) with 1mg/ml lysozyme before sonication. Lysates were centrifuged for 10-30 minutes at 4°C. The GST fusion protein was purified from lysate supernatant using a Microspin GST Purification Module (Amersham, 27-4570-03) according to instructions, and His-tagged protein was purified using an Ni-NTA Spin Kit (Qiagen, 31314) according to instructions for native conditions. Presence of intact protein in the cell pellet before lysis, in lysate supernatant, and after purification was verified by SDS-PAGE followed by staining with Coomassie Blue. Protein concentration was estimated by spectrophotometry of the purified sample.

**Protein Binding Microarrays**

Protein binding microarray experiments were performed as described previously (Berger et al. 2006; Berger and Bulyk 2009). Briefly, PBMs were performed using a custom-designed microarray from Agilent Technologies consisting of ~44,000 features of 60 base pairs, encompassing all possible combinations of 10-mers. Primer extension was performed on the single-stranded slides utilizing a common 24 base pair region to generate a double-stranded DNA microarray. GST-tagged NSY-7 was diluted to a final concentration of 500 nM in PBS, 2% (wt/vol) milk, 51.3 ng/ul salmon testes DNA (Sigma), 0.2 ug/ul BSA (New England Biolabs) and incubated for 1 h at 20ºC. After washing, arrays were incubated with an Alexa488-conjugated anti-GST antibody (Invitrogen) for 1 hr at 20º C. After subsequent washing, DNA-protein interaction was visualized using a Perkin Elmer ScanArray 5000 scanner to detect Alexa488 fluorescence. Arrays were normalized as previously described (Berger et al., 2006), and contiguous and gapped 8-mer enrichment scores as well as position weight matrices were generated using the "Seed-and-Wobble" algorithm (Berger et al., 2006).

**Electromobility Shift Assays**

*Annealing & Labeling.* Double-stranded oligonucleotides were produced by annealing single-stranded DNA fragments, synthesized by Integrated DNA Technologies (IDT). Sequences used were:
(Figure 2.5):

FL: 5′-ttttgttttccaaatcccttaacacaatactaaaaggaa-3′

FLrc: 5′-ttccttttagtattgtgttaagggatttggaaaacaaaa-3′

M1: 5′-ttttgttttccaaatcccttcacacaatactaaaaggaa-3′

M1rc: 5′-ttccttttagtattgtgtgaagggatttggaaaacaaaa-3′

M2: 5′-ttttgttttccaaatccagctggacaatactaaaaggaa-3′

M2rc: 5′-ttccttttagtattgtccagctggatttggaaaacaaaa-3′

(Figure 3.5): 'Mc' indicates methylated cytosine.

nometh_F: 5′-ctagcaacgacgtgccgaggctcgaccagcacgtgctag-3′

nometh_R: 5′-ctagcacgtgctggtcgagcctcggcacgtcgttgctag-3′

mCpT/A/C_F:

5′-ctagMcaacgacgtgMccgaggMctcgaMccagMcacgtgMctag-3′

mCpT/A/C_R:

5′-ctagcacgtgMctggtcgagMcctcggMcacgtcgttgctag-3′

mCpG_F:

5′-ctagcaaMcgaMcgtgccgaggctMcgaccagcaMcgtgctag-3′

mCpG_R:

5′-ctagcaMcgtgctggtMcgagcctcggcaMcgtMcgttgctag-3′

For labeling with γ-$^{32}$P-ATP, 100ng of double-stranded probe was mixed with 1ul of radioactive nucleotide (Perkin Elmer, BLU502A) and incubated for 1 hour at 37$^{o}$C with T4 polynucleotide kinase (New England Biolabs, M0201S). Excess nucleotide was removed using NucTrap Probe Purification Columns (Stratagene, 400701), according to instructions.

*EMSA.* For all experiments, 500ng of fusion protein was incubated with binding buffer (10mM Tris, 1mM EDTA, 100mM KCl, 0.1mM DTT, 5% glycerol), 5ug/ml salmon sperm, and 50ug/ml BSA for 10 minutes at room temperature.

0.25ng labeled probe or 0.25ng labeled probe with 10x, 25x, 50x, 100x, or 200x cold competitor was added, and the reaction mixture was incubated at room temperature for an additional 15 minutes. The entire 20ul binding reaction was then loaded onto a 6% NOVEX DNA Retardation gel (Invitrogen, EC6365BOX) and run for ~1 hour in 0.5x TBE buffer. Experiments were repeated three times.

**Microscopy and Fluorescence Quantification**

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

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 20x or 40x Plan-Neofluar objective, and photographs were taken under a 40x Plan-Neofluar or 63x Plan-Apochromat objective. For quantification, images were collected on the Zeiss Axioplan2 system under standardized detector settings using Metamorph software. A region within the cell body of the AWC neuron was selected and average fluorescence intensity was calculated for the same region in red and green channels. A region of similar size was selected for the contralateral cell of the same animal, and the same data was collected. Data shown represents average values for at least 25 animals of each genotype at each time point.

For visualization of NSY-7::GFP and Gal4::mCherry arrays in the nucleus, animals were examined and images obtained using a 40x/1.2W C-Apochromat water-immersion objective on a Zeiss LSM 510 confocal imaging system, using

the Zeiss LSM 510 v3.2 confocal software. Images were then deconvolved using Huygens deconvolution software and analyzed using ImageJ.

**Developmental Analysis**

To evaluate marker expression at precise stages during postembryonic development, larvae were staged by hatch-off. Late embryos were picked to an NGM plate seeded with the *E. coli* strain OP50. After 30 minutes, just-hatched L1s were transferred to a fresh plate and grown at 20ºC for the specified time. Individuals scored at any given time point were discarded; therefore, a separate set of staged worms was scored at each time point.

**Heat shock assays**

All heat shocks were performed for 2 hours at 33°C. Worms were transferred to 20°C for recovery. For developmental heat shocks, gravid adult animals were allowed to lay eggs at 20°C for one hour and then removed from the plate; progeny were heat shocked 0h, 14h, 20h, or 27h after removal of adults, then scored after reaching adulthood. For adult heat shocks, 30 L4 animals were picked to a seeded plate and heat shocks were performed after they had reached adulthood, ~12 hours later. Each plate contained a mixture of array-positive and array-negative animals. Each assay included a paired, no-heat-shock control. Adult animals were scored at on a Zeiss Axioplan2 or a Zeiss Axio Imager.Z1 imaging system under a 20x Plan-Neofluar objective.

**Behavioral Assays**

Chemotaxis assays were performed as described (Bargmann et al. 1993). Odors were diluted in ethanol and tested at standard concentrations (1:1000 butanone and 1:10,000 2,3-pentanedione). Three independent assays of each strain were conducted for each odor.

**Pheromone assays**

Ascarosides C3, C6, and C9 were added to liquid agar at the concentrations indicated. For negative controls, the same volume of solvent (ethanol) was added to the agar. 10 ml of agar was poured into 6-cm culture dishes and allowed to cool. Plates were then seeded with 100 ml OP50 bacteria and dried in a hood for 1.5-2 hours. These plates were either used immediately or stored overnight at 4ºC. 20 *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 at 25ºC for 4 hours. Animals were scored for presence or absence of GFP in AWC neurons under a 40x Plan-Neofluar objective on a Zeiss Axioplan2 imaging system.

# References


Affolter, M., Percival-Smith, A., Muller, M., Leupin, W., and Gehring, W.J. 1990. DNA binding properties of the purified Antennapedia homeodomain. *Proc Natl Acad Sci U S A* **87**(11): 4093-4097.

Ahringer, J. 2000. NuRD and SIN3 histone deacetylase complexes in development. *Trends Genet* **16**(8): 351-356.

Ashworth, R. and Bolsover, S.R. 2002. Spontaneous activity-independent intracellular calcium signals in the developing spinal cord of the zebrafish embryo. *Brain Res Dev Brain Res* **139**(2): 131-137.

Ballestar, E. and Wolffe, A.P. 2001. Methyl-CpG-binding proteins. Targeting specific gene repression. *Eur J Biochem* **268**(1): 1-6.

Bantignies, F., Grimaud, C., Lavrov, S., Gabut, M., and Cavalli, G. 2003. Inheritance of Polycomb-dependent chromosomal interactions in *Drosophila*. *Genes Dev* **17**(19): 2406-2420.

Bargmann, C.I., Hartwieg, E., and Horvitz, H.R. 1993. Odorant-selective genes and neurons mediate olfaction in *C. elegans*. *Cell* **74**(3): 515-527.

Bauer Huang, S.L., Saheki, Y., VanHoven, M.K., Torayama, I., Ishihara, T., Katsura, I., van der Linden, A., Sengupta, P., and Bargmann, C.I. 2007. Left-right olfactory asymmetry results from antagonistic functions of voltage-activated calcium channels and the Raw repeat protein OLRN-1 in *C. elegans*. *Neural Develop* **2**: 24.

Baugh, L.R. and Sternberg, P.W. 2006. DAF-16/FOXO regulates transcription of *cki-1*/Cip/Kip and repression of *lin-4* during *C. elegans* L1 arrest. *Curr Biol* **16**(8): 780-785.

Ben-Ari, Y., Cherubini, E., Corradetti, R., and Gaiarsa, J.L. 1989. Giant synaptic potentials in immature rat CA3 hippocampal neurones. *J Physiol* **416**: 303-325.

Benard, C. and Hobert, O. 2009. Looking beyond development: maintaining nervous system architecture. *Curr Top Dev Biol* **87**: 175-194.

Bender, L.B., Suh, J., Carroll, C.R., Fong, Y., Fingerman, I.M., Briggs, S.D., Cao, R., Zhang, Y., Reinke, V., and Strome, S. 2006. MES-4: an autosome-associated histone methyltransferase that participates in silencing the X chromosomes in the *C. elegans* germ line. *Development* **133**(19): 3907-3917.

Berger, M.F., Badis, G., Gehrke, A.R., Talukder, S., Philippakis, A.A., Pena-Castillo, L., Alleyne, T.M., Mnaimneh, S., Botvinnik, O.B., Chan, E.T., Khalid, F., Zhang,

W., Newburger, D., Jaeger, S.A., Morris, Q.D., Bulyk, M.L., and Hughes, T.R. 2008. Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. *Cell* **133**(7): 1266-1276.

Berger, M.F. and Bulyk, M.L. 2009. Universal protein-binding microarrays for the comprehensive characterization of the DNA-binding specificities of transcription factors. *Nat Protoc* **4**(3): 393-411.

Berger, M.F., Philippakis, A.A., Qureshi, A.M., He, F.S., Estep, P.W., 3rd, and Bulyk, M.L. 2006. Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. *Nat Biotechnol* **24**(11): 1429-1435.

Bestor, T.H. 2000. The DNA methyltransferases of mammals. *Hum Mol Genet* **9**(16): 2395-2402.

Bienvenu, T. and Chelly, J. 2006. Molecular genetics of Rett syndrome: when DNA methylation goes unrecognized. *Nat Rev Genet* **7**(6): 415-426.

Blencowe, B.J., Ahmad, S., and Lee, L.J. 2009. Current-generation high-throughput sequencing: deepening insights into mammalian transcriptomes. *Genes Dev* **23**(12): 1379-1386.

Bozza, T., Vassalli, A., Fuss, S., Zhang, J.J., Weiland, B., Pacifico, R., Feinstein, P., and Mombaerts, P. 2009. Mapping of class I and class II odorant receptors to glomerular domains by two distinct types of olfactory sensory neurons in the mouse. *Neuron* **61**(2): 220-233.

Brenner, S. 1974. The genetics of Caenorhabditis elegans. *Genetics* **77**(1): 71-94.

Buck, L. and Axel, R. 1991. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* **65**(1): 175-187.

Busturia, A. and Morata, G. 1988. Ectopic expression of homeotic genes caused by the elimination of the Polycomb gene in *Drosophila* imaginal epidermis. *Development* **104**(4): 713-720.

Butcher, R.A., Fujita, M., Schroeder, F.C., and Clardy, J. 2007. Small-molecule pheromones that control dauer development in *Caenorhabditis elegans*. *Nat Chem Biol* **3**(7): 420-422.

Butcher, R.A., Ragains, J.R., and Clardy, J. 2009. An indole-containing dauer pheromone component with unusual dauer inhibitory activity at higher concentrations. *Org Lett* **11**(14): 3100-3103.

Butcher, R.A., Ragains, J.R., Kim, E., and Clardy, J. 2008. A potent dauer pheromone component in *Caenorhabditis elegans* that acts synergistically with other components. *Proc Natl Acad Sci U S A* **105**(38): 14288-14292.

Campbell, G. and Tomlinson, A. 1999. Transducing the Dpp morphogen gradient in the wing of *Drosophila*: regulation of Dpp targets by brinker. *Cell* **96**(4): 553-562.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. 2002. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**(5595): 1039-1043.

Cao, X. and Jacobsen, S.E. 2002a. Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proc Natl Acad Sci U S A* **99 Suppl 4**: 16491-16498.

-. 2002b. Role of the arabidopsis DRM methyltransferases in de novo DNA methylation and gene silencing. *Curr Biol* **12**(13): 1138-1144.

Cardoso, C., Couillault, C., Mignon-Ravix, C., Millet, A., Ewbank, J.J., Fontes, M., and Pujol, N. 2005. XNP-1/ATR-X acts with RB, HP1 and the NuRD complex during larval development in *C. elegans*. *Dev Biol* **278**(1): 49-59.

Carey, M.B. and Matsumoto, S.G. 1999. Spontaneous calcium transients are required for neuronal differentiation of murine neural crest. *Dev Biol* **215**(2): 298-313.

Chahrour, M., Jung, S.Y., Shaw, C., Zhou, X., Wong, S.T., Qin, J., and Zoghbi, H.Y. 2008. MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science* **320**(5880): 1224-1229.

Chan, C.S., Rastelli, L., and Pirrotta, V. 1994. A Polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression. *Embo J* **13**(11): 2553-2564.

Chaney, B.A., Clark-Baldwin, K., Dave, V., Ma, J., and Rance, M. 2005. Solution structure of the K50 class homeodomain PITX2 bound to DNA and implications for mutations that cause Rieger syndrome. *Biochemistry* **44**(20): 7497-7511.

Chang, A.J., Chronis, N., Karow, D.S., Marletta, M.A., and Bargmann, C.I. 2006. A distributed chemosensory circuit for oxygen preference in *C. elegans*. *PLoS Biol* **4**(9): e274.

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

Charney, D.S., Nestler, E.J. 2004. *Neurobiology of Mental Illness*. Oxford University Press, Oxford.

Chen, R.Z., Akbarian, S., Tudor, M., and Jaenisch, R. 2001. Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. *Nat Genet* **27**(3): 327-331.

Chen, S., Saiyin, H., Zeng, X., Xi, J., Liu, X., Li, X., and Yu, L. 2006. Isolation and functional analysis of human HMBOX1, a homeobox containing protein with transcriptional repressor activity. *Cytogenet Genome Res* **114**(2): 131-136.

Chen, W.G., Chang, Q., Lin, Y., Meissner, A., West, A.E., Griffith, E.C., Jaenisch, R., and Greenberg, M.E. 2003. Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. *Science* **302**(5646): 885-889.

Chess, A., Simon, I., Cedar, H., and Axel, R. 1994. Allelic inactivation regulates olfactory receptor gene expression. *Cell* **78**(5): 823-834.

Chi, Y.I. 2005. Homeodomain revisited: a lesson from disease-causing mutations. *Hum Genet* **116**(6): 433-444.

Choksi, S.P., Southall, T.D., Bossing, T., Edoff, K., de Wit, E., Fischer, B.E., van Steensel, B., Micklem, G., and Brand, A.H. 2006. Prospero acts as a binary switch between self-renewal and differentiation in *Drosophila* neural stem cells. *Dev Cell* **11**(6): 775-789.

Chu, D.S., Dawes, H.E., Lieb, J.D., Chan, R.C., Kuo, A.F., and Meyer, B.J. 2002. A molecular link between gene-specific and chromosome-wide transcriptional repression. *Genes Dev* **16**(7): 796-805.

Chuang, C.F. and Bargmann, C.I. 2005. A Toll-interleukin 1 repeat protein at the synapse specifies asymmetric odorant receptor expression via ASK1 MAPKKK signaling. *Genes Dev* **19**(2): 270-281.

Chuang, C.F., Vanhoven, M.K., Fetter, R.D., Verselis, V.K., and Bargmann, C.I. 2007. An innexin-dependent cell network establishes left-right neuronal asymmetry in *C. elegans*. *Cell* **129**(4): 787-799.

Chubb, J.E., Bradshaw, N.J., Soares, D.C., Porteous, D.J., and Millar, J.K. 2008. The DISC locus in psychiatric illness. *Mol Psychiatry* **13**(1): 36-64.

Coburn, C.M. and Bargmann, C.I. 1996. A putative cyclic nucleotide-gated channel is required for sensory development and function in *C. elegans*. *Neuron* **17**(4): 695-706.

Corlew, R., Bosma, M.M., and Moody, W.J. 2004. Spontaneous, synchronous electrical activity in neonatal mouse cortical neurones. *J Physiol* **560**(Pt 2): 377-390.

Couto, A., Alenius, M., and Dickson, B.J. 2005. Molecular, anatomical, and functional organization of the *Drosophila* olfactory system. *Curr Biol* **15**(17): 1535-1547.

Cui, M., Kim, E.B., and Han, M. 2006. Diverse chromatin remodeling genes antagonize the Rb-involved SynMuv pathways in *C. elegans*. *PLoS Genet* **2**(5): e74.

da Graca, L.S., Zimmerman, K.K., Mitchell, M.C., Kozhan-Gorodetska, M., Sekiewicz, K., Morales, Y., and Patterson, G.I. 2004. DAF-5 is a Ski oncoprotein homolog that functions in a neuronal TGF beta pathway to regulate *C. elegans* dauer development. *Development* **131**(2): 435-446.

Daniels, S.A., Ailion, M., Thomas, J.H., and Sengupta, P. 2000. *egl-4* acts through a transforming growth factor-beta/SMAD pathway in *Caenorhabditis elegans* to regulate multiple neuronal circuits in response to sensory cues. *Genetics* **156**(1): 123-141.

Davis, M.W., Hammarlund, M., Harrach, T., Hullett, P., Olsen, S., and Jorgensen, E.M. 2005. Rapid single nucleotide polymorphism mapping in *C. elegans*. *BMC Genomics* **6**: 118.

Dawes, H.E., Berlin, D.S., Lapidus, D.M., Nusbaum, C., Davis, T.L., and Meyer, B.J. 1999. Dosage compensation proteins targeted to X chromosomes by a determinant of hermaphrodite fate. *Science* **284**(5421): 1800-1804.

Decembrini, S., Andreazzoli, M., Vignali, R., Barsacchi, G., and Cremisi, F. 2006. Timing the generation of distinct retinal cells by homeobox proteins. *PLoS Biol* **4**(9): e272.

Deisseroth, K., Bito, H., and Tsien, R.W. 1996. Signaling from synapse to nucleus: postsynaptic CREB phosphorylation during multiple forms of hippocampal synaptic plasticity. *Neuron* **16**(1): 89-101.

Dobritsa, A.A., van der Goes van Naters, W., Warr, C.G., Steinbrecht, R.A., and Carlson, J.R. 2003. Integrating the molecular and cellular basis of odor coding in the *Drosophila* antenna. *Neuron* **37**(5): 827-841.

Draganescu, A. and Tullius, T.D. 1998. The DNA binding specificity of engrailed homeodomain. *J Mol Biol* **276**(3): 529-536.

Dragich, J., Houwink-Manville, I., and Schanen, C. 2000. Rett syndrome: a surprising result of mutation in MECP2. *Hum Mol Genet* **9**(16): 2365-2375.

Dyer, M.A., Livesey, F.J., Cepko, C.L., and Oliver, G. 2003. Prox1 function controls progenitor cell proliferation and horizontal cell genesis in the mammalian retina. *Nat Genet* **34**(1): 53-58.

Eade, K.T. and Allan, D.W. 2009. Neuronal phenotype in the mature nervous system is maintained by persistent retrograde bone morphogenetic protein signaling. *J Neurosci* **29**(12): 3852-3864.

Elmore, T., Ignell, R., Carlson, J.R., and Smith, D.P. 2003. Targeted mutation of a *Drosophila* odor receptor defines receptor requirement in a novel class of sensillum. *J Neurosci* **23**(30): 9906-9912.

Estevez, M., Attisano, L., Wrana, J.L., Albert, P.S., Massague, J., and Riddle, D.L. 1993. The *daf-4* gene encodes a bone morphogenetic protein receptor controlling *C. elegans* dauer larva development. *Nature* **365**(6447): 644-649.

Etchberger, J.F., Flowers, E.B., Poole, R.J., Bashllari, E., and Hobert, O. 2009. Cis-regulatory mechanisms of left/right asymmetric neuron-subtype specification in *C. elegans*. *Development* **136**(1): 147-160.

Fan, G. and Hutnick, L. 2005. Methyl-CpG binding proteins in the nervous system. *Cell Res* **15**(4): 255-261.

Ficz, G., Heintzmann, R., and Arndt-Jovin, D.J. 2005. Polycomb group protein complexes exchange rapidly in living *Drosophila*. *Development* **132**(17): 3963-3976.

Fischer, A., Sananbenesi, F., Wang, X., Dobbin, M., and Tsai, L.H. 2007. Recovery of learning and memory is associated with chromatin remodelling. *Nature* **447**(7141): 178-182.

Fisher, A.L. and Lithgow, G.J. 2006. The nuclear hormone receptor DAF-12 has opposing effects on *Caenorhabditis elegans* lifespan and regulates genes repressed in multiple long-lived worms. *Aging Cell* **5**(2): 127-138.

Fishilevich, E. and Vosshall, L.B. 2005. Genetic and functional subdivision of the *Drosophila* antennal lobe. *Curr Biol* **15**(17): 1548-1553.

Flavell, S.W., Cowan, C.W., Kim, T.K., Greer, P.L., Lin, Y., Paradis, S., Griffith, E.C., Hu, L.S., Chen, C., and Greenberg, M.E. 2006. Activity-dependent regulation of MEF2 transcription factors suppresses excitatory synapse number. *Science* **311**(5763): 1008-1012.

Fraenkel, E., Rould, M.A., Chambers, K.A., and Pabo, C.O. 1998. Engrailed homeodomain-DNA complex at 2.2 A resolution: a detailed view of the interface and comparison with other engrailed structures. *J Mol Biol* **284**(2): 351-361.

Frand, A.R., Russel, S., and Ruvkun, G. 2005. Functional genomic analysis of *C. elegans* molting. *PLoS Biol* **3**(10): e312.

Friedkin, M. and Wood, H.I. 1956. Utilization of thymidine-C14 by bone marrow cells and isolated thymus nuclei. *J Biol Chem* **220**(2): 639-651.

Frokjaer-Jensen, C., Davis, M.W., Hopkins, C.E., Newman, B.J., Thummel, J.M., Olesen, S.P., Grunnet, M., and Jorgensen, E.M. 2008. Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nat Genet* **40**(11): 1375-1383.

Fujiwara, M., Sengupta, P., and McIntire, S.L. 2002. Regulation of body size and behavioral state of *C. elegans* by sensory perception and the EGL-4 cGMP-dependent protein kinase. *Neuron* **36**(6): 1091-1102.

Fuks, F., Hurd, P.J., Wolf, D., Nan, X., Bird, A.P., and Kouzarides, T. 2003. The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *J Biol Chem* **278**(6): 4035-4040.

Fuss, S.H. and Ray, A. 2009. Mechanisms of odorant receptor gene choice in *Drosophila* and vertebrates. *Mol Cell Neurosci* **41**(2): 101-112.

Garaschuk, O., Linn, J., Eilers, J., and Konnerth, A. 2000. Large-scale oscillatory calcium waves in the immature cortex. *Nat Neurosci* **3**(5): 452-459.

Gaudet, J. and Mango, S.E. 2002. Regulation of organogenesis by the *Caenorhabditis elegan*s FoxA protein PHA-4. *Science* **295**(5556): 821-825.

Gehring, W.J., Affolter, M., and Burglin, T. 1994. Homeodomain proteins. *Annu Rev Biochem* **63**: 487-526.

Georgi, L.L., Albert, P.S., and Riddle, D.L. 1990. *daf-1*, a *C. elegans* gene controlling dauer larva development, encodes a novel receptor protein kinase. *Cell* **61**(4): 635-645.

Goldman, A.L., Van der Goes van Naters, W., Lessing, D., Warr, C.G., and Carlson, J.R. 2005. Coexpression of two functional odor receptors in one neuron. *Neuron* **45**(5): 661-666.

Gowher, H., Leismann, O., and Jeltsch, A. 2000. DNA of *Drosophila melanogaster* contains 5-methylcytosine. *Embo J* **19**(24): 6918-6923.

Greer, P.L. and Greenberg, M.E. 2008. From synapse to nucleus: calcium-dependent gene transcription in the control of synapse development and function. *Neuron* **59**(6): 846-860.

Grove, C.A., De Masi, F., Barrasa, M.I., Newburger, D.E., Alkema, M.J., Bulyk, M.L., and Walhout, A.J. 2009. A multiparameter network reveals extensive divergence between *C. elegans* bHLH transcription factors. *Cell* **138**(2): 314-327.

Gu, X., Olson, E.C., and Spitzer, N.C. 1994. Spontaneous neuronal calcium spikes and waves during early differentiation. *J Neurosci* **14**(11 Pt 1): 6325-6335.

Gutierrez, A. and Sommer, R.J. 2004. Evolution of *dnmt-2* and *mbd-2*-like genes in the free-living nematodes *Pristionchus pacificus*, *Caenorhabditis elegans* and *Caenorhabditis briggsae*. *Nucleic Acids Res* **32**(21): 6388-6396.

Guy, J., Gan, J., Selfridge, J., Cobb, S., and Bird, A. 2007. Reversal of neurological defects in a mouse model of Rett syndrome. *Science* **315**(5815): 1143-1147.

Hallem, E.A., Ho, M.G., and Carlson, J.R. 2004. The molecular basis of odor coding in the *Drosophila* antenna. *Cell* **117**(7): 965-979.

Hendrich, B. and Tweedie, S. 2003. The methyl-CpG binding domain and the evolving role of DNA methylation in animals. *Trends Genet* **19**(5): 269-277.

Hirota, J., Omura, M., and Mombaerts, P. 2007. Differential impact of Lhx2 deficiency on expression of class I and class II odorant receptor genes in mouse. *Mol Cell Neurosci* **34**(4): 679-688.

Hobert, O. 2002. PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic *C. elegans*. *Biotechniques* **32**(4): 728-730.

Holliday, R. and Pugh, J.E. 1975. DNA modification mechanisms and gene activity during development. *Science* **187**(4173): 226-232.

Hoppe, R., Breer, H., and Strotmann, J. 2006. Promoter motifs of olfactory receptor genes expressed in distinct topographic patterns. *Genomics* **87**(6): 711-723.

Hoppe, R., Frank, H., Breer, H., and Strotmann, J. 2003. The clustered olfactory receptor gene family 262: genomic organization, promotor elements, and interacting transcription factors. *Genome Res* **13**(12): 2674-2685.

Hsieh, J., Liu, J., Kostas, S.A., Chang, C., Sternberg, P.W., and 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**(22): 2958-2970.

Hu, Q., Kwon, Y.S., Nunez, E., Cardamone, M.D., Hutt, K.R., Ohgi, K.A., Garcia-Bassets, I., Rose, D.W., Glass, C.K., Rosenfeld, M.G., and Fu, X.D. 2008. Enhancing nuclear receptor-induced transcription requires nuclear motor and

LSD1-dependent gene networking in interchromatin granules. *Proc Natl Acad Sci U S A* **105**(49): 19199-19204.

Ideraabdullah, F.Y., Vigneau, S., and Bartolomei, M.S. 2008. Genomic imprinting mechanisms in mammals. *Mutat Res* **647**(1-2): 77-85.

Inoue, T. and Thomas, J.H. 2000a. Suppressors of transforming growth factor-beta pathway mutants in the *Caenorhabditis elegans* dauer formation pathway. *Genetics* **156**(3): 1035-1046.

-. 2000b. Targets of TGF-beta signaling in *Caenorhabditis elegans* dauer formation. *Dev Biol* **217**(1): 192-204.

Jazwinska, A., Kirov, N., Wieschaus, E., Roth, S., and Rushlow, C. 1999. The *Drosophila* gene brinker reveals a novel mechanism of Dpp target gene regulation. *Cell* **96**(4): 563-573.

Johnston, R.J., Jr., Chang, S., Etchberger, J.F., Ortiz, C.O., and Hobert, O. 2005. MicroRNAs acting in a double-negative feedback loop to control a neuronal cell fate decision. *Proc Natl Acad Sci U S A* **102**(35): 12449-12454.

Jovelin, R. 2009. Rapid sequence evolution of transcription factors controlling neuron differentiation in *Caenorhabditis*. *Mol Biol Evol*.

Kennison, J.A. and Tamkun, J.W. 1988. Dosage-dependent modifiers of polycomb and antennapedia mutations in *Drosophila*. *Proc Natl Acad Sci U S A* **85**(21): 8136-8140.

Keshavan, M., Kennedy, J., Murray, R. 2004. *Neurodevelopment and Schizophrenia*. Cambridge University Press, Cambridge.

Kiefer, J.C., Smith, P.A., and Mango, S.E. 2007. PHA-4/FoxA cooperates with TAM-1/TRIM to regulate cell fate restriction in the *C. elegans* foregut. *Dev Biol* **303**(2): 611-624.

Kishi, N. and Macklis, J.D. 2004. MECP2 is progressively expressed in post-migratory neurons and is involved in neuronal maturation rather than cell fate decisions. *Mol Cell Neurosci* **27**(3): 306-321.

Klebes, A., Sustar, A., Kechris, K., Li, H., Schubiger, G., and Kornberg, T.B. 2005. Regulation of cellular plasticity in *Drosophila* imaginal disc cells by the Polycomb group, trithorax group and lama genes. *Development* **132**(16): 3753-3765.

Klose, R.J., Sarraf, S.A., Schmiedeberg, L., McDermott, S.M., Stancheva, I., and Bird, A.P. 2005. DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. *Mol Cell* **19**(5): 667-678.

Koga, M. and Ohshima, Y. 2004. The *C. elegans ceh-36* gene encodes a putative homemodomain transcription factor involved in chemosensory functions of ASE and AWC neurons. *J Mol Biol* **336**(3): 579-587.

Komatsu, H., Mori, I., Rhee, J.S., Akaike, N., and Ohshima, Y. 1996. Mutations in a cyclic nucleotide-gated channel lead to abnormal thermosensation and chemosensation in *C. elegans*. *Neuron* **17**(4): 707-718.

L'Etoile, N.D. and Bargmann, C.I. 2000. Olfaction and odor discrimination are mediated by the *C. elegans* guanylyl cyclase ODR-1. *Neuron* **25**(3): 575-586.

L'Etoile, N.D., Coburn, C.M., Eastham, J., Kistler, A., Gallegos, G., and Bargmann, C.I. 2002. The cyclic GMP-dependent protein kinase EGL-4 regulates olfactory adaptation in *C. elegans*. *Neuron* **36**(6): 1079-1089.

Lanjuin, A. and Sengupta, P. 2002. Regulation of chemosensory receptor expression and sensory signaling by the KIN-29 Ser/Thr kinase. *Neuron* **33**(3): 369-381.

Lanjuin, A., VanHoven, M.K., Bargmann, C.I., Thompson, J.K., and Sengupta, P. 2003. Otx/otd homeobox genes specify distinct sensory neuron identities in *C. elegans*. *Dev Cell* **5**(4): 621-633.

Lans, H. and 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**(3): 1287-1299.

Lans, H., Rademakers, S., and Jansen, G. 2004. A network of stimulatory and inhibitory Galpha-subunits regulates olfaction in *Caenorhabditis elegans*. *Genetics* **167**(4): 1677-1687.

Larsson, M.C., Domingos, A.I., Jones, W.D., Chiappe, M.E., Amrein, H., and Vosshall, L.B. 2004. Or83b encodes a broadly expressed odorant receptor essential for *Drosophila* olfaction. *Neuron* **43**(5): 703-714.

Lesch, B.J., Gehrke, A.R., Bulyk, M.L., and Bargmann, C.I. 2009. Transcriptional regulation and stabilization of left-right neuronal identity in *C. elegans*. *Genes Dev* **23**(3): 345-358.

Lewcock, J.W. and Reed, R.R. 2004. A feedback mechanism regulates monoallelic odorant receptor expression. *Proc Natl Acad Sci U S A* **101**(4): 1069-1074.

Lewis, E.B. 1978. A gene complex controlling segmentation in *Drosophila*. *Nature* **276**(5688): 565-570.

Lin, Y., Bloodgood, B.L., Hauser, J.L., Lapan, A.D., Koon, A.C., Kim, T.K., Hu, L.S., Malik, A.N., and Greenberg, M.E. 2008. Activity-dependent regulation of inhibitory synapse development by Npas4. *Nature* **455**(7217): 1198-1204.

Liu, T., Zimmerman, K.K., and Patterson, G.I. 2004. Regulation of signaling genes by TGFbeta during entry into dauer diapause in *C. elegans*. *BMC Dev Biol* **4**: 11.

Lyko, F., Ramsahoye, B.H., and Jaenisch, R. 2000. DNA methylation in *Drosophila melanogaster*. *Nature* **408**(6812): 538-540.

Majdan, M. and Shatz, C.J. 2006. Effects of visual experience on activity-dependent gene regulation in cortex. *Nat Neurosci* **9**(5): 650-659.

Marhold, J., Rothe, N., Pauli, A., Mund, C., Kuehle, K., Brueckner, B., and Lyko, F. 2004. Conservation of DNA methylation in dipteran insects. *Insect Mol Biol* **13**(2): 117-123.

Martinowich, K., Hattori, D., Wu, H., Fouse, S., He, F., Hu, Y., Fan, G., and Sun, Y.E. 2003. DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science* **302**(5646): 890-893.

Maurange, C., Cheng, L., and Gould, A.P. 2008. Temporal transcription factors and their targets schedule the end of neural proliferation in *Drosophila*. *Cell* **133**(5): 891-902.

Maves, L. and Schubiger, G. 1998. A molecular basis for transdetermination in *Drosophila* imaginal discs: interactions between wingless and decapentaplegic signaling. *Development* **125**(1): 115-124.

McDonel, P., Costello, I., and Hendrich, B. 2009. Keeping things quiet: roles of NuRD and Sin3 co-repressor complexes during mammalian development. *Int J Biochem Cell Biol* **41**(1): 108-116.

McIntyre, J.C., Bose, S.C., Stromberg, A.J., and McClintock, T.S. 2008. Emx2 stimulates odorant receptor gene expression. *Chem Senses* **33**(9): 825-837.

Michaloski, J.S., Galante, P.A., and Malnic, B. 2006. Identification of potential regulatory motifs in odorant receptor genes by analysis of promoter sequences. *Genome Res* **16**(9): 1091-1098.

Mikkola, I., Heavey, B., Horcher, M., and Busslinger, M. 2002. Reversion of B cell commitment upon loss of Pax5 expression. *Science* **297**(5578): 110-113.

Miller, F.D. and Kaplan, D.R. 2007. To die or not to die: neurons and p63. *Cell Cycle* **6**(3): 312-317.

Miyamichi, K., Serizawa, S., Kimura, H.M., and Sakano, H. 2005. Continuous and overlapping expression domains of odorant receptor genes in the olfactory epithelium determine the dorsal/ventral positioning of glomeruli in the olfactory bulb. *J Neurosci* **25**(14): 3586-3592.

Mukhopadhyay, S., Lu, Y., Shaham, S., and Sengupta, P. 2008. Sensory signaling-dependent remodeling of olfactory cilia architecture in *C. elegans*. *Dev Cell* **14**(5): 762-774.

Muller, J. and Kassis, J.A. 2006. Polycomb response elements and targeting of Polycomb group proteins in *Drosophila*. *Curr Opin Genet Dev* **16**(5): 476-484.

Muotri, A.R. and Gage, F.H. 2006. Generation of neuronal variability and complexity. *Nature* **441**(7097): 1087-1093.

Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, A. 1998. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**(6683): 386-389.

Neuhaus, E.M., Gisselmann, G., Zhang, W., Dooley, R., Stortkuhl, K., and Hatt, H. 2005. Odorant receptor heterodimerization in the olfactory system of *Drosophila melanogaster*. *Nat Neurosci* **8**(1): 15-17.

Newburger, D.E. and Bulyk, M.L. 2008. UniPROBE: an online database of protein binding microarray data on protein-DNA interactions. *Nucleic Acids Res*.

Nolan, K.M., Sarafi-Reinach, T.R., Horne, J.G., Saffer, A.M., and Sengupta, P. 2002. The DAF-7 TGF-beta signaling pathway regulates chemosensory receptor gene expression in *C. elegans*. *Genes Dev* **16**(23): 3061-3073.

Nutt, S.L., Heavey, B., Rolink, A.G., and Busslinger, M. 1999. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* **401**(6753): 556-562.

Ohki, I., Shimotake, N., Fujita, N., Jee, J., Ikegami, T., Nakao, M., and Shirakawa, M. 2001. Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA. *Cell* **105**(4): 487-497.

Okkema, P.G. and Krause, M. 2005. Transcriptional regulation. *WormBook*: 1-40.

Owens, D.F. and Kriegstein, A.R. 1998. Patterns of intracellular calcium fluctuation in precursor cells of the neocortical ventricular zone. *J Neurosci* **18**(14): 5374-5388.

Patterson, G.I., Koweek, A., Wong, A., Liu, Y., and Ruvkun, G. 1997. The DAF-3 Smad protein antagonizes TGF-beta-related receptor signaling in the *Caenorhabditis elegans* dauer pathway. *Genes Dev* **11**(20): 2679-2690.

Peckol, E.L., Troemel, E.R., and Bargmann, C.I. 2001. Sensory experience and sensory activity regulate chemosensory receptor gene expression in *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A* **98**(20): 11032-11038.

Peel, A.D., Chipman, A.D., and Akam, M. 2005. Arthropod segmentation: beyond the *Drosophila* paradigm. *Nat Rev Genet* **6**(12): 905-916.

Pierce-Shimomura, J.T., Faumont, S., Gaston, M.R., Pearson, B.J., and Lockery, S.R. 2001. The homeobox gene *lim-6* is required for distinct chemosensory representations in *C. elegans*. *Nature* **410**(6829): 694-698.

Pirrotta, V., Chan, C.S., McCabe, D., and Qian, S. 1995. Distinct parasegmental and imaginal enhancers and the establishment of the expression pattern of the Ubx gene. *Genetics* **141**(4): 1439-1450.

Pokala, N. and Handel, T.M. 2004. Energy functions for protein design I: efficient and accurate continuum electrostatics and solvation. *Protein Sci* **13**(4): 925-936.

Ponger, L. and Li, W.H. 2005. Evolutionary diversification of DNA methyltransferases in eukaryotic genomes. *Mol Biol Evol* **22**(4): 1119-1128.

Poole, R.J. and Hobert, O. 2006. Early embryonic programming of neuronal left/right asymmetry in *C. elegans*. *Curr Biol* **16**(23): 2279-2292.

Ray, A., van der Goes van Naters, W., and Carlson, J.R. 2008. A regulatory code for neuron-specific odor receptor expression. *PLoS Biol* **6**(5): e125.

Ray, A., van Naters, W.G., Shiraiwa, T., and Carlson, J.R. 2007. Mechanisms of odor receptor gene choice in *Drosophila*. *Neuron* **53**(3): 353-369.

Redmond, L., Kashani, A.H., and Ghosh, A. 2002. Calcium regulation of dendritic growth via CaM kinase IV and CREB-mediated transcription. *Neuron* **34**(6): 999-1010.

Ren, P., Lim, C.S., Johnsen, R., Albert, P.S., Pilgrim, D., and Riddle, D.L. 1996. Control of *C. elegans* larval development by neuronal expression of a TGF-beta homolog. *Science* **274**(5291): 1389-1391.

Rens-Domiano, S. and Hamm, H.E. 1995. Structural and functional relationships of heterotrimeric G-proteins. *Faseb J* **9**(11): 1059-1066.

Ressler, K.J., Sullivan, S.L., and Buck, L.B. 1993. A zonal organization of odorant receptor gene expression in the olfactory epithelium. *Cell* **73**(3): 597-609.

Ringrose, L. and Paro, R. 2004. Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu Rev Genet* **38**: 413-443.

Roayaie, K., Crump, J.G., Sagasti, A., and Bargmann, C.I. 1998. The G alpha protein ODR-3 mediates olfactory and nociceptive function and controls cilium morphogenesis in *C. elegans* olfactory neurons. *Neuron* **20**(1): 55-67.

Robertson, K.D. 2005. DNA methylation and human disease. *Nat Rev Genet* **6**(8): 597-610.

Rolink, A.G., Nutt, S.L., Melchers, F., and Busslinger, M. 1999. Long-term in vivo reconstitution of T-cell development by Pax5-deficient B-cell progenitors. *Nature* **401**(6753): 603-606.

Roloff, T.C., Ropers, H.H., and Nuber, U.A. 2003. Comparative study of methyl-CpG-binding domain proteins. *BMC Genomics* **4**(1): 1.

Ross, J.M. and Zarkower, D. 2003. Polycomb group regulation of Hox gene expression in *C. elegans*. *Dev Cell* **4**(6): 891-901.

Rushlow, C., Colosimo, P.F., Lin, M.C., Xu, M., and Kirov, N. 2001. Transcriptional regulation of the *Drosophila* gene zen by competing Smad and Brinker inputs. *Genes Dev* **15**(3): 340-351.

Sagasti, A., Hisamoto, N., Hyodo, J., Tanaka-Hino, M., Matsumoto, K., and Bargmann, C.I. 2001. The CaMKII UNC-43 activates the MAPKKK NSY-1 to execute a lateral signaling decision required for asymmetric olfactory neuron fates. *Cell* **105**(2): 221-232.

Saller, E. and Bienz, M. 2001. Direct competition between Brinker and *Drosophila* Mad in Dpp target gene transcription. *EMBO Rep* **2**(4): 298-305.

Sato, K., Simon, M.D., Levin, A.M., Shokat, K.M., and Weiss, G.A. 2004. Dissecting the Engrailed homeodomain-DNA interaction by phage-displayed shotgun scanning. *Chem Biol* **11**(7): 1017-1023.

Schackwitz, W.S., Inoue, T., and Thomas, J.H. 1996. Chemosensory neurons function in parallel to mediate a pheromone response in *C. elegans*. *Neuron* **17**(4): 719-728.

Schier, A.F. and Gehring, W.J. 1993. Functional specificity of the homeodomain protein fushi tarazu: the role of DNA-binding specificity in vivo. *Proc Natl Acad Sci U S A* **90**(4): 1450-1454.

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., and Cavalli, G. 2007. Genome regulation by polycomb and trithorax proteins. *Cell* **128**(4): 735-745.

Schwappacher, R., Weiske, J., Heining, E., Ezerski, V., Marom, B., Henis, Y.I., Huber, O., and Knaus, P. 2009. Novel crosstalk to BMP signalling: cGMP-dependent kinase I modulates BMP receptor and Smad activity. *Embo J* **28**(11): 1537-1550.

Schwartz, Y.B. and Pirrotta, V. 2007. Polycomb silencing mechanisms and the management of genomic programmes. *Nat Rev Genet* **8**(1): 9-22.

Serizawa, S., Ishii, T., Nakatani, H., Tsuboi, A., Nagawa, F., Asano, M., Sudo, K., Sakagami, J., Sakano, H., Ijiri, T., Matsuda, Y., Suzuki, M., Yamamori, T., and Iwakura, Y. 2000. Mutually exclusive expression of odorant receptor transgenes. *Nat Neurosci* **3**(7): 687-693.

Serizawa, S., Miyamichi, K., Nakatani, H., Suzuki, M., Saito, M., Yoshihara, Y., and Sakano, H. 2003. Negative feedback regulation ensures the one receptor-one olfactory neuron rule in mouse. *Science* **302**(5653): 2088-2094.

Sheng, M., Thompson, M.A., and Greenberg, M.E. 1991. CREB: a Ca(2+)-regulated transcription factor phosphorylated by calmodulin-dependent kinases. *Science* **252**(5011): 1427-1430.

Shirasaki, R. and Pfaff, S.L. 2002. Transcriptional codes and the control of neuronal identity. *Annu Rev Neurosci* **25**: 251-281.

Shykind, B.M., Rohani, S.C., O'Donnell, S., Nemes, A., Mendelsohn, M., Sun, Y., Axel, R., and Barnea, G. 2004. Gene switching and the stability of odorant receptor gene choice. *Cell* **117**(6): 801-815.

Sieburth, D., Ch'ng, Q., Dybbs, M., Tavazoie, M., Kennedy, S., Wang, D., Dupuy, D., Rual, J.F., Hill, D.E., Vidal, M., Ruvkun, G., and Kaplan, J.M. 2005. Systematic analysis of genes required for synapse structure and function. *Nature* **436**(7050): 510-517.

Silva, J.A.C.e. 2008. Autism, a brain developmental disorder: some new pathopysiologic and genetics findings. *Metabolism* **57 supplement 2**(October 2008): S40-S43.

Spitzer, N.C. 2006. Electrical activity in early neuronal development. *Nature* **444**(7120): 707-712.

Sprecher, S.G. and Desplan, C. 2008. Switch of rhodopsin expression in terminally differentiated *Drosophila* sensory neurons. *Nature* **454**(7203): 533-537.

Stellwagen, D. and Shatz, C.J. 2002. An instructive role for retinal waves in the development of retinogeniculate connectivity. *Neuron* **33**(3): 357-367.

Stringham, E.G., Dixon, D.K., Jones, D., and Candido, E.P. 1992. Temporal and spatial expression patterns of the small heat shock (*hsp16*) genes in transgenic *Caenorhabditis elegans*. *Mol Biol Cell* **3**(2): 221-233.

Sulston, J.E. and Horvitz, H.R. 1977. Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev Biol* **56**(1): 110-156.

Suzuki, H., Thiele, T.R., Faumont, S., Ezcurra, M., Lockery, S.R., and Schafer, W.R. 2008. Functional asymmetry in *Caenorhabditis elegans* taste neurons and its computational role in chemotaxis. *Nature* **454**(7200): 114-117.

Tabish, M., Siddiqui, Z.K., Nishikawa, K., and Siddiqui, S.S. 1995. Exclusive expression of *C. elegans osm-3* kinesin gene in chemosensory neurons open to the external environment. *J Mol Biol* **247**(3): 377-389.

Tanaka-Hino, M., Sagasti, A., Hisamoto, N., Kawasaki, M., Nakano, S., Ninomiya-Tsuji, J., Bargmann, C.I., and Matsumoto, K. 2002. SEK-1 MAPKK mediates Ca2+ signaling to determine neuronal asymmetric development in *Caenorhabditis elegans*. *EMBO Rep* **3**(1): 56-62.

Tichy, A.L., Ray, A., and Carlson, J.R. 2008. A new *Drosophila* POU gene, pdm3, acts in odor receptor expression and axon targeting of olfactory neurons. *J Neurosci* **28**(28): 7121-7129.

Torayama, I., Ishihara, T., and Katsura, I. 2007. *Caenorhabditis elegans* integrates the signals of butanone and food to enhance chemotaxis to butanone. *J Neurosci* **27**(4): 741-750.

Torborg, C.L., Hansen, K.A., and Feller, M.B. 2005. High frequency, synchronized bursting drives eye-specific segregation of retinogeniculate projections. *Nat Neurosci* **8**(1): 72-78.

Troemel, E.R., Chou, J.H., Dwyer, N.D., Colbert, H.A., and Bargmann, C.I. 1995. Divergent seven transmembrane receptors are candidate chemosensory receptors in *C. elegans*. *Cell* **83**(2): 207-218.

Troemel, E.R., Kimmel, B.E., and Bargmann, C.I. 1997. Reprogramming chemotaxis responses: sensory neurons define olfactory preferences in *C. elegans*. *Cell* **91**(2): 161-169.

Troemel, E.R., Sagasti, A., and Bargmann, C.I. 1999. Lateral signaling mediated by axon contact and calcium entry regulates asymmetric odorant receptor expression in *C. elegans*. *Cell* **99**(4): 387-398.

Tsunozaki, M., Chalasani, S.H., and Bargmann, C.I. 2008. A behavioral switch: cGMP and PKC signaling in olfactory neurons reverses odor preference in *C. elegans*. *Neuron* **59**(6): 959-971.

Unhavaithaya, Y., Shin, T.H., Miliaras, N., Lee, J., Oyama, T., and Mello, C.C. 2002. MEP-1 and a homolog of the NURD complex component Mi-2 act together to maintain germline-soma distinctions in *C. elegans*. *Cell* **111**(7): 991-1002.

van der Linden, A., Wiener, S., You, Y.J., Kim, K., Avery, L., and 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*.

van der Linden, A.M., Nolan, K.M., and Sengupta, P. 2007. KIN-29 SIK regulates chemoreceptor gene expression via an MEF2 transcription factor and a class II HDAC. *Embo J* **26**(2): 358-370.

Vanhoven, M.K., Bauer Huang, S.L., Albin, S.D., and Bargmann, C.I. 2006. The claudin superfamily protein *nsy-4* biases lateral signaling to generate left-right asymmetry in *C. elegans* olfactory neurons. *Neuron* **51**(3): 291-302.

Vassar, R., Ngai, J., and Axel, R. 1993. Spatial segregation of odorant receptor expression in the mammalian olfactory epithelium. *Cell* **74**(2): 309-318.

Vosshall, L.B., Wong, A.M., and Axel, R. 2000. An olfactory sensory map in the fly brain. *Cell* **102**(2): 147-159.

Wes, P.D. and Bargmann, C.I. 2001. *C. elegans* odour discrimination requires asymmetric diversity in olfactory neurons. *Nature* **410**(6829): 698-701.

Whetstine, J.R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M., and Shi, Y. 2006. Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* **125**(3): 467-481.

White, J.G., Southgate, E., Thomson, J. N. and Brenner, S. . 1986. The Structure of the Nervous System of the Nematode *Caenorhabditis Elegans*. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **314**(1165): 1-340.

Wigle, J.T., Chowdhury, K., Gruss, P., and Oliver, G. 1999. Prox1 function is crucial for mouse lens-fibre elongation. *Nat Genet* **21**(3): 318-322.

Wigle, J.T. and Eisenstat, D.D. 2008. Homeobox genes in vertebrate forebrain development and disease. *Clin Genet* **73**(3): 212-226.

Wu, M., Allis, C.D., Sweet, M.T., Cook, R.G., Thatcher, T.H., and Gorovsky, M.A. 1994. Four distinct and unusual linker proteins in a mitotically dividing nucleus are derived from a 71-kilodalton polyprotein, lack p34cdc2 sites, and contain protein kinase A sites. *Mol Cell Biol* **14**(1): 10-20.

Xu, L., Fong, Y., and Strome, S. 2001. The *Caenorhabditis elegans* maternal-effect sterile proteins, MES-2, MES-3, and MES-6, are associated in a complex in embryos. *Proc Natl Acad Sci U S A* **98**(9): 5061-5066.

Xue, D., Finney, M., Ruvkun, G., and Chalfie, M. 1992. Regulation of the *mec-3* gene by the *C.elegans* homeoproteins UNC-86 and MEC-3. *Embo J* **11**(13): 4969-4979.

Yu, S., Avery, L., Baude, E., and Garbers, D.L. 1997. Guanylyl cyclase expression in specific sensory neurons: a new family of chemosensory receptors. *Proc Natl Acad Sci U S A* **94**(7): 3384-3387.

Zhang, H., Azevedo, R.B., Lints, R., Doyle, C., Teng, Y., Haber, D., and Emmons, S.W. 2003. Global regulation of Hox gene expression in *C. elegans* by a SAM domain protein. *Dev Cell* **4**(6): 903-915.

Zhang, S.J., Zou, M., Lu, L., Lau, D., Ditzel, D.A., Delucinge-Vivier, C., Aso, Y., Descombes, P., and Bading, H. 2009. Nuclear calcium signaling controls expression of a large gene pool: identification of a gene program for acquired neuroprotection induced by synaptic activity. *PLoS Genet* **5**(8): e1000604.

Zhang, T., Sun, Y., Tian, E., Deng, H., Zhang, Y., Luo, X., Cai, Q., Wang, H., Chai, J., and Zhang, H. 2006. RNA-binding proteins SOP-2 and SOR-1 form a novel PcG-like complex in *C. elegans*. *Development* **133**(6): 1023-1033.

Zhu, S., Lin, S., Kao, C.F., Awasaki, T., Chiang, A.S., and Lee, T. 2006. Gradients of the *Drosophila* Chinmo BTB-zinc finger protein govern neuronal temporal identity. *Cell* **127**(2): 409-422.

Zilberman, D. and Henikoff, S. 2007. Genome-wide analysis of DNA methylation patterns. *Development* **134**(22): 3959-3965.