HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript

Howard Hughes Medical Institute

Published as: *Wiley Interdiscip Rev Dev Biol*. 2016 July ; 5(4): 474–498.

A map of terminal regulators of neuronal identity in *C. elegans*

Oliver Hobert

Department of Biological Sciences, Department of Biochemistry and Molecular Biophysics, Columbia University, Howard Hughes Medical Institute, New York, USA

Abstract

Our present day understanding of nervous system development is an amalgam of insights gained from studying different aspects and stages of nervous system development in a variety of invertebrate and vertebrate model systems, with each model system making its own distinctive set of contributions. One aspect of nervous system development that has been among the most extensively studied in the nematode *C. elegans* is the nature of the gene regulatory programs that specify hardwired, terminal cellular identities. I will first summarize a number of maps (anatomical, functional, molecular) that describe the terminal identity of individual neurons in the *C. elegans* nervous system. I then provide a comprehensive summary of regulatory factors that specify terminal identities in the nervous system, synthesizing these past studies into a regulatory map of cellular identities in the *C. elegans* nervous system. This map shows that for three quarters of all neurons in the *C. elegans* nervous system, regulatory factors that control terminal identity features are known. In-depth studies of specific neuron types have revealed that regulatory factors rarely act alone, but rather act cooperatively in neuron-type specific combinations. In most cases examined so far, distinct, biochemically-unlinked terminal identity features are co-regulated via cooperatively acting transcription factors, termed terminal selectors, but there are also cases in which distinct identity features are controlled in a piecemeal fashion by independent regulatory inputs. The regulatory map also illustrates that identity-defining transcription factors are re-employed in distinct combinations in different neuron types. However, the same transcription factor can drive terminal differentiation in neurons that are unrelated by lineage, unrelated by function, connectivity and neurotransmitter deployment. Lastly, the regulatory map illustrates the preponderance of homeodomain transcription factors in the control of terminal identities, suggesting that these factors have ancient, phylogenetically conserved roles in controlling terminal neuronal differentiation in the nervous system.

INTRODUCTION

Sydney Brenner's explicit vision when he established *C. elegans* as a model system was to first draw a series of maps – anatomical, lineage and functional maps – that describe this simple metazoan organism and to then use these maps as a starting point for microbial-style genetic analysis [1,2]. His explicit focus was on the nervous system and his stated goal was to define the genetic programs that instruct the development and function of this most complex of all tissue types [1,2]. And, indeed, we have come a long way. Based on Brenner's initial

contact: or38@columbia.edu.

HHMI Author Manuscript HHMI Author Manuscript HHMI Author Manuscript

efforts, no other nervous system is as extensively characterized on so many different levels – anatomically, functionally and developmentally - as that of *C. elegans*[3].

The nervous system of *C. elegans* is composed of 302 neurons in the hermaphrodite and 385 in the male. The position and shape of each neuron has been precisely described and the connectivity of neurons has been elucidated, both in the hermaphrodite and male (Fig. 1A–C, Fig. 2)[4,5]. In addition to these anatomical maps, easily accessible at www.wormatlas.org [3], there is a comprehensive lineage map that describes the developmental history of every single cell of the animal (Fig. 1D)[6,7]. Both the anatomical and lineage maps have been used with tremendous success to identify mutants in which nervous system anatomy or lineage is disrupted. Among the major findings of initial genetic screens were the identification of molecular cues involved in axon pathfinding (e.g. the *unc-6*/Netrin system)[8,9], heterochronic timer systems involved in neuroblast patterning (involving the first miRNAs to be discovered, *lin-4* and *let-7*[10,11]) and factors involved in synaptic specificity (*unc-4* homeobox gene)[12].

Another map established over the years can be termed a *functional map* of the nervous system. This map – which is still incomplete but nevertheless remarkably advanced – is based mostly on assessing the behavioral consequences of killing individual neurons through laser ablation. Such microsurgical approaches have revealed functions for almost all sensory neurons, most motor neurons and more than half of all interneurons (Table 1). For example, laser ablation of sensory neurons with specific microtubular structures identified their roles as touch receptor neurons [13], ablation of amphid wing neurons revealed their role in olfaction [14], or ablation of the AIY interneurons revealed their function in thermosensory information processing [15]. One of the fascinating themes that emerged over the years is the multifunctionality of many neurons (which is not comprehensively displayed in Table 1). For example, many sensory neurons are polymodal, detecting a wide range of distinct sensory modalities [16–18].

One of many values of these functional maps became apparent when mutant animals were identified from genetic screens that precisely phenocopy the loss of specific neurons. This led to the identification of, for example, sensory receptors that act in functionally defined sensory neurons (e.g. touch receptor proteins [19]; olfactory receptors [20]) or transcription factors that were found to be required for the development of specific neuron types (e.g. the *mec-3* transcription factor, whose loss results in mechanosensory defects similar to those observed upon ablation of touch sensory neurons [21]; the *che-1* transcription factor whose loss results in chemotaxis defects similar to those observed upon laser ablation of the ASE gustatory neurons [22] or the *ttx-1* and *ttx-3* transcription factors whose losses results in thermotaxis defects similar to those observed upon ablation of the AFD or AIY interneurons [23,24]).

With the advent of a number of molecular tools, yet another type of map has gained successively more prominence in the past two decades, a molecular map of the nervous system (Fig. 2). In this review, I shall describe molecular maps of the *C. elegans* nervous system and summarize how important they have been to understand the acquisition of

HHMI Author Manuscript HHMI Author Manuscript HHMI Author Manuscript

terminal neuronal identity features, allowing us to now establish a conglomerate *regulatory map* of the nervous system.

MOLECULAR MAPS OF THE *C. ELEGANS* NERVOUS SYSTEM

Ever since reporter gene technology was introduced into *C. elegans*, several thousand reporter lines have been generated to monitor the expression of specific genes (examples are shown in Fig. 2A). In many cases, reporter genes were generated for genes with suspected functions in the nervous system, such as sensory receptors [25,26], neuropeptides [27,28], ion channels [29], neurotransmitter receptors [30] or putative neuronal adhesion/recognition molecules [31]. As expected, most of these reporter constructs showed neuron-type specific expression profiles. In addition, the reporter-based expression analysis of the many genes characterized for various phenotypes within or outside the nervous system, yielded a tremendous number of additional reagents to monitor neuron-type specific gene expression. Since expression patterns are commonly analyzed in the context of a mature nervous system (larval or early adult stages), in which the identity of neurons can be most readily defined, there has been an inadvertent bias toward the characterization of expression patterns in postmitotic, mature neurons. Even though the precise sites of cellular expression of many reporter constructs (particularly those resulting from large scale expression screens conducted in Vancouver [32]) have not been determined, the site of expression of 868 reporter genes within the nervous system have been analyzed with single neuron resolution by the *C. elegans* community, creating more than 260,000 data points (868×302). These data points can be extracted from Wormbase and are displayed in Fig. 2B. On average, each of these 868 reporters is expressed in 10 neurons (with a range of 1 to 110 neurons; very few reporter genes are expressed exclusively in single neuron classes and those that are encode mostly sensory receptor proteins). And, on average, any given neuron type is associated with the expression of 28 distinct reporter genes (ranging from at least 3 to as many as 133). The extent to which each neuron type can be associated with molecular markers is unprecedented in other models, including flies and mice. While gene expression atlases have been generated in these models [33,34], they usually suffer from a lack of cellular resolution owing to the much more complex cellular anatomy of the fly and mouse nervous systems.

It is important to point out that these molecular maps describe *regulatory endpoints*, that is, terminal effector genes that are continuously expressed throughout the life of the neuron and that most often code for proteins that define the mature, functional and anatomical properties of a postmitotic neuron type (e.g. ion channels, sensory receptors, neurotransmitter receptors and neuropeptides). Equally important, owing to their unsystematic generation by many different labs for many different purposes, these transgenic markers provide a relatively unbiased snapshot of functionally unrelated identity features of a mature neuron.

AN (ALMOST COMPREHENSIVE) NEUROTRANSMITTER MAP

One specific subtype of molecular brain maps is a neurotransmitter map. Neurotransmitter maps are invaluable complements to any anatomical map (specifically, a connectome) as they inform us about the nature of information flow in the nervous system. Like any other organism, the *C. elegans* nervous system utilizes a plethora of distinct neurotransmitter

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript

systems, from fast-acting neurotransmitters (GABA, glutamate, acetylcholine) to aminergic neurotransmitters (dopamine, serotonin, octopamine, tyramine) to more than 200 modulatory neuropeptides [35]. Given the diversity of neurotransmitter systems, the neurotransmitter identity of a neuron is a key molecular identity feature of a neuron. While the identity of GABAergic and aminergic neurons has been known for a while (which amount to ~ 10% of all neuron classes [36,37]), it was only the recent analysis of glutamatergic and cholinergic neurons that brought the coverage of neurotransmitter identities to about 90% of all neurons in the nervous system. Cholinergic and glutamatergic neurotransmitter identities were inferred through the systematic expression analysis of vesicular transporters for acetylcholine (ACh) and glutamate (Glu), demonstrating that these two transmitters are the most broadly used neurotransmitters in the *C. elegans* nervous system (Fig. 2C; ACh: 52 of all 118 neuron classes; Glu 38 of all 118 neuron classes)[38,39]. Notable features of the *C. elegans* neurotransmitter map include the broad usage of ACh and Glu in different neuron types of the nervous system (each utilized by sensory, inter- and motorneurons)[38,39] and its association with specific circuitry (e.g. the ventral nerve cord motor neuron circuit, composed of motor neurons and innervating command interneurons is cholinergic [38]). Notably, from a developmental standpoint, neurons that share the same neurotransmitter show no apparent lineage relationships (Fig. 3).

A REGULATORY MAP OF THE NERVOUS SYSTEM

Much like the anatomical, lineage and functional maps of the worm nervous system, available molecular maps have served as powerful starting points for incisive genetic loss of function analysis. Specifically, reporter genes that are expressed in particular neuron types have offered the opportunity to engage in a bottom-up analysis which seeks to define the regulatory factors that control these terminal neuronal identity features. This type of analysis permitted a significant expansion of the phenotypic analysis of some classical transcription factor mutants. For example, genetic elimination of the LIM homeobox gene *mec-3* phenocopies the loss of mechanosensory neuron function [21], or loss of the LIM homeobox gene *ttx-3* phenocopies the loss of AIY neuron function (in both cases, animals show a characteristic cryophilic thermotaxis phenotype)[23], loss of the zinc (Zn) finger transcription factor *che-1* phenocopies the chemotaxis defects of ASE sensory neuron ablation [22]. However, it is *a priori* not clear whether such phenocopying is the result of the respective transcription factor merely regulating one or a few function-defining genes (e.g. a chemoreceptor protein and/or mechanoreceptor channel) or whether these factors have much more profound effects on the identity of the respective neuron. The ability to phenotype the loss-of-function for a regulatory factor of course very much depends on the availability of descriptors of neuronal identity features. The existence of a relatively fine-grained molecular phenotypic space of individual neuron type (in the form of marker transgenes) has traditionally set *C. elegans* apart from other systems in which the precise impact of a regulatory pathway on the identity of a neuron often cannot be assessed with the same level of precision.

In principle, the identification of genes controlling terminal identity features for specific neurons can be expected to address the following questions: (1) are the multiple molecular features that define a specific neuron type independently regulated by distinct regulatory

HHMI Author Manuscript HHMI Author Manuscript HHMI Author Manuscript

factors? Or are they co-regulated? (2) are there recurrent themes in which the identity of distinct neuron types are controlled? (3) are neurons that share specific functional features, e.g. sensory neurons; or synaptically connected neurons, controlled by the same regulatory factors?

These questions have been extensively addressed over the past decade, either by (1) directly screening for mutants in which specific molecular markers show defective expression; (2) by examination of reporter gene expression in candidate mutants (i.e. loss-of-function mutations in transcription factors expressed in a specific neuron type); or (3) by the analysis of mutants retrieved from behavioral screens (the above-mentioned classic behavioral mutants *che-1, ttx-3, mec-3, unc-3* etc.). The results of dozens of reports ordered by neuron class (sensory, inter and motorneurons) are summarized in Fig. 4, Fig. 5 and Fig. 6. In sum, transcriptional regulators have been identified that control the differentiation of 76 of the 104 *C. elegans* extrapharyngeal neuron classes (73%) which amounts to 217 of all the 282 extrapharyngeal neurons (77%). Each of these transcription factors is a terminal regulator since their expression persists throughout the life of the neuron and in the few cases explicitly examined, they have been found to also be continuously required to maintain the differentiated state (earlier, transiently acting factors are not considered here) [40].

Coregulation versus piecemeal regulation

Coming back to the question posed above - what do the transcription factors that control terminal identity features of a neuron exactly do? As discussed already above, a substantial phenotypic space can be probed for most individual neuron types – from anatomy, to connectivity, to molecular markers. It is particularly informative to consider the large panel of molecular markers, i.e. the molecular maps described above (and illustrated in Fig. 2), which describe a host of biochemically linked features (e.g. enzymes that act in a pathway to synthesize a neurotransmitter), but more importantly, also biochemically unlinked molecular features of a neuron (as illustrated schematically in Fig. 7A). The unlinked nature of many of these molecular markers makes it *a priori* entirely plausible to hypothesize that distinct biochemical features are organized into distinct regulons, i.e. are controlled via distinct sets of transcription factors. Notably, however, this does not appear to be a predominant theme. On the contrary, the in-depth analysis of the differentiation programs from a number of isolated neuron types – the light touch sensory neurons, GABAergic neurons, cholinergic AIY interneurons, glutamatergic ASE sensory neurons or cholinergic ventral nerve cord motor neurons – reveal that scores of molecular markers that are coexpressed by any of these specific neuron types are *co-regulated* by transcription factors expressed in these neurons (examples shown in Fig. 7)[41–45]. In each of these cases, the binding sites for the respective transcription factors are known and in each of these cases, the regulation of the terminal identity markers is thought to be largely direct (Fig. 7). As expected from the profound molecular defects of these neurons, the neurons are not functional and display various anatomical defects, including axonal outgrowth defects and synaptic connectivity defects. Yet in all cases examined, the neurons are still generated and even still express pan-neuronal features. In other words, the adoption of pan-neuronal identity can be genetically separated from the adoption of neuron type-specific identity. The above mentioned transcription factors have been called “Terminal Selectors” due to the profound impact they have on

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript

neuronal identity. The salient features of terminal selectors are summarized in Box 1 (recently reviewed in [46]).

Box 1

Terminal selectors of neuronal identity

Terminal selectors are transcription factors that induce expression of the terminally differentiated properties of mature neuron types [46]. These factors work mostly in combinations, e.g. as heterodimers that cooperatively bind DNA [41,57] to directly control the expression of effector genes that define the functional properties of a neuron type. Terminal selectors are continuously expressed throughout the life of a neuron and are required to not only induce but also maintain the differentiated state of a neuron, a feat achieved mostly through autoregulation (Box 2) [40]. Terminal selectors are not only required to induce specific differentiation programs, but are also sufficient to do so, albeit only in some cellular contexts [41,49,96]. This context dependency is likely dictated by the need of proper cofactors [49], but also by a chromatin environment that may be refractory to terminal selector activity[96].

Box 2

Regulatory network motifs

Three types of network motifs[51] employed in the context of neuronal specification and terminal selector function in *C. elegans*. The autoregulatory feedback loop is employed to ensure maintained terminal selector expression and, hence, maintained terminal effector gene expression[40,97]. Feedforward loops involve intervening transcription factors. Coherent feedforward loops have been found in other systems to filter transient input signals[51], but even though frequently observed[48,49], their function in terminal selector-mediated gene expression control remains to be shown. Incoherent feedforward loops occur in the context of *C. elegans* motor neuron class specification or lateralization of gustatory neuron function[45,66], as discussed in the text.

The vast majority of terminal selector targets are, by definition, terminal effector genes, i.e. genes encoding for proteins involved in defining the neuron-type specific structural and functional features of a neuron (Fig. 7). This sets terminal selectors apart from earlier acting transcription factors whose primary role is to trigger ensuing waves of regulatory factors (these factors are not covered here; see another review that considers these earlier acting factors [47]). Nevertheless, some terminal selector targets encode for regulatory factors that control specific *subroutines*, likely in conjunction with the upstream terminal selector. For example, the *ttx-3/ceh-10* terminal selector complex directly activates the expression of the homeobox gene *ceh-23* in the AIY interneurons and these factors together control expression of the G-protein-coupled receptor *sra-11* [48]. This regulatory architecture is a classic example of a feedforward regulatory motif and is found in other neuron types as well [45,49–51].

Since the in-depth analysis of a number of select cases mentioned above (Fig. 7), co-regulation has been examined in many different neuron types, albeit most often in less depth

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript

(summarized in Fig. 4, Fig. 5, Fig. 6). Much of this analysis was done in the context of mapping glutamatergic and cholinergic neurotransmitter identity throughout the nervous system [38,39]. With this neurotransmitter map in hand, a host of regulatory factors expressed in either glutamatergic or cholinergic neurons were analyzed for how their genetic removal affects the acquisition of the respective neurotransmitter identity. In those cases where an effect was observed, it was tested whether other, unlinked biochemical identity feature are also affected. In most cases, it was found that a factor that regulates neurotransmitter identity also affects the expression of other identity features. In a number of cases, the substantial number of markers tested (>20; Fig. 4,5,6) make a compelling case that factors that regulate neurotransmitter identity regulate large sets of genes expressed in the terminally differentiated neuron, i.e. that many distinct identity features of a terminally differentiated neuron are co-regulated. Even in cases where the sample size of examined marker is relatively small, it seems reasonable to extrapolate the effect to the many other genes expressed in the terminally differentiated neuron.

Even though the concept of co-regulation appears to apply broadly, the cholinergic command interneurons, a group of interconnected neurons that innervate ventral nerve cord motor neurons present a notable exception. In these neurons, the *unc-3* transcription factor, which co-regulates many, independent terminal identity features (including neurotransmitter identity) in ventral nerve cord motor neurons, appears to only control the acquisition of cholinergic neurotransmitter identity for all command interneurons [38]. More than a dozen additional markers expressed in select subsets of command interneurons are not affected in *unc-3* mutants [52]. The additional identity markers (a good number of them encoding for ionotropic Glu receptors) are in turn controlled in a cell-type specific, piecemeal manner by the *unc-42* homeobox gene, the *fax-1* nuclear hormone receptor, or the *cfi-1* ARID-type transcription factor [53–56]. Hence, it appears that in some cases, neuronal identity is controlled in a piecemeal manner through parallel acting identity regulators. However, at this point we cannot exclude that the above mentioned factors (*unc-3, unc-42, fax-1*) are mere subroutine regulators that operate in the context of aforementioned feedforward motifs, downstream of as yet unknown terminal selectors for command interneuron identity.

Other features of the regulatory map

Two other themes of the regulatory map are striking. First, transcription factors are used over and over again in distinct neuron types. Remarkably, four transcription factors are required to specify the identity of 50 distinct neuron classes, almost half of all neuron types in the *C. elegans* nervous system (*unc-3:* 16 neuron classes; *unc-86:* 14 neuron classes; *ceh-14*: 10 neuron classes; *unc-42:* 10 neuron classes). Note that these four transcription factors are deeply conserved (*unc-3* = EBF/Collier; *unc-86* = Brn3; *ceh-14* = Lhx3/4, *unc-42* = Prop1). In each neuron class these factors appear to act in combination with other factors and this combinatorial code appears to provide specificity. This combinatorial coding scheme is displayed in Fig. 8. For example, the POU homeobox gene *unc-86* acts as a regulator of terminal identity in 13 distinct neuron classes and it appears to operate in conjunction with specific cofactors in distinct neuron classes, such as the LIM homeobox gene *mec-3* in touch sensory neurons [57], the *ceh-14* LIM homeobox gene in a number of distinct interneurons [39], the *pag-3* Zn finger transcription factors in one specific interneuron type [49], or the *ttx-3* LIM

HHMI Author Manuscript HHMI Author Manuscript HHMI Author Manuscript

homeobox gene in a neurosecretory sensory neuron [58]. When reused, the same transcription factor may not necessarily perform as a terminal selector: it can work as a terminal selector in one neuron type or as a sub-routine regulator in another. For example, *ceh-14* operates as a terminal selector in 10 different neuron types [39,59,60], but operates only as a sub-routine regulator (controlling neuropeptide expression) in the BDU neurons [49].

Another theme is the striking preponderance of employment of homeodomain transcription factors. While these transcription factors make up only ~10% of all transcription factors in the *C. elegans* genome, a vast number of identified terminal regulators are homeodomain transcription factors and those that are not most frequently collaborate with a homeodomain transcription factor (homeodomain transcription factors are underlined in Fig. 4, Fig. 5, Fig. 6). Re-usage of the same transcription factors, as well as the preponderance of homeodomain transcription factors may indicate the evolutionary history of cell types in the nervous system. Homeodomain proteins may have had very early roles in specifying ancestral neuron types.

Are there other obvious, underlying themes or logical patterns by which transcription factors are used and re-used? If one considers the lineage history of each individual neuron type, no obvious patterns emerge. First, if one just considers molecular features, it is evident that the neurotransmitter identity and other molecular features of a neuron do not obviously correlate with lineage (Fig. 3). For example, dopaminergic neurons, which are molecularly and functionally highly related, show little lineage relationship. Second, transcription factor expression profiles also show a lack of correlation with lineage. This is illustrated by mapping the expression pattern of four transcription factors that control the terminal identity of at least 50 distinct neuron types on the lineage diagram (Fig. 3). The color-coding in this illustration also shows that neurotransmitter identity generally does not correlate with transcription factor usage; i.e., the same transcription factor can operate in neurons of different neurotransmitter identities. The only exception is the *unc-3* transcription factor, which represents a remarkable case of a transcription factors not only being committed to defining only cholinergic neuron identity, but doing so in the context of a specific circuit, the VNC motor circuit [38].

DIVERSIFICATION BY REPRESSORS

As summarized in Fig. 4,5,6 not every factor that controls terminal neuron differentiation is an inducer, i.e. a likely activator of some, many, or all terminal features of a neuron. In fact, one of the earliest identified regulatory factors in *C. elegans* is the *unc-4* homeobox gene [12]. *unc-4* mutants were initially characterized as synaptic specificity mutants, in which the synaptic input to a specific motor neuron subtype (VA) is altered in a manner that resembles that of another motor neuron subtype (VB)[12]. *unc-4* has since been shown to repress not just VB-specific synaptic wiring patterns in the VA neurons but other identity features of the VB neurons as well [61]. Additional repressor proteins (e.g., *cog-1* or *vab-7*) have since been identified whose loss results in de-repression or entire reversions of neuronal identity programs [62,63]. The repressor nature of proteins like UNC-4 or COG-1 is indicated by their proven or probable direct association with the co-repressor protein UNC-37, (orthologous to *Drosophila* Groucho), which recruits histone deacetylases to repress gene expression [63,64].

HHMI Author Manuscript HHMI Author Manuscript HHMI Author Manuscript

Intriguingly, the repressor proteins mentioned above exert their negative effect (either directly or indirectly) on terminal effector genes that are themselves direct targets of terminal selectors. For example, the acetylcholine receptor subunit *acr-5* is expressed in B-type motor neurons and is a direct target of *unc-3* in these neurons [45]. In A-type motor neurons, however, *acr-5* is repressed by *unc-4*, and *unc-4* is itself a target of *unc-3* [45,61]. *unc-3* therefore ensures that one of its direct targets is only expressed in a subtype of motor neurons, by inducing the expression of a subtype-specific repressor, *unc-4.* This regulatory configuration constitutes an *incoherent feedforward motif* [51].

The principle of modifying a terminal selector-induced differentiation program through selective repression in neuronal subtypes has been observed in another intriguing neurobiological context, namely the development of left/right asymmetric neuronal subtype specification. The left and right ASE neurons are the two main gustatory neurons in the worm and their terminal differentiation program is directly controlled by the *che-1* terminal selector [65]. While the majority of terminal features of ASE are expressed in both ASE neurons, a subset of putative chemoreceptor proteins, encoded by the *gcy* genes, are expressed exclusively in the left or the right ASE neuron. *gcy* genes are also direct targets of *che-1* but the ability of *che-1* to activate *gcy* genes is restricted by ASEL- or ASER-subtype specific regulatory factors [66]. Two of these repressors are the miRNA *lsy-6* (ASEL-specific) and its target, the homeobox gene *cog-1* (ASER-specific). Both *lsy-6* and *cog-1* are also direct targets of *che-1* [67,68].

The overall logic of incoherent feedforward regulation is to diversify a *ground state* into various sub-states. In an evolutionary context, one could envision an ancestral state in which multiple neurons were induced by the same transcription factor; the recruitment of a repressor in subsets of these neurons may have then diversified the spectrum of targets of the original inducer.

REGULATION OF PANNEURONAL FEATURES

One aspect of neuronal differentiation that has been remarkably unexplored in most model organisms is the acquisition of panneuronal features. A recent study addressed this problem in some depth, using a panel of genes that are expressed in all neurons throughout the nervous system (mostly, but not exclusively, genes involved in the synaptic vesicle cycle)[69]. Some surprising themes emerged from the analysis of the regulation of more than a dozen pan-neuronal genes. To appreciate these themes it is important to first consider the *cis*-regulatory architecture of genes that are expressed in a neuron-type specific manner. As schematically indicated in Fig. 9, *cis*-regulatory control regions of neuron-type specific genes are organized in a modular manner, composed of response elements to neuron-type specific terminal selectors. For example, the choline reuptake transporter *cho-1* (involved in clearance of the breakdown product of ACh at cholinergic synapses) contains separable modules with validated binding sites for TTX-3/CEH-10, required and sufficient for expression in the cholinergic AIY interneurons and a module with a binding site for UNC-3, which is required and sufficient for expression of *cho-1* in ventral nerve cord motor neurons [45]. Mutation of these sites in the context of fosmid-based reporters indicates the requirement of these single sites for expression in specific neuron types. This theme

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript

reiterates over many of the other neuron-type specific terminal effector genes tested. The *cis*-regulatory control regions of pan-neuronal genes are organized in a strikingly distinct manner, characterized by redundant, independently acting regulatory inputs (Fig. 9) [69]. Viewed from an evolutionary perspective, this regulatory architecture indicates that pan-neuronal genes have accumulated responsiveness to many different regulatory factors expressed in a terminally differentiated neuron, perhaps with the purpose of ensuring robustness of gene expression.

HOW IS THE REGULATORY MAP ESTABLISHED?

While terminal differentiation programs have been described in many neuron types throughout the *C. elegans* nervous system, much less is know about earlier neuronal patterning events in *C. elegans*. One obvious question is how the expression of terminal selector-type transcription factors is restricted to specific cell types and therefore, how the regulatory map described above is established. An earlier review has covered a few aspects of neuronal lineage control in *C. elegans* [47]. I will focus here instead on three aspects of early neuronal patterning, some of which directly relating to the control of terminal selectors.

The binary Wnt system

In the 1990s, Ralf Schnabel, Jim Priess and others discovered an intriguing binary patterning system that operates during *C. elegans* embryogenesis [70,71]. A non-canonical Wnt signaling system results in an anterior/posterior difference in the activity of transcriptional outputs of the Wnt system [72]. This system has subsequently been explicitly shown to also be involved late in embryonic nervous system patterning, directly regulating the expression of terminal selectors of neuronal identity [73]. This work, previously reviewed in [74] provided the first link of terminal patterning of neuronal identity to earlier patterning events, demonstrating a transition from *transient regulatory states* to *terminal, maintained regulatory states*. Recent work describes how this system operates in distinct neuroblasts [75].

Due to its binary nature, the anterior/posterior patterning system can obviously not act alone to specify the complex expression patterns of terminal selectors. The Wnt output regulators must interact with other transcription factors. The above-mentioned studies identified several such factors as being involved in defining terminal selector expression [73,75]. The multiplicity of regulatory inputs into terminal selector loci has led to the suggestion of an *hourglass regulatory architecture*, in which terminal selectors are nodes that integrate various transient regulatory inputs into a stable, maintained regulatory state [76].

Differences and similarities in early patterning – the bHLH and SoxB/C cases

Specific types of basic helix-loop-helix transcription factors operate as proneural genes throughout the animal kingdom and they have also been identified in *C. elegans*. For example, in one case, a proneural gene *hlh-14* controls the neuronal (vs. hypodermal) identity of a lineage that produces the gustatory ASE neurons [77]. As expected, *hlh-14* is expressed earlier in the lineage than the terminal selector of ASE identity (the *che-1* transcription factor), but its expression persists long enough to make it conceivable that

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript

hlh-14 may be involved in directly controlling *che-1* expression, perhaps in conjunction with the nuclear hormone receptor *nhr*-67, another upstream regulator of *che-1* expression [78]. Similarly, the postembryonically generated touch receptor neurons AVM and PVM fail to be generated in animals lacking the proneural gene *lin-32/*Atonal [79], which may be involved in directly or indirectly activating expression of the terminal selector of AVM/PVM identity, *unc-86* [80]. Whether proneural bHLH genes are required for neuronal fate induction in all neurons of the nervous system (a notion that is surprisingly unexplored in other systems) is currently being investigated.

A recent study shows that one aspect of early neuronal patterning in *C. elegans* displays a remarkable difference to early neuronal patterning in other organisms [81]. The HMG box transcription factor SOX-2*,* a member of the SoxB subfamily, is deeply conserved sequence-wise throughout the animal kingdom. Its function has been addressed in a wide range of animals from deuterostomes (chordates, echinoderms, hemichordates) and protostomes (Ecdysozoa, Lophotrochozoa) to even earlier diverging animal lineages such as cnidarians. In all these cases *sox-2* was shown to be expressed in neuronal precursors (upstream of the above-mentioned proneural bHLH genes) and whenever functional evidence was available, a role for *sox-2* in nervous system development was elucidated (discussed in Ref. [81]). Curiously, *sox-2* function in *C. elegans* is different. *sox-2* is not broadly expressed in neuronal precursors and *sox-2* has no function in embryonic neurogenesis [81]. Again in contrast to other organisms, the SoxC-type *sox-3* gene also has no role in embryonic neurogenesis [81]. *C. elegans sox-2* seems to rather be involved in driving terminal differentiation of select neuron types, as well as in the specification of a subset of postembryonic neuroblast [81,82]. The absence of a neurogenic *sox-2* function in *C. elegans* may relate to what *sox-2* does in other organisms – maintaining the pro-neuronal developmental potential of neuroepithelial cells during phases of cell proliferation in the developing nervous system. Such expansion may not occur in *C. elegans* embryogenesis.

Redundancy of early regulators

It is remarkable how few regulators of early neuronal patterning have been retrieved through genetic mutant screens in *C. elegans*. At first sight, a relatively trivial explanation could be essential pleiotropic functions of regulatory factors that may hinder retrieval from genetic screens. However, hypomorphic alleles of essential early patterning genes or even regulatory null alleles (i.e. loss of a gene only in specific cells due to mutations in *cis*-regulatory elements) have been retrieved in screens for terminal differentiation mutants. For example, a mutant allele of the essential Distalless-like homeobox gene *ceh-43* disrupts dopaminergic neuron differentiation and is defined by a deletion of a distal enhancer required for *ceh-43* expression in dopaminergic neurons [83]. Also, RNAi-based screens have been conducted for early regulator of ASE lineage specification and differentiation, recovering little more than the above-mentioned, and somewhat expected proneural bHLH factor for this lineage, *hlh-14* [77]. One very intriguing possibility for the paucity of early patterning genes is suggested by an older study [84] and a more recent study [85]. Priess and colleagues found that early patterning in the ABa lineage, which generates about half of the *C. elegans* nervous system, is controlled by two redundantly acting T-box genes, *tbx-37* and *tbx-38* [84]. Both genes are close paralogs and both need to be genetically eliminated to observe patterning

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript

defects. Notably, T-box genes expanded significantly in *C. elegans*, and many come in closely paralogous pairs [86]. It will be most interesting to determine whether it is a common theme that T-box factors act redundantly during early patterning decisions.

An intriguing, more recent example of redundancy shows that two unrelated homeobox genes, the Pitx-type homeobox gene *unc-30* and the Otx-type homeobox gene *ceh-36* act redundantly in early neuronal lineage decisions [85]. Both genes act as terminal selectors in defined neuronal cell types [44,87], but apparently also moonlight as patterning factors at earlier time points in lineages where their expression overlaps. It is tantalizing to think that early patterning in *C. elegans* is highly buffered through redundantly acting factors.

CONCLUDING REMARKS

More than half a century after Sydney Brenner first formulated his *C. elegans* research program [1,2], it is deeply satisfying to see how the *C. elegans* field has moved toward fulfilling Sydney Brenner's vision of using the tools of the microbial genetic trade, in combination with his visionary map making efforts, to understand how a nervous system develops. One reason this research program has been particularly successful in the nervous system is the lack of pleiotropies associated with regulators of terminal neuronal identity. In fact, such lack of pleiotropies were recognized by Brenner very early in his analysis of behavioral mutants by electron micrographical analysis [88]. Molecular, functional and expression studies subsequently revealed that most transcription factors that are components of the regulatory map are exclusively expressed in the nervous system, i.e. they do not have what could be essential functions elsewhere which could have prevented their retrieval from forward genetic screens. Within the nervous system many of these factors also appear to be committed to controlling terminal differentiation, without being involved in earlier patterning events. One exception to this notion are the *unc-30* and *ceh-36* terminal selectors, which, apart from their terminal selector function in distinct neuron types, have redundant function in controlling earlier lineage decisions [85]. These findings illustrate the importance of not solely relying on forward genetic screens, but also taking into account expression patterns and hence possibly redundant gene functions.

One major future challenge lies in connecting terminal regulators of neuronal identity to early lineage patterning events. Using automated lineage tracing in combination with RNAi-mediated gene knockdown, Du et al. have recently identified a large number of genes with roles in early lineage specification [89]. How these genes are coupled to the induction of expression of terminal selectors requires future exploration.

In addition to studying early neuronal patterning events, a great number of questions remain at the level of terminal differentiation programs. For example, the nature of the combinatorial codes of transcription factors that control the differentiated state should be more extensively characterized, perhaps with a specific focus on the ~100 homeobox genes encoded by the *C. elegans* genome. Importantly, we need a much more comprehensive understanding of the extent to which terminal differentiation programs are controlled in a co-regulatory manner by terminal selector-type transcription factors versus being regulated in a piecemeal manner by distinct cohorts of transcription factors. One important step in this

HHMI Author Manuscript HHMI Author Manuscript HHMI Author Manuscript

direction needs to be a systematic transcriptome analysis of each and every neuron in the worm. This is now technically feasible and it is easy to predict that the worm, after having been the first organism with a lineage, connectome and genome, will also be the first with a complete expression atlas with single cell resolution. The regulatory programs that control panneuronal genes (e.g. synaptic vesicle machinery) need to be better understood. The plasticity of the terminally differentiated state will also need to be examined on a much more detailed level. For example, how do postmitotic, embryonically generated neurons alter their differentiated state at defined postembryonic time points? How does circuit activity impinge on the differentiated state of specific neuronal circuit components? How does the sexual identity of the organism impinge on neuronal differentiation? Many of these questions can be addressed by continuing to pursue Sydney Brenner's vision, building higher resolution and now multi-dimensional maps (gene expression maps at distinct time points, under distinct external conditions), and by then exploiting the unique strengths of the *C. elegans* model system to genetically dissect these molecular maps through classic mutant analysis.

Acknowledgments

I thank Daniela Raciti and Wen Chen at Wormbase for extracting expression data from Wormbase, Lori Glenwinkel for building the resulting molecular map, Nuria Flames and Marie Gendrel for sharing unpublished data, Ev Yemini and Nuria Flames for comments on the manuscript. Stephen Larsen at Openworm and Chris Gove at WormBase kindly provided the image in Fig. 1A. Work in my laboratory is funded by the NIH and HHMI.

References

1. Brenner S. The genetics of Caenorhabditis elegans. Genetics. 1974; 77:71–94. [PubMed: 4366476]
2. Brenner, S. Foreword. In: Wood, WB., editor. The Nematode Caenorhabditis elegans. Cold Spring Harbor Laboratory Press; 1988.
3. Hall, DH.; Altun, Z. C. Elegans Atlas. Cold Spring Harbor Laboratory Press; 2007.
4. White JG, Southgate E, Thomson JN, Brenner S. The structure of the nervous system of the nematode *Caenorhabditis elegans*. Philosophical Transactions of the Royal Society of London B. Biological Sciences. 1986; 314:1–340. [PubMed: 22462104]
5. Jarrell TA, et al. The connectome of a decision-making neural network. Science. 2012; 337:437–44. [PubMed: 22837521]
6. Sulston JE, Horvitz HR. Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. Dev Biol. 1977; 56:110–56. [PubMed: 838129]
7. Sulston JE, Schierenberg E, White JG, Thomson JN. The embryonic cell lineage of the nematode Caenorhabditis elegans. Dev Biol. 1983; 100:64–119. [PubMed: 6684600]
8. Hedgecock EM, Culotti JG, Hall DH. The unc-5, unc-6, and unc-40 genes guide circumferential migrations of pioneer axons and mesodermal cells on the epidermis in C. elegans. Neuron. 1990; 4:61–85. [PubMed: 2310575]
9. Wadsworth WG, Bhatt H, Hedgecock EM. Neuroglia and pioneer neurons express UNC-6 to provide global and local netrin cues for guiding migrations in C. elegans. Neuron. 1996; 16:35–46. [PubMed: 8562088]
10. Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell. 1993; 75:843–54. [PubMed: 8252621]
11. Reinhart BJ, et al. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature. 2000; 403:901–6. [PubMed: 10706289]
12. Miller DM, et al. C. elegans unc-4 gene encodes a homeodomain protein that determines the pattern of synaptic input to specific motor neurons. Nature. 1992; 355:841–5. [PubMed: 1347150]
13. Chalfie M, Sulston J. Developmental genetics of the mechanosensory neurons of Caenorhabditis elegans. Dev Biol. 1981; 82:358–70. [PubMed: 7227647]

HHMI Author Manuscript HHMI Author Manuscript HHMI Author Manuscript

14. Bargmann CI, Hartwieg E, Horvitz HR. Odorant-selective genes and neurons mediate olfaction in C. elegans. Cell. 1993; 74:515–27. [PubMed: 8348618]
15. Mori I, Ohshima Y. Neural regulation of thermotaxis in Caenorhabditis elegans. Nature. 1995; 376:344–8. [PubMed: 7630402]
16. Kaplan JM, Horvitz HR. A dual mechanosensory and chemosensory neuron in Caenorhabditis elegans. Proc Natl Acad Sci U S A. 1993; 90:2227–31. [PubMed: 8460126]
17. Beverly M, Anbil S, Sengupta P. Degeneracy and neuromodulation among thermosensory neurons contribute to robust thermosensory behaviors in Caenorhabditis elegans. J Neurosci. 2011; 31:11718–27. [PubMed: 21832201]
18. Chatzigeorgiou M, et al. Specific roles for DEG/ENaC and TRP channels in touch and thermosensation in C. elegans nociceptors. Nat Neurosci. 2010; 13:861–8. [PubMed: 20512132]
19. Huang M, Chalfie M. Gene interactions affecting mechanosensory transduction in Caenorhabditis elegans. Nature. 1994; 367:467–70. [PubMed: 7509039]
20. Sengupta P, Chou JH, Bargmann CI. odr-10 encodes a seven transmembrane domain olfactory receptor required for responses to the odorant diacetyl. Cell. 1996; 84:899–909. [PubMed: 8601313]
21. Way JC, Chalfie M. mec-3, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in C. elegans. Cell. 1988; 54:5–16. [PubMed: 2898300]
22. Uchida O, Nakano H, Koga M, Ohshima Y. The C. elegans che-1 gene encodes a zinc finger transcription factor required for specification of the ASE chemosensory neurons. Development. 2003; 130:1215–24. [PubMed: 12588839]
23. Hobert O, et al. Regulation of interneuron function in the C. elegans thermoregulatory pathway by the ttx-3 LIM homeobox gene. Neuron. 1997; 19:345–57. [PubMed: 9292724]
24. Satterlee JS, et al. Specification of Thermosensory Neuron Fate in C. elegans Requires ttx-1, a Homolog of otd/Otx. Neuron. 2001; 31:943–56. [PubMed: 11580895]
25. Troemel ER, Chou JH, Dwyer ND, Colbert HA, Bargmann CI. Divergent seven transmembrane receptors are candidate chemosensory receptors in C. elegans. Cell. 1995; 83:207–18. [PubMed: 7585938]
26. Ortiz CO, et al. Searching for neuronal left/right asymmetry: genomewide analysis of nematode receptor-type guanylyl cyclases. Genetics. 2006; 173:131–49. [PubMed: 16547101]
27. Nathoo AN, Moeller RA, Westlund BA, Hart AC. Identification of neuropeptide-like protein gene families in Caenorhabditiselegans and other species. Proc Natl Acad Sci U S A. 2001; 98:14000–5. [PubMed: 11717458]
28. Li C, Kim K, Nelson LS. FMRFamide-related neuropeptide gene family in Caenorhabditis elegans. Brain Res. 1999; 848:26–34. [PubMed: 10612695]
29. Salkoff L, et al. Evolution tunes the excitability of individual neurons. Neuroscience. 2001; 103:853–9. [PubMed: 11301195]
30. Tsalik EL, et al. LIM homeobox gene-dependent expression of biogenic amine receptors in restricted regions of the C. elegans nervous system. Dev Biol. 2003; 263:81–102. [PubMed: 14568548]
31. Aurelio O, Hall DH, Hobert O. Immunoglobulin-domain proteins required for maintenance of ventral nerve cord organization. Science. 2002; 295:686–90. [PubMed: 11809975]
32. Hunt-Newbury R, et al. High-throughput in vivo analysis of gene expression in Caenorhabditis elegans. PLoS Biol. 2007; 5:e237. [PubMed: 17850180]
33. Sunkin SM, et al. Allen Brain Atlas: an integrated spatio-temporal portal for exploring the central nervous system. Nucleic Acids Res. 2013; 41:D996–D1008. [PubMed: 23193282]
34. Jenett A, et al. A GAL4-driver line resource for Drosophila neurobiology. Cell Rep. 2012; 2:991–1001. [PubMed: 23063364]
35. Hobert O. The neuronal genome of Caenorhabditis elegans. WormBook. 2013:1–106. [PubMed: 24081909]
36. McIntire SL, Jorgensen E, Horvitz HR. Genes required for GABA function in Caenorhabditis elegans. Nature. 1993; 364:334–7. [PubMed: 8332190]

HHMI Author Manuscript HHMI Author Manuscript HHMI Author Manuscript

37. Chase DL, Koelle MR. Biogenic amine neurotransmitters in C. elegans. WormBook. 2007:1–15. [PubMed: 18050501]
38. Pereira L, et al. A cellular and regulatory map of the cholinergic nervous system of. Elife. 2015; 4
39. Serrano-Saiz E, et al. Modular Control of Glutamatergic Neuronal Identity in C. elegans by Distinct Homeodomain Proteins. Cell. 2013; 155:659–673. [PubMed: 24243022]
40. Deneris ES, Hobert O. Maintenance of postmitotic neuronal cell identity. Nat Neurosci. 2014; 17:899–907. [PubMed: 24929660]
41. Wenick AS, Hobert O. Genomic cis-Regulatory Architecture and trans-Acting Regulators of a Single Interneuron-Specific Gene Battery in C. elegans. Dev Cell. 2004; 6:757–70. [PubMed: 15177025]
42. Zhang Y, et al. Identification of genes expressed in C. elegans touch receptor neurons. Nature. 2002; 418:331–5. [PubMed: 12124626]
43. Etchberger JF, et al. The molecular signature and cis-regulatory architecture of a C. elegans gustatory neuron. Genes Dev. 2007; 21:1653–74. [PubMed: 17606643]
44. Cinar H, Keles S, Jin Y. Expression profiling of GABAergic motor neurons in Caenorhabditis elegans. Curr Biol. 2005; 15:340–6. [PubMed: 15723795]
45. Kratsios P, Stolfi A, Levine M, Hobert O. Coordinated regulation of cholinergic motor neuron traits through a conserved terminal selector gene. Nat Neurosci. 2011; 15:205–14. [PubMed: 22119902]
46. Hobert O. Terminal selectors of neuronal identity. Curr Top Dev Biol. 2016 in press.
47. Hobert O. Neurogenesis in the nematode Caenorhabditis elegans. WormBook. 2010:1–24. [PubMed: 20891032]
48. Altun-Gultekin Z, et al. A regulatory cascade of three homeobox genes, ceh-10, ttx-3 and ceh-23, controls cell fate specification of a defined interneuron class in C. elegans. Development. 2001; 128:1951–69. [PubMed: 11493519]
49. Gordon PM, Hobert O. A Competition Mechanism for a Homeotic Neuron Identity Transformation in C. elegans. Dev Cell. 2015; 34:206–19. [PubMed: 26096732]
50. Johnston RJ Jr, et al. An unusual Zn-finger/FH2 domain protein controls a left/right asymmetric neuronal fate decision in C. elegans. Development. 2006; 133:3317–28. [PubMed: 16887832]
51. Alon U. Network motifs: theory and experimental approaches. Nat Rev Genet. 2007; 8:450–61. [PubMed: 17510665]
52. Caicedo A, Pereira E, Margolskee RF, Roper SD. Role of the G-protein subunit alpha-gustducin in taste cell responses to bitter stimuli. J Neurosci. 2003; 23:9947–52. [PubMed: 14586025]
53. Wightman B, Ebert B, Carmean N, Weber K, Clever S. The C. elegans nuclear receptor gene fax-1 and homeobox gene unc-42 coordinate interneuron identity by regulating the expression of glutamate receptor subunits and other neuron-specific genes. Dev Biol. 2005; 287:74–85. [PubMed: 16183052]
54. Baran R, Aronoff R, Garriga G. The C. elegans homeodomain gene unc-42 regulates chemosensory and glutamate receptor expression. Development. 1999; 126:2241–51. [PubMed: 10207148]
55. Brockie PJ, Madsen DM, Zheng Y, Mellem J, Maricq AV. Differential expression of glutamate receptor subunits in the nervous system of Caenorhabditis elegans and their regulation by the homeodomain protein UNC-42. J Neurosci. 2001; 21:1510–22. [PubMed: 11222641]
56. Shaham S, Bargmann CI. Control of neuronal subtype identity by the C. elegans ARID protein CFI-1. Genes Dev. 2002; 16:972–83. [PubMed: 11959845]
57. Xue D, Tu Y, Chalfie M. Cooperative interactions between the Caenorhabditis elegans homeoproteins UNC-86 and MEC-3. Science. 1993; 261:1324–8. [PubMed: 8103239]
58. Zhang F, et al. The LIM and POU homeobox genes ttx-3 and unc-86 act as terminal selectors in distinct cholinergic and serotonergic neuron types. Development. 2014; 141:422–35. [PubMed: 24353061]
59. Kagoshima H, Kohara Y. Co-expression of the transcription factors CEH-14 and TTX-1 regulates AFD neuron-specific genes gcy-8 and gcy-18 in C. elegans. Dev Biol. 2015; 399:325–36. [PubMed: 25614239]

HHMI Author Manuscript HHMI Author Manuscript HHMI Author Manuscript

60. Van Buskirk C, Sternberg PW. Paired and LIM class homeodomain proteins coordinate differentiation of the C. elegans ALA neuron. Development. 2010; 137:2065–74. [PubMed: 20501595]
61. Winnier AR, et al. UNC-4/UNC-37-dependent repression of motor neuron-specific genes controls synaptic choice in Caenorhabditis elegans. Genes Dev. 1999; 13:2774–86. [PubMed: 10557206]
62. Esmaeili B, Ross JM, Neades C, Miller DM 3rd, Ahringer J. The C. elegans even-skipped homologue, vab-7, specifies DB motoneurone identity and axon trajectory. Development. 2002; 129:853–62. [PubMed: 11861469]
63. Chang S, Johnston RJ Jr, Hobert O. A transcriptional regulatory cascade that controls left/right asymmetry in chemosensory neurons of C. elegans. Genes Dev. 2003; 17:2123–37. [PubMed: 12952888]
64. Pflugrad A, Meir JY, Barnes TM, Miller DM 3rd. The Groucho-like transcription factor UNC-37 functions with the neural specificity gene unc-4 to govern motor neuron identity in C. elegans. Development. 1997; 124:1699–709. [PubMed: 9165118]
65. Hobert O. Development of left/right asymmetry in the Caenorhabditis elegans nervous system: from zygote to postmitotic neuron. Genesis. 2014; 52:528–43. [PubMed: 24510690]
66. Etchberger JF, Flowers EB, Poole RJ, Bashllari E, Hobert O. Cis-regulatory mechanisms of left/right asymmetric neuron-subtype specification in C. elegans. Development. 2009; 136:147–60. [PubMed: 19060335]
67. O'Meara MM, et al. Cis-regulatory Mutations in the Caenorhabditis elegans Homeobox Gene Locus cog-1 Affect Neuronal Development. Genetics. 2009; 181:1679–1686. [PubMed: 19189954]
68. Sarin S, et al. Genetic Screens for Caenorhabditis elegans Mutants Defective in Left/Right Asymmetric Neuronal Fate Specification. Genetics. 2007; 176:2109–30. [PubMed: 17717195]
69. Stefanakis N, Carrera I, Hobert O. Regulatory Logic of Pan-Neuronal Gene Expression in C. elegans. Neuron. 2015; 87:733–50. [PubMed: 26291158]
70. Kaletta T, Schnabel H, Schnabel R. Binary specification of the embryonic lineage in Caenorhabditis elegans. Nature. 1997; 390:294–8. [PubMed: 9384382]
71. Lin R, Hill RJ, Priess JR. POP-1 and anterior-posterior fate decisions in C. elegans embryos. Cell. 1998; 92:229–39. [PubMed: 9458047]
72. Mizumoto K, Sawa H. Two betas or not two betas: regulation of asymmetric division by beta-catenin. Trends Cell Biol. 2007; 17:465–73. [PubMed: 17919911]
73. Bertrand V, Hobert O. Linking asymmetric cell division to the terminal differentiation program of postmitotic neurons in C. elegans. Dev Cell. 2009; 16:563–75. [PubMed: 19386265]
74. Bertrand V, Hobert O. Lineage programming: navigating through transient regulatory states via binary decisions. Curr Opin Genet Dev. 2010; 20:362–8. [PubMed: 20537527]
75. Murgan S, et al. Atypical Transcriptional Activation by TCF via a Zic Transcription Factor in C. elegans Neuronal Precursors. Dev Cell. 2015; 33:737–45. [PubMed: 26073017]
76. Hobert O. Regulation of terminal differentiation programs in the nervous system. Annu Rev Cell Dev Biol. 2011; 27:681–96. [PubMed: 21985672]
77. Poole RJ, Bashllari E, Cochella L, Flowers EB, Hobert O. A Genome-Wide RNAi Screen for Factors Involved in Neuronal Specification in Caenorhabditis elegans. PLoS Genet. 2011; 7:e1002109. [PubMed: 21698137]
78. Sarin S, Antonio C, Tursun B, Hobert O. The C. elegans Tailless/TLX transcription factor nhr-67 controls neuronal identity and left/right asymmetric fate diversification. Development. 2009; 136:2933–44. [PubMed: 19641012]
79. Zhang Y, Emmons SW. Specification of sense-organ identity by a Caenorhabditis elegans Pax-6 homologue. Nature. 1995; 377:55–9. [PubMed: 7659160]
80. Baumeister R, Liu Y, Ruvkun G. Lineage-specific regulators couple cell lineage asymmetry to the transcription of the Caenorhabditis elegans POU gene unc-86 during neurogenesis. Genes Dev. 1996; 10:1395–410. [PubMed: 8647436]
81. Vidal B, et al. C. elegans SoxB genes are dispensable for embryonic neurogenesis but required for terminal differentiation of specific neuron types. Development. 2015; 142:2464–77. [PubMed: 26153233]

HHMI Author Manuscript HHMI Author Manuscript HHMI Author Manuscript

82. Alqadah A, et al. Postmitotic diversification of olfactory neuron types is mediated by differential activities of the HMG-box transcription factor SOX-2. EMBO J. 2015; 34:2574–89. [PubMed: 26341465]
83. Doitsidou M, et al. A combinatorial regulatory signature controls terminal differentiation of the dopaminergic nervous system in C. elegans. Genes Dev. 2013; 27:1391–1405. [PubMed: 23788625]
84. Good K, et al. The T-box transcription factors TBX-37 and TBX-38 link GLP-1/Notch signaling to mesoderm induction in C. elegans embryos. Development. 2004; 131:1967–78. [PubMed: 15056620]
85. Walton T, et al. The Bicoid class homeodomain factors ceh-36/OTX and unc-30/PITX cooperate in C. elegans embryonic progenitor cells to regulate robust development. PLoS Genet. 2015; 11:e1005003. [PubMed: 25738873]
86. Ruvkun G, Hobert O. The Taxonomy of Developmental Control in Caenorhabditis elegans. Science. 1998; 282:2033–2041. [PubMed: 9851920]
87. Lanjuin A, VanHoven MK, Bargmann CI, Thompson JK, Sengupta P. Otx/otd Homeobox Genes Specify Distinct Sensory Neuron Identities in C. elegans. Dev Cell. 2003; 5:621–33. [PubMed: 14536063]
88. Brenner S. The Genetics of Behavior. Br Med Bull. 1973; 29:269–271. [PubMed: 4807330]
89. Du Z, et al. The Regulatory Landscape of Lineage Differentiation in a Metazoan Embryo. Dev Cell. 2015; 34:592–607. [PubMed: 26321128]
90. Varshney LR, Chen BL, Paniagua E, Hall DH, Chklovskii DB. Structural properties of the Caenorhabditis elegans neuronal network. PLoS Comput Biol. 2011; 7:e1001066. [PubMed: 21304930]
91. Margeta MA, Wang GJ, Shen K. Clathrin adaptor AP-1 complex excludes multiple postsynaptic receptors from axons in C. elegans. Proc Natl Acad Sci U S A. 2009; 106:1632–7. [PubMed: 19164532]
92. Duggan A, Chalfie M. Control of neuronal development in Caenorhabditis elegans. Curr Opin Neurobiol. 1995; 5:6–9. [PubMed: 7773007]
93. Hsieh YW, Alqadah A, Chuang CF. Asymmetric neural development in the Caenorhabditis elegans olfactory system. Genesis. 2014; 52:544–54. [PubMed: 24478264]
94. Aurelio O, Boulin T, Hobert O. Identification of spatial and temporal cues that regulate postembryonic expression of axon maintenance factors in the C. elegans ventral nerve cord. Development. 2003; 130:599–610. [PubMed: 12490565]
95. Etchberger JF, Hobert O. Vector-free DNA constructs improve transgene expression in C. elegans. Nat Methods. 2008; 5:3. [PubMed: 18165801]
96. Tursun B, Patel T, Kratsios P, Hobert O. Direct conversion of C. elegans germ cells into specific neuron types. Science. 2011; 331:304–8. [PubMed: 21148348]
97. Hobert O. Maintaining a memory by transcriptional autoregulation. Curr Biol. 2011; 21:R146–7. [PubMed: 21334290]
98. Bargmann CI. Chemosensation in C. elegans. WormBook. 2006:1–29. [PubMed: 18050433]
99. Von Stetina SE, Treinin M, Miller DM. The Motor Circuit. Int Rev Neurobiol. 2006; 69:125–167. [PubMed: 16492464]
100. Sawin ER, Ranganathan R, Horvitz HR. C. elegans locomotory rate is modulated by the environment through a dopaminergic pathway and by experience through a serotonergic pathway. Neuron. 2000; 26:619–31. [PubMed: 10896158]
101. Hills T, Brockie PJ, Maricq AV. Dopamine and glutamate control area-restricted search behavior in Caenorhabditis elegans. J Neurosci. 2004; 24:1217–25. [PubMed: 14762140]
102. Bargmann CI, Horvitz HR. Control of larval development by chemosensory neurons in Caenorhabditis elegans. Science. 1991; 251:1243–6. [PubMed: 2006412]
103. Chalasani SH, et al. Dissecting a circuit for olfactory behaviour in Caenorhabditis elegans. Nature. 2007; 450:63–70. [PubMed: 17972877]
104. McIntire SL, Jorgensen E, Kaplan J, Horvitz HR. The GABAergic nervous system of Caenorhabditis elegans. Nature. 1993; 364:337–41. [PubMed: 8332191]

HHMI Author Manuscript HHMI Author Manuscript HHMI Author Manuscript

105. Gray JM, Hill JJ, Bargmann CI. A circuit for navigation in Caenorhabditis elegans. Proc Natl Acad Sci U S A. 2005; 102:3184–91. [PubMed: 15689400]
106. Bendena WG, et al. A Caenorhabditis elegans allatostatin/galanin-like receptor NPR-9 inhibits local search behavior in response to feeding cues. Proc Natl Acad Sci U S A. 2008; 105:1339–42. [PubMed: 18216257]
107. Chalfie M, et al. The neural circuit for touch sensitivity in Caenorhabditis elegans. J Neurosci. 1985; 5:956–64. [PubMed: 3981252]
108. Van Buskirk C, Sternberg PW. Epidermal growth factor signaling induces behavioral quiescence in Caenorhabditis elegans. Nat Neurosci. 2007; 10:1300–7. [PubMed: 17891142]
109. Sanders J, et al. The Caenorhabditis elegans interneuron ALA is (also) a high-threshold mechanosensor. BMC Neurosci. 2013; 14:156. [PubMed: 24341457]
110. Tsalik EL, Hobert O. Functional mapping of neurons that control locomotory behavior in Caenorhabditis elegans. J Neurobiol. 2003; 56:178–97. [PubMed: 12838583]
111. Coates JC, de Bono M. Antagonistic pathways in neurons exposed to body fluid regulate social feeding in Caenorhabditis elegans. Nature. 2002; 419:925–9. [PubMed: 12410311]
112. Gray JM, et al. Oxygen sensation and social feeding mediated by a C. elegans guanylate cyclase homologue. Nature. 2004; 430:317–22. [PubMed: 15220933]
113. Piggott BJ, Liu J, Feng Z, Wescott SA, Xu XZ. The neural circuits and synaptic mechanisms underlying motor initiation in C. elegans. Cell. 2011; 147:922–33. [PubMed: 22078887]
114. Greer ER, Perez CL, Van Gilst MR, Lee BH, Ashrafi K. Neural and molecular dissection of a C. elegans sensory circuit that regulates fat and feeding. Cell Metab. 2008; 8:118–31. [PubMed: 18680713]
115. Hart AC, Sims S, Kaplan JM. Synaptic code for sensory modalities revealed by C. elegans GLR-1 glutamate receptor. Nature. 1995; 378:82–5. [PubMed: 7477294]
116. Hardaker LA, Singer E, Kerr R, Zhou G, Schafer WR. Serotonin modulates locomotory behavior and coordinates egg-laying and movement in Caenorhabditis elegans. J Neurobiol. 2001; 49:303–313. [PubMed: 11745666]
117. Hutter H. Extracellular cues and pioneers act together to guide axons in the ventral cord of C. elegans. Development. 2003; 130:5307–18. [PubMed: 13129845]
118. Macosko EZ, et al. A hub-and-spoke circuit drives pheromone attraction and social behaviour in C. elegans. Nature. 2009; 458:1171–5. [PubMed: 19349961]
119. Kratsios P, et al. Transcriptional coordination of synaptogenesis and neurotransmitter signaling. Curr Biol. 2015; 25:1282–95. [PubMed: 25913400]
120. Kennerdell JR, Fetter RD, Bargmann CI. Wnt-Ror signaling to SIA and SIB neurons directs anterior axon guidance and nerve ring placement in C. elegans. Development. 2009; 136:3801–10. [PubMed: 19855022]
121. Biron D, Wasserman S, Thomas JH, Samuel AD, Sengupta P. An olfactory neuron responds stochastically to temperature and modulates Caenorhabditis elegans thermotactic behavior. Proc Natl Acad Sci U S A. 2008; 105:11002–7. [PubMed: 18667708]
122. Li W, Kang L, Piggott BJ, Feng Z, Xu XZ. The neural circuits and sensory channels mediating harsh touch sensation in Caenorhabditis elegans. Nat Commun. 2011; 2:315. [PubMed: 21587232]
123. Hallem EA, Sternberg PW. Acute carbon dioxide avoidance in Caenorhabditis elegans. Proc Natl Acad Sci U S A. 2008; 105:8038–43. [PubMed: 18524955]
124. Li W, Feng Z, Sternberg PW, Xu XZ. A C. elegans stretch receptor neuron revealed by a mechanosensitive TRP channel homologue. Nature. 2006; 440:684–7. [PubMed: 16572173]
125. Lee H, et al. Nictation, a dispersal behavior of the nematode Caenorhabditis elegans, is regulated by IL2 neurons. Nat Neurosci. 2012; 15:107–12. [PubMed: 22081161]
126. Waggoner LE, Zhou GT, Schafer RW, Schafer WR. Control of alternative behavioral states by serotonin in Caenorhabditis elegans. Neuron. 1998; 21:203–14. [PubMed: 9697864]
127. Chang HC, Paek J, Kim DH. Natural polymorphisms in C. elegans HECW-1 E3 ligase affect pathogen avoidance behaviour. Nature. 2011; 480:525–9. [PubMed: 22089131]

HHMI Author Manuscript HHMI Author Manuscript HHMI Author Manuscript

128. Hilliard MA, Bargmann CI, Bazzicalupo P. C. elegans responds to chemical repellents by integrating sensory inputs from the head and the tail. Curr Biol. 2002; 12:730–4. [PubMed: 12007416]

129. Liu S, Schulze E, Baumeister R. Temperature- and touch-sensitive neurons couple CNG and TRPV channel activities to control heat avoidance in Caenorhabditis elegans. PLoS One. 2012; 7:e32360. [PubMed: 22448218]

130. Nelson MD, et al. The neuropeptide NLP-22 regulates a sleep-like state in Caenorhabditis elegans. Nat Commun. 2013; 4:2846. [PubMed: 24301180]

131. Flavell SW, et al. Serotonin and the neuropeptide PDF initiate and extend opposing behavioral states in C. elegans. Cell. 2013; 154:1023–35. [PubMed: 23972393]

132. Chatzigeorgiou M, Schafer WR. Lateral facilitation between primary mechanosensory neurons controls nose touch perception in C. elegans. Neuron. 2011; 70:299–309. [PubMed: 21521615]

133. Turek M, Lewandrowski I, Bringmann H. An AP2 transcription factor is required for a sleep-active neuron to induce sleep-like quiescence in C. elegans. Curr Biol. 2013; 23:2215–23. [PubMed: 24184105]

134. Flames N, Hobert O. Gene regulatory logic of dopamine neuron differentiation. Nature. 2009; 458:885–9. [PubMed: 19287374]

135. Gendrel M, Hobert O. Transcription factor combinations specifying the GABAergic nervous system of C. elegans. 2016 in preparation.

136. Gramstrup Petersen J, et al. EGL-13/SoxD specifies distinct O2 and CO2 sensory neuron fates in Caenorhabditis elegans. PLoS Genet. 2013; 9:e1003511. [PubMed: 23671427]

137. Qin H, Powell-Coffman JA. The Caenorhabditis elegans aryl hydrocarbon receptor, AHR-1, regulates neuronal development. Dev Biol. 2004; 270:64–75. [PubMed: 15136141]

138. Feng G, et al. Developmental stage-dependent transcriptional regulatory pathways control neuroblast lineage progression. Development. 2013; 140:3838–47. [PubMed: 23946438]

139. Sarafi-Reinach TR, Melkman T, Hobert O, Sengupta P. The lin-11 LIM homeobox gene specifies olfactory and chemosensory neuron fates in C. elegans. Development. 2001; 128:3269–81. [PubMed: 11546744]

140. Kim K, Colosimo ME, Yeung H, Sengupta P. The UNC-3 Olf/EBF protein represses alternate neuronal programs to specify chemosensory neuron identity. Dev Biol. 2005; 286:136–48. [PubMed: 16143323]

141. Prasad BC, et al. unc-3, a gene required for axonal guidance in Caenorhabditis elegans, encodes a member of the O/E family of transcription factors. Development. 1998; 125:1561–8. [PubMed: 9502737]

142. Gonzalez-Barrios M, et al. Cis- and trans-regulatory mechanisms of gene expression in the ASJ sensory neuron of Caenorhabditis elegans. Genetics. 2015; 200:123–34. [PubMed: 25769980]

143. Alqadah A, et al. Postmitotic diversification of olfactory neuron types is mediated by differential activities of the HMG-box transcription factor SOX-2. EMBO J. 2015

144. Sagasti A, et al. The CaMKII UNC-43 activates the MAPKKK NSY-1 to execute a lateral signaling decision required for asymmetric olfactory neuron fates. Cell. 2001; 105:221–32. [PubMed: 11336672]

145. Nokes EB, et al. Cis-regulatory mechanisms of gene expression in an olfactory neuron type in Caenorhabditis elegans. Dev Dyn. 2009; 238:3080–92. [PubMed: 19924784]

146. Koga M, Ohshima Y. The C. elegans ceh-36 gene encodes a putative homemodomain transcription factor involved in chemosensory functions of ASE and AWC neurons. J Mol Biol. 2004; 336:579–87. [PubMed: 15095973]

147. Rojo Romanos T, Petersen JG, Riveiro AR, Pocock R. A novel role for the zinc-finger transcription factor EGL-46 in the differentiation of gas-sensing neurons in Caenorhabditis elegans. Genetics. 2015; 199:157–63. [PubMed: 25395666]

148. Brandt JP, et al. A single gene target of an ETS-family transcription factor determines neuronal CO2-chemosensitivity. PLoS ONE. 2012; 7:e34014. [PubMed: 22479504]

149. Zheng C, Diaz-Cuadros M, Chalfie M. Hox Genes Promote Neuronal Subtype Diversification through Posterior Induction in Caenorhabditis elegans. Neuron. 2015; 88:514–27. [PubMed: 26539892]

HHMI Author Manuscript HHMI Author Manuscript HHMI Author Manuscript

150. Sze JY, Zhang S, Li J, Ruvkun G. The C. elegans POU-domain transcription factor UNC-86 regulates the tph-1 tryptophan hydroxylase gene and neurite outgrowth in specific serotonergic neurons. Development. 2002; 129:3901–11. [PubMed: 12135927]
151. Kage E, et al. MBR-1, a novel helix-turn-helix transcription factor, is required for pruning excessive neurites in Caenorhabditis elegans. Curr Biol. 2005; 15:1554–9. [PubMed: 16139210]
152. Schmid C, Schwarz V, Hutter H. AST-1, a novel ETS-box transcription factor, controls axon guidance and pharynx development in C. elegans. Dev Biol. 2006; 293:403–13. [PubMed: 16584723]
153. Wightman B, Baran R, Garriga G. Genes that guide growth cones along the C. elegans ventral nerve cord. Development. 1997; 124:2571–80. [PubMed: 9216999]
154. Cameron S, Clark SG, McDermott JB, Aamodt E, Horvitz HR. PAG-3, a Zn-finger transcription factor, determines neuroblast fate in C. elegans. Development. 2002; 129:1763–74. [PubMed: 11923211]
155. Clark SG, Chiu C. C. elegans ZAG-1, a Zn-finger-homeodomain protein, regulates axonal development and neuronal differentiation. Development. 2003; 130:3781–94. [PubMed: 12835394]
156. Wacker I, Schwarz V, Hedgecock EM, Hutter H. zag-1, a Zn-finger homeodomain transcription factor controlling neuronal differentiation and axon outgrowth in C. elegans. Development. 2003; 130:3795–805. [PubMed: 12835395]
157. Hobert O, Tessmar K, Ruvkun G. The Caenorhabditis elegans lim-6 LIM homeobox gene regulates neurite outgrowth and function of particular GABAergic neurons. Development. 1999; 126:1547–1562. [PubMed: 10068647]
158. Lickteig KM, et al. Regulation of neurotransmitter vesicles by the homeodomain protein UNC- 4 and its transcriptional corepressor UNC-37/groucho in Caenorhabditis elegans cholinergic motor neurons. J Neurosci. 2001; 21:2001–14. [PubMed: 11245684]
159. Prasad B, Karakuzu O, Reed RR, Cameron S. unc-3-dependent repression of specific motor neuron fates in Caenorhabditis elegans. Dev Biol. 2008; 323:207–15. [PubMed: 18817768]
160. Jin Y, Hoskins R, Horvitz HR. Control of type-D GABAergic neuron differentiation by C. elegans UNC-30 homeodomain protein. Nature. 1994; 372:780–3. [PubMed: 7997265]
161. Eastman C, Horvitz HR, Jin Y. Coordinated transcriptional regulation of the unc-25 glutamic acid decarboxylase and the unc-47 GABA vesicular transporter by the Caenorhabditis elegans UNC-30 homeodomain protein. J Neurosci. 1999; 19:6225–34. [PubMed: 10414952]
162. Petersen SC, et al. A transcriptional program promotes remodeling of GABAergic synapses in Caenorhabditis elegans. J Neurosci. 2011; 31:15362–75. [PubMed: 22031882]
163. Lloret C, et al. Deep homology of a genetic programs controlling serotonergic neuron differentiation in nematodes and mammals. 2016 in preparation.
164. Richard JP, et al. Direct in vivo cellular reprogramming involves transition through discrete, non-pluripotent steps. Development. 2011; 138:1483–92. [PubMed: 21389048]
165. Kim J, et al. The Evolutionarily Conserved LIM Homeodomain Protein LIM-4/LHX6 Specifies the Terminal Identity of a Cholinergic and Peptidergic C. elegans Sensory/Inter/Motor Neuron-Type. PLoS Genet. 2015; 11:e1005480. [PubMed: 26305787]
166. Von Stetina SE, et al. UNC-4 represses CEH-12/HB9 to specify synaptic inputs to VA motor neurons in C. elegans. Genes Dev. 2007; 21:332–46. [PubMed: 17289921]
167. Melkman T, Sengupta P. Regulation of chemosensory and GABAergic motor neuron development by the C. elegans Aristaless/Arx homolog alr-1. Development. 2005; 132:1935–49. [PubMed: 15790968]
168. Zhou HM, Walthall WW. UNC-55, an Orphan Nuclear Hormone Receptor, Orchestrates Synaptic Specificity among Two Classes of Motor Neurons in Caenorhabditis elegans. J Neurosci. 1998; 18:10438–10444. [PubMed: 9852581]

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript

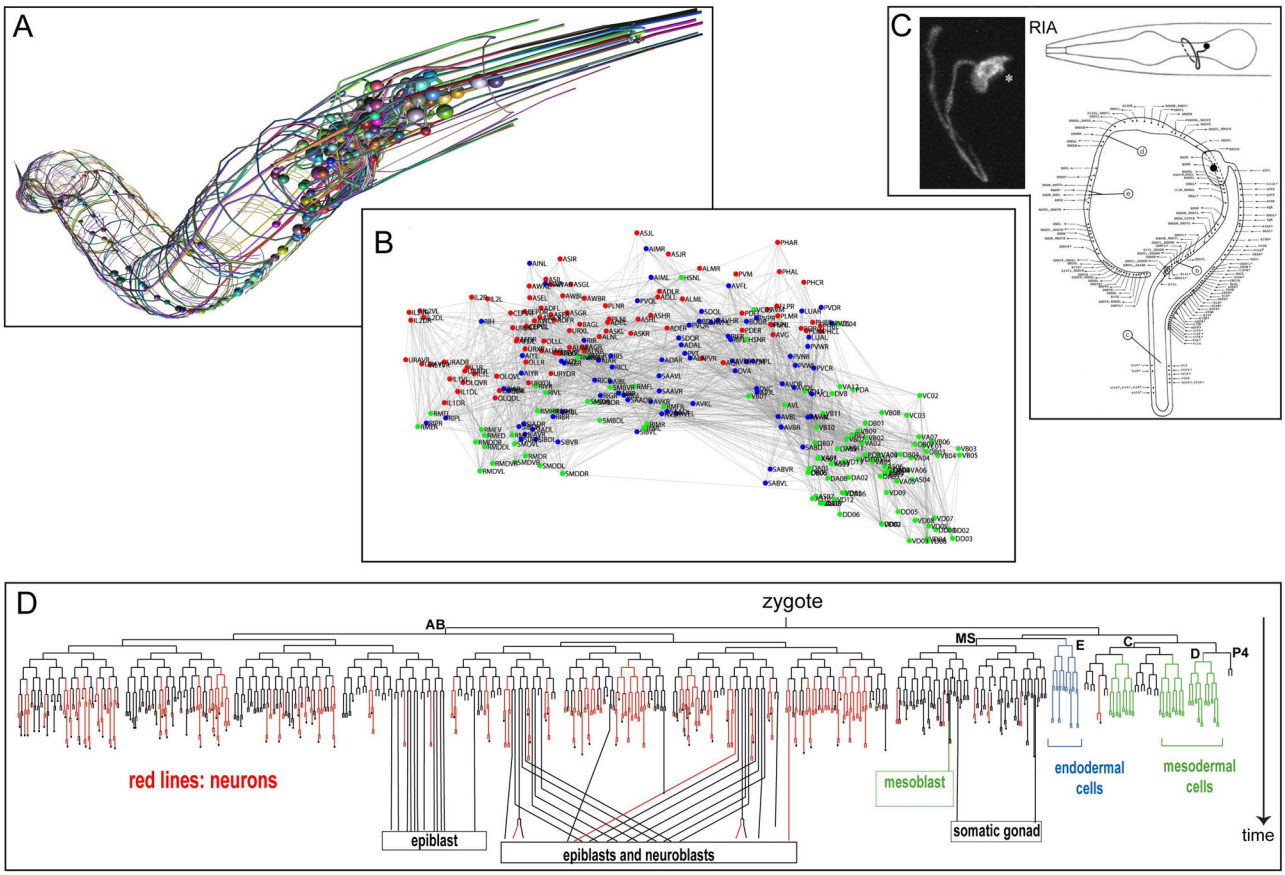


Fig. 1. An anatomical and lineage map of the *C. elegans* nervous system

A: Illustrations of the entire nervous system and fascicles, kindly provided by Openworm.org

B: Connectome. Reproduced from [90]

C: One example of a single neuron type, the glutamatergic RIA interneurons, labeled with gfp (reproduced from [91]) and its synaptic connectivity (reproduced from [4]).

D: Lineage of cells generated in the embryo. Red lines indicate neurons.

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript

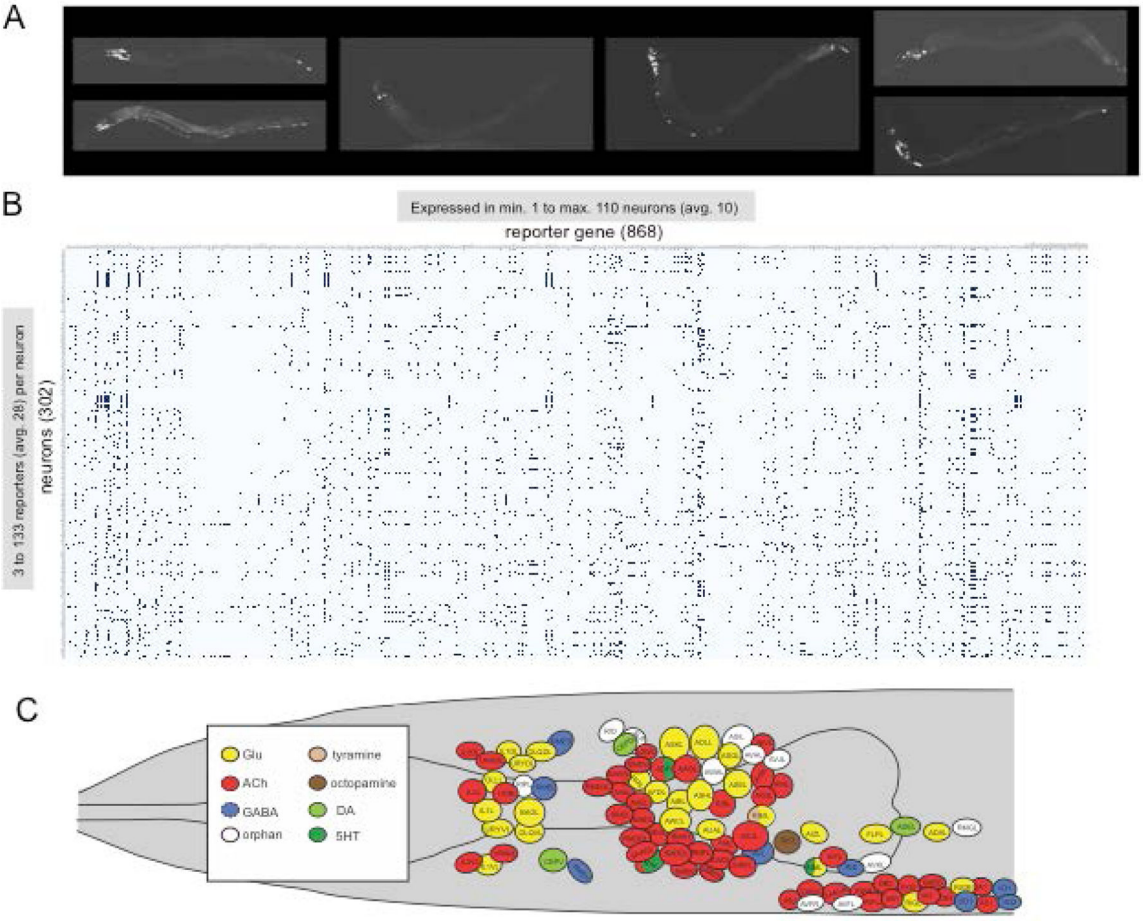


Fig. 2. Molecular maps of the *C. elegans* nervous system

A: Examples of transgenic worms expressing *gfp* reporter transgenes. From [41].

B: Manually curated expression patterns of transgenes throughout the C. elegans nervous system, kindly extracted from www.wormbase.org by Daniela Raciti and Wen Chen and organized by Lori Glenwinkel. Note that these transgenes may not display the complete expression pattern of the respective gene, but they nevertheless serve as invaluable read outs of the regulatory state of a neuron.

C: Schematic map of neurotransmitter usage in the *C. elegans* head. From [38].

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript

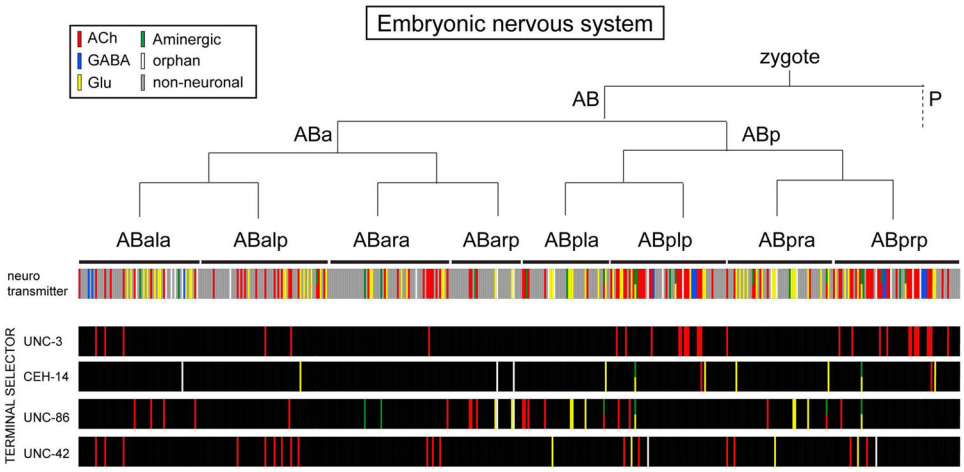
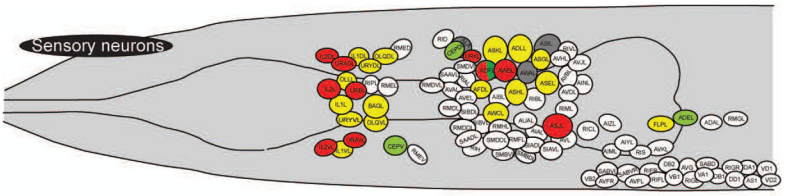


Fig. 3. Lack of correlation of lineage with neurotransmitter identity or transcription factor expression

Each bar represents an individual cell.

HHMI Author Manuscript HHMI Author Manuscript HHMI Author Manuscript



Sensory	Number	Neurotransmitter	Inducers	Repressor	Reference
ADE	2	DA	*ast-1 (7/7), ceh-43 (5/5)*		83,134
ADF	2	ACh & 5HT			
ADL	2	Glu	*lin-11 (1/1)*		39
AFD	2	Glu	*ttx-1 (7/7), ceh-14 (7/7)*		39,59
ALA	1	GABA	*ceh-14 (6/6), ceh-17 (6/6)*		60,135
ALM	2	Glu	*unc-86, mec-3 (71)**		42
ALN	2	ACh			
AQR	1	Glu	*unc-86 (1/1), egl-13 (2/2); ahr-1 (1/1)*		136-138
ASE	2	Glu	*che-1 (22/22 + cilium), ceh-36, die-1, lim-6 *****	*cog-1*	22,43
ASG	2	Glu	*lin-11(5/5), ceh-37(1/1)*		39,139
ASH	2	Glu	*unc-42 (5/5 + cilium)*		39,54
ASI	2	unknown	*unc-3 (6/23)*	*unc-3*	140,141 ***
ASJ	2	ACh	*sptf-1 (4/4)*		142
ASK	2	Glu	*ttx-3 (4/4 + cilium)*		39
AVM	1	Glu	*unc-86, mec-3 (71)**		42
AWA	2	unknown	*odr-7 (1/1)*		20
AWB	2	ACh	*lim-4 (4/4 + cilium), sox-2 (5/6)*		143-145
AWC	2	Glu	*ceh-36(4/4), sox-2 (6/7), die-1, nsy-7, hmbx-1 *****		39,87,143,146
BAG	2	Glu	*ets-5 (5/5), ceh-37 (1/1), egl-13 (6/6), egl-46 (3/6)*		136,147,148
CEP	4	DA	*ast-1 (7/7), ceh-43 (5/5)*		83,134
FLP	2	Glu	*unc-86, mec-3*	*egl-44, egl-46*	
IL1	6	Glu	*sox-2 (2/2), vab-3 (1/1)*		81
IL2	6	ACh	*unc-86 (6/6 + cilium), cfi-1 (6/6)*		58
OLL	2	Glu	*sox-2 (3/3), vab-3 (4/4)*		81
OLQ	4	Glu			
PDE	2	DA	*ceh-43 (5/5); ceh-20 (5/5)*		83,134
PHA	2	Glu	*ceh-14 (3/3 + cilium)*		39
PHB	2	Glu	*ceh-14 (3/3+ cilium)*		39
PHC	2	Glu	*ceh-14 (3/3)*		39
PLM	2	Glu	*unc-86, mec-3 (71)**	*egl-5*	42,149
PLN	2	ACh			
PQR	1	Glu	*unc-86 (2/2), egl-13 (2/2), ahr-1(1/1)*		136-138
PVD	2	Glu	*unc-86, mec-3*		
PVM	1	unknown	*unc-86, mec-3 (71)**		42
URA	4	ACh	*unc-86 (2/2), cfi-1 (2/2)*		58
URB	2	ACh	*unc-86 (1/1)*		58
URX	2	ACh	*unc-86, ahr-1 (3/3), egl-13 (4/4)*		136
URY	4	Glu	*vab-3 (3/3)*		39
38	87		33/38 sensory neuron classes		
			75/87 total sensory neurons		

Fig. 4. Terminal regulators of sensory neurons

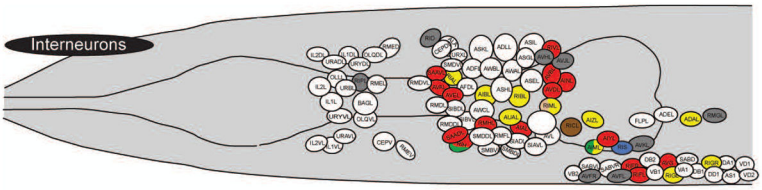
A: Schematic worm with only sensory neurons colored.

B: Tabular summary of terminal regulators of neuronal identity. “Terminal regulators” refers to the key property of these transcription factors: they are expressed in mature neurons throughout their lifetime, likely a reflection of their continuous role in maintaining the differentiated, terminal state. Early or transiently acting regulators are not shown. Only extrapharyngeal neurons are shown. Black font: most/all tested markers affected (terminal selector). Blue font: only subsets of markers affected. (x/y) indicates x markers out of y markers tested show defective expression. Homeodomain transcription factors are underlined. * indicates that in the case of the touch neurons, expression profiling has identified at least 71 *mec-3*-dependent genes. *unc-86* is known to regulate *mec-3* expression and cooperate with *mec-3* to control touch neuron-expressed genes [92]. ** indicates that the green labeled factors control only a small subset of left/right asymmetrically expressed chemoreceptors [65,93], but not bilateral identity of these neurons. *** indicates our own unpublished data. “Cilium” indicates that the respective regulators also control cilium structure, as assessed by dye-filling defects observed in the respective mutants.

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript



Inter	Number	Neurotransmitter	Inducers	Reference
ADA	2	Glu		
AIA	2	ACh	*ttx-3 (8/8)*	48,58
AIB	2	Glu	*unc-42 (4/4)*	*
AIM	2	Glu + 5HT	*unc-86 (3/3), ceh-14 (3/3)*	39,150,151
AIN	2	ACh		
AIY	2	ACh	*ttx-3, ceh-10 (42/42)*	41,48
AIZ	2	Glu	*unc-86 (3/3)*	39
AUA	2	Glu	*ceh-6 (4/4)*	39
AVA	2	ACh	*unc-3 (2/11), unc-42 (3/8), fax-1 (2/6)*	53-55
AVB	2	ACh	*unc-3 (2/3)*	38
AVD	2	ACh	*unc-3 (2/7), unc-42 (2/7), cfi-1 (2/5)*	38,53,55,56
AVE	2	ACh	*unc-3 (2/8), unc-42 (3/8), fax-1 (3/6)*	38,53-55
AVF	2	unknown		
AVG	1	ACh	*lin-11 (3/3), ast-1 (2/2)*	117,152
AVH	2	unknown		
AVJ	2	unknown		
AVK	2	unknown	*unc-42 (2/3), fax-1 (2/3)*	153
BDU	2	unknown	*pag-3 (10/10), unc-86 (10/10), ceh-14 (3/10)*	49
CAN	2	unknown MA		
DVA	1	ACh	*unc-3 (2/8)*	38
DVC	1	Glu	*ceh-14 (2/2)*	39
LUA	2	Glu		
PVC	2	ACh	*ceh-14 (2/2), cfi-1 (2/2), unc-3 (2/7)*	38,56
PVN	2	ACh	*unc-3 (2/2), ceh-14 (2/2)*	38
PVP	2	ACh	*lin-11 (3/3), unc-30 (3/3)*	38,117
PVQ	2	Glu	*pag-3* (1/1), *zag-1* (1/1)	154-156
PVR	1	Glu	*unc-86 + ceh-14 (4/4)*	39
PVT	1	unknown	*ceh-14 (5/10), lim-6 (5/10) **	
PVW	2	unknown		
RIA	2	Glu		
RIB	2	GABA(+ ACh)		
RIC	2	Octopamine		
RID	1	unknown	*lim-4 (4/6)*	30
RIF	2	ACh		
RIG	2	Glu		
RIH	1	ACh & 5HT	*unc-86 (5/5)*	38,150
RIP	2	unknown		
RIR	1	ACh		
RIS	1	GABA	*nhr-67 (5/5), lim-6 (8/8)*	30,135,157
RIV	2	ACh	*unc-42 (2/2)*	38
SAA	4	ACh	*sox-3 (3/3)*	81
SDQ	2	ACh		
42	77		26/42 interneuron classes	
			46/77 total interneurons	

Fig. 5. Terminal regulators of interneurons

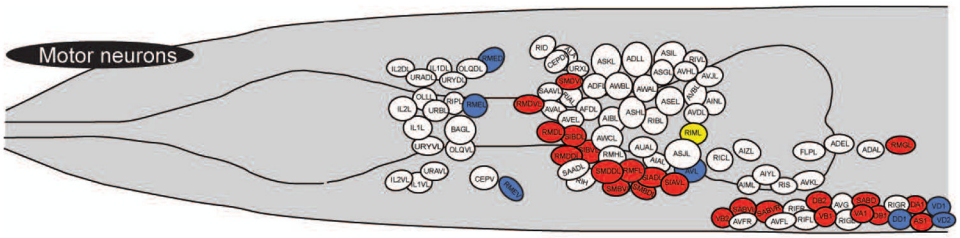
A: Schematic worm with only interneurons colored.

B: Tabular summary of terminal regulators of neuronal identity. See legend to Fig.4 for an explanation of colors and numbers. * indicates our own unpublished data. ** *lim-6* and *ceh-14* regulate, in conjunction with the *lin-14* heterochronic gene, temporally controlled but not continuously expressed genes in PVT [94].

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript



Motor	Number	Neurotransmitter	Inducers	Repressor	Reference
AS	11	ACh	*unc-3 (17/22)*		45
AVL	1	GABA	*nhr-67 (4/4), lim-6 (5/5)*		135,157
DA	9	ACh	*unc-3 (26/31)*	*unc-4*	45,61,119,158,159
DB	7	ACh	*unc-3 (27/32)*	*vab-7*	45,62,119,159
DD	6	GABA	*unc-30 (10/10)*	*irx-1*	44,160-162
DVB	1	GABA	*lim-6 (1/5)*		135,157
HSN	2	ACh & 5HT	*unc-86 (12/12), sem-4 (12/12)*		163
PDA	1	ACh	*unc-3 (5/5)*		38,164
PDB	1	ACh	*unc-3 (2/2)*		38
RIM	2	Glu + Tyr			
RMD	6	ACh	*unc-42 (3/3)*		38
RME	4	GABA	*nhr-67 (6/6), ceh-10 (4/4), tab-1 (4/4)*	*ahr-1*	78,135
RMF	2	ACh			
RMG	2	unknown			
RMH	2	ACh			
SAB	3	ACh	*unc-3 (10/10)*		119
SIA	4	ACh			
SIB	4	ACh			
SMB	4	ACh	*lim-4 (5/5)*		165
SMD	4	ACh	*unc-42 (2/2)*		38
VA	12	ACh	*unc-3 (23/26)*	*unc-4*	45,61,119,158,159
VB	11	ACh	*unc-3 (26/29)*	*ceh-12*	45,119,159,166
VC	6	ACh			
VD	13	GABA	*unc-30 (10/10)*	*unc-55, alr-1*	44,160-162,167,168
24	118		17/24 motor neuron classes		
			96/118 total motor neurons		

Fig. 6. Terminal regulators of motor neurons

A: Schematic worm with only motor neurons colored.

B: Tabular summary of terminal regulators of neuronal identity. See legend to Fig.4 for an explanation of colors and numbers.

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript

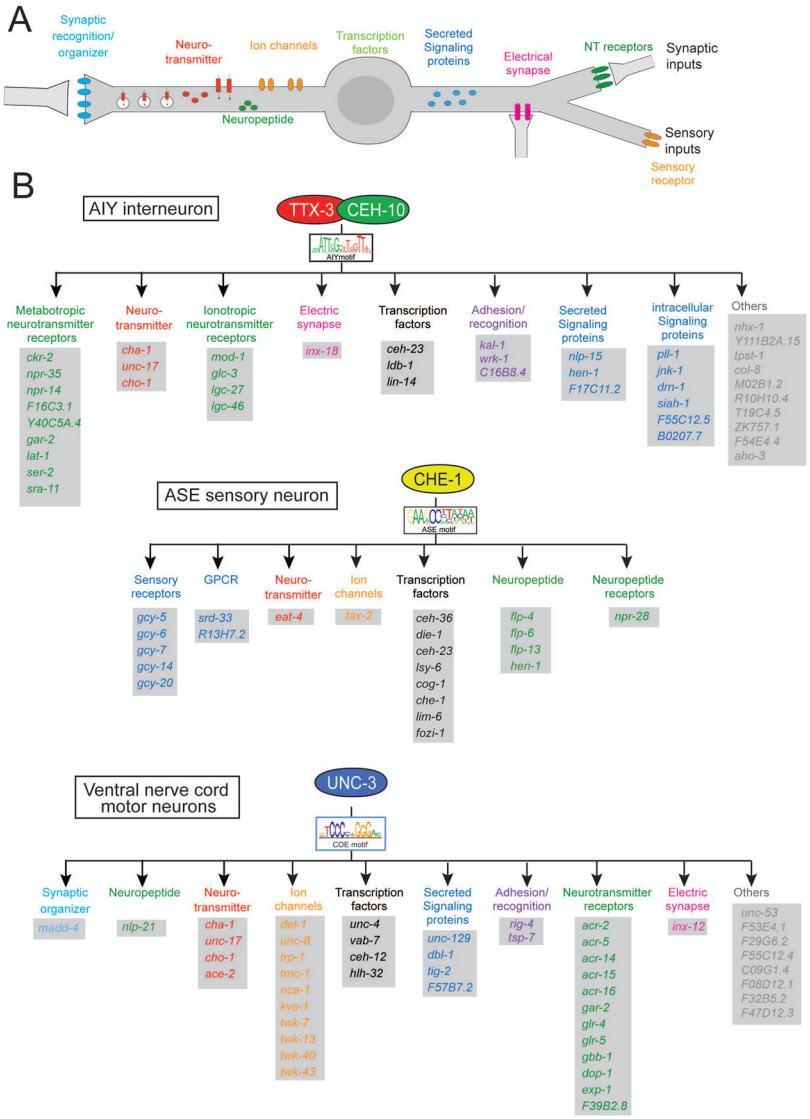


Fig. 7. Terminal selector regulons

A: Schematic illustration of terminal identity features that are controlled by terminal selector-type transcription factors. These identity features are continuously expressed throughout the life of a neuron.

B: Three examples of terminal selector regulons. All genes shown here were shown to be direct targets of the indicated terminal selectors [41,45,95]. In addition to the genes shown here, scores of additional genes have been identified as being expressed in the respective neuron types and containing binding sites for the respective regulators, but these additional genes have not been validated for terminal selector-dependence.

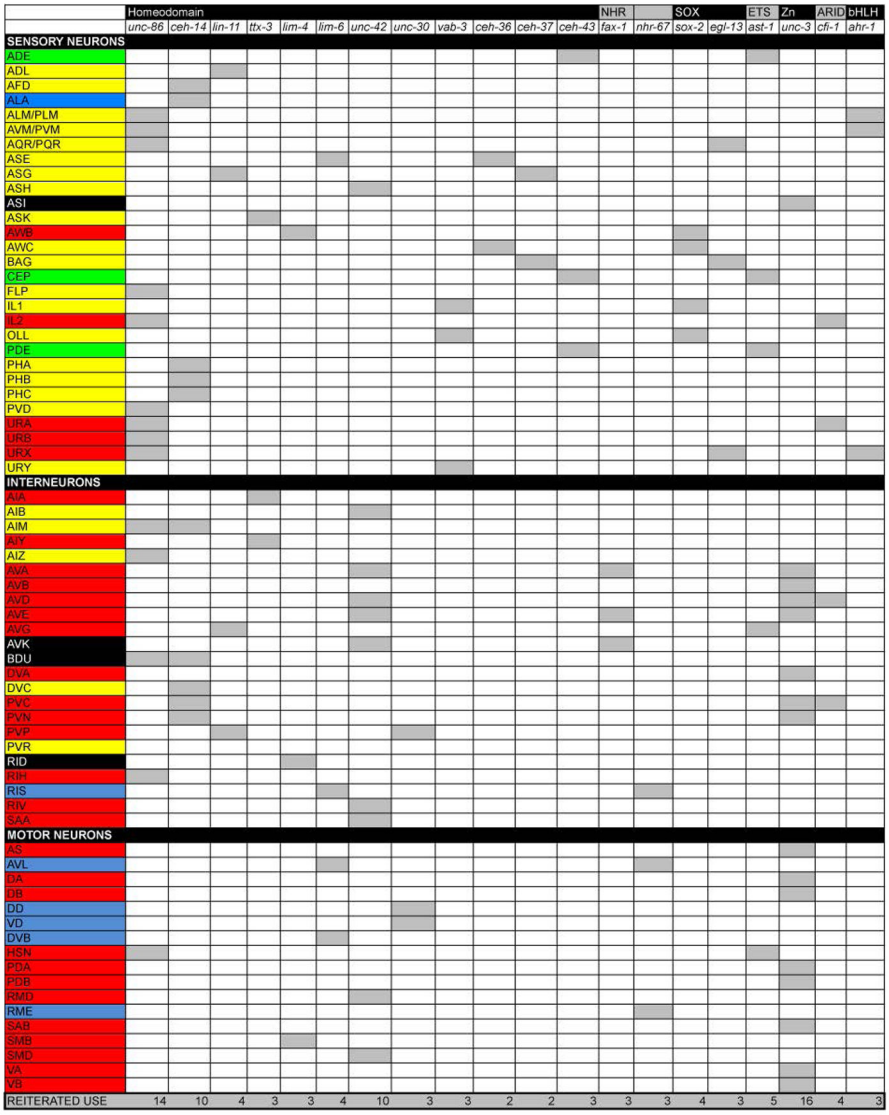


Fig. 8. Re-deployment of regulators of terminal neuronal identity in distinct neuron types

Data is extracted from Fig.4,5,6. Coloring scheme indicates neurotransmitter identity: yellow = Glu, red = ACh, green = aminergic, black = unknown. Only transcription factors that operate in >1 class are shown

HHMI Author Manuscript HHMI Author Manuscript HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript

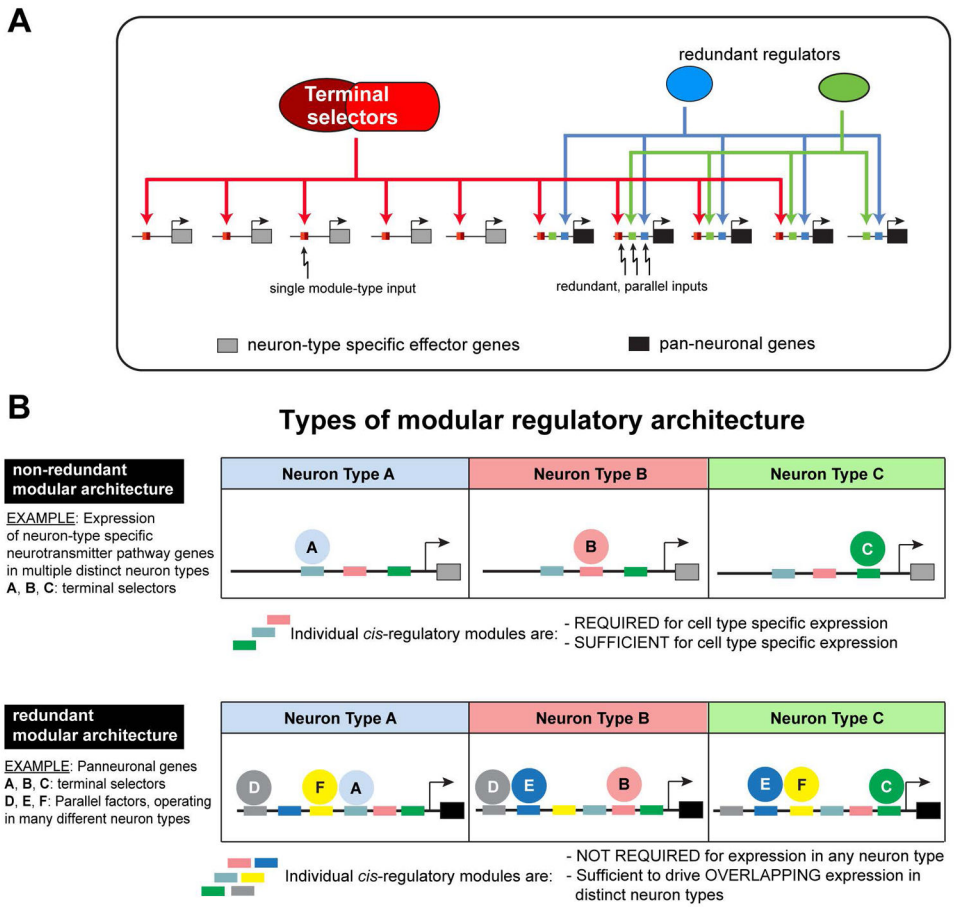


Fig. 9. Regulation of pan-neuronal identity

This figure is reproduced from [69].

HHMI Author Manuscript HHMI Author Manuscript HHMI Author Manuscript

Table 1

Functional map of the extrapharyngeal nervous system

Not every behavior associated with a given cell is shown. Usually the first identified behavior is shown. For more in-depth coverage see [98,99]

Sensory neuron	FUNCTION	Reference	Inter-neuron	FUNCTION	Reference	Motor neuron	FUNCTION	Reference
ADE	mechanosensory	100,101	ADA			AS		
ADF	dauer formation	102	AIA	processing sensory information	103	AVL	defecation	104
ADL	olfaction (repulsive)	25	AIB	reversal, pausing, integrating sensory information	103,105,106	DA	locomotion	107
AFD	Thermosensation	15	AIM	sexual attraction		DB	locomotion	107
ALA	sleep, mechanosensory	108,109	AIN			DD	locomotion	104
ALM	mechanosensory	13	AIY	locomotion, integrating sensory information	15,110	DVB	defecation	104
ALN			AIZ	locomotion, processing sensory information	15,110	HSN	egg-laying	
AQR	aggregation, O2 sensation	111,112	AUA	aggregation	111	PDA		
ASE	gustation	14	AVA	command interneurons	107,113	PDB		
ASG	gustation	14,102	AVB	command interneurons	107,113	RIM	reversal behavior, feeding circuit	105,114
ASH	mechanosensory, gustation	16	AVD	command interneurons	107,113	RMD	head movement, foraging	105,115
ASI	dauer formation, feeding state, chemosensory, thermosensory	14,17,102	AVE	command interneurons	107,113	RME	head movement	104
ASJ	dauer formation and recovery	102	AVF	egg-laying/locomotion	116	RMF		
ASK	gustation	14	AVG	pioneer/guidepost	117	RMG	aggregation	118
AVM	mechanosensory	13	AVH			RMH		
AWA	olfaction	14	AVJ			SAB	head movement	119
AWB	olfaction	14	AVK			SIA	guidepost	120
AWC	olfaction, thermosensation	14,121	BDU	mechanosensory	122	SIB	guidepost	120
BAG	CO2 sensation	123	CAN	fluid homeostasis		SMB	sinusoidal locomotion	105
CEP	mechanosensory	100,101	DVA	stretch reception	124	SMD	omega turns	105
FLP	mechanosensory	16	DVC			VA	locomotion	107
IL1	head bends, foraging	115	LUA	pausing		VB	locomotion	107
IL2	nictation behavior	125	PVC	command interneurons	107	VC	egg-laying	126
OLL	pathogen avoidance behavior	127	PVN			VD	locomotion	104

HHMI Author Manuscript HHMI Author Manuscript HHMI Author Manuscript

Sensory neuron	FUNCTION	Reference	Inter-neuron	FUNCTION	Reference	Motor neuron	FUNCTION	Reference
OLQ	mechanosensory, head bends, foraging	16,115	PVP					
PDE	mechanosensory	100,101	PVQ					
PHA	gustation, mechanosensory	122,128	PVR					
PHB	gustation, mechanosensory	122,128	PVT	pioneer/guidepost	31			
PHC	heat avoidance	122,129	PVW					
PLM	mechanosensory	13	RIA	thermosensory processing, quiescence	15,130			
PLN			RIB	reversal behavior	105,110			
PQR	aggregation, O2 sensation	111,112	RIC	feeding circuit	114			
PVD	mechanosensory, thermosensory	18	RID					
PVM	mechanosensory	13	RIF	dwelling	131			
URA			RIG					
URB			RIH	processes nose touch	132			
URX	O2 sensation	111,112	RIP	coupling touch and pharyngeal pumping	107			
URY			RIR					
			RIS	quiescence	133			
			RIV	omega turns	105			
			SAA					
			SDQ					
38			42			24		