

cnd-1/NeuroD1 Functions with the Homeobox Gene *ceh-5*/Vax2 and Hox Gene *ceh-13*/labial To Specify Aspects of RME and DD Neuron Fate in *Caenorhabditis elegans*

Wendy Aquino-Nunez,^{1,2} Zachery E. Mielko,^{1,3} Trae Dunn, Elise M. Santorella,⁴ Ciara Hosea,⁵ Lauren Leitner, Derrica McCalla, Claire Simms, Wendy M. Verola,⁶ Sharanya Vijaykumar, and Martin L. Hudson⁷

Department of Molecular and Cellular Biology, Kennesaw State University, Kennesaw, GA 30144

ORCID IDs: 0000-0001-8348-2145 (W.A.-N.); 0000-0002-2628-035X (M.L.H.)

ABSTRACT Identifying the mechanisms behind neuronal fate specification are key to understanding normal neural development in addition to neurodevelopmental disorders such as autism and schizophrenia. *In vivo* cell fate specification is difficult to study in vertebrates. However, the nematode *Caenorhabditis elegans*, with its invariant cell lineage and simple nervous system of 302 neurons, is an ideal organism to explore the earliest stages of neural development. We used a comparative transcriptome approach to examine the role of *cnd-1*/NeuroD1 in *C. elegans* nervous system development and function. This basic helix-loop-helix transcription factor is deeply conserved across phyla and plays a crucial role in cell fate specification in both the vertebrate nervous system and pancreas. We find that *cnd-1* controls expression of *ceh-5*, a Vax2-like homeobox class transcription factor, in the RME head motoneurons and PVQ tail interneurons. We also show that *cnd-1* functions redundantly with the Hox gene *ceh-13*/labial in defining the fate of DD1 and DD2 embryonic ventral nerve cord motoneurons. These data highlight the utility of comparative transcriptomes for identifying transcription factor targets and understanding gene regulatory networks.

KEYWORDS

NeuroD1
cell fate
specification
Hox genes
homeobox genes
RNA-seq

Accurate control of gene expression is fundamental for the development and function of the central nervous system (CNS). Defects in CNS gene expression underlie many neurodevelopmental disorders, indicating a critical need for further study (Basu *et al.* 2009; Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014). Gene expression is controlled by combinations of transcription factors that work in conjunction with chromatin remodeling complexes to promote or inhibit RNA polymerase access to the genome (Clapier and Cairns 2009). According to Waddington's model of cellular differentiation, cell fate progressively refines over multiple rounds of cell division, first to tissue-type progenitors, then exiting the cell cycle to take on a tissue-specific terminal fate (Waddington 1957). The transcription factors and chromatin remodeling complexes required for these narrowing rounds of fate specification are not well understood.

The basic Helix-Loop-Helix (bHLH) super family of proneural transcription factors includes Atonal, NeuroD, Neurogenin, and Achaete/Scute, and has broad roles in nervous system development (Baker and Brown 2018). This family of bHLH transcription factors

function as either homodimers or heterodimers and bind to E-box sequences of the motif CANNTG. In *Drosophila*, the bHLH family acts in neural cell fate specification and neurogenesis, with different family members having roles in external sensory organ formation, chordotonal organ development, and others (Lee 1997).

Vertebrate *neurogenic differentiation 1* (NeuroD1) is a bHLH transcription factor that has a role in the transcriptional activation of proneural genes (Wang and Baker 2015). In addition, NeuroD1 is expressed abundantly in the brain after terminal fate specification, which suggests a secondary role in nervous system homeostasis and/or neural maturation and survival (Miyata *et al.* 1999). Ectopic expression of NeuroD1 in *Xenopus* embryos can convert non-neural ectodermal cells into fully differentiated neurons, indicating the potential role of NeuroD1 as a neural differentiator factor (Lee *et al.* 1995; Lee 1997). Humans bearing homozygous NeuroD1 mutations showed severe cerebral hypoplasia and developmental delay, in addition to defects in pancreatic β -cell maturation and islet formation, demonstrating the importance of this gene in nervous system and pancreatic development (Rubio-Cabezas *et al.* 2010).

In the mouse, NeuroD1 is essential for the generation of granule cells in the hippocampus and the cerebellum (Miyata *et al.* 1999; D'Amico *et al.* 2013). Despite extensive research on the role of NeuroD1 in cell fate specification and nervous system development (Seo *et al.* 2007; Pataskar *et al.* 2016), a comprehensive list of NeuroD1 targets has not been compiled and many questions on its role in neural development remain unanswered.

The nematode *Caenorhabditis elegans*, with its invariant cell lineage and well-defined nervous system, is an excellent model to study cell lineage determination and terminal fate specification (Sulston and Horvitz 1977; Sulston *et al.* 1983). Once a neuroblast exits the cell cycle, it needs to extend growth cones and axons through the extra-cellular matrix, find its appropriate pre- and post-synaptic partners, assemble a synapse, gap junction, or neuromuscular junction, then package the various proteins required for synaptic transmission (Chisholm *et al.* 2016; Jin and Qi 2018). This leads to two major questions. First, what cascade of transcription factors is required to specify interim cell fates, prior to final specification of a neuron? Second, does a single transcription factor control the fate of a single neuron, or is terminal fate specified in a combinatorial manner, with multiple transcription factors controlling different aspects of the final cell fate? Extensive work has identified a battery of transcription factors known as “terminal selectors”, which are required for terminal fate specification in *C. elegans* neurons (Hobert 2016 and references therein). These transcription factors generally act in a combinatorial fashion to specify cell fates, although individual transcription factors may specify the fate of multiple cells that are unrelated by cell lineage, type, or circuit. Terminal selectors typically have autoregulatory properties, in that they positively regulate their own transcription to maintain neuronal identity throughout the life of an organism. In addition, they either directly or indirectly control the expression of “terminal effector” genes, which are required for that neuron’s post-mitotic function, for instance neurotransmitter biosynthesis, packaging, and release. Despite this depth of knowledge, the “proneural” transcription factors that act up-stream of terminal selector genes are not well described.

The *C. elegans* bHLH transcription factor *cnd-1* is orthologous to the human NeuroD1 gene and is one of the earliest proneural genes to be activated during *C. elegans* embryonic development (Hallam *et al.* 2000). However, the only reported defects seen in *cnd-1* loss-of-function mutants are a relatively mild back-coiler phenotype caused by misspecification of 2-3 dorsal D (DD) motorneurons required for

inhibitory GABAergic neuromuscular innervation, in addition to axon guidance and synapse remodeling defects in the remaining D neurons (Hallam *et al.* 2000). To gain a better understanding of *CND-1*'s role during *C. elegans* neural development, we performed an RNA-seq assay comparing embryonic wild type and *cnd-1(ju29)* mutant transcriptomes. We find that *CND-1* positively regulates the expression of homeobox transcription factor *ceh-5/Vax2* in the head RME and tail PVQ neurons. We also confirm that *CND-1* is required for the generation of *cnd-1* expressing cells during ventral nerve cord fate specification. Finally, we show that *cnd-1* functions in parallel with the Hox gene *ceh-13/labial* to specify a subset of embryonic DD class ventral nerve cord motorneuron fates.

MATERIALS AND METHODS

Strains and maintenance

C. elegans strains were grown on nematode growth medium plates (NGM Lite) at 20° according to Brenner (1974). Bristol N2 strain was used as wild type and all analyses were conducted at 20°. The following alleles were used in this study: LGIII *cnd-1(ju29)*, *cnd-1(gk718)* and *ceh-13(sw1)/qC1 [dpy-19(e1259) glp-1(q339)]*. Integrated transgenes used were *juls76 [unc-25p::GFP + lin-15(+)]*, *kyIs39 [sra-6p::GFP + lin-15(+)]*, *lhIs5 [unc-25p::mCherry]*, *otIs356 [rab-3p(prom1)::2xNLS::TagRFP]*, *pkIs586 [gpa-9p::GFP + dpy-20(+)]*, and *stIs10055 [cnd-1p::his-24::mCherry + unc-119(ed3)]*. Extra-chromosomal arrays used in this study were *leEx2489 [ceh-5p::GFP + unc-119(+)]* and *dbEx724 [flp-6p::tax-2(cDNA)::SL2::GFP + lin-15(+)]*. The *cnd-1(gk718)* allele was identified by the *C. elegans* deletion mutant consortium (2012). All mutants were outcrossed at least twice prior to analysis. Table S1 shows details of strains generated during the course of this study including strain numbers and sources.

A *cnd-1(gk718) ceh-13(sw1)/qC1* line was built by crossing *gk718/+* males into *sw1/qC1*, keeping lines that did not give rise to the *qC1* dumpy/sterile phenotype (genotype *gk718 +/- sw1*), selecting for *gk718* homozygous animals (uncoordinated phenotype) then screening for the embryonic lethal *sw1* phenotype (parent genotype *gk718 sw1/gk718 +*). Two recombinants were identified from 200 *gk718* animals screened, consistent with the 1 map unit distance between *cnd-1* and *ceh-13* on LG III. These lines were rebalanced over the *qC1* chromosomal inversion prior to further characterization.

RNA extraction

Embryos were isolated from gravid worms grown in liquid culture as described previously (Hudson *et al.* 2006). Total RNA was extracted using RiboZol (AMRESCO) and followed the vendor’s protocol except that embryos were frozen in liquid nitrogen prior to grinding. Embryonic tissue was added to 1 mL of RiboZol and 500 μ L aliquoted into two 5 PRIME Phase Lock Gel Tubes. At the isopropanol stage, 3 μ L of 20ng/ml Glycogen was added to improve RNA pellet visualization. RNA quality control was assayed by measuring the A260/A280 ratio using a Thermo Scientific NanoDrop and via an Agilent Bioanalyzer. All samples used for RNA-Seq had an RNA integrity number (RIN) above 9.2. Triplicate *cnd-1(ju29)* and N2 wild type samples were sent to the University of Kansas Genome Sequencing Core for RNA-seq library construction (Illumina TruSeq v2), bar-coded, pooled, and sequenced in a single lane on an Illumina HiSeq 2500 system (high output, single read 100bp sequencing).

RNA-Seq expression analysis pipeline

Gene expression abundance were obtained using previous published RNA-seq analysis workflows. Briefly, read quality was determined

Copyright © 2020 Aquino-Nunez *et al.*

doi: <https://doi.org/10.1534/g3.120.401515>

Manuscript received April 27, 2020; accepted for publication June 26, 2020; published Early Online June 29, 2020.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplemental material available at figshare: <https://doi.org/10.25387/g3.12567827>.

¹These authors contributed equally to this work.

²Present Addresses: Molecular Biosciences Graduate Program, University of Kansas, Lawrence, KS 66045

³Graduate Program in Genetics and Genomics, Duke University Medical Center, Durham, NC 27710

⁴Biomedical Sciences Graduate Program, University of Connecticut Health Sciences Center, Farmington, CT 06030

⁵Development, Disease Models & Therapeutics Graduate Program, Baylor College of Medicine, Houston, TX 77030

⁶Medical Program, Trinity School of Medicine, VC0100, St. Vincent

⁷Corresponding author: Department of Molecular and Cellular Biology, Kennesaw State University, Building 12, Room SC507, 370 Paulding Avenue, Kennesaw, GA 30144. E-mail: mhudson28@kennesaw.edu

using FastQC 0.11.5 (Andrews 2010). Reads were aligned back to the *C. elegans* reference genome (wbcel235) using Hisat2-2.1.0 (Kim *et al.* 2015). Samtools 1.5 (htslib 1.4.1) was used for file conversion and sorting (Li *et al.* 2009). Stringtie 1.3.3.b was used to assemble and quantify transcripts (Pertea *et al.* 2016). The script prepDE.py was used to generate gene and transcript count matrix files that were compatible with DESeq2 1.16.1 (R3.4.1), which was used to identify differentially expressed genes (Love *et al.* 2014). All programs were run locally.

Identification of differential isoform usage

This was performed using DEXSeq (Anders *et al.* 2017). Raw RNA-seq reads were mapped with HiSat2 to the *C. elegans* annotated reference genome obtained from Ensembl. A python script provided by the DEXSeq package was used to combine all isoforms of a gene into one global schematic representation with marked intron/exon boundaries. The analysis was performed following the developers' recommendations with the following inclusion "- s" for no strand specific reads.

Characterization of the *ken2* deletion allele

2.0X Taq RED Master Mix Kit (Apex Bio Research Products) was used to amplify the 3' end of the *ptrn-1* gene. Reverse primers at approximately 1kb intervals were used in combination with a

single forward primer within the *ptrn-1* coding region to amplify the *ptrn-1* 3' untranslated region (UTR). Primer R9, located 10kb downstream of the F1 forward primer were the only pair which amplified *cnd-1(ju29)* genomic DNA, and gave a 3kb amplicon. Sanger sequencing was used to confirm the *ken2* breakpoints, which corresponds to a 6908bp deletion/80bp insertion. Coordinates of the *ken2* breakpoints along with insertion and flanking sequences are; LG X

```
AGAGATACACATGTTTTGGTGGCTTTGTAGAAACCAG-  
TACACGCGCATTTTCACCTACTTTTTT-  
TATTTTTTCCGTTTCTTTCTGTTTCTAATTTTGCAGATT  
(17,022,165)/
```

```
CTAATTTTGCAGATTCCGGTGTCTCCGAGGTTTTT-  
TAAATCGGTGGGCAGGTGGAAATATTTTGTCA-  
TAGTTTTTTCGAAG / (17,015,257)
```

```
TATCAGTTTGTCCATAAGTTTTTGTAC-  
TATTTTTTTTTTGAATAAATTTATTTCTCTCAAGCGA-  
CAAGTAGTACTATTCACACAAGTATTCACCATTAGTGT.
```

Note that the 3' breakpoint lies in a transposon sequence and is difficult to identify via BLAST search. The *ken2* deletion removes all of *F35B3.1*, *F35B3.10*, *F35B3.4*, and around 250bp of the *ptrn-1* 3' untranslated region. Primer sequences used for PCR characterization of the *ken2* allele were as follows and show expected wild type amplicon size:

ptrn-1F1 (gtgaccaaatcaaccgtg); *ptrn-1R1* (ttgcctcagtcatttgg, 593bp); *ptrn-1R2* (ctttaaggaaggcatgggatg, 1164bp); *ptrn-1R4* (gtcaaggcatcaagtggtag, 2245bp); *ptrn-1R5* (ttgaataatgcctcttaagtgaatg, 3670bp); *ptrn-1R6* (agtaacttctgagccac, 4717bp); *ptrn-1R7* (gtggaacctgaagtgaataatgg, 6116bp); *ptrn-1R8* (caaaacctctgccacg, 7119bp); *ptrn-1R9* (gaagtagtagatgacgatgcatg, 9896bp).

Plasmid constructs

Plasmids were constructed using Gateway Technology (Invitrogen). Target gene promoters were PCR amplified from genomic DNA using Phusion polymerase (Thermo Scientific), gel purified, A-overhangs added by incubating with Taq RED Master Mix Kit (Apex Bio Research Products), cloned into the pCR8/GW/TOPO vector (Thermo Scientific)

following the manufacturer's specifications then transformed into NEB 5-alpha Competent *E. coli* (New England BioLabs). Plasmid digestion was used to confirm the forward orientation of the amplicon. Promoters were recombined into the promoterless Gateway GFP expression vector pCZGY32 vector (a kind gift from Yishi Jin, UC San Diego) using Gateway LR Clonase II Enzyme Mix (Invitrogen) following the vendor's instructions and transformed into NEB 5-alpha Competent *E. coli*. Plasmid digestion was used to confirm identity. Constructs were microinjected into the germlines of either wild type or *stIs10055[cnd-1p::his-24::mCherry unc-119(+)]* worms at 20ng/μL in conjunction with either *unc-122::GFP* (coelomocyte::GFP) or *ttx-3p::RFP* co-injection plasmids and pBlueScript II to a final concentration of 50ng/μL. F1 hermaphrodites that expressed the co-injection marker were single plated and screened to obtain stable lines.

cDNA synthesis and RT-qPCR

Total RNA from mixed staged embryos was treated with DNAase (New England Biolabs) and cleaned using the RNA Clean & Concentrator kit (Zymo Research). First-strand synthesis was done using the RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific). *cdc-42* transcript levels were used to normalize for differences in input cDNA. Three or four biological samples, in triplicate, were run on a LightCycler 480 real-time PCR system and relative expression ratios were calculated according to Pfaffl (2001).

Dye-filling assay

Animals were washed off plates, washed three times with M9 buffer then incubated for one hour in 100 ng/μl Vybrant DiD cell-labeling solution (Invitrogen) as described previously (Schultz and Gumieny 2012). After destaining for one hour on NGMLite plates seeded with OP50 *E. coli*, animals were imaged by confocal microscopy at 40x magnification for DiD uptake (647nm excitation) in the amphid neurons and GFP expression (488nm excitation).

Microscopy

Well-fed worms grown under standard conditions were used for expression pattern characterization. Images were captured on either a Zeiss LSM 700 confocal microscope, Zeiss Axiovision compound microscope, or Olympus BX61 compound microscope. Expression patterns in L1 larvae were imaged within 1 hr of hatching. For quantitative imaging of *cnd-1p::his-24::mCherry* expression in wild type and *cnd-1* mutant backgrounds, 2-4 cell embryos were isolated from gravid hermaphrodites and mounted on a bead pad (Murray *et al.* 2008). After 6-7 hr, when the embryos reached comma stage, a single image stack (to eliminate possible photobleaching) was captured for each embryo using a Zeiss LSM 700 confocal microscope at 40x magnification under identical settings, ensuring no detector saturation. Images were processed using Fiji (Schindelin *et al.* 2012) using a Z-project - Sum Slices workflow which rendered the summated stacks as 32-bit images. Pixel values and counts for the whole image (1024 × 1024 pixels) were obtained using the Analyze - Histogram function (using the pixel value range) then processed in Microsoft Excel (sum [pixel value x pixel count]).

Individual DD neuron identities was inferred by imaging *unc-25p::GFP* in L1 larvae then measuring the nose-to-neuron, neuron-to-neuron, or neuron-to-tail distances along the anterior-posterior body axis using the segmented line function in Fiji. DD neuron identity in *cnd-1(gk718)* and *ceh-13(sw1)* mutant combinations were mapped to the closest relative location in wild type L1 larvae. Total ventral

nerve cord cells were determined by double-labeling with *juIs76* [*unc-25p::GFP*] and *otIs356* [*rab-3p(prom1)::2xNLS::TagRFP*] then counting from DD1 or the posterior end of the terminal pharyngeal bulb (if DD1 was absent) to the anus. In wild type, this included most of the 22 DA, DB and DD class motorneurons, plus some cells in the retrovesicular ganglion and tail region.

Neuroanatomy

Cell identity was confirmed by crossing transgenic arrays into previously characterized strains then imaging as above. Table S1 shows the strains used for cell identification.

Statistical analysis

Student *t*-tests were performed in Microsoft Excel, Mann-Whitney tests were performed in R, and two-tailed Fisher Exact tests were performed using GraphPad QuickCalcs. Graphs were generated in Excel or SAS. RNA-seq significance values reported in the text are not adjusted for false discovery. Benjamini-Hochberg corrections for multiple comparisons are available in the supplemental data. Bonferroni corrections were applied to other data where appropriate.

Data availability

All reagents are available on request. Table S1 shows strains generated in this study. Tables S2 and S3 show lists of significantly down-regulated and up-regulated genes in the *cmd-1(ju29)* comparative transcriptome. Figure S1 shows a volcano plot summarizing the *cmd-1(ju29)* comparative transcriptome data. Figure S2 shows DEXseq hits identified in this work along with any expression validations. Figure S3 compares predicted CND-1 wild type and *ju29* mutant protein sequences. Figure S4 shows *cmd-1p::his-24::mCherry* and *unc-25p::GFP* co-localization. Raw and processed transcriptome files generated in this study are publicly available via the Gene Expression Omnibus (Edgar *et al.* 2002), accession number GSE125051 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE125051>). Supplemental material available at figshare: <https://doi.org/10.25387/g3.12567827>.

RESULTS

cmd-1 controls the expression of multiple genes During embryogenesis

Previous studies demonstrated that the proneural transcription factor *cmd-1* is active early in embryogenesis, with expression first being seen at the 14-cell stage and persisting until just prior to hatching (Hallam *et al.* 2000; Murray *et al.* 2012). *cmd-1* reporter gene expression decreases rapidly beyond the L1 stage but persists at low levels in adult head and ventral neurons (Kratsios *et al.* 2011). To gain a better understanding of how CND-1 controls gene expression during early nervous system development, we performed RNA-seq on RNA isolated from three samples each of N2 wild type and *cmd-1(ju29)* mutant mixed stage embryos and analyzed the data using the DSeq2 package (Love *et al.* 2014). *cmd-1(ju29)* is a G-to-A transition in the splice acceptor of intron 2 and behaves as a strong loss-of-function recessive allele (Hallam *et al.* 2000). *cmd-1(ju29)* mutant embryos show 105 genes with significantly lower transcript levels ($P < 0.05$) when compared to wild type (Table S2) and 46 genes with significant higher transcript levels (Table S3, Figure S1). Table 1 shows the top 40 most significant hits sorted by up- vs. down-regulation and *p*-value. Surprisingly, only a single transcription factor gene, *ceh-5*, was identified in the down-regulated dataset, whereas three transcription factors genes (*nhr-68*, *nhr-77*, and *cmd-1* itself) were found in the up-regulated dataset. This suggests that *cmd-1* functions close to

the end of a transcriptional regulatory cascade during *C. elegans* embryogenesis. This is in contrast to *ngn-1*/neurogenin, which controls expression of at least eight downstream transcription factors (Christensen *et al.* 2020).

We also analyzed our data using DEXseq, a variant of the DSeq2 workflow (Anders *et al.* 2017). This analysis compares data sets by aligning individual sequence blocks (exons, alternative transcriptional start sites, and alternative splice sites) and is a sensitive way to identify splice or transcriptional variants between two datasets. Using this approach, *aak-2*, *srw-85*, and *ptrn-1* were found to have at least one significantly different transcript block in *cmd-1(ju29)* mutants when compared to wild type (Figure S2). *aak-2* belongs to the AMP-activated protein kinase (AMPK) family and has roles in DAF-2-mediated insulin signaling, lifespan, and temperature-dependent dauer larva formation (Apfeld *et al.* 2004; Hardie 2014). In *cmd-1* mutants, the first exon of an internally transcribed *aak-2c* variant is expressed at a significantly higher level ($P < 0.05$), suggesting that CND-1 may repress this internally transcribed variant in wild type animals (Figure S2A).

DEXseq identification of *ptrn-1* transcript differences were resolved by visual inspection of the *ptrn-1* genomic locus using Integrated Genome Viewer (Robinson *et al.* 2011; Thorvaldsdottir *et al.* 2013). This revealed loss of all gene transcription in the 7-8kb region immediately downstream of *ptrn-1*. Genomic PCR coupled with Sanger sequencing confirmed this to be a novel 6906bp deletion/80bp insertion allele that removes most of the *ptrn-1* 3' UTR including the poly-adenylation signal, along with three downstream genes (F35B3.1, F35B3.4 and F35B3.10), of which F35B3.1 and F35B3.4 were also differentially expressed in our transcriptome (Figure S2I-L and Table S2). F35B3.10 codes for a predicted snRNA whose transcript was not represented in either wild type or *cmd-1(ju29)* datasets. *ptrn-1* codes for a known neuronal microtubule stabilizing protein (Chuang *et al.* 2014; Marcette *et al.* 2014; Richardson *et al.* 2014), so it is possible that this *ptrn-1(ken2)* deletion may enhance the *cmd-1(ju29)* uncoordinated phenotype and be selected for during out-cross. To control for this possibility, we performed validation assays using the *cmd-1(gk718)* mutation, which is a large deletion allele, predicted to be a null mutant, and was shown by genomic PCR not to contain the *ken2* deletion (Figure S2K).

cmd-1 controls ceh-5 expression in a subset of neurons

Our comparative transcriptome showed that *ceh-5* was significantly down-regulated in *cmd-1(ju29)* mutants when compared to wild type (Table 1). *ceh-5* is the *C. elegans* ortholog of the mammalian transcription factor ventral anterior homeobox 2 gene (*Vax2*), which is required for correct dorsoventral patterning of the eye (Takeuchi *et al.* 2003; Liu *et al.* 2008; Alfano *et al.* 2011). A sequence search across the *ceh-5* locus found a single candidate *cmd-1* CATATG E-box binding site about 50bp 5' to the *ceh-5* translational start site (Figure 1A). To better understand the role of *cmd-1* in controlling *ceh-5* expression, we used a *ceh-5p::GFP* reporter gene to compare expression patterns in wild type and *cmd-1* mutants (Figure 1, Table 2) (Reece-Hoyes *et al.* 2007). In wild type L1 larvae, *ceh-5p::GFP* showed robust expression in head muscles, a subset of head neurons (including the RME neurons), and five or six cells in the tail including the PVQL/R neurons (Figure 1B). Weak *ceh-5p::GFP* expression was also seen in the coelomocytes (asterisks, Figure 1B) and the pharyngeal terminal bulb. In *cmd-1(gk718)* mutants, *ceh-5p::GFP* expression was lost from many head neurons including the RMEs, and also the tail neurons (Figure 1C), but was retained in some head muscles and also the coelomocytes. We used quantitative PCR to further validate *ceh-5* transcript levels; these were significantly lower in

■ **Table 1** Top 40 differentially expressed genes in the *cnd-1(ju29)* RNA-seq dataset based on P-value. (A) Down-regulated genes; (B) up-regulated genes. Gene ID, gene identity; gene name, commonly used gene name/cosmid name; base mean, mean of normalized counts for that gene; log2 fold-change, log2 change in gene expression level when compared to wild type; P-value, significance; P-adj, significance with Benjamini-Hochberg adjustment for false discovery rate

A. <i>cnd-1(ju29)</i> transcriptome down-regulated genes (most significant p - value)						
Gene_ID	Gene Name	Base mean	log2 fold change	P-value	P-adj	
WBGene00018031	<i>F35B3.4</i>	402	-1.7	6.5E-22	1.76E-17	
WBGene00005832	<i>srw-85</i>	428	-1.2	3.7E-20	4.96E-16	
WBGene00014955	<i>Y102A5C.6</i>	158	-1.0	2.6E-10	1.19E-06	
WBGene00010212	<i>fbxa-192</i>	262	-0.9	2.6E-08	7.00E-05	
WBGene00044213	<i>Y102A5C.36</i>	131	-0.9	1.2E-07	2.96E-04	
WBGene00014954	<i>Y102A5C.5</i>	66	-0.9	7.1E-09	2.40E-05	
WBGene00007454	<i>C08F11.7</i>	54	-0.9	3.6E-10	1.38E-06	
WBGene00015990	<i>C18H2.3</i>	120	-0.9	1.3E-16	1.15E-12	
WBGene00014454	<i>MTCE.7</i>	1917	-0.7	8.3E-06	0.016	
WBGene00010958	<i>ndfl-4</i>	70	-0.7	8.4E-06	0.016	
WBGene00010209	<i>fbxa-191</i>	31	-0.7	1.1E-08	3.38E-05	
WBGene00007201	<i>exos-4.1</i>	947	-0.7	3.1E-05	0.056	
WBGene00014472	<i>MTCE.33</i>	2190	-0.6	6.5E-04	0.93	
WBGene00016953	<i>C55C3.3</i>	65	-0.5	1.1E-03	1.00	
WBGene00015044	<i>cyp-34A9</i>	82	-0.5	1.0E-03	1.00	
WBGene00016506	<i>abhd-5.1</i>	95	-0.5	1.9E-03	1.00	
WBGene00006650	<i>tts-1</i>	1948	-0.5	2.8E-04	0.42	
WBGene00000754	<i>col-181</i>	39	-0.5	3.1E-03	1.00	
WBGene00000430	<i>ceh-5</i>	332	-0.5	6.2E-03	1.00	
WBGene00014672	<i>C08F11.6</i>	27	-0.4	7.8E-04	1.00	
WBGene00077585	<i>T01G5.8</i>	7	-0.3	5.7E-05	0.096	
WBGene00202498	<i>Y60C6A.2</i>	35	-0.3	2.0E-03	1.00	
WBGene00022013	<i>Y60C6A.1</i>	59	-0.3	2.1E-03	1.00	
WBGene00002013	<i>hsp-12.6</i>	164	-0.2	4.4E-03	1.00	
WBGene00012790	<i>Y43D4A.4</i>	6	-0.2	1.4E-04	0.22	
WBGene00015549	<i>C06G3.3</i>	24	-0.2	2.9E-03	1.00	
WBGene00008396	<i>D1086.9</i>	27	-0.2	7.2E-03	1.00	
WBGene00017371	<i>sre-39</i>	4	-0.2	2.2E-03	1.00	
WBGene00020178	<i>T02H6.8</i>	4	-0.1	3.5E-03	1.00	
WBGene00045311	<i>Y57G11C.57</i>	3	-0.1	5.7E-03	1.00	
WBGene00044293	<i>K08D12.7</i>	3	-0.1	6.2E-03	1.00	
WBGene00044390	<i>ZK177.11</i>	3	-0.1	7.0E-03	1.00	
WBGene00011429	<i>T04C12.7</i>	7	-0.1	7.8E-03	1.00	
B. <i>cnd-1(ju29)</i> transcriptome up-regulated genes (most significant p - value)						
gene_ID	name	base mean	log2 fold change	P-value	P-adj	
WBGene00018031	<i>lgc-34</i>	1214	1.3	4.30E-16	2.91E-12	
WBGene00005832	<i>F35B3.3</i>	41	0.9	2.37E-11	1.28E-07	
WBGene00014955	<i>ctc-3</i>	66934	0.7	6.72E-06	0.015	
WBGene00010212	<i>ZC21.10</i>	882	0.5	0.003	1.00	
WBGene00044213	<i>cnd-1</i>	958	0.5	0.004	1.00	
WBGene00008677	<i>col-171</i>	8	0.1	0.007	1.00	
WBGene00004567	<i>C32E8.4</i>	3	0.1	0.007	1.00	

both *cnd-1(ju29)* and *cnd-1(gk718)* RNA samples, with *gk718* showing lower *ceh-5* transcript levels when compared to *ju29*, suggesting that CND-1(*ju29*) may retain some function (Figure 1D). Together, these data indicate that CND-1 is necessary for *ceh-5p::GFP* expression in a subset of cells including the RME head and PVQ tail neurons, although it is not known if these cells are lost, changing fate, or are merely losing *ceh-5* reporter gene expression.

CND-1 controls the fate of some *cnd-1*-expressing cells during embryonic nervous system development

The *cnd-1* locus is comprised of three exons spanning a 1.5kb region of chromosome III (Figure 2A). As mentioned previously, the *cnd-1(ju29)*

allele used in our transcriptome analysis is a G-to-A transition in the splice acceptor of intron 2 and was predicted to force a splice onto a non-canonical splice acceptor leading to a frame shift (Hallam *et al.* 2000). When *cnd-1* RNA-seq data were viewed using Integrated Genome Viewer, we confirmed that *cnd-1(ju29)* transcripts showed the G-to-A transition at the *ju29* base change and also a 1bp shift in the splice acceptor (Figure 2B, red column at the start of exon 3; Figure 2C; Figure S3). In addition, around 20% of transcripts show inclusion of intron 2, presumably because the *ju29* mutation creates a weak splice acceptor. Figure 2B shows representative read depth across the *cnd-1* locus. In terms of raw reads and when normalized to Fragments per Kilobase Million (FPKM), *cnd-1* transcript levels are

almost twice as high in *cnd-1(ju29)* compared to wild type, suggesting that CND-1 may be partially responsible for regulating its own transcript levels via transcriptional repression.

To further explore these data, we performed quantitative confocal microscopy on wild type and *cnd-1* mutant comma stage embryos carrying an integrated *cnd-1p::his-24::mCherry* transcriptional reporter gene (Murray *et al.* 2012). The comma stage is easily identified during embryonic development and provides a defined time point to quantitatively compare mCherry expression levels. Contrary to what RNA-seq revealed, we found that *cnd-1p::mCherry* reporter gene expression was significantly lower in *cnd-1* mutants when compared to wild type (Figure 2D, $P < 0.005$, $n = 34$, 25, and 30 embryos analyzed for wild type, *cnd-1(ju29)* and *cnd-1(gk718)* mutants respectively). The reason for this discrepancy was unclear but may be assay dependent. For instance, inclusion of intron 2 within the *cnd-1(ju29)* mRNA transcript may increase its representation within the RNA-seq dataset, causing it to appear up-regulated. Alternatively, it may increase the stability of *cnd-1(ju29)* transcripts, causing an apparent up-regulation of RNA levels.

To clarify the role of *cnd-1* during nervous system development, we also analyzed *cnd-1* reporter gene expression in L1 larvae. Only three types of ventral cord motorneuron (cholinergic DA and DB, and GABAergic DD class) are born during embryogenesis and can be easily assayed in young L1 larvae using genetically encoded reporter genes. Previous data showed that *cnd-1(ju29)* mutants exhibit variable loss of all three embryonic motor neuron types (Hallam *et al.* 2000). We examined *cnd-1p::mCherry* expression in early L1 larvae, within an hour of hatching, counting nuclei in the retrovesicular ganglion and also the ventral nerve cord (Figure 3). The retrovesicular ganglion is a linear cluster of cells located on the ventral midline of the worm immediately posterior to the pharynx, and contains the anterior most DA, DB and DD motorneurons (DA1, DB1, DB2 and DD1), along with eight additional cells. Wild type animals showed an average of nine *cnd-1p::mCherry* nuclei in the retrovesicular ganglion and 10 in the ventral nerve cord ($n = 26$). As there are 22 motorneurons in early L1 larvae, this suggests that only a subset

express the *cnd-1* reporter gene. Based on their location along the ventral nerve cord, co-labeling with *unc-25p::GFP* (a known DD neuron marker), and by corroborating against single cell RNA-seq expression data (Packer *et al.* 2019), we tentatively conclude that *cnd-1p::mCherry* is expressed in DA1-5, DB1, DB3, and DD1-6 (Table 3 and Figure S4). In *cnd-1(ju29)* mutants, the average number of *cnd-1*-positive cells in the retrovesicular ganglion increased to 10 ($P < 0.005$), although the number of cells in the ventral nerve cord dropped dramatically to around six ($n = 19$, $P < 0.001$). In *cnd-1(gk718)* mutants, the number of *cnd-1*-positive cells dropped to an average of seven in the retrovesicular ganglion and four in the ventral nerve cord ($P < 0.0005$). When comparing *cnd-1p::mCherry* expression between the *cnd-1(ju29)* and *cnd-1(gk718)* backgrounds, both retrovesicular ganglion and ventral nerve cord cell counts are significantly different ($P < 0.0005$ for retrovesicular ganglion cells, and $P < 0.005$ for ventral nerve cord cells). Based on these reporter gene studies, we conclude that CND-1 is required for the fate specification of a subset of embryonic ventral nerve cord neurons, confirming data reported by Hallam *et al.* (2000). In addition, the above data suggests that the CND-1(*ju29*) protein retains some activity, or in some contexts behaves in a neomorphic manner to affect the developmental outcome of CND-1-dependent cell fates (Figure S4). For this reason, all remaining analyses were performed using the *cnd-1(gk718)* allele.

***cnd-1* and *ceh-13* are co-expressed in a subset of ventral cord motorneurons**

Previous work using RNA-seq analysis of FACS-isolated *cnd-1p::mCherry*-positive cells showed that Hox gene *ceh-13*/labial transcripts were enriched in *cnd-1*-expressing cells when compared to negative controls (Burdick *et al.* 2016). The *ceh-13* locus has multiple consensus CND-1 E-box binding sites (Figure 4A), raising the possibility that *cnd-1* may control aspects of *ceh-13* transcription. In addition, *ceh-13* mutants exhibit loss of ventral cord motorneurons similar to that seen in *cnd-1* mutants (Stefanakis *et al.* 2015). This led us to investigate the relationship between *cnd-1* and *ceh-13* in controlling embryonic motorneuron cell fate specification.

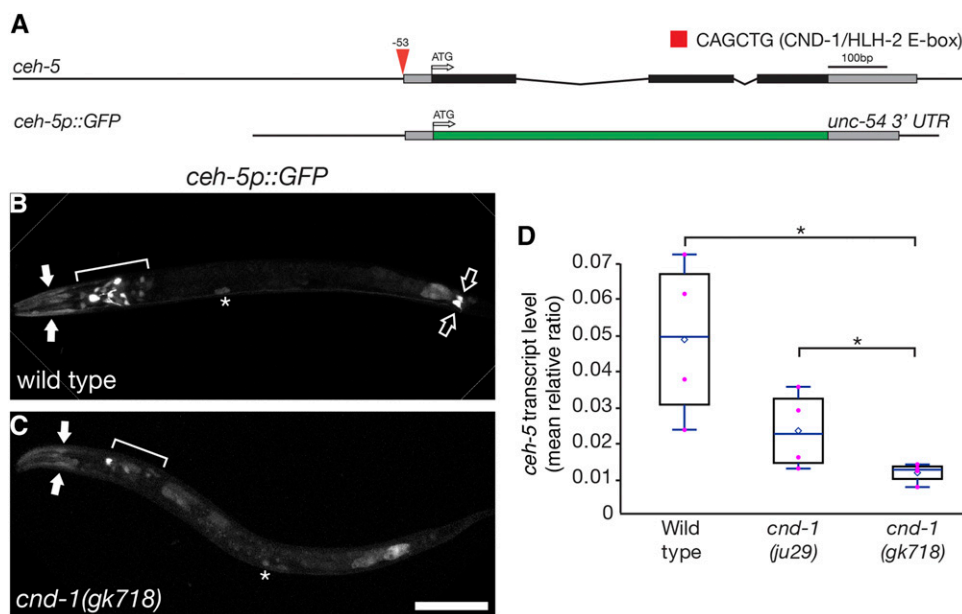


Figure 1 *cnd-1* controls *ceh-5* expression in a subset of neurons. (A) Schematic of the *ceh-5* genomic region showing predicted CND-1/HLH-2 binding site (Grove *et al.* 2009), and structure of the *ceh-5* reporter gene used in this study. (B, C) *ceh-5p::GFP* reporter gene expression in wild type (B) and *cnd-1(gk718)* mutants (C). Filled arrows, head muscles; open arrows, PVQL/R neurons; asterisks, coelomocytes. Bracketed regions show RME plus other head neurons. Scale bar = 25 μ m. (D) Box and whisker plot showing average quantitative RT-PCR levels of *ceh-5* mRNA transcript in wild type, *cnd-1(ju29)* and *cnd-1(gk718)* mutants. Open diamond, average; box shows median, first, and third quartiles. Whiskers show data extremes in $1.5 \times$ interquartile range. Data are relative to *cdc-42* mRNA. * $P < 0.025$, Student's t-test with Bonferroni correction for multiple comparisons.

Table 2 Summary of *ceh-5p::GFP* expression in L1 larvae. (A) Average number of head neurons, head muscle bundles, and tail neurons observed in wild type and *cnd-1(gk718)* mutants respectively (+/- standard error of the mean). ** $P < 0.01$, Mann-Whitney *U*-test with continuity correction. (B) Percentage of animals showing expression in other tissues. Note that the GFP reporter stain used in this assay was an extra-chromosomal array and showed some expression variability between animals. ** $P < 0.01$, Fisher Exact test

A. <i>ceh-5p::GFP</i> expression in neurons and head muscles			
Strain	# head neurons	# head muscles	# tail neurons
wild type (n = 11)	12.5 (1.6)	4.0 (0)	2.2 (0.3)
<i>cnd-1(gk718)</i> (n = 5)	4.8 (0.8)**	2.2 (0.7)**	0.0 (0)**
B. <i>ceh-5p::GFP</i> expression in other tissues			
Strain	pharynx	coelomocytes	gut
wild type (n = 11)	100%	73%	100%
<i>cnd-1(gk718)</i> (n = 5)	20%**	60%	100%

We examined *ceh-13p::GFP* and *cnd-1p::mCherry* expression in the retrovesicular ganglion and ventral nerve cords of wild type, *cnd-1(gk718)*, *ceh-13(sw1)/qC1* heterozygotes, and *ceh-13(sw1)* homozygous mutant L1 larvae (Figure 4B-G). *ceh-13(sw1)* is a 1.5kb deletion of that removes most of intron 1, and all of exons 2 and 3 (Figure 4A) and behaves as a recessive null allele (Brunschwig *et al.* 1999). *ceh-13(sw1)* homozygous animals are 97% embryonic lethal, with the remaining 3% of surviving larvae showing strong body morphology defects. We used this phenotype to identify L1 stage *sw1* homozygotes to analyze for *ceh-13p::GFP* and *cnd-1p::mCherry* expression. In wild type and *ceh-13(sw1)/qC1* heterozygotes, *cnd-1p::mCherry* and *ceh-13p::GFP* showed a complex and partially overlapping expression pattern in the retrovesicular ganglion and ventral nerve cord, with around two cells co-expressing *cnd-1* and *ceh-13* in the ganglion and

an average of 4.1 (wild type) and 4.7 cells (*ceh-13/qC1*) co-expressing in the ventral nerve cord (Figure 4B-O, Tables 4 and 5). In *cnd-1(gk718)* and *ceh-13(sw1)* homozygotes, the average number of cells co-expressing each marker in the ventral nerve cord dropped significantly to 3.0 and 2.8 for *cnd-1* and *ceh-13* respectively. Both *cnd-1* and *ceh-13* homozygous mutants also showed a significant reduction in cells that expressed *cnd-1p::mCherry* only, but not *ceh-13p::GFP*. Based on the similarity in phenotypes shown, we conclude that both *cnd-1* and *ceh-13* have roles in controlling a subset of ventral nerve cord cell fates during embryogenesis. We note that the ganglion and ventral cord cell counts in this double reporter gene assay were slightly different from the data reported in Figure 3. However, the data in Figure 3 was captured on a confocal microscope whereas the data in Figure 4 was captured using epifluorescence,

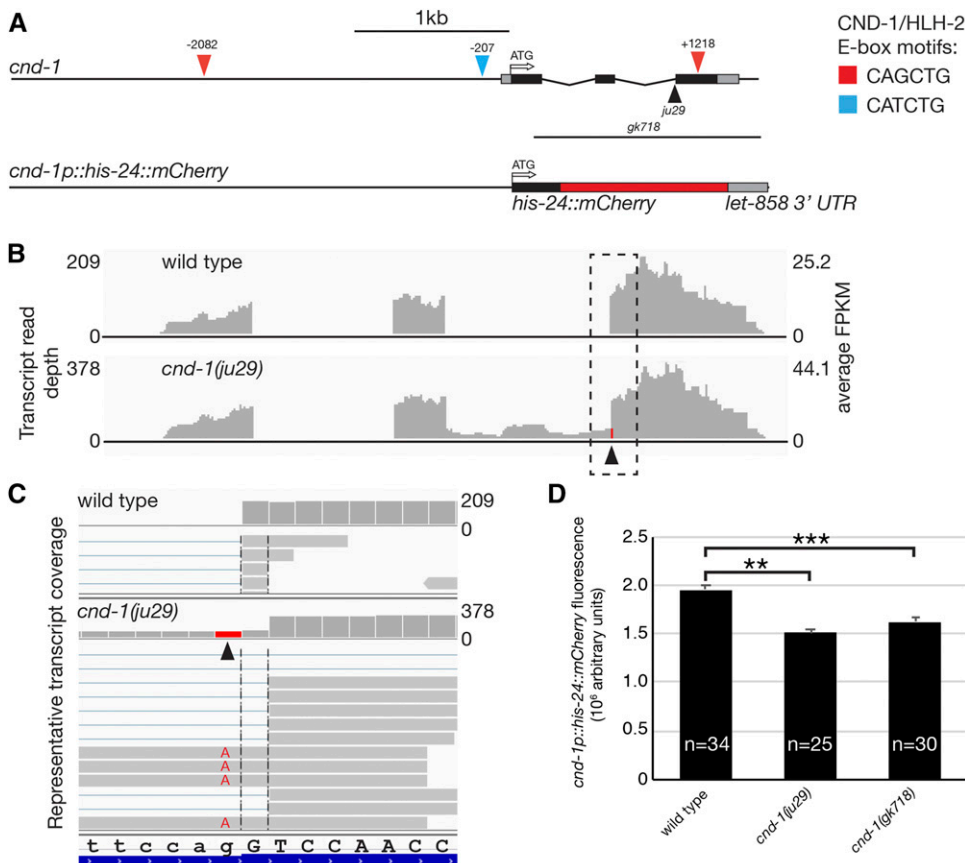
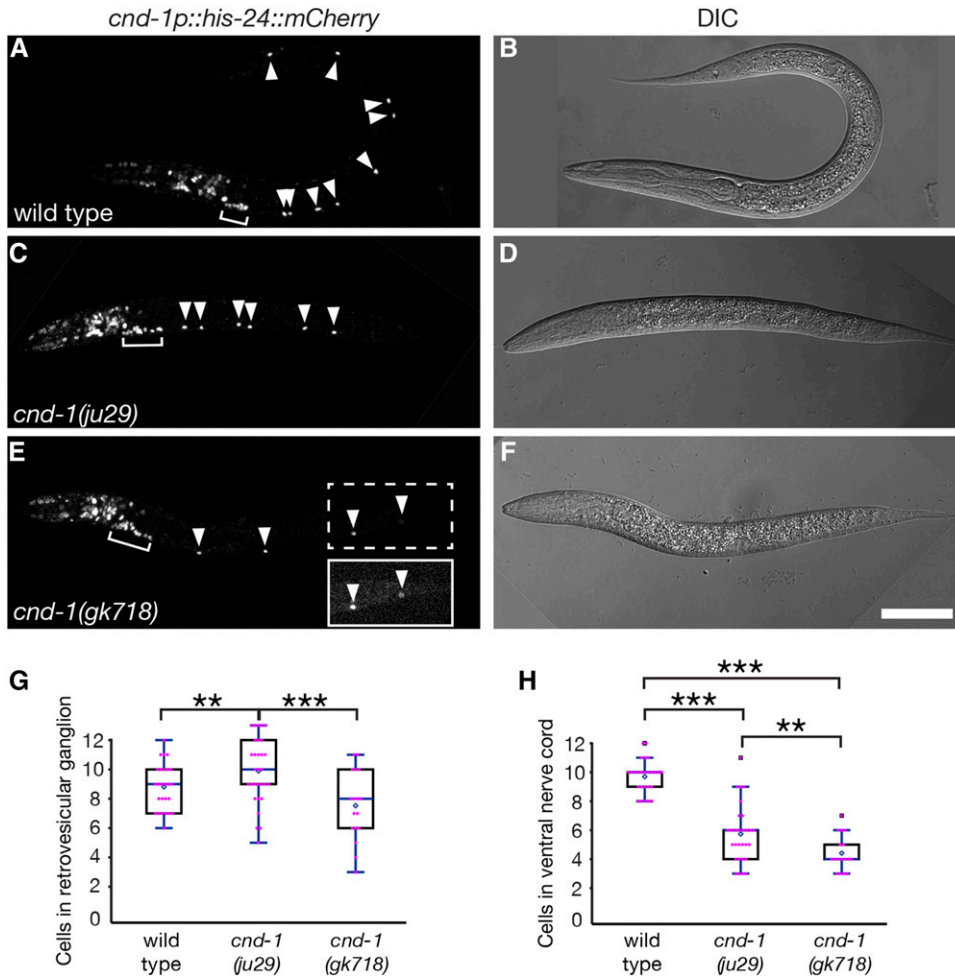


Figure 2 *cnd-1(ju29)* transcript levels are significantly up-regulated when compared to wild type, although embryonic *cnd-1p::mCherry* fluorescence is reduced in *cnd-1* mutants. (A) Schematic of the *cnd-1* genomic region showing location of predicted CND-1/HLH-2 binding sites, and the *stls10055* [*cnd-1p::his-24::mCherry*] reporter gene used extensively in this study. (B) Integrated Genome Viewer output of wild type and *cnd-1(ju29)* RNA-seq reads showing raw transcript depth and average Fold Per Kilobase Million coverage. Arrowhead shows location of the *ju29* mutation. (C) Inset of boxed region in (B), showing representative reads, the *ju29* G-to-A mutation, and the non-canonical 3' splice acceptor used in the *ju29* mutant. (D) Quantitative fluorescence of *cnd-1p::his-24::mCherry* expression in comma-stage embryos in wild type, *cnd-1(ju29)* and *cnd-1(gk718)* mutant embryos. Error bars show standard error of the mean. ** $P < 0.005$, *** $P < 0.0005$, Student's t-test with Bonferroni correction for multiple comparisons.



which gives slightly lower resolution and may have led to an undercount of cells that were directly adjacent to or behind each other.

***cnd-1* and *ceh-13* function redundantly to induce DD1 and DD2 motorneuron fate**

The similarity in *cnd-1* and *ceh-13* loss-of-function phenotypes and their effect on each other's reporter gene expression suggest that they may function together, either to cross-regulate each other or to specify ventral nerve cord motor neuron fate. To clarify this, we used an *unc-25p::GFP* reporter gene (Jin *et al.* 1999) to examine DD motorneuron fate in *cnd-1(gk718)* and *ceh-13(sw1)* single mutants, and *cnd-1(gk718) ceh-13(sw1)* double mutant L1 larvae (Figure 5). *unc-25p::GFP* is expressed in the six DD neurons (annotated in color, Figure 5A), in addition to the four RME head neurons (asterisk,

Figure 5A). We also plotted the cell body location of each DD neuron relative to the nose and tail tip, to establish which cells were more sensitive to loss of *cnd-1* or *ceh-13*. Wild type animals showed GFP expression in all six DD neurons (Figure 5B, H), in agreement with previous studies (Jin *et al.* 1999). However, *cnd-1(gk718)* mutants showed an average of 2.5 DD neurons (Figure 5C, H, $n = 25$, $P < 0.001$), with DD1 and DD2 being retained and DD3-6 being lost. L1 larvae of *ceh-13(sw1)/qC1* or *qC1* genotype (*i.e.*, those with wild type morphology) showed GFP expression in all 6 DD neurons ($n = 21$ larvae scored). In contrast, *ceh-13(sw1)* homozygous animals (identified by body morphology defects) showed on average four DD neurons, with DD1 generally being present but with variable loss of DD2-6 (Figure 5E, H, $n = 18$, $P < 0.001$). *cnd-1 ceh-13/qC1* balanced double mutants again showed a wild type DD neuron induction

■ **Table 3** *cnd-1* and *unc-25* markers co-express in sub-sets of ventral cord motorneurons. $N = 21$ wild type animals scored. Cells are listed from left-to-right in the anterior-to-posterior order they appear in L1 larvae. *In two animals, we saw an additional cell body (tentatively identified as RIGL) between DB3 and DA2. #Three animals had *cnd-1*-positive cells at the DB5 location. However, we suspect these are animals where DA4 and DB5 switched position during development

Ventral cord neuron	DB2	DD1	DB1	DA1	DB3	RIGL*	DA2	DD2	DA3	DB4	DA4	DD3
<i>unc-25p::GFP</i>	0	21	0	0	0	0	0	21	0	0	0	21
<i>cnd-1p::mCherry</i>	0	21	20	21	18	2	21	21	21	0	18	21
Ventral cord neuron	DB5	DA5	DD4	DB6	DA6	DD5	DB7	DA7	DD6	DA8	DA9	
<i>unc-25p::GFP</i>	0	0	21	0	0	21	0	0	21	0	0	
<i>cnd-1p::mCherry</i>	3#	21	21	0	0	21	0	0	21	0	0	

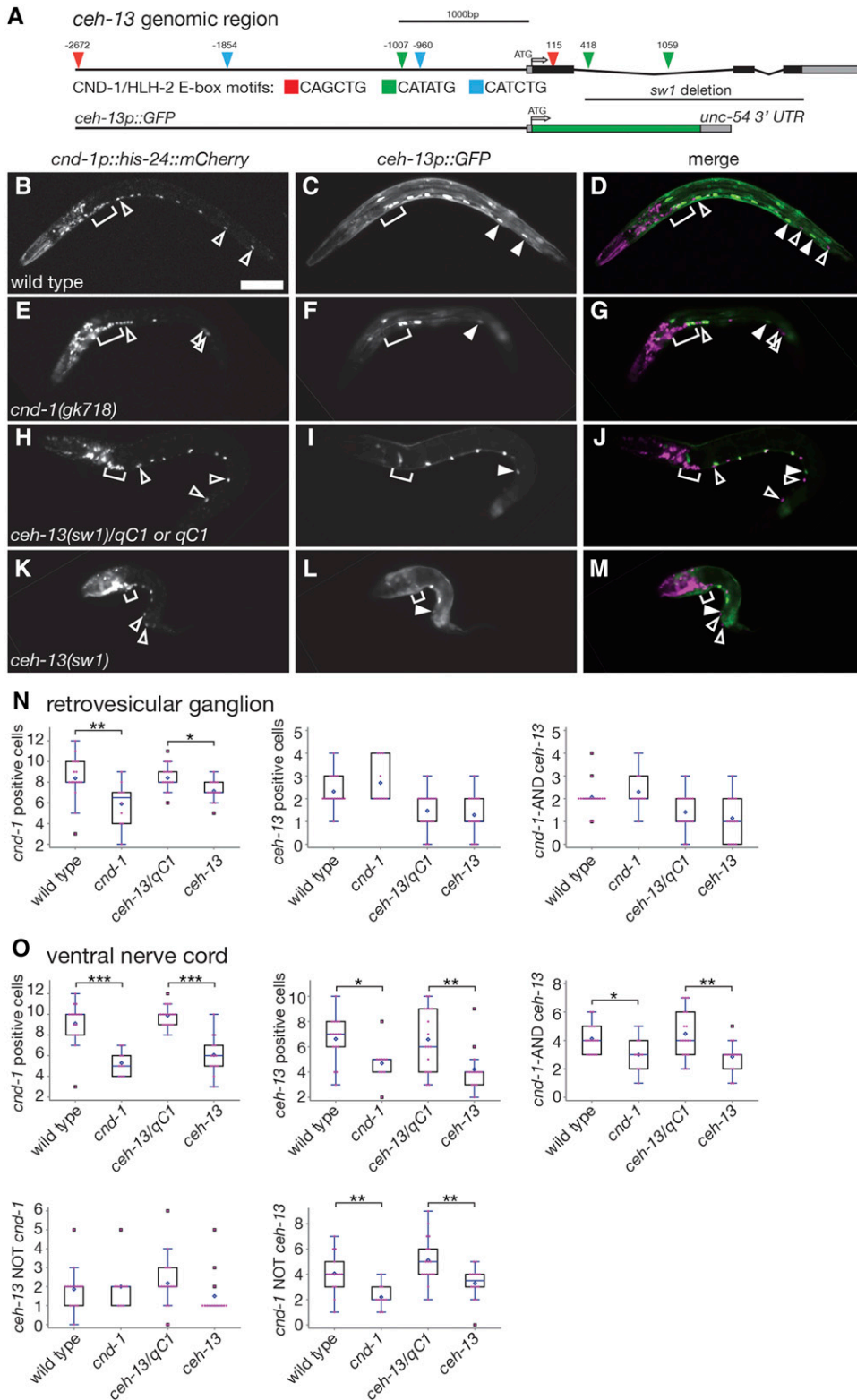


Figure 4 *ceh-13* and *cnd-1* show partially overlapping expression in the ventral nerve cord. (A) Schematic of the *ceh-13* genomic region showing predicted CND-1/HLH-2 binding sites, location of the *ceh-13*(*sw1*) mutation, and structure of the *ceh-13* reporter gene used in this study. (B-M) Expression patterns of *cnd-1p::his-24::mCherry* and *ceh-13p::GFP* in wild type, *cnd-1*(*gk718*), *ceh-13*(*sw1*)/*qC1* (or *qC1*), and *ceh-13*(*sw1*) mutants. While most VNC cells express both *cnd-1* and *ceh-13*, a distinct subset does not, and multiple cells are absent in both *cnd-1* and *ceh-13* mutants. Open arrowheads, *cnd-1* expressing cells only; filled arrowheads, *ceh-13* expressing cells only; bracket, retrovesicular ganglion. Scale bar in (B) = 25 μ m. (N, O) Box and whisker plots of *cnd-1* and *ceh-13* reporter genes showing average cell counts in the retrovesicular ganglia and ventral nerve cord respectively. Open diamond shows the average; box shows median, first, and third quartiles; whiskers show data extremes in 1.5 \times interquartile range with outliers shown beyond. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$, Mann-Whitney U -test with continuity correction.

pattern consistent with both *cnd-1* and *ceh-13* displaying recessive phenotypes (Figure 5F, H). However, 12/20 *cnd-1*(*gk718*) *ceh-13*(*sw1*) homozygous double mutants showed no DD neuron fate induction, with the remaining animals showing variable induction of one or more DD neurons (Figure 5G, H, $n = 20$, $P < 0.001$). One animal with

body morphology defects had six DD neurons. We speculate that this may have been a *cnd-1* *ceh-13*/*qC1* heterozygous or *qC1* homozygous animal that happened to have a body morphology defect. Overall, this suggests that *cnd-1* is primarily required for fate induction of DD3 through DD6, with *ceh-13* playing only a minor role in this process.

■ **Table 4** *cnd-1* and *ceh-13* markers co-express in mid-body DA and DD ventral cord motorneurons. Percentage of *cnd-1p::mCherry* and *ceh-13p::GFP*-positive neurons scored each for ventral nerve cord cell assayed. N = 10 wild type animals scored. Cells are listed from left-to-right in the anterior-to-posterior order they appear in L1 larvae. Strains that were highly mosaic for the *ceh-13p::GFP* reporter gene were omitted from this analysis

	DA2	DD2	DA3	DA4	DD3	DA5	DD4	DA6	DD5	DA7	DD6
<i>cnd-1p::mCherry</i>	100	100	100	100	100	100	100	0	100	0	100
<i>ceh-13p::GFP</i>	50	50	80	90	60	100	40	90	0	80	0

In contrast, *cnd-1* and *ceh-13* have redundant roles in DD1 and probably DD2 fate induction, with loss of either transcription factor still allowing robust fate specification. Note that distinguishing between DD2 and DD3 neurons was difficult in some animals, so DD2 fate is likely to be under-counted in *ceh-13* mutants.

***cnd-1* and *ceh-13* are redundantly required for setting ventral nerve cord cell number**

Our data above indicate that both *cnd-1* and *ceh-13* have roles in regulating expression of *unc-25p::GFP*, a reporter gene for GABAergic terminal fate specification of DD motorneurons. However, loss of *unc-25p::GFP* expression does not necessarily mean loss of neural fate induction. For instance, DA2-DA5 cholinergic motorneurons share the same great-grandmother cell that gives rise to the DD neurons, raising the possibility that DD cells might switch fates in *cnd-1* or *ceh-13* mutants (Sulston *et al.* 1983). To address this possibility, we examined DD fate induction using *unc-25p::GFP* in the presence of pan-neuronal terminal fate marker *rab-3p(prom1)::2xNLS::TagRFP* (Stefanakis *et al.* 2015). Wild type animals (n = 37) showed an average of 24 ventral nerve cord cells, of which six were *unc-25p::GFP*-positive consistent with our previous data (Figure 6A - C). However, in *cnd-1(gk718)* mutants (n = 46), only 20 ventral cord cells were observed (Figure 6D - F, $P < 0.001$ compared to wild type). In addition, they showed a parallel loss of around four DD cells similar to Figure 5C and 5H. *ceh-13(sw1)* homozygous mutants (n = 24) showed similar phenotypes, with 21 ventral cord cells counted compared to 25 in *sw1/qC1* balanced heterozygotes (n = 33) (Figure 6G - L, $P < 0.001$). Finally, *cnd-1(gk718) ceh-13(sw1)* double homozygous mutants (n = 20) showed around 18 cells in the ventral cord when compared to an average of 24 cells in *gk718 sw1/qC1* balanced heterozygotes (n = 24) (Figure 6M - R). Figure 6S and 6T summarize average DD neuron and ventral cord motorneuron counts for the above assays.

Overall, the difference in ventral cord neuron count parallels the loss of DD neurons observed in *cnd-1* and *ceh-13* mutant backgrounds and argues against a change of DD neuron fate to another neuronal cell type.

DISCUSSION

The role of *cnd-1* as a proneural transcription factor

Our comparative transcriptome data expand on *cnd-1*'s role as a proneural transcription factor, identifying the homeobox gene *ceh-5* as a novel downstream target of *cnd-1* (Figure 1). We find that *cnd-1* controls *ceh-5* gene expression in the RME neurons, suggesting that *ceh-5* may function as a terminal selector transcription factor in this cell type. Analysis of publicly available single-cell RNA-seq data allows us to contextualize the relationship between *cnd-1* and *ceh-5* (Packer *et al.* 2019). Figure 7 shows sub-lineages of single-cell RNA-seq data visualized using the Viscello package. Onset of *cnd-1* expression in the RME parent cell lineages (Figure 7A) occurs at the same time as onset of *ceh-5* expression (Figure 7B). However, our *ceh-5p::GFP* reporter gene data suggests that *cnd-1* controls *ceh-5* expression in the RME neurons. It may be that the single-cell RNA-seq data lacks the temporal resolution to define when one transcription factor is transcribed relative to another. Alternatively, *cnd-1* and *ceh-5* may function collaboratively to maintain *ceh-5* expression, for instance in other head neuron and muscle cell types, where we see *ceh-5p::GFP* expression drop but is not eliminated in *cnd-1* mutants. It should be noted that some aspects of RME neuron fate appear to be preserved in *cnd-1* mutants as they continue to express *unc-25p::GFP* (a known RME marker gene), even when *unc-25p::GFP* is lost in posterior DD neurons (Figure 5 and 6). This suggests that *ceh-5* may control a sub-module of RME terminal fate but not the actual fate of the neuron itself. The co-expression of *cnd-1* and *ceh-13* in terminal

■ **Table 5** Summary of *cnd-1p::mCherry* and *ceh-13p::GFP* expression in L1 larvae. (A, B) Average number of cells showing reporter gene expression in the retrovesicular ganglion and ventral nerve cord respectively (+/- standard error of the mean). * $P < 0.05$; ** $P < 0.01$, *** $P < 0.001$, Mann-Whitney U-test with continuity correction (*cnd-1* compared against wild type and *ceh-13* compared against *ceh-13/qC1*). The *ceh-13p::GFP* reporter stain used in this assay was an extra-chromosomal array and showed some expression variability between animals

A. Cells in retrovesicular ganglion						
	N	<i>cnd-1p::mCherry</i>	<i>ceh-13p::GFP</i>	<i>cnd-1</i> AND <i>ceh-13</i>	<i>ceh-13</i> NOT <i>cnd-1</i>	<i>cnd-1</i> NOT <i>ceh-13</i>
wild type	16	8.4(0.5)	2.3(0.2)	2.1(0.2)	—	—
<i>cnd-1(gk718)</i>	10	5.9(0.7) **	2.7(0.3)	2.3(0.3)	—	—
<i>ceh-13(sw1)/qC1</i> or <i>qC1</i>	17	8.5(0.4)	1.3(0.2)	1.3(0.2)	—	—
<i>ceh-13(sw1)</i>	14	6.9(0.3) *	1.2(0.3)	1.0(0.3)	—	—
B. Cells in ventral nerve cord						
wild type	16	9.1(0.5)	6.6(0.4)	4.1(0.3)	1.9(0.3)	4.1(0.4)
<i>cnd-1(gk718)</i>	10	5.3(0.4) ***	4.7(0.5) *	3.0(0.4) *	2.0(0.4)	2.2(0.3) **
<i>ceh-13(sw1)/qC1</i> or <i>qC1</i>	17	10.0(0.2)	6.5(0.6)	4.7(0.4)	1.4(0.4)	4.8(0.5)
<i>ceh-13(sw1)</i>	14	5.8(0.4) ***	4.2(0.5) **	2.8(0.3) **	1.6(0.3)	3.3(0.4) **

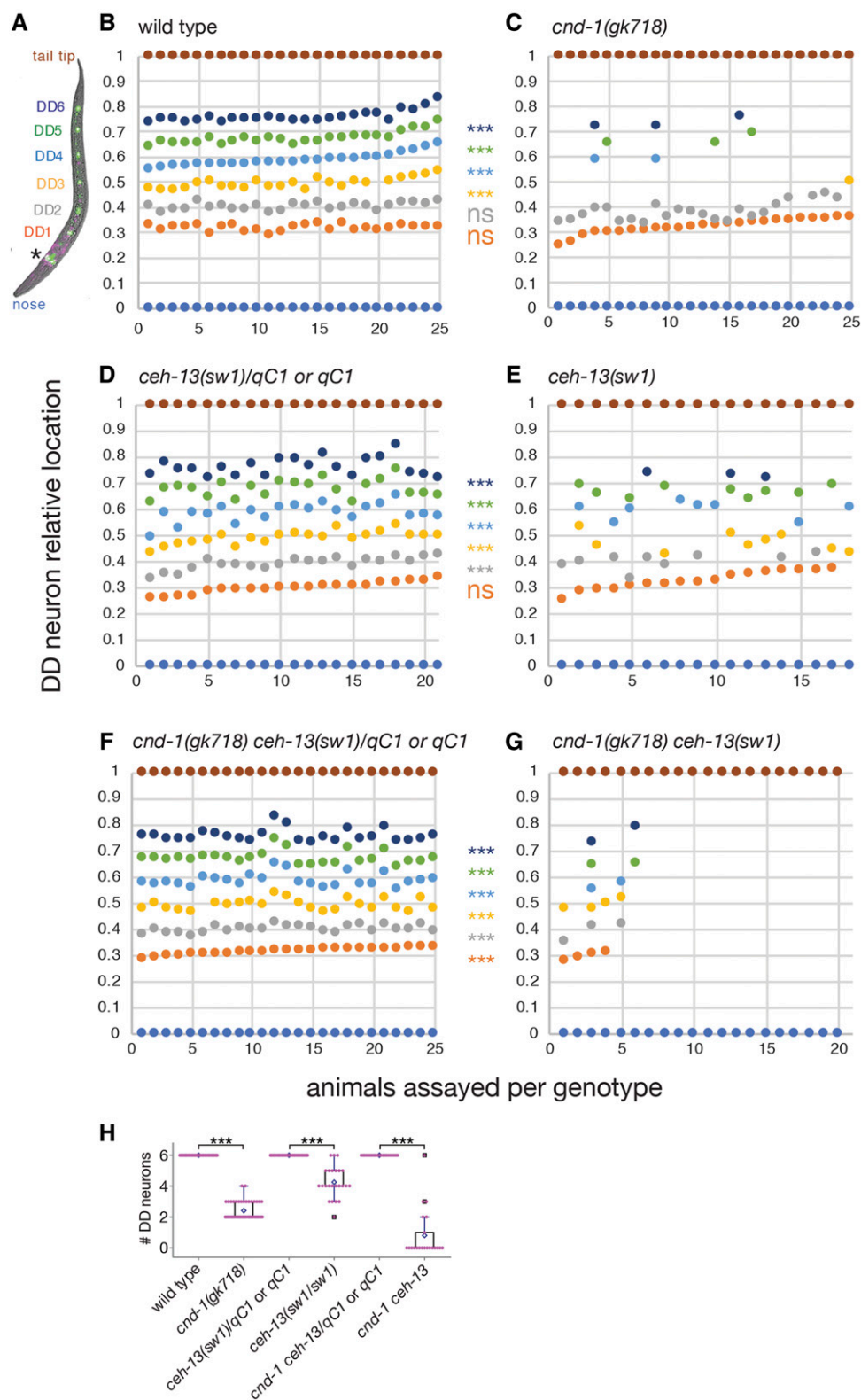
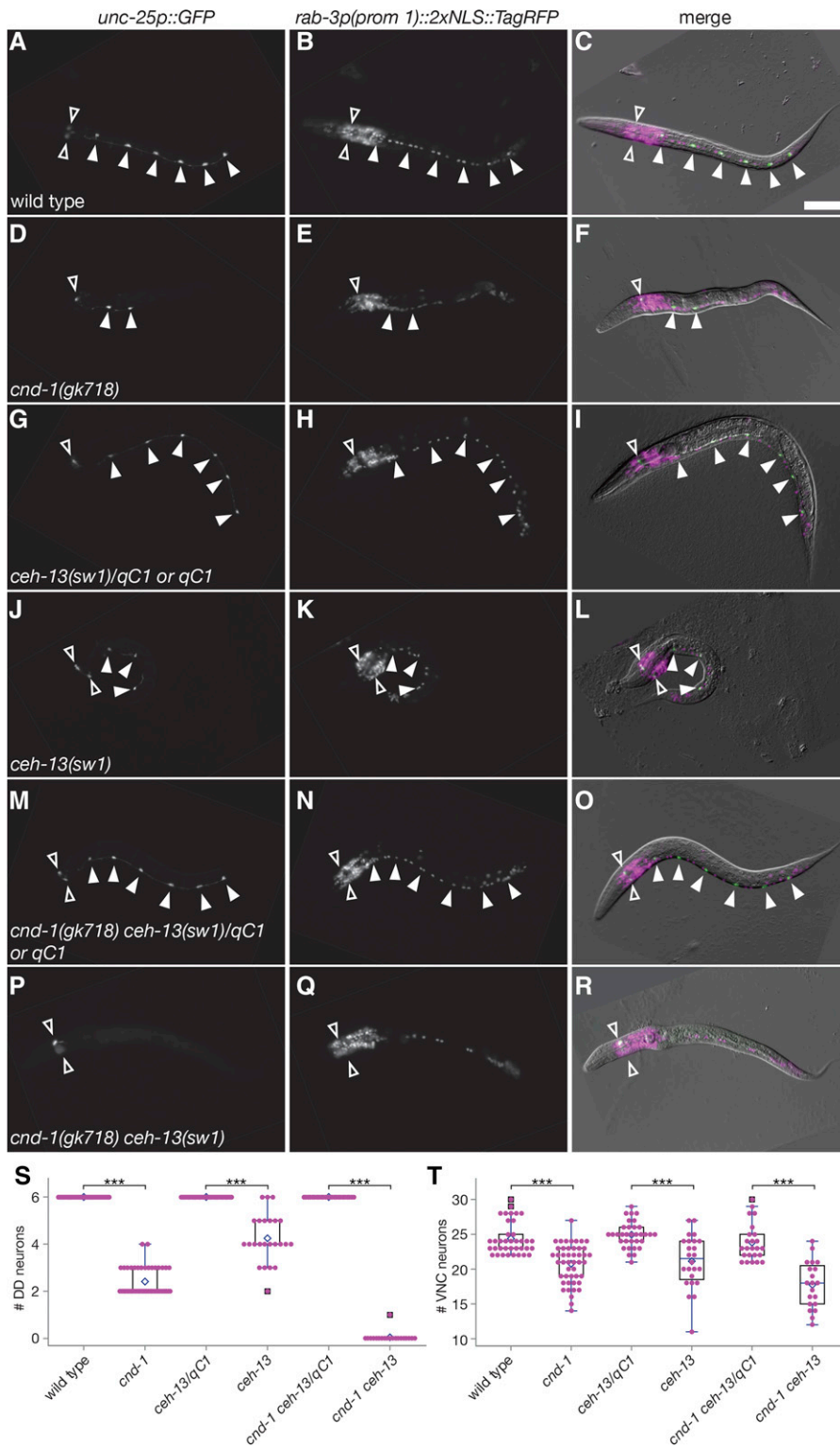


Figure 5 *cnd-1* and *ceh-13* function redundantly to control a subset of DD motoneuron cell fates. (A) Representative image of *unc-25p::GFP* and *cnd-1p::his-24::mCherry* expression in an L1 larva showing DD neuron color key. Asterisk shows RME head neurons. (B-G) Plots of DD neuron relative location in (B) wild type, (C) *cnd-1(gk718)*, (D) *ceh-13(sw1)/qC1*, (E) *ceh-13(sw1)*, (F) *cnd-1(gk718) ceh-13(sw1)/qC1*, and (G) *cnd-1(gk718) ceh-13(sw1)* L1 larvae. Body morphology defects were used to identify *sw1* or *gk718 sw1* homozygous animals. *** $P < 0.001$, Fisher exact test, presence vs. absence of each DD neuron type between control and experimental group (ns = not significant). (H) Box and whisker plots showing average number of DD neurons in the above strains. Open diamond shows the average; box shows median, first, and third quartiles; whiskers show data extremes in $1.5 \times$ interquartile range with outliers shown beyond. *** $P < 0.001$, Mann-Whitney U -test with continuity correction.

fate cells such as the RMEs may be predictive of *cnd-1*'s ability to control *ceh-5* expression in other neurons. Viscello data shows that these two transcription factors are co-expressed in RIVL/R, FLPL/R and PVQL/R neurons (Packer *et al.* 2019). Our data shows that *cnd-1(gk718)* mutants lose *ceh-5p::GFP* expression in PVQ neurons,

supporting this hypothesis. Overall, our data adds to previously published work, placing *cnd-1* as a proneural transcription factor upstream of *ceh-5*, *unc-3*, *unc-4*, and *unc-47*, to control aspects of RME, PVQ, DA, DB, and DD neuron fate respectively (Miller *et al.* 1992; Jin *et al.* 1994; Prasad *et al.* 1998; Kratsios *et al.* 2011).



CND-1 regulation of *cnd-1*-expressing cells

Our RNA-seq data shows that *cnd-1* transcription appears to be up-regulated in *cnd-1(ju29)* animals (Table 1B, Figure 2B and C). While there is evidence of intron inclusion, it does not appear to be sufficient to explain the almost twofold increase in *cnd-1(ju29)* transcript levels. We postulate that the intron inclusion may positively affect transcript stability, leading to higher levels of transcript

for longer. We used two separate reporter gene assays to further explore the role of *cnd-1* in nervous system development. First, quantitative imaging of *cnd-1p::his24::mCherry* in comma-stage embryos shows reduced *cnd-1* expression in the two *cnd-1* mutant alleles examined. Second, *cnd-1p::his24::mCherry* and *unc-25p::GFP* assays in ventral cord motoneurons reveals significantly lower cell counts in *cnd-1* mutants, with *cnd-1(ju29)* mutants displaying strong DD

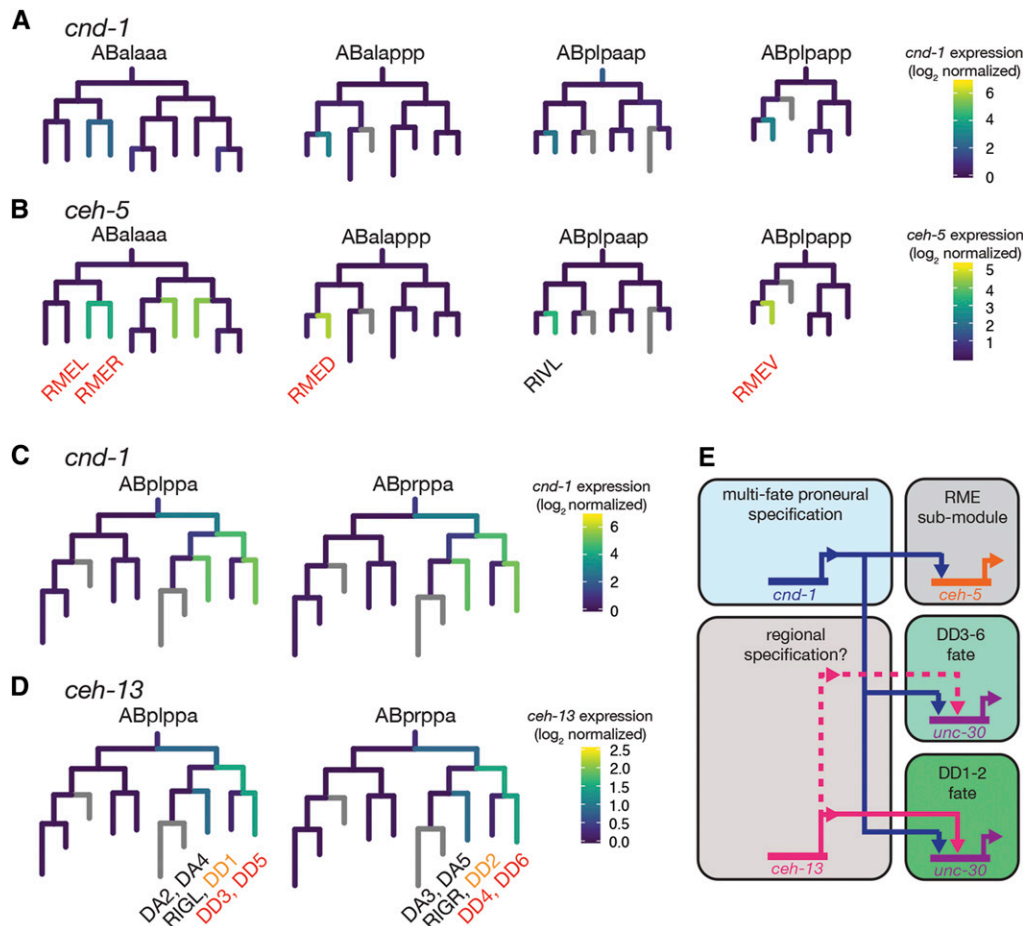


Figure 7 Summary of *cnd-1*'s role in controlling a *ceh-5*-dependent RME sub-module and DD motorneuron fate. (A - D) Single-cell RNA-seq expression lineages of (A, C) *cnd-1*, (B) *ceh-5*, and (D) *ceh-13* expression. Data derived from Packer *et al.* (2019) and visualized using the Viscello data display tool (<https://cello.shinyapps.io/celegans/>). Co-expression of *cnd-1* and *ceh-5* transcripts occurs in all RME-class neurons. We find that *ceh-5::GFP* expression in RME neurons is lost in *cnd-1(gk718)* mutants suggesting that *cnd-1* is responsible for driving *ceh-5* expression in these cells. Based on the above expression overlap, we predict that *cnd-1* also controls *ceh-5p::GFP* expression in RIVL/R. Note that *ceh-5* does not control RME neuron fate, because those cells can still be visualized using an *unc-25p::GFP* reporter gene. Similarly, *cnd-1* and *ceh-13* are co-expressed in DD3-6 motoneurons but not in DD1 and DD2. *unc-25p::GFP* expression in DD3-6 is primarily controlled by *cnd-1*, with a weak contribution from *ceh-13*. However, *unc-25p::GFP* expression in DD1 and DD2 is redundantly controlled by both *cnd-1* and *ceh-13*. An alternative interpretation is that both *cnd-1* and *ceh-13* are required for successful induction of ABpl/rppapp fates. Loss of these genes may mean that this cell division is lost, leading to a default anterior fate that permits aspects of DA2-6 to be specified normally but leads to loss of all DD neurons. (E) model summarizing *cnd-1*, *ceh-5*, and *ceh-13* function in the control of RME sub-module transcription and DD neuron fate specification.

neuron induction defects, although not as strong as those seen in *cnd-1(gk718)* mutants. This indicates that *cnd-1* is required for the fate specification of cells that normally express *cnd-1*. Whether this is via a self-regulatory mechanism is not known.

neuron induction defects, although not as strong as those seen in *cnd-1(gk718)* mutants. This indicates that *cnd-1* is required for the fate specification of cells that normally express *cnd-1*. Whether this is via a self-regulatory mechanism is not known.

CND-1 functions redundantly with CEH-13 to specify DD1 and DD2 cell fate

Our data corroborate previous work showing that loss-of-function in the Hox gene *ceh-13/labial* leads to loss of ventral cord motorneurons in a manner similar to that seen in *cnd-1* mutants (Hallam *et al.* 2000; Stefanakis *et al.* 2015). While *ceh-13* is not significantly different in our whole embryo RNA-seq dataset, a previous RNA-seq study, using Fluorescence-Activated Cell Sorting to enrich for *cnd-1*-labeled embryonic cells, revealed that *ceh-13* transcripts are up-regulated in that tissue relative to background (Burdick *et al.* 2016). This led us to investigate the genetic interaction between these two highly conserved transcription factors with regard to DD neuron fate specification (Figures 5 and 6). While both genes show similar loss-of-function DD neuron phenotypes, there are subtle differences. In *cnd-1(gk718)* mutants, cell fate specification of DD1 and DD2 ventral motorneurons appears normal, whereas DD3-6 are generally eliminated (Figure 5B and C). However, *ceh-13(sw1)* mutants show a weaker, variable loss of DD3-6, but with robust induction of DD1 and perhaps DD2. Analysis of *cnd-1(gk718)*

ceh-13(sw1) double mutants reveals a striking synergy with almost complete loss of all DD neurons (Figure 5F and G). This is not due to a change of cell fate, as analysis of all ventral nerve cord cells shows a corresponding cell count difference that mirrors the loss of DD neurons (Figure 6). Based on total ventral cord cell counts, we tentatively conclude that DA and DB motorneurons are not obviously affected by loss of *cnd-1* and/or *ceh-13*. This is in contrast to previously published work showing less cholinergic DA and DB ventral cord neurons when labeled by *acr-2p::YFP* (Hallam *et al.* 2000). One possible explanation for this discrepancy may be the transgenes used to label the cells. Hallam *et al.* (2000) used *acr-2p::YFP* and *unc-25p::GFP* (which are driven from a cholinergic receptor and glutamate decarboxylase promoter sequences respectively) to label all embryonic motorneurons and reported that some *cnd-1(ju29)* ventral cord cells lacked expression of both reporters. Despite this, those cells were apparently present, based on the stereotyped location of cell bodies and nuclei identified using Differential Interference Contrast microscopy. The reporter gene used in our assay was driven from a *rab-3* promoter element that was previously shown to label all neurons apart from the CAN associated neurons (Stefanakis *et al.* 2015). It may be that all DA and DB type neurons are born in *cnd-1* and/or *ceh-13* mutants but lack a complete battery of terminal selectors leading to apparent loss of identity, depending on the reporter gene used to label that cell.

Figure 7C and D shows the relationship between *cnd-1* and *ceh-13* transcript expression as identified via single-cell RNA-seq (Packer *et al.* 2019). Similar to the relationship between *cnd-1* and *ceh-5*, there does not appear to be any temporal sequence in their expression. Intriguingly, this visualization shows loss of *cnd-1* and *ceh-13* expression in the DD1 and DD2 mother cells (ABplppappa and ABprppappa respectively). We speculate that this renders DD1 and DD2 resistant to changes in either *cnd-1* or *ceh-13* expression, such that both cells are correctly specified in either single mutant background. Perhaps loss of both *cnd-1* and *ceh-13* promotes premature cell cycle exit in DD mother cells (as postulated by Hallam *et al.* 2000), preventing any DD neurons (and presumably RIGL and RIGR) from being born. While the DA2-5 mother cells express *cnd-1* and *ceh-13*, their grandparents (ABplppappa and ABprppappa) only express *cnd-1* (and at a lower level than the posterior daughter). This raises the possibility that this anterior branch of the lineage is less sensitive to these transcription factors and may give rise to a default set of cell fates, which means that DA2-5 are born whereas DD1-6 are lost. Figure 7E summarizes our analysis on the genetic interactions between *cnd-1* and *ceh-13* and the sub-fate terminal selector transcription factors *ceh-5* in RME and *unc-30* in DD neurons respectively. *unc-30* was previously reported to control the GABAergic neurotransmission module of DD and RME neurons (Eastman *et al.* 1999). While our RNA-seq assay does not reveal significant changes in *unc-30* expression, it may have been below the threshold for statistical significance.

ACKNOWLEDGMENTS

This work is dedicated to the memory of John Salerno, a great colleague and mentor, who passed away on December 25th, 2015. We thank Brian Ackley for providing feedback on this work prior to publication. We thank Ian Hope for the *ceh-5p::GFP* reporter strain and Parmida Jamshidi for building the *ceh-13(sw1)/qC1; cnd-1p::mCherry* strain. We also thank Yishi Jin for the pCZGY32 promoterless Gateway-GFP plasmid. We are grateful to Erik Lundquist and Jenny Hackett at the University of Kansas Genome Sequencing Core (GSC) for sequencing library construction, Illumina sequencing, and advice on RNA handling. KU GSC is supported by the NIH National Institute of General Medical Sciences under award number P20GM103638. The *cnd-1(gk718)* mutant was provided by the *C. elegans* Reverse Genetics Core Facility at the University of British Columbia, which is part of the international *C. elegans* Gene Knockout Consortium. Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). MLH is funded by NIH grant 1R15NS100632-01, and from grants from Kennesaw State University's Office of the Vice President for Research. WAN was funded by NIH grant R25GM111565 (PI Jonathan McMurry). Instrumentation used in this work was funded by NSF Major Research Instrument grants DBI-1624654, DBI-1337791, and DBI-1229237.

LITERATURE CITED

Alfano, G., I. Conte, T. Caramico, R. Avellino, B. Arnò *et al.*, 2011 Vax2 regulates retinoic acid distribution and cone opsin expression in the vertebrate eye. *Development* 138: 261–271. <https://doi.org/10.1242/dev.051037>

Anders, S., A. Reyes, and W. Huber, 2017 Detecting differential usage of exons from RNA-seq data. *Genome Res.* 22: 2008–2017. <https://doi.org/10.1101/gr.133744.111>

Andrews, S., 2010 FastQC: a quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

Apfeld, J., G. O. Connor, T. Mcdonagh, P. S. Distefano, and R. Curtis, 2004 The AMP-activated protein kinase AAK-2 links energy levels and insulin-like signals to lifespan in *C. elegans*. *Genes Dev.* 18: 3004–3009. <https://doi.org/10.1101/gad.1255404>

Baker, N. E., and N. L. Brown, 2018 All in the family: proneural bHLH genes and neuronal diversity. *Development* 145: 1–9. <https://doi.org/10.1242/dev.159426>

Basu, S. N., R. Kollu, and S. Banerjee-Basu, 2009 AutDB: a gene reference resource for autism research. *Nucleic Acids Res.* 37: D832–D836. <https://doi.org/10.1093/nar/gkn835>

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

Brunschwig K, C. Wittmann, R. Schnabel, T. R. Bürglin, H. Tobler, *et al.*, 1999 Anterior organization of the *Caenorhabditis elegans* embryo by the labial-like Hox gene *ceh-13*. *Development* 126: 1537–1546.

Burdick, J., T. Walton, E. Preston, A. Zacharias, A. Raj *et al.*, 2016 Overlapping cell population expression profiling and regulatory inference in *C. elegans*. *BMC Genomics* 17: 159. <https://doi.org/10.1186/s12864-016-2482-z>

C. elegans Deletion Mutant Consortium 2012 Large-scale screening for targeted knockouts in the *Caenorhabditis elegans* genome. *G3 (Bethesda)* 2: 1415–1425.

Chisholm, A. D., H. Hutter, Y. Jin, and W. G. Wadsworth, 2016 The Genetics of Axon Guidance and Axon Regeneration in *Caenorhabditis elegans*. *WormBook* 204: 849–882.

Christensen, E. L., A. Beasley, J. Radchuk, Z. E. Mielko, E. Preston *et al.*, 2020 *ngn-1/neurogenin* activates transcription of multiple terminal selector transcription factors in the *Caenorhabditis elegans* nervous system. *G3 (Bethesda)* 10: 1949–1962. <https://doi.org/10.1534/g3.120.401126>

Chuang, M., A. Goncharov, S. Wang, K. Oegema, Y. Jin *et al.*, 2014 The microtubule minus-end-binding protein patronin/ PTRN-1 is required for axon regeneration in *C. elegans*. *Cell Rep.* 9: 874–883. <https://doi.org/10.1016/j.celrep.2014.09.054>

Clapier, C. R., and B. R. Cairns, 2009 The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* 78: 273–304. <https://doi.org/10.1146/annurev.biochem.77.062706.153223>

D'Amico, L. A., D. Boujard, and P. Coumilleau, 2013 The neurogenic factor NeuroD1 is expressed in post-mitotic cells during juvenile and adult xenopus neurogenesis and not in progenitor or radial glial cells. *PLoS One* 8: e66487. <https://doi.org/10.1371/journal.pone.0066487>

Eastman, C., H. R. Horvitz, and Y. Jin, 1999 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.* 19: 6225–6234. <https://doi.org/10.1523/JNEUROSCI.19-15-06225.1999>

Edgar, R., M. Domrachev, and A. E. Lash, 2002 Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30: 207–210. <https://doi.org/10.1093/nar/30.1.207>

Grove, C. A., F. De Masi, M. I. Barrasa, D. E. Newburger, M. J. Alkema *et al.*, 2009 A multiparameter network reveals extensive divergence between *C. elegans* bHLH transcription factors. *Cell* 138: 314–327. <https://doi.org/10.1016/j.cell.2009.04.058>

Hallam, S., E. Singer, D. Waring, and Y. Jin, 2000 The *C. elegans* NeuroD homolog *cnd-1* functions in multiple aspects of motor neuron fate specification. *Development* 127: 4239–4252.

Hardie, D. G., 2014 AMP-activated protein kinase: a key regulator of energy balance with many roles in human disease. *J. Intern. Med.* 276: 543–559. <https://doi.org/10.1111/joim.12268>

Hobert, O., 2016 Terminal Selectors of Neuronal Identity. *Curr. Top. Dev. Biol.* 116: 455–475. <https://doi.org/10.1016/bs.ctdb.2015.12.007>

Hudson, M. L., T. Kinnunen, H. N. Cinar, and A. D. Chisholm, 2006 *C. elegans* Kallmann syndrome protein KAL-1 interacts with syndecan and glypican to regulate neuronal cell migrations. *Dev. Biol.* 294: 352–365. <https://doi.org/10.1016/j.ydbio.2006.02.036>

Jin, Y., R. Hoskins, and H. R. Horvitz, 1994 Control of type-D GABAergic neuron differentiation by *C. elegans* UNC-30 homeodomain protein. *Nature* 372: 780–783. <https://doi.org/10.1038/372780a0>

- Jin, Y., E. Jorgensen, E. Hartwig, and H. R. Horvitz, 1999 The *C. elegans* gene *unc-25* encodes glutamic acid decarboxylase and is required for synaptic transmission but not synaptic development. *J. Neurosci.* 19: 539–548. <https://doi.org/10.1523/JNEUROSCI.19-02-00539.1999>
- Jin, Y., and Y. B. Qi, 2018 Building stereotypic connectivity: mechanistic insights into structural plasticity from *C. elegans*. *Curr. Opin. Neurobiol.* 48: 97–105. <https://doi.org/10.1016/j.conb.2017.11.005>
- Kim, D., B. Langmead, and S. L. Salzberg, 2015 HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12: 357–360. <https://doi.org/10.1038/nmeth.3317>
- Kratsios, P., A. Stolfi, M. Levine, and O. Hobert, 2012 Coordinated regulation of cholinergic motor neuron traits through a conserved terminal selector gene. *Nat. Neurosci.* 15: 205–214. <https://doi.org/10.1038/nn.2989>
- Lee, J. E., 1997 Basic helix-loop-helix genes in neural development. *Curr. Opin. Neurobiol.* 7: 13–20. [https://doi.org/10.1016/S0959-4388\(97\)80115-8](https://doi.org/10.1016/S0959-4388(97)80115-8)
- Lee, J. E., S. M. Hollenberg, L. Snider, D. L. Turner, N. Lipnick *et al.*, 1995 Conversion of xenopus ectoderm into neurons by NeuroD, a basic helix-loop-helix protein. *Science* 268: 836–844. <https://doi.org/10.1126/science.7754368>
- Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The Sequence Alignment / Map format and SAMtools. *Bioinformatics* 25: 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Liu, M., Y. Liu, Y. Liu, G. Lupo, L. Lan *et al.*, 2008 A Role for Xvax2 in Controlling Proliferation of Xenopus Ventral Eye and Brain Progenitors. *Dev. Dyn.* 237: 3387–3393. <https://doi.org/10.1002/dvdy.21763>
- Love, M. I., W. Huber, and S. Anders, 2014 Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15: 1–21. <https://doi.org/10.1186/s13059-014-0550-8>
- Marcette, J. D., J. J. Chen, and M. L. Nonet, 2014 The *Caenorhabditis elegans* microtubule minus-end binding homolog PTRN-1 stabilizes synapses and neurites. *eLife* 2014: 1–16.
- Miller, D. M. I., M. M. Shen, C. E. Shamu, T. R. Bürglin, G. Ruvkun *et al.*, 1992 *C. elegans unc-4* Gene Encodes a Homeodomain Protein That Determines the Pattern of Synaptic Input to Specific Motor Neurons. *Nature* 355: 841–845. <https://doi.org/10.1038/355841a0>
- Miyata, T., T. Maeda, and J. E. Lee, 1999 NeuroD is required for differentiation of the granule cells in the cerebellum and hippocampus. *Genes Dev.* 13: 1647–1652. <https://doi.org/10.1101/gad.13.13.1647>
- Murray, J. I., Z. Bao, T. J. Boyle, M. E. Boeck, L. Barbara *et al.*, 2008 Automated analysis of embryonic gene expression with cellular resolution in *C. elegans*. *Nat. Methods* 5: 703–709. <https://doi.org/10.1038/nmeth.1228>
- Murray, J. I., T. J. Boyle, E. Preston, D. Vafeados, B. Mericle *et al.*, 2012 Multidimensional regulation of gene expression in the *C. elegans* embryo. *Genome Res.* 22: 1282–1294. <https://doi.org/10.1101/gr.131920.111>
- Packer, J. S., Zhu, Q., Huynh, C., Sivaramakrishnan, P., Preston, E. *et al.*, 2019 A lineage-resolved molecular atlas of *C. elegans* embryogenesis at single-cell resolution. *Science* 365 <https://doi.org/10.1126/science.aax1971>
- Pataskar, A., J. Jung, P. Simalowski, F. Noack, F. Calegari *et al.*, 2016 NeuroD1 reprograms chromatin and transcription factor landscapes to induce the neuronal program. *EMBO J.* 35: 24–45. <https://doi.org/10.15252/emboj.201591206>
- Pertea, M., D. Kim, G. M. Pertea, J. T. Leek, and S. L. Salzberg, 2016 Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* 11: 1650–1667. <https://doi.org/10.1038/nprot.2016.095>
- Pfaffl, M. W., 2001 A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29: e45. <https://doi.org/10.1093/nar/29.9.e45>
- Prasad, B. C., B. Ye, R. Zackhary, K. Schrader, G. Seydoux *et al.*, 1998 *unc-3*, a gene required for axonal guidance in *Caenorhabditis elegans*, encodes a member of the O / E family of transcription factors. *Development* 125: 1561–1568.
- Reece-Hoyes, J. S., J. Shingles, D. Dupuy, C. A. Grove, A. J. M. Walhout *et al.*, 2007 Insight into transcription factor gene duplication from *Caenorhabditis elegans* Promoterome-driven expression patterns. *BMC Genomics* 8: 27. <https://doi.org/10.1186/1471-2164-8-27>
- Richardson, C. E., K. A. Spilker, J. G. Cueva, J. Perrino, M. B. Goodman *et al.*, 2014 PTRN-1, a microtubule minus end-binding CAMSAP homolog, promotes microtubule function in *Caenorhabditis elegans* neurons. *eLife* 2014: 1–22.
- Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014 Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511: 421–427. <https://doi.org/10.1038/nature13595>
- Robinson, J. T., H. Thorvaldsdottir, W. Winckler, M. Guttman, E. S. Lander *et al.*, 2011 Integrative Genomics Viewer. *Nat. Biotechnol.* 29: 24–26. <https://doi.org/10.1038/nbt.1754>
- Rubio-Cabezas, O., J. A. Minton, I. Kantor, D. Williams, S. Ellard *et al.*, 2010 Homozygous mutations in NEUROD1 are responsible for a novel syndrome of permanent neonatal diabetes and neurological abnormalities. *Diabetes* 59: 2326–2331. <https://doi.org/10.2337/db10-0011>
- Schindelin, J., I. Arganda-Carreras, E. Frise, V. Kaynig, T. Pietzsch *et al.*, 2012 Fiji - an Open Source platform for biological image analysis. *Nat. Methods* 9: 676–682. <https://doi.org/10.1038/nmeth.2019>
- Schultz, R. D., and T. L. Gumienny, 2012 Visualization of *Caenorhabditis elegans* Cuticular Structures Using the Lipophilic Vital Dye DiI. *J. Vis. Exp.* 59: e3362.
- Seo, S., J. Lim, D. Yellajoshiyula, L. Chang, and K. L. Kroll, 2007 Neurogenin and NeuroD direct transcriptional targets and their regulatory enhancers. *EMBO J.* 26: 5093–5108. <https://doi.org/10.1038/sj.emboj.7601923>
- Stefanakis, N., I. Carrera, and O. Hobert, 2015 Regulatory Logic of Pan-Neuronal Gene Expression in *C. elegans*. *Neuron* 87: 733–750. <https://doi.org/10.1016/j.neuron.2015.07.031>
- Sulston, J. E., and H. R. Horvitz, 1977 Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev. Biol.* 56: 110–156. [https://doi.org/10.1016/0012-1606\(77\)90158-0](https://doi.org/10.1016/0012-1606(77)90158-0)
- Sulston, J. E., E. Schierenberg, J. G. White, and J. N. Thomson, 1983 The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* 100: 64–119. [https://doi.org/10.1016/0012-1606\(83\)90201-4](https://doi.org/10.1016/0012-1606(83)90201-4)
- Take-uchi, M., J. D. W. Clarke, and S. W. Wilson, 2003 Hedgehog signalling maintains the optic stalk-retinal interface through the regulation of Vax gene activity. *Development* 130: 955–968. <https://doi.org/10.1242/dev.00305>
- Thorvaldsdottir, H., J. Robinson, and J. P. Mesirov, 2013 Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* 14: 178–192. <https://doi.org/10.1093/bib/bbs017>
- Waddington, C. H., 1957 *The Strategy of the Genes. A Discussion of Some Aspects of Theoretical Biology*, Allen & Unwin, London.
- Wang, L., and N. E. Baker, 2015 E proteins and ID proteins: Helix-loop-helix partners in development and disease. *Dev. Cell* 35: 269–280. <https://doi.org/10.1016/j.devcel.2015.10.019>

Communicating editor: J. Kim