The transcription factor ELT-2 positively and negatively impacts direct target genes to modulate the *Caenorhabditis elegans* intestinal transcriptome

David C. King[*†] and Robert T.P. Williams[*†], Izabella R. Mastroianni[*], Jessica Hill[*], Nicolai W. Apenas[‡], Gabriela Ramirez[§], and Erin Osborne Nishimura[*, **]

[*]Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, CO, USA 80523

[†] Authors contributed equally

[‡]Current affiliation – Experimental Immunology Branch, National Cancer Institute, US National Institutes of Health, Bethesda, MD, USA

[§] Current affiliation – Center for Vector-Borne Infectious Diseases, Colorado State University, Fort Collins, CO, USA

** Corresponding author: erin.nishimura@colostate.edu

RUNNING TITLE: ELT-2 binding dynamics in *C. elegans*

KEYWORDS: *C. elegans,* intestine, ELT-2, ELT-7, ChIP-seq, transcription factor, organogenesis, master regulator

CORRESPONDENCE:

Erin Osborne Nishimura
1870 Campus Delivery
Colorado State University
Fort Collins, CO 80523

erin.nishimura@colostate.edu

ABSTRACT:

The transcription factor ELT-2 is the major factor required for *Caenorhabditis elegans* intestinal development. ELT-2 expression initiates in embryos, promoting development, but persists after hatching through larval and adult stages, where it contributes to diverse intestinal functions. How ELT-2's regulatory role changes over development is largely unexplored. We sought to determine whether ELT-2 target gene preference changes over developmental time and whether those changes underlie transcriptome dynamics. We analyzed stage-specific ELT-2 ChIP-seq assays to identify sets of dynamically changing ELT-2 peaks. Embryo-specific and larval-specific ELT-2 peaks were less common than peaks that accumulated in the L3 stage or progressively throughout development. To assess how ELT-2 occupancy related to transcriptional output, we compared patterns of ELT-2 occupancy over developmental time to intestinal expression dynamics and sets of ELT-2-dependent genes. ELT-2 occupancy predicted gene expression changes at larval but not embryonic stages, suggesting that ELT-2 promotes transcriptional output in some cases. However, we also found that ELT-2 negatively regulated a subset of its direct target genes. Repressed target genes were over-represented for neuronal and gonad functions, signifying that ELT-2 may protect the intestine from these tissue fates. Furthermore, we observed that ELT-2 repressed its own promoter in a negative feedback loop to stabilize the *elt-2* gene's expression to a set amount. These findings illustrate that ELT-2 exerts both positive and negative regulatory control - turning on some target genes to promote intestinal function while down-regulating others to prevent or reduce their transcriptional output.

INTRODUCTION:

Transcription factors (TFs) work together to form gene regulatory networks (GRNs) that direct the expression of target genes (Davidson *et al.* 2003; Levine and Davidson 2005; Emmert-Streib *et al.* 2014). These GRNs form positive, negative, feed-forward, and auto-inhibitory connections to orchestrate downstream transcriptional responses critical for development and health (Alon 2007; Murray *et al.* 2012; Oestreich and Weinmann 2012; Delás and Briscoe 2020). The *Caenorhabditis elegans* intestine is a model for understanding how GRNs contribute to organogenesis (Maduro and Rothman 2002; McGhee 2007; Dimov and Maduro 2019). In embryos, the intestinal GRN initiates with maternally loaded transcription factors and culminates in the expression of the GATA transcription factor ELT-2 (McGhee 2007; Maduro 2019). Historically, ELT-2 was heralded as the intestinal master regulator for the following reasons: 1) ELT-2 loss leads to a dramatic intestinal phenotype, 2) ELT-2 misexpression ectopically induces intestinal features, 3) ELT-2 promotes its own expression, and 4) 1) ELT-2's TGATAA binding sites reside in the promoters of almost all intestine-expressed genes in both embryonic and larval stages. (Fukushige *et al.* 1998; McGhee *et al.* 2007, 2009; Riddle *et al.* 2013, 2016; Du *et al.* 2016; Choi *et al.* 2017). However, a more extensive network of TFs supports ELT-2 by preceding ELT-2 in intestinal specification, working with it during cell commitment and acting downstream of it in intestinal differentiation. This is a departure from the strict definition of a master regulator TF as an independent driver of organogenesis and illustrates the larger context of the intestinal GRN in which ELT-2 operates. A missing aspect of the *C. elegans* intestinal GRN model is how it evolves over time. Though many TFs within the described embryonic network are transient, ELT-2 and its partner ELT-7 remain expressed for the duration of the worm's lifespan. Whether their roles remain consistent or change over developmental is unclear.

The GRN underlying *C. elegans* intestinal development is well established. The intestine arises clonally from a single cell, the E cell, and ultimately produces a 20-cell-long tube that extends from the animal's pharynx to its rectum. At the 4-cell stage, specification of the endomesoderm occurs when a non-canonical Wnt-signaling pathway responds to positional cues to alleviate repression of the maternal TF, SKN-1 (Maduro *et al.* 2005b). The endoderm, or E-cell, is determined through SKN-1-dependent activation of a series of GATA-family transcription factors which bind similar regulatory DNA sequences

(WGATAR or GTATACTY) (Broitman-Maduro *et al.* 2005; Wiesenfahrt *et al.* 2015). Beginning in the E-cell (8-cell stage of embryogenesis), transient pulses of GATA pairs, first MED-1/MED-2 then END-1/END-3, lead to the eventual expression of the final pair ELT-2 and ELT-7 (Maduro *et al.* 2005a; Sommermann *et al.* 2010; Choi *et al.* 2017; Dimov and Maduro 2019; Maduro 2019). ELT-2 and ELT-7 are partially overlapping TFs whose expression is sustained through embryonic, larval, and adult stages, only declining in aged, post-fertile worms (McGhee *et al.* 2009; Sommermann *et al.* 2010; Mann *et al.* 2016; Dineen *et al.* 2018). Upon initial ELT-2 expression, the intestine continues to divide and grow, undergoing a further 3-4 rounds of cell division. During these divisions, the E lineage ingresses, aligns into two rows down the center, fuses, and finally creates a lumen to form the alimentary canal (McGhee 2007).

During larval and adult stages, ELT-2 continues to mark intestinal identity and contributes to diverse intestinal processes. ELT-2 is required for optimal performance of digestion, immunity, detoxification, and aging (Elliott *et al.* 2010; Head and Aballay 2014; Roh *et al.* 2014; Head *et al.* 2016; Mann *et al.* 2016; Yang *et al.* 2016; Su *et al.* 2020; Zárate-Potes *et al.* 2020), but whether the role of ELT-2 on other functions is direct or indirect remains unknown in many cases.

Indeed, the intestinal transcriptome changes dramatically over time as only 20% of genes are shared between embryonic and larval stages, as previously reported (McGhee *et al.* 2009). These changes underscore the multiple functions of the intestine upon hatching. *C. elegans* intestines coordinate digestion, nutrient absorption, metabolism, fat storage, immune response, yolk production, genetic aging, insulin signaling, detoxification, and environmental responses like avoidance behavior and communication (Kimble and Sharrock 1983; An and Blackwell 2003; Libina *et al.* 2003; Martinez-Finley and Aschner 2011; Ludewig and Schroeder 2013; Chun *et al.* 2017; Lee and Mylonakis 2017). Because the TGATAA sites to which ELT-2 binds are over-represented in the promoters of genes expressed in both embryonic and larval stages, ELT-2 was initially suggested to activate all intestinal genes.

However, ChIP-seq assays of ELT-2 have revealed that ELT-2 binds only a small fraction of potential binding sites, echoing a general theme in the TF field that sequence alone does not dictate TF occupancy (Carr and Biggin 1999; Liu *et al.* 2006; Yang *et al.* 2006; McGhee *et al.* 2009; Wiesenfahrt *et al.* 2015; Mann *et al.* 2016; Kudron *et al.* 2017). Chromatin, flanking sequences, long-range interactions, and combinatorial binding also contribute to the final set of genomic loci a TF inhabits. To further complicate matters, TFs differ in the degree to which their physical presence at a promoter site leads to functional transcriptional activation (Ucar *et al.* 2009). By integrating genome-wide ELT-2 binding with transcriptomic data, we are interested in defining both how ELT-2's role as a transcriptional regulator changes over the course of the intestine's life history and whether or how ELT-2 shapes the intestinal transcriptome.

How might ELT-2's dynamics over time correlate to intestinal transcriptome changes? On the one hand, a high correlation between ELT-2 occupancy and transcriptional output would suggest ELT-2 plays an instructive role, promoting transcriptional activity via higher occupancy. In this scenario, fluctuating levels of ELT-2 could be accounted for by chromatin landscape, differences in ELT-2 sequence preference, changing ELT-2 concentrations, or cooperative binding partners. In support of this concept, the Drosophila TF Bicoid promotes anterior to posterior transcriptional differences through concentration-dependent activation at different promoters (Driever *et al.* 1989). In contrast, ELT-2 may occupy all intestinal genes at all stages of development but only catalyze transcriptional activation in combination with other, stage-specific TFs. This combinatorial activity is typical of *C. elegans* (Hobert 2016). Similarly, vitellogenin (yolk protein) transcriptional output integrates information for intestine-specificity (via ELT-2 binding) with TFs for sex and worm size (Goszczynski *et al.* 2016). Finally, it is possible that ELT-2 is highly placed in the hierarchy of the intestinal GRN, and a TF downstream of ELT-2, or one that is independent of ELT-2, accounts for stage-specific intestinal gene expression. In this case, we would predict ELT-2 to be absent from many intestine-expressed genes. These models are

simplifications. Combinations of these cases may be occurring, and different behaviors could be relevant to different subsets of target genes.

To determine which model ELT-2 utilizes to regulate the intestine GRN, we analyzed a publicly available set of ELT-2 ChIP-seq samples performed at three developmental stages. We compared the resulting ELT-2 occupancy dynamics to intestine transcriptional dynamics and *elt-2* loss-of-function transcriptome profiling. We identified a general trend in which ELT-2 occupancy correlated to target gene expression, indicative of the instructive model above. However, we also found evidence that ELT-2 negatively regulates subsets of directs targets, including its own promoter.

A growing number of ChIP-seq datasets are available that characterize TF occupancy over time through the modERN Resource (model organism Encyclopedia of Regulatory Networks). Therefore, the approaches we undertook here may serve as a general roadmap to the study of other transcription factors that contribute to numerous *C. elegans* GRNs (Kudron *et al.* 2017).

MATERIALS AND METHODS:

C. elegans strains and culture conditions

All worm strains were maintained as described (Stiernagle 2006) and cultured at 20°C on NGM plates unless otherwise stated. Wild-type strain N2 (Bristol) was used. Transgenic strains used in this study are below:

OP56 unc-119(ed3) III; wgIs56 [elt-2::TY1::EGFP::3xFLAG(92C12) + unc-119(+)].
OP563 unc-119(tm4063) III; wgIs563 [cebp-1::TY1::EGFP::3xFLAG + unc-119(+)].
JM149 caIs71[elt-2p::GFP::HIS-2B::unc-54 3'UTR + rol-6(su1006)].
JM259 elt-7(tm840) V; caIs71[elt-2p::GFP::HIS-2B::unc-54 3'UTR + rol-6(su1006)].

Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440).

Computational Analysis

Detailed descriptions of computational analytics are listed in Supplemental Methods including 1) ELT-2 ChIP-seq peak calling, 2) quantification of ETL-2 ChIP-seq binding signal, 3) identification of developmentally dynamic ELT-2 ChIP-seq peaks, 4) relative ELT-2 occupancy and clustering, 5) ChIP-seq peak DNA motif analysis, 6) mapping ELT-2 peaks to *C. elegans* genes, 7) gene ontology analysis, 8) developmental intestine-specific transcriptome analysis, 9) correlation of gene clusters and ELT-2 occupancy types, and 10) transcriptional response to single and double *elt-2;elt-7* mutations. All in-house scripts are available (https://github.com/meekrob/kw-onish-elt-2).

RNAi treatment

To measure the effect of ELT-2 depletion, we used *E. coli* feeding strains engineered to produce double-stranded RNA cognate to *elt-2* transcript. Freshly starved worms were chunked to large NGM/OP50 plates to grow for 48 hours. Synchronized embryos were isolated from mixed stage populations through hypochlorite treatment (Stiernagle 2006). To visualize the effect of ELT-2 depletion on CEBP-1 expression (Figure 6), we grew synchronized embryos to L3 stage by incubation at 20°C for 48 hours before RNAi exposure. L3 stage worms were transferred to NGM plates seeded with RNAi-inducing *E. coli* and incubated for an additional 48 hours until gravid. Synchronized L1s were collected by transferring RNAi treated gravid adults to 50 ul drops of M9 and culturing for 24 hours. To image the response of *elt-2* promoter across developmental stages to ELT-2 depletion, we modified the above procedure by exposing worms to RNAi feeding strains 24 hours before the queried developmental stage (Figure 7, Figure S6). Negative control experiments utilized feeding strains containing L4440 empty vector. RNAi plasmids were sequence verified before experimental use.

Fluorescence Microscopy

Response of reporter gene expression to ELT-2 depletion was visualized using fluorescence microscopy. Worms were paralyzed with 25 mM sodium azide and mounted on microscope slides prepared with 5% agar pads and sealed with VALAP. RNAi-treated worms were imaged on a DeltaVision Elite inverted microscope equipped with an Olympus PLAN APO 60× (1.42 NA, PLAPON60XOSC2) objective, an Insight SSI 7-Color Solid State Light Engine and SoftWorx software (Applied Precision) using 1 µm z-stacks. For *elt-2* promoter experiments, images were collected using a Keyence BZ-X800 fluorescence microscope equipped with a Keyence PLAN APO 40x (0.95 NA, BZ-PA40) objective using 1 µm z-stacks. Imaging settings were kept constant between RNAi treatments and developmental stages per strain. Maximum Z projections were generated, and intestine fluorescence reporter was measured in ImageJ. Signal quantification was restricted to the intestine by utilizing a DIC reference image. Analysis and plots were generated using custom R scripts. Where ratios of reporter intensity are shown against controls with a p-value and confidence interval, a *t*-test of the ratio of means was performed using the *ttestratio* function in R package *mratios*, with the null hypothesis of a ratio of 1. Confidence intervals are 95% standard error around the ratio of means (Djira *et al.* 2020).

Single-Molecule RNA Fluorescence in situ Hybridization (smFISH)

smFISH was used to confirm intestine specificity of bioinformatically identified intestine genes. Transcripts were probed in the strain JM149 (*elt-2*P::GFP::H2B). Transcripts were probed in L1 developmental stage using smFISH as previously described (Parker *et al.* 2021) with minor modification. Briefly, samples were freeze cracked by adding 1 mL of pre-chilled acetone and incubated in liquid nitrogen for 1 min. Samples were then immediately transferred to a -20°C freezer for 10 min fixation. Post-fixation, samples were washed and hybridized with smFISH probes for *aat-6* or *pgp-1* transcript. Probes were designed using Biosearch Technologies RNA FISH custom probe designer and conjugated to CAL Fluor Red 610. All samples were imaged on a DeltaVision Elite inverted microscope described above.

RESULTS:

Genome-wide ELT-2 occupancy changes over developmental time

We set out to determine how genome-wide ELT-2 occupancy changed with developmental stage. To date, several ELT-2 ChIP-seq studies have been performed (Wiesenfahrt *et al.* 2015; Mann *et al.* 2016; Kudron *et al.* 2017). Of these, the modERN project conducted ELT-2 ChIP-seq in late embryo, L1, and L3 developmental stages using an anti-GFP antibody in strains expressing ELT-2::GFP driven from an integrated array (Figure 1A) (Kudron *et al.* 2017). Because an in-depth analysis of their findings had yet to be published, we capitalized on their data to determine the dynamics of ELT-2 occupancy over time.

Overall, replicates of ChIP-seq samples clustered hierarchically, depending on developmental stage and distinct from input controls (Figure 1B). We identified 11,015 total ELT-2 peaks over the entire dataset, of which 3,874, 4,948, and 9,990 peaks were identified at late embryo, L1, and L3 stages, respectively. Many peaks were shared between stages, and 20.6 % (2,265 peaks) were common to all three (Figure 1C). The majority of ELT-2 peaks occurred within 5 kb upstream to (51.6%) or within (27.6%) transcribed regions (Figure 1D). ELT-2 peaks were evenly distributed in the genome, proportional with chromosome size (Figure S1A) and showed a preference for chromosome centers mirroring the characteristic gene density of *C. elegans* (Barnes *et al.* 1995; Consortium* 1998)(Figure S1B).

To identify clusters of ELT-2 peaks with shared occupancy dynamics, we first filtered out non-dynamic peaks in which ELT-2's occupancy did not change over developmental time (See Supplemental

Methods) and retained peaks in which ELT-2's presence or peak height was dynamic. This retained 95% of identified peaks.

To organize ELT-2 peaks into occupancy clusters of shared dynamic behaviors, we performed K-means clustering. Four clusters yielded patterns that show intuitive developmental trends with ~78% sum of squares explained, suggesting high clustering efficacy (Figure 2A). This identified an "embryo" cluster with highest ELT-2 peak height at embryonic stages (276 peaks), a "larval cluster" with highest occupancy sustained through L1 and L3 stages (1,963 peaks), an "L3 cluster" with peak occupancy in the L3 stage (2,797), and an "increasing cluster" in which ELT-2 occupancy increased over time (5,428 peaks). An inspection of genome browser tracks visually verified the ChIP-seq binding dynamics (Figure 2B). Overall, the identified dynamic ELT-2 occupancy clusters represent locales of ELT-2 binding at which ELT-2 association changes over time in distinctive patterns.

We are interested in the degree to which ELT-2 binding dynamics correlates to mRNA abundance of nearby genes. To explore this, we first linked ELT-2 peaks to nearby genes using a nearest neighbor approach (Figure 1D, Supplemental Methods). Of the identified 7,006 associated genes, 77% (5,387 genes) mapped to a single type of ELT-2 occupancy cluster (Figure S3). For simplicity, genes associated with a single type of ELT-2 occupancy cluster were used in downstream analysis.

To determine whether ELT-2 occupancy clusters were associated with genes of differing biological functions for different dynamic clusters, we performed Gene Ontology (GO) enrichment analysis on annotated function, tissue expression, and phenotype (Angeles-Albores *et al.* 2016), (Figure S4). Overall, ELT-2 peaks in all clusters associated with genes from a wide array of GO categories without meaningful patterns emerging from any single group. Terms denoting specific intestinal functions were enriched in "larval occupancy" cluster. These include toxin and infection-related processes, as well as fat metabolism. This is expected, as the larval stages are the first in which the intestine is presented with food and exposed to pathogens. Tissue categories associated with ELT-2 peaks of different dynamic clusters were varied, representing a mix of not only intestine but neuronal, muscle and germline tissue categories. Since ELT-2 is not known to be expressed outside the intestine, this suggests a regulatory role of ELT-2 in the targeting of non-intestine genes.

Increasing ELT-2 occupancy correlates with increasing transcription of nearby genes in larval and adult stages

To determine how ELT-2 occupancy dynamics associate with transcriptional output, we assessed ELT-2 occupancy clusters for their expression dynamics, focusing on intestine-specific genes. Consistent with ELT-2's role as a transcriptional activator, it is possible that ELT-2 occupancy is *instructive* for transcriptional output with higher ELT-2 occupancy correlating with higher transcriptional output. If so, the "increasing cluster" of ELT-2 binding would be predictive of genes with increasing expression over time. Conversely, ELT-2 may only be permissive for expression, a logical interpretation should ELT-2 occupancy remain constant at intestine-expressed genes, independent of their dynamics. In this case, stage-specific expression would be regulated by other means. Finally, it is plausible that ELT-2 may not occupy many intestine genes suggesting their expression is indirect or independent of ELT-2.

To assess how intestine-specific genes behave over developmental time, we would ideally use an intestine-specific transcriptome dataset collected across developmental stages. However, this comprehensive resource does not exist. Instead, several intestine-specific RNA-seq datasets are available at one or, at most, two developmental stages (Pauli *et al.* 2005; Spencer *et al.* 2011; Blazie *et al.* 2015; Hashimshony *et al.* 2015). Conversely, high-resolution RNA-seq time course data is available on whole worms (Baugh *et al.* 2003; Boeck *et al.* 2016). We exploited these available resources by determining the set of reproducibly intestine-exclusive genes and filtering for their transcriptional dynamics out of whole-

worm time course assays (Boeck *et al.* 2016) (Figure 3A, Supplemental Methods). This resulted in 5,516 genes whose global expression dynamics reflect their intestine-specific expression due to their exclusion in all other tissues (File S4).

To determine the accuracy of our intestine-specific set of genes, we selected from our curated intestine-specific gene list *aat-1 and pgp-1* for visual inspection using single molecule Fluorescence In Situ Hybridization (smFISH). Both yielded high intestine enrichment validating our intestine-enrichment filtering procedure (Figure 3D).

Our set of intestine-specific genes was enriched for ELT-2 occupancy as 57.7 % contained an ELT-2 binding site in a proximal region (3,182 bound genes of 5,516 total), a greater fraction than ELT-2's binding genome-wide (33.9 %, 6,944 bound genes of 20,470 genes total) (Figure S5A). This high propensity illustrates that ELT-2 is an important intestinal regulator but is not the only factor driving intestinal expression.

To determine the relationship between intestine expression dynamics and ELT-2 occupancy categories, we compared the RNA-seq behaviors of the intestine-specific gene set to ELT-2 binding dynamics of nearby peaks. For simplicity, we restricted our analysis to genes that mapped to a single-ELT-2 occupancy cluster category (2,412 genes, 75.8 % of ELT-2 bound intestine-specific genes) with a majority bound by "increasing cluster" ELT-2 occupancy (Figure S5B). Hierarchical clustering was used to determine expression dynamics for intestine-specific, ELT-2-bound genes. This identified four distinct sets of intestine-specific genes classified based on the developmental stages corresponding to their highest relative expression: "embryonic", "L1", "post-embryonic" and "adult" expression (Figure 3B, S5C, Supplemental Methods).

To quantify the association between developmentally dynamic intestine-specific genes and ELT-2 occupancy statistically, we employed Chi-square statistics (Figure 3C). We found that genes with their highest relative mRNA abundance in embryos were under-represented in all four ELT-2 occupancy categories, suggesting that ELT-2 occupancy does ***not*** correlate with the expression of these genes. However, genes with peak expression in L1 stage, high expression in larval and adult stages (post-embryonic expression), or high expression in adult stages were associated with the ELT-2 occupancy of "increasing", "larval" and "L3" categories, respectively. This suggests ELT-2 occupancy is instructive for transcriptional output of intestine genes at these stages. Notably, genes with highest relative expression during the young adult stage were enriched for peaks from the "L3" occupancy cluster. This is interesting as we had no ChIP-seq datasets in adult worms, leading to the speculation that adding adult-specific ELT-2 ChIP-seq may contribute to a stronger trend. Overall, our results suggest that ELT-2 occupancy dynamics are predictive of gene expression dynamics in larval stages, but the relationship does not hold in embryonic stages. In the absence of a comprehensive dataset tracking the intestine transcriptome over time, our analysis serves as a working proxy for intestine-specific gene expression.

ELT-2 target genes are positively and negatively regulated

To better understand the role of ELT-2 on target gene transcriptional output, we employed a previous transcriptomics study that assessed gene expression changes upon *elt-2* loss (Dineen *et al.* 2018). The study included depletion of both *elt-2* and its paralog, *elt-7,* individually or in combination. ELT-2 and ELT-7 work in an unequal partnership to shape the intestinal transcriptome. Whereas *elt-2* loss-of-function is larval lethal due to collapsed intestinal lumens, *elt-7* mutant worms are superficially normal and only show a 7-minute delay in hatching compared to wild type (Fukushige *et al.* 1998; Sommermann *et al.* 2010; Dineen *et al.* 2018). In contrast, *elt-2; elt-7* double mutants have enhanced larval lethality characterized by a misshapen, discontinuous intestinal lumen (Figure 4A). This relationship implies that they each have some distinct roles, and that ELT-2 can direct intestine development in the absence of *elt-7,* whereas the reverse is not true. ELT-2 and ELT-7 are therefore said to be partially redundant, yet the

mechanistic details of their relationship are still unresolved, and it is unclear why ELT-2 is uniquely required.

Previous work measuring the transcriptional response of *elt-2* and *elt-7* deletion, either together or separately, identified six sets of gene responses (Figure S6A). A large set of genes are dependent on ELT-2 for their expression, in keeping with the model that ELT-2 acts principally as a transcriptional activator. However, a similarly large set of genes responds to *elt-2* loss with over-expression, suggesting they depend on ELT-2 for their negative regulation, either directly or indirectly. In that study, the authors hypothesized indirect effects, possibly even occurring external to the intestine (Dineen *et al.* 2018). However, several such genes did harbor ELT-2 ChIP-seq peaks at the L3 stage of development, raising the possibility that ELT-2 represses some targets directly.

Given that we now have the L1 stage ChIP-seq data, we wanted to determine the proportion of genes in each set associated with L1 stage ELT-2 binding. Overall, the results largely reinforced the previous findings conducted using L3 ELT-2 ChIP-seq. Genes that depended on ELT-2 and ELT-7 for activation individually (SET6, 36.0%, p-value = 1.7e-03) or in combination (SET4, 50.4%, p-value = 6.9e-05) had higher proportions of ELT-2 occupancy than ELT-2 occupancy across all expression sets (31.4%). Genes with implied ELT-2 or ELT-7 repression dynamics had lower propensities for direct ELT-2 binding (SET1, 17%, SET2, 21%). This suggested that the negative regulation imparted by ELT-2 was largely indirect. Conversely, the set with highest proportion of ELT-2 occupancy corresponded to genes that were apparently repressed by ELT-2 and "overcompensated" for ELT-7 (SET3, 53%, p-value = 4.7e-20). Together, these results suggest that repressed genes were either direct or indirect targets of ELT-2 regulation (Figure S6B).

To focus on direct transcriptional responses of ELT-2, we filtered the Dineen et al. RNA-seq dataset for genes that exhibited L1-stage ELT-2 occupancy (Kudron *et al.* 2017). To organize genes by their dependence on ELT-2 or ELT-7, we then performed hierarchical clustering on Z-score row-normalized read counts. This identified four response classes (Figure 4B, S5C) that largely resembled the previous analysis (Figure S6B) (Dineen *et al.* 2018). The first two target gene classes, A1 and A2, represent genes whose expression depended on ELT-2 with or without an ELT-7 impact for their activation. The remaining two target gene classes exhibited dynamics indicative of ELT-2 having a negative regulatory impact. Target Gene Class B genes became over-abundant in the absence of *elt-2* in an ELT-7-dependent manner, implying that ELT-7 is required for their over-expression. In contrast, Target Gene Class C became over-abundant in the absence of *elt-2* and remained high in *elt-7*'s absence, suggesting their overexpression required an as-yet unknown factor.

To assess whether members in each target class of ELT-2 and ELT-7 regulated genes were associated with distinct ELT-2 occupancy clusters, we merged this data to the ELT-2 occupancy clusters we established (from Figure 2) (Figure 4C) (Kudron *et al.* 2017; Dineen *et al.* 2018). Target Gene Class A1 was enriched for the "larval" ELT-2 occupancy cluster. These are genes that depended on ELT-2 for L1 stage activation, were independent of ELT-7 and had increased ELT-2 occupancy in both L1 and L3 stages. We found a lack of enrichment for either Target Gene Class A2 or Target Gene Class B genes with any ELT-2 occupancy cluster. This observation may be due to both classes being co-regulated by ELT-7. Finally, Target Gene Class C genes were most enriched in the "L3" ELT-2 occupancy cluster, suggesting that Target Gene Class C genes have higher ELT-2 occupancy at L3 stages when compared to L1 stage ELT-2 occupancy.

What accounts for different target genes organizing into different target gene classes with respect to ELT-2 and ELT-7 loss? It is possible that ELT-2 regulates genes indirectly through the activity of its downstream transcription factors. To explore this possibility, we filtered our set of ELT-2 target genes (Figure 4B) for transcription factors or genes with DNA binding activity (as annotated in wTF3.0) (Figure S7A) (Bass *et al.* 2016). We identified transcription factors in each of the four Target Gene Class categories, including *zip-10*, *ets-4*, and *pqm-1*. It is possible these TF's work in combination with ELT-2

and ELT-7 or that they transduce the regulatory impact of ELT-2 and ELT-7 to target genes indirectly within the intestine. To support this, we measured the number of ELT-2 target TFs with intestine-specific expression (File S4). We determined that 85% of ELT-2 regulated target TFs (23 of 27 total TFs) were annotated for intestine-specific expression (Figure S7A).

It is possible that genes in each target class distinguish themselves by the precise or extended TGATAA motif ELT-2 and/or ELT-7 occupy, a possibility supported by recent evidence that flanking sequences influence ELT-2's impact on transcriptional output (Lancaster and McGhee 2020). Alternatively, it is possible neighboring binding motifs of auxiliary TF's distinguish their impact. To determine whether we could identify hallmarks of such distinguishing DNA regulatory sequences, we performed motif analysis in the vicinity of ELT-2 peaks for each target gene class. We identified a robust TGATAA motif in each Target Gene Class with negligible differences in flanking nucleotide composition (Figure S7B). We also identified motifs matching the DNA binding sequence of transcription factor EOR-1 (AGAGACGGAGA) in Target Gene Class A1, B, and C genes near ELT-2 peaks (Figure S7C). Finally, we identified a motif in Target Gene Class A1 genes, matching the DAF-12 DNA binding site ($[GT]_7$) (Figure S7D). Together these data suggest that ELT-2, EOR-1, and DAF-12 may work together to regulate ELT-2 target genes with differential transcriptional outcomes.

ELT-2 negatively regulates non-intestine genes

Since the transcriptomic response to *elt-2* and *elt-7* depletion, together and separately, was measured in whole worms, we wanted to determine if genes in each Target Gene Class were enriched for intestinal tissue origin. To determine this, we performed *C. elegans* Tissue Ontology enrichment analysis (TEA) for genes within each Target Gene Class (Angeles-Albores *et al.* 2016). Tissue ontology terms were considered significant if more than 5 genes from the input set associated with each term and if they registered a hypergeometric test adjusted p-value of less than 0.05 (Figure 5A). Target Gene Class A1 and A2, the two classes that are activated by ELT-2 and ELT-7, were associated with intestine expressed genes. Target Gene Class A2 was additionally over-represented in head mesodermal cell, muscular system, PVD, and outer labial sensillum categories. In contrast, the two target gene classes that illustrated a negative regulatory effect by ELT-2 were not enriched in intestinal tissue terms, suggesting that ELT-2 is successful in its repressive role in enough of these genes to reduce their levels within the intestine. Target Gene Class B was over-represented in categories associated with cephalic sensillum, anterior ganglion, and cephalic sheath cell expression, all neuronal associated to some degree. Target Gene Class C was over-represented for genes associated with germline, reproductive system, corpus, reproductive tract, and AVK neuron expression, tissues found in germline, pharynx, and neurons.

To further probe the relationship between a gene's dependence on ELT-2 and its intestine-specificity, we compared the target gene classes to time-course RNA-seq data that we striated for intestine-specific versus non-intestinal categories in Figure 3 through Chi-square statistics (Figure 5B). This analysis revealed that Target Gene Class A1 was strongly associated with L1-expressed genes, as expected since this analysis was restricted to the L1 stage. Target Gene Class A2 genes did not strongly associate with any expression cluster. Target Gene Class B was associated with post-embryonic intestine expressed genes, suggesting that the identified neuronal genes in Target Gene Class B may have intestine-specific expression at developmental stages beyond L1. Target Gene Class C was strongly associated with non-intestinal expression. Together, these findings suggest that ELT-2 plays a dual role in both positive expression of intestine-associated genes and repressing expression of non-intestinal genes within the intestine.

ELT-2 negatively regulates expression of the neuronal transcription factor CEBP-1 in the intestine

Our analysis suggests that ELT-2 promotes intestine-specific genes within the intestine while suppressing genes whose activity is associated with other tissues. The gene *cebp-1* (*CCAAT/enhancer-binding protein*) is among the genes putatively repressed by ELT-2 for the purpose of decreasing inappropriate neuronal characteristics in intestinal cells. *cebp-1* encodes a bZIP class transcription factor expressed in hypodermis, pharynx, muscle, intestine, and neuronal cells where it is essential for axon regeneration, stress response, and intestinal immune response (Yan *et al.* 2009; McEwan *et al.* 2016; Yang *et al.* 2017). The *cebp-1* promoter is bound by ELT-2 with increasing occupancy over developmental time and *cebp-1* mRNA becomes more abundant when *elt-2* is absent (Figure 6A). We wanted to determine whether the over-expression of *cebp-1* upon loss of *elt-2* occurred throughout the body, suggesting an indirect response to stress or injury, or whether *cebp-1* expression was intestine-specific and could further support the model that *cebp-1* over-expression was due to the relief of direct ELT-2 repression.

To determine if *cebp-1* over-expression was intestine-specific, we used CEBP-1::GFP worms depleted of ELT-2 using *elt-2* RNAi and imaged by fluorescence microscopy. ELT-2 knockdown was validated by observing a reduction in ELT-2::GFP signal compared to control worms (Figure 6D, E). We observed a clear increase in CEBP-1::GFP reporter activity specifically in intestine cells in response to ELT-2 depletion (Figure 6B, C), further supporting the model that ELT-2 plays an active role in repressing neuronal programs in intestinal cells.

ELT-2 negatively regulates its own promoter

ELT-2 regulates its own expression, and the occupancy of ELT-2 protein at its own gene promoter has been extensively characterized (Fukushige *et al.* 1999; Wiesenfahrt *et al.* 2015). The reported purpose of this regulation is to sustain ELT-2 expression throughout the worm's lifetime through autoactivation until it eventually declines in old age. Previous studies were unable to discern how *elt-2* promoter output responded to *elt-2* depletion as the mutations and rescue constructs used to manipulate ELT-2 confounded that analysis (Dineen *et al.* 2018). However, by using an *elt-2* promoter reporter transgene driving histone H2B fused to GFP, we could study *elt-2* promoter activity divorced from production of ELT-2 protein (Figure 7A). If ELT-2 autoactivates its own promoter, we expect to observe reduced *elt-2* promoter activity upon ELT-2 depletion. Conversely, we observed ~5-fold over-expression of GFP upon depletion of ELT-2 protein, suggesting that ELT-2 negatively regulates *elt-2* promoter activity (Figure 7B-D).

This observation resembled the response of Target Gene Class B and C genes (Figure 4B). To determine if ELT-2 falls into Target Gene Class B or C genes, we measured *elt-2* promoter activity in an *elt-7* genetic deletion background. We observed that the *elt-2* promoter is no longer overactivated when both ELT-2 and ELT-7 are absent (Figure 7B-D), thereby determining ELT-2 is itself within Target Gene Class B. Additionally, we observed this phenomenon is not restricted to the L1 stage but similarly observed in L3 and adult stages (Figure S8).

Together, these results suggest that ELT-2 protein negatively regulates its own promoter. Additionally, it suggests that ELT-7 is responsible for over-activating the *elt-2* promoter when relieved of ELT-2 negative control. Intriguingly, *elt-2* is still expressed to a low level when both ELT-2 and ELT-7 are absent, implying that an additional factor contributes to *elt-2* transcription (Figure 7C, D). Together, these findings show that the wildtype level of *elt-2* mRNA production is affected by both positive and negative inputs. This example further emphasizes combinatorial control by both ELT-2 and ELT-7, and illustrates that an additional unknown factor participates in the intestine GRN.

DISCUSSION:

ELT-2 is the major transcription factor that marks intestinal tissue. ELT-2 production commences in embryos where its role is well studied, yet it remains present during larval and adult stages where its impact is more nebulous. Indeed, larval and adult stage intestines perform digestion and metabolism characteristic of most animal intestines, but also additional functions are overseen by separate organs in other animals, such as yolk production, insulin signaling, regulation of developmental aging, immune response, stress response, and detoxification (Kimble and Sharrock 1983; An and Blackwell 2003; Libina *et al.* 2003; Martinez-Finley and Aschner 2011; Ludewig and Schroeder 2013; Head and Aballay 2014; Block *et al.* 2015; Keith *et al.* 2016; Mann *et al.* 2016; Chun *et al.* 2017; Lee and Mylonakis 2017; Zárate-Potes *et al.* 2020). ELT-2 is required for optimal performance of many of these sub-functions, but whether the role of ELT-2 with respect to these processes is direct or indirect remains unknown in most cases.

Our goal was to determine whether ELT-2's genome-wide occupancy remained consistent over developmental time or changed in stage-specific ways associated with the intestine's changing transcriptome. This work nicely synergized with previous studies that characterized transcriptome changes due to *elt-2* loss but could not differentiate between direct and indirect effects without the binding information provided here (Dineen *et al.* 2018).

ELT-2 shows hallmarks of the instructive mode of gene regulatory control

In envisioning how ELT-2 might associate with its genome-wide targets, we focused on three major hypotheses. In the instructive model, a correlation between dynamic ELT-2 occupancy and changing transcriptional output would suggest that increased ELT-2 binding could promote higher expression of target genes. In contrast, the permissive model predicted ELT-2's static presence at target genes would be independent of their output, suggesting ELT-2 could mark intestine genes as competent for activation, but robust transcription could only be achieved in combination with stage-specific components. Finally, the upstream model posited that ELT-2 would occupy a small minority of intestine-specific genes with dynamics largely driven by downstream TFs or those independent from its influence. Though these models were simplifications and not strictly mutually exclusive, they helped to direct our thinking and data interpretation.

Overall, we identified ELT-2 ChIP-seq peaks that were differentially changing over developmental time. A greater number of ELT-2 peaks were recovered at the L3 stage than in either the embryonic or L1 stages. This led to most ELT-2 peaks organizing into dynamic occupancy clusters in which ELT-2 was maximal in the L3 stage either abruptly ("L3 cluster") or after trending upward ("increasing cluster"). It is noteworthy that places of ELT-2 residence often illustrated lower or sub-threshold ELT-2 peaks at other time points (Figure S2), possibly suggesting that regions of ELT-2 binding have distinguishing characteristics such as low nucleosome occupancy or low levels of ELT-2 association.

We found a positive association between increasing ELT-2 occupancy and increasing target gene expression, supporting the instructive model of ELT-2 behavior, at least in larval stages. However, in embryonic stages, this trend did not hold as ELT-2 peaks with maximal occupancy in the embryo were not predictive of high embryonic transcription. We assessed these associations through an imperfect method by filtering mRNA dynamics for a subset of genes with well-defined intestine-specificity. If, instead, we could have used a true intestine-specific RNA-seq time course dataset, it is possible that a pattern would emerge for the embryo-specific stages, though it is impossible to speculate which it would be. Through a separate analysis, we also characterized roughly 400 genes targeted by ELT-2 that did not follow any of our predicted patterns of ELT-2 activation, and instead were subject to negative regulation by ELT-2.

Admittedly, it is challenging to interpret the meaning of differential ELT-2 occupancies as the underlying molecular characteristics captured by ChIP-seq peak height are still unclear. Evidence supports a general model that TFs bind to DNA sequence motifs in favorable chromatin contexts where they recruit RNA Polymerase II or transcription preinitiation complex (PIC) components to stimulate transcription of nearby genes (Haberle and Stark 2018). Though higher occupancy levels are linked, in many cases, to higher transcriptional output (Senecal *et al.* 2014), TF occupancy is often studied as a binary output. Indeed, interpreting the molecular meaning of differential ChIP-seq signals can be complicated. Increasing ChIP-seq occupancy may reflect a greater amount of TF present at the genomic locus within the sample, or it could represent a more favorable conformation for crosslinking or recovery. Also, ChIP-seq assays are performed on large populations of cells and are therefore unable to disentangle a uniform increase in TF occupancy across a population from fluctuations of TF occupancy in different sub-populations either due to timing dynamics, diverse cell types, or varying cell states. In these cases, temporal fluctuations that result in longer TF residency time can manifest as higher TF occupancy within the population. To further complicate matters, robust TF occupancy alone is sometimes not sufficient for transcriptional activation. An association with other factors, a critical phosphorylation event, or TF dynamics may be required to stimulate recruitment of RNA Polymerase II following successful TF binding (Meijsing *et al.* 2009; Ucar *et al.* 2009; Lickwar *et al.* 2012; Kulik *et al.* 2021).

Despite these uncertainties, we did find positive associations between ELT-2 occupancy and the transcriptional output of nearby genes in larval stages, suggesting that a cause or effect relationship exists. We are interested in the molecular characteristics that underlie ELT-2's behavior. We predict that changing cooperative partners, different flanking sequences in the DNA motifs, or an evolving chromatin landscape may influence ELT-2's dynamics. Indeed, transcriptional output of ELT-2 target genes increases with increasing ELT-2 affinity for XXGATAXX motifs (Lancaster and McGhee 2020). We took the first steps in exploring these potential models. Using computational approaches, we identified that the TFs DAF-12 and EOR-1 occur in close association to ELT-2 peaks, and it will be interesting to further investigate their relationship. Furthermore, we assessed the sequence motifs underlying and flanking ELT-2-bound loci at different stages but were unable to ascertain meaningful nucleotide differences above a reasonable significance. Finally, an analysis of changing chromatin landscapes, though interesting, would require a dataset of chromatin profiling specific to intestinal tissue at different developmental stages. Even if such a resource existed, it would be challenging to disentangle cause from effect relationships between chromatin landscape and ELT-2 binding, and this analysis remains to be conducted.

ELT-2 negatively regulates a subset of target genes

To date, experimental evidence has suggested that ELT-2 is the intestine master regulator largely responsible for intestine gene activation (Fukushige *et al.* 1998, 2003, 2005; Kalb *et al.* 2002; Oskouian *et al.* 2005; McGhee *et al.* 2007, 2009; Sommermann *et al.* 2010). A previous transcriptional study reported that ELT-2 negatively regulates two subsets target genes but was unable to discern direct from indirect targets (Dineen *et al.* 2018). Here, we combined that study with ChIP-seq data from the appropriate developmental stage to verify that ELT-2 negatively regulates 400 of its direct targets (Figure 4B). Of these, we verified that the target gene *cebp-1* is negatively regulated by ELT-2 within the intestine. Given the role of CEBP-1 in neuronal differentiation, we posit that ELT-2 prevents neuronal characteristics from developing by repressing this gene. In a similar vein, a previous investigation revealed that depletion of *elt-2* and *elt-7* results in de-repression of the valve-cell-specific gene *cdf-1* within intestinal cells (Sommermann *et al.* 2010), suggesting this may be a general mechanism to protect the intestine from ectopic cellular characteristics.

We hypothesize that negative regulation by ELT-2 could occur through competition within the intestinal GRN or direct repression by ELT-2 (Figure 8). In our dataset, Class B Target Genes were over-

expressed upon *elt-2* loss but depended on ELT-7 for that over-expression. Their behavior suggested that ELT-7 was responsible for activating Class B genes in ELT-2's absence. In contrast, Class C genes were ELT-7 independent, suggesting a missing component of the intestinal GRN was responsible for their activation. On the one hand, it is possible that competition between ELT-2, ELT-7, and the unknown TF could account for these behaviors, especially given ELT-2 and ELT-7's shared DNA binding specificity (Wiesenfahrt *et al.* 2015; Dineen *et al.* 2018). That is, ELT-2 may out-compete ELT-7 when present to activate transcription of target genes, albeit at a lower level than ELT-7 could (Figure 8A). In this case, loss of ELT-2 would result in a hyper-activation of target genes by ELT-7. Alternatively, ELT-2 may negatively regulate transcription through direct repression, likely in association with repressive complexes or repressive chromatin regulators (Figure 8B).

Interestingly, a precedent for GATA factors acting as dual activators and repressors has been established in both yeast and mammals, and examples of competition and direct repression exist in those systems (Bresnick *et al.* 2010; Zheng and Blobel 2010; Block and Shapira 2015; Fujiwara 2017). Over-expression of human GATA factors GATA-4, GATA-5, and GATA-6 can inhibit tetracycline-catalyzed induction by out-competing higher activity tetracycline activators due to naturally occurring GATA sites within the tetO promoter (Gould and Chernajovsky 2004). In contrast, numerous examples of direct repression have been documented. In humans, GATA factors catalyze cell differentiation, can act as both tumor suppressors or oncogenes, and serve as critical markers of cancer onset and progression (Fujikura *et al.* 2002; Bresnick *et al.* 2010; Chou *et al.* 2010; Zheng and Blobel 2010; Fujiwara 2017). Human GATA-1 positively regulates subsets of direct targets through the recruitment of histone acetyltransferase complexes (CBP/p300) (Blobel *et al.* 1998) but represses others through PU.1, the co-repressor FOG-1, or the NuRD co-repressor complex (Hong *et al.* 2005; Burda *et al.* 2010). In yeast, GATA factors function as metabolic switches responding to changing levels of amino acid availability. Remarkably, the yeast GATA factor Gaf1 responds to amino acid depletion by activating amino acid biosynthetic genes through RNA Pol II activation while simultaneously repressing tRNA transcription through RNA Pol III repression (Rodríguez-López *et al.* 2020). While there is ample evidence for both competitive relationships and direct repression by GATA factors in other systems, further experimentation will be required to determine which applies to ELT-2 and ELT-7.

Developmental implications for ELT-2's dual role in gene regulatory control

We hypothesize that ELT-2's dual role as an activator and repressor is critical to its role in promoting intestinal cell fate. ELT-2 could promote intestine differentiation by simultaneously activating intestinal programs while suppressing non-intestinal ones. Ectopic expression of ELT-2 is competent to induce transdifferentiation (redirection of one cell fate to another evidenced by molecular markers) of pharynx and transorganogenesis (redirection of organ development from one type to another) of the somatic gonad into intestine-like tissues after the multipotency-to-commitment transition (Riddle *et al.* 2013, 2016). These findings are consistent with our observations that ELT-2 negatively regulates pharynx and germline associated genes in Target Gene Class C. In the pharynx, ectopic expression of either *end-3*, *elt-7,* or *elt-2* leads to pharynx expression of intestine terminal differentiation marker *ifb-2*, detection of birefringent gut granules, and loss of pharynx specific markers *ceh-22* and *myo-2* but does not change shape into intestine (Riddle *et al.* 2013, 2016). In a more dramatic scenario, pulsed ectopic ELT-7 expression in L3 or L4 stage worms leads to stable ELT-2 expression in the uterus and spermatheca tissues of the hermaphrodite somatic gonad. The somatic gonad morphs into an intestine-like endotube marked by intermediate filament *ifb-2* and highly organized microvilli, abolishing normal uterus function (Riddle *et al.* 2016). Interestingly, muscle and neuron tissues were not competent for transdifferentiation or transorganogenesis, which were tissues that were similarly negatively regulated among Class B and C Target Genes. It is interesting to speculate whether the process of transorganogenesis depends on repressive TF activities to canalize cells into the new tissue identity.

Our results support the idea that ELT-2 may play a direct repressive function. Were this the case, ELT-2 would likely interact with co-repressors or repressive machinery in the cell. Of these repressive complexes, the DREAM complex is required to repress germline genes within the intestine, including the P-granule gene *pgl-1.* In *C. elegans*, the DREAM complex (dimerization partner, RB-like, E2F, and multi-vulval class B) is a chromatin regulator of germline genes that is active within somatic cells (Wang *et al.* 2005; Petrella *et al.* 2011) in which disruption leads to ectopic expression of germline genes preferentially in the intestine. The ability of both ELT-2 and DREAM complex to repress germline genes within the intestine suggests their relationship warrants further investigation

We found that ELT-2 protein is capable of negatively regulating its own promoter. Previous work established that ELT-2 binds its own promoter, measured through microscopy, ChIP-seq, and in vitro binding assays (Hawkins and McGhee 1995; Wiesenfahrt *et al.* 2015). The finding that ELT-2 regulates itself through positive feedback was established through ectopic expression of ELT-2 leading to ectopic *elt-2* promoter reporter activity (Fukushige *et al.* 1998; Wiesenfahrt *et al.* 2015). In seemingly direct contrast, we found that loss of ELT-2 led to an upregulation of *elt-2* promoter reporter activity in the intestine. Additionally, we identified that simultaneous depletion of ELT-2 and ELT-7 reduced ELT-2 promoter activity back to wildtype levels (Figure 7). This suggested that ELT-7 is primarily responsible for the *elt-2* promoter's over-expression in the absence of ELT-2. Since ELT-2 can ectopically activate ELT-7 (Sommermann *et al.* 2010), we suggest that the initial evidence for ELT-2 positive feedback may have been observed indirectly through ELT-7 activation by ELT-2. Finally, the observation that *elt-2* promoter activity is wild-type in the absence of both ELT-2 and ELT-7 protein suggests another TF may contribute to its sustained regulation (Dineen *et al.* 2018).

Evolutionary implications in the intestinal GRN

Ultimately, ELT-2 works within the context of a larger GRN that shapes the changing intestinal transcriptome. It is interesting that multiple TFs in the intestine bind the same DNA binding sequence centered around GATA nucleotides. The intestine-specific GATA factors, END-1, END-3, MED-1, MED-2, ELT-2, and ELT-7 as well as the non-GATA factor PQM-1 all bind these similar motifs (Hawkins and McGhee 1995; Lowry and Atchley 2000; Broitman-Maduro *et al.* 2005; Tepper *et al.* 2013). Their convergence provides ample opportunities for feed-forward loops, functional overlap, and robustness (Maduro *et al.* 2005a; Raj *et al.* 2010; Boeck *et al.* 2011; Maduro 2019). We are interested in how these networks morph over developmental time and why ELT-2 is uniquely required within this system. The evolutionary implications for how these networks are selected and maintained are also of paramount interest within the field.

It is intriguing that the intestinal GRN is comprised of successive pairs of GATA factors that turn on each downstream pair during embryogenesis through both immediate and feed-forward mechanisms. Of these, the MED-1/MED-2 duo and the END-1/END-3 duo arose from gene duplication events (Gillis *et al.* 2008) whereas ELT-2 and ELT-7 are more distantly related. Specifically, both the END-1/END-3 and ELT-2/ELT-7 pairs are examples of partial redundancy, an asymmetrically overlapping TF relationship that is widespread but not well understood (Sommermann *et al.* 2010; Dineen *et al.* 2018). Partially redundant TFs predominate in *C. elegans* cell fate specification and are common across metazoans (Woollard 2005; Boeck *et al.* 2011; Tintori *et al.* 2016). For example, the yeast stress regulators Msn2 and Msn4 activate transcription with distinct dynamics either together and apart, such that one acts as a switch and the other fine-tunes gene expression levels (Schmitt and McEntee 1996; AkhavanAghdam *et al.* 2016). Similarly, the mammalian immune response factors STAT5A and STAT5B and the members of the NFAT family overlap in unequal ways to regulate gene expression outputs (Yissachar *et al.* 2012; Villarino *et al.* 2016).

Networks that employ overlapping paralogs use them to buffer against noise and to promote developmental robustness. The END1/3 and MED1/2 gene pairs of the embryonic intestinal GRN guard against phenotypic stochasticity (Choi *et al.* 2017) and lead to more precise timing of ELT-2 expression

(Nair *et al.* 2013). In essence, these overlapping TFs work together to ensure the stereotyped embryonic development for which *C. elegans* is so widely known. ELT-2 and ELT-7 may similarly contribute to system-wide robustness and precision. However, other relationships between them are also possible. Unlike END1/3 and MED1/2, ELT-2 and ELT-7 are more distantly related, and ELT-2 contains a longer, extended N-terminus, thereby allowing for additional functionality. Altogether, the *C. elegans* intestinal GRN continues to be an intriguing system. It will be important to understand how asymmetrically overlapping TFs are evolutionarily shaped and how they impact development. Future insights into how the system changes over the course of developmental progression will also lend valuable insight to how GRNs morph throughout an organism's lifespan.

ACKNOWLEDGMENTS:

We thank Dr. James McGhee, Michelle Kudron, Valerie Reinke, and Wormbase for protocols, equipment, reagents, strains, advice, productive discussions, feedback regarding the manuscript, and commentary. Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). This work utilized resources from the University of Colorado Boulder Research Computing Group, which is supported by the National Science Foundation (awards ACI-1532235 and ACI-1532236), the University of Colorado Boulder, and Colorado State University. This work utilized microscopy resources from NIH grant 1S10 OD025127 and support from the CSU Microscope Imaging Network. Some figure elements were created in BioRender.

COMPETING INTERESTS:

The authors declare that they have no competing interests.

FUNDING:

Funder	Grant reference number	Author
National Institutes of Health	R35GM124877	PI: Erin Osborne Nishimura
Boettcher Webb-Waring		PI: Erin Osborne Nishimura
Bridge to Doctorate at Colorado State University	1612513	Robert T.P. Williams

Any opinions, findings, conclusions or recommendations expressed are those of the authors and do not reflect the views of the funding agencies. The funding had no role in the generation of this manuscript.

AUTHOR ORCIDS:

David C. King	https://orcid.org/0000-0003-1601-1612
Robert T.P. Williams	https://orcid.org/0000-0002-3384-212X
Izabella R. Mastroianni	https://orcid.org/0000-0003-3281-2807A
Jessica Hill	https://orcid.org/0000-0002-3323-8379
Nicolai W. Apenas	https://orcid.org/0000-0003-1134-1890
Gabriela Ramirez	https://orcid.org/0000-0002-2818-6711
Erin Osborne Nishimura	https://orcid.org/0000-0002-4313-4573

FIGURE LEGENDS:

Figure 1. ELT-2 target specificity at three key developmental stages. (A) Experimental design. ELT-2 ChIP-seq was performed at three different stages of development: Late Embryo (LE), Larval Stage 1 (L1) and Larval Stage 3 (L3) (Kudron *et al.* 2017). (B) Signal correlation between samples. Correlation matrix illustrates ELT-2 ChIP-seq replicates clustered together (Pearson correlation greater than 0.91) indicating reproducibility (Table S2). Input controls clustered together independent of stage. (C) Peaks called reproducibly for each stage. Intersections are formed by overlapping regions being called in more than one stage. (D) Peak-to-gene mapping. The pie chart describes the number and proportion of peaks mapping to the indicated region of an annotated gene, as described by the diagram below it.

Figure 2. ELT-2 binding patterns cluster into four major classes based on binding dynamics. (A) Dynamically changing ELT-2 ChIP-seq peaks were organized by k-means clustering based on their relative dynamics across the three developmental time points. Row-wise normalization (subtracting the row mean and dividing by row standard deviation) yields ELT-2 ChIP-seq peak enrichment over background. (B) Genome browser tracks illustrate examples of representative genes harboring peaks from each of the four clusters: *lin-48* of the embryo cluster; *clec-44* in the larval cluster; K01F9.2 in the L3 cluster; *aat-6* of the increasing cluster. ELT-2 ChIP-seq peak regions are shown as brown rectangles. Replicates 1 and 2 are illustrated in two different colors with overlapping regions in green. Incidents of TGATAA are depicted as horizontal lines. Gene models are in blue below each track. All tracks are set to the same scale.

Figure 3. Increasing ELT-2 occupancy correlates with increasing transcription of nearby genes in larval and adult stages. (A) Summary of data processing analysis. Whole worm transcriptomic data was subset for intestine-specific genes (Boeck *et al.* 2016). Intestine-specific genes were compiled from embryo (Spencer *et al.* 2011; Hashimshony *et al.* 2015), L2 (Spencer *et al.* 2011), L4 (Pauli *et al.* 2005) and mixed stage (Blazie *et al.* 2015) datasets. (B) *(left)* Developmentally dynamic intestine expressed genes organize into four clusters. Hierarchical clustering was based on Euclidian distance of expression values for Z-score normalized transcript abundance between developmental stages using the ward.D2 method. *(right)* Annotation of ELT-2 dynamic occupancy cluster type associated with each gene. (C)

Enrichment measurement of genes in each intestine expression category to ELT-2 occupancy cluster with Chi-square statistics. Pearson residuals were utilized to evaluate if a given combination was enriched or depleted in a single ELT-2 occupancy type relative to all other genes in the analysis. One-tailed p-values were calculated for Pearson residuals and adjusted for multiple test correction with the Benjamini-Hochberg method. (D) Intestine specificity for genes identified by the filtering scheme outlined in (A) was verified by smFISH. Intestine nuclei were visualized through reporter *elt-2promoter::H2B::GFP* (green). smFISH probes targeting *pgp-1* mRNA *(left)* or *aat-6* mRNA *(right)* were visualized in fixed L1 stage animals (magenta). DNA (DAPI, blue) is also shown. Scale bars represent 10 microns.

Figure 4. ELT-2 target genes are positively and negatively regulated. (A) Diagram of loss-of-function phenotypes observed in *elt-7* and *elt-2* mutants either together or separately. (B) *(left)* Whole L1 transcriptomic data was measured in WT (N2), *elt-7 (tm840)*, *elt-2 (ca15)*, and *elt-7 (tm840); elt-2 (ca15)* genetic backgrounds (Dineen *et al.* 2018). Genes were subset for differential transcript abundance in any pairwise comparison and further filtered for L1 stage ELT-2 occupancy measured through ChIP-seq (n = 973 genes). Transcript abundance was Z-score normalized across genetic backgrounds (Row normalized read counts). Hierarchical clustering of Z-score normalized transcript abundance was performed using complete-linkage clustering of Spearman correlation distances. This identified four gene clusters of transcriptional response to deletion of either *elt-7* mutation alone, *elt-2* mutation alone, or *elt-7* and *elt-2* double mutation. Identified clusters are titled Target Gene Class A1, Target Gene Class A2, Target Gene Class B, and Target Gene Class C. *(right)* Annotation of L1 stage ELT-2 ChIP-seq peaks and ELT-2 occupancy type for individual transcripts. (C) Enrichment measurement of genes in each intestine expression category to ELT-2 occupancy cluster with Chi-square statistics, as in Figure 3.

Figure 5. ELT-2 negatively regulates non-intestine genes. (A) Tissue Ontology terms measured for genes in each Target Gene Class identified in Figure 4. Tissue ontology terms were considered significant if more than 5 genes from the input set associated with each term and if the term registered a hypergeometric test adjusted p-value (q-value) of less than 0.05. Multiple test correction was performed with the Benjamini-Hochberg method. Tissue ontology terms were ranked by performing -log10 transformation of the q-value. Point size represents the number genes in the input set that corresponds to the given Tissue Ontology term. General tissue terms were manually assigned post-hoc, encompassing intestine (green), pharynx (brown), muscle (purple), neurons and support cells (blue) and germline (red) tissue systems. (B) Enrichment measurement of genes in each Target Gene Class to genes in the Intestine Development Clusters identified in Figure 3 using Chi-square statistics, as in Figure 3.

Figure 6. ELT-2 negatively regulates intestine expression of neuronal transcription factor CEBP-1 in the intestine. (A) Genome browser tracks of the *cebp-1* genomic locus. ELT-2 ChIP-seq tracks display increasing ELT-2 occupancy over developmental time at the ELT-2 peak region (pink) that is located within the *cebp-1* promoter. *(below)* RNA-seq tracks display increased transcript abundance in *elt-2 (ca15)* genetic background compared to wildtype (WT, N2) (Dineen *et al.* 2018). (B) Representative image of CEBP-1::GFP signal (wgIs563 [*cebp-1promoter::cebp-1::GFP*] allele) (green) in immobilized L1 stage animals under control RNAi (L4440, below) or *elt-2* RNAi (above) treatments. Differential interference contrast microscopy (DIC) is also shown. Small puncta in control RNAi image corresponds to birefringent gut granules. (C) Quantification of CEBP-1::GFP signal in *C. elegans* intestines. (control RNAi n = 52 worms, *elt-2* RNAi n = 63 worms, p-value = 2.3e-07 as calculated by Student's t-test). (D) Representative images of ELT-2::GFP reporter (wgIs56 [*elt-2promoter::ELT-2::GFP*] allele) (green) in immobilized L1 stage animals during control RNAi (L4440, below) or *elt-2* RNAi (above) treatments. Small puncta in elt-2 RNAi image corresponds to birefringent gut granules. **(E)** Quantification of ELT-2 signal in *C. elegans* intestines. (control RNAi n = 94 worms, *elt-2* RNAi n = 120 worms, p-value = 2.6e-

03 as calculated by Student's t-test). For (C) and (E), GFP measurements were collected from background subtracted maximum Z projections from three biological replicates. Intensity measurements were normalized for intestine area. Box and whisker plots represent the distribution of data points. All points represent a single worm; filled points represent outliers.

Figure 7. ELT-2 negatively regulates its own promoter. (A) Diagram of *elt-2* promoter reporter (caIS71 [*elt-2promoter::H2B::GFP*]) used in this analysis. (B) Representative images of *elt-2* promoter activity in immobilized L1 stage animals. *elt-2* promoter reporter fluorescence was measured in *elt-7* wildtype or *elt-7 (tm840)* loss-of-function genetic background. Control RNAi (L4440) or *elt-2* RNAi depletion were performed. Consistent imaging exposure was used between treatments and genetic backgrounds. The intestine is outlined by a white dotted line. (C) Quantification of the *elt-2* promoter reporter's fluorescence intensity. This analysis included intestine GFP signal quantification, background subtraction, and area normalization using ImageJ. Student's t-test was used to determine statistical significance of measured fluorescence signal. Data represents 30 worms per treatment for three biological replicates. Box and whisker plots display the data point spread. Individual measurements are overlaid as points. (D) Comparison of *elt-2* promoter reporter activity for *elt-2* RNAi conditions with or without *elt-7* activity. Horizontal dotted line indicates relative fluorescence of 1. Error bars represent the t-test 95% confidence interval for the ratio of means between *elt-2* RNAi and control RNAi measurements.

Figure 8. Potential models explaining Target Class B Gene regulation. (A) Target Class B Genes are direct targets of ELT-2 that become over-expressed when *elt-2* gene products are depleted in an ELT-7 dependent fashion. Under wildtype conditions, ELT-2 may out-compete ELT-7 for access to the promoter. When *elt-2* is absent, it is possible that ELT-2 recruitment could lead to elevated, unregulated levels of RNA Polymerase II transcription. (B) Alternatively, ELT-2 and ELT-7 may have additive effects on the transcription of their shared target genes with ELT-2 repressing transcription and ELT-7 activating it. In *elt-2's* absence, its repressive role would be alleviated leading to greater transcriptional output.

REFERENCES:

AkhavanAghdam Z., J. Sinha, O. P. Tabbaa, and N. Hao, 2016 Dynamic control of gene regulatory logic by seemingly redundant transcription factors. Elife 5: e18458. https://doi.org/10.7554/elife.18458

Alon U., 2007 Network motifs: theory and experimental approaches. Nat Rev Genet 8: 450–461. https://doi.org/10.1038/nrg2102

An J. H., and T. K. Blackwell, 2003 SKN-1 links C. elegans mesendodermal specification to a conserved oxidative stress response. Gene Dev 17: 1882–1893. https://doi.org/10.1101/gad.1107803

Angeles-Albores D., R. Y. N. Lee, J. Chan, and P. W. Sternberg, 2016 Tissue enrichment analysis for C. elegans genomics. Bmc Bioinformatics 17: 366. https://doi.org/10.1186/s12859-016-1229-9

Barnes T. M., Y. Kohara, A. Coulson, and S. Hekimi, 1995 Meiotic recombination, noncoding DNA and genomic organization in Caenorhabditis elegans. Genetics 141: 159–79.

Bass J. I. F., C. Pons, L. Kozlowski, J. S. Reece-Hoyes, S. Shrestha, *et al.*, 2016 A gene-centered C. elegans protein–DNA interaction network provides a framework for functional predictions. Mol Syst Biol 12: 884. https://doi.org/10.15252/msb.20167131

Baugh L. R., A. A. Hill, D. K. Slonim, E. L. Brown, and C. P. Hunter, 2003 Composition and dynamics of the Caenorhabditis elegans early embryonic transcriptome. Development 130: 889–900. https://doi.org/10.1242/dev.00302

Blazie S. M., C. Babb, H. Wilky, A. Rawls, J. G. Park, *et al.*, 2015 Comparative RNA-Seq analysis reveals pervasive tissue-specific alternative polyadenylation in Caenorhabditis elegans intestine and muscles. Bmc Biol 13: 4. https://doi.org/10.1186/s12915-015-0116-6

Blobel G. A., T. Nakajima, R. Eckner, M. Montminy, and S. H. Orkin, 1998 CREB-binding protein cooperates with transcription factor GATA-1 and is required for erythroid differentiation. Proc National Acad Sci 95: 2061–2066. https://doi.org/10.1073/pnas.95.5.2061

Block D. H. S., K. Twumasi-Boateng, H. S. Kang, J. A. Carlisle, A. Hanganu, *et al.*, 2015 The Developmental Intestinal Regulator ELT-2 Controls p38-Dependent Immune Responses in Adult C. elegans. Plos Genet 11: e1005265. https://doi.org/10.1371/journal.pgen.1005265

Block D. H., and M. Shapira, 2015 GATA transcription factors as tissue-specific master regulators for induced responses. Worm 4: e1118607. https://doi.org/10.1080/21624054.2015.1118607

Boeck M. E., T. Boyle, Z. Bao, J. Murray, B. Mericle, *et al.*, 2011 Specific roles for the GATA transcription factors end-1 and end-3 during C. elegans E-lineage development. Dev Biol 358: 345–55. https://doi.org/10.1016/j.ydbio.2011.08.002

Boeck M. E., C. Huynh, L. Gevirtzman, O. A. Thompson, G. Wang, *et al.*, 2016 The time-resolved transcriptome of C. elegans. Genome Res 26: 1441–1450. https://doi.org/10.1101/gr.202663.115

Bresnick E. H., H.-Y. Lee, T. Fujiwara, K. D. Johnson, and S. Keles, 2010 GATA switches as developmental drivers. J Biological Chem 285: 31087–93. https://doi.org/10.1074/jbc.r110.159079

Broitman-Maduro G., M. F. Maduro, and J. H. Rothman, 2005 The Noncanonical Binding Site of the MED-1 GATA Factor Defines Differentially Regulated Target Genes in the C. elegans Mesendoderm. Dev Cell 8: 427–433. https://doi.org/10.1016/j.devcel.2005.01.014

Burda P., P. Laslo, and T. Stopka, 2010 The role of PU.1 and GATA-1 transcription factors during normal and leukemogenic hematopoiesis. Leukemia 24: 1249–1257. https://doi.org/10.1038/leu.2010.104

Carr A., and M. D. Biggin, 1999 A comparison of in vivo and in vitro DNA-binding specificities suggests a new model for homeoprotein DNA binding in Drosophila embryos. Embo J 18: 1598–1608. https://doi.org/10.1093/emboj/18.6.1598

Choi H., G. Broitman-Maduro, and M. F. Maduro, 2017 Partially compromised specification causes stochastic effects on gut development in C. elegans. Dev Biol 427: 49–60. https://doi.org/10.1016/j.ydbio.2017.05.007

Chou J., S. Provot, and Z. Werb, 2010 GATA3 in development and cancer differentiation: Cells GATA have it! J Cell Physiol 222: 42–49. https://doi.org/10.1002/jcp.21943

Chun H., A. K. Sharma, J. Lee, J. Chan, S. Jia, *et al.*, 2017 The Intestinal Copper Exporter CUA-1 Is Required for Systemic Copper Homeostasis in Caenorhabditis elegans * ♦. J Biol Chem 292: 1–14. https://doi.org/10.1074/jbc.m116.760876

Consortium* T. C. elegans S., 1998 Genome Sequence of the Nematode C. elegans: A Platform for Investigating Biology. Science 282: 2012–2018. https://doi.org/10.1126/science.282.5396.2012

Davidson E. H., D. R. McClay, and L. Hood, 2003 Regulatory gene networks and the properties of the developmental process. Proc National Acad Sci 100: 1475–1480. https://doi.org/10.1073/pnas.0437746100

Delás M. J., and J. Briscoe, 2020 Repressive interactions in gene regulatory networks: When you have no other choice. Curr Top Dev Biol 139: 239–266. https://doi.org/10.1016/bs.ctdb.2020.03.003

Dimov I., and M. F. Maduro, 2019 The C. elegans intestine: organogenesis, digestion, and physiology. Cell Tissue Res 377: 383–396. https://doi.org/10.1007/s00441-019-03036-4

Dineen A., E. O. Nishimura, B. Goszczynski, J. H. Rothman, and J. D. McGhee, 2018 Quantitating transcription factor redundancy: The relative roles of the ELT-2 and ELT-7 GATA factors in the C. elegans endoderm. Dev Biol 435: 150–161. https://doi.org/10.1016/j.ydbio.2017.12.023

Djira G., M. Hasler, D. Gerhard, and L. S. and F. Schaarschmidt, 2020 *mratios: Ratios of Coefficients in the General Linear Model. R package version 1.4.2.*

Driever W., G. Thoma, and C. Nüsslein-Volhard, 1989 Determination of spatial domains of zygotic gene expression in the Drosophila embryo by the affinity of binding sites for the bicoid morphogen. Nature 340: 363–367. https://doi.org/10.1038/340363a0

Du L., S. Tracy, and S. A. Rifkin, 2016 Mutagenesis of GATA motifs controlling the endoderm regulator elt-2 reveals distinct dominant and secondary cis-regulatory elements. Dev Biol 412: 160–70. https://doi.org/10.1016/j.ydbio.2016.02.013

Elliott S. L., C. R. Sturgeon, D. M. Travers, and M. C. Montgomery, 2010 Mode of bacterial pathogenesis determines phenotype in elt-2 and elt-7 RNAi Caenorhabditis elegans. Dev Comp Immunol 35: 521–4. https://doi.org/10.1016/j.dci.2010.12.008

Emmert-Streib F., M. Dehmer, and B. Haibe-Kains, 2014 Gene regulatory networks and their applications: understanding biological and medical problems in terms of networks. Frontiers Cell Dev Biology 2: 38. https://doi.org/10.3389/fcell.2014.00038

Fujikura J., E. Yamato, S. Yonemura, K. Hosoda, S. Masui, *et al.*, 2002 Differentiation of embryonic stem cells is induced by GATA factors. Gene Dev 16: 784–789. https://doi.org/10.1101/gad.968802

Fujiwara T., 2017 GATA Transcription Factors: Basic Principles and Related Human Disorders. Tohoku J Exp Medicine 242: 83–91. https://doi.org/10.1620/tjem.242.83

Fukushige T., M. G. Hawkins, and J. D. McGhee, 1998 The GATA-factor elt-2 is essential for formation of the Caenorhabditis elegans intestine. Dev Biol 198: 286–302. https://doi.org/10.1016/s0012-1606(98)80006-7

Fukushige T., M. J. Hendzel, D. P. Bazett-Jones, and J. D. McGhee, 1999 Direct visualization of the elt-2 gut-specific GATA factor binding to a target promoter inside the living Caenorhabditis elegans embryo. Proc National Acad Sci 96: 11883–11888. https://doi.org/10.1073/pnas.96.21.11883

Fukushige T., B. Goszczynski, H. Tian, and J. D. McGhee, 2003 The Evolutionary Duplication and Probable Demise of an Endodermal GATA Factor in Caenorhabditis elegans. Genetics 165: 575–588. https://doi.org/10.1093/genetics/165.2.575

Fukushige T., B. Goszczynski, J. Yan, and J. D. McGhee, 2005 Transcriptional control and patterning of the pho-1 gene, an essential acid phosphatase expressed in the C. elegans intestine. Dev Biol 279: 446–461. https://doi.org/10.1016/j.ydbio.2004.12.012

Gillis W. Q., B. A. Bowerman, and S. Q. Schneider, 2008 The evolution of protostome GATA factors: Molecular phylogenetics, synteny, and intron/exon structure reveal orthologous relationships. Bmc Evol Biol 8: 112. https://doi.org/10.1186/1471-2148-8-112

Goszczynski B., V. V. Captan, A. M. Danielson, B. R. Lancaster, and J. D. McGhee, 2016 A 44 bp intestine-specific hermaphrodite-specific enhancer from the C. elegans vit-2 vitellogenin gene is directly regulated by ELT-2, MAB-3, FKH-9 and DAF-16 and indirectly regulated by the germline, by daf-2/insulin signaling and by the TGF-β/Sma/Mab pathway. Dev Biol 413: 112–27. https://doi.org/10.1016/j.ydbio.2016.02.031

Gould D. J., and Y. Chernajovsky, 2004 Endogenous GATA Factors Bind the Core Sequence of the tetO and Influence Gene Regulation with the Tetracycline System. Mol Ther 10: 127–138. https://doi.org/10.1016/j.ymthe.2004.04.011

Haberle V., and A. Stark, 2018 Eukaryotic core promoters and the functional basis of transcription initiation. Nat Rev Mol Cell Bio 19: 621–637. https://doi.org/10.1038/s41580-018-0028-8

Hashimshony T., M. Feder, M. Levin, B. K. Hall, and I. Yanai, 2015 Spatiotemporal transcriptomics reveals the evolutionary history of the endoderm germ layer. Nature 519: 219–222. https://doi.org/10.1038/nature13996

Hawkins M. G., and J. D. McGhee, 1995 elt-2, a Second GATA Factor from the Nematode Caenorhabditis elegans ∗. J Biol Chem 270: 14666–14671. https://doi.org/10.1074/jbc.270.24.14666

Head B., and A. Aballay, 2014 Recovery from an Acute Infection in C. elegans Requires the GATA Transcription Factor ELT-2. Plos Genet 10: e1004609. https://doi.org/10.1371/journal.pgen.1004609

Head B. P., A. O. Olaitan, and A. Aballay, 2016 Role of GATA transcription factor ELT-2 and p38 MAPK PMK-1 in recovery from acute P. aeruginosa infection in C. elegans. Virulence 8: 261–274. https://doi.org/10.1080/21505594.2016.1222334

Hobert O., 2016 A map of terminal regulators of neuronal identity in Caenorhabditis elegans. Wiley Interdiscip Rev Dev Biology 5: 474–498. https://doi.org/10.1002/wdev.233

Hong W., M. Nakazawa, Y. Chen, R. Kori, C. R. Vakoc, *et al.*, 2005 FOG-1 recruits the NuRD repressor complex to mediate transcriptional repression by GATA-1. Embo J 24: 2367–2378. https://doi.org/10.1038/sj.emboj.7600703

Kalb J. M., L. Beaster-Jones, A. P. Fernandez, P. G. Okkema, B. Goszczynski, *et al.*, 2002 Interference Between the PHA-4 and PEB-1 Transcription Factors in Formation of the Caenorhabditis elegans Pharynx. J Mol Biol 320: 697–704. https://doi.org/10.1016/s0022-2836(02)00555-7

Keith S. A., S. K. Maddux, Y. Zhong, M. N. Chinchankar, A. A. Ferguson, *et al.*, 2016 Graded Proteasome Dysfunction in Caenorhabditis elegans Activates an Adaptive Response Involving the Conserved SKN-1 and ELT-2 Transcription Factors and the Autophagy-Lysosome Pathway. Plos Genet 12: e1005823. https://doi.org/10.1371/journal.pgen.1005823

Kimble J., and W. J. Sharrock, 1983 Tissue-specific synthesis of yolk proteins in Caenorhabditis elegans. Dev Biol 96: 189–196. https://doi.org/10.1016/0012-1606(83)90322-6

Kudron M. M., A. Victorsen, L. Gevirtzman, L. W. Hillier, W. W. Fisher, *et al.*, 2017 The modERN Resource: Genome-Wide Binding Profiles for Hundreds of Drosophila and

Caenorhabditis elegans Transcription Factors. Genetics 208: genetics.300657.2017. https://doi.org/10.1534/genetics.117.300657

Kulik M., M. Bothe, G. Kibar, A. Fuchs, S. Schöne, *et al.*, 2021 Androgen and glucocorticoid receptor direct distinct transcriptional programs by receptor-specific and shared DNA binding sites. Nucleic Acids Res 49: gkab185-. https://doi.org/10.1093/nar/gkab185

Lancaster B. R., and J. D. McGhee, 2020 How affinity of the ELT-2 GATA factor binding to cis-acting regulatory sites controls C. elegans intestinal gene transcription. Dev Camb Engl dev.190330. https://doi.org/10.1242/dev.190330

Lee K., and E. Mylonakis, 2017 An Intestine-Derived Neuropeptide Controls Avoidance Behavior in Caenorhabditis elegans. Cell Reports 20: 2501–2512. https://doi.org/10.1016/j.celrep.2017.08.053

Levine M., and E. H. Davidson, 2005 Gene regulatory networks for development. P Natl Acad Sci Usa 102: 4936–4942. https://doi.org/10.1073/pnas.0408031102

Libina N., J. R. Berman, and C. Kenyon, 2003 Tissue-Specific Activities of C. elegans DAF-16 in the Regulation of Lifespan. Cell 115: 489–502. https://doi.org/10.1016/s0092-8674(03)00889-4

Lickwar C. R., F. Mueller, S. E. Hanlon, J. G. McNally, and J. D. Lieb, 2012 Genome-wide protein–DNA binding dynamics suggest a molecular clutch for transcription factor function. Nature 484: 251–255. https://doi.org/10.1038/nature10985

Liu X., C.-K. Lee, J. A. Granek, N. D. Clarke, and J. D. Lieb, 2006 Whole-genome comparison of Leu3 binding in vitro and in vivo reveals the importance of nucleosome occupancy in target site selection. Genome Res 16: 1517–1528. https://doi.org/10.1101/gr.5655606

Lowry J. A., and W. R. Atchley, 2000 Molecular Evolution of the GATA Family of Transcription Factors: Conservation Within the DNA-Binding Domain. J Mol Evol 50: 103–115. https://doi.org/10.1007/s002399910012

Ludewig A. H., and F. C. Schroeder, 2013 Ascaroside signaling in C. elegans. Wormbook 1–22. https://doi.org/10.1895/wormbook.1.155.1

Maduro M. F., and J. H. Rothman, 2002 Making Worm Guts: The Gene Regulatory Network of the Caenorhabditis elegans Endoderm. Dev Biol 246: 68–85. https://doi.org/10.1006/dbio.2002.0655

Maduro M. F., R. J. Hill, P. J. Heid, E. D. Newman-Smith, J. Zhu, *et al.*, 2005a Genetic redundancy in endoderm specification within the genus Caenorhabditis. Dev Biol 284: 509–522. https://doi.org/10.1016/j.ydbio.2005.05.016

Maduro M. F., J. J. Kasmir, J. Zhu, and J. H. Rothman, 2005b The Wnt effector POP-1 and the PAL-1/Caudal homeoprotein collaborate with SKN-1 to activate C. elegans endoderm development. Dev Biol 285: 510–523. https://doi.org/10.1016/j.ydbio.2005.06.022

Maduro M. F., 2019 Evolutionary Dynamics of the SKN-1 → MED → END-1,3 Regulatory Gene Cascade in Caenorhabditis Endoderm Specification. G3 Genes Genomes Genetics 10: g3.400724.2019. https://doi.org/10.1534/g3.119.400724

Mann F. G., E. L. V. Nostrand, A. E. Friedland, X. Liu, and S. K. Kim, 2016 Deactivation of the GATA Transcription Factor ELT-2 Is a Major Driver of Normal Aging in C. elegans. Plos Genet 12: e1005956. https://doi.org/10.1371/journal.pgen.1005956

Martinez-Finley E. J., and M. Aschner, 2011 Revelations from the Nematode Caenorhabditis elegans on the Complex Interplay of Metal Toxicological Mechanisms. J Toxicol 2011: 895236. https://doi.org/10.1155/2011/895236

McEwan D. L., R. L. Feinbaum, N. Stroustrup, W. Haas, A. L. Conery, *et al.*, 2016 Tribbles ortholog NIPI-3 and bZIP transcription factor CEBP-1 regulate a Caenorhabditis elegans intestinal immune surveillance pathway. Bmc Biol 14: 105. https://doi.org/10.1186/s12915-016-0334-6

McGhee J., 2007 The C. elegans intestine. Wormbook 1–36. https://doi.org/10.1895/wormbook.1.133.1

McGhee J. D., M. C. Sleumer, M. Bilenky, K. Wong, S. J. McKay, *et al.*, 2007 The ELT-2 GATA-factor and the global regulation of transcription in the C. elegans intestine. Dev Biol 302: 627–645. https://doi.org/10.1016/j.ydbio.2006.10.024

McGhee J. D., T. Fukushige, M. W. Krause, S. E. Minnema, B. Goszczynski, *et al.*, 2009 ELT-2 is the predominant transcription factor controlling differentiation and function of the C. elegans intestine, from embryo to adult. Dev Biol 327: 551–565. https://doi.org/10.1016/j.ydbio.2008.11.034

Meijsing S. H., M. A. Pufall, A. Y. So, D. L. Bates, L. Chen, *et al.*, 2009 DNA Binding Site Sequence Directs Glucocorticoid Receptor Structure and Activity. Science 324: 407–410. https://doi.org/10.1126/science.1164265

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

Nair G., T. Walton, J. I. Murray, and A. Raj, 2013 Gene transcription is coordinated with, but not dependent on, cell divisions during C. elegans embryonic fate specification. Development 140: 3385–3394. https://doi.org/10.1242/dev.098012

Oestreich K. J., and A. S. Weinmann, 2012 Master regulators or lineage-specifying? Changing views on CD4+ T cell transcription factors. Nat Rev Immunol 12: 799–804. https://doi.org/10.1038/nri3321

Oskouian B., J. Mendel, E. Shocron, M. A. Lee, H. Fyrst, *et al.*, 2005 Regulation of Sphingosine-1-phosphate Lyase Gene Expression by Members of the GATA Family of Transcription Factors*. J Biol Chem 280: 18403–18410. https://doi.org/10.1074/jbc.m410928200

Pauli F., Y. Liu, Y. A. Kim, P.-J. Chen, and S. K. Kim, 2005 Chromosomal clustering and GATA transcriptional regulation of intestine-expressed genes in C. elegans. Development 133: 287–295. https://doi.org/10.1242/dev.02185

Petrella L. N., W. Wang, C. A. Spike, A. Rechtsteiner, V. Reinke, *et al.*, 2011 synMuv B proteins antagonize germline fate in the intestine and ensure C. elegans survival. Development 138: 1069–1079. https://doi.org/10.1242/dev.059501

Raj A., S. A. Rifkin, E. Andersen, and A. van Oudenaarden, 2010 Variability in gene expression underlies incomplete penetrance. Nature 463: 913–8. https://doi.org/10.1038/nature08781

Riddle M. R., A. Weintraub, K. C. Q. Nguyen, D. H. Hall, and J. H. Rothman, 2013 Transdifferentiation and remodeling of post-embryonic C. elegans cells by a single transcription factor. Dev Camb Engl 140: 4844–9. https://doi.org/10.1242/dev.103010

Riddle M. R., E. A. Spickard, A. Jevince, K. C. Q. Nguyen, D. H. Hall, *et al.*, 2016 Transorganogenesis and transdifferentiation in C. elegans are dependent on differentiated cell identity. Dev Biol 420: 136–147. https://doi.org/10.1016/j.ydbio.2016.09.020

Rodríguez-López M., S. Gonzalez, O. Hillson, E. Tunnacliffe, S. Codlin, *et al.*, 2020 The GATA Transcription Factor Gaf1 Represses tRNAs, Inhibits Growth, and Extends Chronological Lifespan Downstream of Fission Yeast TORC1. Cell Reports 30: 3240-3249.e4. https://doi.org/10.1016/j.celrep.2020.02.058

Roh H. C., I. Dimitrov, K. Deshmukh, G. Zhao, K. Warnhoff, *et al.*, 2014 A modular system of DNA enhancer elements mediates tissue-specific activation of transcription by high dietary zinc in C. elegans. Nucleic Acids Res 43: 803–16. https://doi.org/10.1093/nar/gku1360

Schmitt A. P., and K. McEntee, 1996 Msn2p, a zinc finger DNA-binding protein, is the transcriptional activator of the multistress response in Saccharomyces cerevisiae. Proc National Acad Sci 93: 5777–5782. https://doi.org/10.1073/pnas.93.12.5777

Senecal A., B. Munsky, F. Proux, N. Ly, F. E. Braye, *et al.*, 2014 Transcription Factors Modulate c-Fos Transcriptional Bursts. Cell Reports 8: 75–83. https://doi.org/10.1016/j.celrep.2014.05.053

Sommermann E. M., K. R. Strohmaier, M. F. Maduro, and J. H. Rothman, 2010 Endoderm development in Caenorhabditis elegans: The synergistic action of ELT-2 and -7 mediates the specification→differentiation transition. Dev Biol 347: 154–166. https://doi.org/10.1016/j.ydbio.2010.08.020

Spencer W. C., G. Zeller, J. D. Watson, S. R. Henz, K. L. Watkins, *et al.*, 2011 A spatial and temporal map of C. elegans gene expression. Genome Res 21: 325–341. https://doi.org/10.1101/gr.114595.110

Stiernagle T., 2006 Maintenance of C. elegans. Wormbook 1–11. https://doi.org/10.1895/wormbook.1.101.1

Su L., T. Zhao, H. Li, H. Li, X. Su, *et al.*, 2020 ELT-2 promotes O-GlcNAc transferase OGT-1 expression to modulate Caenorhabditis elegans lifespan. J Cell Biochem 121: 4898–4907. https://doi.org/10.1002/jcb.29817

Tepper R. G., J. Ashraf, R. Kaletsky, G. Kleemann, C. T. Murphy, *et al.*, 2013 PQM-1 Complements DAF-16 as a Key Transcriptional Regulator of DAF-2-Mediated Development and Longevity. Cell 154: 676–690. https://doi.org/10.1016/j.cell.2013.07.006

Tintori S. C., E. Osborne Nishimura, P. Golden, J. D. Lieb, and B. Goldstein, 2016 A Transcriptional Lineage of the Early C. elegans Embryo. Dev Cell 38: 430–444. https://doi.org/10.1016/j.devcel.2016.07.025

Ucar D., A. Beyer, S. Parthasarathy, and C. T. Workman, 2009 Predicting functionality of protein–DNA interactions by integrating diverse evidence. Bioinformatics 25: i137–i144. https://doi.org/10.1093/bioinformatics/btp213

Villarino A., A. Laurence, G. W. Robinson, M. Bonelli, B. Dema, *et al.*, 2016 Signal transducer and activator of transcription 5 (STAT5) paralog dose governs T cell effector and regulatory functions. Elife 5: e08384. https://doi.org/10.7554/elife.08384

Wang D., S. Kennedy, D. Conte, J. K. Kim, H. W. Gabel, *et al.*, 2005 Somatic misexpression of germline P granules and enhanced RNA interference in retinoblastoma pathway mutants. Nature 436: 593–597. https://doi.org/10.1038/nature04010

Wiesenfahrt T., J. Y. Berg, E. O. Nishimura, A. G. Robinson, B. Goszczynski, *et al.*, 2015 The function and regulation of the GATA factor ELT-2 in the C. elegans endoderm. Dev Camb Engl 143: 483–91. https://doi.org/10.1242/dev.130914

Woollard A., 2005 Gene duplications and genetic redundancy in C. elegans. Wormbook 1–6. https://doi.org/10.1895/wormbook.1.2.1

Yan D., Z. Wu, A. D. Chisholm, and Y. Jin, 2009 The DLK-1 Kinase Promotes mRNA Stability and Local Translation in C. elegans Synapses and Axon Regeneration. Cell 138: 1005–1018. https://doi.org/10.1016/j.cell.2009.06.023

Yang A., Z. Zhu, P. Kapranov, F. McKeon, G. M. Church, *et al.*, 2006 Relationships between p63 Binding, DNA Sequence, Transcription Activity, and Biological Function in Human Cells. Mol Cell 24: 593–602. https://doi.org/10.1016/j.molcel.2006.10.018

Yang W., K. Dierking, P. C. Rosenstiel, and H. Schulenburg, 2016 GATA transcription factor as a likely key regulator of the Caenorhabditis elegans innate immune response against gut pathogens. Zoology 119: 244–253. https://doi.org/10.1016/j.zool.2016.05.013

Yang Y., L. Liu, I. Naik, Z. Braunstein, J. Zhong, *et al.*, 2017 Transcription Factor C/EBP Homologous Protein in Health and Diseases. Front Immunol 8: 1612. https://doi.org/10.3389/fimmu.2017.01612

Yissachar N., T. S. Fischler, A. A. Cohen, S. Reich-Zeliger, D. Russ, *et al.*, 2012 Dynamic response diversity of NFAT isoforms in individual living cells. Mol Cell 49: 322–30. https://doi.org/10.1016/j.molcel.2012.11.003

Zárate-Potes A., W. Yang, B. Pees, R. Schalkowski, P. Segler, *et al.*, 2020 The C. elegans GATA transcription factor elt-2 mediates distinct transcriptional responses and opposite infection outcomes towards different Bacillus thuringiensis strains. Plos Pathog 16: e1008826. https://doi.org/10.1371/journal.ppat.1008826

Zheng R., and G. A. Blobel, 2010 GATA Transcription Factors and Cancer. Genes Cancer 1: 1178–1188. https://doi.org/10.1177/1947601911404223

Figure 1

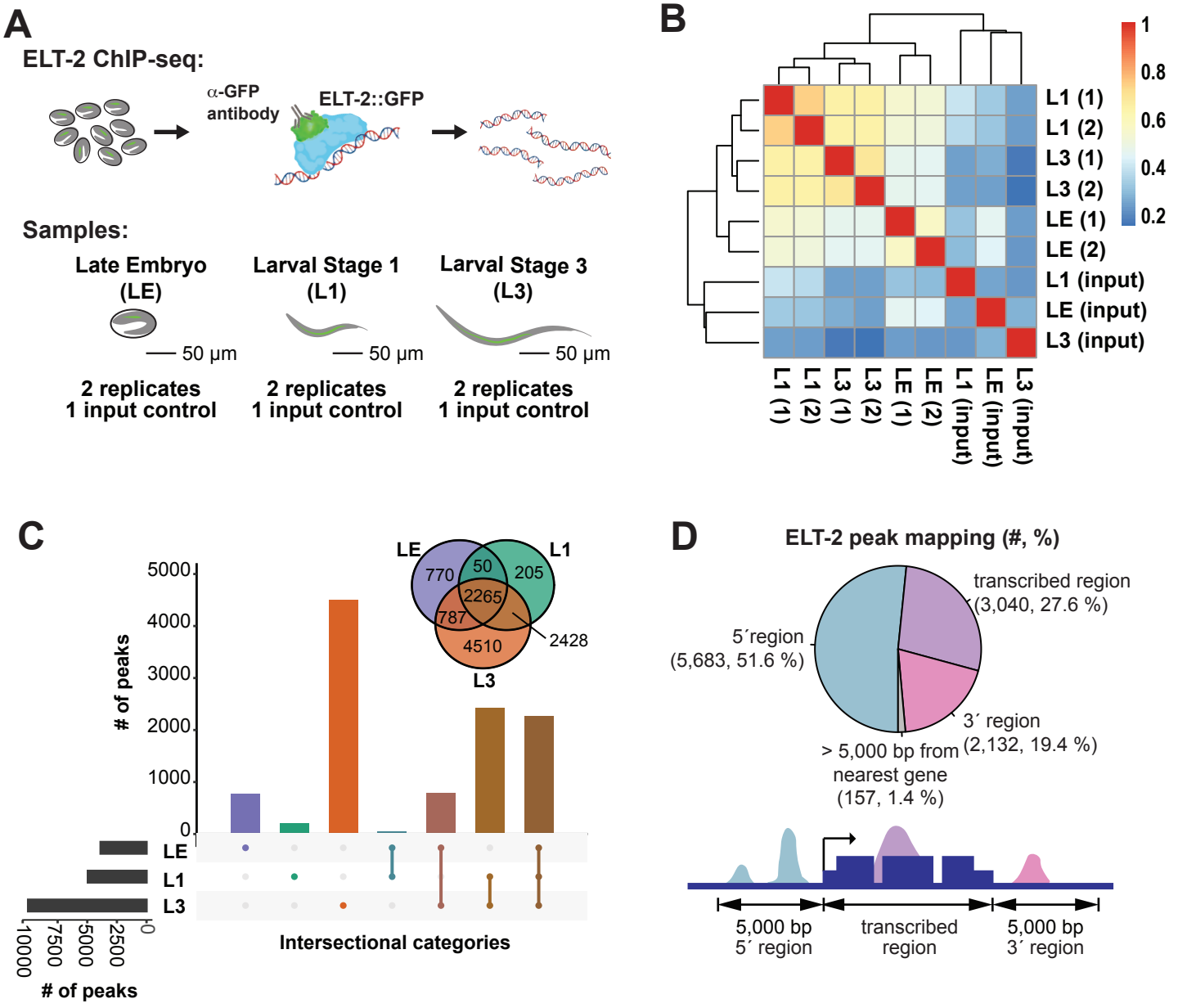

A
ELT-2 ChIP-seq:
α-GFP antibody
ELT-2::GFP
Samples:
Late Embryo (LE)
Larval Stage 1 (L1)
Larval Stage 3 (L3)
50 µm
50 µm
50 µm
2 replicates 1 input control
2 replicates 1 input control
2 replicates 1 input control
B
L1 (1)
L1 (2)
L3 (1)
L3 (2)
LE (1)
LE (2)
L1 (input)
LE (input)
L3 (input)
1
0.8
0.6
0.4
0.2
C
LE
L1
L3
770
50
205
2265
787
4510
2428
# of peaks
5000
4000
3000
2000
1000
0
10000
7500
5000
2500
0
Intersectional categories
D
ELT-2 peak mapping (#, %)
transcribed region (3,040, 27.6 %)
5´region (5,683, 51.6 %)
3´ region (2,132, 19.4 %)
> 5,000 bp from nearest gene (157, 1.4 %)
5,000 bp 5´ region
transcribed region
5,000 bp 3´ region

Figure 2

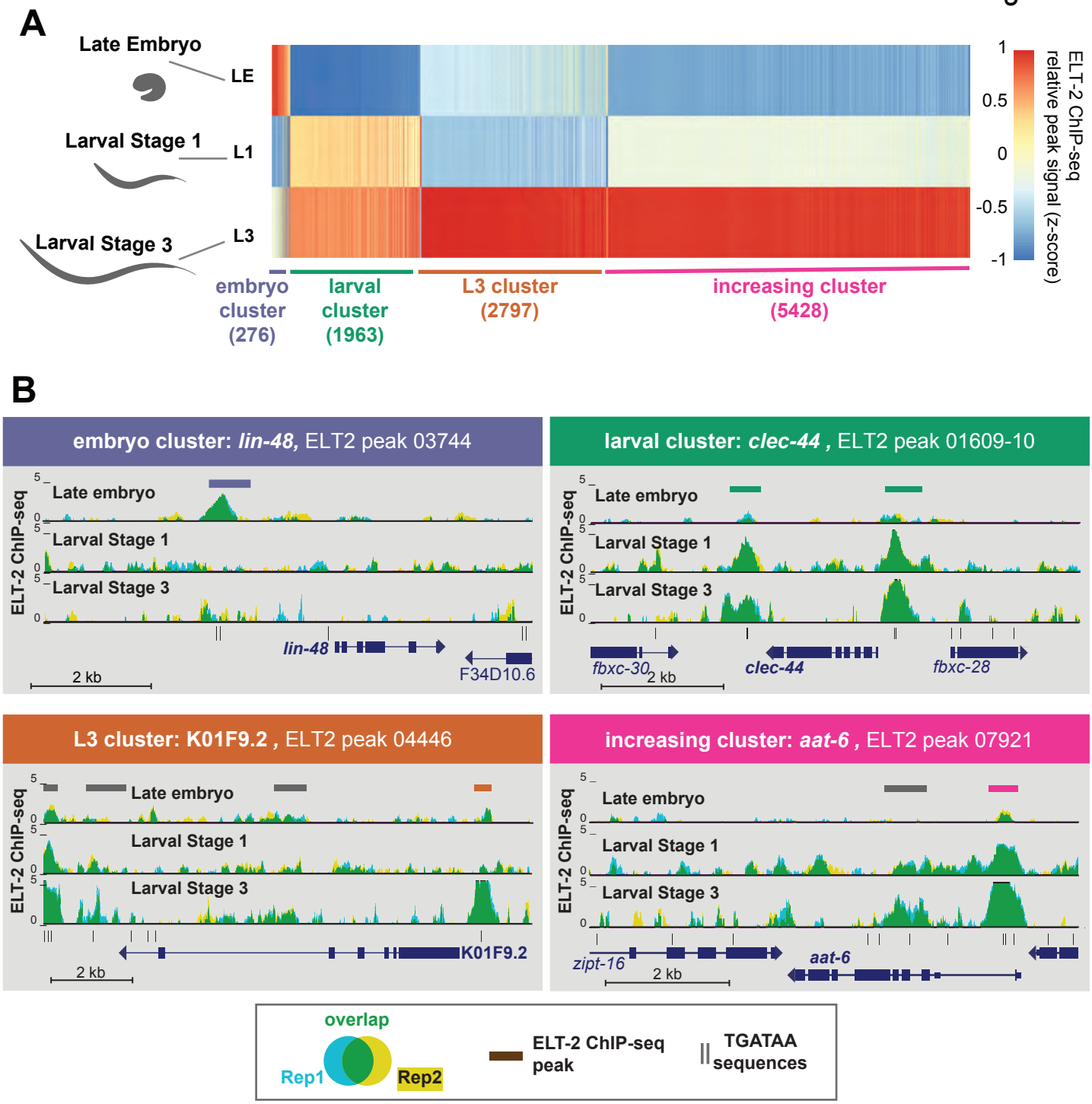

A
Late Embryo
LE
Larval Stage 1
L1
Larval Stage 3
L3
embryo cluster (276)
larval cluster (1963)
L3 cluster (2797)
increasing cluster (5428)
ELT-2 ChIP-seq relative peak signal (z-score)
1
0.5
0
-0.5
-1
B
embryo cluster: lin-48, ELT2 peak 03744
larval cluster: clec-44 , ELT2 peak 01609-10
L3 cluster: K01F9.2 , ELT2 peak 04446
increasing cluster: aat-6 , ELT2 peak 07921
ELT-2 ChIP-seq
Late embryo
Larval Stage 1
Larval Stage 3
lin-48
F34D10.6
fbxc-30
clec-44
fbxc-28
K01F9.2
zipt-16
aat-6
2 kb
overlap
Rep1
Rep2
ELT-2 ChIP-seq peak
TGATAA sequences

Figure 3

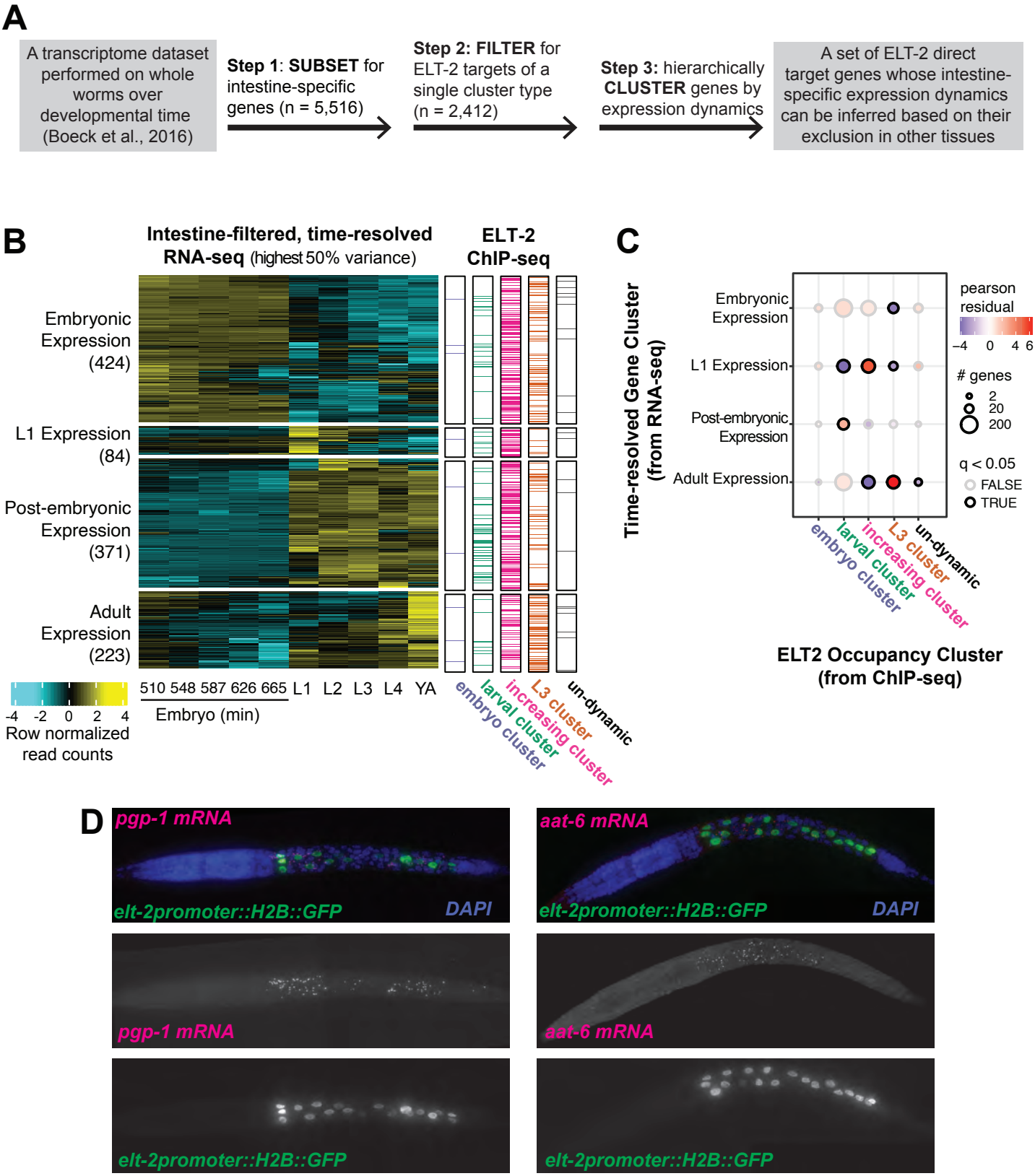

A
A transcriptome dataset performed on whole worms over developmental time (Boeck et al., 2016)
Step 1: SUBSET for intestine-specific genes (n = 5,516)
Step 2: FILTER for ELT-2 targets of a single cluster type (n = 2,412)
Step 3: hierarchically CLUSTER genes by expression dynamics
A set of ELT-2 direct target genes whose intestine-specific expression dynamics can be inferred based on their exclusion in other tissues
B
Intestine-filtered, time-resolved RNA-seq (highest 50% variance)
ELT-2 ChIP-seq
Embryonic Expression (424)
L1 Expression (84)
Post-embryonic Expression (371)
Adult Expression (223)
-4 -2 0 2 4
Row normalized read counts
510 548 587 626 665 L1 L2 L3 L4 YA
Embryo (min)
embryo cluster
larval cluster
increasing cluster
L3 cluster
un-dynamic
C
Time-resolved Gene Cluster (from RNA-seq)
Embryonic Expression
L1 Expression
Post-embryonic Expression
Adult Expression
pearson residual
−4 0 4 6
genes
2
20
200
q < 0.05
FALSE
TRUE
embryo cluster
larval cluster
increasing cluster
L3 cluster
un-dynamic
ELT2 Occupancy Cluster (from ChIP-seq)
D
pgp-1 mRNA
elt-2promoter::H2B::GFP
DAPI
aat-6 mRNA
elt-2promoter::H2B::GFP
DAPI
pgp-1 mRNA
aat-6 mRNA
elt-2promoter::H2B::GFP
elt-2promoter::H2B::GFP

Figure 4

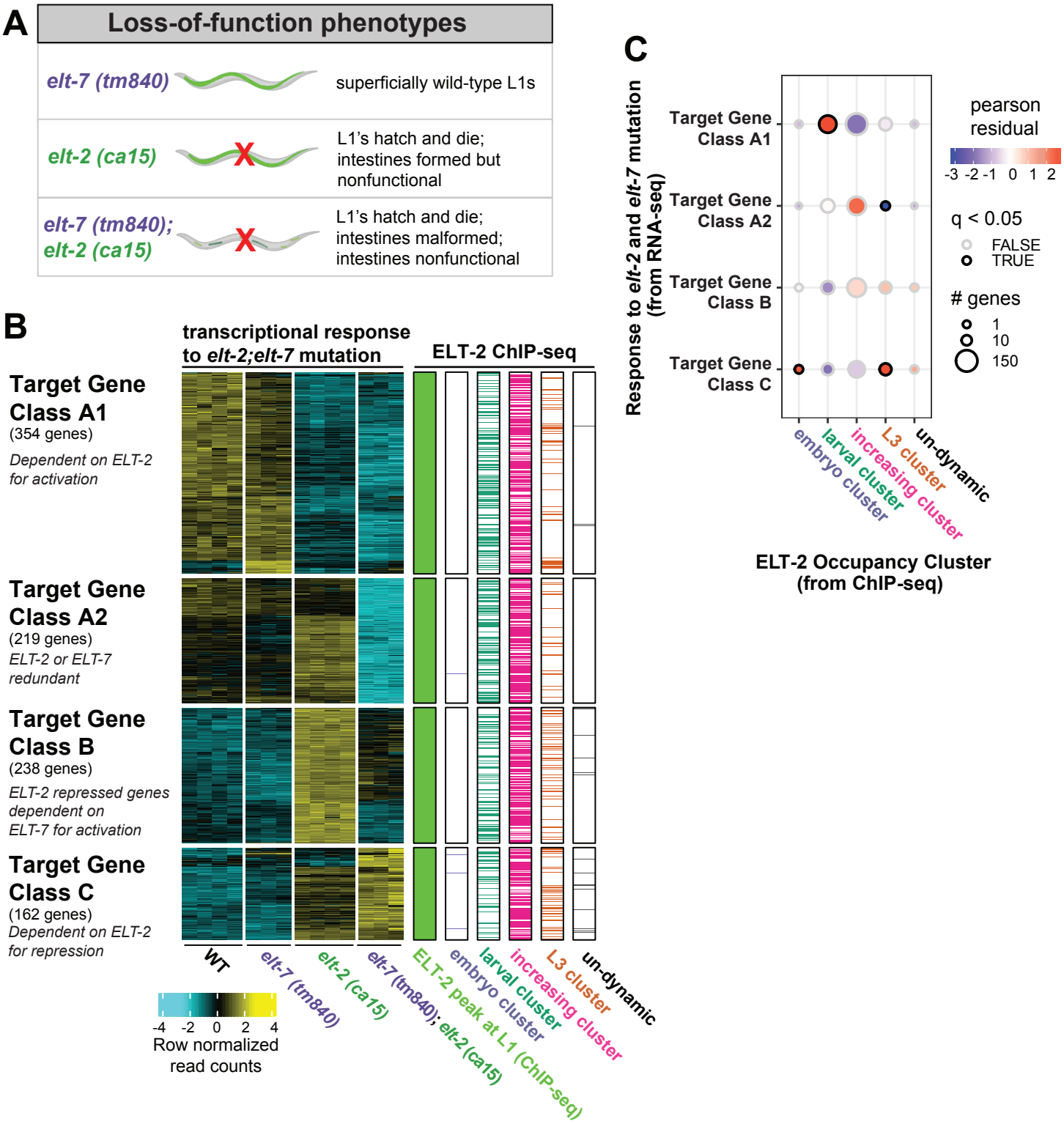

A
Loss-of-function phenotypes
elt-7 (tm840)
superficially wild-type L1s
elt-2 (ca15)
L1's hatch and die; intestines formed but nonfunctional
elt-7 (tm840); elt-2 (ca15)
L1's hatch and die; intestines malformed; intestines nonfunctional
B
transcriptional response to elt-2;elt-7 mutation
ELT-2 ChIP-seq
Target Gene Class A1
(354 genes)
Dependent on ELT-2 for activation
Target Gene Class A2
(219 genes)
ELT-2 or ELT-7 redundant
Target Gene Class B
(238 genes)
ELT-2 repressed genes dependent on ELT-7 for activation
Target Gene Class C
(162 genes)
Dependent on ELT-2 for repression
WT
elt-7 (tm840)
elt-2 (ca15)
elt-7 (tm840); elt-2 (ca15)
ELT-2 peak at L1 (ChIP-seq)
embryo cluster
larval cluster
increasing cluster
L3 cluster
un-dynamic
-4 -2 0 2 4
Row normalized read counts
C
Response to elt-2 and elt-7 mutation (from RNA-seq)
Target Gene Class A1
Target Gene Class A2
Target Gene Class B
Target Gene Class C
embryo cluster
larval cluster
increasing cluster
L3 cluster
un-dynamic
ELT-2 Occupancy Cluster (from ChIP-seq)
pearson residual
-3 -2 -1 0 1 2
q < 0.05
FALSE
TRUE
genes
1
10
150

Figure 5

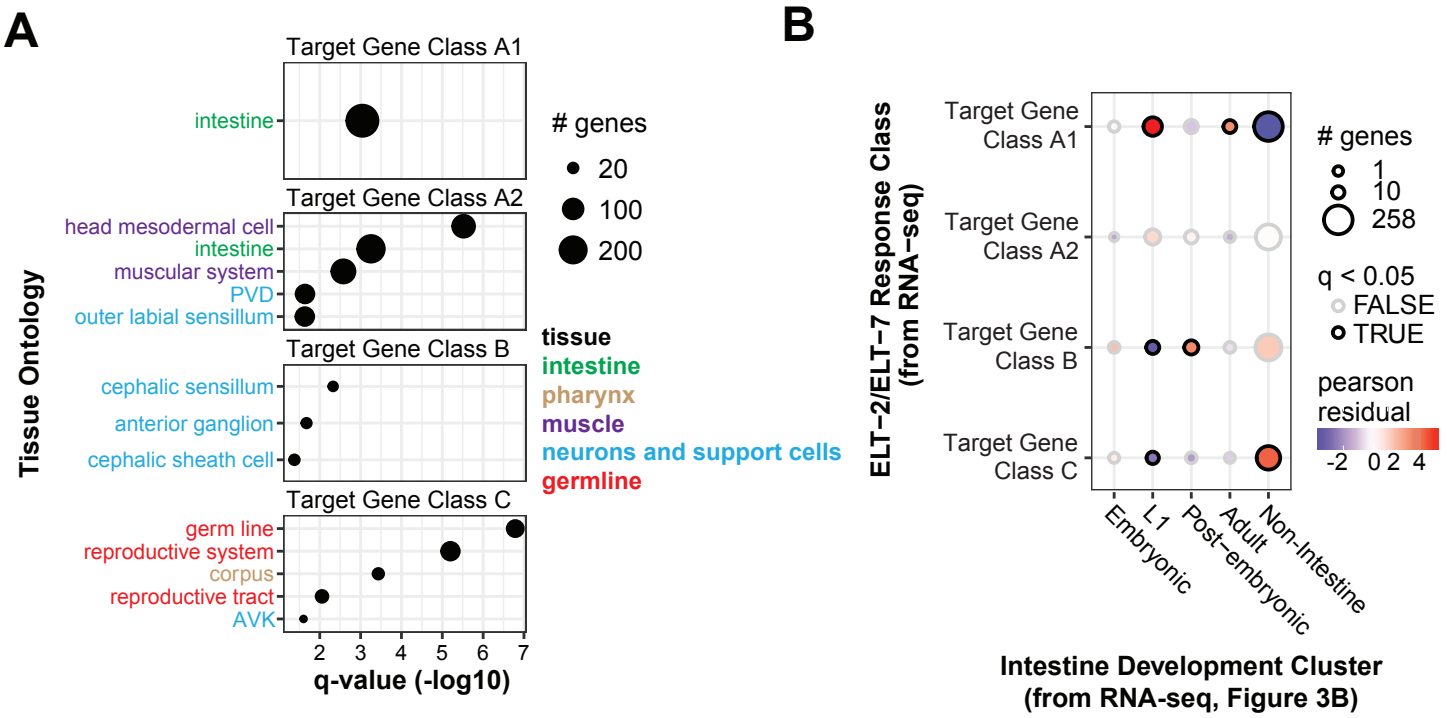

A
Target Gene Class A1
intestine
Target Gene Class A2
head mesodermal cell
intestine
muscular system
PVD
outer labial sensillum
Target Gene Class B
cephalic sensillum
anterior ganglion
cephalic sheath cell
Target Gene Class C
germ line
reproductive system
corpus
reproductive tract
AVK
Tissue Ontology
2 3 4 5 6 7
q-value (-log10)
genes
20
100
200
tissue
intestine
pharynx
muscle
neurons and support cells
germline
B
ELT−2/ELT−7 Response Class
(from RNA−seq)
Target Gene Class A1
Target Gene Class A2
Target Gene Class B
Target Gene Class C
Embryonic
L1
Post−embryonic
Adult
Non−Intestine
Intestine Development Cluster
(from RNA-seq, Figure 3B)
genes
1
10
258
q < 0.05
FALSE
TRUE
pearson residual
-2 0 2 4

Figure 6

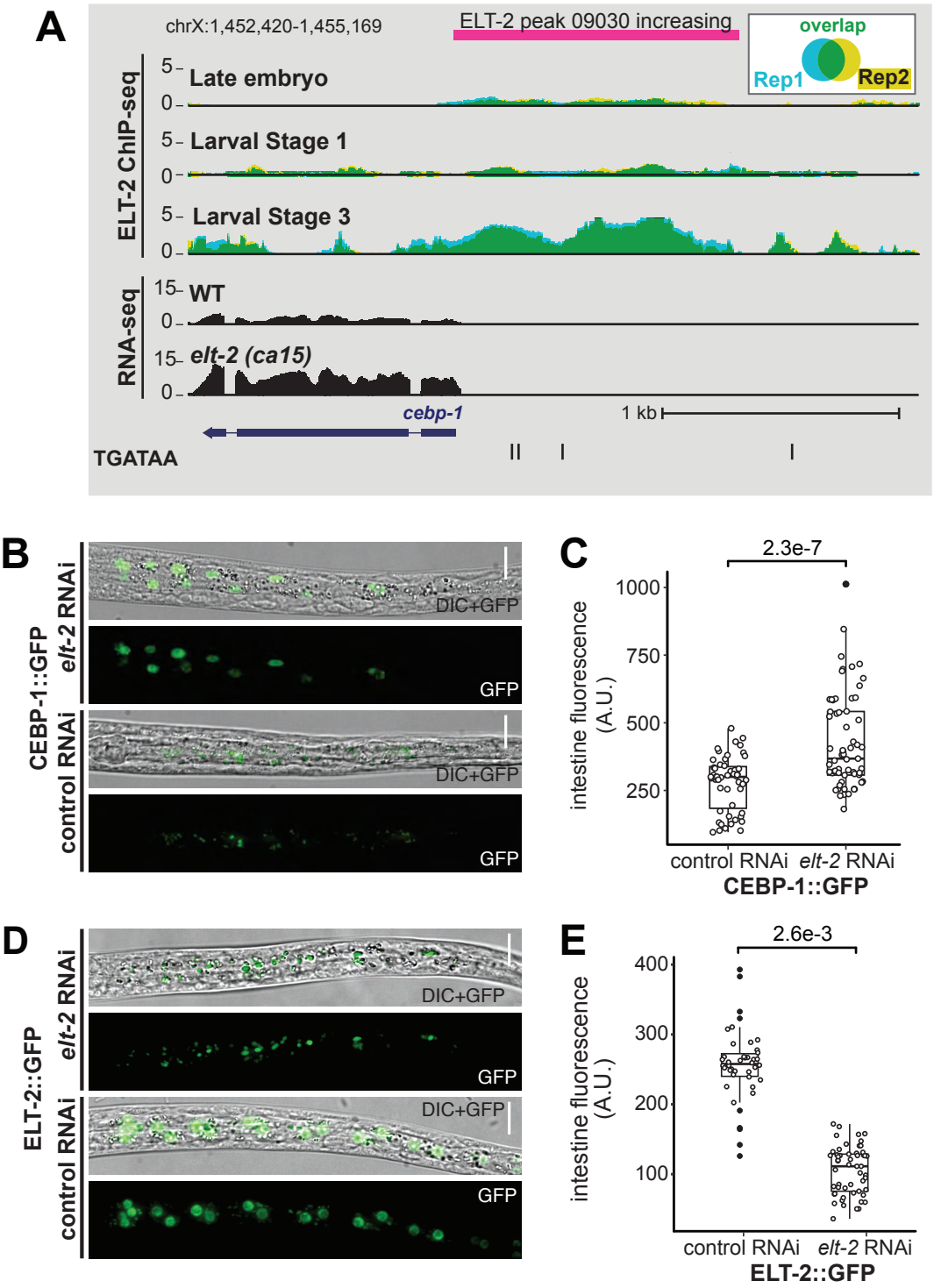

A
chrX:1,452,420-1,455,169
ELT-2 peak 09030 increasing
overlap
Rep1
Rep2
ELT-2 ChIP-seq
Late embryo
Larval Stage 1
Larval Stage 3
RNA-seq
WT
elt-2 (ca15)
cebp-1
1 kb
TGATAA
B
CEBP-1::GFP
elt-2 RNAi
control RNAi
DIC+GFP
GFP
C
2.3e-7
intestine fluorescence (A.U.)
control RNAi
elt-2 RNAi
CEBP-1::GFP
D
ELT-2::GFP
E
2.6e-3
ELT-2::GFP

Figure 7

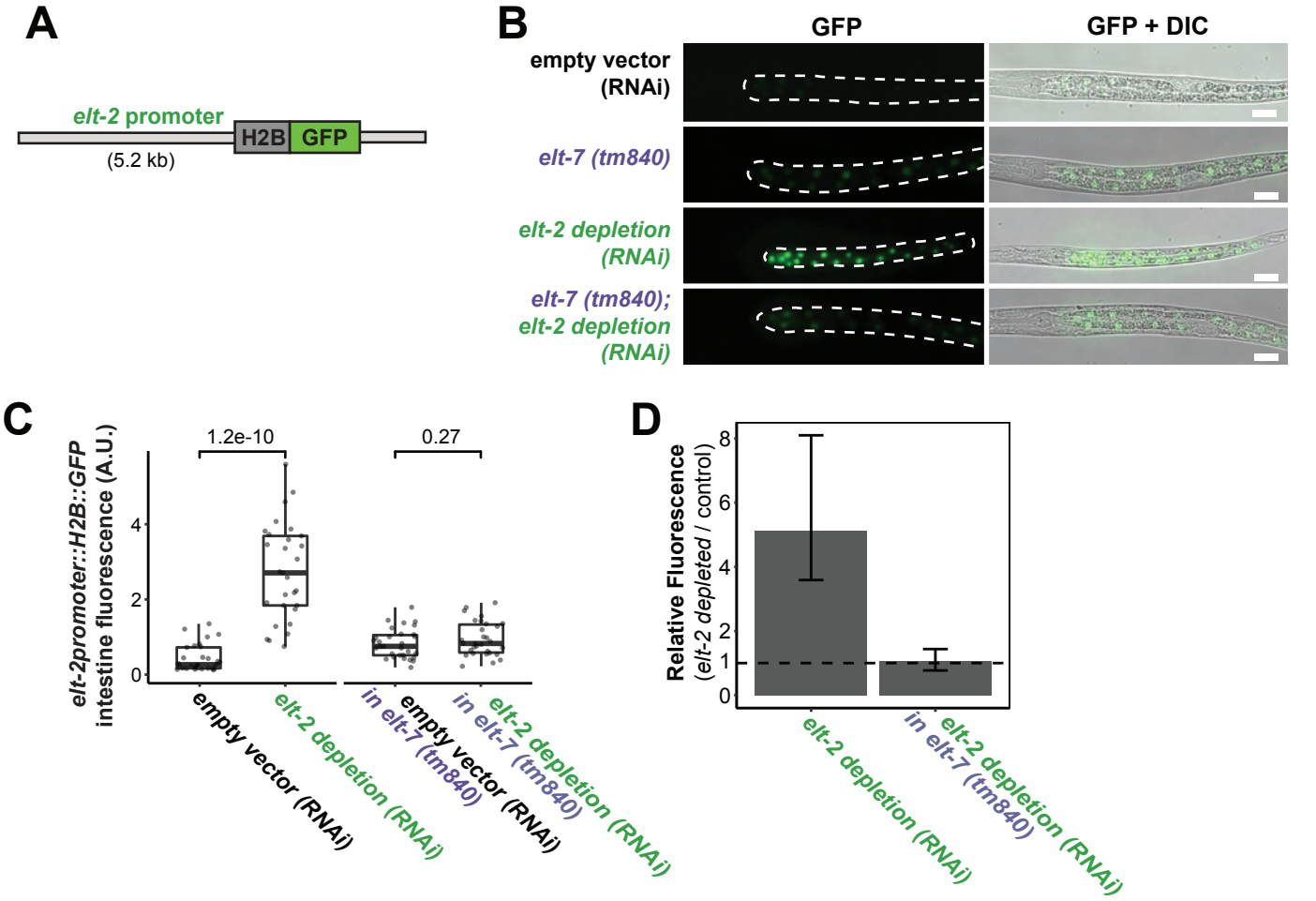
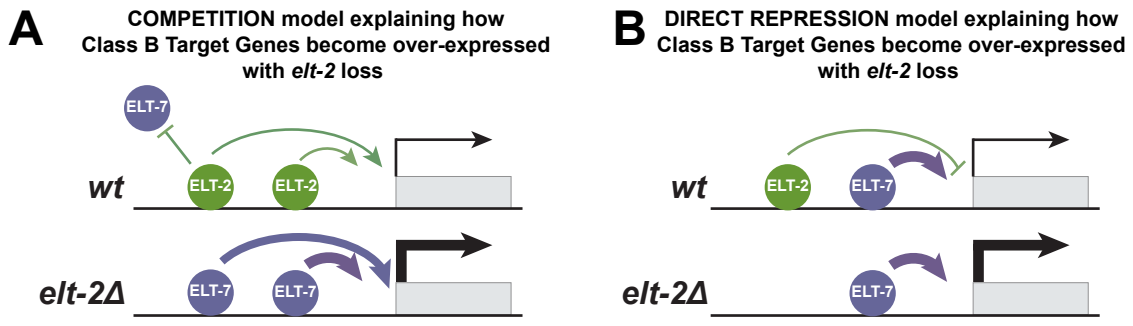

A
elt-2 promoter
H2B
GFP
(5.2 kb)
B
GFP
GFP + DIC
empty vector (RNAi)
elt-7 (tm840)
elt-2 depletion (RNAi)
elt-7 (tm840); elt-2 depletion (RNAi)
C
elt-2promoter::H2B::GFP intestine fluorescence (A.U.)
1.2e-10
0.27
0
2
4
empty vector (RNAi)
elt-2 depletion (RNAi)
in elt-7 (tm840)
empty vector (RNAi)
in elt-7 (tm840)
elt-2 depletion (RNAi)
D
Relative Fluorescence
(elt-2 depleted / control)
0
1
2
4
6
8
elt-2 depletion (RNAi)
in elt-7 (tm840)
elt-2 depletion (RNAi)

Figure 8


A
COMPETITION model explaining how
Class B Target Genes become over-expressed
with elt-2 loss
ELT-7
wt
ELT-2
ELT-2
elt-2Δ
ELT-7
ELT-7
B
DIRECT REPRESSION model explaining how
Class B Target Genes become over-expressed
with elt-2 loss
wt
ELT-2
ELT-7
elt-2Δ
ELT-7

Supplemental Materials

The transcription factor ELT-2 positively and negatively impacts direct target genes to modulate the *Caenorhabditis elegans* intestinal transcriptome

David C. King[*†] and Robert T.P. Williams[*†], Izabella R. Mastroianni[*], Jessica Hill[*], Nicolai W. Apenas[‡], Gabriela Ramirez[§], and Erin Osborne Nishimura[*, **]

Supplemental Figures & Legends:

Supplemental Figure 1. Genomic features of ELT-2 ChIP-seq peaks. Distribution of peaks genome-wide and by chromosome**. A.** Number of peaks versus chromosome length. A smoothed 95% confidence interval (standard error) is plotted in gray. **B.** Density of peaks along chromosome position. Density of peaks is higher in chromosome centers.

Supplemental Figure 2. The four developmental trends for Late Embryo (LE) and Larval Stages 1 (L1) and 3 (L3) by peak intensity. A. The k-means cluster centers for the indicated stages and cluster. **B.** Deeptools plot of read-depth normalized signal intensity (blue to red) of peaks in the indicated stage and cluster. Peaks are organized vertically and centered in a 2kb window. **C.** Average values of the peaks in (B) represented by a smoothed density profile per base +/- 1kb

Supplemental Figure 3. Variety of developmental peak types assigned to genes. UpSet plot of genes that are unique and shared between ELT-2 occupancy types. Filled dots indicate the specific combination of peak types that are counted in the bar above it. Bars represent the number of genes having at least one of the indicated type (the number of peaks of each type mapping to a given gene is not represented). The blocks of vertical bars show the number of genes mapping to exactly 1, 2, 3 and 4 types of peaks, respectively. Bar heights below 25 are not labeled. Counts greater than 0 but below 5 are given a uniform bar height for visualization purposes.

Supplemental Figure 4. Enriched functional annotations of genes assigned to unique ELT-2 occupancy types. For each panel, the most enriched terms are shown for each dynamic class versus the union of all dynamic classes. The p-value is determined by the hypergeometric test and adjusted to q-value through Benjamini-Hochberg multiple test correction. Terms are included when represented by more than 5 genes and q-value < 0.05. **(A)** Gene ontology. **(B)** Phenotype Ontology. **(C)** Tissue Ontology.

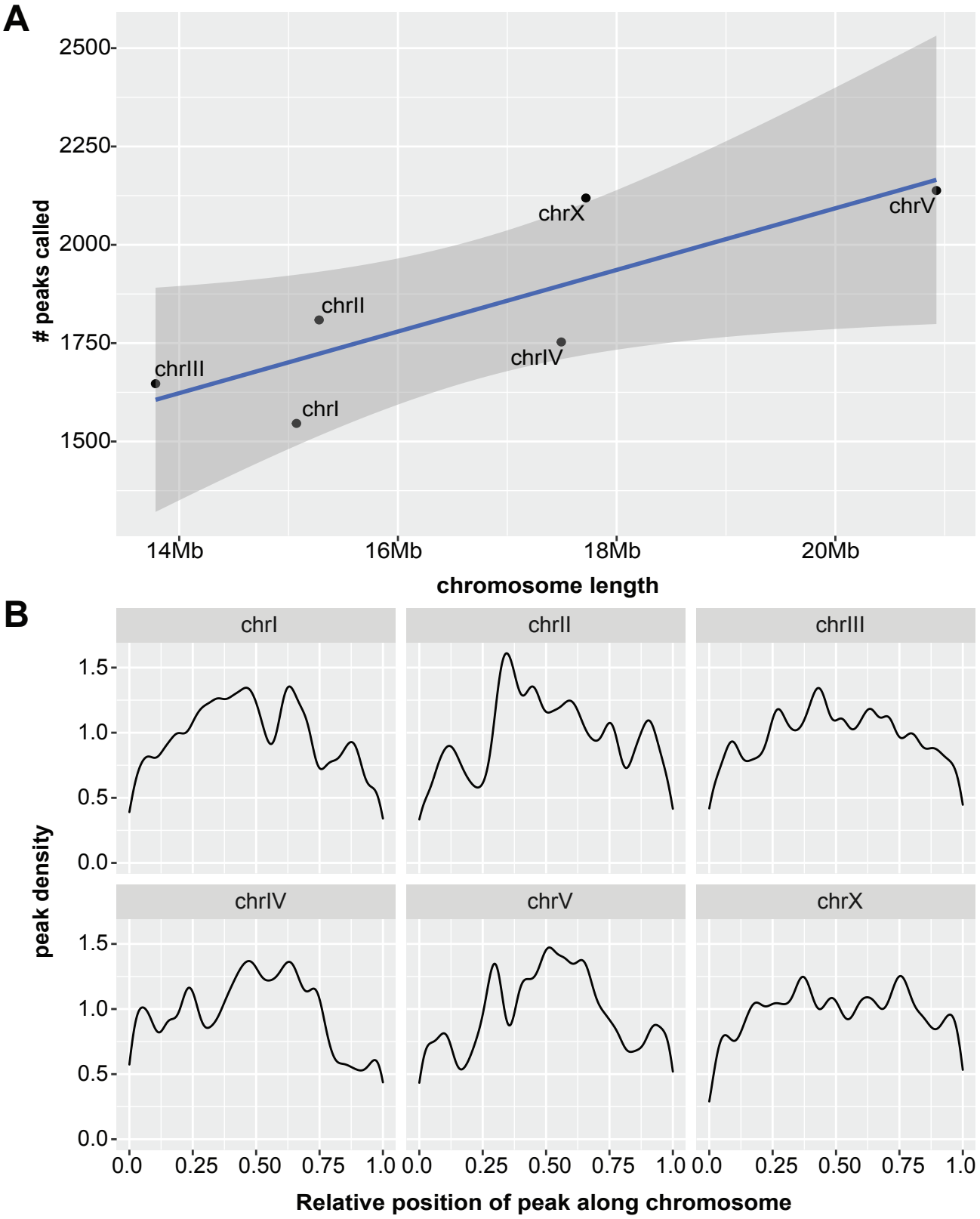
Supplemental Figure 5. ELT-2 binding and dynamic expression of intestine-specific genes. (A) Percentage of genes annotated with ELT-2 binding in any developmental stage for either intestine-specific genes or all *C. elegans* genes. (B) *(left)* Percentage of intestine-specific genes bound by unique or multiple ELT-2 occupancy types. (right) Per-occupancy percentage of intestine specific genes with unique ELT-2 occupancy **(C)** Scree plot analysis of dynamically expressed intestine-specific genes was used to determine the number of developmental expression clusters. Hierarchical clustering was performed on Euclidian distance of row Z-score normalized intestine-specific expression values using the ward.D2 method. The x-axis represents the number of clusters and y-axis represents the clustering height for the given subcluster. Vertical dotted line (red) represents the number of clusters used for further analysis (see Figure 3B).

Supplemental Figure 6. Genes with transcriptional response to *elt-2/elt-7* deletion are enriched or depleted for ELT-2 binding. (A) *(left)* Heatmap representation of genes differentially expressed any pairwise comparison between wildtype (N2), *elt-7 (tm840)*, *elt-2 (ca15)*, and *elt-7 (tm840); elt-2 (ca15)* genetic backgrounds. Transcript read counts were Z-score normalized between genetic backgrounds. The transcriptional response of each gene to deletion of *elt-2* or *elt-7*, together or separately, were hierarchically clustered using complete-linkage clustering of Spearman correlation distances. Clustering yielded six distinct sets of transcriptional responses (SET1-SET6). *(right)* Annotation of L1 stage ELT-2 binding (green) measured through ChIP-seq. **(B)** The percentage of genes within each SET assigned with an L1 stage ELT-2 ChIP-seq peak. Error bars represent the 95% confidence intervals calculated through binomial statistics. Vertical dotted line (red) represents the percentage of genes bound in all sets (SET1 through SET6, 31.4%).

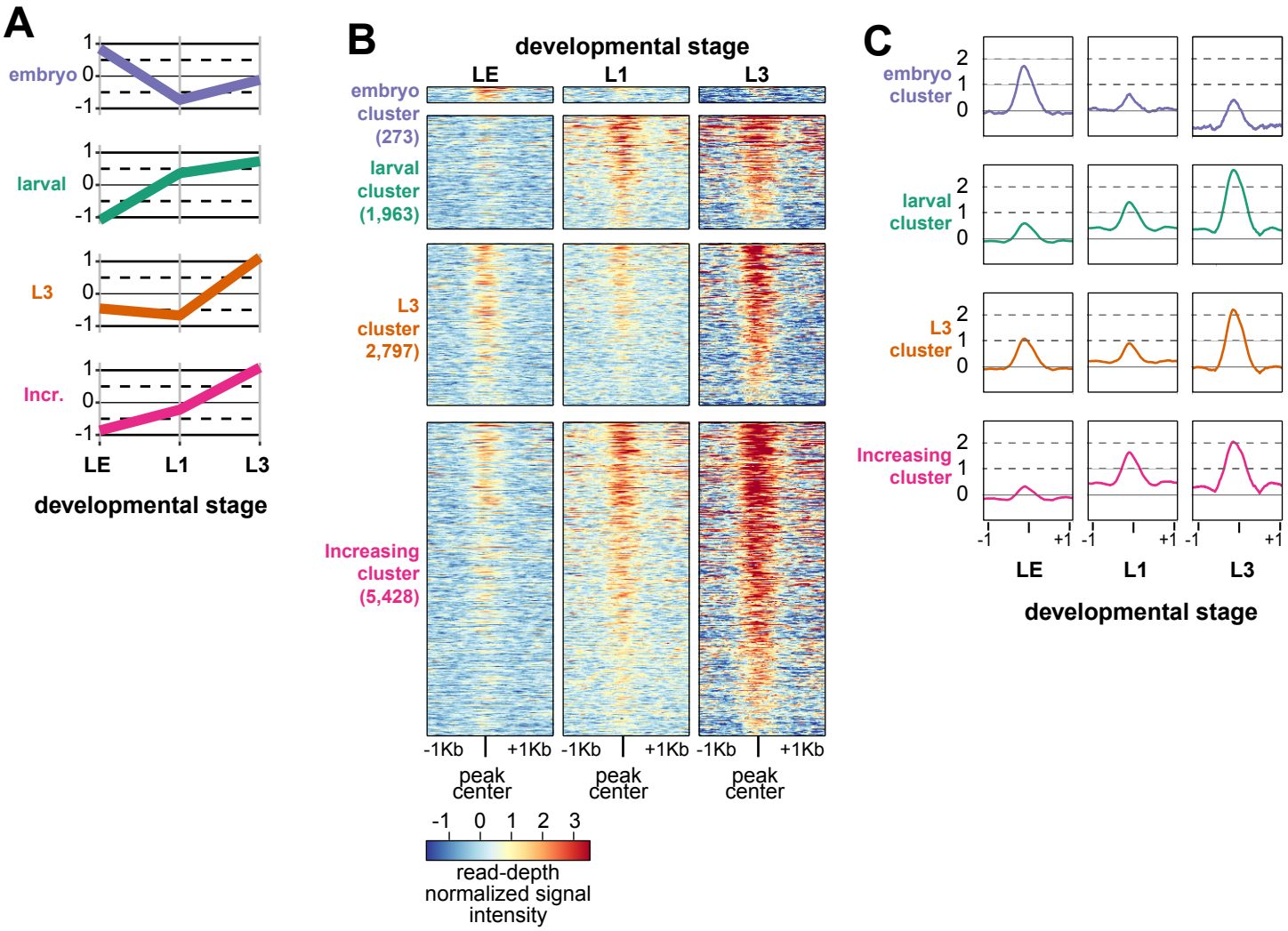
Supplemental Figure 7. Identification of intestine GRN components downstream and parallel to ELT-2 targeting. (A) *(left)* Hierarchical clustering heatmap representing ELT-2 bound transcription factor response to *elt-2* and *elt-7* mutation. Bound and differentially expressed genes were subset for known or putative transcription factors (wTF3.0). *(right)* Annotation of intestine-specific expression (blue, Table S4) and L1 stage ELT-2 ChIP-seq peak (green). (**B)** Enriched DNA motifs in ELT-2 peaks targeting bound and differentially expressed Target Class genes. MEME-ChIP was run on ELT-2 peaks of Target Class Genes A1 through C individually. The complementary sets of peaks were used as background for the discriminant objective function of MEME-ChIP. For instance, if A1 was the foreground set, the background set was a new collection of peaks A2, B, and C. Discriminant TGATAA motifs are shown for each set versus the rest. **(C)** DNA motifs corresponding to EOR-1 (E-value = 3.25e-7) were identified in Target Class A1, B and C. **(D)** DNA motif corresponding to DAF-12 (E-value = 8.37e-9) was identified exclusively in Target Class A1.

Supplemental Figure 8. Negative regulation of the *elt-2* promoter is observed in L1, L3 and adult stages. (A) Block diagram of the *elt-2* promoter reporter allele used in this study.
(B) Quantification of elt-2 promoter intestine fluorescence in either control RNAi (L4440) or *elt-2* RNAi conditions. Measurements were performed in L1, L3 and adult stages. GFP measurements were collected from background subtracted maximum Z projections from three biological replicates and normalized for intestine area (n = 30 worms for all stages and RNAi treatments). Points represent a single worm and are colored based on biological replicate. Data distribution is represented by box and whisker plots. Student's t-test was used for significance testing. Plot y-axis is log-scale. **(C)** Relative fluorescence of *elt-2* promoter reporter activity measured in *elt-2* RNAi over control RNAi in L1, L3 and adult developmental stages. Error bars represent the t-test 95% confidence interval for the ratio of means between *elt-2* RNAi and control RNAi measurements.

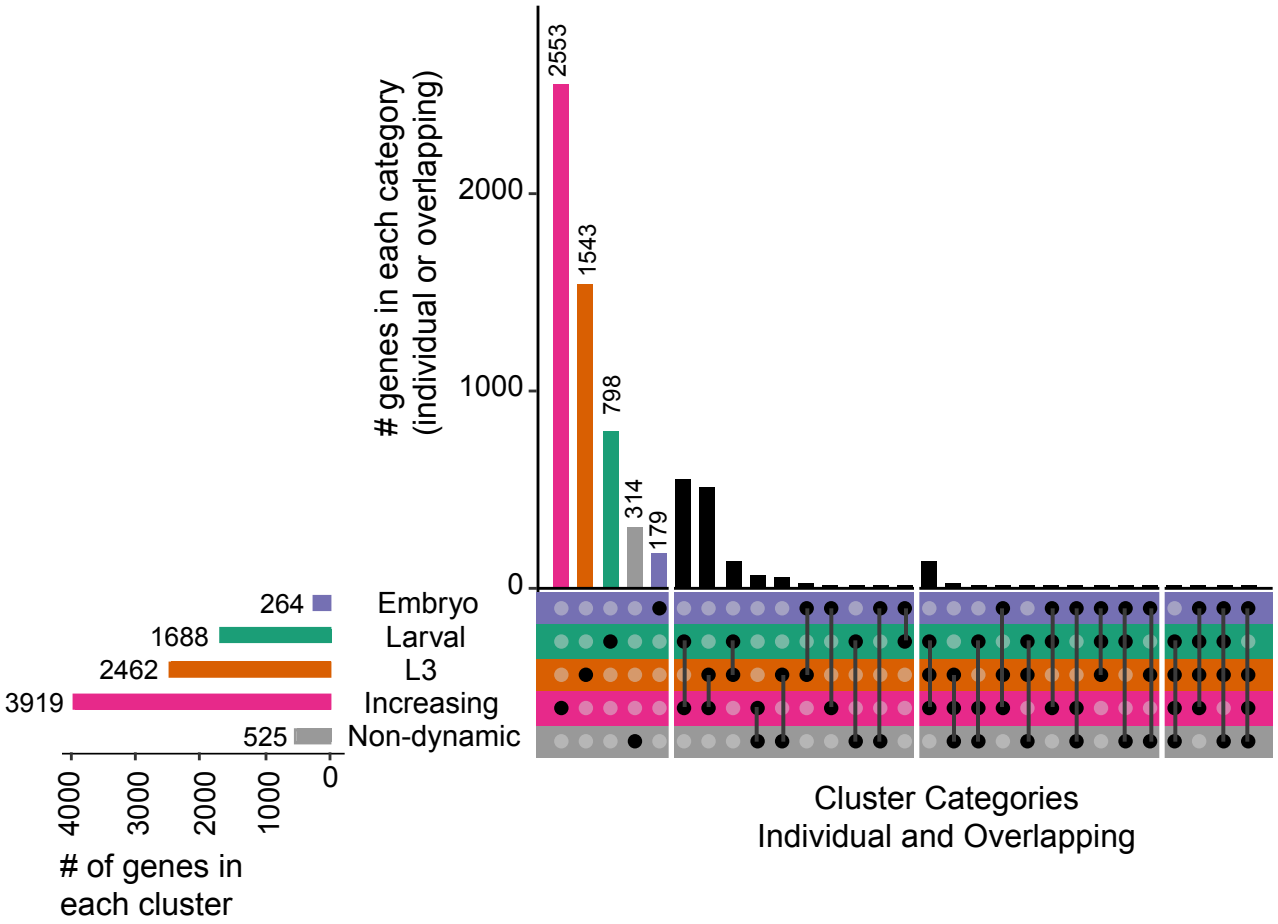
Supplemental Figure 1



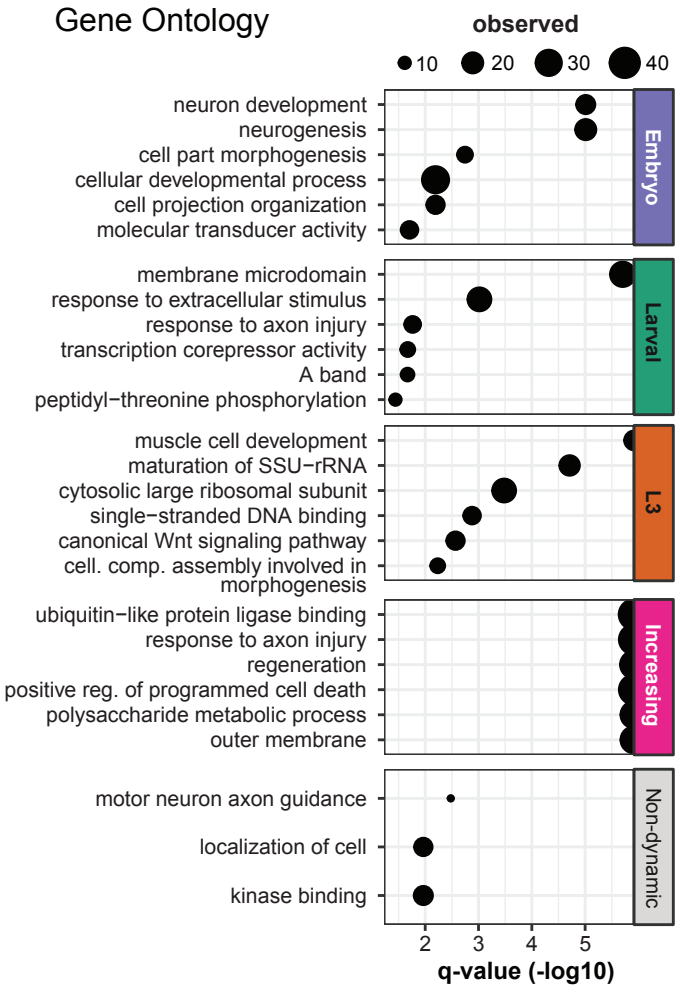
Supplemental Figure 2

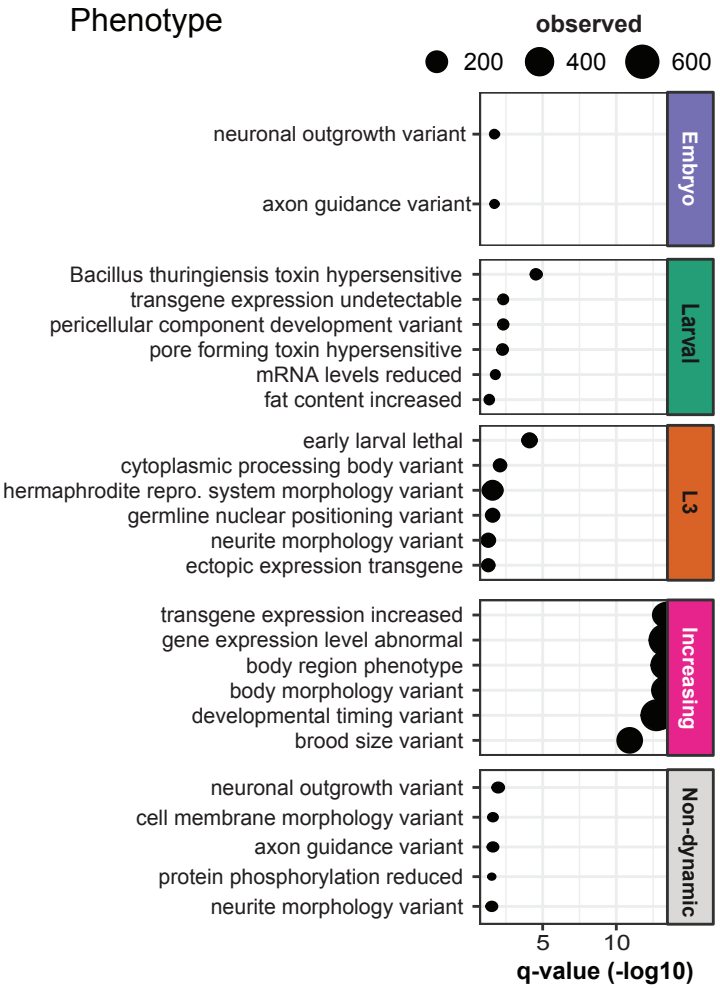


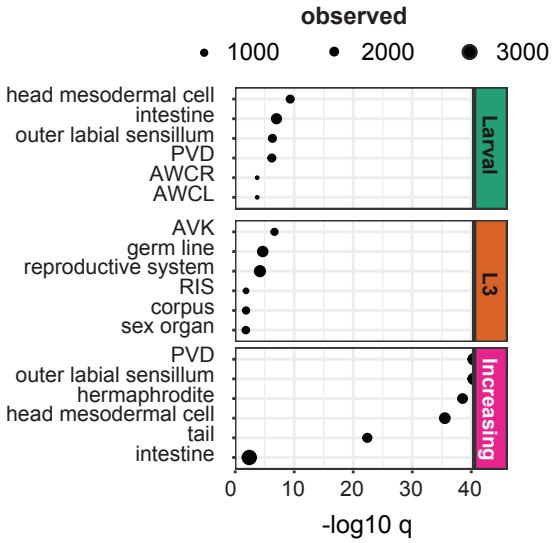
Supplemental Figure 3



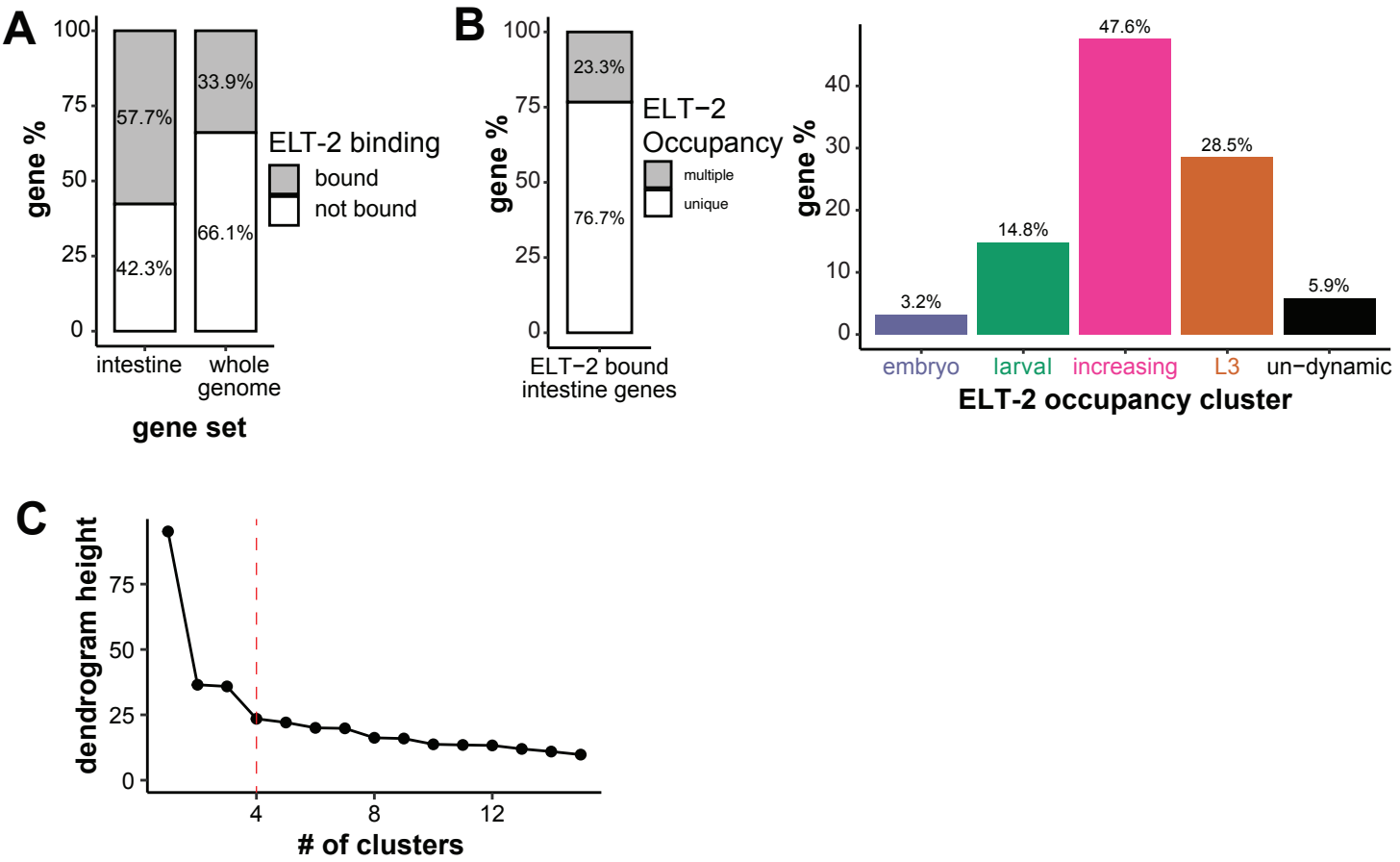
Supplemental Figure 4



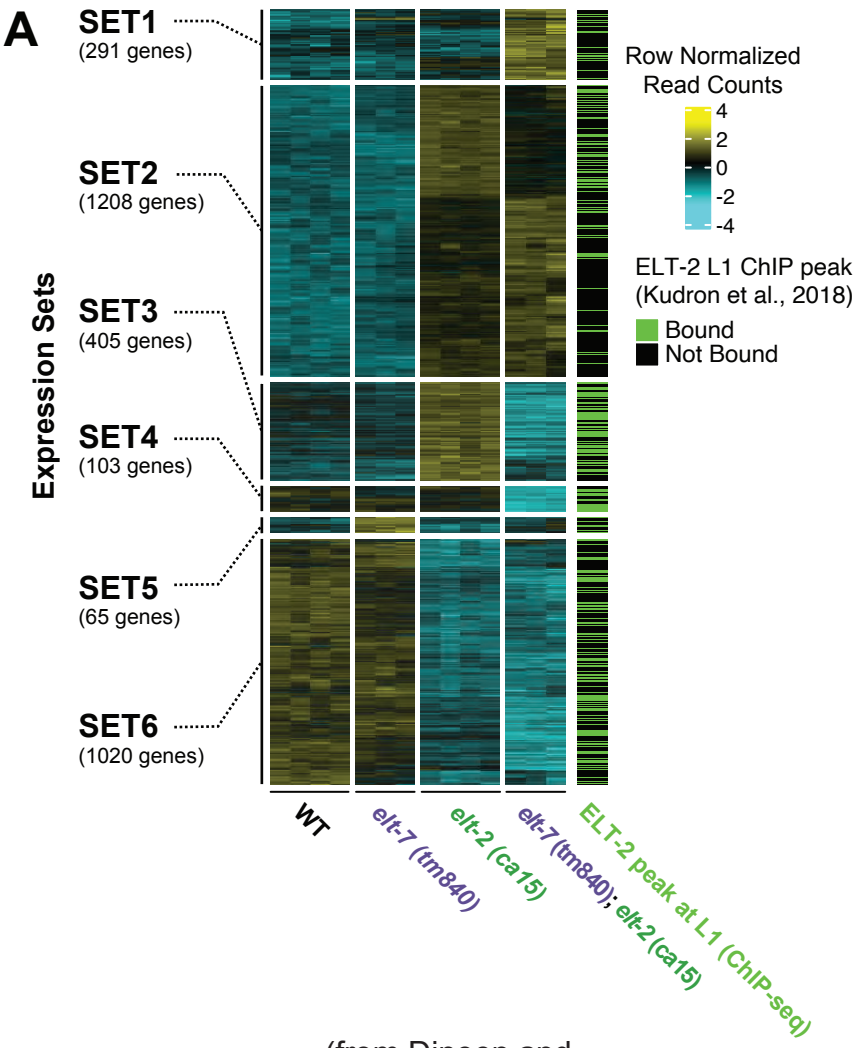




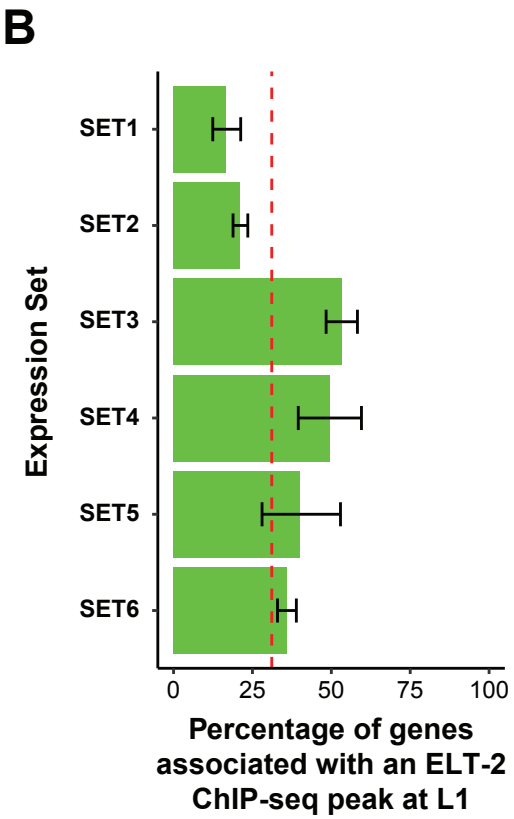
Supplemental Figure 5



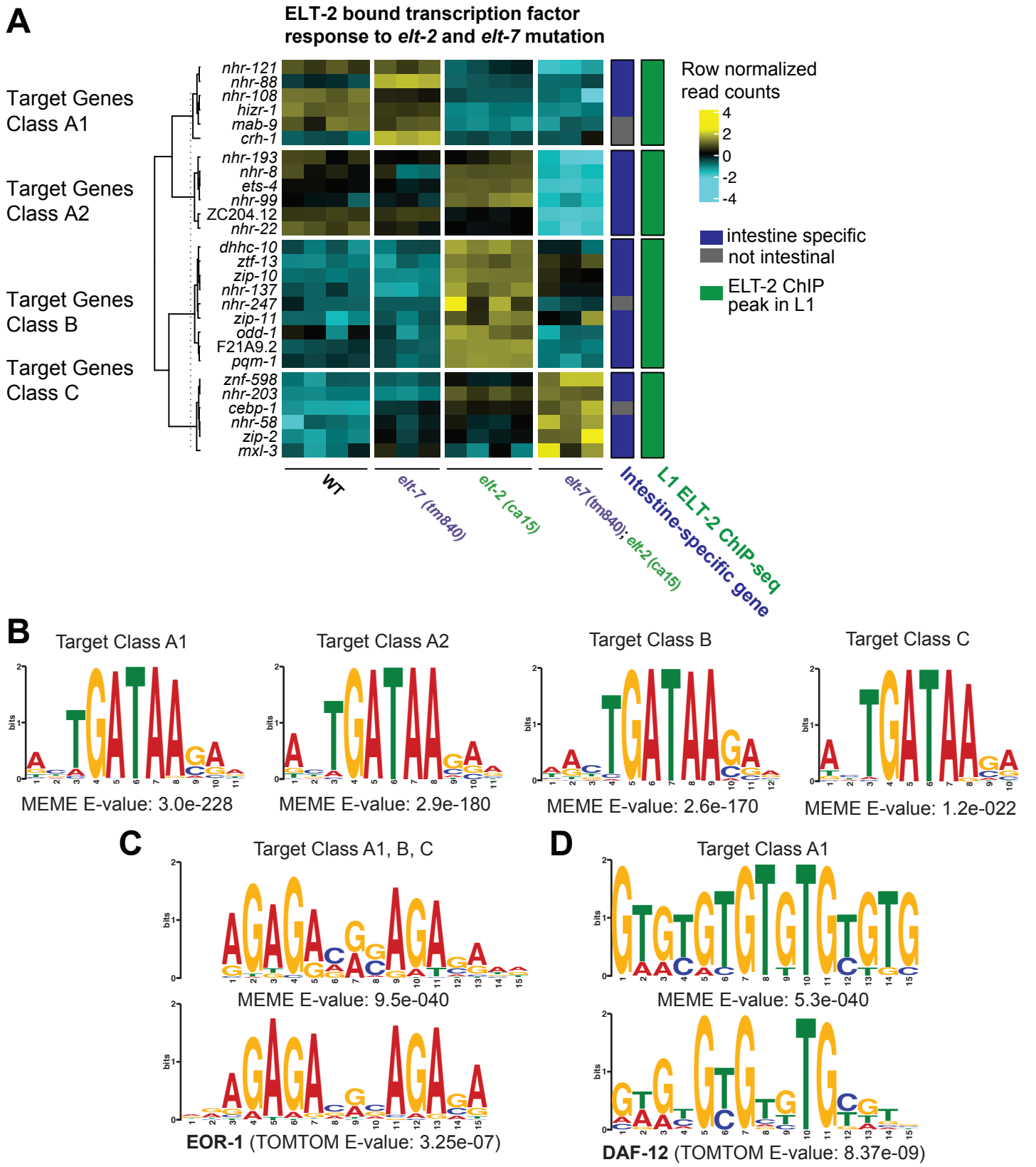
Supplemental Figure 6



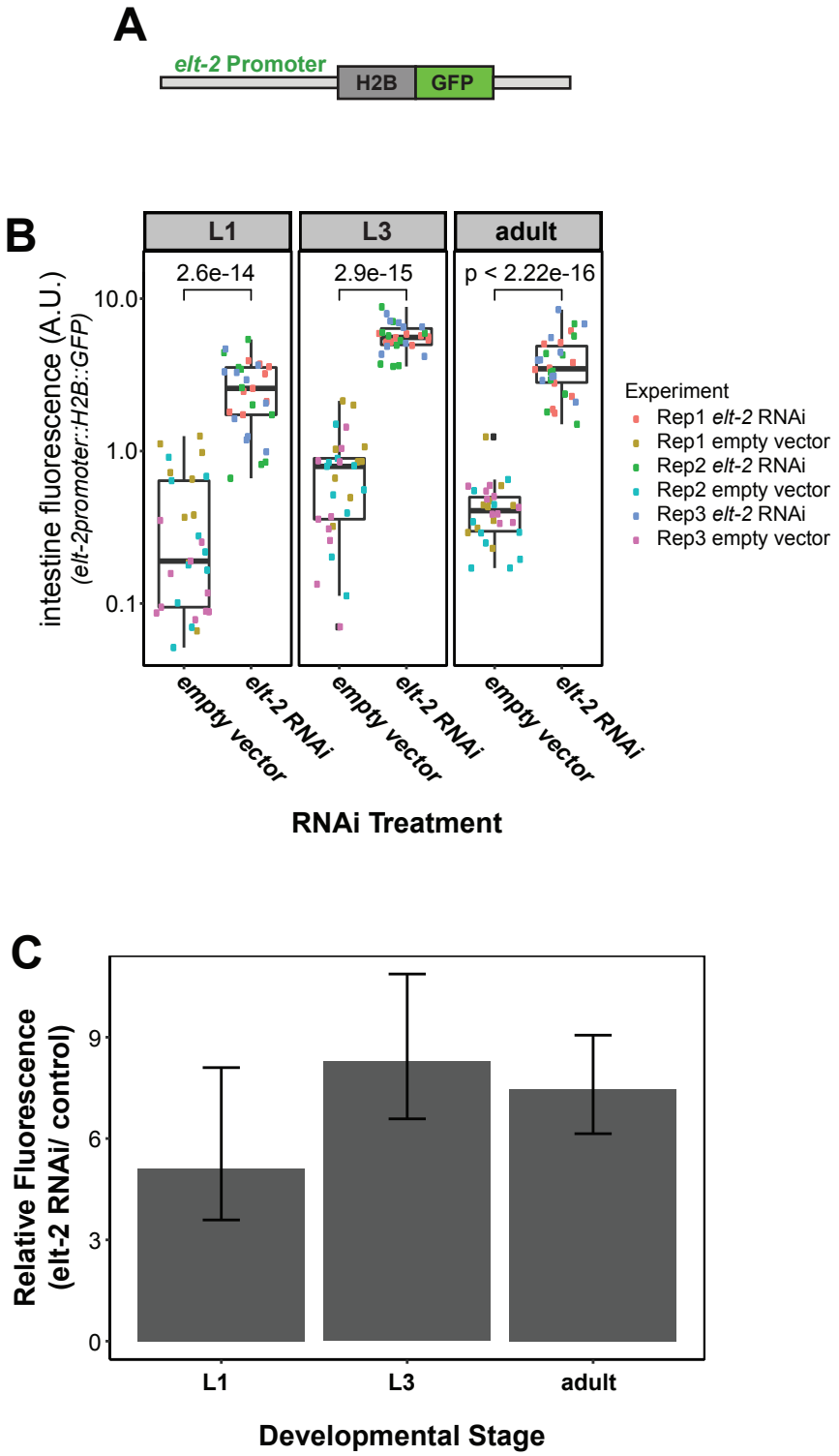
(from Dineen and
Osborne Nishimura et al., 2018)



Supplemental Figure 7



Supplemental Figure 8



Supplemental Materials & Methods:

Repository:

Data and analysis associated with this project are located here: https://github.com/meekrob/kw-onish-elt-2.

ELT-2 ChIP-seq peak calling

Alignments of the ChIP-seq reads to the ce11 genome assembly (WBcel235, February 2013) were produced by the modENCODE/modERN project (Kudron *et al.* 2017) *et al.* 2017). Data accession IDs are listed in Table S1 and downloaded from https://www.encodeproject.org/files/.

Alignments of biological replicates were compared to the alignment of the corresponding input control using the peak caller SPP version 1.16.0, with default settings (Kharchenko *et al.* 2008). This peak caller evaluates positive/negative strand peak profiles and evaluates them against a randomization model of tag densities based on the input control experiment. This procedure allows for controlling the false discovery rate at .01. The replicates between peaks called by SPP were analyzed using the IDR procedure version 2.0.4.2 using an irreproducible discovery rate of .05 (Li *et al.* 2011). SAM and BAM alignments were managed using samtools version 1.11 (Li *et al.* 2009).

Quantification of ELT-2 ChIP-seq binding signal

To integrate ELT-2 binding signal from replicates and compare them between different stages, we processed ChIP-seq reads as follows. For each control and biological replicate, the number of aligned reads were counted in 100 bp sliding windows across the genome, dividing by the total number of reads in the replicate. The script (bedToBw.sh) used bedtools version 2.29.2, bedgraphToBigWig version 377, GNU Awk (GNU Awk 4.0.2) and sort (GNU coreutils 8.22).

Identification of developmentally dynamic ELT-2 ChIP-seq peaks

To determine a set of high-confidence peaks for each stage and compare them between stages, we merged reproducible peak calls across stages. SPP peak calls between replicates were associated by overlap during IDR. Pairs of peaks between replicates whose signal passes the IDR threshold, the comprehensive range was taken (a basewise union) of the genomic region called in either replicate. This produced a set of high-confidence peaks for a given stage: 1 set each for Late Embryo, Larval Stage 1, and Larval Stage 3. These 3 sets were then merged by a basewise union using bx-python version 0.8.9 to produce the set of 11,015 regions that have a reproducible peak called in *any* of the three stages. For each of these regions, the maximum value of the binding signal (described above) in individual replicates was taken for the region using a program built on javaGenomicsToolKit (https://github.com/meekrob/java-genomics-toolkit/tree/257dd113e2).

Relative ELT-2 occupancy and clustering

We sought a representative set of values for each peak that reflects the relative occupancy of ELT-2 in Late Embryo, Larval Stage 1, and Larval Stage 3. To accomplish this, replicates were averaged to yield a 11,015 x 3 matrix. To express the values as relative occupancy, the rows were standardized by subtracting row means and dividing by row standard deviation.

The IDR-validated peak calls from Late Embryo contribute 3,389 intervals to the 11,085 total union. Although the signal of each replicate is read-depth normalized, the Late Embryo peaks show a lower signal in the called intervals, (Figure S2B Deeptools plot), presumably reflecting more concentrated signal within the called peaks of the larval time points. We wish to include these points in order to observe the LE/L1/L3 dynamic. The exclusion threshold at the .05 quantile of the overall variance is denoted $V_{.05}$=.147, includes 273 embryo-high peaks, and retains the LE timepoint for the rest of the 10,812 dynamic peaks that emphasize a larval or increasing trend.

The magnitude of that threshold can be better conceptualized when applying it to benchmark patterns which characterize the clusters. Consider, on a log scale, {1,0,0} for embryo-high, {0,1,1} for larval-high, {0,0,1} for L3 high, and {-1,0,1} for increasing. In log base 2 (the scale of the signal data), every

integral difference reflects 2 fold change. For clarity, we here reverse the log2 application and simply express the effect of the $V_{.05}$ threshold in terms of the ratio of LE to L1 to L3 (LE:L1:L3) as it would apply to the benchmark patterns listed above..
For embryo-high, $V_{.05}$ = *variance*({x, 0, 0}) yields $2^x = 1.61$, thus 1.61 : 1 : 1 benchmark pattern of change. In other words, LE is 1.61 fold greater than both L1 and L3 timepoints in the benchmark pattern. This fold change also applies to larval-high {0,1,1} and L3-high {0,0,1}, whereas increasing is $V_{.05}$= *variance*({-x,0,x}) yields $2^x = 1.31$, thus $1 : 1.31 : 1.31^2$ fold-change pattern. In other words, each time point is 1.31 times greater than the one preceding it, and L3 is $1.31^2 = 1.71$ times greater than LE in the simple pattern. However, 94% of the increasing peaks are at the benchmark of $1 : 1.61 : 1.61^2$, or rather, 1 : 1.6 : 2.59.

We then applied *k-means* (R *stats* package) clustering for different values of *k*. Scree plots showed no dramatic break between choices for the value of *k* beginning at 3, although k=4 A smaller value, *k=4*, better captured the general trends while still explaining of the 78% variance.

ELT-2 ChIP-seq peak DNA motif analysis

DNA motifs were discovered using the enrichment methods provided by the MEME Suite [ref]. Specifically, the program meme-chip isolates a 100bp center of the sequences defined by the peaks callers and uses them as the foreground sequences to pass to MEME and DREME to learn motifs defined by positition specific probability matrices (PSPMs) and regular expressions, respectively. The background sequences were either generated by shuffling the order of nucleotides in the foreground set or were supplied directly to find differences between sets or clusters. In both cases, 1st order Markov models (dinucleotide frequencies) served as probabilistic background against which the EM algorithm fit enriched models, and to evaluate a expected number of matches in the randomized sequences (e-value). For TGATAA it is X, for *eor-1* it is Y, and for *nhr-177* it is Z. Motifs were identified using the TOMTOM (Gupta et al. 2007)software, which uses the euclidean distance between the enriched motif model and libraries such as JASPAR to score similarity, and an emprical probability distribution to rate the significance of the match within its library.

Mapping ELT-2 peaks to *C. elegans* genes

Set operations and metadata were managed with the BioConductor R package *GenomicRanges* (Lawrence et al., 2013). The genomic locations of the called peaks were associated with *C. elegans* annotation freeze WS271, including only non-mitochondrial protein-coding genes. This was retrieved using the BioConductor R package *BioMart* (Durinck *et al.* 2009). Annotations between peaks and genes were completed by the BioConductor R package *ChIPpeakAnno* (Zhu *et al.* 2010). The mapping was run with different settings to assign overlap to a given gene, or nearest gene within 5KB. Peaks further than 5KB to any annotated gene start or end were assigned "unmapped".

Gene ontology analysis

Tissue of origin Gene Ontology terms for gene sets were evaluated using python package Tissue Enrichment Analysis (Angeles-Albores *et al.* 2016). This package assigns tissue-relevant terms for groups of genes utilizing a trimmed set of well-defined criteria for nematode-specific and tissue related gene ontology terms. The terms are based on WormBase expression information, and employ the hypergeometric test.

Developmental intestine-specific transcriptome analysis

Intestine specific transcriptome data was accessed by downloading supplemental files from the original publications (Table S3). Higher intestine expression stringency cutoff values was used to ensure intestine specificity. The union of intestine genes between all five datasets analyzed resulted in 5,516 genes (File S4).

Correlation of gene clusters and ELT-2 occupancy types

Categorical variables of gene expression cluster were evaluated for enrichment or depletion in ELT-2 occupancy type through chi-squared statistics. Genes were filtered for those with a single ELT-2 occupancy type. Contingency tables were generated for number of genes in an expression category with a given ELT-2 occupancy type (Table S4, S5). Pearson residuals were utilized to evaluate if a given combination were enriched or depleted in a single ELT-2 occupancy type relative to all other genes in the analysis. One-tailed p-values were calculated for Pearson residuals and adjusted for multiple test correction with the Benjamini-Hochberg method.