

RESEARCH ARTICLE

Feedforward regulatory logic controls the specification-to-differentiation transition and terminal cell fate during *Caenorhabditis elegans* endoderm development

Chee Kiang Ewe^{1,‡}, Erica M. Sommermann^{1,*‡}, Josh Kenchel^{2,3}, Sagen E. Flowers¹, Morris F. Maduro⁴, Pradeep M. Joshi¹ and Joel H. Rothman^{1,2,§}

ABSTRACT

The architecture of gene regulatory networks determines the specificity and fidelity of developmental outcomes. We report that the core regulatory circuitry for endoderm development in *Caenorhabditis elegans* operates through a transcriptional cascade consisting of six sequentially expressed GATA-type factors that act in a recursive series of interlocked feedforward modules. This structure results in sequential redundancy, in which removal of a single factor or multiple alternate factors in the cascade leads to a mild or no effect on gut development, whereas elimination of any two sequential factors invariably causes a strong phenotype. The phenotypic strength is successfully predicted with a computational model based on the timing and levels of transcriptional states. We found that one factor in the middle of the cascade, END-1, which straddles the distinct events of specification and differentiation, functions in both processes. Finally, we reveal roles for key GATA factors in establishing spatial regulatory state domains by repressing other fates, thereby defining boundaries in the digestive tract. Our findings provide a paradigm that could account for the genetic redundancy observed in many developmental regulatory systems.

KEY WORDS: Gene regulatory network, GATA factor, Cell specification, Differentiation, Endoderm, *C. elegans*

INTRODUCTION

Development is driven by the progressive deployment of transcriptional programs or gene regulatory networks (GRNs), which direct specification of progenitor cells to engender different specialized cell types within a lineage and subsequent lockdown of a unique differentiated state (Boveri, 1899; Davidson and Levine,

2008; Rothman and Jarriault, 2019). This sequential restriction of cell identity and developmental potential through dynamic changes in transcriptional states was anticipated by Waddington's epigenetic landscape, a graphic metaphor describing canalization and robustness during development (Waddington, 1957).

The generation of diverse animal forms relies largely on a common genetic toolkit. GATA transcription factors (TFs) play a conserved role in the development of diverse cell types, including those of the endoderm, the first of the three germ layers to have evolved during the late Precambrian era (Hashimshony et al., 2015; Rodaway and Patient, 2001). In the diploblastic phyla Cnidarians and Poriferans, GATA factors are specifically expressed in the endoderm or in endoderm-related cells, suggesting that these transcriptional regulators may have driven the development of the endoderm germ layer and gastrulation at the dawn of metazoan evolution (Martindale et al., 2004; Nakanishi et al., 2014). This association of GATA factors with endoderm development persists throughout metazoan phylogeny, generally through successive deployment of multiple GATA factors. The sequential use of GATA factors is particularly striking in *Caenorhabditis elegans*, in which a cascade of six GATA factor-like TFs regulates specification and differentiation of the endoderm (Maduro, 2017; Maduro and Rothman, 2002; McGhee, 2007). In *Drosophila*, the GATA factor Serpent specifies endodermal fate and activates the expression of a second GATA factor, dGATAe, which is essential for the terminal differentiation of the intestine. These factors can act across wide phylogenetic spans, as first demonstrated with the *C. elegans* GATA factor END-1, which ectopically activates endoderm development when expressed in the prospective ectoderm of *Xenopus* (Shoichet et al., 2000). Similarly, overexpression of *serpent* or *dGATAe* causes ectopic endoderm differentiation in non-endodermal lineages in *Drosophila*, as well as in *Xenopus*, further supporting the functional conservation of the GATA factors (Murakami et al., 2005; Okumura et al., 2005). In sea urchin, Blimp1/Krox1 activates *otx1*, the product of which activates *gatae* expression (Davidson et al., 2002). *Gatae* in turn provides a positive input to *otx1*, in addition to activating the transcriptional program for endoderm development, thereby forming a stable circuit in the GRN (Davidson et al., 2002; Yuh et al., 2004). Accordingly, knocking down *gatae* severely blocks endoderm development and gastrulation (Davidson et al., 2002).

The endoderm in *C. elegans*, which arises from a single blastomere formed at the eight-cell stage known as the E cell (Boveri, 1899; Sulston et al., 1983), provides a highly tractable system for investigating the mechanisms of cell specification, differentiation and organogenesis. This progenitor cell gives rise to a clone of 20 cells comprising the intestine, which are arranged in nine rings (int1-9) spanning the length of the animal (Fig. S1A). Endoderm development is driven by three pairs of duplicated genes encoding

¹Department of MCD Biology and Neuroscience Research Institute, University of California Santa Barbara, Santa Barbara, CA 93106, USA. ²Program in Biomolecular Science and Engineering, University of California Santa Barbara, Santa Barbara, CA 93106, USA. ³Chemical and Biomolecular Engineering Department, University of California Los Angeles, Los Angeles, CA 90095, USA. ⁴Molecular, Cell and Systems Biology Department, University of California Riverside, Riverside, CA 92521, USA.

*Present address: Amgen, Thousand Oaks, CA 91320, USA

‡These authors contributed equally to this work

§Author for correspondence (rothman@lifesci.ucsb.edu)

ORCID: C.K.E., 0000-0003-1973-1308; E.M.S., 0000-0003-2345-0306; J.K., 0000-0001-6378-6723; S.E.F., 0000-0002-7818-2188; M.F.M., 0000-0001-6257-7612; P.M.J., 0000-0002-4220-0559; J.H.R., 0000-0002-6844-1377

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

GATA-like TFs: the divergent MED-1 and MED-2 factors (hereby referred to as MED-1/2), and the canonical END-1/3 and ELT-2/7 factors. The maternally provided SKN-1 TF (a homologue of mammalian Nrf proteins) activates MED-1 and MED-2, which initiate specification of mesodermal fate in the EMS blastomere (Bowerman et al., 1992). In the anterior daughter of the EMS, blastomere the MS cell, the Wnt effector POP-1/Tcf represses expression of *end-1/3*, and MED-1/2 activate *tbx-35*, the product of which specifies mesodermal fate. In the E cell, which is the posterior daughter of the EMS blastomere, a triply redundant signaling system (Wnt, Src and MAPK) leads to phosphorylation of POP-1 by the nemo-like kinase LIT-1 (Bei et al., 2002; Maduro et al., 2002; Meneghini et al., 1999; Shin et al., 1999; Thorpe et al., 1997). This modified POP-1 is partially excluded from the nucleus of the E blastomere and the remaining nuclear POP-1 is converted from a repressor to an activator of endodermal fate (Maduro et al., 2005b; Shetty et al., 2005; Owraghi et al., 2010). Together with MED-1/2, Wnt-signaled POP-1 activates genes encoding the transiently expressed endoderm specification factors END-1 and END-3, which, in turn, activate the expression of ELT-7 and ELT-2, orthologues of GATA4, GATA5 and GATA6, which direct endoderm development in vertebrates (Fig. S1B). Expression of the ELT factors is sustained throughout the remainder of the life of the animal via a positive autoregulatory loop that ‘locks’ the differentiated state of the intestine (Maduro, 2017; Maduro and Rothman, 2002; McGhee, 2007). Whereas *elt-7* loss-of-function mutants do not show a discernible phenotype, animals lacking ELT-2 arrest at the earliest larval stage (L1) owing to a severely obstructed gut (Sommermann et al., 2010). Nonetheless, *elt-2(-)* animals contain a well-defined intestinal lumen and the intestinal cells appear to be fully differentiated (Fukushige et al., 1998; Sommermann et al., 2010). However, in the absence of both ELT-2 and ELT-7, the intestinal lumen is completely abolished and differentiation appears to proceed in only a subset of the endoderm-derived cells (Sommermann et al., 2010). This suggests that ELT-2 and ELT-7 function redundantly to mediate morphological differentiation of the intestine, and, in the absence of ELT-2/7, additional input(s) promote a bistable switch in the endodermal differentiation program (Dineen et al., 2018; Sommermann et al., 2010).

In this study, we sought to critically distinguish two alternative regulatory structures for the action of the GATA factors (paired redundant tiers versus sequential feedforward modules; Fig. 1A,B) in the endoderm GRN by two complementary approaches: genetic analysis of mutation combinations and computational modeling of the regulatory cascade. Both approaches strongly support an architecture consisting of a series of interlocking feedforward loops, creating ‘sequential redundancy’ in the cascade and culminating in the rapid lockdown of cell fate. We further report that END-1 acts at a transition point, participating in both specification and differentiation. Finally, we demonstrate the important roles of GATA factors in safeguarding intestinal cell fate and defining the boundaries of the digestive tract. Overall, our findings reveal the basis for the extensive genetic redundancy in the regulatory circuitry for endoderm development and how this GRN architecture dictates forward-driven cell specification and differentiation during embryogenesis, providing a paradigm for understanding the widespread redundancy of regulatory factors observed in many systems.

RESULTS

The driving goal of this study was to distinguish between two alternative models for the action of the cascade of TFs underlying

the endoderm GRN in *C. elegans*. Further, we sought to assess the basis for the transition between specification and differentiation, and the extent to which factors involved in the former also participate in the latter. Finally, we investigated the genetic basis underlying spatial partition of regulatory states in the digestive tract.

Sequential redundancy suggests feedforward regulatory circuitry in the endoderm GRN

The endoderm GRN was originally interpreted as functioning through sequential pairs of redundant factors acting in three distinct tiers (two redundant MEDs acting upon two redundant ENDS, which in turn act upon two redundant ELTs) (Fig. 1A) (Maduro and Rothman, 2002). This redundancy is evident from the phenotypes of single mutants of most of the genes expressed in the *C. elegans* endoderm GRN, which show either no overt phenotype [single *med* mutants, *end-1(-)* and *elt-7(-)*] or a weakly penetrant phenotype [*end-3(-)*]. Fully penetrant loss of endoderm specification is observed only when both *med* genes or both *end* genes are removed in pairs, leading the E blastomere to adopt the fate of the C mesodermal progenitor. Similarly, although the *elt-2(-)* mutant shows a penetrant larval arrest immediately after hatching and a dysfunctional gut, morphological differentiation of gut cells appears largely normal in these mutants; a strong gut differentiation defect requires removal of both *elt-2* and *elt-7* (Fukushige et al., 1998; Sommermann et al., 2010), indicating their redundancy in the third tier of paired redundancy.

Although these findings were consistent with the three-tiered model of redundant TF pairs, temporally resolved transcriptomic analyses revealed that the GATA factor-encoding genes are sequentially activated in the mesoderm regulatory pathway (Fig. S2) (Baugh et al., 2003; Tintori et al., 2016). This observation is consistent with the alternative possibility that, rather than pairs of factors acting together at specific tiers in the cascade (Fig. 1A), each factor is redundant with its immediate upstream or downstream factor in a continuously sequential cascade (‘sequential redundancy’) of feedforward regulatory steps (Fig. 1B). For example, as is often seen in many similarly forward-driven biological switches, each factor may act through a feedforward motif in which a given factor activates both its immediate target gene and the target of that gene (Fig. 1B) (Mangan and Alon, 2003).

To test these two alternative models – tiered versus feedforward – we systematically analyzed a series of double mutants using null mutations of all genes and compared the penetrance of lethality and extent of gut development (Tables S1 and S3). We found that unlike the mild or undetectable phenotypes observed with single mutants, removing pairs of genes that are expressed at sequential steps in the cascade invariably resulted in severely diminished viability and gut defects. Eliminating the functions of early acting factors abrogated normal E specification, as observed in *med-2(-); med-1(-)*, *end-3(-); med-1(-)* and *end-1(-) end-3(-)* double mutants. In these cases, the endoderm precursor failed to adopt the correct identity, leading to gastrulation defects and excessive epidermis, resulting in severely disrupted embryonic morphogenesis and lethality (Goldstein and Nance, 2020). In contrast, for the ‘sequential’ double mutants that did not prevent E cell specification, as evidenced by the presence of at least some differentiated gut cells and no dramatic impact on overall embryonic morphogenesis, gut lumen morphology showed periodic interruptions, with apparently undifferentiated gaps (Fig. 1C,D). This effect was associated with sporadic expression and, in some cases, complete absence of gut-specific rhabditiin granules (Clokey and Jacobson, 1986) (Fig. 1E,F). We found that expression of the gut-specific intermediate filament B-2 (IFB-2) was strongly diminished in the double mutants (Fig. S3A).

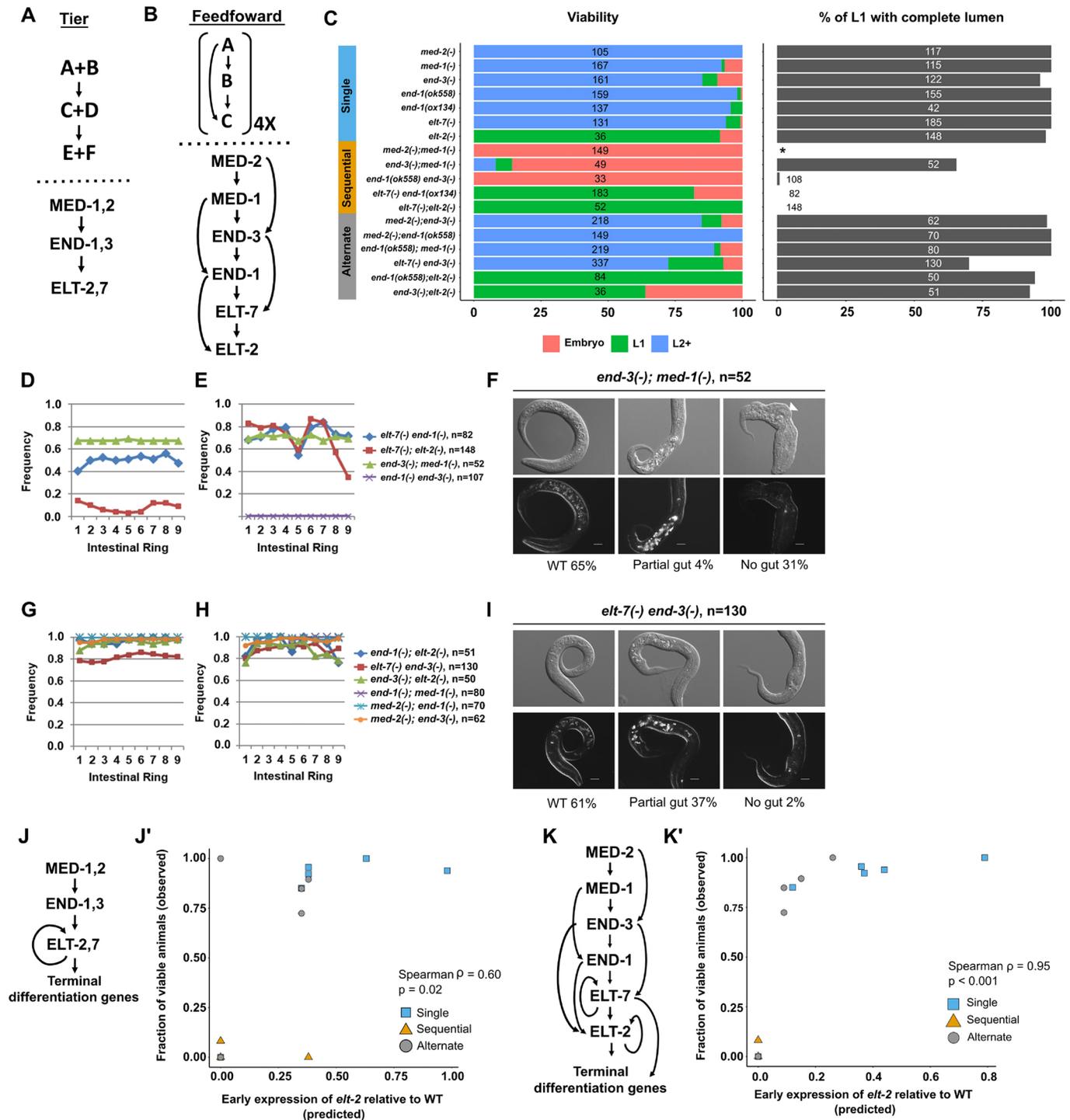


Fig. 1. Recursive feedforward loops in the endoderm GRN. (A,B) Model for (A) tiered and (B) feedforward architecture in the endoderm GRN. (C) Comparison of viability and gut lumen morphology in mutant combinations lacking one or two GATA factors in the endoderm GRN. We note that the two loss-of-function alleles of *end-1*, *ox134* and *ok558* (see Table S1), do not result in significant loss in viability (Fisher's exact test; $P=0.31$) or other discernible phenotypes. We used *ox134* in subsequent analyses unless stated otherwise (see Table S3). The total number of animals scored for each genotype is indicated. No L1 animals were scored for *med-2(-);med-1(-)* (indicated by *) as all double mutants underwent embryonic arrest owing to misspecification of MS (and E in the majority of but not all embryos). (D,E,G,H) The frequencies of visible lumen (D,G) and gut granules (E,H) are dramatically reduced in double mutants lacking sequential GATA pairs (D,E) compared with those lacking alternate members (G,H) of the endoderm cascade. (F,I) *end-3(-);med-1(-)* shows a more severe defect in intestinal differentiation than *elt-7(-) end-3(-)*. The white arrowhead indicates epidermal defect in *end-3(-);med-1(-)*, resulting from the E→C transformation. WT, wild type. Scale bars: 10 μ m. (J) Tiered model of the endoderm GRN with the terminal positive feedback loop between the ELTs which lock down gut differentiation. (J') Computed *elt-2* expression levels in the tiered model. (K) The sequential feedforward model of the endoderm GRN. (K') Computed *elt-2* expression levels in the feedforward model. All mutants show reduced *elt-2* expression relative to WT. *med-2(-);med-1(-)* double mutants show severe embryonic lethality despite *elt-2* activation owing to a fully penetrant MS→C misspecification. We excluded *med-2(-);med-1(-)* from the analyses shown in J' and K'.

Moreover, reduced apical junction molecule-1 (AJM-1) expression indicated gut epithelialization defects in the double mutants of sequential members in the endoderm GRN (Fig. S3B).

Further supporting the sequential redundancy model, the removal of two genes that function at ‘alternate’ steps in the endoderm cascade, thereby retaining an intervening factor acting between them as predicted in the feedforward model, did not result in the strong synergistic effect that we observed with the removal of pairs of sequentially expressed genes (Fig. 1C,G,H; Fig. S4). We observed normal expression of IFB-2 and AJM-1 along the length of the gut region in all such ‘alternate’ double mutants (Fig. S4). Moreover, although the *end-3(-);med-1(-)* mutants, in which sequential genes in the cascade are removed, showed severe embryonic lethality as a result of E→C misspecification, most (89.5%; $n=219$) of the *end-1(-);med-1(-)* alternate mutants developed into fertile adults (Fig. 1C). This observation is consistent with the reported divergent roles of END-1 and END-3 in endoderm specification (Boeck et al., 2011). Additionally, the alternate *elt-7(-) end-3(-)* mutants were largely viable, in contrast to the *end-3(-);med-1(-)* sequential mutants, which showed a strongly penetrant embryonic lethal phenotype (Fig. 1C). Of the hatched L1 larvae, 31% of *end-3(-);med-1(-)* sequential mutant animals exhibited no overt signs of gut differentiation at all, whereas only 2% of *elt-7(-) end-3(-)* alternate mutants completely lacked gut (Fisher’s exact test; $P<0.001$) (Fig. 1F,I). However, many *elt-7(-) end-3(-)* alternate mutants contained a partially differentiated gut (Fig. 1I). We suggest that the developmental defects observed in the most affected animals may result from suboptimal expression of *end-1* in the absence of *end-3* (Maduro et al., 2007; Raj et al., 2010). It is striking to note that some *end-3(-);med-1(-)* animals survived to adulthood; this incomplete penetrance may be attributable to sufficient activation of *end-1* by POP-1 in a minor fraction of the embryos (Raj et al., 2010).

Although none of the single mutants showed a discernible phenotype, animals lacking both END-1 and ELT-7 were completely inviable, and the resultant arrested L1 larvae showed a striking gut differentiation defect (Fig. 1C; Fig. S3). The double-mutant worms contained patches of apparently differentiated gut, as evidenced by expression of gut granules and immunoreactive IFB-2, similar to *elt-7(-);elt-2(-)* double mutants, which exhibited an all-or-none block to differentiation event along the length of the animals (Fig. 1D,E; Fig. S3) (Sommermann et al., 2010). We note that, unlike *med-2(-);med-1(-)* and *end-1(-) end-3(-)* double mutants, in which E adopts a C fate, both *elt-7(-);elt-2(-)* and *elt-7(-) end-1(-)* mutant combinations did not exhibit any sign of cell fate transformation (highly defective embryonic morphogenesis), consistent with the failure to establish a ‘post-specification’ differentiated state in these mutants. However, the differentiation defect observed in *elt-7(-) end-1(-)* double mutants appeared to be somewhat milder than that in *elt-7(-);elt-2(-)* mutants: unlike the latter, we observed defined, albeit sporadic, lumen and brush border, and the undifferentiated patches were more frequently interspersed with differentiated patches (Fig. 1D,E; Fig. S3). This partially penetrant block to endoderm differentiation in the *elt-7(-) end-1(-)* double mutants suggests a potential role for END-3 in activating *elt-2* and gut development, as discussed below.

Analysis of the two classes of double-mutant combinations – (1) sequential and (2) alternate – revealed similar effects within, but not between each class (Fig. 1C), wherein sequential double mutants were invariably more severely affected than alternate double mutants. Together, our data support sequential redundancy throughout the cascade, reflecting a recursive series of feedforward regulatory steps.

It is conceivable that such a feedforward system creates a forward-driven, rapidly deployed switch that ensures timely and robust cell fate commitment during embryogenesis.

Finally, although MED-1 and MED-2 differ by only two amino acids located outside of the DNA-binding domain and appear to perform redundant functions (Maduro et al., 2001), we were able to resolve differences in their actions, consistent with their sequential action in the cascade. In *med-2(-);med-1(-)* double mutants, all embryos died as a result of MS→C cell fate transformation (Fig. 1C), as was also the case in embryos lacking their activator, SKN-1 (Bowerman et al., 1992). In both *skn-1(-)* mutants and *med-2(-);med-1(-)* double mutants, the E cell similarly adopts a C fate in some, but not all, arrested embryos (Maduro et al., 2007). Consistent with complete redundancy, we found that removing either *med-1* or *med-2* alone did not substantially impact the expression of the endogenously tagged mNeonGreen::END-3 reporter (Fig. S5). However, our genetic analyses revealed distinguishable contributions of the two nearly identical paralogues: whereas *med-2(-);end-3(-)* double mutants showed essentially normal guts, *med-1(-);end-3(-)* double mutants showed partially defective gut development, consistent with MED-2 function preceding that of MED-1 (Fig. 1; Figs S3 and S4). Indeed, we note that, based on lineage-resolved single-cell RNA-sequencing data (Tintori et al., 2016), *med-2* transcripts are undetectable by the eight-cell stage, whereas *med-1* expression persists briefly in the endoderm precursors, suggesting that the two paralogous genes are differentially regulated.

Computational model predicts phenotypes of sequential and alternative double mutants

We took a complementary approach to testing the two alternative GRN architectures (tiered versus feedforward) (Fig. 1A,B) through a computational strategy. We constructed a mathematical model based on the network topology, in which the interactions between the GATA factors, as well as the additional POP-1-dependent activation of *end-3* and *end-1*, were written as a series of ordinary differential equations and the model parameters were determined by fitting to published transcriptomics data (Fig. S2) (Baugh et al., 2003; Tintori et al., 2016) using a custom algorithm that followed an iterative least-squares method. We then carried out *in silico* perturbations of the two models of GRN (see Materials and Methods, and Supplementary Materials and Methods) to provide predictions of these effects on the relative timing and levels of *elt-2* activation as a proxy for the final output of the network. We found that, unlike the feedforward model, the tiered model fit poorly to the transcriptomics data (Table S4). Additionally, the computed *elt-2* expression levels in the tiered model only weakly correlated with the phenotypes of the mutant combinations (Spearman’s Rank Correlation $\rho=0.60$; $P=0.02$) (Fig. 1J,J’). In contrast, when the feedforward model was tested, our computed results predicted that *elt-2* expression would occur with delayed onset in all single mutants and in mutants lacking ‘alternate’ pairs of GATA factors (Fig. 1K,K’); however, *elt-2* expression was predicted to be completely abrogated in double mutants in which ‘sequential’ members of the endoderm cascade were removed (Fig. 1K,K’), consistent with their pronounced developmental defects (Fig. 1C). To validate the performance of our computational model experimentally, we measured the expression of endogenously tagged mNeonGreen::ELT-2 in selected mutant combinations and compared the experimentally determined expression data to the computationally predicted *elt-2* levels (Fig. S6A-C). We found a striking and strong correlation between the measured *elt-2* expression levels and the computationally predicted values

obtained from the feedforward model (Spearman's Rank Correlation $\rho=1.00$; $P=0.017$) (Fig. S6C). Moreover, the predicted *elt-2* expression levels strongly correlated with the experimentally observed penetrance of the phenotypes of the single and multiple mutants (Spearman's Rank Correlation $\rho=0.95$; $P<0.001$) (Fig. 1C,K,K'): single and alternate double mutant combinations, which showed no or weak phenotypes, were predicted to express high levels of *elt-2* early, whereas sequential double mutants, which showed strong defects in gut development and were inviable, were predicted to not express *elt-2* at substantial levels at its normal time of onset. These findings, based on computational modeling with gene expression data, coupled with experimental validation, bolster the recursive feedforward model for the GATA-factor cascade.

Synergistic requirements and cross-regulatory interactions of END-1, ELT-7 and ELT-2

As described above, both *elt-7(-)* and *elt-2(-)* single mutants contained a fully differentiated gut, which showed a contiguous lumen from the pharynx to the rectum and which was surrounded by cells of normal differentiated morphology. However, *elt-7(-);elt-2(-)* double mutants invariably lacked both a defined gut lumen and at least some overtly differentiated gut cells, i.e. an apparent sporadic, all-or-none, block to gut differentiation along the length of the animal (Fig. 1C-E, Fig. 2A; Figs S3 and S7) (Sommermann et al., 2010). Although differentiation was highly defective in the absence of ELT-2 and ELT-7, patches of well-differentiated gut were nonetheless evident. Moreover, many terminal differentiation genes continue to express in the absence of ELT-2 and ELT-7 (Dineen et al., 2018). For example, eliminating the functions of ELT-2 and ELT-7 had little effect on the expression of *act-5* (Fig. S8), a gene encoding an actin isoform required for microvilli formation (Dineen et al., 2018; MacQueen et al., 2005). These observations suggest that at least one additional factor, in addition to the ELTs, may activate gut differentiation. One such candidate is END-1, which acts in specification of E-cell identity immediately upstream of the *elt* genes. Indeed, although *end-1(-)* and *elt-7(-)* mutants were both phenotypically silent, we found that the *elt-7(-)end-1(-)* double mutant showed extensive gut differentiation defects, with sporadic expression of rhabditiin granules (Fig. 1C-E, Fig. 2B,B') and IFB-2 (Fig. 2C,D; Fig. S3), as well as reduced number of differentiated gut cells (Fig. 2E; Fig. S9), reminiscent of *elt-7(-);elt-2(-)* animals.

We further tested whether END-1 could account for the residual gut-promoting activity by removing it from *elt-7(-);elt-2(-)* animals. Strikingly, simultaneously eliminating END-1, ELT-7 and ELT-2 resulted in apparent abolishment of intestinal differentiation (Fig. 2F-I). We found that 19.6% of *elt-7(-)end-1(-);elt-2(-)* triple mutants underwent embryonic arrest and the remainder died as L1 larvae ($n=143$). These animals exhibited no morphological signs of gut differentiation, such as gut granules (Fig. 2F-H; Fig. S10A,C), no detectable intestinal brush border or lumen (Fig. 2F-I; Fig. S10) and no cells expressing the gut-specific peptidase transporter OPT-2 (Fig. 2J; Fig. S11), suggesting a total block to gut differentiation. We found that knocking out *end-1* strongly reduced the expression of *act-5* (Fig. 2K,L) and *act-5* expression was further downregulated in *elt-7(-)end-1(-)* and in *elt-7(-)end-1(-);elt-2(RNAi)* animals (Fig. 2K; Fig. S12), suggesting that END-1, ELT-7 and ELT-2 act collaboratively and redundantly to mediate *act-5* expression, and that END-1 may compensate for the loss of ELT-2 and ELT-7 inputs.

To investigate whether END-3, like END-1, promotes gut morphological differentiation, we knocked down *elt-7* in *end-*

3(-);elt-2(-) double mutants and found that the arrested animals showed profound gut differentiation defects, as we observed in *end-1(-);elt-2(-);elt-7(-)* triple mutants (Fig. 2F-I; Fig. S10). Moreover, *act-5* was strongly downregulated in *end-3(-)* mutants (Fig. S13). Although these results may implicate a role of END-3 in driving differentiation, we cannot rule out that these effects were mediated primarily through END-1, as END-3 acts upstream to activate END-1 and *end-1* expression is downregulated in the absence of END-3 (Maduro et al., 2007; Raj et al., 2010). For this reason, we hereafter focused specifically on END-1 function in the specification-to-differentiation transition.

The notion for a role of END-1 as the gut-promoting factor acting in the absence of ELTs is challenged by the finding that its expression in wild-type embryos is transient to the degree that the protein is largely undetectable by the 16E embryonic stage (Li et al., 2019; Zhu et al., 1997); nonetheless, gut differentiation seen in the differentiated patches in *elt-7(-);elt-2(-)* mutants appears strong and robust late in development. These findings led us to hypothesize that ELT-2 and/or ELT-7 may normally repress *end-1* expression through negative feedback and that, in the absence of the ELTs, *end-1* may be upregulated and drive differentiation. Indeed, our preliminary results revealed that the expression of the *end-1* endogenous protein fused to a reporter was modestly elevated in 8E embryos when *elt-2* was knocked down by RNAi (Fig. 3A). Moreover, the expression levels of *end-1* were upregulated in both 4E and 8E embryos when ELT-7 and ELT-2 were simultaneously depleted (Fig. 3B). Although these findings provide initial evidence for the hypothesized feedback inhibition, the effect was rather weak, as we did not observe an obvious longer-term perdurance of *end-1* expression.

Taken together, our results suggest that END-1 (and possibly END-3) is poised at the interface between specification and differentiation in the cascade, placing it at the crux of this major developmental transition. END-1, acting with END-3, regulates specification of the E lineage, whereas END-1, acting with ELT-7 and ELT-2, controls differentiation of the intestine.

ELT-2 and ELT-7 collaborate to safeguard intestinal cell fate

Given that END-1 straddles the transition from specification of the endoderm progenitor and differentiation of the gut, we sought to investigate whether specification and differentiation involve distinct regulatory events. To do so, we examined non-endodermal (epidermal and pharyngeal) gene activity when differentiation was impaired at a point after this transition was thought to occur. When improperly specified, E undergoes extensive lineage conversion to either a C-like mesectodermal fate (when SKN-1, MED-1/2 or END-1/3 are removed), thereby improperly generating epidermis, or an MS-like mesodermal fate (for example, in the absence of Wnt signaling), thereby improperly giving rise to pharyngeal tissue (details described in Fig. 4A). In *end-1(-)end-3(-)* double mutants, misspecification of the E into a C-like cell caused severe morphological defects as a result of the supernumerary epidermal cells and consequent deformation of the epidermis (Fig. 4B; Fig. S14) (Maduro et al., 2005a). In contrast, the *elt-7(-)end-1(-);elt-2(-)* triple mutant animals, which, like the *end-1(-)end-3(-)* double mutants, do not form a discernible gut, were nonetheless not substantially defective for overall body morphogenesis (Fig. S14), implying that E was not mis-specified. Consistent with this observation, unlike the *end-1(-)end-3(-)* double mutant animals, *end-1(-);elt-2(-);elt-7(RNAi)* larvae contained a wild-type number of epidermal cells, further supporting the view that E→C

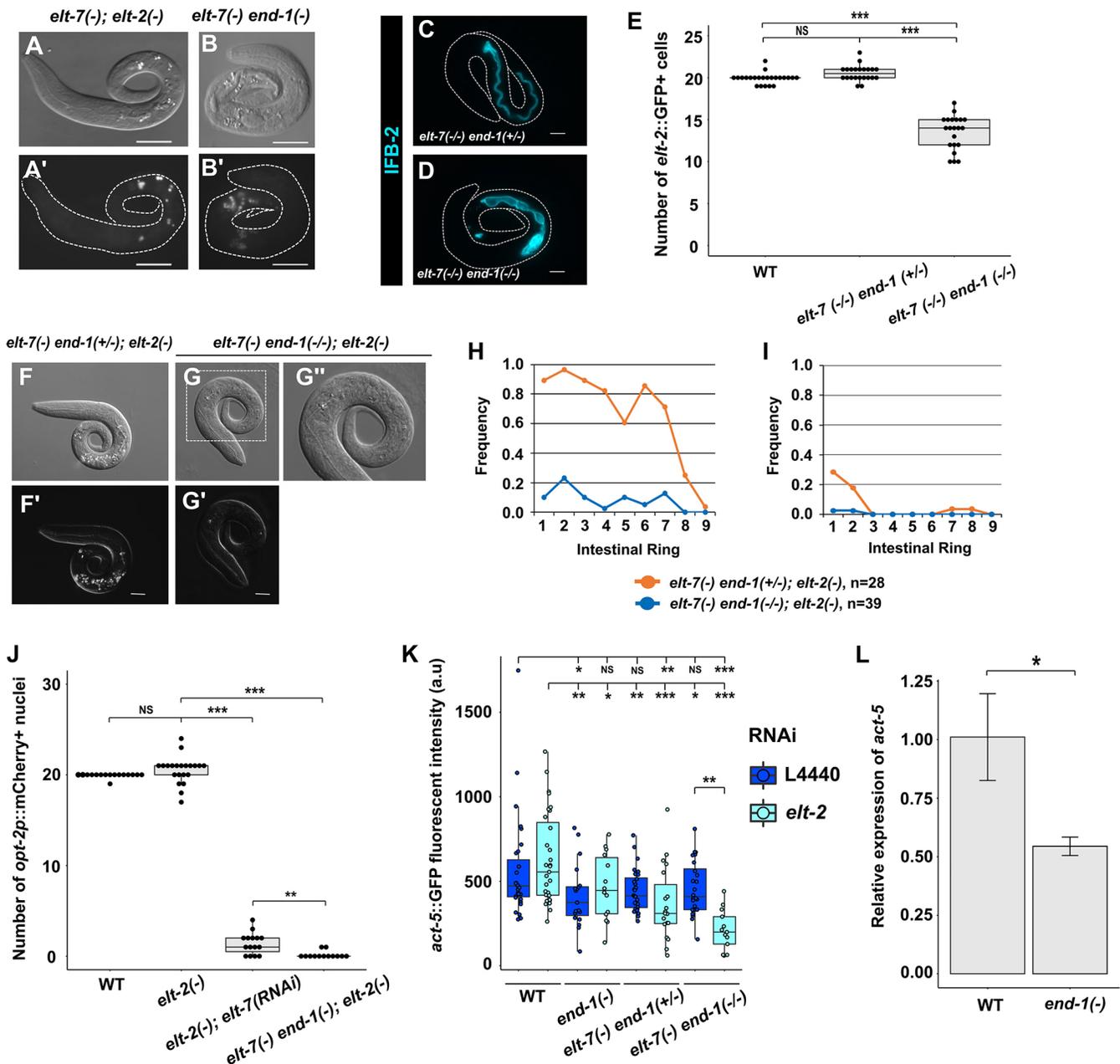


Fig. 2. Synergistic actions of END-1, ELT-7 and ELT-2 mediate morphological differentiation of endoderm. (A,B) The *elt-7(-) end-1(-)* double mutant contains a defective gut with sporadic patches of rhabditiin granules interspersed with apparently undifferentiated regions ($n=82$), similar to that seen in *elt-7(-); elt-2(-)* ($n=148$). Scale bars: 20 μm . (C,D) A balanced *elt-7(-/-) end-1(+/-)* larva shows uniform expression of IFB-2 (*kcl56*) along the length of the animal ($n=40$), whereas the *elt-7(-/-) end-1(-/-)* larva shows sporadic expression of the transgene ($n=42$). Scale bars: 10 μm . (E) WT and *elt-7(-/-) end-1(+/-)* animals contain ~20 intestinal cells, marked by the *elt-2::GFP* reporter. The number of differentiated gut cells is reduced in *elt-7(-/-) end-1(-/-)* (mean of 13.5 cells). (F,G) Representative micrographs of *elt-7(-/-) end-1(+/-); elt-2(-)* and *elt-7(-/-) end-1(-/-); elt-2(-)* mutants. *elt-7(-/-) end-1(+/-); elt-2(-)* animals (F,F') lack a visible lumen and contain sporadic birefringent granules ($n=28$). *elt-7(-/-) end-1(-/-); elt-2(-)* animals (G,G') exhibit no signs of endoderm differentiation at all ($n=39$). (G'') Magnified view of the boxed region shown in G. Scale bars: 10 μm . (H,I) The average frequencies of gut granules (H) and lumen (I) are reduced in *elt-7(-/-) end-1(-/-); elt-2(-)* compared with *elt-7(-/-) end-1(+/-); elt-2(-)*. (J) WT and *elt-2(-)* animals show similar numbers of differentiated intestinal cells, although the variance in *elt-2(-)* is significantly increased (F -test, $P<0.001$). The number of *opt-2p::mCherry* (*irSi24*)-expressing cells is reduced in *elt-2(-); elt-7(RNAi)* animals. No gut cells were detected in the vast majority of *elt-7(-) end-1(-); elt-2(-)* triple mutants. (K) The expression of *act-5::GFP* translational reporter (*jyls13*) in various mutant combinations. a.u., arbitrary units. (L) *act-5* is downregulated in *end-1(-)* mutants compared to WT, as detected by RT-qPCR. *act-1* was used as the internal reference. Three replicates were performed for each genotype. Error bars represent standard deviation. For box plots, boxes represent the 25–75th percentiles and the median is indicated. The whiskers show the $1.5\times$ interquartile range. Statistical significance was determined by non-parametric Kruskal–Wallis test and pairwise Wilcoxon tests with Benjamini–Hochberg correction (E,J), parametric one-way ANOVA followed by pairwise two-tailed unpaired t -tests with Benjamini–Hochberg correction (K) and two-tailed unpaired t -test (L). NS, not significant ($P>0.05$); * $P\leq 0.05$; ** $P\leq 0.01$; *** $P\leq 0.001$.

misspecification did not occur (Fig. 4B). These results demonstrate that END-3 alone was sufficient to promote endoderm specification over C-like mesectodermal cell fate.

Although the foregoing results suggest that the E lineage was not mis-specified when multiple late components in the cascade were eliminated, we were surprised to observe mis-expression of the

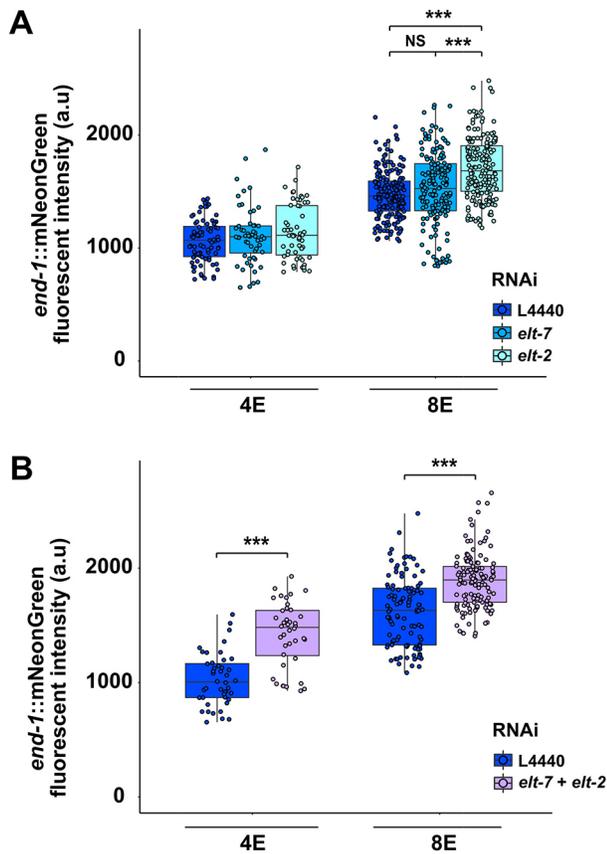


Fig. 3. ELT-2 antagonizes *end-1* expression. (A) Expression of endogenously tagged *end-1* increases in 8E embryos upon *elt-2* RNAi treatment. (B) Knocking down both *elt-7* and *elt-2* further elevates END-1 expression in both 4E and 8E embryos. Boxes represent the 25-75th percentiles and the median is indicated. The whiskers show the 1.5× interquartile range. a.u., arbitrary units. Statistical significance was determined by the Kruskal–Wallis test and pairwise Wilcoxon tests with Benjamini–Hochberg correction (A) and Wilcoxon tests (B). NS, not significant ($P>0.05$); *** $P\leq 0.001$.

pharyngeal muscle-specific myosin gene, *myo-2*, in the gut region of many *elt-7(-);elt-2(-)* and *elt-7(-) end-1(-);elt-2(-)* mutants (Fig. 4C,D). We further confirmed this observation using a different *myo-2* transgenic reporter (Fig. 4E). Additionally, we found that knocking down *elt-7* in *end-1(-);elt-2(-)* mutants caused ectopic expression of *ceh-22*, which encodes a pharynx-specific NK-2-type homeodomain protein, in the otherwise undifferentiated gut (Fig. S15). We reasoned that this inappropriate expression of pharyngeal genes was unlikely to result from an extensive E→MS transformation but may reflect later errors in the fidelity of differentiation for the following reasons: (1) Wnt/MAPK/Src signaling was unperturbed in *elt-7(-);elt-2(-)* and *elt-7(-) end-1(-);elt-2(-)*, such that the polarization of E and MS in the early embryo was expected to proceed as normal, and (2) the E→MS transformation in cell fate invariably led to profound embryonic lethality as the endodermal progenitors failed to migrate properly during gastrulation, resulting in a highly defective morphogenesis (see, for example, the ‘Mom’ phenotype in Thorpe et al., 1997). We did not observe such a phenotype in either *elt-7(-);elt-2(-)* double mutants or *elt-7(-) end-1(-);elt-2(-)* triple mutants. Finally, we found that *elt-7(-);elt-2(-)* animals contained an average of ~14 cells expressing the *elt-2* transcriptional reporter, indicating initiation of the endoderm developmental program (Fig. S7), and

a small number of these cells expressed a late gut marker (the *opt-2* reporter) (Fig. 2J; Fig. S11). Collectively, these results suggest that in *elt-7(-);elt-2(-)* animals, E and its descendants are initially specified, but fail to maintain their proper terminally differentiated fate following specification, resulting in sporadic ‘transdifferentiation’ of some of the cells.

PHA-4/FoxA is the organ selector gene that specifies pharyngeal identity, regulating *myo-2* and *ceh-22* among thousands of other targets in the pharynx (Mango et al., 1994; Zhong et al., 2010). We found that the ectopic expression of *myo-2* in *elt-7(-);elt-2(-)* double mutants was at least partially suppressed in *pha-4(RNAi)* animals (Fig. 4F). In wild-type animals, PHA-4 is expressed at high levels in the pharynx and rectum and at low levels in the intestine (Fig. 4G,G’) (Horner et al., 1998; Kalb et al., 1998). It was previously found that ELT-2 positively regulates *pha-4*, as forced expression of *elt-2* causes widespread activation of *pha-4* (Kalb et al., 1998). Paradoxically, *pha-4* transcript levels were elevated in *elt-2(-)* single mutants and *elt-7(-);elt-2(-)* double mutants (Fig. S16). Using a CRISPR-tagged PHA-4::GFP endogenous reporter, we showed that PHA-4 was indeed upregulated in the intestine of *elt-2(-)* animals (Fig. 4H,H’), and depleting *elt-7* in *elt-2(-)* animals further enhanced this effect (Fig. 4I,I’). These findings suggest that ELT-2 serves dual roles as both an activator and a repressor of *pha-4* depending on its expression levels. Thus, it appears that upregulation of *pha-4* in the midgut in the absence of ELT-2 and ELT-7 activates sporadic ectopic pharyngeal gene activity. Supporting our model, PHA-4 target genes have been shown to be regulated in part by PHA-4 binding affinity and occupancy (Fakhouri et al., 2010; Gaudet et al., 2004). Taken together, our results show that the boundaries of regulator state domains along the digestive tract of *C. elegans* are established, at least partly, by transcriptional repression mediated by ELT-2/7 in the intestine (Fig. 4J).

END-1 and ELT-7 establish the boundary between the valve and intestinal tubes

The foregoing results suggest that the core regulators involved in gut differentiation (END-1, ELT-2 and ELT-7) regulate faithful differentiation of cells in the digestive tract. The pharynx and the intestine are linked by the pharyngeal-intestinal valve (vpi), which consists of six cells arranged into three rings (Rasmussen et al., 2013) (Fig. 5A). In wild-type worms, *ajm-1::GFP* (*jcls1* transgene) is strongly expressed through the adherens junctions lining the lumen of the pharynx and the vpi, and the expression sharply drops to low levels starting at the anterior-most ring of the intestine, and remains low throughout the entire length of the animal (Köppen et al., 2001; Sommermann et al., 2010). However, although *end-1(-/-)* and *elt-7(-/-) end-1(+/-)* animals showed wild-type *ajm-1* expression pattern, the *ajm-1* signal was markedly elevated in the anterior intestinal terminus of *elt-7(-/-) end-1(-/-)* animals (Fig. 5B-E). Additionally, we observed ectopic expression of two vpi markers, *cation diffusion facilitator-1* (*cdf-1*) (Fig. 5F-H) and *heavy chain, unconventional myosin-1* (*hum-1*) (Fig. 5I-K), in the anterior terminus of *end-1(-/-) elt-7(-/-)* larvae.

It has previously been shown that worms lacking *elt-2*, similar to *elt-7(-/-) end-1(-/-)* animals, exhibit striking caudal extension of the valve cell markers (Sommermann et al., 2010). We found that ELT-2 was expressed at wild-type levels in *elt-7(-/-) end-1(-/-)* larvae, owing to its positive autoregulation (Fig. 5L). This suggests that the expansion of vpi gene expression in the intestine that we observed in *elt-7(-/-) end-1(-/-)* animals may be independent of

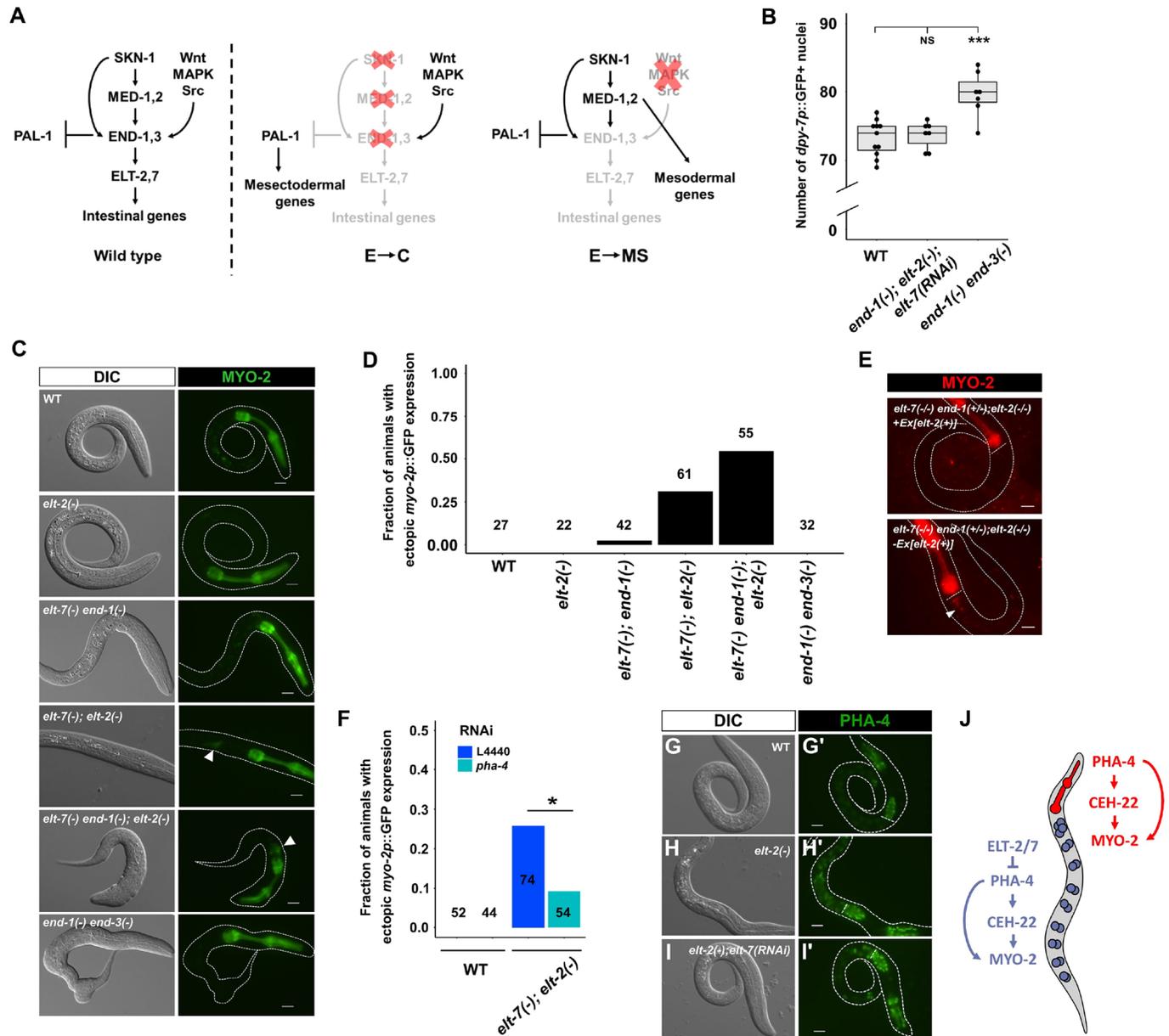


Fig. 4. ELT-2 and ELT-7 repress pharyngeal gene expression in the intestine. (A) E→C or E→MS binary fate choice during early embryogenesis. PAL-1 is required for the specification of the C blastomere, which gives rise to epidermis and body wall muscles (Hunter and Kenyon, 1996). Maternally provided *pal-1* is specifically translated in EMS and P₂. In MS and E, PAL-1 activity is blocked by TBX-35 and END-1/3, respectively (Broitman-Maduro et al., 2006). Thus, depleting TBX-35, END-1/3 or their upstream activators, MED-1/2 and SKN-1, causes excess skin and muscle owing to the misspecification of MS and/or E into C, the somatic descendant of P₂ (Hunter and Kenyon, 1996). As EMS divides, Wnt/MAPK/Src signaling from P₂ polarizes EMS, leading to the phosphorylation of POP-1 and activation of *end-1/3* in E, but not in MS. When the polarizing signal from P₂ is disrupted, POP-1 is unphosphorylated and MED-1/2 instead activate the development of MS, which gives rise to the posterior pharynx and body wall muscles (Maduro and Rothman, 2002; Maduro et al., 2002; Rocheleau et al., 1999; Shin et al., 1999). (B) Both WT and *end-1(-); elt-2(-); elt-7(RNAi)* larvae contain ~73 epidermal cells, whereas *end-1(-) end-3(-)* larvae contain ~80 epidermal cells marked by *dpy-7p::GFP*. Boxes represent the 25-75th percentiles and the median is indicated. The whiskers show the 1.5× interquartile range. NS, not significant ($P>0.05$); *** $P\leq 0.001$ (parametric one-way ANOVA followed by pairwise two-tailed unpaired *t*-tests with Benjamini–Hochberg correction). (C) Representative differential interference contrast (DIC) and fluorescence micrographs showing different mutation combinations expressing *myo-2p::GFP*. Arrowheads indicate ectopic expression of *myo-2* in *elt-7(-); elt-2(-)* and *elt-7(-) end-1(-); elt-2(-)*. (D) The frequency of animals showing mis-expression of *myo-2*, as shown in C. The number of animals scored for each genotype is indicated. (E) Fluorescence micrographs showing expression of *myo-2p::mCherry* present on *tmC12*, which balances *end-1(-)*. An *elt-7(-/-) end-1(+/-); elt-2(-/-)* animal that has lost the *elt-2(+)* rescuing array shows ectopic expression of *myo-2* in the midgut (arrowhead) ($n=50$). The white horizontal lines mark the posterior end of the pharynx. (F) Knocking down *pha-4* partially rescues ectopic expression of *myo-2* in *elt-2(-); elt-2(-)* animals. * $P\leq 0.05$ (Fisher's exact test). (G-I) The expression of the endogenously tagged *pha-4* reporter in WT (G, G') ($n=120$), *elt-2(-)* (H, H') ($n=40$) and *elt-2(-); elt-7(RNAi)* (I, I') ($n=30$) animals. The white horizontal lines in G'-I' mark the posterior end of the pharynx. Exposure time: 195 ms. All scale bars: 10 μ m. (J) Model of spatial repression and fate exclusion in the digestive tract.

ELT-2 function. It is currently unclear whether the expression of vpi reporters in the intestine reflects bona fide transformation of gut cells into valve-like cells or aberrant development of vpi and

mispositioning of excess valve cells. Regardless, our results demonstrate the important roles of intestinal GATA factors in the development of a properly patterned digestive tract.

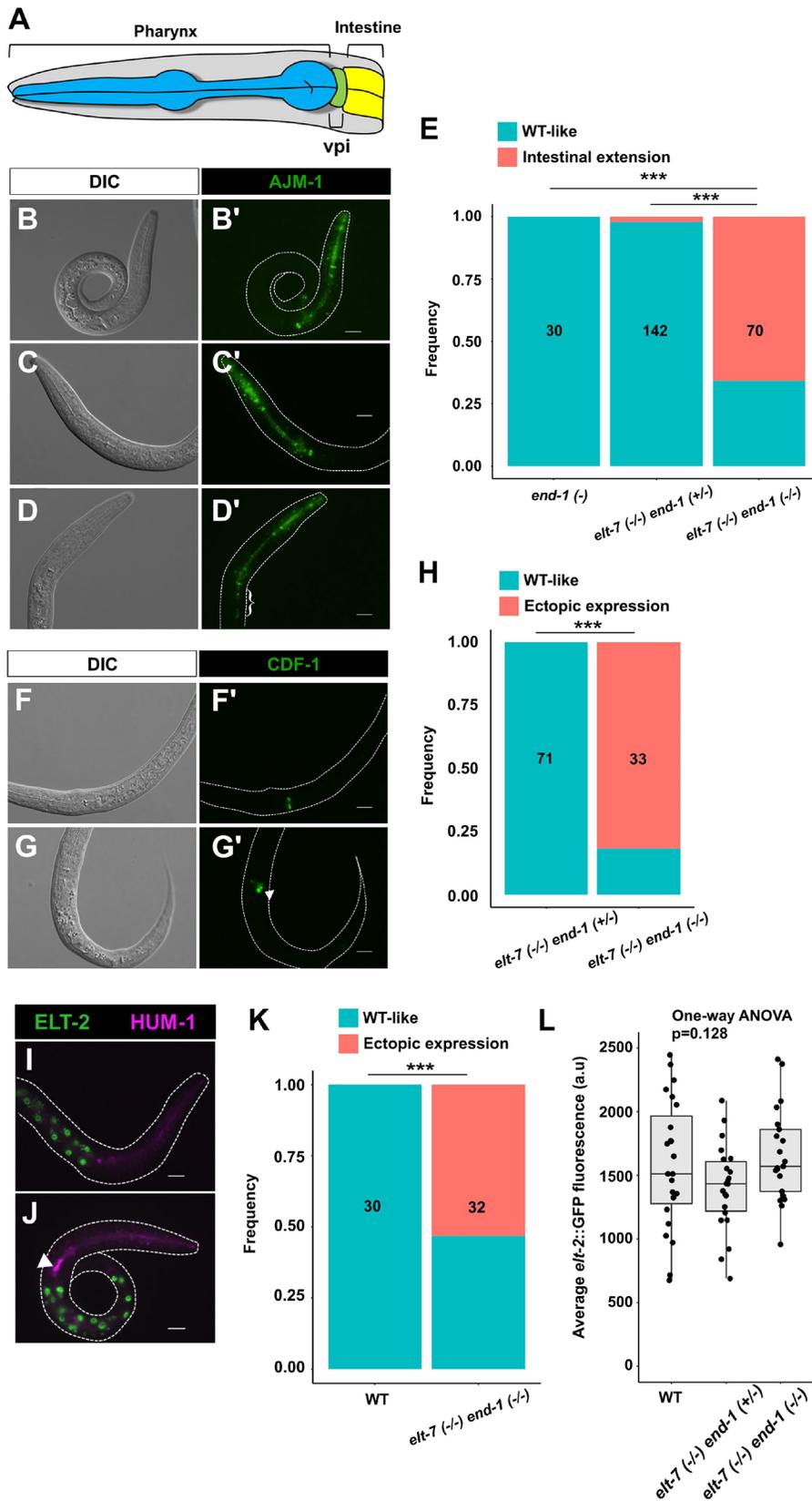


Fig. 5. ELT-7 and END-1 function synergistically to repress ectopic expression of valve cell markers in the anterior gut. (A) Schematic of the anatomy of the pharynx, vpi and intestine. (B-D) The expression of the *jcls1*[*ajm-1*::*GFP*] transgene appears similar to wild-type in *end-1*(-) (B,B') and *elt-7*(-/-) *end-1*(+/-) (C,C') animals. *elt-7*(-) *end-1*(-) (D,D') shows caudal expansion of *jcls1* expression into the anterior gut (curly bracket). (E) The frequency of animals exhibiting ectopic expression of *jcls1*. (F,F') *cdf-1*::*GFP* expression is restricted to the vpi in *elt-7*(-) *end-1*(+/-) animals. (G,G') Ectopic expression of *cdf-1* reporter is observed in *elt-7*(-/-) *end-1*(-/-) animals (arrowhead). (H) The frequency of animals with ectopic *cdf-1*::*GFP* expression. (I) HUM-1 is highly expressed in the vpi in a wild-type animal as revealed by a CRISPR-tagged endogenous reporter. (J) Ectopic expression of *hum-1* is observed in *elt-7*(-) *end-1*(-) (arrowhead). The intestinal cells are marked by *elt-2*::*GFP*. (K) Frequency of animals with ectopic *hum-1*::*RFP* expression. (L) The expression of ELT-2 is not altered in *elt-7*(-/-) *end-1*(+/-) and *elt-7*(-/-) *end-1*(-/-), compared with wild-type L1 larvae. a.u., arbitrary units. All scale bars: 10 μ m. For panels E,H,K, the number of animals scored is indicated in each graph. *** P ≤0.001 (Fisher's exact test).

DISCUSSION

The development of the *C. elegans* endoderm provides a powerful system to study the regulatory logic underlying cell specification and differentiation. In this study, we report four major findings:

- (1) the hierarchical organization of and feedforward regulatory relationship between GATA factors promote rapid lockdown of endodermal cell fate during *C. elegans* early embryogenesis;
- (2) END-1 participates in both specification and differentiation and

mediates a smooth regulatory state transition; (3) ELT-2 and ELT-7 repress expression of *pha-4* in the midgut to establish the regulatory state boundary between the pharynx and the intestine; and (4) END-1, ELT-7 and ELT-2 repress the characteristics of vpi cell fate at the anterior gut terminus, further defining the spatial domains of the foregut and midgut. Our study therefore provides an important insight into the regulatory circuits that direct the specification-to-differentiation transition and subsequent restriction and maintenance of differentiation patterns during development.

Architecture of the *C. elegans* endoderm regulatory cascade

This study provides a comprehensive overview of the core endoderm regulatory cascade, with six GATA factors acting through reiterated, sequential feedforward loops (Fig. 6). At the top of the cascade, maternally provided SKN-1 turns on the *med* and the *end* genes (Maduro et al., 2001, 2005b). Although the MED-1 and MED-2 protein sequences are nearly identical, we found distinguishable contributions between the two paralogs, with MED-2 acting upstream of MED-1. Indeed, embryos lacking MED-1 show a weaker loss-of-gut phenotype than those lacking MED-2 when SKN-1 function is impaired (Maduro et al., 2007). Moreover, *med-2* is expressed slightly earlier than *med-1* (Maduro et al., 2007; Tintori et al., 2016), suggesting that the two genes are differentially regulated, as we have observed in this study.

In the E blastomere, SKN-1 and the MEDs collaboratively activate END-3. MED-1/2 and END-3 in turn activate END-1, the next step in the cascade. Although the two endoderm-specifying factors, END-1 and END-3, perform largely overlapping functions, our data suggest that END-3 alone is sufficient to direct endoderm specification and suppress improper mesectodermal development, consistent with the lack of a detectable phenotype in *end-1(-)* single mutants. Interestingly, we found that END-1, which is poised at the interface between specification and differentiation, works synergistically with ELT-2 and ELT-7 to promote endoderm differentiation (Fig. 6). Supporting our model, *in vitro* gel-shift assays demonstrate the

binding of END-1, ELT-7 and ELT-2 to TGATAA sites, which are highly enriched in the promoters of intestinal genes (Du et al., 2016; McGhee et al., 2009; Wiesenfahrt et al., 2016). Remarkably, END-1 is able to initiate endoderm differentiation in *Xenopus* embryos, demonstrating its role as a potent organ selector (Shoichet et al., 2000). Hence, it appears that specification and differentiation involve a bona fide handoff of regulatory events by END-1.

Our results indicate that multiple factors (END-1, ELT-7 and ELT-2) act synergistically to promote endoderm morphological differentiation, challenging the notion that ELT-2 is the dominant ‘organ identity factor’ for intestinal differentiation (Fukushige et al., 1998; McGhee, 2007; McGhee et al., 2007, 2009). *elt-2(-)* animals produce an apparently complete, though defective, organ and express most gut-specific genes. The essential function for ELT-2 in widespread gut morphological differentiation is revealed only when its strongly synergistic action with ELT-7 function is eliminated (our observations) (Sommermann et al., 2010). Although the majority of hatched *end-1(-);elt-2(-)* larvae contain a continuous lumen, many *elt-7(-) end-3(-)* animals show significant gut defects, demonstrating that ELT-2 alone is insufficient to drive robust gut development. ELT-7, when expressed under the control of the *end-1* and *elt-2* promoters, can replace all other GATA factors in the GRN (Dineen et al., 2018). Additionally, overexpression of ELT-7 may cause widespread transdifferentiation of fully differentiated post-mitotic cells, showing that ELT-7 is a potent driver of intestinal differentiation (Riddle et al., 2013, 2016).

Regulatory logic of a developmental GRN

Our results support the conclusion that the *C. elegans* endoderm GRN comprises a recursive series of feedforward steps, culminating in rapid terminal differentiation (Fig. 6). Each GATA factor in the cascade receives redundant activating inputs acting through an ‘OR’ logic gate. Consequently, any single mutation in the regulatory cascade is largely phenotypically silent, with the exception of *elt-2(-)*; however, even *elt-2(-)* animals contain what appears to be a well-differentiated, albeit dysfunctional, intestine. Coherent feedforward motifs of the type that we observed reiteratively in the endoderm GRN are ubiquitous in developmental GRNs. Such a network configuration appears to result in a rapid response to an activating signal and a delayed response when the inputs are removed (termed sign-sensitive delay), thereby prolonging the effect of a transient activator (Mangan and Alon, 2003). Additionally, such feedforward loops are effective at buffering the system against stochastic noise, thereby ensuring developmental robustness (Chepyala et al., 2016; Gui et al., 2016; Maduro, 2015). This design principle appears to be crucial to ensure timely and robust activation of *elt-7* and *elt-2*, as we and others have shown (Boeck et al., 2011; Dineen et al., 2018; Maduro et al., 2015). Timely onset of *elt-2* appears to be crucial, as its delay in early embryos has been shown to cause sustained metabolic defects in larvae, despite later reattainment of wild-type ELT-2 levels (Maduro et al., 2015; Choi et al., 2017).

Our results suggest that END-1 (and potentially END-3) mediates an all-or-none switch in the differentiation program in the absence of ELT-2 and ELT-7. END-1 straddles both specification and differentiation, being buttressed by END-3 upstream and ELT-2/7 downstream. The only difference in these END-1 functions appears to be its timing of action and partnership with another regulatory factor. Regulatory nodes in early specification can indeed directly control morphogenetic events in various contexts (Davidson, 2010; Zhu and Rosenfeld, 2004). Interestingly, although END-1 expression declines shortly after gastrulation, we observed a persistent reduction in *act-5* expression in postembryonic larvae

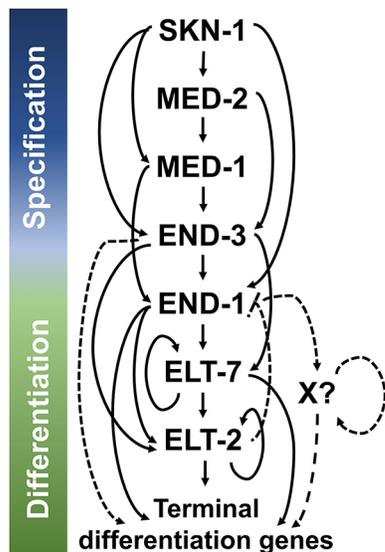


Fig. 6. Current model for the *C. elegans* endoderm GRN. Solid lines indicate known interactions identified biochemically or implied genetically, whereas dashed lines represent proposed interactions. SKN-1 initiates the GATA-driven cascade. END-1 and END-3 are also regulated by non-GATA TFs, including SPTF-3 (Sullivan-Brown et al., 2016), PAL-1 (Maduro et al., 2005b), PLP-1 (Witze et al., 2009) and POP-1 (Maduro et al., 2005b), which are omitted from this model for simplicity. See text for details.

lacking *end-1*. One possible explanation for this observation, and the finding that patches of well-differentiated gut arise in *elt-7(-); elt-2(-)* double mutants long after END-1 becomes undetectable, might be that END-1 targets an unidentified factor (indicated by X in Fig. 6) that directs differentiation in at least a subset of endodermal progenitors. Like ELT-2/7, expression of the putative factor X would be expected to be maintained through a positive autoregulatory loop, thereby continuing to promote gut differentiation after the expression of END-1 has subsided (Fig. 6) (Sommermann et al., 2010). Alternatively, END-1 may directly activate differentiation gene batteries in early embryos, and that regulatory state might be sustained through the propagation of epigenetic memory. This priming mechanism has recently been demonstrated in the specification of the ASE sensory neurons in *C. elegans* (Charest et al., 2020). Two transiently expressed T-box factors, TBX-37 and TBX-38, lock their target, *lsy-6*, in a transcriptionally active state during early embryogenesis, priming it for activation in restricted neuronal lineage (Charest et al., 2020). In mammals, Pax-7 initiates myogenic specification and differentiation. It is of relevance to note that many enhancers of the target genes of Pax-7 retain epigenetic signatures and remain active even in the absence of the initial activator (Zhang et al., 2020).

Our preliminary results showed that knocking down *elt-2* causes a slight but significant increase in *end-1* expression in early embryos; however, we did not observe obvious perdurance of END-1 when ELT-2 and ELT-7 were depleted, although we cannot rule out the possibility that the modest negative feedback we observed resulted from incomplete RNAi penetrance. Nevertheless, we propose that this modest feedback inhibition may function to facilitate the transition of regulatory states and ensure that the developmental process moves inexorably forward. For example, repression of an early cardiac specification factor, Bmp2, by the Nkx2-5 homeodomain factor is necessary for proper morphological development of the heart in mice (Prall et al., 2007). Such transcriptional repression is also frequently used to install spatial subdivision of regulatory states (Davidson, 2010). As we have shown above, ELT-2, ELT-7 and END-1 may repress alternate cell fates in the midgut and define the boundaries of the digestive tract. It is noteworthy that structurally similar regulatory circuits are repeatedly deployed in different biological networks while performing similar functions. Thus, the functional output of a GRN depends not only on the specificity of the TFs, but also the underlying circuit architecture (Davidson, 2010; Peter, 2020).

Rapid rewiring of the endoderm GRN in *Caenorhabditis*

How might a regulatory system of the type described here evolve? Effectors acting on terminal differentiation gene batteries, such as ELT-2 and PHA-4, are widely conserved across the animal kingdom, whereas the upstream inputs into GRNs appear to be recent innovations that arose during the radiation of the *Elegans* supergroup within the *Caenorhabditis* genus (Maduro, 2020). The *end* and *med* genes have been proposed to have arisen from the duplication of *elt-2*. Hence, gene duplication, coupled with *cis*-regulatory changes, may have led to the emergence of new circuitry and rewiring of the endoderm GRN in nematodes.

Intercalation of the MEDs and ENDS in the cascade may serve to buffer the system against environmental variation and developmental noise by freeing ELT-2 from direct control of SKN-1, which has been shown to play conserved pleiotropic roles in stress response and lifespan regulation (reviewed by Ewe et al., 2020, 2021). As we and others have shown, robust induction of *elt-2* is crucial to ensure the viability and fitness of the animals (Maduro

et al., 2015; Raj et al., 2010). Moreover, the deployment of MEDs and ENDS in the sequential hierarchy may allow canalization of the endoderm lineage by rapidly establishing its regulatory state in the E blastomere (Peter and Davidson, 2011). Consequently, this structure may enable increased developmental speed and early specification of the founder cells in *Caenorhabditis* species.

MATERIALS AND METHODS

C. elegans cultivation and genetics

Worm strains were cultured using standard procedures (Brenner, 1974) and all experiments were performed at room temperature (20–23°C). All genetic manipulations were performed according to standard techniques (Fay, 2013). *him-5(-)* or *him-8(-)* mutations were introduced into some strains to generate males used in crosses. Mutations and transgenes were validated by PCR and sequencing. Some lethal mutations were balanced using structurally defined balancers with fluorescent and phenotypic markers (Dejima et al., 2018). *elt-2(cal5)* was balanced with an extrachromosomal transgenic array, *irEx404 [unc-119::CFP, elt-2(+)]*. Animals that had lost the *elt-2(+)* rescuing array were CFP-negative and underwent developmental arrest (Sommermann et al., 2010). The mutations in the GATA genes are described in Table S1 (see also www.wormbase.org) and a complete list of strains used in this study is provided in Table S3. A summary of PCR primers used to detect mutations is given in Table S2.

Immunofluorescence analysis

Antibody staining with methanol-acetone fixation was performed as previously described (Sommermann et al., 2010). The antibodies MH27 (1:1000, AB_531819, Developmental Studies Hybridoma Bank), MH33 (1:1000, AB_528311, Developmental Studies Hybridoma Bank) and 455-2A4 (1:1000, AB_2618114, Developmental Studies Hybridoma Bank) were used to detect AJM-1, IFB-2 and ELT-2, respectively. Alexa Fluor 488 goat anti-mouse secondary antibody (Abcam; ab150113) was used at 1:1000 dilution.

RNAi

RNAi feeding clones were obtained from the Ahringer (Kamath et al., 2003) or the Vidal (Rual et al., 2004) libraries. The bacterial strains were inoculated overnight at 37°C in LB media containing 50 µg/ml ampicillin. The culture was then diluted 1:10 and incubated for an additional 4 h. Next, 1 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to the bacterial culture and 100 µl was seeded into 35 mm nematode growth medium (NGM) agar plates containing 1 mM IPTG and 25 µg/ml carbenicillin. For simultaneous knockdown of *elt-2* and *elt-7*, the two bacterial strains, each expressing dsRNA for one gene, were concentrated and resuspended in 1 ml of LB at a 1:1 ratio before seeding the NGM plates. Seeded plates were allowed to dry for 48 h before use. Next, 10–20 L4 animals were placed on the RNAi plates. After 24 h, the animals were transferred to fresh RNAi plates and allowed to lay eggs. The progeny were then collected for analyses.

Imaging and fluorescence quantification

The animals were immobilized using 10 mM levamisole and mounted on 4% agarose pads. Images were acquired, typically at 60×, using a Nikon Eclipse Ti-E inverted microscope fitted with an ORCA-Flash2.8 camera. For expression studies, a maximum intensity z-projection was generated with the Nikon NIS-Elements AR v4.13.05. Images were then analyzed using ImageJ or Imaris v9.7.2.

Reverse transcription quantitative PCR (RT-qPCR)

RNA was extracted from synchronized L1 animals using Monarch Total RNA Miniprep Kit (TS010S, New England Biolabs). cDNA synthesis was performed using SuperScript III First-Strand Synthesis SuperMix (Thermo Fisher Scientific, 18080400). qPCR was performed using BioRad CFX96 Real-Time System. Each 15 µl reaction contained cDNA, primers and PowerUp SYBR Green Master Mix (Thermo Fisher Scientific, A25743). The data were analyzed using the standard $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). The primer sequences used were: *act-1*,

5'-TCCATTGTCGGAAGACCACG-3' and 5'-GGTGACGATACCGT-GCTCAA-3'; *act-5*, 5'-GTCACACTACACCGTTCCAATC-3' and 5'-GTGAGGATCTTCATCATGTAGTCG-3'.

Modeling endoderm gene regulatory circuits

The topology of the gene circuits with temporal information was written as a system of differential equations, with expression of each factor dependent on the concentration of its activators (see Supplementary Materials and Methods and Table S4). The gene cascade is initiated by SKN-1, which was modeled as a square wave in the EMS blastomere. Similarly, the positive inputs of (phosphorylated) POP-1 into *end-3* and *end-1* were modeled as a square wave in the E blastomere (23 mins after the four-cell stage). Model runs were calculated as time-discretized Euler approximations (Hahn, 1991) with time steps of 0.01 s. An iterative least-squares algorithm following a modified Gauss–Newton method (Ruhe, 1979; Yip et al., 2010) was used to fit the model parameters to published transcriptomics data (Baugh et al., 2003; Tintori et al., 2016). The performance of the final model was then evaluated by comparing the predicted phenotypes of the single mutants with published results (Boeck et al., 2011; Dineen et al., 2018; Maduro et al., 2005a, 2015). Finally, predictions of *elt-2* activation (a readout for the commitment to E fate) in the mutant combinations were generated by holding the concentration of knockout genes at zero and otherwise running the model as described. The source code for this analysis is available at https://github.com/RothmanLabCode/endoderm_GRN_model.

Statistics and figure preparation

Statistical analyses were performed using R software v3.4.1 (<https://www.r-project.org/>). The specific statistical tests are reported in the figure legends. Plots were generated using R package ggplot2 or Microsoft Excel. Figures were assembled in Inkscape v0.92.4 (<https://inkscape.org/>).

Acknowledgements

We thank members of Rothman lab for helpful advice and feedback, and Cricket Wood for excellent technical support. We thank Dr Irene Chen (University of California Los Angeles) for her helpful guidance and Dr Zhuo Du (University of Chinese Academy of Sciences) for providing strain SYS549. We would also like to thank Dr Rubayn Goh, Dr Francesca Zappa, Nerea Muniozguren and Soham Chowdhury for technical assistance. We acknowledge the use of the UCSB NRI-MCDB Microscopy Facility supported by the National Institutes of Health (NIH) Grant S10OD010610-01A1, as well as WormBase. Some nematode strains used in this work were provided by the *Caenorhabditis* Genetics Center (CGC), which is funded by NIH Office of Research Infrastructure Programs Grant P40 OD010440.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: C.K.E., E.M.S., P.M.J., J.H.R.; Formal analysis: C.K.E., E.M.S., J.K.; Investigation: C.K.E., E.M.S., J.K.; Resources: C.K.E., E.M.S., S.E.F., M.F.M.; Writing - original draft: C.K.E.; Writing - review & editing: C.K.E., E.M.S., J.K., S.E.F., M.F.M., J.H.R.; Supervision: J.H.R.; Funding acquisition: J.H.R.

Funding

This work was supported by the National Institutes of Health grants (R01GM141493 and 1R01HD081266) to J.H.R. and the National Science Foundation (Division of Integrative Organismal Systems, 1258054) to M.F.M. Open Access funding provided by University of California, Santa Barbara. Deposited in PMC for immediate release.

Data availability

The source code for the GRN computational model is available on https://github.com/RothmanLabCode/endoderm_GRN_model.

Peer review history

The peer review history is available online at <https://journals.biologists.com/dev/article-lookup/doi/10.1242/dev.200337>.

References

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

Bei, Y., Hogan, J., Berkowitz, L. A., Soto, M., Rocheleau, C. E., Pang, K. M., Collins, J. and Mello, C. C. (2002). SRC-1 and Wnt signaling act together to specify endoderm and to control cleavage orientation in early *C. elegans* embryos. *Dev. Cell* **3**, 113–125. doi:10.1016/S1534-5807(02)00185-5

Boeck, M. E., Boyle, T., Bao, Z., Murray, J., Mericle, B. and Waterston, R. (2011). Specific roles for the GATA transcription factors end-1 and end-3 during *C. elegans* E-lineage development. *Dev. Biol.* **358**, 345–355. doi:10.1016/j.ydbio.2011.08.002

Boveri, T. (1899). *Die Entwicklung von Ascaris megalcephala mit besonderer Rücksicht auf die Kernverhältnisse*. Jena: Gustav Fischer (Opitz Library).

Bowerman, B., Eaton, B. A. and Priess, J. R. (1992). *skn-1*, a maternally expressed gene required to specify the fate of ventral blastomeres in the early *C. elegans* embryo. *Cell* **68**, 1061–1075. doi:10.1016/0092-8674(92)90078-Q

Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94. doi:10.1093/genetics/77.1.71

Broitman-Maduro, G., Lin, K. T.-H., Hung, W. W. K. and Maduro, M. F. (2006). Specification of the *C. elegans* MS blastomere by the T-box factor TBX-35. *Development* **133**, 3097–3106. doi:10.1242/dev.02475

Charest, J., Daniele, T., Wang, J., Bykov, A., Mandlbauer, A., Asparuhova, M., Röhsner, J., Gutiérrez-Pérez, P. and Cochella, L. (2020). Combinatorial action of temporally segregated transcription factors. *Dev. Cell* **55**, 483–499.e7. doi:10.1016/j.devcel.2020.09.002

Chepyala, S. R., Chen, Y.-C., Yan, C.-C. S., Lu, C.-Y. D., Wu, Y.-C. and Hsu, C. P. (2016). Noise propagation with interlinked feed-forward pathways. *Sci. Rep.* **6**, 23607. doi:10.1038/srep23607

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

Clokey, G. V. and Jacobson, L. A. (1986). The autofluorescent “lipofuscin granules” in the intestinal cells of *Caenorhabditis elegans* are secondary lysosomes. *Mech. Ageing Dev.* **35**, 79–94. doi:10.1016/0047-6374(86)90068-0

Davidson, E. H. (2010). Emerging properties of animal gene regulatory networks. *Nature* **468**, 911–920. doi:10.1038/nature09645

Davidson, E. H. and Levine, M. S. (2008). Properties of developmental gene regulatory networks. *Proc. Natl. Acad. Sci. USA* **105**, 20063–20066. doi:10.1073/pnas.0806007105

Davidson, E. H., Rast, J. P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C.-H., Minokawa, T., Amore, G., Hinman, V., Arenas-Mena, C. et al. (2002). A genomic regulatory network for development. *Science* **295**, 1669–1678. doi:10.1126/science.1069883

Dejima, K., Hori, S., Iwata, S., Suehiro, Y., Yoshina, S., Motohashi, T. and Mitani, S. (2018). An aneuploidy-free and structurally defined balancer chromosome toolkit for *caenorhabditis elegans*. *Cell Rep.* **22**, 232–241. doi:10.1016/j.celrep.2017.12.024

Dineen, A., Osborne Nishimura, E., Goszczynski, B., Rothman, J. H. and McGhee, J. D. (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. doi:10.1016/j.ydbio.2017.12.023

Du, L., Tracy, S. and Rifkin, S. A. (2016). Mutagenesis of GATA motifs controlling the endoderm regulator *elt-2* reveals distinct dominant and secondary cis-regulatory elements. *Dev. Biol.* **412**, 160–170. doi:10.1016/j.ydbio.2016.02.013

Ewe, C. K., Torres Cleuren, Y. N. and Rothman, J. H. (2020). Evolution and developmental system drift in the endoderm gene regulatory network of *Caenorhabditis* and other nematodes. *Front. Cell Dev. Biol.* **8**, 170. doi:10.3389/fcell.2020.00170

Ewe, C. K., Alok, G. and Rothman, J. H. (2021). Stressful development: integrating endoderm development, stress, and longevity. *Dev. Biol.* **471**, 34–48. doi:10.1016/j.ydbio.2020.12.002

Fakhouri, T. H. I., Stevenson, J., Chisholm, A. D. and Mango, S. E. (2010). Dynamic chromatin organization during foregut development mediated by the organ selector gene PHA-4/FoxA. *PLoS Genet.* **6**, e1001060. doi:10.1371/journal.pgen.1001060

Fay, D. S. (2013). *Classical Genetic Methods*, pp. 1–58. WormBook.

Fukushige, T., Hawkins, M. G. and McGhee, J. D. (1998). The GATA-factor *elt-2* is essential for formation of the *Caenorhabditis elegans* intestine. *Dev. Biol.* **198**, 286–302. doi:10.1016/S0012-1606(98)80006-7

Gaudet, J., Muttumu, S., Horner, M. and Mango, S. E. (2004). Whole-genome analysis of temporal gene expression during foregut development. *PLoS Biol.* **2**, e352. doi:10.1371/journal.pbio.0020352

Goldstein, B. and Nance, J. (2020). *Caenorhabditis elegans* gastrulation: a model for understanding how cells polarize, change shape, and journey toward the center of an embryo. *Genetics* **214**, 265–277. doi:10.1534/genetics.119.300240

Gui, R., Liu, Q., Yao, Y., Deng, H., Ma, C., Jia, Y. and Yi, M. (2016). Noise decomposition principle in a coherent feed-forward transcriptional regulatory loop. *Front. Physiol.* **7**, 600. doi:10.3389/fphys.2016.00600

Hahn, G. D. (1991). A modified Euler method for dynamic analyses. *Int. J. Numer. Methods Eng.* **32**, 943–955. doi:10.1002/nme.1620320502

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

- Horner, M. A., Quintin, S., Domeier, M. E., Kimble, J., Labouesse, M. and Mango, S. E. (1998). pha-4, an HNF-3 homolog, specifies pharyngeal organ identity in *Caenorhabditis elegans*. *Genes Dev.* **12**, 1947. doi:10.1101/gad.12.13.1947
- Hunter, C. P. and Kenyon, C. (1996). Spatial and temporal controls target pal-1 blastomere-specification activity to a single blastomere lineage in *C. elegans* embryos. *Cell* **87**, 217-226. doi:10.1016/S0092-8674(00)81340-9
- Kalb, J. M., Lau, K. K., Goszczynski, B., Fukushige, T., Moons, D., Okkema, P. G. and McGhee, J. D. (1998). pha-4 is Ce-fkh-1, a fork head/HNF-3alpha,beta,gamma homolog that functions in organogenesis of the *C. elegans* pharynx. *Development* **125**, 2171-2180. doi:10.1242/dev.125.12.2171
- Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M. et al. (2003). Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* **421**, 231-237. doi:10.1038/nature01278
- Köppen, M., Simske, J. S., Sims, P. A., Firestein, B. L., Hall, D. H., Radice, A. D., Rongo, C. and Hardin, J. D. (2001). Cooperative regulation of AJM-1 controls junctional integrity in *Caenorhabditis elegans* epithelia. *Nat. Cell Biol.* **3**, 983-991. doi:10.1038/ncb1101-983
- Li, X., Zhao, Z., Xu, W., Fan, R., Xiao, L., Ma, X. and Du, Z. (2019). Systems properties and spatiotemporal regulation of cell position variability during embryogenesis. *Cell Rep.* **26**, 313-321.e7. doi:10.1016/j.celrep.2018.12.052
- Livak, K. J. and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔCT} method. *Methods* **25**, 402-408. doi:10.1006/meth.2001.1262
- MacQueen, A. J., Baggett, J. J., Perumov, N., Bauer, R. A., Januszewski, T., Schriefer, L. and Waddle, J. A. (2005). ACT-5 is an essential *Caenorhabditis elegans* actin required for intestinal microvilli formation. *Mol. Biol. Cell* **16**, 3247-3259. doi:10.1091/mbc.e04-12-1061
- Maduro, M. F. (2015). Developmental robustness in the *Caenorhabditis elegans* embryo. *Mol. Reprod. Dev.* **82**, 918-931. doi:10.1002/mrd.22582
- Maduro, M. F. (2017). Gut development in *C. elegans*. *Semin. Cell Dev. Biol.* **66**, 3-11. doi:10.1016/j.semcdb.2017.01.001
- Maduro, M. F. (2020). Evolutionary dynamics of the SKN-1/Med/end-1,3 regulatory gene cascade in *Caenorhabditis* endoderm specification. *G3 Genes Genomes Genet* **10**, 333-356. doi:10.1534/g3.119.400724
- Maduro, M. F. and Rothman, J. H. (2002). Making worm guts: the gene regulatory network of the *caenorhabditis elegans* endoderm. *Dev. Biol.* **246**, 68-85. doi:10.1006/dbio.2002.0655
- Maduro, M. F., Meneghini, M. D., Bowerman, B., Broitman-Maduro, G. and Rothman, J. H. (2001). Restriction of mesendoderm to a single blastomere by the combined action of SKN-1 and a GSK-3β homolog is mediated by MED-1 and -2 in *C. elegans*. *Mol. Cell* **7**, 475-485. doi:10.1016/S1097-2765(01)00195-2
- Maduro, M. F., Lin, R. and Rothman, J. H. (2002). Dynamics of a developmental switch: recursive intracellular and intranuclear redistribution of *caenorhabditis elegans* POP-1 parallels Wnt-inhibited transcriptional repression. *Dev. Biol.* **248**, 128-142. doi:10.1006/dbio.2002.0721
- Maduro, M. F., Hill, R. J., Heid, P. J., Newman-Smith, E. D., Zhu, J., Priess, J. R. and Rothman, J. H. (2005a). Genetic redundancy in endoderm specification within the genus *Caenorhabditis*. *Dev. Biol.* **284**, 509-522. doi:10.1016/j.ydbio.2005.05.016
- Maduro, M. F., Kasmir, J. J., Zhu, J. and Rothman, J. H. (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. doi:10.1016/j.ydbio.2005.06.022
- Maduro, M. F., Broitman-Maduro, G., Mengarelli, I. and Rothman, J. H. (2007). Maternal deployment of the embryonic SKN-1→MED-1,2 cell specification pathway in *C. elegans*. *Dev. Biol.* **301**, 590-601. doi:10.1016/j.ydbio.2006.08.029
- Maduro, M. F., Broitman-Maduro, G., Choi, H., Carranza, F., Wu, A. C.-Y. and Rifkin, S. A. (2015). MED GATA factors promote robust development of the *C. elegans* endoderm. *Dev. Biol.* **404**, 66-79. doi:10.1016/j.ydbio.2015.04.025
- Mangan, S. and Alon, U. (2003). Structure and function of the feed-forward loop network motif. *Proc. Natl. Acad. Sci. USA* **100**, 11980-11985. doi:10.1073/pnas.2133841100
- Mango, S. E., Lambie, E. J. and Kimble, J. (1994). The pha-4 gene is required to generate the pharyngeal primordium of *Caenorhabditis elegans*. *Development* **120**, 3019-3031. doi:10.1242/dev.120.10.3019
- Martindale, M. Q., Pang, K. and Finnerty, J. R. (2004). Investigating the origins of triploblasty: 'mesodermal' gene expression in a diploblastic animal, the sea anemone *Nematostella vectensis* (phylum, Cnidaria; class, Anthozoa). *Development* **131**, 2463-2474. doi:10.1242/dev.01119
- McGhee, J. D. (2007). *The C. elegans Intestine*, pp. 1-36. WormBook.
- McGhee, J. D., Sleumer, M. C., Bilenky, M., Wong, K., McKay, S. J., Goszczynski, B., Tian, H., Krich, N. D., Khattria, J., Holt, R. A. et al. (2007). The ELT-2 GATA-factor and the global regulation of transcription in the *C. elegans* intestine. *Dev. Biol.* **302**, 627-645. doi:10.1016/j.ydbio.2006.10.024
- McGhee, J. D., Fukushige, T., Krause, M. W., Minnema, S. E., Goszczynski, B., Gaudet, J., Kohara, Y., Bossinger, O., Zhao, Y., Khattria, J. 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. doi:10.1016/j.ydbio.2008.11.034
- Meneghini, M. D., Ishitani, T., Carter, J. C., Hisamoto, N., Ninomiya-Tsuji, J., Thorpe, C. J., Hamill, D. R., Matsumoto, K. and Bowerman, B. (1999). MAP kinase and Wnt pathways converge to downregulate an HMG-domain repressor in *Caenorhabditis elegans*. *Nature* **399**, 793-797. doi:10.1038/21666
- Murakami, R., Okumura, T. and Uchiyama, H. (2005). GATA factors as key regulatory molecules in the development of *Drosophila* endoderm. *Dev. Growth Differ.* **47**, 581-589. doi:10.1111/j.1440-169X.2005.00836.x
- Nakanishi, N., Sogabe, S. and Degnan, B. M. (2014). Evolutionary origin of gastrulation: Insights from sponge development. *BMC Biol.* **12**, 26. doi:10.1186/1741-7007-12-26
- Okumura, T., Matsumoto, A., Tanimura, T. and Murakami, R. (2005). An endoderm-specific GATA factor gene, dGATAe, is required for the terminal differentiation of the *Drosophila* endoderm. *Dev. Biol.* **278**, 576-586. doi:10.1016/j.ydbio.2004.11.021
- Owraghi, M., Broitman-Maduro, G., Luu, T., Roberson, H. and Maduro, M. F. (2010). Roles of the Wnt effector POP-1/TCF in the *C. elegans* endomesoderm specification gene network. *Dev. Biol.* **340**, 209-221. doi:10.1016/j.ydbio.2009.09.042
- Peter, I. S. (2020). The function of architecture and logic in developmental gene regulatory networks. In *Current Topics in Developmental Biology* (ed. I. S. Peter), pp. 267-295. Academic Press. doi:10.1016/bs.ctdb.2020.04.001
- Peter, I. S. and Davidson, E. H. (2011). Evolution of gene regulatory networks controlling body plan development. *Cell* **144**, 970-985. doi:10.1016/j.cell.2011.02.017
- Prall, O. W. J., Menon, M. K., Solloway, M. J., Watanabe, Y., Zaffran, S., Bajolle, F., Biben, C., McBride, J. J., Robertson, B. R., Chaulet, H. et al. (2007). An Nkx2-5/Bmp2/Smad1 negative feedback loop controls heart progenitor specification and proliferation. *Cell* **128**, 947-959. doi:10.1016/j.cell.2007.01.042
- Raj, A., Rifkin, S. A., Andersen, E. and van Oudenaarden, A. (2010). Variability in gene expression underlies incomplete penetrance. *Nature* **463**, 913-918. doi:10.1038/nature08781
- Rasmussen, J. P., Feldman, J. L., Reddy, S. S. and Priess, J. R. (2013). Cell interactions and patterned intercalations shape and link epithelial tubes in *C. elegans*. *PLoS Genet.* **9**, e1003772. doi:10.1371/journal.pgen.1003772
- Riddle, M. R., Weintraub, A., Nguyen, K. C. Q., Hall, D. H. and Rothman, J. H. (2013). Transdifferentiation and remodeling of post-embryonic *C. elegans* cells by a single transcription factor. *Development* **140**, 4844-4849. doi:10.1242/dev.103010
- Riddle, M. R., Spickard, E. A., Jevince, A., Nguyen, K. C. Q., Hall, D. H., Joshi, P. M. and Rothman, J. H. (2016). Transorganogenesis and transdifferentiation in *C. elegans* are dependent on differentiated cell identity. *Dev. Biol.* **420**, 136-147. doi:10.1016/j.ydbio.2016.09.020
- Rocheleau, C. E., Yasuda, J., Shin, T. H., Lin, R., Sawa, H., Okano, H., Priess, J. R., Davis, R. J. and Mello, C. C. (1999). WRM-1 activates the LIT-1 protein kinase to transduce anterior/posterior polarity signals in *C. elegans*. *Cell* **97**, 717-726. doi:10.1016/S0092-8674(00)80784-9
- Rodaway, A. and Patient, R. (2001). Mesendoderm: an ancient germ layer? *Cell* **105**, 169-172. doi:10.1016/S0092-8674(01)00307-5
- Rothman, J. and Jarriault, S. (2019). Developmental plasticity and cellular reprogramming in *caenorhabditis elegans*. *Genetics* **213**, 723-757. doi:10.1534/genetics.119.302333
- Rual, J.-F., Ceron, J., Koreth, J., Hao, T., Nicot, A.-S., Hirozane-Kishikawa, T., Vandenhaute, J., Orkin, S. H., Hill, D. E., van den Heuvel, S. et al. (2004). Toward improving *caenorhabditis elegans* phenotype mapping with an ORFeome-based RNAi library. *Genome Res.* **14**, 2162-2168. doi:10.1101/gr.2505604
- Ruhe, A. (1979). Accelerated Gauss-Newton algorithms for nonlinear least squares problems. *BIT* **19**, 356-367. doi:10.1007/BF01930989
- Shetty, P., Lo, M.-C., Robertson, S. M. and Lin, R. (2005). *C. elegans* TCF protein, POP-1, converts from repressor to activator as a result of Wnt-induced lowering of nuclear levels. *Dev. Biol.* **285**, 584-592. doi:10.1016/j.ydbio.2005.07.008
- Shin, T. H., Yasuda, J., Rocheleau, C. E., Lin, R., Soto, M., Bei, Y., Davis, R. J. and Mello, C. C. (1999). MOM-4, a MAP kinase kinase-related protein, activates WRM-1/LIT-1 kinase to transduce anterior/posterior polarity signals in *C. elegans*. *Mol. Cell* **4**, 275-280. doi:10.1016/S1097-2765(00)80375-5
- Shoichet, S. A., Malik, T. H., Rothman, J. H. and Shivasani, R. A. (2000). Action of the *Caenorhabditis elegans* GATA factor END-1 in *Xenopus* suggests that similar mechanisms initiate endoderm development in ecdysozoa and vertebrates. *Proc. Natl. Acad. Sci. USA* **97**, 4076-4081. doi:10.1073/pnas.97.8.4076
- Sommermann, E. M., Strohmaier, K. R., Maduro, M. F. and Rothman, J. H. (2010). Endoderm development in *Caenorhabditis elegans*: The synergistic action of ELT-2 and -7 mediates the specification→differentiation transition. *Dev. Biol.* **347**, 154-166. doi:10.1016/j.ydbio.2010.08.020
- Sullivan-Brown, J. L., Tandon, P., Bird, K. E., Dickinson, D. J., Tintori, S. C., Heppert, J. K., Meserve, J. H., Trogden, K. P., Orlowski, S. K., Conlon, F. L. et al. (2016). Identifying regulators of morphogenesis common to vertebrate

- neural tube closure and *Caenorhabditis elegans* gastrulation. *Genetics* **202**, 123-139. doi:10.1534/genetics.115.183137
- Sulston, J. E., Schierenberg, E., White, J. G. and Thomson, J. N.** (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64-119. doi:10.1016/0012-1606(83)90201-4
- Thorpe, C. J., Schlesinger, A., Carter, J. C. and Bowerman, B.** (1997). Wnt signaling polarizes an early *C. elegans* blastomere to distinguish endoderm from mesoderm. *Cell* **90**, 695-705. doi:10.1016/S0092-8674(00)80530-9
- Tintori, S. C., Osborne Nishimura, E., Golden, P., Lieb, J. D. and Goldstein, B.** (2016). A transcriptional lineage of the early *C. elegans* embryo. *Dev. Cell* **38**, 430-444. doi:10.1016/j.devcel.2016.07.025
- Waddington, C.** (1957). *The Strategy of the Genes*. London: George Allen and Unwin.
- Wiesenfahrt, T., Berg, J. Y., Osborne Nishimura, E., Robinson, A. G., Goszczynski, B., Lieb, J. D. and McGhee, J. D.** (2016). The function and regulation of the GATA factor *ELT-2* in the *C. elegans* endoderm. *Development* **143**, 483-491. doi:10.1242/dev.130914
- Witze, E. S., Field, E. D., Hunt, D. F. and Rothman, J. H.** (2009). *C. elegans* *pur alpha*, an activator of *end-1*, synergizes with the Wnt pathway to specify endoderm. *Dev. Biol.* **327**, 12-23. doi:10.1016/j.ydbio.2008.11.015
- Yip, K. Y., Alexander, R. P., Yan, K.-K. and Gerstein, M.** (2010). Improved reconstruction of in silico gene regulatory networks by integrating knockout and perturbation data. *PLoS One* **5**, e8121. doi:10.1371/journal.pone.0008121
- Yuh, C.-H., Dorman, E. R., Howard, M. L. and Davidson, E. H.** (2004). An *otx* cis-regulatory module: a key node in the sea urchin endomesoderm gene regulatory network. *Dev. Biol.* **269**, 536-551. doi:10.1016/j.ydbio.2004.02.025
- Zhang, N., Mendieta-Esteban, J., Magli, A., Lijja, K. C., Perlingeiro, R. C. R., Marti-Renom, M. A., Tsirigos, A. and Dynlacht, B. D.** (2020). Muscle progenitor specification and myogenic differentiation are associated with changes in chromatin topology. *Nat. Commun.* **11**, 6222. doi:10.1038/s41467-020-19999-w
- Zhong, M., Niu, W., Lu, Z. J., Sarov, M., Murray, J. I., Janette, J., Raha, D., Sheaffer, K. L., Lam, H. Y. K., Preston, E. et al.** (2010). Genome-wide identification of binding sites defines distinct functions for *caenorhabditis elegans* *PHA-4/FOXA* in development and environmental response. *PLoS Genet.* **6**, e1000848. doi:10.1371/journal.pgen.1000848
- Zhu, X. and Rosenfeld, M. G.** (2004). Transcriptional control of precursor proliferation in the early phases of pituitary development. *Curr. Opin. Genet. Dev.* **14**, 567-574. doi:10.1016/j.gde.2004.08.006
- Zhu, J., Hill, R. J., Heid, P. J., Fukuyama, M., Sugimoto, A., Priess, J. R. and Rothman, J. H.** (1997). *end-1* encodes an apparent GATA factor that specifies the endoderm precursor in *Caenorhabditis elegans* embryos. *Genes Dev.* **11**, 2883-2896. doi:10.1101/gad.11.21.2883