

Review

Evolutionary Change in Gut Specification in *Caenorhabditis* Centers on the GATA Factor ELT-3 in an Example of Developmental System Drift

Gina Broitman-Maduro and Morris F. Maduro * 

Department of Molecular, Cell, and Systems Biology, University of California-Riverside, Riverside, CA 92521, USA; ginam@ucr.edu

* Correspondence: mmaduro@ucr.edu

Abstract: Cells in a developing animal embryo become specified by the activation of cell-type-specific gene regulatory networks. The network that specifies the gut in the nematode *Caenorhabditis elegans* has been the subject of study for more than two decades. In this network, the maternal factors SKN-1/Nrf and POP-1/TCF activate a zygotic GATA factor cascade consisting of the regulators MED-1,2 → END-1,3 → ELT-2,7, leading to the specification of the gut in early embryos. Paradoxically, the MED, END, and ELT-7 regulators are present only in species closely related to *C. elegans*, raising the question of how the gut can be specified without them. Recent work found that ELT-3, a GATA factor without an endodermal role in *C. elegans*, acts in a simpler ELT-3 → ELT-2 network to specify gut in more distant species. The simpler ELT-3 → ELT-2 network may thus represent an ancestral pathway. In this review, we describe the elucidation of the gut specification network in *C. elegans* and related species and propose a model by which the more complex network might have formed. Because the evolution of this network occurred without a change in phenotype, it is an example of the phenomenon of Developmental System Drift.

Keywords: cell specification; *C. elegans*; intestine; gene network evolution; developmental system drift



Citation: Broitman-Maduro, G.; Maduro, M.F. Evolutionary Change in Gut Specification in *Caenorhabditis* Centers on the GATA Factor ELT-3 in an Example of Developmental System Drift. *J. Dev. Biol.* **2023**, *11*, 32. <https://doi.org/10.3390/jdb11030032>

Academic Editor: Simon J. Conway

Received: 2 June 2023

Revised: 4 July 2023

Accepted: 6 July 2023

Published: 8 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Gene regulatory networks drive early embryonic development in animals [1]. Such networks evolve over time, undergoing changes in *cis*-regulation, duplications, and functional divergence, producing differences in development, or perhaps enabling phenotype plasticity [2–5]. We consider changes that involve making a simpler network more complex or that rewire major regulatory interactions (Figure 1). A surprising finding has been that changes in gene networks can occur without any apparent change in phenotype, a phenomenon termed Developmental System Drift (DSD) [6]. In these cases, networks undergo distinct evolutionary modifications but nonetheless produce the same phenotype, illustrating that there is more than one way to produce a particular developmental endpoint.

Nematodes of the genus *Caenorhabditis* exhibit examples of DSD [7,8]. We have studied the evolution of an early cell specification event in the nematode *C. elegans* and its relatives in the genus [9–16]. Here, we will review how the intestinal progenitor, E, is specified through the intersection of maternal factors, an extrinsic Wnt signal, and a gene cascade of GATA-type transcription factors. We will describe differences in the network in related species and propose a model for how the more complex network found in *C. elegans* might have evolved through changes in the function of the GATA factor ELT-3.

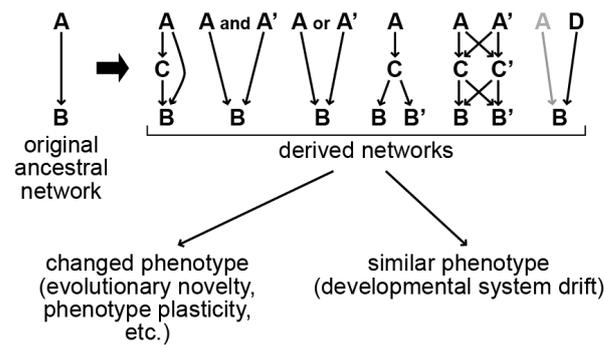


Figure 1. Models for gene network rewiring. We consider a simple gene network in which one factor, A, activates the expression of another factor, B. Networks could evolve through changes in genes by duplication (e.g., gene A giving rise to paralogues A and A'), addition of new regulators (e.g., regulators C and D), changes in *cis*-regulation (new regulatory arrows), and through replacement of factors by others (on the top right, in which regulator D becomes an activator of B followed by loss of regulatory input by A) [14]. Such changes can produce a novel phenotype, but they can also occur without causing any apparent phenotypic change, resulting in Developmental System Drift.

2. The Gut in *C. elegans*

The nematode *C. elegans* has been a model system for developmental studies for nearly a half-century [17]. Its small size, short generation time, and ability to do genetics and genome manipulation have made it possible to decipher genetic mechanisms of development [18]. Early *C. elegans* embryos undergo holoblastic cell divisions in a stereotyped pattern that is nearly invariant from animal to animal [19]. Early cells, or blastomeres, are specified very early; developing cells become progressively restricted in their fate potential by the activity of maternal, then zygotic, factors that drive cell type-specific patterns of gene expression [20,21].

The 8-cell stage blastomere E is the progenitor of the endoderm, which forms the midgut or intestine, as shown in Figure 2a [19]. The specification of E begins before its birth, when the mother cell of E, called EMS, undergoes a cell-cell interaction with its posterior neighbor, P₂ [22]. After its birth, the E cell undergoes several rounds of mitosis to produce the 20-cell intestine [19]. Even as the E descendants are undergoing mitosis, the gut primordium becomes arranged in a series of morphogenetic events to form an organized tube that will function as the intestine upon hatching [23]. Remarkably, the pattern of cell division of the E lineage can be altered considerably without affecting the morphogenesis of the intestine to form a functional organ [24,25]. Hence, following early specification, gut development can be called regulative, i.e., able to compensate for changes in the numbers of cells despite the relatively fixed number of 20–22 that are observed in the wild type [19,26].

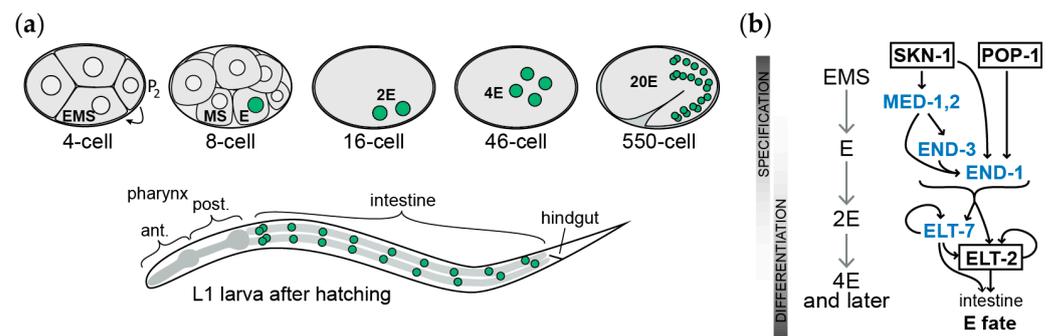


Figure 2. Origin of the gut in *Caenorhabditis* embryogenesis and the core gene specification network. (a) Origin of the gut, from the birth of the E cell at the 8-cell stage, through to the 20 cells present in the L1 larva after hatching, a developmental time of 14 h at 25 °C [19]. The nuclei of the E lineage are colored green. (b) Sequential activation of transcription factors in gut specification and early differentiation through developmental time. Maternal factors SKN-1 and POP-1 activate a cascade of upstream GATA factors (in blue), ultimately impinging on the activation of the GATA factor ELT-2.

3. Elucidation of the Gut Specification Network: Maternal Factors

A summary of the core factors in gut specification appears in Table 1. Early insights into the genetic mechanisms of blastomere specification came from screens for maternal-effect lethal mutations that changed the cell types produced by early blastomeres [27,28]. These identified two major maternal regulators that participate in gut specification. The first, SKN-1, is a bZIP/homeodomain factor that is essential for morphogenesis, pharynx specification, and gut specification most of the time [27,29]. The second, POP-1, is a TCF-like regulator that is essential for preventing specification of gut from MS, the sister cell of E, and which contributes, in parallel with SKN-1, to specification of gut from E [28,30]. Both SKN-1 and POP-1 have functions in other contexts. SKN-1 functions in stress responses and aging, similar to its mammalian orthologue, Nrf [31]. Other screens identified components of a Wnt/MAPK/Src signal that are required for the P₂-to-EMS signal that specifies the gut and which act upstream of POP-1 activity [30,32–37]. A weak input by the Caudal orthologue PAL-1 can be detected when gut specification is partially compromised [38]. Other weaker regulatory inputs into gut specification have been found that result from more global effects on embryonic gene expression. These include the Sp1 orthologue SPTF-3, the histone deacetylase HDA-1, the p300-like factor CBP-1, the Pur-alpha orthologue PLP-1, the endopeptidase TASP-1, the Prion-like-(Q/N-rich)-domain-bearing protein PQN-82, and other factors that can modulate gene expression epigenetically [39–43]. Finally, other factors have been identified that are important for restricting the activity of SKN-1 to MS and E, such as the RNA-binding proteins MEX-1 and PIE-1 [44–46].

Table 1. Core genes in endoderm specification in *Caenorhabditis*.

Gene	Loss-of-Function Phenotype in <i>C. elegans</i>	Loss-of-Function Phenotype in <i>C. briggsae</i> *	Loss-of-Function Phenotype in <i>C. angaria</i> *
Maternal factors			
<i>skn-1/Nrf</i>	Loss of MS, E fates [27]	Loss of MS, E fates [12]	No phenotype [9]
<i>pop-1/TCF</i>	Excess gut from MS [28]	Loss of gut [12]	Loss of gut [9]
Zygotic GATA factors			
<i>med-1, med-2</i>	Loss of MS, E fates [47]	n.d.	n/a
<i>end-1, end-3</i>	Loss of gut [13]	Loss of gut [13]	n/a
<i>elt-3</i>	No phenotype [48]	n.d.	Loss of gut [9]
<i>elt-7</i>	No phenotype; enhances <i>elt-2(-)</i> [49]	n.d.	n/a
<i>elt-2</i>	Incomplete gut differentiation [50]	n.d.	Incomplete gut differentiation [9]

* n.d.—not determined; n/a—not applicable (no orthologous genes).

4. Elucidation of the Gut Specification Network: Zygotic Factors

The zygotic genes that act downstream of these maternal factors were identified through a combination of forward and reverse genetics. The main specifiers of the gut progenitor E are the partially redundant GATA factors END-1 and END-3. Both are transiently expressed in the early E lineage, with *end-3* being activated slightly earlier than *end-1* [13,51–54]. Both were initially identified through large deficiencies that removed both *end* genes as well as hundreds of other genes, such that the resulting embryos did not hatch [54]. Deletion of only *end-1* and *end-3* results in a fully penetrant loss of gut specification and mostly embryonic arrest, with a small proportion of the embryos hatching as arrested larvae [16]. Instead of producing gut, the E blastomere is transformed into a C-like progenitor, producing ectopic muscle and hypodermis instead of gut [13,24,54]. Single mutants show mild effects. Loss of *end-1* has no detectable phenotype, while loss of *end-3* results in a loss of robust embryonic gut specification, altering the number of gut cells produced by the E cell from 0 to more than 30 [24,55].

The terminal regulator that directs gut fate is the GATA factor ELT-2. It was first identified as a regulator of *ges-1*, a gene encoding a gut-specific esterase that likely functions in digestion [56]. Whereas loss of the *end* genes results in a penetrant loss of all gut, deletion of *elt-2* is compatible with initial gut specification because *ges-1* is still expressed in an *elt-2* null, and animals arrest as larvae with variable amounts of hallmarks of gut differentiation, such as gut granules and luminal tissue [50]. Expression of *elt-2* is initiated by the END factors, and its expression continues throughout adulthood, maintained by positive autoregulation [50,57–59]. An adjacent gene to *elt-2*, *elt-4*, was identified that encodes an apparent nonfunctional partial duplication of *elt-2* and will not be discussed further [60].

ELT-2 shares function with another GATA factor, ELT-7, which was identified by its structural similarity to the other endodermal GATA factors [49]. While deletion of *elt-7* has no detectable phenotype, loss of *elt-7* in an *elt-2* null background increases the severity of the gut morphogenesis phenotype apparent in single *elt-2* mutants [49]. Like *elt-2*, *elt-7* is also expressed in the E lineage around the time of E specification and through adulthood.

All of END-1, END-3, ELT-2, and ELT-7 share similar structures as GATA factors and recognize similar HGATAR binding sites [11,58,59,61,62]. Consistent with this, forced individual overexpression throughout the early embryo is sufficient to promote widespread gut specification [13,16,49,50,63]. Forced overexpression of ELT-7 can also reprogram differentiated cells toward a gut fate [64]. This ability is not a widespread property of GATA factors, as overexpression of ELT-1 and ELT-3A, two hypodermis-specific factors, does not promote gut fate [50,65–67]. Interestingly, all of *end-1*, *end-3*, and *elt-7* can be replaced with

a single multicopy transgene that fuses the *end-1* promoter to the coding region of either *elt-2* or *elt-7* [42,59,65].

The last of the key regulators in gut specification were identified by a search of the genome sequence of *C. elegans* for additional GATA factors, which identified an unusual pair of unlinked, intronless, near-identical coding regions [47]. The two genes, *med-1* and *med-2*, encode divergent GATA factors that were later found to bind a target site with the core sequence RGTATAC, distinct from the canonical HGATAR site [61,68,69]. The *med* factors are expressed in the early E lineage as well as that of its sister cell, MS, and are first transcribed in the mother of both cells, EMS [47,52]. RNAi targeting both genes, or double null mutants, arrest as dead embryos lacking pharynx and muscle (tissues made by MS descendants), and a fraction of the time these also lack gut as well [47,55,70]. The *med* genes are themselves directly activated by SKN-1 [47]. In turn, the MED factors bind the promoters of several genes that are activated in the early MS and E lineages, including *end-1* and *end-3* [68,71].

As shown in Figure 2b, the genes described above can be assembled into a core gene network that describes the specification of gut in *C. elegans* through space and time, from maternal factors (SKN-1 and POP-1), early zygotic factors (MED-1,2), E-specific specification factors (END-1 and END-3), and those that maintain the differentiated state (ELT-2 and ELT-7). The transiently expressed factors upstream of ELT-2,7 can be called primarily specification factors, although even in the absence of ELT-2,7, some features of intestinal differentiation are still apparent in *elt-2,7(-)* larvae [49]. In turn, ELT-2,7, which are expressed throughout the life span and directly regulate terminal genes, are primarily differentiation factors [65,72]. They can contribute to specification, as replacing *end-1*, *end-3*, and *elt-7* with *end-1promoter::ELT-2* or *end-1promoter::ELT-7* transgenes can replace these upstream factors [59,65]. Clearly, there is enough structural similarity among these endodermal GATA factors to permit functional overlap.

Why are there so many factors? The structural similarity among the endodermal GATA factors suggests that functional redundancy contributes to making gut specification robust. Genetic studies have established that the multiple regulators in the network do work this way. Many individual factors can be deleted individually with no apparent phenotypic consequence on gut specification, such as single mutations in *med-1*, *med-2*, *end-1*, or *elt-7* [16,24,49,55]. Deletion of the remaining factors, either alone or in combination to make double mutants, results in a “stochastic” specification: Despite being genetically identical, such embryos no longer robustly specify the gut, so that some embryos make gut and others do not, and those that make gut usually have abnormal numbers of gut cells [24,55]. Examples of such genetic backgrounds include *skn-1* mutants, *end-3* mutants, *med-1; end-3* and *end-1; elt-7* double mutants, and strains engineered to remove the MED-dependent regulatory input into *end-1,3* [13,24,27,55,73]. This probabilistic, binary output of gut specification was found to result from stochastic variation in gene expression of the remaining upstream components, coupled with a threshold of required numbers of transcripts, resulting in a reduced probability of timely activation of *elt-2* [24,74]. In the case of stochastic specification that affects only the E lineage, the gut can still support development to viability. However, in such animals, the intestine is often abnormal in both the number of cells and abnormally increased storage of lipids, presumably indicative of a metabolic defect [24]. Hence, the multiple factors in the network have likely been maintained by selection for both robust development and function of the gut.

5. Evolution of Gut Specification in *C. briggsae*

The phylogeny of the *Caenorhabditis* genus consists of more than 50 identified species [75–77]. *C. elegans* and its close relatives form the Elegans supergroup of species, which consists of the Elegans and Japonica groups (Figure 3). The nematode *C. briggsae*, a close relative of *C. elegans* within the Elegans group, was the second nematode to have its genome sequenced after *C. elegans* [78,79]. *C. briggsae* is hermaphroditic like *C. elegans*, has

nearly the same embryonic development, and is amenable to many methods available in *C. elegans*, including RNA interference, so it has been a useful comparative species [80–83].

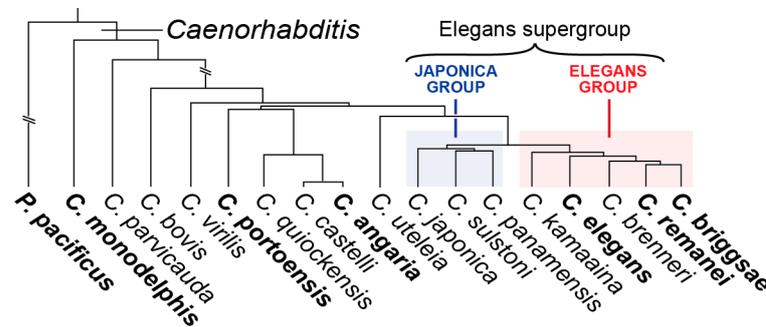


Figure 3. Simplified phylogeny of a subset of species in *Caenorhabditis* from prior publications [75–77,84]. The location of *C. virilis* is shown as deduced by the maximum likelihood as reported in [75]. The species referred to in this review are in boldface.

The zygotic gene network that specifies gut in *C. briggsae* appears to be very similar to that of *C. elegans* (Figure 4). The *C. briggsae* genome contains orthologues of the *med* and *end* genes [10,12,13]. These have similar embryonic expression in *C. briggsae* as their *C. elegans* counterparts, and the *C. briggsae med* genes can rescue *med-1,2* double mutants of *C. elegans* to complete viability [10]. Overexpression of *Cbr-end-3.1* in *C. elegans* can promote widespread gut specification [13]. In *C. briggsae*, RNAi targeting the *C. briggsae end-1* and *end-3* orthologues (the latter of which exists as two nearly identical paralogues, *Cbr-end-3.1* and *Cbr-end-3.2*) results in a penetrant absence of gut, similar to the loss of *C. elegans end-1* and *end-3* [13]. Binding sites for SKN-1 are found upstream of the *C. briggsae med* orthologues, and similarly, MED binding sites are found in the putative regulatory regions of the *C. briggsae end-1* and *end-3* orthologues [15]. Together, these results suggest that gut specification is highly conserved between *C. elegans* and *C. briggsae*. The *med* factors from *C. remanei*, a closely related male-female species, are also similarly expressed in *C. remanei* and able to rescue *med-1,2* mutants of *C. elegans* to viability, suggesting extended conservation throughout the Elegans group [10].

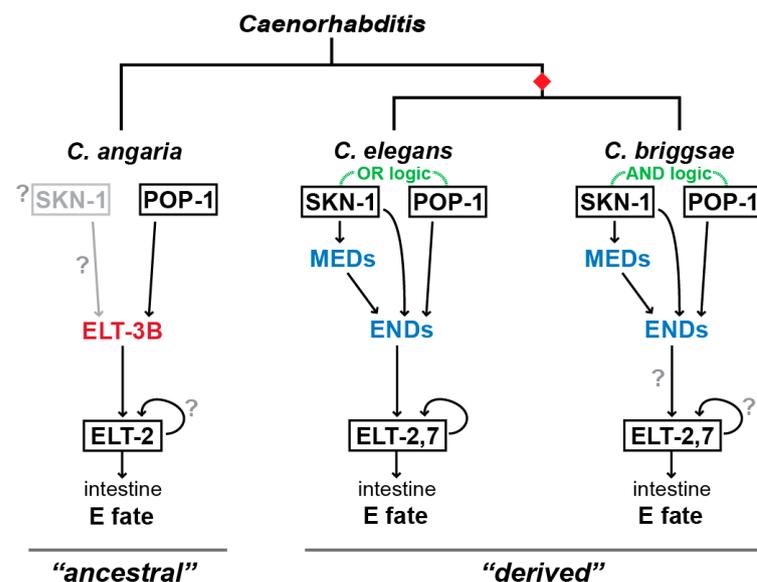


Figure 4. Simplified gut specification network among *C. elegans*, *C. briggsae*, and *C. angaria*. The simpler network of the latter species likely represents the ancestral form of the network in the genus, with the other two representing a derived state that originated at the base of the Elegans supergroup (red diamond).

The maternal factors SKN-1 and POP-1 also have clear orthologues in *C. briggsae* [12]. When RNAi experiments were tried with these two factors, surprising differences were obtained. In *C. elegans*, loss of *skn-1* by RNAi or mutation resulted in arrested embryos, but some 20% of these still contain gut, primarily due to input from POP-1 [12,27,30,38]. In the case of *Cbr-skn-1(RNAi)*, arrested embryos were obtained, but these were largely absent of any gut [12]. In *C. elegans*, loss of *pop-1* by RNAi or mutation results in a transformation of MS, the sister cell of E, into an E-like cell, resulting in excess gut [28,38]. However, in *C. briggsae*, the opposite result was obtained. Knockdown of *Cbr-pop-1* by RNAi resulted in the complete absence of the gut and a transformation of the E cell into an MS-like precursor [12,85]. The conclusion was that whereas in *C. elegans*, SKN-1 and POP-1 follow an “OR” type of regulatory logic, where either is sufficient for specification of at least some gut, in *C. briggsae*, these factors follow an ‘AND’ logic, where loss of either factor eliminates gut. This type of change in regulatory logic is an example of DSD [6]. Smaller-scale DSD in gut specification has been found among wild isolates of *C. elegans*, suggesting that the gut specification network has evolved to be robust to these [73].

A change in the promoters of the *C. briggsae end* genes may be responsible for the difference in phenotype of loss of *pop-1* between *C. briggsae* and *C. elegans*. In *C. elegans*, the promoter of *end-1* contains two MED sites, while that of *end-3* contains four MED sites [15,68]. The *C. briggsae* orthologous genes carry only a single apparent MED site in each [15]. When the number of MED sites is reduced in single-copy chromosomal versions of the *C. elegans end* genes, even though gut is specified robustly, depletion of *pop-1* results in the absence of gut [86]. Hence, it is possible to “turn *C. elegans* into *C. briggsae*” with respect to the *pop-1(RNAi)* phenotype by reducing MED regulatory input into *end* expression. Further supporting the importance of MED regulatory input in causing the excess gut in *C. elegans pop-1* mutants, triple mutant *med-1,2; pop-1* embryos also lack gut, showing that in the absence of POP-1, gut specification becomes dependent upon MED input [55]. Taken together, these results strongly suggest that evolutionary changes in the MED binding sites in the *end-1* and *end-3* genes contributed to DSD in gut specification between *C. elegans* and *C. briggsae*.

6. Gut Specification Outside of the Elegans Supergroup by the ELT-3 GATA Factor

When high-quality genome sequences became available for additional species in *Caenorhabditis*, it became possible to search for orthologues of the endodermal GATA factors [75]. Within the Elegans supergroup, the MEDs and ENDS were highly conserved, sharing blocks of amino acid conservation both within and outside of their DNA binding domains, as well as in intron/exon structure and putative *cis*-regulatory sites in their upstream flanking regions [15]. A surprising finding was that despite the essentiality of these genes in *C. elegans* and their apparent ubiquity in closely related species, the MED/END factors are absent from the genomes of *Caenorhabditis* species outside of the Elegans supergroup and from more distant nematodes [11,15]. This finding showed that the specification of the gut in these distant species must involve different factors other than MEDs or ENDS. The nature of these was unknown, although it was suggested that perhaps ELT-2 could act as the ancestral endoderm specifier, with the additional factors in *C. elegans* having arisen by a process of repeated duplication and divergence, explaining the structural similarity among the endodermal GATA factors [15,87].

Our recent work in the species *C. angaria*, which lies outside of the Elegans supergroup, revealed that the upstream portion of gut specification in this species consists not of a multifactor GATA cascade but merely of the single GATA factor ELT-3 [9]. In this study, we showed that a *C. angaria elt-2* translational reporter gene is expressed similarly to *C. elegans elt-2* when introduced into *C. elegans* as a transgene. The transgene can functionally complement the larval lethality of an *elt-2; elt-7* double null mutant to viability, showing that it can act in the maintenance of gut differentiation. Consistent with the same role for *elt-2* in *C. angaria*, RNAi targeting *Can-elt-2* resulted in penetrant larval lethality in which the intestine was only partially differentiated, a phenotype similar to loss of *C. elegans elt-2*

or loss of both *C. elegans elt-2* and *elt-7* [49,50]. Unexpectedly, the expression of *C. angaria elt-2* in *C. elegans* was completely dependent on END-3 and END-1, despite the absence of genes encoding these factors in the *C. angaria* genome. This suggested that within *C. angaria*, activation of *Can-elt-2* depends on a different GATA factor.

Identification of this GATA factor was straightforward, as only ELT-1, ELT-3, and ELT-5, three primarily hypodermal GATA factors in *C. elegans*, have orthologues in *C. angaria* in addition to the endodermal ELT-2 factor [9,48,88,89]. By examining expression using single-molecule inexpensive FISH (smiFISH), we found that all three show similar embryonic expression to the corresponding genes in *C. elegans*, while *Can-elt-3* showed an additional domain of expression in the early E lineage, from the time of its birth until the 4E stage [9,90]. This pattern overlaps that of *end-3* and *end-1* together, suggesting that ELT-3 fulfills the roles of these factors in *C. angaria* [51,52,74]. Consistent with this, RNAi or a null mutation in *Can-elt-3* resulted in penetrant loss of gut with a phenotype resembling *C. elegans end-1,3* double mutants [9,16].

We tested for the putative roles of SKN-1 and POP-1 in *C. angaria* after identifying their orthologues. Depletion of *Can-pop-1* resulted in a similar phenotype to *C. briggsae pop-1(RNAi)*, namely embryonic arrest and the absence of gut [9,12]. Consistent with a role for Can-POP-1 in the activation of *Can-elt-3*, expression of *Can-elt-3* was abolished in *Can-pop-1(RNAi)* embryos. In contrast, no phenotype for *Can-skn-1(RNAi)* was found, even though transcripts for *Can-skn-1* were successfully depleted by the RNAi treatment [9]. These results suggest that the gut specification network in *C. angaria* is a simpler network in which maternal POP-1 acts upstream of Can-ELT-3, which then activates *Can-elt-2* (Figure 4). Consistent with ELT-3 being widespread as the gut specification factor outside of the *Elegans* supergroup, we found that orthologues of *elt-3* are expressed in the early E lineage in *C. portoensis* and the stem species *C. monodelphis*, the latter of which is considered a basal outgroup for the rest of the known species in the genus [9]. We thus proposed the most parsimonious hypothesis, which is that gut specification in *C. angaria* reflects an “ancestral” mode of gut specification in the genus, while species in the *Elegans* supergroup, such as *C. elegans* and *C. briggsae*, have a “derived” mode that originated at the base of that group.

7. Short and Long Isoforms of ELT-3 Have Different Functions

C. elegans ELT-3 was identified by its sequence similarity to other GATA factors [48]. Although the gene was found to be expressed primarily in hypodermal cells, suggesting it may encode a tissue identity factor, a null mutant of *elt-3* showed no apparent developmental phenotype [66]. Forced overexpression of *C. elegans elt-3* throughout early embryos can drive hypodermal specification, but not endoderm specification [50,66]. ELT-3 and the nuclear hormone receptor NHR-25 have been placed in a gene network that drives hypodermal specification downstream of the GATA factor ELT-1 [89,91]. Further studies have shown that *elt-3* is required for gene expression changes, presumably in the hypodermis, that result from exposure of *C. elegans* to oxidative and other environmental stresses [91–94]. Given that the DNA-binding domains of *C. elegans* and *C. angaria* ELT-3 are identical, the finding that *C. angaria* ELT-3 has a role in the specification of endoderm was therefore unexpected because we also detected *Can-elt-3* in the embryonic hypodermis [9].

The apparent paradox was resolved by the finding that *elt-3* encodes two isoforms in both species: A long isoform, ELT-3B, and a short isoform, ELT-3A [9,67,95]. The two isoforms are encoded by differential transcriptional initiation, such that transcripts for both carry the same 3' end, but the longer isoform has an additional two exons at its 5' end (Figure 5). Both mRNAs are *trans*-spliced to SL1 [9,95,96]. At the time that *elt-3* was first described, only the shorter ELT-3A isoform was known [48]. In our studies in *C. angaria*, we distinguished expression of the two isoforms using an smiFISH probe that was specific for the longer 5' sequences present in the mRNA that encodes Can-ELT-3B and found that the long isoform is expressed only in the early E lineage [9]. To test the hypothesis that Can-ELT-3B is endoderm-specific, we forced the overexpression of Can-ELT-3B or

Can-ELT-3A throughout *C. elegans* early embryos. Only Can-ELT-3B promoted endoderm, as expected. More recently, we tested whether *C. elegans* ELT-3B could promote endoderm specification when ectopically expressed, and it too has this ability [67].

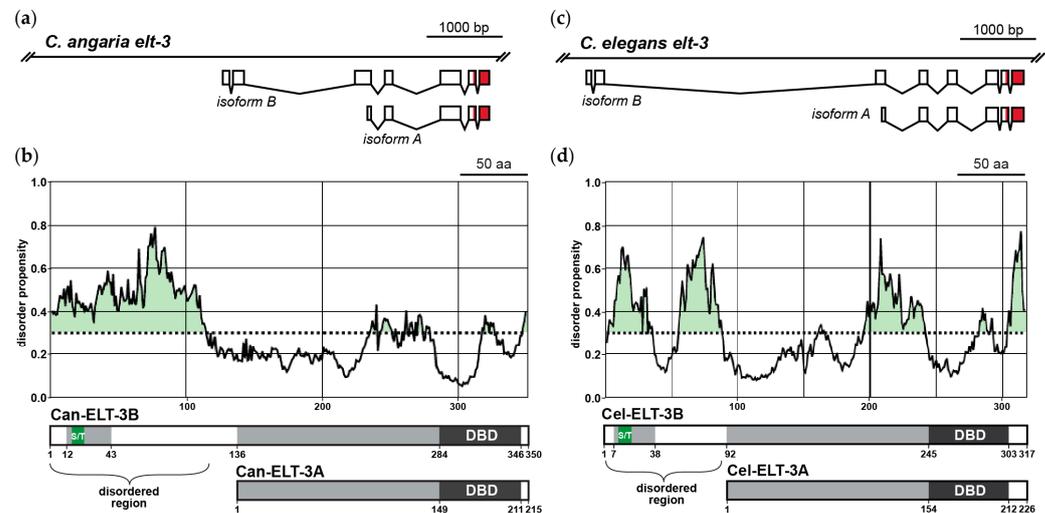


Figure 5. *C. angaria* ELT-3B differs from ELT-3A by having an extended amino region that is intrinsically disordered. (a) Diagram showing arrangement of exons in the *C. angaria elt-3* gene and the transcripts corresponding to the B and A isoforms [67]. (b) Use of the fDpNn (putative function- and linker-based Disorder Prediction using a deep neural network) algorithm to predict regions of intrinsic disorder in Can-ELT-3B [97]. The results show that the region of highest disorder (with a score of 0.3 or greater) lies in the amino-terminal 118 amino acids, which are within the 135 amino acids specific to the long isoform of Can-ELT-3B [67]. The structures of Can-ELT-3B and Can-ELT-3A are shown beneath the chart, with numbers indicating amino acid positions from the start. The DNA-binding domains (DBD) are shaded dark. Regions of weaker similarity and a serine-threonine region (S/T) are shown after [67]. Similar diagrams are shown for *C. elegans elt-3* in (c,d).

We further tested whether Can-ELT-3B could rescue gut specification in *C. elegans* when expressed in the correct spatiotemporal context [9]. We constructed a transgene in which Can-ELT-3B::CFP expression is activated by the *end-3* promoter and introduced this transgene into a *C. elegans* strain lacking *end-1*, *end-3*, and *elt-7*. The embryonic lethality of this strain could be rescued by the *end-3promoter::Can-ELT-3B::CFP* transgene. We also confirmed that forced early E lineage expression of *C. elegans* ELT-3B was able to rescue the gut specification of an *end-1*, *end-3* double mutant [67]. Hence, even though there is no apparent role for *C. elegans elt-3* in endoderm specification, there remains a cryptic ability of the ELT-3B isoform to activate *elt-2*.

As a test of whether the minimal *C. angaria* gut specification network could replace that of *C. elegans*, we introduced both the *end-3promoter::Can-ELT-3B::CFP* and *Can-ELT-2::GFP* transgenes into a quadruple mutant *elt-7 end-1, end-3, and elt-2* background, and found that ~50% of animals could be rescued to complete viability [9]. This result shows that Can-ELT-3B and Can-ELT-2 together are indeed sufficient to drive gut specification and differentiation in *C. elegans*, implying that the main zygotic endoderm specification pathway in *C. angaria* consists of only *Can-elt-3B* and *Can-elt-2*.

8. How Do ELT-3B and ELT-3A Regulate Different Target Genes?

Our recent findings show that ELT-3B, but not ELT-3A, from either *C. angaria* or *C. elegans* can promote endoderm specification in *C. elegans*. The gene and protein structures are similar between the two species (Figure 5A–D). Despite sharing the same DNA-binding domain, these two isoforms must therefore activate different target genes. In humans and mice, the diversity of GATA factor function is often encoded in protein regions upstream of the DNA-binding domains, which mediate protein–protein interactions [98]. If this were

true for ELT-3B, we might expect to see a high degree of protein conservation between the ELT-3B-specific parts of Can-ELT-3B and Cel-ELT-3B. However, these regions share only a weakly conserved region that is serine- and threonine-rich, suggesting this portion of ELT-3B may be intrinsically disordered [67]. Using the fIDPnn computational tool for disorder prediction, we find that the amino-terminal 118 amino acids of the 135 that are specific to Can-ELT-3B and the amino-terminal 88 amino acids of the 91 that are specific for Cel-ELT-3B are intrinsically disordered (Figures 4D and 5B) [97]. While the amino terminus could be responsible for protein differences due to DNA-binding activity, post-translational modification, stability, or localization, it could be that it is the intrinsic disorder itself that contributes to the ability of ELT-3B to activate endoderm target genes. Intrinsic disorder may allow some transcription factors to become more concentrated at regulatory regions in liquid–liquid phase separations (LLPSs), which could increase the ability of a factor to activate gene expression [99,100]. An intriguing hypothesis, therefore, is that the amino terminus of ELT-3B within *Caenorhabditis* enables an intrinsic change in its state that allows it to interact with multiple dispersed target sites more efficiently (Figure 6). This would mean that ELT-3B-responsive target genes are controlled by “super-enhancers” that permit more rapid initiation of transcription [100]. This might be expected for a factor that has to act within a very short time in the early embryo to drive specification.

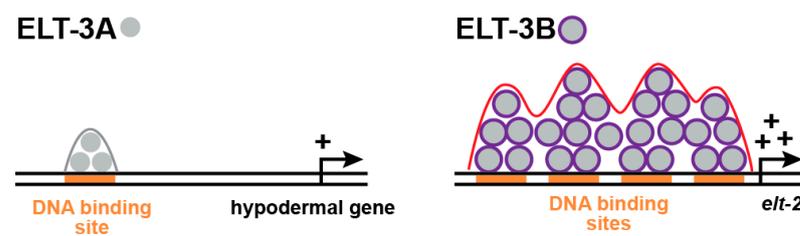


Figure 6. Speculative difference in protein states between the smaller ELT-3A factor and the larger ELT-3B factor. Here, amino-terminal regions of intrinsic disorder favor cooperative interactions with clusters of binding sites in an endodermal target gene like *elt-2*, permitting a more rapid activation of transcription. Modeled after a similar diagram in a prior work [100].

9. How Did the Gut Specification Network Change in *Caenorhabditis*?

Work in other systems reveals many mechanisms by which gene regulatory networks change by rewiring, although these are usually associated with changes in phenotype (e.g., body plan) or ecology (e.g., life history) [2,101]. In the case of the differences between *C. elegans* and *C. angaria*, there do not appear to be gross differences in either phenotype or ecology, which is typical for Developmental System Drift [6]. Our understanding of gut specification outside of the *Elegans* supergroup is in its early stages compared to our understanding of the network in *C. elegans*. Results from other species studied so far, *C. briggsae* and *C. remanei* within the *Elegans* supergroup and *C. angaria*, *C. portoensis*, and *C. monodelphis* outside of it, nonetheless suggest changes that occurred to produce the more complex gut specification network from a simpler network that is likely to have been the ancestral state. Here, we propose a speculative sequence of events that may have led to the derived network in the *Elegans* supergroup species by predicting one of many possible patterns of gene duplication and divergence and changes in *cis*-regulation, mechanisms that drive DSD [102]. Our inability to be more definitive results mostly from the absence of known “intermediate” species, for example, lacking only one of the MEDs, END-1 or END-3. Either such species have not yet been identified (or had their genomes sequenced), lineages preserving these states are now extinct, or else the changes that must have occurred took place over a relatively short evolutionary timescale. We propose that the network may have evolved mostly from the bottom upwards in a series of steps, as shown in Figure 7. Such a retrograde pattern of evolution has been proposed as a mechanism by which such gene networks arise in general [103].

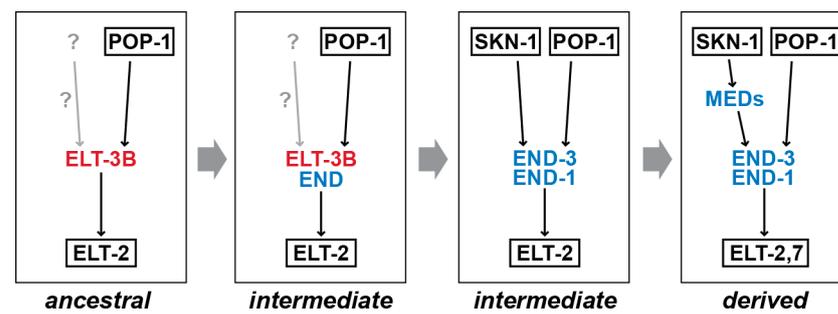


Figure 7. Possible stepwise evolution for the rewiring of endoderm specification within the *Caenorhabditis* genus. Several steps had to occur between the presumed simpler ancestral network and the derived network, and this diagram presents one of many possible patterns of gene duplication and changes in *cis*-regulation. We previously proposed a pattern of network expansion by retrograde evolution, starting from the direct regulation of *elt-2* by a SKN-1 factor [15].

Of the maternal factors SKN-1 and the Wnt effector POP-1/TCF, we know only that POP-1 has a conserved role in contributing to the activation of gut specification across the genus. In *C. briggsae* and *C. angaria*, knockdown of *pop-1* results in the absence of gut, while in *C. elegans*, simultaneous knockdown of *skn-1* is required to see this contribution, but it is nonetheless there [9,30,38,85]. Since POP-1 acts downstream of the Wnt/ β -catenin asymmetry pathway interaction between P₂ and EMS, it is reasonable to hypothesize that this signal and the regulation of POP-1 are conserved at least throughout the genus [32,36,104].

A role for SKN-1 in gut specification is detectable in both *C. elegans* and *C. briggsae*; however, we could not find a role for the only SKN-1 orthologue in *C. angaria* [9,27,38]. As such, the contribution of SKN-1 to gut specification outside of the *Elegans* supergroup is not yet known. As SKN-1 directly activates the *med* genes in EMS, some mechanism must act to delay activation of E specification factors like *end-3* until EMS has divided. We previously proposed that this may occur through lower affinity sites in the *end* genes [15]. SKN-1/Nrf has an evolutionarily conserved role in stress response across animals, suggesting that it could have been co-opted into specification late in evolution in *Caenorhabditis* [31]. In an interesting convergence of function, ELT-3 and SKN-1 have a shared role in such responses post-embryonically [92]. If so, regulatory interactions between SKN-1 and *elt-3* could have been co-opted into gut specification.

As mentioned earlier, other maternal regulators such as PAL-1/Caudal and SPTF-3/Sp1 have been found in *C. elegans* that play at least some role in gut specification, as revealed by the ability of knockdowns in these to increase the severity of a gutless phenotype in partially compromised specification backgrounds [38,39].

ELT-2 is likely to be conserved as the terminal gut factor, as it is present in most nematode species related to *Caenorhabditis* [11]. Within the genus, the ability of the *C. angaria* gene to replace the *C. elegans* gene confirms that *elt-2* maintains functional *cis*-regulation of its promoter as well as the ability of its product, ELT-2, to activate downstream target genes [9]. Outside of the genus, an *elt-2* orthologue from the vertebrate parasite *H. contortus* can direct gut specification when forcibly overexpressed in *C. elegans*, consistent with widespread conservation [105].

The origin of the early, transiently expressed “specifier” gene(s) has been open for speculation for years. Prior work had suggested that the END factors arose as duplicates of *elt-2* [15,87]. This idea was proposed because, until recently, no other GATA factors were known that could specify gut outside of the *Elegans* supergroup. We now know that ELT-3 is likely to be an ancestral gut specifier, at least in *Caenorhabditis*. Since ELT-3 is structurally similar to END-1 and END-3, which specify gut in both *C. elegans* and *C. briggsae*, it is logical to propose that a prototype *end* gene arose as a duplication of *elt-3*. Indeed, we noted in a prior work that the *end* genes are more similar to *elt-3* than they are to *elt-2* [13]. Because *end-3* and *end-1* orthologues are found within tens of kilobase pairs of each other across many species in the *Elegans* supergroup, it is likely that the prototype *end* gene underwent

duplication to make *end-1* and *end-3* [15]. During this process, the two genes together must have been able to replace the gut-specific function of *elt-3*, resulting in the loss of *elt-3* expression in the early E lineage. The *end* genes maintained their responsiveness to POP-1, as this was already present in *elt-3*. Maternal SKN-1 was co-opted into the specification of MS and E at some point, reinforcing input from POP-1 upstream of the *ends*. Regulatory sites for SKN-1 are still found in the extant *end* genes but could not be found in *Can-elt-3*, suggesting that this regulation arose after *end-1,3* and was not present ancestrally [9,11,15].

The MED GATA factors are found only in the Elegans supergroup and act immediately downstream of SKN-1 in *C. elegans* [15,47]. While MED-1 and MED-2 are not essential for E specification in *C. elegans*, they are essential for specification of tissues derived from MS, the sister of E, and the MEDs appear to activate early E or MS lineage expression of a number of genes [47,70]. Genomic regions containing individual *med* genes from *C. briggsae* or *C. remanei* can functionally replace the loss of *med-1,2* in *C. elegans*, showing they have maintained conservation of expression downstream of SKN-1 and function of the MED gene products [10]. As MED-like factors are not found outside of the Elegans supergroup, these factors likely arose at the base of the supergroup [15]. Exactly when they might have arisen relative to maternal input from SKN-1 is not clear. However, because direct SKN-1 regulation of the *end* genes is genetically detectable in *C. elegans*, we propose that the MEDs arose more recently and became intercalated between SKN-1 and *end-1,3* [15,38]. The same intercalation between SKN-1 and factors that specify MS could have occurred at this time as well [47,70].

Finally, the ELT-7 factor has structural similarity to ELT-3 and plays a reinforcing role in gut development in parallel with END-1 and ELT-2 [49,65,73]. Whether this factor arose later is not known; however, extant ELT-7 orthologues are structurally more like the ENDS in terms of protein size and hence likely to have arisen from an *end* or *elt-3* gene [15]. However, its loss to two species within the Elegans supergroup suggests that it may be dispensable at some level [15]. Other changes in the endoderm network within the Elegans supergroup suggest that the network is highly dynamic in terms of genome evolution. This includes changes in the copy number of the *med* and *end-3* genes, as well as changes in *cis*-regulatory sites [15]. The Elegans supergroup is, curiously, the only clade of the genus in which hermaphroditism is known to have evolved [76]. Indeed, genomic studies across the genus suggest that genome evolution is more rapid in hermaphroditic species [106].

10. Examples of DSD in *Caenorhabditis* and Other Animals

Several examples of DSD are known in *Caenorhabditis* and other animals are reviewed elsewhere [7,8]. Whereas the type of DSD we have uncovered in endoderm specification is largely due to changes in transcription factor gene networks, others have been found to involve changes at multiple molecular mechanisms. In one of the defining examples of DSD, the specification of the precursor cells that form the hermaphrodite vulva has been the subject of study in *Caenorhabditis* and related nematodes, including *Pristionchus* [107,108]. In both species, the same set of three hypodermal cells will form the vulva primordium. In *C. elegans*, LIN-3/EGF (Epidermal Growth Factor) signaling from a cell outside the equivalence group, the anchor cell (AC), specifies the primary fate in the closest cell, while the neighboring cells receive less EGF signal and lateral signaling by LIN-12/Notch, which specifies the secondary fate [109]. Among different *Caenorhabditis* species, cryptic variation is observed in how EGF, Notch, and Wnt pathways work cooperatively to ensure robust vulval specification [110].

In an example of gene network rewiring in the evolution of vulva specification, a 59-bp *cis*-regulatory module upstream of *lin-3* contains three *cis*-regulatory sites that are required for normal expression in the anchor cell [108]. In *C. angaria*, only one of these sites is present, but it is nonetheless sufficient to drive normal expression of *lin-3* when used as a transgene in *C. elegans*, hypothesized to be due to an unknown factor that can activate that module in both species [108]. We identified conserved putative *cis*-regulatory sites in the *med* and *end* genes across the Elegans supergroup, but comparative functional analyses have not yet

been performed [15]. The *elt-2* promoter sequences have been compared among a subset of these species to identify GATA site clusters in *C. elegans* by conservation, but not for cross-species comparative functional studies [58,59]. As multiple redundant factors appear to be at work in endoderm specification, it is likely that some species may show different phenotypes from *C. elegans* when individual factors are knocked out. Similarly, induction of gut by the P₂-to-EMS cell signal involves Wnt, MAP kinase, and Src tyrosine kinase pathways working in parallel [35,111,112]. Hence, there are opportunities for future work on cryptic variation in endoderm specification in both cell signaling and regulation of gene expression across the genus.

In the more distantly related species, *Pristionchus pacificus*, vulva specification exhibits significant changes in how fates are assigned, despite conservation of the underlying primary and secondary vulval fates with *Caenorhabditis* [107]. In this species, both the anchor cell and the larger somatic gonad provide an inductive signal for the primary fate, while two other cells, the mesoblast M and a more posterior hypodermal cell, induce the secondary fate [107]. This predicts that in more distant species, even larger-scale differences in gut specification may be found compared with those occurring within *Caenorhabditis*. Intriguingly, an expansion of *elt-3*-like genes is found in *Pristionchus*, perhaps hinting at an independent radiation of endodermal GATA factors [11]. This possibility is currently being investigated.

One well-known example of change in a developmental gene network outside of nematodes involves that which determines segment identity in *Drosophila* and related species. Compared with *Drosophila*, the underlying gene network has undergone rapid evolution, particularly in its upstream inputs [113]. In cyclorrhaphan flies, a group that includes *Drosophila*, the maternal anterior specification factor Bicoid arose from a duplication of Hox3 at the base of the group [114,115]. Outside of Cyclorrhapha, different factors appear to play a similar role as *Bicoid* [116–119]. The replacement of other upstream regulators by *Bicoid* resembles the replacement of ELT-3 by END-1/3 in *C. elegans*. However, whereas early embryos of *Caenorhabditis* are almost indistinguishable from one another in their embryonic development, Cyclorrhapha embryonic development shows a wide variation, for example in the speed of embryogenesis and in the length of the germ-band [81,120–122].

Another example of DSD in a gene network is found in tunicate larvae, in the development of *Ciona* compared with that of the distant relative *Molgula* [123]. These two species share similar cell lineages in development, despite hundreds of millions of years since the common ancestor [124]. Within the developing motor ganglion of the two species, expression of several regulatory genes, including the homeobox genes *Dmbx* and *Vsx*, is conserved [123,125]. However, the upstream regulators have undergone a change in expression and regulatory logic. This is revealed in the ability of *Dmbx* cis-regulatory sequences from *Molgula* to drive proper spatiotemporal expression in *Ciona* but not the reverse, implying DSD [123]. The authors proposed that slight differences in developmental time, with *Molgula* development occurring 10% faster, added to the constraint that might be imposed by fixed cell lineages, may have imposed a selection for DSD-type changes to keep neural specification robust [124].

11. Additional Questions for Future Work

11.1. Gut Specification in Other Species in the Genus

The apparent rapid changes in genes involved in gut specification suggest that there are likely to be other examples of either network rewiring or cryptic changes in the relative contribution of different paralogous genes, especially among species in the *Elegans* supergroup, for which complex patterns of gene duplication were observed [15]. Although such changes are probably not likely to be as major as the rewiring that expanded *elt-3* and produced several paralogous GATA factors, there could be examples of the gain or loss of specific regulators. Indeed, among the 20 *Elegans* supergroup species examined, two appeared to lack *elt-7* orthologues [15]. Alternatively, there may be some *Elegans* supergroup species that have retained expression of *elt-3* in the early E lineage, for example, or have

further rewired their gut specification networks by additional duplication and divergence or changes in *cis*-regulation [15]. A further challenge will be to identify genes that have been co-opted into the pathway that are not GATA factors. Single-cell transcriptomics methods have advanced considerably in recent years, which might permit the identification of such factors [52,126]. The gold standard in determining whether factors expressed in the E lineage participate in specification is to delete them by mutation or test for a phenotype by RNAi, as we did for *elt-3* in *C. angaria* and *end-1,3* in *C. elegans* and *C. briggsae*, and show that gut specification is abolished [9,13,16]. Methods in species other than *C. elegans* and *C. briggsae* are less well developed, which may limit such functional studies [127].

11.2. Function of ELT-3B in *C. Elegans* and Other Species in the *Elegans* Supergroup

The ability of ELT-3B from *C. angaria* and *C. elegans* to activate *elt-2* creates an opportunity to study *cis*-regulation of *elt-2*, as prior work on *elt-2* regulation has considered only the known END-1,3 and ELT-2,7 factors [58,59]. Is the activation of *elt-2* by ELT-3B distinctly different? What microevolutionary changes have occurred in the DNA-binding domains, and how does the amino terminus affect DNA binding, if at all?

There is also the question of what ELT-3B could be doing in *C. elegans*. Its cryptic ability to promote *elt-2* expression and gut specification when forcibly expressed in early embryos, strongly suggests that ELT-3B has been maintained by selection for some other function outside of specification. An analysis of the expression of each isoform in *C. elegans* might suggest what the possible role of ELT-3B might be, as it is likely to be cell-autonomous. Prior work suggested functions for ELT-3 in the adult intestine, although others have ruled out such a function [128,129]. With new results showing that ELT-3B can activate *elt-2*, it remains possible that a function for ELT-3B is conditional, perhaps increasing expression of *elt-2* under stressful conditions [67]. Recent studies on the roles of ELT-3 in stress responses have considered only the effect of alleles that are predicted to affect the DNA-binding domain, which would affect both isoforms [92–94].

11.3. Evolution of ELT-3 Orthologues and Isoforms

To our knowledge, the dual functions of ELT-3B in endoderm and ELT-3A in ectoderm represent the first identification of a canonical *C. elegans* GATA factor with isoforms that have roles in distinct germ layers [9]. It is not known how widespread the long/short isoforms of *elt-3* genes may be, as this has not yet been examined [11]. The upstream exons in *elt-3B*, being smaller, less conserved, and upstream of a large intron, might make identification by gene prediction algorithms more difficult. A more detailed analysis of *elt-3* genes within and outside of *Caenorhabditis* would reveal whether the long isoform is a recent innovation or more widely conserved, and hence whether ELT-3B as an ancestral gut specification factor is more widespread. Curiously, when it came to the evolution of GATA factors in general, orthologues of *elt-3* appeared to have undergone increases in copy number in several species [11]. In our study on *Can-elt-3*, we found that only one of two *C. monodelphis* *elt-3* paralogues is expressed in the early E lineage [9]. Therefore, it may be that different functions for *elt-3* genes, in addition to gut specification, could evolve through gene duplication and divergence.

11.4. Amenability of Gut Specification to Rewiring

Much of the research in the *C. elegans* gut specification network has been directed at understanding how the network responds to perturbations, such as the removal of regulators by mutation, modification of *cis*-regulatory sites, and replacement of factors by others [9,24,59,67,74,86,87]. In general, these results have shown that the network is highly amenable to being rewired. To perform functional studies of putative networks from other nematode species, *C. elegans* can be used as a test system in which to synthetically replace the network with that of other species, as we did with *C. angaria* [9].

11.5. What Caused the Rewiring of Endoderm Specification?

Figure 7 above provides a speculative model for the stepwise conversion of the putative ancestral gut specification network into the extant network in *C. elegans*. However, the selective pressures that might have led to the rewiring of the network are still unresolved. Haag and True [102] considered two mechanisms that drive DSD that are pertinent to the rewiring of a transcriptional factor network. One of these mechanisms is the duplication of an ancestral gene to result in two subfunctionalized paralogues [130]. In endoderm specification, this mechanism accounts for the origin of the *end-1* and *end-3* genes from a prototype *end-1* gene, or *elt-3* itself. Once *end-1* and *end-3* had diversified in function, it would be difficult to lose either of them. Consistent with this, subtle differences in binding affinities of END-1 vs. END-3 on the *elt-2* promoter have been observed [58,59]. Such differences would be consistent with END-1 and END-3 paralogues having co-evolved with subsets of ancestral ELT-3 target genes. The overexpression of ELT-3B in our cross-species studies would likely overcome such differences, so subtle differences in target genes would not be apparent [9]. The driver for the radiation of factors may have been that increased numbers of regulators result in higher robustness to environmental fluctuations or more rapid development. In some insects, for example, the A/P patterning gene networks evolved more complexity through changes in upstream regulators, which may have allowed more robust patterning under pressure for more rapid development [117,121]. It is not clear whether more rapid development could be the driving force in endoderm development with *Caenorhabditis*, although some of the fastest-developing nematodes are found in the clade that includes this genus [131]. The speed of embryogenesis between *C. angaria* and *C. elegans* is about the same, although there are minor differences in particular milestones [132]. It is also not apparent that the simpler network in *C. angaria* is in any way less “robust” than the more complex network in *C. elegans*. Under laboratory conditions, *C. angaria* animals grow well, especially considering that they are a male-female species, which tend to suffer from inbreeding depression when cultured for long periods [127]. We also do not know if there are additional regulators in *C. angaria* that play roles analogous to the MED, END-1, and END-3 factors but that are not GATA factors. These factors could make the gene network more robust in *C. angaria*.

A second mechanism that drives gene networks DSD involves different types of selection acting on the direct targets of a pleiotropic transcription factor [102,133,134]. Here, the factor activates two distinctly different target genes. Directional selection acts on the *cis*-regulation of one of the targets to increase its expression, while the second target undergoes stabilizing selection. This drives changes in the second gene to maintain its expression levels, providing a selection that may ultimately result in the recruitment of a different activator [133]. DSD would thus be observed as a change in regulation of this gene from the ancestral activator. In the case of endoderm specification in *Caenorhabditis*, this would require that the ancestral *elt-3* gene function in two processes. Indeed, *C. elegans* ELT-3 functions in the post-embryonic hypodermis, where it is required to activate genes that respond dynamically to oxidative stress as well as to modulate expression of hypodermal genes involved in cuticle structure [92,93]. We observed expression of *C. angaria elt-3* in the embryonic hypodermis, similar to expression of *elt-3* *C. elegans* [9,48]. Hence, it is likely that ancestral ELT-3 activates both *elt-2* in the early endoderm and stress response genes in the hypodermis. Over time, directional selection acting on its role in the hypodermis would drive stabilizing selection on its role in activation of *elt-2*. This could result in the reassignment of gut specification to what became *end-1,3*. Once gut specification was uncoupled from stress response, the genes involved were free to undergo more rapid evolution. Hence, it is possible that both gene duplication/subfunctionalization and pleiotropy of ancestral ELT-3 led to the radiation of gut specification genes in *Caenorhabditis*, driven by directional selection for ELT-3's role in stress responses.

Another possible mechanism driving DSD in this system may be related to the development of the E lineage itself. Although we have been concerned primarily with activation of the gut-specific transcriptome via ELT-2, it may be that the expanded factors in *C. elegans* fulfilled a need to directly connect the timing of cell divisions with specification. At least in *C. elegans*, the early E lineage undergoes a slowing down of its cell cycle relative to its sister lineage MS to allow the two E daughter cells to move into the embryo during gastrulation [19,135]. With stratified transcription factors activated at slightly different times, more precise timing of the cell cycle could be achieved than might be possible with only a single factor. There is evidence that connects the gut specification network with the cell cycle in *C. elegans*. Early E-specific expression of a *wee1* kinase gene, *wee-1.1*, is directed by MED-1,2 [68]. Loss of *wee-1.1* by mutation causes changes in the E cell cycle without affecting differentiation, and there are other putative cell cycle regulators activated in the early E lineage [136]. Other perturbations in the E specification pathway, such as loss of *end-3* by mutation, lead to profound changes in the pattern of cell divisions arising from the E cell, causing the generation of supernumerary gut cells [24,137]. Gain-of-function mutations in *cdc-25.1*, another cell cycle regulator, can increase the number of E descendants [138,139]. These results provide evidence that the transcription factors that direct gut specification also control the cell cycle and that cell cycle factors are critical for proper patterning of the E lineage. Hence, the additional layers in the expanded network in *C. elegans* may have permitted fine-tuning of the cell cycle as compared with the simpler network in *C. angaria*.

12. Conclusions

Endoderm specification in *Caenorhabditis* is now established as a model for the study of DSD. The availability of genome sequences from more species and newer genetic tools will enable the testing of hypotheses for how DSD might have occurred in this system. For example, single-cell transcriptomics approaches offer a way of exhaustively identifying factors that could be important for cell specification [52,126,140]. Cell-by-cell descriptions of gene expression will be a useful resource, especially for evolutionary comparisons, and these could be complemented with perturbation studies and cross-species expression tests to identify changes in gut specification genes. With the knowledge of at least one key regulator of E specification outside of the *Elegans* supergroup, the stage is set in this system to perform directed functional studies aimed at understanding mechanisms by which the network might have evolved, which could be generalizable to the evolution of gene networks across many systems.

Author Contributions: Conceptualization, M.F.M. and G.B.-M.; methodology, M.F.M. and G.B.-M.; writing—original draft preparation, M.F.M.; writing—review and editing, M.F.M. and G.B.-M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Davidson, E.H.; Levine, M.S. Properties of developmental gene regulatory networks. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 20063–20066. [[CrossRef](#)]
2. Erwin, D.H.; Davidson, E.H. The evolution of hierarchical gene regulatory networks. *Nat. Rev. Genet.* **2009**, *10*, 141–148. [[CrossRef](#)] [[PubMed](#)]
3. Kittelmann, S.; Preger-Ben Noon, E.; McGregor, A.P.; Frankel, N. A complex gene regulatory architecture underlies the development and evolution of cuticle morphology in *Drosophila*. *Curr. Opin. Genet. Dev.* **2021**, *69*, 21–27. [[CrossRef](#)] [[PubMed](#)]

4. Chevin, L.M.; Leung, C.; Le Rouzic, A.; Uller, T. Using phenotypic plasticity to understand the structure and evolution of the genotype-phenotype map. *Genetica* **2022**, *150*, 209–221. [[CrossRef](#)] [[PubMed](#)]
5. Fusco, G.; Minelli, A. Phenotypic plasticity in development and evolution: Facts and concepts. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* **2010**, *365*, 547–556. [[CrossRef](#)] [[PubMed](#)]
6. True, J.R.; Haag, E.S. Developmental system drift and flexibility in evolutionary trajectories. *Evol. Dev.* **2001**, *3*, 109–119. [[CrossRef](#)]
7. Haag, E.S. The same but different: Worms reveal the pervasiveness of developmental system drift. *PLoS Genet.* **2014**, *10*, e1004150. [[CrossRef](#)]
8. Haag, E.S.; Fitch, D.H.A.; Delattre, M. From “the Worm” to “the Worms” and Back Again: The Evolutionary Developmental Biology of Nematodes. *Genetics* **2018**, *210*, 397–433. [[CrossRef](#)]
9. Broitman-Maduro, G.; Sun, S.; Kikuchi, T.; Maduro, M.F. The GATA factor ELT-3 specifies endoderm in *Caenorhabditis angaria* in an ancestral gene network. *Development* **2022**, *149*, dev200984. [[CrossRef](#)]
10. Coroian, C.; Broitman-Maduro, G.; Maduro, M.F. Med-type GATA factors and the evolution of mesendoderm specification in nematodes. *Dev. Biol.* **2005**, *289*, 444–455. [[CrossRef](#)]
11. Eurmsirilerd, E.; Maduro, M.F. Evolution of Developmental GATA Factors in Nematodes. *J. Dev. Biol.* **2020**, *8*, 27. [[CrossRef](#)] [[PubMed](#)]
12. Lin, K.T.; Broitman-Maduro, G.; Hung, W.W.; Cervantes, S.; Maduro, M.F. Knockdown of SKN-1 and the Wnt effector TCF/POP-1 reveals differences in endomesoderm specification in *C. briggsae* as compared with *C. elegans*. *Dev. Biol.* **2009**, *325*, 296–306. [[CrossRef](#)] [[PubMed](#)]
13. Maduro, M.; Hill, R.J.; Heid, P.J.; Newman-Smith, E.D.; Zhu, J.; Priess, J.; Rothman, J. Genetic redundancy in endoderm specification within the genus *Caenorhabditis*. *Dev. Biol.* **2005**, *284*, 509–522. [[CrossRef](#)]
14. Maduro, M.F. Structure and evolution of the *C. elegans* embryonic endomesoderm network. *Biochim. Et Biophys. Acta* **2009**, *1789*, 250–260. [[CrossRef](#)]
15. Maduro, M.F. Evolutionary Dynamics of the SKN-1→MED→END-1,3 Regulatory Gene Cascade in *Caenorhabditis* Endoderm Specification. *G3* **2020**, *10*, 333–356. [[CrossRef](#)] [[PubMed](#)]
16. Owraghi, M.; Broitman-Maduro, G.; Luu, T.; Roberson, H.; Maduro, M.F. Roles of the Wnt effector POP-1/TCF in the *C. elegans* endomesoderm specification gene network. *Dev. Biol.* **2010**, *340*, 209–221. [[CrossRef](#)]
17. Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **1974**, *77*, 71–94. [[CrossRef](#)] [[PubMed](#)]
18. Corsi, A.K.; Wightman, B.; Chalfie, M. A Transparent Window into Biology: A Primer on *Caenorhabditis elegans*. *Genetics* **2015**, *200*, 387–407. [[CrossRef](#)]
19. Sulston, J.E.; Schierenberg, E.; White, J.G.; Thomson, J.N. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **1983**, *100*, 64–119. [[CrossRef](#)]
20. Bowerman, B. Maternal control of pattern formation in early *Caenorhabditis elegans* embryos. *Curr. Top. Dev. Biol.* **1998**, *39*, 73–117.
21. Maduro, M.F. Cell fate specification in the *C. elegans* embryo. *Dev. Dyn.* **2010**, *239*, 1315–1329. [[CrossRef](#)]
22. Goldstein, B. Induction of gut in *Caenorhabditis elegans* embryos. *Nature* **1992**, *357*, 255–257. [[CrossRef](#)]
23. Asan, A.; Raiders, S.A.; Priess, J.R. Morphogenesis of the *C. elegans* Intestine Involves Axon Guidance Genes. *PLoS Genet.* **2016**, *12*, e1005950. [[CrossRef](#)] [[PubMed](#)]
24. Choi, H.; Broitman-Maduro, G.; Maduro, M.F. Partially compromised specification causes stochastic effects on gut development in *C. elegans*. *Dev. Biol.* **2017**, *427*, 49–60. [[CrossRef](#)] [[PubMed](#)]
25. Lee, Y.U.; Son, M.; Kim, J.; Shim, Y.H.; Kawasaki, I. CDC-25.2, a *C. elegans* ortholog of cdc25, is essential for the progression of intestinal divisions. *Cell Cycle* **2016**, *15*, 654–666. [[CrossRef](#)] [[PubMed](#)]
26. Lawrence, P.A.; Levine, M. Mosaic and regulative development: Two faces of one coin. *Curr. Biol. CB* **2006**, *16*, R236–R239. [[CrossRef](#)]
27. Bowerman, B.; Eaton, B.A.; Priess, J.R. *skn-1*, a maternally expressed gene required to specify the fate of ventral blastomeres in the early *C. elegans* embryo. *Cell* **1992**, *68*, 1061–1075. [[CrossRef](#)] [[PubMed](#)]
28. Lin, R.; Thompson, S.; Priess, J.R. *pop-1* encodes an HMG box protein required for the specification of a mesoderm precursor in early *C. elegans* embryos. *Cell* **1995**, *83*, 599–609. [[CrossRef](#)] [[PubMed](#)]
29. Bowerman, B.; Draper, B.W.; Mello, C.C.; Priess, J.R. The maternal gene *skn-1* encodes a protein that is distributed unequally in early *C. elegans* embryos. *Cell* **1993**, *74*, 443–452. [[CrossRef](#)]
30. Shetty, P.; Lo, M.C.; Robertson, S.M.; Lin, R. *C. elegans* TCF protein, POP-1, converts from repressor to activator as a result of Wnt-induced lowering of nuclear levels. *Dev. Biol.* **2005**, *285*, 584–592. [[CrossRef](#)]
31. Blackwell, T.K.; Steinbaugh, M.J.; Hourihan, J.M.; Ewald, C.Y.; Isik, M. SKN-1/Nrf, stress responses, and aging in *Caenorhabditis elegans*. *Free. Radic. Biol. Med.* **2015**, *88*, 290–301. [[CrossRef](#)] [[PubMed](#)]
32. Rocheleau, C.E.; Downs, W.D.; Lin, R.; Wittmann, C.; Bei, Y.; Cha, Y.H.; Ali, M.; Priess, J.R.; Mello, C.C. Wnt signaling and an APC-related gene specify endoderm in early *C. elegans* embryos. *Cell* **1997**, *90*, 707–716. [[CrossRef](#)] [[PubMed](#)]
33. Rocheleau, C.E.; Yasuda, J.; Shin, T.H.; Lin, R.; Sawa, H.; Okano, H.; Priess, J.R.; Davis, R.J.; Mello, C.C. WRM-1 activates the LIT-1 protein kinase to transduce anterior/posterior polarity signals in *C. elegans*. *Cell* **1999**, *97*, 717–726. [[CrossRef](#)]

34. Schlesinger, A.; Shelton, C.A.; Maloof, J.N.; Meneghini, M.; Bowerman, B. Wnt pathway components orient a mitotic spindle in the early *Caenorhabditis elegans* embryo without requiring gene transcription in the responding cell. *Genes Dev.* **1999**, *13*, 2028–2038. [[CrossRef](#)] [[PubMed](#)]
35. Shin, T.H.; Yasuda, J.; Rocheleau, C.E.; Lin, R.; Soto, M.; Bei, Y.; Davis, R.J.; Mello, C.C. MOM-4, a MAP kinase kinase kinase-related protein, activates WRM-1/LIT-1 kinase to transduce anterior/posterior polarity signals in *C. elegans*. *Mol. Cell.* **1999**, *4*, 275–280. [[CrossRef](#)]
36. Thorpe, C.J.; Schlesinger, A.; Carter, J.C.; Bowerman, B. Wnt signaling polarizes an early *C. elegans* blastomere to distinguish endoderm from mesoderm. *Cell* **1997**, *90*, 695–705. [[CrossRef](#)]
37. Bhambhani, C.; Ravindranath, A.J.; Mentink, R.A.; Chang, M.V.; Betist, M.C.; Yang, Y.X.; Koushika, S.P.; Korswagen, H.C.; Cadigan, K.M. Distinct DNA binding sites contribute to the TCF transcriptional switch in *C. elegans* and *Drosophila*. *PLoS Genet.* **2014**, *10*, e1004133. [[CrossRef](#)]
38. Maduro, M.F.; Kasmir, J.J.; Zhu, J.; Rothman, J.H. The Wnt effector POP-1 and the PAL-1/Caudal homeoprotein collaborate with SKN-1 to activate *C. elegans* endoderm development. *Dev. Biol.* **2005**, *285*, 510–523. [[CrossRef](#)]
39. Sullivan-Brown, J.L.; Tandon, P.; Bird, K.E.; Dickinson, D.J.; Tintori, S.C.; Heppert, J.K.; Meserve, J.H.; Trogden, K.P.; Orłowski, S.K.; Conlon, F.L.; et al. Identifying Regulators of Morphogenesis Common to Vertebrate Neural Tube Closure and *Caenorhabditis elegans* Gastrulation. *Genetics* **2016**, *202*, 123–139. [[CrossRef](#)]
40. Calvo, D.; Victor, M.; Gay, F.; Sui, G.; Luke, M.P.; Dufourcq, P.; Wen, G.; Maduro, M.; Rothman, J.; Shi, Y. A POP-1 repressor complex restricts inappropriate cell type-specific gene transcription during *Caenorhabditis elegans* embryogenesis. *Embo. J.* **2001**, *20*, 7197–7208. [[CrossRef](#)]
41. Witze, E.S.; Field, E.D.; Hunt, D.F.; Rothman, J.H. *C. elegans* pur alpha, an activator of end-1, synergizes with the Wnt pathway to specify endoderm. *Dev. Biol.* **2009**, *327*, 12–23. [[CrossRef](#)]
42. Wiesenfahrt, T.; Duanmu, J.; Snider, F.; Moerman, D.; Au, V.; Li-Leger, E.; Flibotte, S.; Parker, D.M.; Marshall, C.J.; Nishimura, E.O.; et al. A Strategy To Isolate Modifiers of *Caenorhabditis elegans* Lethal Mutations: Investigating the Endoderm Specifying Ability of the Intestinal Differentiation GATA Factor ELT-2. *G3* **2018**, *8*, 1425–1437. [[CrossRef](#)]
43. Ewe, C.K.; Torres Cleuren, Y.N.; Flowers, S.E.; Alok, G.; Snell, R.G.; Rothman, J.H. Natural cryptic variation in epigenetic modulation of an embryonic gene regulatory network. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 13637–13646. [[CrossRef](#)] [[PubMed](#)]
44. Mello, C.C.; Draper, B.W.; Krause, M.; Weintraub, H.; Priess, J.R. The pie-1 and mex-1 genes and maternal control of blastomere identity in early *C. elegans* embryos. *Cell* **1992**, *70*, 163–176. [[CrossRef](#)]
45. Mello, C.C.; Schubert, C.; Draper, B.; Zhang, W.; Lobel, R.; Priess, J.R. The PIE-1 protein and germline specification in *C. elegans* embryos. *Nature* **1996**, *382*, 710–712. [[CrossRef](#)] [[PubMed](#)]
46. Schnabel, R.; Weigner, C.; Hutter, H.; Feichtinger, R.; Schnabel, H. mex-1 and the general partitioning of cell fate in the early *C. elegans* embryo. *Mech. Dev.* **1996**, *54*, 133–147. [[CrossRef](#)]
47. Maduro, M.F.; Meneghini, M.D.; Bowerman, B.; Broitman-Maduro, G.; Rothman, J.H. Restriction of mesendoderm to a single blastomere by the combined action of SKN-1 and a GSK-3beta homolog is mediated by MED-1 and -2 in *C. elegans*. *Mol. Cell.* **2001**, *7*, 475–485. [[CrossRef](#)] [[PubMed](#)]
48. Gilleard, J.S.; Shafi, Y.; Barry, J.D.; McGhee, J.D. ELT-3: A *Caenorhabditis elegans* GATA factor expressed in the embryonic epidermis during morphogenesis. *Dev. Biol.* **1999**, *208*, 265–280. [[CrossRef](#)]
49. Sommermann, E.M.; Strohmaier, K.R.; Maduro, M.F.; Rothman, J.H. Endoderm development in *Caenorhabditis elegans*: The synergistic action of ELT-2 and -7 mediates the specification → differentiation transition. *Dev. Biol.* **2010**, *347*, 154–166. [[CrossRef](#)]
50. Fukushige, T.; Hawkins, M.G.; McGhee, J.D. The GATA-factor elt-2 is essential for formation of the *Caenorhabditis elegans* intestine. *Dev. Biol.* **1998**, *198*, 286–302. [[CrossRef](#)]
51. Baugh, L.R.; Hill, A.A.; Slonim, D.K.; Brown, E.L.; Hunter, C.P. Composition and dynamics of the *Caenorhabditis elegans* early embryonic transcriptome. *Development* **2003**, *130*, 889–900. [[CrossRef](#)]
52. Tintori, S.C.; Osborne Nishimura, E.; Golden, P.; Lieb, J.D.; Goldstein, B. A Transcriptional Lineage of the Early *C. elegans* Embryo. *Dev. Cell* **2016**, *38*, 430–444. [[CrossRef](#)] [[PubMed](#)]
53. Hashimshony, T.; Feder, M.; Levin, M.; Hall, B.K.; Yanai, I. Spatiotemporal transcriptomics reveals the evolutionary history of the endoderm germ layer. *Nature* **2015**, *519*, 219–222. [[CrossRef](#)] [[PubMed](#)]
54. Zhu, J.; Hill, R.J.; Heid, P.J.; Fukuyama, M.; Sugimoto, A.; Priess, J.R.; Rothman, J.H. end-1 encodes an apparent GATA factor that specifies the endoderm precursor in *Caenorhabditis elegans* embryos. *Genes Dev.* **1997**, *11*, 2883–2896. [[CrossRef](#)] [[PubMed](#)]
55. Maduro, M.F.; Broitman-Maduro, G.; Mengarelli, I.; Rothman, J.H. Maternal deployment of the embryonic SKN-1 → MED-1,2 cell specification pathway in *C. elegans*. *Dev. Biol.* **2007**, *301*, 590–601. [[CrossRef](#)]
56. Hawkins, M.G.; McGhee, J.D. elt-2, a second GATA factor from the nematode *Caenorhabditis elegans*. *J. Biol. Chem.* **1995**, *270*, 14666–14671. [[CrossRef](#)]
57. Fukushige, T.; Hendzel, M.J.; Bazett-Jones, D.P.; McGhee, J.D. Direct visualization of the elt-2 gut-specific GATA factor binding to a target promoter inside the living *Caenorhabditis elegans* embryo. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11883–11888. [[CrossRef](#)]
58. Du, L.; Tracy, S.; Rifkin, S.A. Mutagenesis of GATA motifs controlling the endoderm regulator elt-2 reveals distinct dominant and secondary cis-regulatory elements. *Dev. Biol.* **2016**, *412*, 160–170. [[CrossRef](#)]
59. Wiesenfahrt, T.; Berg, J.Y.; Nishimura, E.O.; Robinson, A.G.; Goszczynski, B.; Lieb, J.D.; McGhee, J.D. The Function and Regulation of the GATA Factor ELT-2 in the *C. elegans* Endoderm. *Development* **2015**, *143*, 483–491. [[CrossRef](#)]

60. Fukushige, T.; Goszczynski, B.; Tian, H.; McGhee, J.D. The Evolutionary Duplication and Probable Demise of an Endodermal GATA Factor in *Caenorhabditis elegans*. *Genetics* **2003**, *165*, 575–588. [[CrossRef](#)]
61. Lowry, J.A.; Atchley, W.R. Molecular evolution of the GATA family of transcription factors: Conservation within the DNA-binding domain. *J. Mol. Evol.* **2000**, *50*, 103–115. [[CrossRef](#)]
62. Weirauch, M.T.; Yang, A.; Albu, M.; Cote, A.G.; Montenegro-Montero, A.; Drewe, P.; Najafabadi, H.S.; Lambert, S.A.; Mann, I.; Cook, K.; et al. Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* **2014**, *158*, 1431–1443. [[CrossRef](#)] [[PubMed](#)]
63. Zhu, J.; Fukushige, T.; McGhee, J.D.; Rothman, J.H. Reprogramming of early embryonic blastomeres into endodermal progenitors by a *Caenorhabditis elegans* GATA factor. *Genes Dev.* **1998**, *12*, 3809–3814. [[CrossRef](#)]
64. Riddle, M.R.; Weintraub, A.; Nguyen, K.C.; Hall, D.H.; Rothman, J.H. Transdifferentiation and remodeling of post-embryonic *C. elegans* cells by a single transcription factor. *Development* **2013**, *140*, 4844–4849. [[CrossRef](#)]
65. Dineen, A.; Osborne Nishimura, E.; Goszczynski, B.; Rothman, J.H.; McGhee, J.D. Quantitating transcription factor redundancy: The relative roles of the ELT-2 and ELT-7 GATA factors in the *C. elegans* endoderm. *Dev. Biol.* **2018**, *435*, 150–161. [[CrossRef](#)] [[PubMed](#)]
66. Gilleard, J.S.; McGhee, J.D. Activation of hypodermal differentiation in the *Caenorhabditis elegans* embryo by GATA transcription factors ELT-1 and ELT-3. *Mol. Cell. Biol.* **2001**, *21*, 2533–2544. [[CrossRef](#)]
67. Broitman-Maduro, G.; Maduro, M.F. The long isoform of the *C. elegans* ELT-3 GATA factor can specify endoderm when overexpressed. *Micropublication Biol.* **2023**, *2023*. [[CrossRef](#)]
68. Broitman-Maduro, G.; Maduro, M.F.; Rothman, J.H. The noncanonical binding site of the MED-1 GATA factor defines differentially regulated target genes in the *C. elegans* mesendoderm. *Dev. Cell* **2005**, *8*, 427–433. [[CrossRef](#)]
69. Lowry, J.A.; Gamsjaeger, R.; Thong, S.Y.; Hung, W.; Kwan, A.H.; Broitman-Maduro, G.; Matthews, J.M.; Maduro, M.; Mackay, J.P. Structural analysis of MED-1 reveals unexpected diversity in the mechanism of DNA recognition by GATA-type zinc finger domains. *J. Biol. Chem.* **2009**, *284*, 5827–5835. [[CrossRef](#)] [[PubMed](#)]
70. Broitman-Maduro, G.; Owrighi, M.; Hung, W.; Kuntz, S.; Sternberg, P.W.; Maduro, M. The NK-2 class homeodomain factor CEH-51 and the T-box factor TBX-35 have overlapping function in *C. elegans* mesoderm development. *Development* **2009**, *176*, 2735–2746. [[CrossRef](#)] [[PubMed](#)]
71. Maduro, M.F.; Lin, R.; Rothman, J.H. Dynamics of a developmental switch: Recursive intracellular and intranuclear redistribution of *Caenorhabditis elegans* POP-1 parallels Wnt-inhibited transcriptional repression. *Dev. Biol.* **2002**, *248*, 128–142. [[CrossRef](#)] [[PubMed](#)]
72. McGhee, J.D.; Fukushige, T.; Krause, M.W.; Minnema, S.E.; Goszczynski, B.; Gaudet, J.; Kohara, Y.; Bossinger, O.; Zhao, Y.; Khattri, J.; et al. ELT-2 is the predominant transcription factor controlling differentiation and function of the *C. elegans* intestine, from embryo to adult. *Dev. Biol.* **2009**, *327*, 551–565. [[CrossRef](#)]
73. Ewe, C.K.; Sommermann, E.M.; Kenchel, J.; Flowers, S.E.; Maduro, M.F.; Joshi, P.M.; Rothman, J.H. Feedforward regulatory logic controls the specification-to-differentiation transition and terminal cell fate during *Caenorhabditis elegans* endoderm development. *Development* **2022**, *149*, dev200337. [[CrossRef](#)] [[PubMed](#)]
74. Raj, A.; Rifkin, S.A.; Andersen, E.; van Oudenaarden, A. Variability in gene expression underlies incomplete penetrance. *Nature* **2010**, *463*, 913–918. [[CrossRef](#)] [[PubMed](#)]
75. Stevens, L.; Felix, M.A.; Beltran, T.; Braendle, C.; Caurcel, C.; Fausett, S.; Fitch, D.; Frezal, L.; Gosse, C.; Kaur, T.; et al. Comparative genomics of 10 new *Caenorhabditis* species. *Evol. Lett.* **2019**, *3*, 217–236. [[CrossRef](#)]
76. Kiontke, K.C.; Felix, M.A.; Ailion, M.; Rockman, M.V.; Braendle, C.; Penigault, J.B.; Fitch, D.H. A phylogeny and molecular barcodes for *Caenorhabditis*, with numerous new species from rotting fruits. *BMC Evol. Biol.* **2011**, *11*, 339. [[CrossRef](#)]
77. Stevens, L.; Rooke, S.; Falzon, L.C.; Machuka, E.M.; Momanyi, K.; Murungi, M.K.; Njoroge, S.M.; Odinga, C.O.; Ogendo, A.; Ogola, J.; et al. The Genome of *Caenorhabditis bovis*. *Curr. Biol. CB* **2020**, *30*, 1023–1031.e4. [[CrossRef](#)]
78. Stein, L.D.; Bao, Z.; Blasiar, D.; Blumenthal, T.; Brent, M.R.; Chen, N.; Chinwalla, A.; Clarke, L.; Clee, C.; Coghlan, A.; et al. The Genome Sequence of *Caenorhabditis briggsae*: A Platform for Comparative Genomics. *PLoS Biol.* **2003**, *1*, E45. [[CrossRef](#)]
79. Consortium, T.C.e.S. Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science* **1998**, *282*, 2012–2018.
80. Zhao, Z.; Boyle, T.J.; Bao, Z.; Murray, J.I.; Mericle, B.; Waterston, R.H. Comparative analysis of embryonic cell lineage between *Caenorhabditis briggsae* and *Caenorhabditis elegans*. *Dev. Biol.* **2008**, *314*, 93–99. [[CrossRef](#)]
81. Memar, N.; Schiemann, S.; Hennig, C.; Findeis, D.; Conradt, B.; Schnabel, R. Twenty million years of evolution: The embryogenesis of four *Caenorhabditis* species are indistinguishable despite extensive genome divergence. *Dev. Biol.* **2019**, *447*, 182–199. [[CrossRef](#)] [[PubMed](#)]
82. Winston, W.M.; Sutherlin, M.; Wright, A.J.; Feinberg, E.H.; Hunter, C.P. *Caenorhabditis elegans* SID-2 is required for environmental RNA interference. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10565–10570. [[CrossRef](#)] [[PubMed](#)]
83. Gupta, B.P.; Johnsen, R.; Chen, N. Genomics and biology of the nematode *Caenorhabditis briggsae*. In *WormBook: The Online Review of C. elegans Biology*; 2007; pp. 1–16. Available online: <https://pubmed.ncbi.nlm.nih.gov/18050493/> (accessed on 1 June 2023). [[CrossRef](#)]
84. Slos, D.; Sudhaus, W.; Stevens, L.; Bert, W.; Blaxter, M. *Caenorhabditis monodelphis* sp. n.: Defining the stem morphology and genomics of the genus *Caenorhabditis*. *BMC Zool.* **2017**, *2*, 15. [[CrossRef](#)]

85. Zhao, Z.; Flibotte, S.; Murray, J.I.; Blick, D.; Boyle, T.J.; Gupta, B.; Moerman, D.G.; Waterston, R.H. New tools for investigating the comparative biology of *Caenorhabditis briggsae* and *C. elegans*. *Genetics* **2010**, *184*, 853–863. [CrossRef]
86. Maduro, M.F.; Broitman-Maduro, G.; Choi, H.; Carranza, F.; Chia-Yi Wu, A.; Rifkin, S.A. MED GATA factors promote robust development of the *C. elegans* endoderm. *Dev. Biol.* **2015**, *404*, 66–79. [CrossRef] [PubMed]
87. Wiesenfahrt, T.; Osborne Nishimura, E.; Berg, J.Y.; McGhee, J.D. Probing and rearranging the transcription factor network controlling the *C. elegans* endoderm. *Worm* **2016**, *5*, e1198869. [CrossRef]
88. Koh, K.; Rothman, J.H. ELT-5 and ELT-6 are required continuously to regulate epidermal seam cell differentiation and cell fusion in *C. elegans*. *Development* **2001**, *128*, 2867–2880. [CrossRef]
89. Page, B.D.; Zhang, W.; Steward, K.; Blumenthal, T.; Priess, J.R. ELT-1, a GATA-like transcription factor, is required for epidermal cell fates in *Caenorhabditis elegans* embryos. *Genes Dev.* **1997**, *11*, 1651–1661. [CrossRef]
90. Tsanov, N.; Samacoits, A.; Chouaib, R.; Traboulsi, A.M.; Gostan, T.; Weber, C.; Zimmer, C.; Zibara, K.; Walter, T.; Peter, M.; et al. smiFISH and FISH-quant—A flexible single RNA detection approach with super-resolution capability. *Nucleic Acids Res.* **2016**, *44*, e165. [CrossRef]
91. Shao, J.; He, K.; Wang, H.; Ho, W.S.; Ren, X.; An, X.; Wong, M.K.; Yan, B.; Xie, D.; Stamatoyannopoulos, J.; et al. Collaborative regulation of development but independent control of metabolism by two epidermis-specific transcription factors in *Caenorhabditis elegans*. *J. Biol. Chem.* **2013**, *288*, 33411–33426. [CrossRef] [PubMed]
92. Hu, Q.; D'Amora, D.R.; MacNeil, L.T.; Walhout, A.J.M.; Kubiseski, T.J. The Oxidative Stress Response in *Caenorhabditis elegans* Requires the GATA Transcription Factor ELT-3 and SKN-1/Nrf2. *Genetics* **2017**, *206*, 1909–1922. [CrossRef]
93. Mesbahi, H.; Pho, K.B.; Tench, A.J.; Leon Guerrero, V.L.; MacNeil, L.T. Cuticle Collagen Expression Is Regulated in Response to Environmental Stimuli by the GATA Transcription Factor ELT-3 in *Caenorhabditis elegans*. *Genetics* **2020**, *215*, 483–495. [CrossRef] [PubMed]
94. Tabarraei, H.; Waddell, B.M.; Raymond, K.; Murray, S.M.; Wang, Y.; Choe, K.P.; Wu, C.W. CCR4-NOT subunit CCF-1/CNOT7 promotes transcriptional activation to multiple stress responses in *Caenorhabditis elegans*. *Aging Cell* **2023**, *22*, e13795. [CrossRef]
95. Li, R.; Ren, X.; Ding, Q.; Bi, Y.; Xie, D.; Zhao, Z. Direct full-length RNA sequencing reveals unexpected transcriptome complexity during *Caenorhabditis elegans* development. *Genome Res.* **2020**, *30*, 287–298. [CrossRef] [PubMed]
96. Krause, M.; Hirsh, D. A trans-spliced leader sequence on actin mRNA in *C. elegans*. *Cell* **1987**, *49*, 753–761. [CrossRef]
97. Hu, G.; Katuwawala, A.; Wang, K.; Wu, Z.; Ghadermarzi, S.; Gao, J.; Kurgan, L. fIDPnn: Accurate intrinsic disorder prediction with putative propensities of disorder functions. *Nat. Commun.* **2021**, *12*, 4438. [CrossRef]
98. Romano, O.; Miccio, A. GATA factor transcriptional activity: Insights from genome-wide binding profiles. *IUBMB Life* **2020**, *72*, 10–26. [CrossRef]
99. Wang, B.; Zhang, L.; Dai, T.; Qin, Z.; Lu, H.; Zhou, F. Liquid-liquid phase separation in human health and diseases. *Signal Transduct. Target. Ther.* **2021**, *6*, 290. [CrossRef]
100. Hnisz, D.; Shrinivas, K.; Young, R.A.; Chakraborty, A.K.; Sharp, P.A. A Phase Separation Model for Transcriptional Control. *Cell* **2017**, *169*, 13–23. [CrossRef]
101. Kim, J.; Kim, I.; Han, S.K.; Bowie, J.U.; Kim, S. Network rewiring is an important mechanism of gene essentiality change. *Sci. Rep.* **2012**, *2*, 900. [CrossRef] [PubMed]
102. Haag, E.S.; True, J.R. Developmental System Drift. In *Evolutionary Developmental Biology: A Reference Guide*; Nuño de la Rosa, L., Müller, G.B., Eds.; Springer International Publishing: Cham, Switzerland, 2021; pp. 99–110. [CrossRef]
103. Wilkins, A.S. Moving up the hierarchy: A hypothesis on the evolution of a genetic sex determination pathway. *BioEssays News Rev. Mol. Cell. Dev. Biol.* **1995**, *17*, 71–77. [CrossRef]
104. Lin, R.; Hill, R.J.; Priess, J.R. POP-1 and anterior-posterior fate decisions in *C. elegans* embryos. *Cell* **1998**, *92*, 229–239. [CrossRef] [PubMed]
105. Couthier, A.; Smith, J.; McGarr, P.; Craig, B.; Gilleard, J.S. Ectopic expression of a *Haemonchus contortus* GATA transcription factor in *Caenorhabditis elegans* reveals conserved function in spite of extensive sequence divergence. *Mol. Biochem. Parasitol.* **2004**, *133*, 241–253. [CrossRef]
106. Stevens, L.; Moya, N.D.; Tanny, R.E.; Gibson, S.B.; Tracey, A.; Na, H.; Chitrakar, R.; Dekker, J.; Walhout, A.J.M.; Baugh, L.R.; et al. Chromosome-Level Reference Genomes for Two Strains of *Caenorhabditis briggsae*: An Improved Platform for Comparative Genomics. *Genome Biol. Evol.* **2022**, *14*, evac042. [CrossRef]
107. Sommer, R.J. As good as they get: Cells in nematode vulva development and evolution. *Curr. Opin. Cell. Biol.* **2001**, *13*, 715–720. [CrossRef]
108. Barkoulas, M.; Vargas Velazquez, A.M.; Peluffo, A.E.; Felix, M.A. Evolution of New cis-Regulatory Motifs Required for Cell-Specific Gene Expression in *Caenorhabditis*. *PLoS Genet.* **2016**, *12*, e1006278. [CrossRef] [PubMed]
109. Sternberg, P.W. Vulval development. In *WormBook: The Online Review of C. elegans Biology*; 2005; pp. 1–28. Available online: <https://pubmed.ncbi.nlm.nih.gov/18050418/> (accessed on 1 June 2023). [CrossRef]
110. Felix, M.A. Cryptic quantitative evolution of the vulva intercellular signaling network in *Caenorhabditis*. *Curr. Biol. CB* **2007**, *17*, 103–114. [CrossRef] [PubMed]

111. Bei, Y.; Hogan, J.; Berkowitz, L.A.; Soto, M.; Rocheleau, C.E.; Pang, K.M.; Collins, J.; Mello, C.C. SRC-1 and Wnt signaling act together to specify endoderm and to control cleavage orientation in early *C. elegans* embryos. *Dev. Cell* **2002**, *3*, 113–125. [[CrossRef](#)]
112. Thorpe, C.J.; Schlesinger, A.; Bowerman, B. Wnt signalling in *Caenorhabditis elegans*: Regulating repressors and polarizing the cytoskeleton. *Trends Cell Biol.* **2000**, *10*, 10–17. [[CrossRef](#)]
113. Davis, G.K.; Patel, N.H. Short, long, and beyond: Molecular and embryological approaches to insect segmentation. *Annu. Rev. Entomol.* **2002**, *47*, 669–699. [[CrossRef](#)]
114. Lemke, S.; Stauber, M.; Shaw, P.J.; Rafiqi, A.M.; Prell, A.; Schmidt-Ott, U. *Bicoid* occurrence and *Bicoid*-dependent hunchback regulation in lower cyclorrhaphan flies. *Evol. Dev.* **2008**, *10*, 413–420. [[CrossRef](#)] [[PubMed](#)]
115. Stauber, M.; Prell, A.; Schmidt-Ott, U. A single *Hox3* gene with composite *bicoid* and *zerknüllt* expression characteristics in non-Cyclorrhaphan flies. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 274–279. [[CrossRef](#)] [[PubMed](#)]
116. Lynch, J.A.; Brent, A.E.; Leaf, D.S.; Pultz, M.A.; Desplan, C. Localized maternal *orthodenticle* patterns anterior and posterior in the long germ wasp *Nasonia*. *Nature* **2006**, *439*, 728–732. [[CrossRef](#)]
117. McGregor, A.P. How to get ahead: The origin, evolution and function of *bicoid*. *BioEssays News Rev. Mol. Cell. Dev. Biol.* **2005**, *27*, 904–913. [[CrossRef](#)] [[PubMed](#)]
118. Schroder, R. The genes *orthodenticle* and *hunchback* substitute for *bicoid* in the beetle *Tribolium*. *Nature* **2003**, *422*, 621–625. [[CrossRef](#)]
119. Kotkamp, K.; Klingler, M.; Schoppmeier, M. Apparent role of *Tribolium orthodenticle* in anteroposterior blastoderm patterning largely reflects novel functions in dorsoventral axis formation and cell survival. *Development* **2010**, *137*, 1853–1862. [[CrossRef](#)] [[PubMed](#)]
120. El-Sherif, E.; Lynch, J.A.; Brown, S.J. Comparisons of the embryonic development of *Drosophila*, *Nasonia*, and *Tribolium*. *Wiley Interdiscip. Rev. Dev. Biol.* **2012**, *1*, 16–39. [[CrossRef](#)]
121. Rudolf, H.; Zellner, C.; El-Sherif, E. Speeding up anterior-posterior patterning of insects by differential initialization of the gap gene cascade. *Dev. Biol.* **2020**, *460*, 20–31. [[CrossRef](#)]
122. Brauchle, M.; Kiontke, K.; MacMenamin, P.; Fitch, D.H.; Piano, F. Evolution of early embryogenesis in rhabditid nematodes. *Dev. Biol.* **2009**, *335*, 253–262. [[CrossRef](#)]
123. Lowe, E.K.; Stolfi, A. Developmental system drift in motor ganglion patterning between distantly related tunicates. *EvoDevo* **2018**, *9*, 18. [[CrossRef](#)]
124. Stolfi, A.; Lowe, E.K.; Racioppi, C.; Ristatore, F.; Brown, C.T.; Swalla, B.J.; Christiaen, L. Divergent mechanisms regulate conserved cardiopharyngeal development and gene expression in distantly related ascidians. *eLife* **2014**, *3*, e03728. [[CrossRef](#)] [[PubMed](#)]
125. Stolfi, A.; Levine, M. Neuronal subtype specification in the spinal cord of a protovertebrate. *Development* **2011**, *138*, 995–1004. [[CrossRef](#)] [[PubMed](#)]
126. Packer, J.S.; Zhu, Q.; Huynh, C.; Sivaramakrishnan, P.; Preston, E.; Dueck, H.; Stefanik, D.; Tan, K.; Trapnell, C.; Kim, J.; et al. A lineage-resolved molecular atlas of *C. elegans* embryogenesis at single-cell resolution. *Science* **2019**, *365*, eaax1971. [[CrossRef](#)] [[PubMed](#)]
127. Nuez, I.; Felix, M.A. Evolution of susceptibility to ingested double-stranded RNAs in *Caenorhabditis* nematodes. *PLoS ONE* **2012**, *7*, e29811. [[CrossRef](#)]
128. Tonsaker, T.; Pratt, R.M.; McGhee, J.D. Re-evaluating the role of ELT-3 in a GATA transcription factor circuit proposed to guide aging in *C. elegans*. *Mech. Ageing Dev.* **2012**, *133*, 50–53. [[CrossRef](#)]
129. Budovskaya, Y.V.; Wu, K.; Southworth, L.K.; Jiang, M.; Tedesco, P.; Johnson, T.E.; Kim, S.K. An *elt-3/elt-5/elt-6* GATA transcription circuit guides aging in *C. elegans*. *Cell* **2008**, *134*, 291–303. [[CrossRef](#)]
130. Force, A.; Lynch, M.; Pickett, F.B.; Amores, A.; Yan, Y.L.; Postlethwait, J. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **1999**, *151*, 1531–1545. [[CrossRef](#)]
131. Schierenberg, E. Three sons of fortune: Early embryogenesis, evolution and ecology of nematodes. *BioEssays News Rev. Mol. Cell. Dev. Biol.* **2001**, *23*, 841–847. [[CrossRef](#)]
132. Macchietto, M.; Angdembey, D.; Heidarpour, N.; Serra, L.; Rodriguez, B.; El-Ali, N.; Mortazavi, A. Comparative Transcriptomics of *Steinernema* and *Caenorhabditis* Single Embryos Reveals Orthologous Gene Expression Convergence during Late Embryogenesis. *Genome Biol. Evol.* **2017**, *9*, 2681–2696. [[CrossRef](#)]
133. Johnson, N.A.; Porter, A.H. Evolution of branched regulatory genetic pathways: Directional selection on pleiotropic loci accelerates developmental system drift. *Genetica* **2007**, *129*, 57–70. [[CrossRef](#)]
134. Pavlicev, M.; Wagner, G.P. A model of developmental evolution: Selection, pleiotropy and compensation. *Trends Ecol. Evol.* **2012**, *27*, 316–322. [[CrossRef](#)] [[PubMed](#)]
135. Nance, J.; Lee, J.Y.; Goldstein, B. Gastrulation in *C. elegans*. In *WormBook: The Online Review of C. elegans Biology*; 2005; pp. 1–13. Available online: <https://pubmed.ncbi.nlm.nih.gov/18050409/> (accessed on 1 June 2023). [[CrossRef](#)]
136. Robertson, S.M.; Medina, J.; Lin, R. Uncoupling different characteristics of the *C. elegans* E lineage from differentiation of intestinal markers. *PLoS ONE* **2014**, *9*, e106309. [[CrossRef](#)] [[PubMed](#)]
137. Boeck, M.E.; Boyle, T.; Bao, Z.; Murray, J.; Mericle, B.; Waterston, R. Specific roles for the GATA transcription factors *end-1* and *end-3* during *C. elegans* E-lineage development. *Dev. Biol.* **2011**, *358*, 345–355. [[CrossRef](#)]
138. Clucas, C.; Cabello, J.; Bussing, I.; Schnabel, R.; Johnstone, I.L. Oncogenic potential of a *C. elegans cdc25* gene is demonstrated by a gain-of-function allele. *Embo. J.* **2002**, *21*, 665–674. [[CrossRef](#)] [[PubMed](#)]

139. Kostic, I.; Roy, R. Organ-specific cell division abnormalities caused by mutation in a general cell cycle regulator in *C. elegans*. *Development* **2002**, *129*, 2155–2165. [[CrossRef](#)] [[PubMed](#)]
140. Hashimshony, T.; Wagner, F.; Sher, N.; Yanai, I. CEL-Seq: Single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep.* **2012**, *2*, 666–673. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.