Molecular Reproduction and Development

Developmental Robustness in the <I>C. elegans</I> Embryo

Journal:	*Molecular Reproduction and Development*
Manuscript ID:	MRD-15-0169.R1
Wiley - Manuscript type:	Review Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Maduro, Morris; University of California, Riverside, Biology
Keywords:	Endoderm, C. elegans, Robustness, Gene Regulatory Networks

Developmental Robustness in the *C. elegans* Embryo

Morris F. Maduro*

*Department of Biology, University of California, Riverside, Riverside, CA 92521

University of California, Riverside

Riverside, CA

USA 92521

mmaduro@ucr.edu

Running title: ROBUSTNESS OF EMBRYONIC DEVELOPMENT

Abbreviations: bZIP, basic leucine zipper domain; GFP, green fluorescent protein; GRN, gene regulatory network; MAPK, mitogen-activated protein kinase; smFISH, single-molecule fluorescent in situ hybridization

Funding: National Science Foundation Grant IOS#1258054

SUMMARY

Developmental robustness is the ability of an embryo to develop normally despite many sources of variation, from differences in the environment to stochastic cell-to-cell differences in gene expression. The nematode, *Caenorhabditis elegans*, exhibits an additional layer of robustness: Unlike most other animals, the embryonic pattern of cell divisions is nearly identical from animal to animal. The endoderm (gut) lineage is a model for the study of developmental robustness: The juvenile gut has a simple anatomy, consisting of 20 cells that are derived from a single cell, E; and the gene regulatory network (GRN) that controls E specification shares features in common with developmental GRNs in many systems, including genetic redundancy, parallel pathways and feed forward loops. While early studies were more concerned with identifying the genes in the network, recent work has turned to understanding how the endoderm produces a robust developmental output in the face of many sources of variation. Genetic control exists at three levels of endoderm development: Progenitor specification, the cell divisions within the developing gut, and maintenance of gut differentiation. Recent findings show that specification genes regulate all three of these aspects of gut development, and that mutant embryos can experience a "partial" specification state in which some, but not all, of the E descendants adopt a gut fate. Ongoing studies using newer quantitative and genome-wide methods promise further insights into how developmental GRNs buffer variation.

INTRODUCTION

Embryos of triploblastic animals are incredible macromolecular machines. Cells descend from the zygote through multiple mitotic divisions, precursor cells become specified, and ultimately tissue-

specific patterns of gene expression will drive morphogenesis and differentiation. Despite widely varying conditions, embryos of many species are able to form very similar juveniles with minimal phenotype variability. This phenomenon can be called *developmental robustness* and it is synonymous with *canalization*, the notion that only a finite number of developmental trajectories are possible (Waddington, 1942): Small variations are buffered so that cells become "canalized" toward a stable, consistent outcome.

In discussing robustness, biologists recognize three sources of variation: Environmental variation, genetic variation and stochastic noise (Felix and Wagner, 2008). Environmental variation includes differences in nutrition and temperature. Genetic variation includes effects of different alleles segregating in a population as a result of mutations or recombination. Stochastic noise is the result of cell-to-cell (and embryo-to-embryo) differences in protein levels caused by inherent variability in the many steps of gene expression, including the availability of transcription factors, DNA accessibility, RNA polymerase initiation rates, mRNA degradation, and rate of translation (Chalancon et al., 2012). While stochastic variation is useful in certain contexts, such as in a population of cells that are experiencing a novel stress, during development the effect of noise must be minimized (Macneil and Walhout, 2011). The ability of embryos to buffer sources of variation will determine both the consistency of form and function from one individual to the next, and the success of the species over time.

While robustness helps development, it actually works against the goals of a developmental geneticist. Genes are classically identified through mutations that result in a phenotype – that is, one must observe a *loss* of robustness in order to assign a function to a gene. Hence, the first genes identified in a developmental process tend to be the most important core regulators at the 'top' of a hierarchical network. Once a core set of genes is known, it becomes possible to identify more genes by other methods, such as by reverse genetics, analysis of expression, protein-DNA interactions, or by the

ability of knockdowns to synergize with weaker alleles of known mutants. Over time, a comprehensive regulatory network can be assembled that accounts for development of a particular tissue, from its embryonic progenitors to differentiated cells. In turn, this allows the effects of upstream perturbations to be studied on all downstream target genes.

The features that make *C. elegans* a good model system also make it useful for the study of developmental robustness. Many sources of variation can be minimized. Animals can be fed the same food source and be grown at the same temperature. Females are capable of self-fertilization, so that self-progeny of homozygous strains will be genetically identical. While mutations can occur during development or in the germline, these are likely to be rare, and their effects can be minimized by regularly obtaining reference strains from frozen storage (Brenner, 1974). Compared with most animals, *C. elegans* embryonic development is robust in an additional way: Cell divisions follow a highly stereotyped pattern that is nearly identical among wild-type embryos (Sulston et al., 1983). This stereotyped cell lineage means that even minor changes to early development can be detected by changes in cell division pattern or the tissue types made by specific cells. The combination of hermaphrodite genetics and a stereotyped lineage means that the major sources of variation in development are stochastic noise in gene expression, and any introduced genetic differences.

Several recent reviews cover gene regulatory networks and robustness in general (Macneil and Walhout 2011), mechanisms by which expression noise and networks work together (Chalancon et al., 2012), developmental robustness in other systems (Gursky et al., 2012), and robustness from an evolutionary perspective in *Caenorhabditis* (Felix and Wagner, 2008). Rather than duplicate these efforts, this review will describe how the *C. elegans* embryo serves as a model for studies of developmental robustness, focusing on the gene network that controls development of the endoderm

(gut) as this network is the best-understood. I will describe gut development from the perspective of genetic mechanisms that assure its robustness, and pose questions for future exploration.

C. elegans EMBRYONIC DEVELOPMENT

During *C. elegans* embryogenesis, a single-celled zygote develops into a 558-cell juvenile over 14 h at 25°C (Fig. 1) (Sulston et al., 1983). Embryogenesis takes place entirely within an eggshell made primarily of chitin (Olson et al., 2012). During the first half of development, the zygote undergoes mitotic divisions to generate hundreds of genetically identical cells, while in the second half, the embryo undergoes morphogenesis and elongation to adopt the shape of a worm (Sulston et al., 1983). Wild-type embryos follow nearly the same pattern of cell divisions from the zygote. This nearly-invariant pattern can be represented by a cell lineage diagram as shown in Fig. 1. The early cleavages generate six founder cells that will each generate a unique set of tissues among their descendants (Sulston et al., 1983). The founder cell E generates the 20 cells that form the larval intestine (the endoderm). As another example, the sister cell of E, called MS, generates primarily mesodermal cell types, including 31 pharynx cells and 28 muscle cells (Sulston et al., 1983).

Developing *C. elegans* embryos demonstrate some variability in cell positions, particularly in the early embryo, although later development appears to be very similar (Schnabel et al., 1997). Overall embryonic development is robust to temperatures between 12°C and 25°C, but cell divisions take longer as temperature is lowered (Schnabel et al., 1997; Prasad et al., 2011; Neves et al., 2015). The related species *C. briggsae* has a similar range for embryo viability (14°C-27°C), and even though *C. elegans* and *C. briggsae* are separated by 18-100 million years, they share a nearly identical embryonic cell lineage (Sulston et al., 1983; Stein et al., 2003; Cutter, 2008; Zhao et al., 2008). Therefore, embryogenesis is

robust to slight differences in positions of cells in the early embryo, environment (i.e. temperature) and to the accumulation of genetic changes over evolutionary time.

Unlike the early embryos of vertebrates, development of *C. elegans* is mosaic: If an embryonic cell is killed with a laser microbeam, or it fails to adopt the correct identity, its normal descendants will be missing from the larva (Sulston et al., 1983). Because the embryo does not compensate for missing or altered cells, changes to the cell lineage persist until the end of development and even adulthood. The earlier a change occurs to the lineage, the more descendants will be affected; hence, mutations affecting the earliest cell specification events will affect more descendants and usually cause overall developmental arrest. Such is the case with strong mutations that affect gut specification, which were among the first embryonic developmental mutants to be identified (Bowerman et al., 1992; Lin et al., 1995).

THE ENDODERM GENE REGULATORY NETWORK

Maternal Factors

Early *C. elegans* development requires many maternally-supplied gene products. Screens for maternal-effect lethal mutations affecting gut identified the gene *skn-1* (Bowerman et al., 1992). Embryos lacking *skn-1* activity arrest as a clump of cells lacking the normal descendants of MS, and specify gut only 20-30% of the time (Bowerman et al., 1992). In such embryos, the MS blastomere, and E, if gut is absent, produce epidermal and muscle cells, similar to the C founder cell. Consistent with SKN-1 being at the top of a gene network that specifies both MS and E, the SKN-1 protein is a transcription factor similar to bZIP proteins, and it is found in EMS, the mother of MS and E (Bowerman et al., 1993; Blackwell et al., 1994).

Two other maternal genes contribute directly to specification of gut in addition to *skn-1*. The first, *pop-1*, encodes a protein that functions in specification of many cells throughout development (Lin et al., 1998). POP-1 is TCF/LEF transcription factor that works with the co-activator SYS-1/β-catenin to contribute to E specification (Shetty et al., 2005; Phillips et al., 2007). The contribution of POP-1/SYS-1 to E specification results from transduction of a Wnt/MAPK/Src signaling pathway that occurs between EMS and its posterior neighbor P_2 (shown as an arrow on the four-cell embryo in Fig. 1) (Rocheleau et al., 1997; Thorpe et al., 1997; Bei et al., 2002). Although POP-1 contributes to gut specification in E in parallel with SKN-1, its role in E specification is not apparent unless *skn-1* activity is first eliminated: Loss of *skn-1* and *pop-1* together results in a higher number of embryos missing gut than loss of *skn-1* alone, while loss of *pop-1* by itself does not affect gut specification (Lin et al., 1995; Maduro et al., 2005b; Shetty et al., 2005). The second maternal gene, *pal-1*, encodes a transcription factor similar to those in the Caudal family of homeodomain proteins (Hunter and Kenyon, 1996). Like *pop-1*, the contribution of *pal-1* is apparent only when gut specification has been partially compromised. Simultaneous loss of *skn-1*, *pop-1* and *pal-1* results in complete elimination of gut (Maduro et al., 2005b). There are many other maternal factors that affect gut specification, but these generally affect the activity of SKN-1 or POP-1 in E, or cause inappropriate specification of gut in other cells (reviewed in Maduro, 2009).

Zygotic Factors

All three of SKN-1, POP-1 and PAL-1 are present in the early E lineage, though they soon disappear after the E cell divides, suggesting that they activate target genes in the embryo that are responsible for activating gut specification (Lin et al., 1995; Hunter and Kenyon, 1996; Bowerman, 1998). Screens for essential zygotic genes required for gut specification led to identification of *end-1* and *end-3*, which encode GATA-type transcription factors that are expressed in the early E lineage (Zhu et al., 1997;

Maduro et al., 2005a). Mutation of both genes together abolishes gut specification, resulting in arrested embryos and hatched inviable larvae completely lacking gut (Owraghi et al., 2010).

The nearly-identical genes *med-1* and *med-2* were identified by searching the *C. elegans* genome for putative GATA factor genes (Maduro et al., 2001). When *med-1* and *med-2* are both absent, all *med-1,2(-)* embryos are missing the normal descendants of MS, and 50%-85% of embryos make gut (Maduro et al., 2001; Goszczynski and McGhee, 2005; Maduro et al., 2007). Hence, like *skn-1*, the *med-1,2* genes are required for MS fate more than for E fate. Consistent with their role in specification of both MS and E, *med-1* and *med-2* are expressed in EMS, the mother of MS and E, and their gene products appear in MS and E (Maduro et al., 2001; Robertson et al., 2004; Maduro et al., 2007).

The GATA factor ELT-2 was identified as a regulator of the gut-specific esterase gene *ges-1* (Hawkins and McGhee, 1995). Expression of *elt-2* occurs later than the *med* and *end* genes, and unlike these genes, occurs for the lifespan of the animal. Loss of *elt-2* does not prevent specification, but causes gut differentiation to fail, resulting in arrested larvae with abnormal intestine (Fukushige et al., 1998). This is consistent with an essential role for ELT-2 in maintenance of intestine fate (McGhee et al., 2009). A related GATA factor gene, *elt-7*, was identified by sequence, and found to have similar expression to *elt-2*, although *elt-7* expression starts earlier (Sommermann et al., 2010). Loss of *elt-7* can enhance the phenotype of loss of *elt-2*, though loss of *elt-7* alone has no phenotype. A partial duplication of *elt-2*, called *elt-4*, does not appear to contribute a detectable function to gut development (Fukushige et al., 2003).

The E Specification Genes Form a Simple Network

The aforementioned regulators have been assembled into a Gene Regulatory Network (GRN) as shown in Fig. 2. Most of the proposed regulatory interactions have been confirmed experimentally, suggesting that the model reflects the real network topology (Fukushige et al., 1999; Maduro et al.,

2001; Maduro et al., 2002; Shetty et al., 2005). Many features of the network are apparent that suggest mechanisms by which endoderm specification is made robust.

Robustness to genetic change: Gene Duplication/Redundancy. The zygotic endodermal regulators are present as pairs of duplicated genes: *med-1,2*, *end-1,3* and *elt-2,7*. Each gene pair exhibits overlapping function, as mutation of one of the two has a weaker effect than mutation of both genes together, as shown in Table 1 (Maduro et al., 2005a; Maduro et al., 2007; Sommermann et al., 2010). Therefore, the endoderm GRN is robust to genetic variation because many genes make both overlapping and additive contributions to gut specification. Redundancy of this type has been seen in other lineages in the *C. elegans* embryo, suggesting it is a general strategy for network robustness in cell specification GRNs (Good et al., 2004; Baugh et al., 2005; Fukushige et al., 2006; Broitman-Maduro et al., 2009; Walton et al., 2015).

Robustness to noise: Network architecture. The genes in the endoderm network participate in a mostly linear cascade with a defined hierarchy. Upstream regulators interact with only the most immediate downstream tier and the one after that, in what is termed a *feed-forward loop* (Davidson, 2010). For example, MED-1,2 activates *end-3*, and both MED-1,2 and END-3 activate *end-1* (Broitman-Maduro et al., 2005; Maduro, 2006; Maduro et al., 2007). The feed-forward loop is widespread in developmental GRNs, as it favors stabilization of a regulatory state that drives development forward (Davidson, 2010). A second important motif is the *positive feedback loop* that occurs with the terminal regulator *elt-2*: Once *elt-2* is activated, it maintains its own expression (Fukushige et al., 1999). In embryos lacking *end-1* and *end-3* (which never specify endoderm), forced expression of *elt-2* is sufficient to initiate the endoderm differentiation program in most, if not all, embryonic cells (Owraghi et al., 2010). This suggests that END-1, END-3 and ELT-7 initiate activation of *elt-2*, which is then maintained by autoregulation. As *end-1*, *end-3*, *elt-7* and *elt-2* are all GATA-type transcription factors, it is likely that these regulators act through

similar *cis*-regulatory sites in the *elt-2* promoter, which is a further source of redundancy built into the network.

Another type of network topology that gives rise to robustness is *distributed robustness* (Felix and Wagner, 2008). In this paradigm, multiple different factors work in parallel to achieve the same goal. Distributed robustness is a property of the signaling pathways that specify the fate of the vulval precursors in *C. elegans* (Felix and Wagner, 2008). In the embryonic endoderm network, the maternal factors demonstrate this type of robustness, as SKN-1, PAL-1 and POP-1 converge on *end-1* and *end-3* independently of one another (Maduro et al., 2005b; Shetty et al., 2005; Maduro et al., 2015). The signal between P_2 and EMS also occurs via maternal, parallel ligands and receptors (Rocheleau et al., 1997; Thorpe et al., 1997; Bei et al., 2002). In contrast, the downstream zygotic genes in gut development exhibit a linear set of interactions reinforced by feed forward loops, as described above (Maduro, 2009).

Characterization of mutants that affect the endoderm has revealed three levels at which robustness breaks down: The E lineage cell division pattern, acquisition of gut fate by the E descendants, and differentiation of intestine cells. These are described below, in terms of the trait of the endoderm that loses robustness, the types of genetic perturbation that cause variability of the trait, and the mechanism by which robustness is lost.

GUT FUNCTION IS ROBUST TO CHANGES IN THE E LINEAGE PATTERN

As shown in Fig. 1, the E progenitor follows nearly the same pattern of cell divisions in wild-type embryos to generate 20 cells by the end of embryogenesis (Sulston et al., 1983). This pattern of divisions is highly reproducible among *C. elegans* embryos, and even in other related species, which suggests that there has been a strong selection for its maintenance (Sulston et al., 1983; Zhao et al., 2008). But the

pattern is not completely invariant, as a small number of newly-hatched larvae contain one extra gut cell, making 21 intestinal cells in total (Sulston and Horvitz, 1977; Sulston et al., 1983; Leung et al., 1999). Such rare (but detectable) variations in the lineage are seen in other tissues. For example, in the post-embryonic development of the vulva, rare phenotype variation occurs under stress or in particular wild-type genetic backgrounds (Felix and Wagner, 2008).

Two types of genetic perturbation can significantly increase the number of cell divisions within the E lineage, resulting in extra cells (hyperplasia) of the intestine. The first are mutations directly affecting cell cycle progression. The clearest examples are gain-of-function mutations in *cdc-25.1* that result in extra cell divisions specifically in the E lineage, resulting in over 35 gut cells by the end of embryonic development (Clucas et al., 2002; Kostic and Roy, 2002). Surprisingly, the presence of the extra cells affects neither the differentiation of E descendants into gut, nor the function of the adult intestine. The apparently normal phenotype of adults with intestinal hyperplasia suggests that intestine function does not depend strongly on the number of cells in the gut, or that there is some compensatory mechanism that suppresses the effects of extra cells. As a result, intestine function is robust to changes in the pattern of cell divisions.

The second type of perturbation that causes gut hyperplasia is partial loss of gut specification. All endoderm specification mutants that cause at least some loss of gut specification also cause some embryos to experience hyperplasia. For example, approximately one-third of *end-3(-)* embryos generate over 20 gut nuclei from E (Maduro et al., 2007; Maduro et al., 2015). Hyperplasia likely results from a loss of cell cycle control due to a transformation to an alternate fate (see below for details). Embryos exhibiting gut hyperplasia generally develop to adulthood, however, there is evidence that hyperplasia resulting from perturbation of the specification GRN correlates with defects in adult gut function (also

discussed below) (Maduro et al., 2015).

ROBUSTNESS OF GUT SPECIFICATION: WHEN DOES THE E CELL ADOPT A NON-GUT FATE?

Specification is compromised by loss of multiple endoderm genes together

As discussed earlier, the ability of the endoderm GRN to robustly specify a gut fate likely derives from the parallel nature of many of the upstream regulatory interactions, and the existence of some components as pairs of genes with overlapping function (Fig. 2 and Table 1). The network is robust to loss of some components individually, such as *med-1* or *end-1*, as null mutations in these genes do not result in a recognizable phenotype (Maduro et al., 2005a; Maduro et al., 2007; Boeck et al., 2011). Similarly, loss of *pop-1* or *pal-1* individually does not result in significant loss of endoderm, although other lineages become affected and hence embryos undergo developmental arrest (Lin et al., 1995; Hunter and Kenyon, 1996; Maduro et al., 2005b).

When multiple components of the endoderm GRN become compromised, embryos become less likely to specify gut (Table 1). For example, among genetically identical *med-1; end-3* double mutant embryos from the same mother, some embryos lack gut (and hence undergo developmental arrest as late embryos or first-stage larvae), while others make a gut and develop to adulthood (Maduro et al., 2007). A similar effect occurs when *skn-1* activity is lost, because expression of *med-1, med-2, end-1* and *end-3* is highly dependent on SKN-1 (Fig. 2) (Bowerman et al., 1992; Maduro et al., 2005b; Maduro et al., 2007; Raj et al., 2010). In the case of loss of *skn-1*, the ability to specify gut becomes stochastic because expression of *end-1* becomes stochastic: Those embryos that fortuitously accumulate enough *end-1* mRNA can specify gut and activate *elt-2*, as described below (Raj et al., 2010). Similarly, loss of components of the Wnt/MAPK/Src signal between P_2 and EMS compromises gut specification because POP-1 becomes a repressor of gut fate, preventing MED-1,2 from activating *end-1,3* (Lin et al., 1998;

Shetty et al., 2005; Maduro et al., 2007; Maduro et al., 2015); the partial loss of gut conferred by mutation of *mom-2*/Wnt is shown by example in Table 1. As with loss of *skn-1*, gut specification becomes stochastic when Wnt signaling is blocked because of stochastic activation of early E lineage markers including *end-1* (Robertson et al., 2014).

When none of the E descendants make gut, E adopts an alternate fate, usually that of either C or MS. Loss of activity of SKN-1, MED-1,2 or END-1,3 results in a transformation of E into a C-like cell, resulting in E producing muscle and epidermal cells (Bowerman et al., 1992; Zhu et al., 1997; Maduro et al., 2001; Maduro et al., 2005a). In contrast, blockage of components of the overlapping Wnt/MAPK/Src signal between P_2 and EMS results in a transformation of E to an MS-like precursor, producing pharynx and body muscle (Rocheleau et al., 1997; Thorpe et al., 1997; Bei et al., 2002). E can adopt characteristics of both MS and C, as occurs when *end-1,3* and *pop-1* are blocked simultaneously (Owraghi et al., 2010). Hence, it is possible to view robust E specification as a requirement to activate [END-1+END-3] to high enough levels to overcome mechanisms that specify an alternate fate.

E specification regulates the cell cycle

In normal development, the E daughter cells slow their cell cycle to allow their internalization at gastrulation, while in gut specification mutants, the E daughter cells often divide early on the ventral surface of the embryo (Bowerman et al., 1992; Zhu et al., 1997). This suggests that E specification slows down the cell cycle. The MS and C lineages both exhibit faster cell divisions than the E lineage, and both make more descendants in the embryo: C makes 46 cells, while MS makes 63 (Sulston et al., 1983). Therefore, perturbations of E specification likely cause an increase in the number of cell divisions because of the loss of cell cycle slowing and a transformation to the MS or C fates (Bowerman et al., 1992; Zhu et al., 1997; Maduro et al., 2001; Maduro et al., 2005a; Boeck et al., 2011; Robertson et al., 2014). As all the upstream factors ultimately regulate *end-1* and *end-3*, it can be inferred that END-

1+END-3 together regulate cell cycle genes in the E lineage, although this has not yet been demonstrated. A connection between specification and the cell cycle is known for the WEE-1.1 kinase, whose product regulates cell cycle progression in other systems (Kellogg, 2003). The *wee-1.1* gene is expressed in the early E lineage, its promoter has binding sites for MED-1,2, and early E lineage expression of a *wee-1.1* reporter is reduced by perturbation of Wnt/MAPK components (Wilson et al., 1999; Broitman-Maduro et al., 2005; Robertson et al., 2014). Paradoxically, alteration of some cell cycle components can affect E specification. The β-TrCP (beta-transducin repeat containing protein) orthologue LIN-23 regulates CDC-25.1 abundance and also affects gut specification (Hebeisen and Roy, 2008; Segref et al., 2010). Loss of *cdc-25.1* by RNAi can also result in an absence of gut, though this is likely due to an effect on cell cycles prior to the birth of E (Clucas et al., 2002).

Gut specification is not an "all-or-none" phenomenon

By definition, when gut specification fails to occur in an embryo, intestine tissue does not get made. Differentiated intestine cells produce birefringent gut granules that can be easily visualized under polarized light, facilitating identification of embryos that have gut or lack gut (Fig. 3A) (Bowerman et al., 1992). The earliest screens to identify specification genes recovered embryonic lethal mutations, in which embryos contained differentiated tissue, but which were unable to hatch (Priess and Thomson, 1987). When *skn-1* was first identified, arrested *skn-1* embryos were described as either having intestinal cells or not based on the presence of a patch of gut granules in terminally-arrested embryos (Bowerman et al., 1992). This binary characterization persisted through subsequent studies examining phenotypes in the Wnt/MAPK/Src pathways, as these phenotypes also result in failure of morphogenesis (Rocheleau et al., 1997; Thorpe et al., 1997; Bei et al., 2002).

A molecular model for binary gut fate choice was advanced when analysis of mRNAs in the endoderm GRN revealed that the number of *end-1* transcripts, quantified by single-molecule fluorescent

in situ hybridization (smFISH), could predict whether or not gut specification had occurred in a *skn-1* mutant embryo (Raj et al., 2010). Although smFISH requires fixation of embryos, *end-1* and *elt-2* mRNAs could be detected simultaneously during a time window when both are present in the early E lineage. It was found that while wild-type embryos achieve a certain level of *end-1* transcripts, *skn-1* mutant embryos have highly variable amounts of *end-1* mRNA, from almost none to amounts similar to wild-type (Raj et al., 2010). Hence, the apparent binary specification of gut in *skn-1* mutant embryos could be attributed to highly stochastic *end-1* mRNA (and hence of END-1 protein) coupled to a threshold requirement for END-1 for *elt-2* activation and commitment to a gut fate (Raj et al., 2010). As *skn-1* mutants were previously classified as having gut or not, it was thought that this binary choice occurs in the E cell, as diagrammed in Fig. 4A.

Specification is not appropriately described as being all or none, however. It has recently been shown that under particular conditions, some (but not all) of the E descendants can adopt a gut fate, a phenotype that can be called "partial specification" or "hypomorphic specification" (Maduro et al., 2015). The isogenic embryos shown in Fig. 3A are mutant for *end-1* and *end-3*, and carry a chromosomal *end-1* transgene in which the MED-1,2 binding sites have been mutated (Maduro et al., 2015). In this strain, most terminal embryos lack gut granules, many have a normal-looking patch of gut (black arrowheads), and some have tiny patches of gut (white arrowheads). These different-sized regions correspond to varying numbers of gut cells as analyzed by gut-specific transgene reporters. Most perturbations that cause loss of gut specification part of the time (i.e. between 0% and 100) result in production of at least some embryos that have *fewer* than 20 gut nuclei (hypoplasia), in addition to embryos with *greater* than 20 gut nuclei (hyperplasia) (Maduro et al., 2007; Robertson et al., 2014; Maduro et al., 2015). The stronger the defect in gut specification, measured as the proportion of embryos made that have no gut, the more likely that hypoplasia will be seen over hyperplasia among embryos making gut (Maduro et al., 2007; Maduro et al., 2015). Paradoxically, both types of embryos

co-occur with increased cell divisions within the E lineage. In the case of gut hyperplasia, most of these E descendants adopt a gut fate, while in the case of hypoplasia, fewer descendants of E have adopted a gut fate. Figs. 3B-3J show examples of embryos in a partial gut specification strain in which all, some, or none of the E descendants have adopted a gut fate.

A model for partial specification of E

Partial gut specification can be explained if a binary fate choice can also occur stochastically in the descendants of E (Fig. 4B). When [END-1+END-3] become compromised, activation of *elt-2* becomes delayed and stochastic, so commitment to gut fate occurs in some E lineage descendants, but not necessarily all. Although a more comprehensive analysis of the dynamics of *elt-2* expression in such embryos has yet to be reported, variability in onset of *elt-2*::GFP expression was observed in partial gut specification strains as predicted by this model (Maduro et al., 2015), and computer modeling suggests that stochastic, later specification can easily generate the type of partial guts observed in late-stage embryos (Farhad Ghamsari and MM, unpublished).

ROBUSTNESS OF DIFFERENTIATION: IF GUT IS MADE, DOES IT FUNCTION NORMALLY?

Intestine differentiation requires *elt-2* but not *elt-7*

The main regulator of intestine fate is ELT-2 (McGhee et al., 2009). Both *elt-2* and a related gene, *elt-7*, are expressed from the time of specification through adulthood (Fukushige et al., 1998; Sommermann et al., 2010). Loss of *elt-2* is compatible with gut specification, however intestine differentiation is not properly executed, and animals arrest as L1 larvae with various intestine defects (Fukushige et al., 1998). As described above, loss of *elt-7* by itself results in no detectable phenotype, however a double mutant of *elt-2* and *elt-7* results in a stronger phenotype than loss of *elt-2* alone,

suggesting that ELT-7 plays a reinforcing role (likely in parallel with END-1 and END-3) in *elt-2* activation (Sommermann et al., 2010). Given the multiple inputs that activate *end-1* and *end-3*, it is surprising that the maintenance of intestinal fate is not at all robust to the loss of a single gene (*elt-2*). One explanation may be that it may be difficult to recruit new pathways into maintenance of intestine fate. There are thousands of genes that appear to be intestine-enriched and regulated by ELT-2, and co-option of a new regulator into endoderm differentiation would require that thousands of target genes acquire regulatory input from such a new factor (Maduro, 2006; Pauli et al., 2006; McGhee et al., 2007).

Another interpretation of the requirement for *elt-2* in gut fate is that robustness in the endoderm has evolved only in the earliest steps of the network to assure a highly invariant level of *elt-2* expression downstream. Quantification of mRNA showed that *elt-2* appears to have a relatively constant number of transcripts per cell in the early embryo (Raj et al., 2010). Although variation in *elt-2* expression has not been examined in adults, expression of a single-copy reporter for *hsp-16.2* in the intestine is surprisingly similar from animal to animal (75%-88% correlation) (Mendenhall et al., 2015), suggesting that expression variation is generally low in the adult gut. Hence, *C. elegans* may not have evolved mechanisms to accommodate variation in gut gene expression, relying instead on upstream components to assure consistency of expression of *elt-2.*

Partial E specification compromises adult gut function

As mentioned above, consistent output of embryonic development is synonymous with the existence of only a finite number of stable end states (Waddington, 1942). A recent work in *Drosophila* provides cellular-level support for the idea that even when a major upstream regulator is lost, later embryonic cells themselves are driven to stable differentiated states. When the key maternal regulator *bcd* was blocked in early embryos, cell fates nonetheless adopted fates present in wild-type embryos,

and not novel fates, as revealed by patterns of downstream regulator gene expression (Staller et al., 2015). There is evidence that this is also true in the *C. elegans* endoderm. When E specification is partially compromised, three markers – expression of reporters for *elt-2* and the differentiation marker *pept-2*, and gut granules – appear in those E descendants that make gut (Maduro et al., 2015). This suggests that cells that adopt a gut fate in such embryos are comparable to wild-type gut cells. Furthermore, when embryos do not activate both *end-1* and *end-3*, the E cell undergoes a complete transformation into the C lineage, as shown by following the early cell divisions of the transformed E cell (which become similar to those of C), and by appearance of markers of C fates (PAL-1-dependent muscle and epidermis) in descendants of E (Zhu et al., 1997; Maduro et al., 2005a; Owraghi et al., 2010). However, recent findings suggest that, just as specification is not an "all-or-none" phenomenon, activation of differentiation may not be self-correcting for earlier defects in specification, even when an apparently functional embryonic intestine is made.

To examine intestine *function* downstream of partially compromised gut specification, endoderm specification must be compromised in a way that leaves other lineages unaffected, so that gut function in surviving animals can be evaluated in a complete larva or adult. However, most perturbations in E specification either affect other lineages (resulting in embryonic or larval arrest), cause complete loss of gut, or cause only weak effects (Bowerman et al., 1992; Lin et al., 1995; Rocheleau et al., 1997; Thorpe et al., 1997; Rocheleau et al., 1999; Maduro et al., 2001; Maduro et al., 2005a; Maduro et al., 2005b). An engineered strain that meets this requirement is one in which *end-1* and *end-3* activity derives solely from single-copy *end-1* and *end-3* transgenes with mutated MED-1,2 binding sites (Maduro et al., 2015). This strain exhibits gut specification in 75% of embryos, and most embryos are able to survive to adulthood with at least a wild-type number of gut cells. However, these animals exhibit a developmental delay and fat accumulation that resembles the phenotypes of adults subjected to caloric restriction during larval development (Palgunow et al., 2012; Maduro et al., 2015).

Similar phenotypes are observed among surviving adults of a *med-1; end-3* double mutant strain, which exhibits gut specification in 42% of embryos (Maduro et al., 2007; Maduro et al., 2015). These results suggest that proper adult gut function is not robust to perturbations of the early endoderm lineage, even though an apparently normal gut has been made. In turn, it can be inferred that proper gut function is assured by the robustness of gut specification.

There are several mechanisms that could explain why partial specification might result in reduced quality of differentiation. First, many of the adults that survive partial specification demonstrate intestinal hyperplasia, suggesting that extra gut cells may be the cause of poor-quality differentiation (Maduro et al., 2015). However, the average amount of hyperplasia, ~5 extra intestinal nuclei per adult, is much milder than what is seen in *cdc-25.1* gain-of-function mutants (~25 extra nuclei), and major defects in gut function were not seen in the latter (Kostic and Roy, 2002; Maduro et al., 2015). Second, there is a delay in *elt-2* activation in partial gut specification strains, suggesting that late activation of differentiation misses an optimal window for activation of other unknown factors by *end-1* and *end-3* (Boeck et al., 2011; Maduro et al., 2015). Finally, it may be that weak initiation does not allow *elt-2* to achieve the same level of postembryonic expression as in wild-type. In other systems, positive autoregulation can be vulnerable to weaker initial input (Maeda and Sano, 2006), suggesting that robust, high-level initiation of *elt-2* is essential for proper differentiation and intestine function. Loss of stable [END-1+END-3] might therefore propagate into weak and unstable activation of *elt-2* that could persist for many cell generations.

A framework that connects cell cycle, specification and differentiation

A model is presented in Fig. 5 that relates the levels of END-1+END-3 with the development of the gut. At maximal END activity, robust specification occurs and E produces 20 descendants and a

normal intestine. When all END activity is abolished, the E cell fails to generate intestinal cells and all animals arrest development at the end of embryogenesis or immediately after hatching. At intermediate levels, the E lineage, E specification and intestine function all become compromised due to reduced and stochastic activation of END target genes. In most of these embryos, the number of E descendants increases, due to a failure of the ENDs to slow down the cell cycle. If all of these E descendants adopt an intestinal fate, the result is a functional gut with (possible) concomitant defects in adult intestine function. If only some of these adopt a gut fate, the result is hypoplasia of the gut, and an inability to survive embryonic development. Cell cycle regulator defects within the E lineage can induce hyperplasia but these generally do not influence specification.

Evolutionary Changes in the Endoderm Gene Regulatory Network

The embryonic cell lineage, which includes the cell division patterns in the intestine, is nearly identical between *C. elegans* and the related species *C. briggsae* (Zhao et al., 2008). However, there is evidence that the underlying gene regulatory network is subject to more rapid evolution compared with the differentiation network. For example, the *med* GATA factor genes appear to be rapidly-evolving, with many more *med* genes in *C. remanei* and *C. briggsae* (Coroian et al., 2005). Perhaps more surprising is the observation that knockdown of *skn-1* and *pop-1* has different effects on endoderm specification in *C. briggsae* as compared with *C. elegans* (Lin et al., 2009). These differences reveal that regulatory interactions in the endoderm GRN can change over evolutionary time without any effect on the specification of E and the development of the gut. Changes in underlying GRNs among related species, in the absence of phenotype change, are generally called *developmental system drift* (True and Haag, 2001). In this paradigm, the ability of multiple pathways to drive the same processes allows them to acquire different relative importance over evolutionary time. When examined closely, underlying

regulatory mechanisms can be quite different, although they robustly cause the same developmental outcome, and usually involve at least some regulators in common. Examples of developmental system drift in *C. elegans* and its close relatives include changes in the core sex determination pathway, particularly in the modifications that lead to hermaphroditism (Ellis and Lin, 2014), and in the development of the vulva (Felix and Barkoulas, 2012). Given enough evolutionary distance, larger changes become apparent: In the endoderm network, while *elt-2* is conserved, the *med-1,2* and *end-1,3* genes appear to be absent from genomes outside of the Elegans supergroup of *Caenorhabditis* species, suggesting that early gut specification outside this group does not involve GATA factors (Christian Turner and MM, unpublished observations) (Kiontke et al., 2011). Presumably, different zygotic regulators exist in these other species that perform the function of the MEDs and ENDs. Identifying these may reveal cryptic mechanisms that may be partially conserved in *C. briggsae* and could account for the differences in *skn-1* and *pop-1* knockdown with *C. elegans*. The existence of multiple pathways in gut specification, therefore, may be a way to achieve robustness of development while providing the means by which developmental system drift can occur over evolutionary time.

FUTURE PROSPECTS

Although the proposed structure of the *C. elegans* endoderm Gene Regulatory Network has remained largely unchanged in over a decade, there is still much to be learned. For example, newer methods allow quantitative measurements to be made of gene expression at the level of individual mRNAs (Raj et al., 2010) and protein (Raj et al., 2010; Tanenbaum et al., 2014; Mendenhall et al., 2015). Embryonic cell lineages can be followed semi-automatically using fluorescent reporters (Bao et al., 2006). It is also possible to define the transcriptome of individual embryonic cells (Hashimshony et al., 2012). High-resolution analysis of the effect of specific *cis*-regulatory sites could be made by introducing

promoter mutations into genes in their native chromosomal context (Waaijers and Boxem, 2014). These newer approaches could be used in combinations to probe gene expression variation under normal conditions and as a result of specific network perturbations; thus, it should be possible to determine the roles of all the known factors in controlling variation at the level of gene products, and in overall phenotype.

There remain unanswered questions about the architecture of the endoderm gene regulatory network and how embryos assure robustness of intestinal development. For example, are there more regulators in the endoderm specification network that act to reinforce cell fate and differentiation? Do different perturbations of the endoderm GRN all cause the same type of partial gut phenotypes? What happens to non-gut derivatives of E in partial specification strains? What subtle phenotypic effects result from the presence of extra gut cells? Is expression noise in the endoderm specification network buffered by microRNAs as occurs in other developmental contexts (Burke et al., 2015; Schmiedel et al., 2015)? How do differences in genetic background (e.g. from wild isolates of *C. elegans*) or in environment (e.g. temperature) influence the proportion of embryos that make gut in partial specification mutants (Francesconi and Lehner, 2014; Vu et al., 2015)? What is the role of more general gene expression factors in buffering gene expression noise (Burga et al., 2011)? How is gene expression noise different in other species as a result of developmental system drift (True and Haag, 2001)? And finally, how do the E specification genes control the robust pattern of cell divisions in the E lineage?

CONCLUSIONS AND PERSPECTIVES

The *C. elegans* endoderm GRN is a hierarchical network of regulators that exhibits a broad array of inputs in its earlier steps that gives way to a more linear topology that is reinforced by feed-forward loops, and in the case of the gut fate determinant *elt-2*, positive autoregulation. Through mutational

analysis, it is apparent that the structure of the network determines how all of its components contribute to robustness at the level of the cell cycle pattern of the E lineage, the acquisition of gut fate by the E descendants, and expression of differentiation genes. A surprising result has been that adult intestinal defects can result from "hypomorphic" specification, challenging the notion that cell fates always canalize toward stable "normal" fates (Waddington, 1942). This result has important implications for the study of development in any metazoan system, and the programming of stem cells, and in specialized instances of transdifferentiation (Ladewig et al., 2013). Cells that acquire a (new) differentiated state, especially in an *in vitro* system, may overtly acquire the properties of a particular tissue type, for example, by expression of a tissue-specific transcription factor. However, the tissue may not exhibit completely normal function. It is therefore important to confirm a correct differentiated state by additional means, such as by phenotype, or through genome-wide profiling of the transcriptome or proteome. This may be especially important for induced pluripotent stem cells, which can vary in the quality of their differentiation depending on their lineage of origin (Lee et al., 2014).

Looking forward, newer methods will allow a more complete understanding of how gene expression noise is buffered in developmental GRNs, and determine the generality of the observation that early effects on specification may manifest as subtle (but physiologically detectable) changes in the function of the differentiated tissue. The *C. elegans* endoderm will therefore continue to be a useful model for the study of robustness in developmental gene regulatory networks.

ACKNOWLEDGMENTS

I thank Gina Broitman-Maduro, Hailey Choi, Farhad Ghamsari and Christian Turner for contributing unpublished results; two anonymous reviewers for their helpful comments; and Alex Mendenhall for useful insights. This work was supported by NSF grant IOS#1258054 to M.M.

Figure Legends

Fig. 1. Embryonic origin of the *C. elegans* intestine. Top: Diagrams of embryos at various stages of development. Approximate times are for development at 25˚C. Anterior is to the left and dorsal is up. Nuclei of cells in the endoderm lineage are shown as green circles when the E lineage comprises 1, 8 and 20 cells (1E, 8E and 20E, respectively). After 430 minutes, embryonic divisions within the gut are complete. The embryo hatches at approximately 12 h after fertilization. Bottom: Abbreviated cell lineage showing origin of the founder cells AB, MS, E, C, D and P_4, after Sulston et al., (1983). A vertical line is a cell and a horizontal line is a cell division. The founder cells undergo further divisions. The pattern that generates the 20 cells of the intestine is shown in green. A diagram of a larva shows the intestine along with other components of the digestive tract and the blastomere of origin. ABa is the anterior daughter of AB and ABp is the posterior daughter. The egg and larva are approximately 50 µm and 200 µm long, respectively. Diagrams were adapted from WormAtlas (http://wormatlas.org/) and other sources (Sulston et al., 1983; Maduro and Rothman, 2002).

Fig. 2. Simplified diagram of the Endoderm Gene Regulatory Network (GRN) in *C. elegans*. Green font indicates factors that are contributed maternally, while factors in black are zygotic. Most interactions are known to be direct: SKN-1 to *med-1,2* (Maduro et al., 2001); MED-1,2 to *end-1* and *end-3* (Broitman-Maduro et al., 2005; Maduro et al., 2015); POP-1 on *end-1* and *end-3* (Robertson et al., 2014; Shetty et al., 2005); and PAL-1 on *end-1* and *end-3* (Maduro et al., 2015). Other interactions are implied genetically, including END-3 on *end-1* (Maduro et al., 2007) and SKN-1 on *end-1* and *end-3* (Maduro et al., 2005b). There is also some evidence for a maternal contribution of MED-1,2 (Maduro et al., 2007). Left to right: The wild-type network; network missing END-1 but which results in normal intestine (Maduro et al., 2005a; Boeck et al., 2011); loss of robust specification in the absence of MED-1,2 input

into the ENDs (Maduro et al., 2015); loss of most gut when SKN-1 is absent, which also removes MED-1,2 (Bowerman et al., 1992; Maduro et al., 2001); and a malformed intestine when ELT-2 is absent (Fukushige et al., 1998).

Fig. 3. Evidence for partial specification of gut when the endoderm GRN is perturbed at the level of *end-1,3*. (A) Polarized light showing gut granules (white spots) in a collection of isogenic *end-1,3(-)* embryos partially rescued by an *end-1* transgene in which the MED binding sites have been mutated (Maduro et al., 2015). Most embryos lack gut granules, some have a near-normal amount (yellow arrowheads), and others have very small amounts (white arrowheads). (B-J) Examples of embryos in which all, some or none of the E descendants have adopted a gut fate. (B, C, D) Differential Interference Contrast (DIC) microscopy images of embryos at 8E-12E stage. The cells forming the embryonic intestine are outlined in white dashed lines in B and C. (E, F, G) Stacks of images showing a nuclear-localized *end-3*::mCherry transcriptional reporter in the descendants of the E cell. Nuclei adopting a non-gut identity are outlined by a dashed cyan line. The reporter was made previously (Murray et al., 2008). Because all upstream regulators of *end-3* are present, the reporter is activated in the early E lineage even when no gut is made. (H, I, J) Expression of a nuclear-localized *elt-2*::GFP reporter shows nuclei of cells that have adopted a gut fate. In (H), all E descendants express *elt-2*::GFP; in (I), only some descendants do; and in (J), none of the E descendants express *elt-2*::GFP. In (E-J) the eggshell is outlined in yellow. A *C. elegans* embryo is approximately 50 µm long. In (B-J), anterior is to the left.

Fig. 4. Models for how gut specification can fail. (A) Binary (all-or-none) model in which gut specification either occurs or it does not. (B) A schematic diagram of how "partial" specification leads to fewer gut nuclei due to a later commitment to gut fate by some E descendants and not others. For simplicity, the same pattern of cell divisions is maintained across the various examples.

Fig. 5. Relationship of the activity of END-1+END-3 with gut specification and organogenesis of the intestine. At high levels, specification and development of the intestine are robust. At intermediate levels, development becomes stochastic, and animals can experience the absence of gut (aplasia), changes in gut cell number (hypo- or hyperplasia), or no change (normoplasia). Animals with compromised END-1+END-3 can survive hyperplasia, although they exhibit defects in metabolism. Modified from a prior work (Maduro et al., 2015).

REFERENCES

Bao Z, Murray JI, Boyle T, Ooi SL, Sandel MJ, Waterston RH. 2006. Automated cell lineage tracing in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America 103(8):2707-2712.

Baugh LR, Wen JC, Hill AA, Slonim DK, Brown EL, Hunter CP. 2005. Synthetic lethal analysis of Caenorhabditis elegans posterior embryonic patterning genes identifies conserved genetic interactions. Genome Biology 6(5):R45.

Bei Y, Hogan J, Berkowitz LA, Soto M, Rocheleau CE, Pang KM, Collins J, Mello CC. 2002. SRC-1 and Wnt signaling act together to specify endoderm and to control cleavage orientation in early C. elegans embryos. Developmental Cell 3(1):113-125.

Blackwell TK, Bowerman B, Priess JR, Weintraub H. 1994. Formation of a monomeric DNA binding domain by Skn-1 bZIP and homeodomain elements. Science 266(5185):621-628.

Boeck ME, Boyle T, Bao Z, Murray J, Mericle B, Waterston R. 2011. Specific roles for the GATA transcription factors end-1 and end-3 during C. elegans E-lineage development. Developmental Biology 358(2):345-355.

Bowerman B. 1998. Maternal control of pattern formation in early Caenorhabditis elegans embryos. Curr Top Dev Biol 39:73-117.

Bowerman B, Draper BW, Mello CC, Priess JR. 1993. The maternal gene skn-1 encodes a protein that is distributed unequally in early C. elegans embryos. Cell 74(3):443-452.

Bowerman B, Eaton BA, Priess JR. 1992. skn-1, a maternally expressed gene required to specify the fate of ventral blastomeres in the early C. elegans embryo. Cell 68(6):1061-1075.

Brenner S. 1974. The genetics of Caenorhabditis elegans. Genetics 77(1):71-94.

Broitman-Maduro G, Maduro MF, Rothman JH. 2005. The noncanonical binding site of the MED-1 GATA factor defines differentially regulated target genes in the C. elegans mesendoderm. Developmental Cell 8(3):427-433.

Broitman-Maduro G, Owraghi M, Hung W, Kuntz S, Sternberg PW, Maduro M. 2009. The NK-2 class homeodomain factor CEH-51 and the T-box factor TBX-35 have overlapping function in C. elegans mesoderm development. Development 176:2735-2746.

Burga A, Casanueva MO, Lehner B. 2011. Predicting mutation outcome from early stochastic variation in genetic interaction partners. Nature 480(7376):250-253.

Burke SL, Hammell M, Ambros V. 2015. Robust Distal Tip Cell Pathfinding in the Face of Temperature Stress Is Ensured by Two Conserved microRNAS in Caenorhabditis elegans. Genetics 200(4):1201-1218.

Chalancon G, Ravarani CN, Balaji S, Martinez-Arias A, Aravind L, Jothi R, Babu MM. 2012. Interplay between gene expression noise and regulatory network architecture. Trends in genetics : TIG 28(5):221-232.

Clucas C, Cabello J, Bussing I, Schnabel R, Johnstone IL. 2002. Oncogenic potential of a C.elegans cdc25 gene is demonstrated by a gain-of-function allele. The EMBO journal 21(4):665-674.

Coroian C, Broitman-Maduro G, Maduro MF. 2005. Med-type GATA factors and the evolution of mesendoderm specification in nematodes. Developmental biology 289:444-455.

Cutter AD. 2008. Divergence times in Caenorhabditis and Drosophila inferred from direct estimates of the neutral mutation rate. Mol Biol Evol 25(4):778-786.

Davidson EH. 2010. Emerging properties of animal gene regulatory networks. Nature 468(7326):911-920.

Ellis RE, Lin SY. 2014. The evolutionary origins and consequences of self-fertility in nematodes. F1000prime Reports 6:62.

Felix MA, Barkoulas M. 2012. Robustness and flexibility in nematode vulva development. Trends in genetics : TIG 28(4):185-195.

Felix MA, Wagner A. 2008. Robustness and evolution: concepts, insights and challenges from a developmental model system. Heredity 100(2):132-140.

Francesconi M, Lehner B. 2014. The effects of genetic variation on gene expression dynamics during development. Nature 505(7482):208-211.

Fukushige T, Brodigan TM, Schriefer LA, Waterston RH, Krause M. 2006. Defining the transcriptional redundancy of early bodywall muscle development in C. elegans: evidence for a unified theory of animal muscle development. Genes & Development 20(24):3395-3406.

Fukushige T, Goszczynski B, Tian H, McGhee JD. 2003. The Evolutionary Duplication and Probable Demise of an Endodermal GATA Factor in Caenorhabditis elegans. Genetics 165(2):575-588.

Fukushige T, Hawkins MG, McGhee JD. 1998. The GATA-factor elt-2 is essential for formation of the Caenorhabditis elegans intestine. Developmental Biology 198(2):286-302.

Fukushige T, Hendzel MJ, Bazett-Jones DP, McGhee JD. 1999. Direct visualization of the elt-2 gut-specific GATA factor binding to a target promoter inside the living Caenorhabditis elegans embryo. Proceedings of the National Academy of Sciences of the United States of America 96(21):11883-11888.

Good K, Ciosk R, Nance J, Neves A, Hill RJ, Priess JR. 2004. The T-box transcription factors TBX-37 and TBX-38 link GLP-1/Notch signaling to mesoderm induction in C. elegans embryos. Development 131(9):1967-1978.

Goszczynski B, McGhee JD. 2005. Re-evaluation of the role of the med-1 and med-2 genes in specifying the C. elegans endoderm. Genetics 171(2):545-555.

Gursky VV, Surkova SY, Samsonova MG. 2012. Mechanisms of developmental robustness. Bio Systems.

Hashimshony T, Wagner F, Sher N, Yanai I. 2012. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell Reports 2(3):666-673.

Hawkins MG, McGhee JD. 1995. elt-2, a second GATA factor from the nematode Caenorhabditis elegans. The Journal of Biological Chemistry 270(24):14666-14671.

Hebeisen M, Roy R. 2008. CDC-25.1 stability is regulated by distinct domains to restrict cell division during embryogenesis in C. elegans. Development 135(7):1259-1269.

Hunter CP, Kenyon C. 1996. Spatial and temporal controls target pal-1 blastomere-specification activity to a single blastomere lineage in C. elegans embryos. Cell 87(2):217-226.

Kellogg DR. 2003. Wee1-dependent mechanisms required for coordination of cell growth and cell division. Journal of Cell Science 116(Pt 24):4883-4890.

Kiontke KC, Felix MA, Ailion M, Rockman MV, Braendle C, Penigault JB, Fitch DH. 2011. A phylogeny and molecular barcodes for Caenorhabditis, with numerous new species from rotting fruits. BMC Evolutionary Biology 11:339.

Kostic I, Roy R. 2002. Organ-specific cell division abnormalities caused by mutation in a general cell cycle regulator in C. elegans. Development 129(9):2155-2165.

Ladewig J, Koch P, Brustle O. 2013. Leveling Waddington: the emergence of direct programming and the loss of cell fate hierarchies. Nature reviews Molecular Cell Biology 14(4):225-236.

Lee JH, Lee JB, Shapovalova Z, Fiebig-Comyn A, Mitchell RR, Laronde S, Szabo E, Benoit YD, Bhatia M. 2014. Somatic transcriptome priming gates lineage-specific differentiation potential of human-induced pluripotent stem cell states. Nature Communications 5:5605.

Leung B, Hermann GJ, Priess JR. 1999. Organogenesis of the Caenorhabditis elegans intestine. Developmental Biology 216(1):114-134.

Lin KT, Broitman-Maduro G, Hung WW, Cervantes S, Maduro MF. 2009. Knockdown of SKN-1 and the Wnt effector TCF/POP-1 reveals differences in endomesoderm specification in C. briggsae as compared with C. elegans. Developmental Biology 325(1):296-306.

Lin R, Hill RJ, Priess JR. 1998. POP-1 and anterior-posterior fate decisions in C. elegans embryos. Cell 92(2):229-239.

Lin R, Thompson S, Priess JR. 1995. pop-1 encodes an HMG box protein required for the specification of a mesoderm precursor in early C. elegans embryos. Cell 83(4):599-609.

Macneil LT, Walhout AJ. 2011. Gene regulatory networks and the role of robustness and stochasticity in the control of gene expression. Genome Research 21(5):645-657.

Maduro M, Hill RJ, Heid PJ, Newman-Smith ED, Zhu J, Priess J, Rothman J. 2005a. Genetic redundancy in endoderm specification within the genus Caenorhabditis. Developmental biology 284:509-522.

Maduro MF. 2006. Endomesoderm specification in Caenorhabditis elegans and other nematodes. Bioessays 28(10):1010-1022.

Maduro MF. 2009. Structure and evolution of the C. elegans embryonic endomesoderm network. Biochimica et Biophysica Acta 1789(4):250-260.

Maduro MF, Broitman-Maduro G, Choi H, Carranza F, Chia-Yi Wu A, Rifkin SA. 2015. MED GATA factors promote robust development of the C. elegans endoderm. Developmental Biology 404:66-79.

Maduro MF, Broitman-Maduro G, Mengarelli I, Rothman JH. 2007. Maternal deployment of the embryonic SKN-1-->MED-1,2 cell specification pathway in C. elegans. Developmental Biology 301(2):590-601.

Maduro MF, Kasmir JJ, Zhu J, Rothman JH. 2005b. The Wnt effector POP-1 and the PAL-1/Caudal homeoprotein collaborate with SKN-1 to activate C. elegans endoderm development. Developmental Biology 285:510-523.

Maduro MF, Lin R, Rothman JH. 2002. Dynamics of a developmental switch: recursive intracellular and intranuclear redistribution of Caenorhabditis elegans POP-1 parallels Wnt-inhibited transcriptional repression. Developmental Biology 248(1):128-142.

Maduro MF, Meneghini MD, Bowerman B, Broitman-Maduro G, Rothman JH. 2001. 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 7(3):475-485.

Maduro MF, Rothman JH. 2002. Making worm guts: the gene regulatory network of the Caenorhabditis elegans endoderm. Developmental Biology 246(1):68-85.

Maeda YT, Sano M. 2006. Regulatory dynamics of synthetic gene networks with positive feedback. Journal of Molecular Biology 359(4):1107-1124.

McGhee JD, Fukushige T, Krause MW, Minnema SE, Goszczynski B, Gaudet J, Kohara Y, Bossinger O, Zhao Y, Khattra J, Hirst M, Jones SJ, Marra MA, Ruzanov P, Warner A, Zapf R, Moerman DG, Kalb JM. 2009. ELT-2 is the predominant transcription factor controlling differentiation and function of the C. elegans intestine, from embryo to adult. Developmental Biology 327(2):551-565.

McGhee JD, Sleumer MC, Bilenky M, Wong K, McKay SJ, Goszczynski B, Tian H, Krich ND, Khattra J, Holt RA, Baillie DL, Kohara Y, Marra MA, Jones SJ, Moerman DG, Robertson AG. 2007. The ELT-2 GATA-factor and the global regulation of transcription in the C. elegans intestine. Developmental Biology 302(2):627-645.

Mendenhall AR, Tedesco PM, Sands B, Johnson TE, Brent R. 2015. Single Cell Quantification of Reporter Gene Expression in Live Adult Caenorhabditis elegans Reveals Reproducible Cell-Specific Expression Patterns and Underlying Biological Variation. PloS One 10(5):e0124289.

Murray JI, Bao Z, Boyle TJ, Boeck ME, Mericle BL, Nicholas TJ, Zhao Z, Sandel MJ, Waterston RH. 2008. Automated analysis of embryonic gene expression with cellular resolution in C. elegans. Nature Methods 5(8):703-709.

Neves A, Busso C, Gonczy P. 2015. Cellular hallmarks reveal restricted aerobic metabolism at thermal limits. eLife 2015;4;e04810.

Olson SK, Greenan G, Desai A, Muller-Reichert T, Oegema K. 2012. Hierarchical assembly of the eggshell and permeability barrier in C. elegans. The Journal of Cell Biology 198(4):731-748.

Owraghi M, Broitman-Maduro G, Luu T, Roberson H, Maduro MF. 2010. Roles of the Wnt effector POP-1/TCF in the C. elegans endomesoderm specification gene network. Developmental Biology 340(2):209-221.

Palgunow D, Klapper M, Doring F. 2012. Dietary restriction during development enlarges intestinal and hypodermal lipid droplets in Caenorhabditis elegans. PloS One 7(11):e46198.

Pauli F, Liu Y, Kim YA, Chen PJ, Kim SK. 2006. Chromosomal clustering and GATA transcriptional regulation of intestine-expressed genes in C. elegans. Development 133(2):287-295.

Phillips BT, Kidd AR, 3rd, King R, Hardin J, Kimble J. 2007. Reciprocal asymmetry of SYS-1/beta-catenin and POP-1/TCF controls asymmetric divisions in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America 104(9):3231-3236.

Prasad A, Croydon-Sugarman MJ, Murray RL, Cutter AD. 2011. Temperature-dependent fecundity associates with latitude in Caenorhabditis briggsae. Evolution; International Journal of Organic Evolution 65(1):52-63.

Priess JR, Thomson JN. 1987. Cellular interactions in early C. elegans embryos. Cell 48(2):241-250.

Raj A, Rifkin SA, Andersen E, van Oudenaarden A. 2010. Variability in gene expression underlies incomplete penetrance. Nature 463(7283):913-918.

Robertson SM, Medina J, Lin R. 2014. Uncoupling different characteristics of the C. elegans E lineage from differentiation of intestinal markers. PloS One 9(9):e106309.

Robertson SM, Shetty P, Lin R. 2004. Identification of lineage-specific zygotic transcripts in early Caenorhabditis elegans embryos. Developmental Biology 276(2):493-507.

Rocheleau CE, Downs WD, Lin R, Wittmann C, Bei Y, Cha YH, Ali M, Priess JR, Mello CC. 1997. Wnt signaling and an APC-related gene specify endoderm in early C. elegans embryos. Cell 90(4):707-716.

Rocheleau CE, Yasuda J, Shin TH, Lin R, Sawa H, Okano H, Priess JR, Davis RJ, Mello CC. 1999. WRM-1 activates the LIT-1 protein kinase to transduce anterior/posterior polarity signals in C. elegans. Cell 97(6):717-726.

Schmiedel JM, Klemm SL, Zheng Y, Sahay A, Bluthgen N, Marks DS, van Oudenaarden A. 2015. Gene expression. MicroRNA control of protein expression noise. Science 348(6230):128-132.

Schnabel R, Hutter H, Moerman D, Schnabel H. 1997. Assessing normal embryogenesis in Caenorhabditis elegans using a 4D microscope: variability of development and regional specification. Developmental Biology 184(2):234-265.

Segref A, Cabello J, Clucas C, Schnabel R, Johnstone IL. 2010. Fate specification and tissue-specific cell cycle control of the Caenorhabditis elegans intestine. Molecular Biology of the Cell 21(5):725-738.

Shetty P, Lo MC, Robertson SM, Lin R. 2005. C. elegans TCF protein, POP-1, converts from repressor to activator as a result of Wnt-induced lowering of nuclear levels. Developmental Biology 285(2):584-592.

Sommermann EM, Strohmaier KR, Maduro MF, Rothman JH. 2010. Endoderm development in Caenorhabditis elegans: the synergistic action of ELT-2 and -7 mediates the specification-->differentiation transition. Developmental Biology 347(1):154-166.

Staller MV, Fowlkes CC, Bragdon MD, Wunderlich Z, Estrada J, DePace AH. 2015. A gene expression atlas of a bicoid-depleted Drosophila embryo reveals early canalization of cell fate. Development 142(3):587-596.

Stein LD, Bao Z, Blasiar D, Blumenthal T, Brent MR, Chen N, Chinwalla A, Clarke L, Clee C, Coghlan A, Coulson A, D'Eustachio P, Fitch DH, Fulton LA, Fulton RE, Griffiths-Jones S, Harris TW, Hillier LW, Kamath R, Kuwabara PE, Mardis ER, Marra MA, Miner TL, Minx P, Mullikin JC, Plumb RW, Rogers J, Schein JE, Sohrmann M, Spieth J, Stajich JE, Wei C, Willey D, Wilson RK, Durbin R, Waterston RH. 2003. The Genome Sequence of Caenorhabditis briggsae: A Platform for Comparative Genomics. PLoS Biology 1(2):E45.

Sulston JE, Horvitz HR. 1977. Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. Developmental Biology 56(1):110-156.

Sulston JE, Schierenberg E, White JG, Thomson JN. 1983. The embryonic cell lineage of the nematode Caenorhabditis elegans. Developmental Biology 100(1):64-119.

Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD. 2014. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. Cell 159(3):635-646.

Thorpe CJ, Schlesinger A, Carter JC, Bowerman B. 1997. Wnt signaling polarizes an early C. elegans blastomere to distinguish endoderm from mesoderm. Cell 90(4):695-705.

True JR, Haag ES. 2001. Developmental system drift and flexibility in evolutionary trajectories. Evolution & Development 3(2):109-119.

Vu V, Verster AJ, Schertzberg M, Chuluunbaatar T, Spensley M, Pajkic D, Hart GT, Moffat J, Fraser AG. 2015. Natural Variation in Gene Expression Modulates the Severity of Mutant Phenotypes. Cell 162(2):391-402.

Waaijers S, Boxem M. 2014. Engineering the Caenorhabditis elegans genome with CRISPR/Cas9. Methods 68(3):381-388.

Waddington CH. 1942. Canalization of development and the inheritance of acquired characters. Nature 150:563-565.

Walton T, Preston E, Nair G, Zacharias AL, Raj A, Murray JI. 2015. The Bicoid class homeodomain factors ceh-36/OTX and unc-30/PITX cooperate in C. elegans embryonic progenitor cells to regulate robust development. PLoS Genetics 11(3):e1005003.

Wilson MA, Hoch RV, Ashcroft NR, Kosinski ME, Golden A. 1999. A Caenorhabditis elegans wee1 homolog is expressed in a temporally and spatially restricted pattern during embryonic development. Biochimica et Biophysica Acta 1445(1):99-109.

Zhao Z, Boyle TJ, Bao Z, Murray JI, Mericle B, Waterston RH. 2008. Comparative analysis of embryonic cell lineage between Caenorhabditis briggsae and Caenorhabditis elegans. Developmental Biology 314(1):93-99.

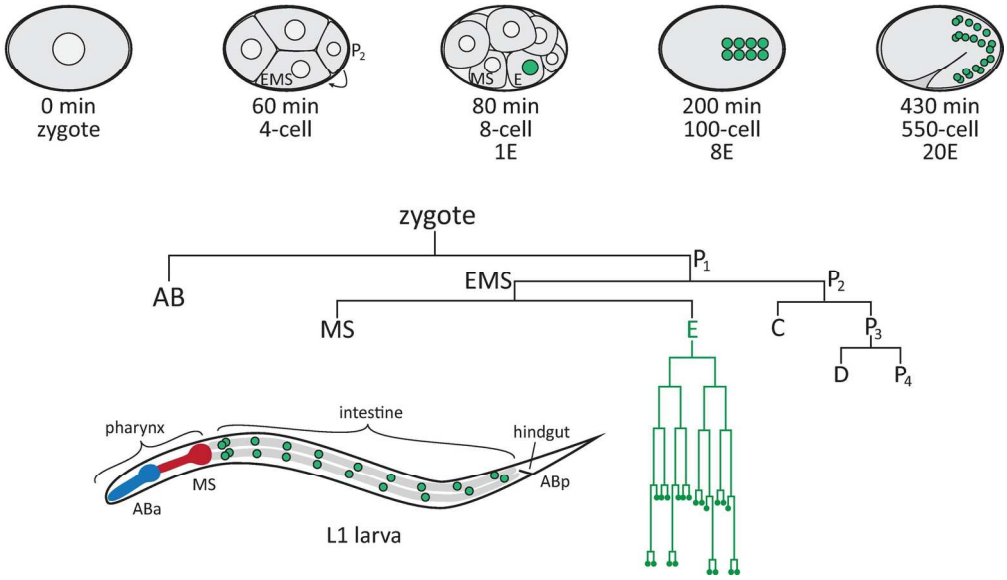
Zhu J, Hill RJ, Heid PJ, Fukuyama M, Sugimoto A, Priess JR, Rothman JH. 1997. end-1 encodes an apparent GATA factor that specifies the endoderm precursor in Caenorhabditis elegans embryos. Genes & Development 11(21):2883-2896.

Table 1. Examples of loss of gene functions that affect robustness of gut specification

Genotype	% of embryos making gut[1]	reference(s)
Weak/no effect		
pop-1	100%	(Lin et al. 1995)
pal-1	100%	(Maduro et al. 2005b)
pop-1; pal-1	100%	(Maduro et al. 2005b)
med-1	100%	(Maduro et al. 2007)
med-2	100%	(Maduro et al. 2007)
end-1	100%	(Maduro et al. 2005a)
elt-7	100%	(Sommermann et al. 2010)
end-3	95%	(Maduro et al. 2005a)
Intermediate effect		
med-1; med-2	50-85%	(Goszczynski and McGhee 2005; Maduro et al. 2015)
med-1; end-3	42%	(Maduro et al. 2007)
Strong effect		
skn-1	20-30%	(Bowerman et al. 1992)
end-3; pop-1	14%	(Maduro et al. 2005a)
mom-2	12%	(Rocheleau et al. 1997; Thorpe et al. 1997)
med-1; med-2; pal-1	7%	(Maduro et al. 2005b)
med-1; med-2; pop-1	6%	(Maduro et al. 2005b)
end-3; mom-2	0%	(Maduro et al. 2005a)
skn-1; mom-2	0%	(Thorpe et al. 1997)
skn-1; pop-1; pal-1	0%	(Maduro et al. 2005b)
end-1 end-3	0%	(Owraghi et al. 2010)

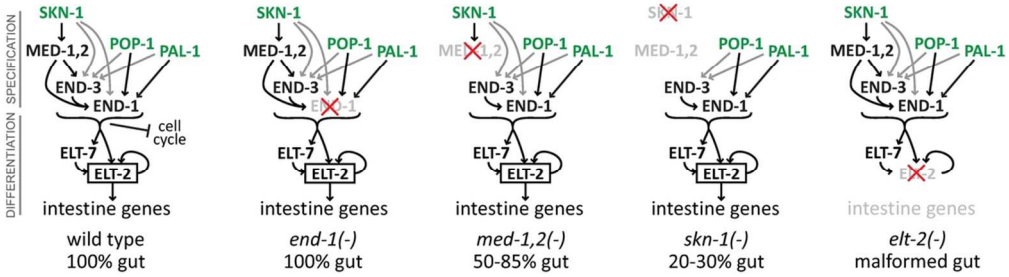
[1]Generally scored by presence of birefringent gut granules (e.g. (Bowerman et al. 1992)).

Figure 1



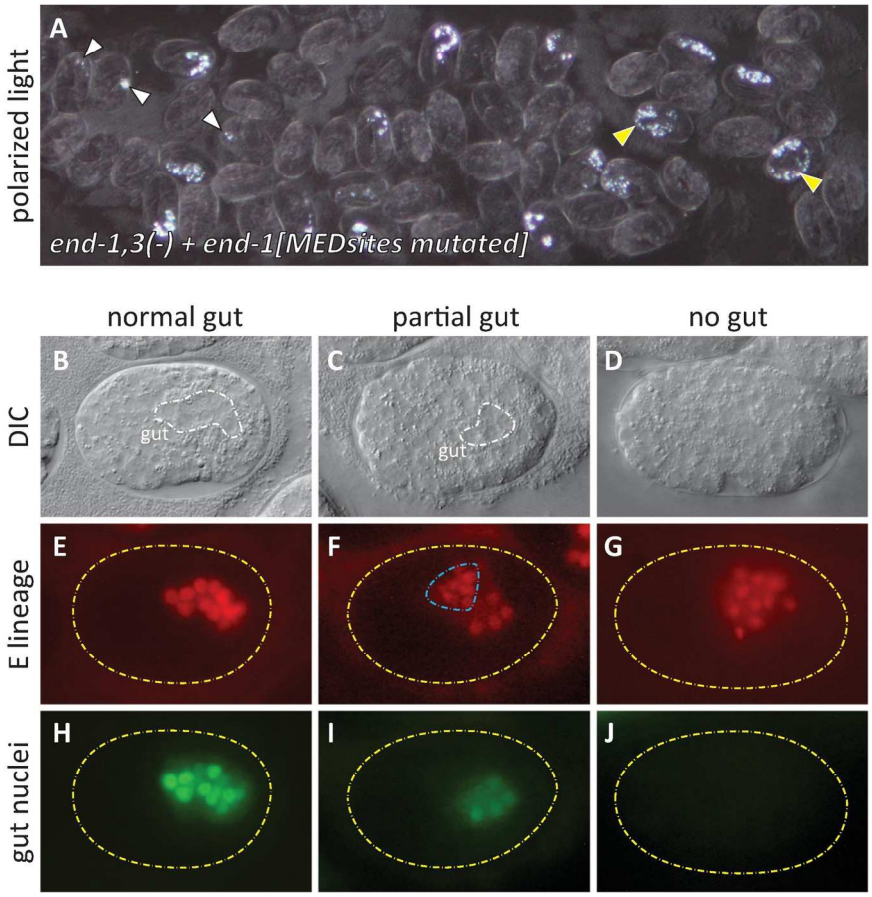
Embryonic origin of the *C. elegans* intestine. Top: Diagrams of embryos at various stages of development. Approximate times are for development at 25°C. Anterior is to the left and dorsal is up. Nuclei of cells in the endoderm lineage are shown as green circles when the E lineage comprises 1, 8 and 20 cells (1E, 8E and 20E, respectively). After 430 minutes, embryonic divisions within the gut are complete. The embryo hatches at approximately 12 h after fertilization. Bottom: Abbreviated cell lineage showing origin of the founder cells AB, MS, E, C, D and P4, after Sulston et al. (1983). A vertical line is a cell and a horizontal line is a cell division. The founder cells undergo further divisions. The pattern that generates the 20 cells of the intestine is shown in green. A diagram of a larva shows the intestine along with other components of the digestive tract and the blastomere of origin. ABa is the anterior daughter of AB and ABp is the posterior daughter. The egg and larva are approximately 50 µm and 200 µm long, respectively. Diagrams were adapted from WormAtlas (http://wormatlas.org/) and other sources (Maduro and Rothman 2002; Sulston et al. 1983).
159x144mm (300 x 300 DPI)

Figure 2



Simplified diagram of the Endoderm Gene Regulatory Network (GRN) in <I>C. elegans. Green font indicates factors that are contributed maternally, while factors in black are zygotic. Most interactions are known to be direct: SKN-1 to *med-1,2* (Maduro et al. 2001); MED-1,2 to *end-1* and *end-3* (Broitman-Maduro et al. 2005; Maduro et al. 2015); POP-1 on end-1 and end-3 (Robertson et al. 2014; Shetty et al. 2005); and PAL-1 on *end-1* and *end-3* (Maduro et al. 2015). Other interactions are implied genetically, including END-3 on *end-1* (Maduro et al. 2007) and SKN-1 on *end-1* and *end-3* (Maduro et al. 2005b). There is also some evidence for a maternal contribution of MED-1,2 (Maduro et al. 2007). Left to right: The wild-type network; network missing END-1 but which results in normal intestine (Boeck et al. 2011; Maduro et al. 2005a); loss of robust specification in the absence of MED-1,2 input into the ENDs (Maduro et al. 2015); loss of most gut when SKN-1 is absent, which also removes MED-1,2 (Bowerman et al. 1992; Maduro et al. 2001); and a malformed intestine when ELT-2 is absent (Fukushige et al. 1998).
99x51mm (300 x 300 DPI)

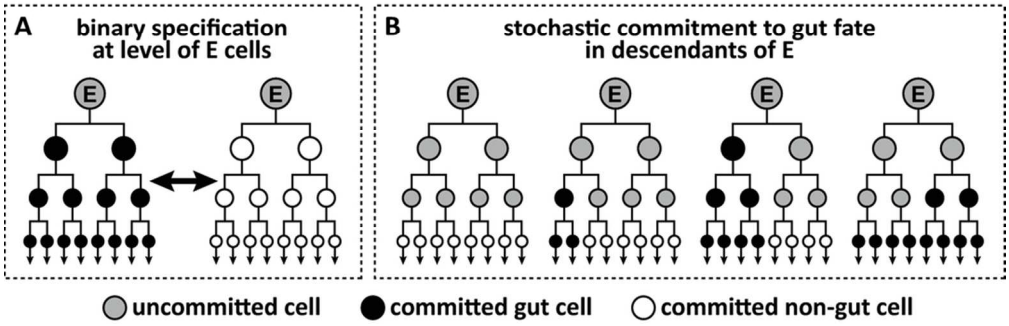
Figure 3



Evidence for partial specification of gut when the endoderm GRN is perturbed at the level of *end-1,3*. (A) Polarized light showing gut granules (white spots) in a collection of isogenic *end-1,3(-)* embryos partially rescued by an end-1 transgene in which the MED binding sites have been mutated (Maduro et al. 2015). Most embryos lack gut granules, some have a near-normal amount (yellow arrowheads), and others have very small amounts (white arrowheads). (B-J) Examples of embryos in which all, some or none of the E descendants have adopted a gut fate. (B, C, D) Differential Interference Contrast (DIC) microscopy images of embryos at 8E-12E stage. The cells forming the embryonic intestine are outlined in white dashed lines in B and C. (E, F, G) Stacks of images showing a nuclear-localized *end-3*::mCherry transcriptional reporter in the descendants of the E cell. Nuclei adopting a non-gut identity are outlined by a dashed cyan line. The reporter was made previously (Murray et al. 2008). Because all upstream regulators of *end-3* are present, the reporter is activated in the early E lineage even when no gut is made. (H, I, J) Expression of a nuclear-localized *elt-2*::GFP reporter shows nuclei of cells that have adopted a gut fate. In (H), all E descendants express *elt-2*::GFP; in (I), only some descendants do; and in (J), none of the E descendants express *elt-*

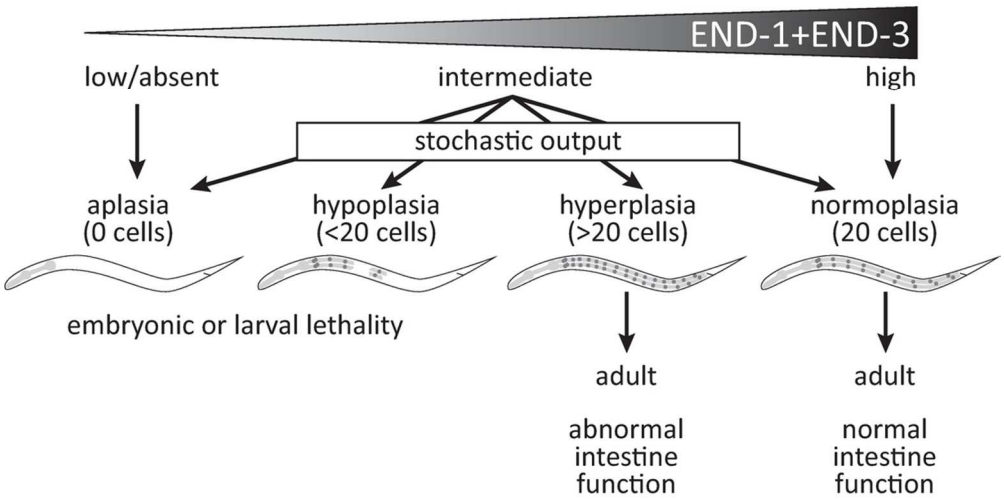
2::GFP. In (E-J) the eggshell is outlined in yellow. A *C. elegans* embryo is approximately 50 μm long. In (B-J), anterior is to the left.
138x191mm (300 x 300 DPI)

Figure 4



Models for how gut specification can fail. (A) Binary (all-or-none) model in which gut specification either occurs or it does not. (B) A schematic diagram of how "partial" specification leads to fewer gut nuclei due to a later commitment to gut fate by some E descendants and not others. For simplicity, the same pattern of cell divisions is maintained across the various examples.
91x43mm (300 x 300 DPI)

Figure 5



Relationship of the activity of END-1+END-3 with gut specification and organogenesis of the intestine. At high levels, specification and development of the intestine are robust. At intermediate levels, development becomes stochastic, and animals can experience the absence of gut (aplasia), changes in gut cell number (hypo- or hyperplasia), or no change (normoplasia). Animals with compromised END-1+END-3 can survive hyperplasia, although they exhibit defects in metabolism. Modified from a prior work (Maduro et al. 2015).
103x75mm (300 x 300 DPI)