Endomesoderm specification in *Caenorhabditis elegans* and other nematodes

Morris F. Maduro

Summary

The endomesoderm gene regulatory network (GRN) of *C. elegans* is a rich resource for studying the properties of cell-fate-specification pathways. This GRN contains both cell-autonomous and cell non-autonomous mechanisms, includes network motifs found in other GRNs, and ties maternal factors to terminal differentiation genes through a regulatory cascade. In most cases, upstream regulators and their direct downstream targets are known. With the availability of resources to study close and distant relatives of *C. elegans*, the molecular evolution of this network can now be examined. Within *Caenorhabditis*, components of the endomesoderm GRN are well conserved. A cursory examination of the preliminary genome sequences of two parasitic nematodes, *Haemonchus contortus* and *Brugia malayi*, suggests that evolution in this GRN is occurring most rapidly for the zygotic genes that specify blastomere identity. *BioEssays* 28:1010–1022, 2006. © 2006 Wiley Periodicals, Inc.

Introduction

Gene regulatory networks embody the regulatory interactions that guide differential gene expression and development in metazoans.[1] While models of GRNs should in principle consider all regulatory mechanisms, in practice they are concerned primarily with transcription factors and intercellular signaling. Work performed over the last 15 years has elucidated the mechanisms that specify the daughters of the *C. elegans* endomesoderm precursor EMS. The result is a detailed GRN that links the earliest-acting maternal factors, through a series of regulators, to terminal differentiation genes. This review will summarize the salient features of the

Department of Biology, University of California, Riverside, Riverside, CA 92521 E-mail: mmaduro@citrus.ucr.edu

Abbreviations: GRN, gene regulatory network; RNAi, RNA-mediated interference; EST, expressed sequence tag.

C. elegans endomesoderm GRN, compare its properties to GRNs elucidated in other systems, and examine its evolution within the nematode phylum.

Origin of the *C. elegans* endomesoderm

The entirety of *C. elegans* embryogenesis occurs in about 14 hours within a chitinous eggshell 50 μm long.[2] The point of sperm entry determines the posterior pole, leading to the asymmetric segregation of determinants in the first and subsequent mitotic divisions.[3–6] The zygote divides several times to produce the six founder cells AB, MS, E, C, D and P_4, each of which generates a specific subset of tissues as summarized in Fig. 1. The first division produces the anterior founder cell AB and its sister cell P_1. Division of P_1 produces the blastomere P_2 (which gives rise to the C, D and P_4 founder cells) and EMS, which will divide to produce MS and E. The E cell generates the entire endoderm, which consists of only 20 cells at hatching, while its sister MS generates 80 primarily mesodermal cells including cells of the posterior half of the feeding organ (the pharynx), one third of the animal's body wall muscles and the somatic gonad.[2] As EMS is a cell that divides to produce endoderm and mesoderm, it can be considered to be an endomesoderm (or mesendoderm) precursor, analogous to the endomesoderm tissue territory of deuterostomes.[1,7]

Overview of the *C. elegans* endomesoderm GRN

A "process diagram" incorporating the salient features of the *C. elegans* endomesoderm GRN is shown in Fig. 2. Specification of MS and E results from two parallel processes that are initiated by maternal gene products. The first is the initiation of a lineage-autonomous gene cascade that specifies EMS daughters as endomesodermal, while the second is a cell–cell interaction that assigns the mesoderm fate to MS and the endoderm fate to E. Downstream of these maternal pathways, embryonic genes generally belong to one of three broad categories as summarized in Fig. 3: (1) transcription factors that are expressed transiently, assigning identity to early blastomeres, (2) transcription factors that specify cells at the tissue or organ level, and are expressed for the duration of the life span, and (3) transcription factors that are activated in differentiated cells. In the following sections, genes whose
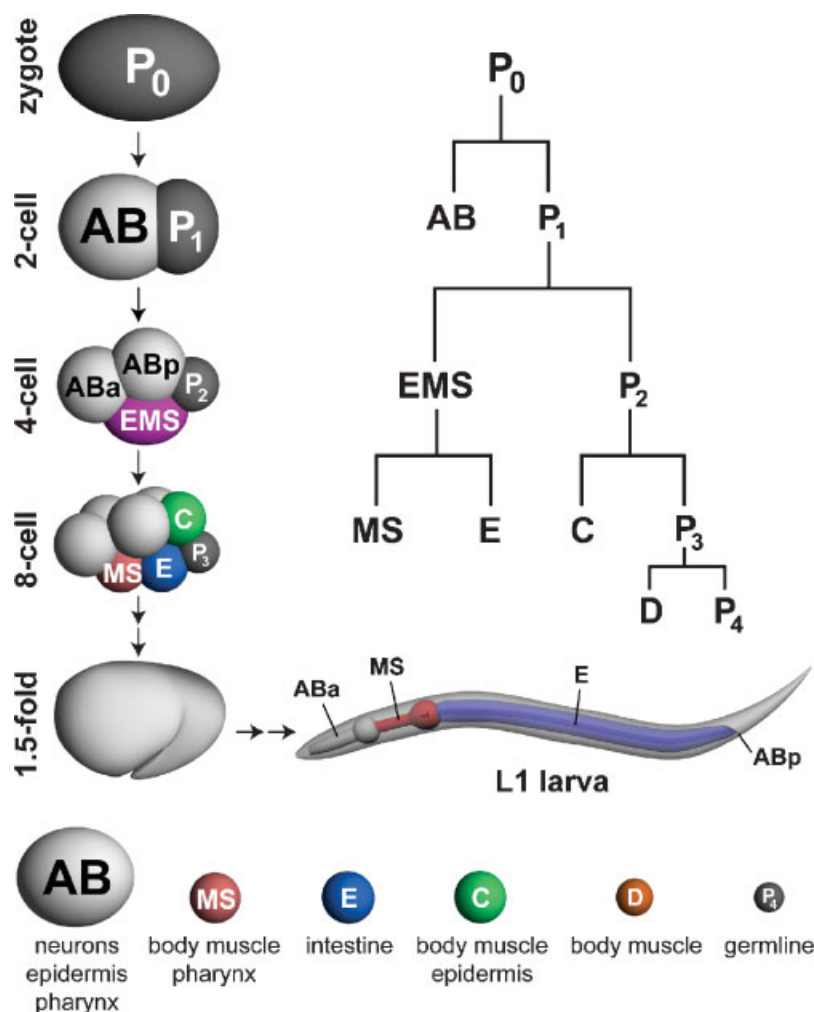
Figure 1. The early *C. elegans* lineage showing the origin of, and major tissue types produced by, the six founder cells.[2] The lineal origins of the digestive tract (anterior and posterior pharynx, intestine and rectum) are shown on a simplified representation of the first-stage juvenile, or L1 larva. For simplicity, additional tissues generated by a minority of AB, MS and C descendants are not indicated.

products function in *C. elegans* endomesoderm specification will be examined in detail. A summary of the genes and their functions is presented in Table 1.

Maternal specification of MS and E

EMS is specified by SKN-1

The gene at the top of the endomesodermal GRN is the maternal gene *skn-1*, which encodes a bZIP/homeodomain transcription factor.[8] Loss of *skn-1* results in the mis-specification of MS in all, and E in the majority ($\sim$70%), of embryos[8] (Figs 4,5D–F). In such embryos, the mis-specified cells produce body muscle and hypodermis, tissues made by the C blastomere, for which a GRN has been recently described.[9] *skn-1* mRNAs are present in oocytes and are translated in a subset of embryonic cells after fertilization,[10]

such that SKN-1 protein is present at high levels in the nuclei of EMS and P_2 at the 4-cell stage.[8,11] The Caudal-like regulator PAL-1, which specifies the C blastomere, is also present in EMS and P_2 at this stage.[12] Like the normal C cell, specification of the ectopic C-like cells produced in *skn-1* mutant embryos requires PAL-1.[12] Therefore, while SKN-1 and PAL-1 coexist in EMS in normal embryos, the MS- and E-promoting functions of SKN-1 are able to override the C-promoting function of PAL-1.[12]

In addition to lacking MS-derived pharynx, *skn-1*(-) embryos also lack anterior, ABa-derived pharynx.[8] The MS cell expresses a Delta-like ligand that allows MS to induce the formation of anterior pharynx from descendants of ABa through the GLP-1/Notch receptor,[13,14] as indicated in the lower left panel of Fig. 2. The E cell is also capable of this induction event, although only MS is in the correct position to
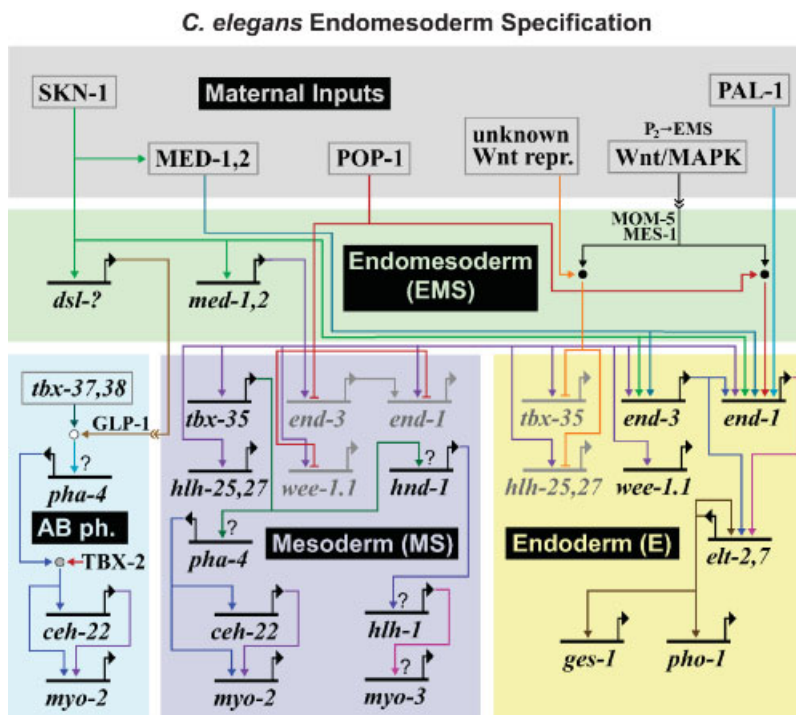
Figure 2. An endomesoderm specification GRN for the early *C. elegans* embryo patterned after the output style generated by BioTapestry.[1] Most other genes known to be activated in the early EMS lineage[32,33,35] are not included, as their roles in cell-fate specification are not known, although *hlh-25/27* and *wee-1.1* have been included as examples.[32] These other genes may function in other processes,[49] such as cell–cell signaling,[13,98] cell movement[80] or timing of the cell cycle.[99,100] Specification of anterior pharynx (AB ph.) is also shown, as this event results from an induction by MS and the activation of similar sets of pharynx-determining genes.[57,64] A horizontal line denotes the *cis*-regulatory region of a gene, with a bent arrow indicating the site of initiation of transcription. An arrow connected to this bent arrow represents the product of the gene. Capital letters denote gene products. Grey text indicates genes that are repressed in the domains indicated. A question mark indicates that the indicated temporal hierarchy is known, but a direct regulatory connection has not been established. Other abbreviations and symbols: $P_2 \rightarrow$ EMS, the cell–cell interaction that initiates the Wnt/MAPK signaling cascade in EMS; *dsl-?*, gene(s) encoding one or more ligands for the GLP-1/Notch receptor;[16] repr., repressor; arrow head inserted in arrow tail, intercellular signaling; gray circle, integration of TBX-2 and PHA-4 input to specify ABa-derived pharynx muscle;[57] black circles, interaction of Wnt/MAPK signal transduction with a nuclear effector;[36] white circle, integration of transcription factor activity and transduction of a GLP-1-dependent signal.[64]

do so in normal embryos.[14,15] The identity of the ligand remains unknown, though it is likely to be one or more of the DSL (Delta/Serrate/Lag) family members encoded in the *C. elegans* genome.[16]

Maternal genes that restrict or modulate SKN-1 activity

Two CCCH zinc finger proteins, MEX-1 and PIE-1, and the glycogen synthase kinase 3 homolog, GSK-3, restrict the activity of SKN-1 to EMS. Loss of *mex-1* function results in the ectopic accumulation of SKN-1 in the AB lineage and the transformation of the AB granddaughters into four MS-like cells (Fig. 4).[11,17] MEX-1 is associated with P-granules, cytoplasmic structures found in germline blastomeres (P_0 through P_4 and the P_4 daughters), suggesting that MEX-1 functions indirectly to prevent SKN-1 accumulation in the AB

lineage.[18] Loss of *pie-1* function results in ectopic specification of SKN-1-dependent fates in the daughters of P_2, causing C to adopt an MS-like fate and P_3 to adopt an E-like fate[17] (Fig. 4). PIE-1 protein is present in the cytoplasm and nuclei of P lineage cells, where it acts to repress transcription.[19] Hence, SKN-1 is prevented from activating EMS development in P_2 because of PIE-1 repression. Finally, loss of *gsk-3* function results in a *skn-1*-dependent transformation of C to an EMS-like cell.[20] Postembryonically, SKN-1 is phosphorylated by GSK-3 in the intestine,[21] suggesting that GSK-3 blocks residual SKN-1 function in C by promoting its degradation.

Embryos lacking maternal *pos-1* or *spn-4* function have many defects, including a failure to specify MS and E.[22,23] Like loss of *skn-1*, both mutants mis-specify MS and E as C-like precursors, although accumulation of SKN-1 protein is not
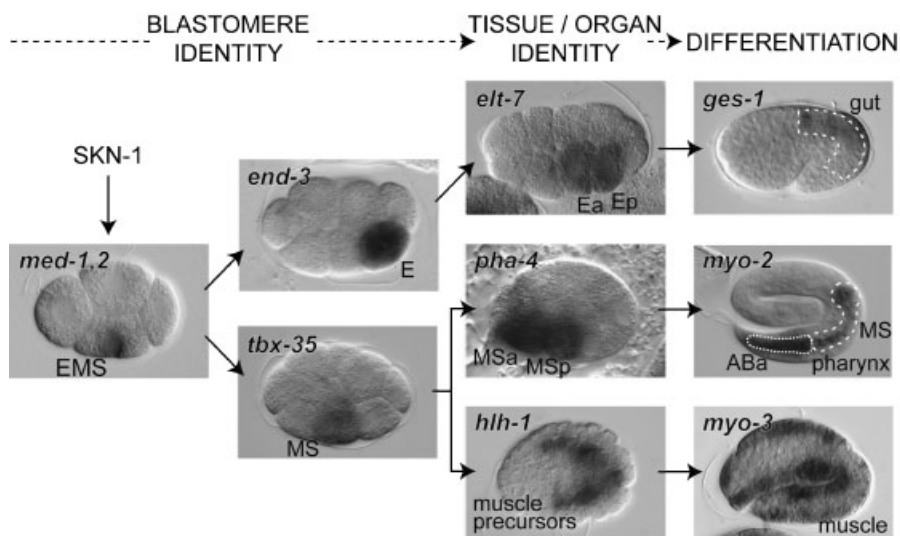
Figure 3. Expression of zygotic *C. elegans* genes detected by in situ hybridization demonstrates three hierarchical classes in the endomesoderm genetic pathway. The *med*s, *end*s and *tbx-35* encode blastomere identity factors, and *elt-7* and *pha-4* encode tissue/organ identity factors as described in the text. *ges-1* encodes a gut-specific esterase,[101] and *myo-2* and *myo-3* encode pharynx-specific and body muscle-specific myosin isoforms, respectively.[102] *C. elegans* embryos are approximately 50μm long.

affected. POS-1 contains two CCCH-type zinc fingers, and SPN-4 contains an RNA-recognition motif.[22–24] Both proteins interact with the 3′UTR of *glp-1* mRNA, suggesting that they regulate multiple maternal mRNAs to provide a permissive, rather than instructive, role in endomesoderm specification.[25]

Zygotic specification of MS and E

MED-1 and MED-2 are targets of SKN-1

SKN-1 activates the genes *med-1* and *med-2* in EMS (Fig. 3).[20] *med-1* and *med-2* encode unlinked, nearly identical, GATA-type transcription factors, and both genes contain a cluster of SKN-1 sites (core sequence RTCAT) in their promoters.[20,26] Recombinant SKN-1 is able to bind these sites in vitro[20] while, in vivo, expression of *skn-1* is both necessary and sufficient for the activation of a *med-1*::GFP reporter.[20] RNAi targeted to both *med* genes results in the production of arrested embryos, all of which lack MS-derived pharynx and body muscle, and approximately 50% of which also lack differentiated endoderm (Fig. 5G–I).[20] As with loss of *skn-1*, mis-specified MS and E cells adopt a C-like fate. Overexpression of *med-1* results in the ectopic accumulation of pharynx and intestine, consistent with the ability of the *med*s to specify MS and E identity. It was proposed, therefore, that the MEDs act downstream of SKN-1 to promote MS and E fates.[20] The partially penetrant endoderm defect (50%) was ascribed to either direct activation of endoderm specification by SKN-1 (i.e. independent of *med-1,2*) or to reduced efficacy of RNAi.

The simple maternal *skn-1* to zygotic *med-1,2* model was challenged by a recent study that showed that homozygous *med-1(-)*; *med-2(-)* embryos were found to be able to specify endoderm 83–100% of the time.[27] A significant maternal contribution of the *med*s was subsequently found to exist that is able to rescue specification of E, but not MS, in these *med-1,2(-)* embryos.[28] Like zygotic expression of *med-1,2*, maternal expression requires the activity of *skn-1* in the mother. When *med-1,2(-)* embryos were produced from mothers in which the maternal contribution of the *med*s was depleted, endoderm specification was found to fail 50% of the time, similar to the results obtained by RNAi. The *med*s, therefore, are expressed both maternally and zygotically.[28] The increased proportion of *skn-1(-)* embryos that fail to make endoderm as compared with *med-1,2(-)* (70% versus 50%) suggests that SKN-1 can activate endoderm specification, at least some of the time, independently of the *med*s.[29]

Specification of E identity by END-1 and END-3

The paralogous genes *end-1* and *end-3* specify E identity.[30] *end-1* and *end-3* are ~30 kbp apart and encode GATA transcription factors that share 48% similarity, consistent with their origin as an ancient duplication.[30] Overexpression of *end-1* or *end-3* can reprogram early embryonic blastomeres into E-like cells, and RNAi targeted to both genes simultaneously results in a strong endoderm defect (Fig. 5M–O) and transformation of E to C.[30] Both *end* promoters interact with MED-1 in vivo and in vitro, confirming direct activation of *end-1,3* by the MEDs.[31,32] Activation of the *end*s is restricted

Table 1. *C. elegans* endomesoderm genes and their functions

C. elegans Gene(s)	Product[a]	Primary function in endomesoderm	Candidate Orthologs[b]	
			Hc	Bm
Maternal genes				
skn-1	bZIP/homeodomain TF[8]	primary activator of endomesoderm, directly activates *med-1,2*	yes	no
pal-1	homeodomain TF[12]	weak positive input into E specification	no	yes
mex-1	CCCH zinc finger[17]	prevents translation of maternal *skn-1* mRNA in the AB lineage	yes	yes
pos-1	CCCH zinc finger[23]	required for SKN-1 function in EMS	yes	yes
spn-4	RNA-binding motif[22]	required for MED-1,2 function in MS,E	yes	yes
glp-1	Notch-like receptor[13]	required for MS induction of ABa-derived pharynx	yes	yes
pie-1	CCCH zinc finger[17]	blocks SKN-1 activity in P_2	yes	yes
gsk-3	glycogen synthase kinase[20]	blocks SKN-1 activity in C	yes	yes
pop-1	TCF/LEF-like TF[15]	blocks *end-1,3* activation in MS, weak positive input into E specification	yes	yes
wrm-1	divergent β-catenin[44]	required for modification of POP-1 in E	no	yes?
Blastomere identity genes				
med-1,2	divergent GATA TFs[20]	specify MS and E fates, activate *end-1,3* and *tbx-35*	no	no
end-1	GATA TF[30]	specifies the E fate	no	yes
end-3	GATA TF[30]	specifies the E fate	no	yes
tbx-35	Tbx TF[36]	specifies the MS fate	no	no
Tissue/organ identity genes				
elt-2	GATA TF[48]	specifies endodermal tissue identity	yes[c]	yes
elt-7	GATA TF[49]	overlapping function with *elt-2*?	no	yes
pha-4	FoxA TF[54]	specifies pharynx identity	yes	yes
hnd-1	HAND bHLH TF[60]	specification of muscle prescursors?	no	no
hlh-1	MyoD-like bHLH TF[59]	specifies muscle identity	no	yes
ceh-22	Nkx2.5 TF[101]	specifies pharynx muscle	yes	yes
Differentiation genes				
myo-2	myosin heavy chain[103]	component of pharynx muscle	yes	yes
myo-3	myosin heavy chain[103]	component of body wall muscle	yes	yes
ges-1	gut esterase[101]	digestive enzyme	yes	yes
pho-1	acid phosphatase[50]	digestive enzyme	yes	yes

[a]TF, transcription factor.
[b]Identification of putative orthologous gene segments in Haemonchus contortus (Hc) or Brugia malayi (Bm) were made by performing a TBLASTN search[104] of the preliminary genome sequence data available at http://www.sanger.ac.uk/cgi-bin/blast/submitblast/h_contortus and http://tigrblast.tigr.org/er-blast/index.cgi?project=bma1, respectively, using C. elegans protein sequences indexed in WormBase (http://www.wormbase.org, release WS160). A candidate ortholog is indicated as 'yes' if the best BLAST alignment had a summed probability P(N) of less than 10^{-3}, 'yes?' if $10^{-3} \leq P(N) \leq 10^{-2}$, or 'no' otherwise.
[c]An elt-2 orthologue in H. contortus was recently described.[105]

to E by the activity of TCF/POP-1, which directly represses the *end*s in MS[31] (see below).

Differences in the end genes

While similar in structure and function, three lines of evidence suggest that the *end* genes are not completely redundant. First, mutants of *end-1* have no apparent phenotype, while a weak defect (5–10% lacking endoderm) is seen in *end-3* mutants.[30] Second, *end-3* transcripts are readily detected early in the E cell cycle (Fig. 3), while those for *end-1* are detected more readily in the E daughters (Ea and Ep).[28] Transcriptional profiling of staged embryos has also detected this temporal difference,[9,33] which suggests that *end-1* and *end-3* respond to slightly different *cis*-regulatory inputs. Third, *end-1* mRNAs accumulate to reduced levels in *end-3(-)* embryos, suggesting that END-3 contributes to *end-1* activa-

tion.[28] Therefore, although *end-1* and *end-3* likely arose by gene duplication, they have undergone changes in their regulation.

Specification of MS identity by TBX-35

The binding sites for MED-1 on the *end* genes were determined biochemically.[32] Unexpectedly, MED-1 was found to recognize a consensus binding site, RAGTATAC, that does not match the GATA factor consensus sequence (HGATAR) expected from the amino acid sequence of its C_4-type zinc finger.[34] A search of the *C. elegans* genome sequence for promoters containing MED site clusters identified 21 genes, the majority of which showed expression in the early MS and/or E lineages by reporter transgene.[32] One of these putative target genes, *tbx-35*, showed an MS-restricted expression pattern (Fig. 3). Another study found *tbx-35* among
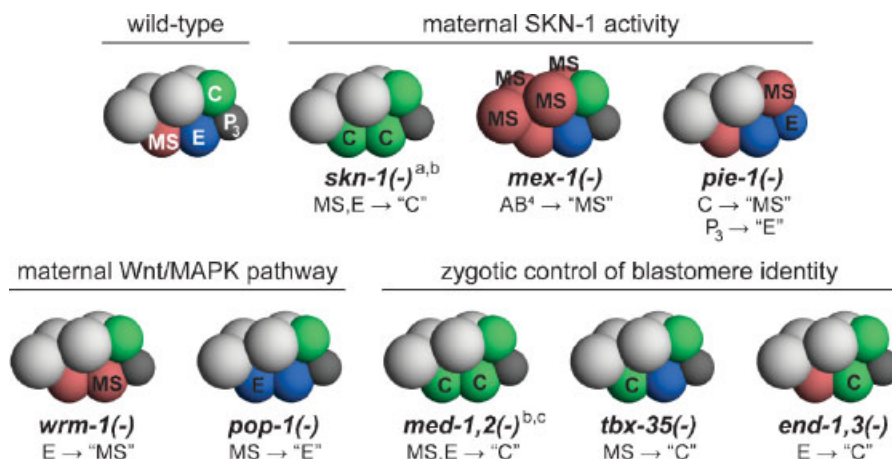
Figure 4. Mis-specification of 8-cell-stage blastomeres in mutant backgrounds. Transformed blastomeres are indicated with black lettering. For cases where different phenotypes are possible, the strongest phenotype is shown. [a]In all *skn-1(-)* embryos, pharynx produced by ABa descendants is also absent. [b]E is specified correctly in some embryos. [c]The *med*s are also expressed maternally.[28]

a larger set of SKN-1-dependent zygotic (*sdz*) genes identified by global transcription profiling.[35]

Embryos homozygous for a *tbx-35* null mutation lack most MS-derived tissues, including pharynx (Fig. 5J–L), and instead show an excess of PAL-1-dependent muscle similar to those made by the C cell.[12,36] Consistent with an apparent MS-to-C transformation, 30% of *tbx-35(-)* embryos express zygotic *pal-1*, a marker for the early C lineage, in both MS and C descendants.[9,36] When *tbx-35* is overexpressed throughout the early embryo, large amounts of pharynx and muscle cells are made.[36] Therefore, TBX-35 is both necessary and sufficient to specify MS identity, fulfilling a function similar to END-1,3 in the E cell.

The maternal Wnt/MAPK pathway: making E different from MS

POP-1: A regulatory switch
E specification requires a cell–cell interaction between P_2 and EMS at the 4-cell stage.[37] In isolation, EMS divides symmetrically to produce two MS-like cells.[37] The components of the P_2–EMS signal are an overlapping Wnt/MAPK signal, with additional signaling through a Src tyrosine kinase pathway.[38–40] Depletion of these components has the same effect as preventing the P_2–EMS interaction, namely a transformation of E into an MS-like cell (shown for the divergent β-catenin gene *wrm-1* in Fig. 4). Only one gene, *pop-1*, is known to show the opposite cell fate change when depleted, namely transformation of MS into E (Fig. 4), which results from derepression of *end-1,3* in MS.[15,31] Loss of *pop-1* is epistatic to mutations in the upstream components, suggesting it is the terminal regulator downstream of the

P_2-EMS interaction.[38] *pop-1* encodes an HMG-domain transcription factor similar to the TCF/LEF family of Wnt effectors.[15] POP-1 is widely expressed and its activity is required for correct specification of many cells that are the product of asymmetric cell divisions along the anterior–posterior axis.[41] While TCF/LEF proteins are known to be important for activating Wnt target genes through interaction with β-catenins (canonical Wnt signaling),[42] the predominant role for POP-1 is as a repressor of endoderm fate in MS.[15] In the absence of Wnt/MAPK signaling, POP-1 recruits a conserved repressor complex to repress the *end*s.[30,31,43] Wnt/MAPK signaling therefore blocks the *end*-repressive function of POP-1 in E, permitting activation of *end-1,3* and specification of an E fate.[31,44,45]

POP-1 does not appear to be the sole mediator of the MS/E regulatory switch, however. In *pop-1(-)* embryos, MS-specific activation of *tbx-35* still occurs even though MS adopts an E-like fate.[36] As *tbx-35* is activated in both E and MS in a Wnt/MAPK(-) background, at least one other cofactor acts, perhaps as a Wnt-dependent repressor in E, to restrict activation of *tbx-35* to MS.[36] The identity of this cofactor is not known.

Positive contribution of PAL-1 and POP-1 to endoderm
Recently, POP-1 and PAL-1 have been shown to function as endoderm activators. Depletion of *pop-1* results in a decrease in *end-1* activation levels,[46] and in a *skn-1(-)* background, loss of *pop-1* increases the proportion of embryos lacking endoderm from 70% to 85%.[29] These results demonstrate a positive, canonical role for POP-1 in E.[29,46] Unexpectedly, the source of the residual gut specification in *skn-1(-)*; *pop-1(-)*
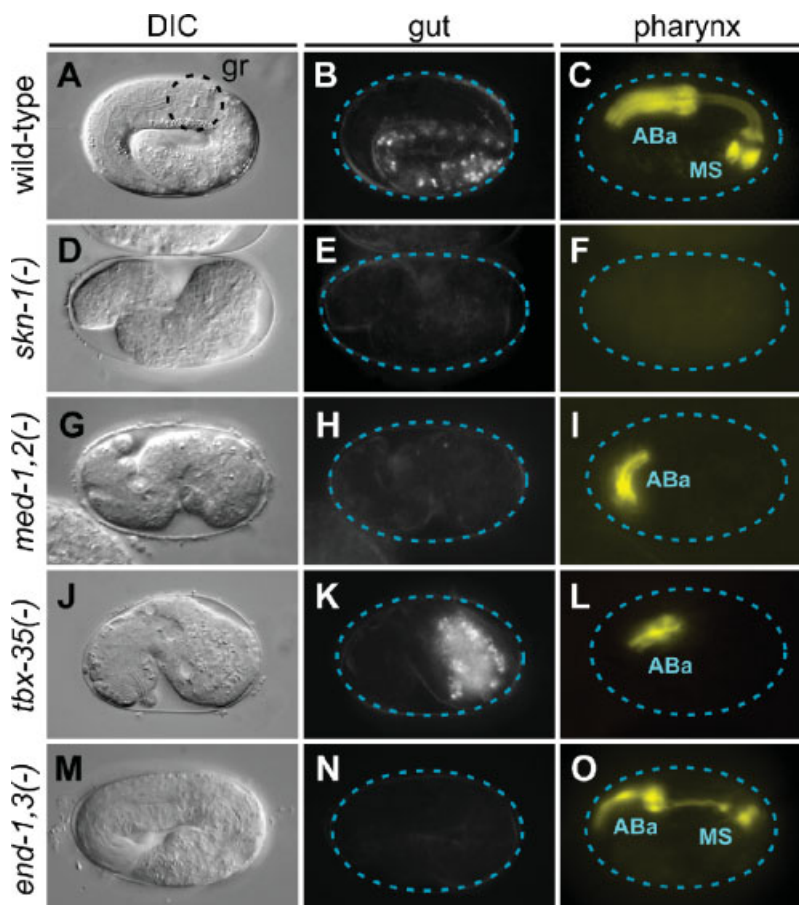
Figure 5. Appearance of wild-type and mutant *C. elegans* embryos. Left column, DIC optics; middle column, birefringent gut granules viewed under polarized light; right column, expression of a *ceh-22*::GFP reporter transgene, which marks pharynx muscle cells[101] (pseudocolored yellow). **A–C:** Wild-type embryo at 3-fold stage, just prior to hatching, with prominent grinder (gr) in MS-derived pharynx. **D–F:** Absence of both pharynx and gut in a *skn-1(-)* embryo. Approximately 30% of such embryos contain gut.[8] **G–I:** Absence of posterior pharynx and gut in a *med-1,2(-)* embryo. Approximately 50% of such embryos contain gut.[20,28] **J–L:** Absence of posterior pharynx in a *tbx-35(-)* embryo.[20,36] **M–O:** Absence of gut in an *end-1,3(-)* embryo.[30] For *skn-1* and *med-1,2* mutant embryos, the strongest phenotype (i.e. embryos showing concomitant loss of endoderm) is shown. Abbreviations: ABa, anterior pharynx; MS, posterior pharynx.

embryos was found to be the C specification gene *pal-1*, as depletion of *skn-1, pop-1* and *pal-1* together resulted in the complete elimination of gut.[29] Hence, zygotic E specification results from multiple, parallel inputs (Fig. 2).

Moving downstream of blastomere fates

Tissue/organ identity genes

Shortly after gastrulation, embryonic development in *C. elegans* undergoes a shift from specification of blastomere identity to specification of tissue/organ identity.[47] Unlike the transient expression shown by the *med*s or *end*s, such genes are generally expressed throughout the remainder of development and adulthood, where they generally work in combination with other cell-type-specific regulators to activate terminal differentiation genes.

ELT-2 establishes intestine identity

Activation of the GATA factor gene *elt-2*, and its paralog *elt-7*, begins in the two E daughters and is subsequently maintained by autoregulation (Figs 2,3).[48,49] ELT-2 appears to be the primary intestinal identity factor: *elt-2(-)* animals produce defective intestine and undergo arrest shortly after hatching, while ectopic overexpression of *elt-2* is sufficient to specify cells as endodermal.[48] ELT-2 directly binds GATA sites in the promoters of the intestine-specific genes *pho-1* and *ges-1*.[50,51] Enrichment of intestine-specific mRNAs identified ∼1750 genes, about half of which contain a GATA site.[52] Hence, it is likely that ELT-2 activates a battery of hundreds of

intestine-specific genes. Like *elt-2*, overexpression of *elt-7* can specify cells as endodermal, suggesting that *elt-2* and *elt-7* share many of the same transcriptional targets.[49] Over-expression of *end-1* or *end-3* results in ectopic activation of both *elt-2* and *elt-7*, while *end-1,3(-)* embryos fail to activate these genes.[30,31,53] The simplest interpretation of these results is that END-1,3 activate *elt-2* and *elt-7* to specify intestinal fate.[30]

Specification of MS-derived tissues

MS descendants include progenitors of distinctly different cell types, such as pharynx and body muscle.[2] Hence, different tissue/organ identity genes must be activated in the early MS lineage. One of these, the pharynx-specific transcription factor FoxA/PHA-4, is at the top of an extensive GRN that regulates pharynx development in descendants of both MS and ABa.[54] *pha-4* transcripts are first detected in the MS daughters (Fig. 3), suggesting that *pha-4* is a direct target of TBX-35. In specifying development of various tissue types present within the pharynx, PHA-4 works with other cell-type-specific regulators. For example, PHA-4 and Nkx2.5/CEH-22 activate the pharynx muscle gene *myo-2*.[55,56] For production of ABa-derived pharynx muscle, there is an additional requirement for the Tbx factor TBX-2.[57]

The MyoD homolog HLH-1 is a muscle-identity factor.[58] Loss of *hlh-1* leads to muscle defects and paralysis,[58] while ectopically expressed *hlh-1* is sufficient to respecify cells as muscle progenitors.[59] The related gene *hnd-1* is expressed in myogenic precursors in the early MS, C and D lineages, upstream of *hlh-1*, suggesting that *hnd-1* and *hlh-1* work together to promote muscle development.[9,60]

While mesodermal tissue/organ factors such as PHA-4 and HLH-1 are known, the regulatory interactions that link MS specification to their activation are only implied (Fig. 2). Hence, current research is aimed at identification of putative *cis*-regulatory sites through which TBX-35 (and other putative early MS factors[32,35]) might act on *pha-4* and *hlh-1*, and elucidation of the possible role of Wnt/MAPK signaling in distinguishing pharynx from muscle precursors.

Conserved use of regulators in other animals

What features of the *C. elegans* endomesoderm GRN, if any, are conserved with other metazoans? To begin with, many transcription factors used in the *C. elegans* endomesoderm have orthologs with similar roles in other animals. For example, the vertebrate myogenic regulatory factor MyoD specifies muscle fates,[61] similar to the *C. elegans* muscle factor HLH-1,[9,59] suggesting their function has been evolutionarily conserved. In *Drosophila* and vertebrates, Tbx regulators function in heart development.[62,63] In *C. elegans*, specification of MS-derived pharynx requires TBX-35,[36] while production of ABa-derived pharynx requires TBX-37 and

TBX-38.[36,64] The *C. elegans* pharynx is a muscular, contractile organ similar to the heart. Consistent with a potential homology of these organs, vertebrate Nkx2.5 is able to functionally replace its *C. elegans* ortholog, the pharynx-specific gene *ceh-22*.[65]

The use of GATA factors to pattern endomesoderm is similarly conserved. In *Drosophila*, transient expression of the GATA factor *serpent* specifies endoderm progenitors, while the downstream regulator dGATAe activates genes that function in differentiated endoderm,[66] reminiscent of the *C. elegans* END-1,3 → ELT-2,7 GATA factor cascade. When *end-1* is expressed in *Xenopus* animal caps, they are respecified as endoderm,[67] consistent with the existence of a conserved subclass of "endoderm" GATA factors. Finally, the use of GATA factors and Tbx genes together also occurs in zebrafish, where *gata5* and the Tbx gene *no tail* specify the endomesoderm territory prior to gastrulation.[68]

While these examples are striking, the vastly different ways in which the body plans of protostomes and deuterostomes are established suggests that the similar roles of these regulators are the result of convergent evolution, rather than homology. It has been proposed that, in the protostome-deuterostome ancestor, regulators such as MyoD or a prototypical GATA factor may have had a more direct role in activating targets in particular tissues.[69] Over evolutionary time, such factors would be promoted up a regulatory hierarchy as different morphogenetic programs were added in by intercalation.[69] Hence, across different phyla, similar use of regulators for early progenitor specification may be indicative of their ancestral function as tissue-specific differentiation factors in the last common bilaterian ancestor.

Conserved network architecture

What generalizations can be made from the study of the *C. elegans* endomesoderm GRN? A systematic study of the types of gene interactions that occur in yeast GRNs identified several "network motifs" that occur at high frequency, suggesting that they confer a strong selective advantage.[70] The *C. elegans* endomesoderm GRN contains examples of six types of network motif (Fig. 6). For example, specification of MS occurs by a linear cascade of factors that form a "regulatory chain." Such motifs are thought to be the simplest means by which gene activation events are ordered in a temporal sequence.[70] Temporal waves of gene expression are observed in genome-wide transcriptome analyses of the early *C. elegans* embryo, suggesting that this type of motif is widely used.[9,33] More complex types of circuits, such as the "feed-forward" and "multi-input" motifs, are also found when the genes that specify endoderm are considered (Fig. 6). These same network motifs have also been seen in the GRN controlling mouse erythroid lineage specification.[71] Hence, the *C. elegans* endomesoderm network embodies features of GRNs found in diverse organisms, although
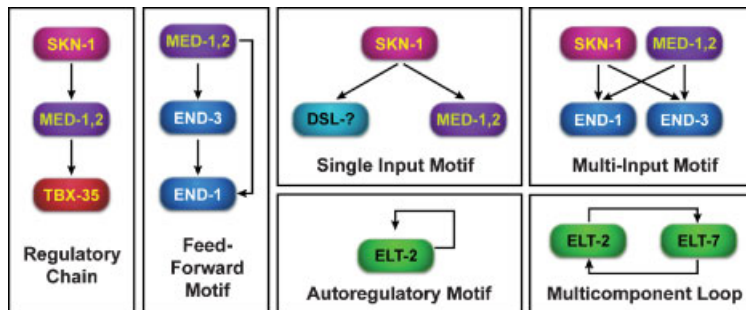
Figure 6. Regulatory motifs identified in other GRNs[70] can be found in portions of the *C. elegans* endomesoderm GRN. An arrow indicates that the product of one gene activates expression of another gene.

connections among specific types of regulators are generally not conserved.

Comparison of the *C. elegans* and sea urchin endomesoderm GRNs

Can the *C. elegans* GRN be used to study how gene networks evolve? One way to begin to make generalizations about GRN evolution is to compare similar networks in extant systems. Here, the natural comparison is between the *C. elegans* endomesoderm GRN and that of the sea urchin *S. purpuratus*.[72] Both networks presumably originated from a common GRN that existed in the protostome–deuterostome ancestor over 600 million years ago.[73] At first glance, and perhaps not unexpectedly, both networks make use of multiple maternally contributed components, which activate transient expression of intermediate regulators, leading to stable expression of downstream factors. Both organisms establish an "endomesoderm" tissue territory that subdivides into mesoderm and endoderm. In sea urchins, endomesoderm consists of many cells while, in *C. elegans*, this has been reduced to a single cell, EMS. As noted earlier, an endomesoderm territory is also seen in vertebrates, suggesting it may be an ancient germ layer unto itself.[7]

Given the evolutionary separation of nematodes and echinoderms, and considering the notion that the common use of regulators might very well be convergent, are there any deeper similarities that can be extracted? It would be expected that maternal components, being most directly tied to reproductive lifestyle and developmental speed, would be highly divergent. For example, *C. elegans* might rely more heavily on a large dowry of maternal gene products to direct rapid, asymmetric cleavages. Indeed, no similarity between the maternal components used in *C. elegans* or *S. purpuratus* is apparent. Looking downstream, many of the same types of embryonic factors are found, such as components of the widely used Notch and Wnt signaling pathways. One can also find examples of what appear to be conserved regulatory

interactions. For example, in *S. purpuratus*, *GataE* acts upstream of *Brachyury* and *FoxA*.[1] In *C. elegans*, a similar regulatory hierarchy is found among *med-1,2*, *tbx-35* and *pha-4*, which encode respective orthologs of these factors.[20,36,54] However, additional inputs regulate all three genes in *S. purpuratus*, including the orthodenticle-related factor Otx, whose orthologs in *C. elegans* function in the nervous system.[74] Moreover, the MEDs represent highly diverged members of the GATA factor family.[32] Hence, finding conserved substructure, or so-called GRN kernels,[75] may be a function of how one defines "conserved".

As stated above, however, the *C. elegans* endomesoderm GRN as shown in Fig. 2 is largely incomplete. This is primarily because a systematic study of the interactions among all known *C. elegans* endomesoderm genes has not been performed. The simplicity of some regulatory interactions also implies that this network is incomplete. For example, some genes (e.g. *med-1,2* or *ceh-22*) have only a single regulatory input. In general, *cis*-regulatory modules integrate inputs of multiple types to produce a regulatory output that is novel, and more precise in space and time, as compared to the inputs themselves.[76] Hence, deeper similarities between the *C. elegans* and *S. purpuratus* endomesoderm GRNs may emerge as more is learned in *C. elegans*.

Evolution of the endomesoderm GRN in nematodes

Within echinoderms, inter-species comparisons have revealed conserved features of the sea urchin endomesoderm network, including conserved GRN kernels and *cis*-regulatory sites.[77] Similar studies within nematodes, therefore, may identify aspects of the *C. elegans* endomesoderm GRN that are flexible, and those that have been conserved, over evolutionary time.

The current molecular phylogeny of nematodes is divided into five broad clades.[78] Clade V includes Rhabditoidea (rhabditids), to which *C. elegans* belongs, and a vertebrate parasite taxon,

Strongyloidea.[79] Within this clade, studies of post-embryonic development in the rhabditids *Oscheius* and *Pristionchus* have revealed surprising flexibility in the signaling mechanisms used to establish post-embryonic cell fates when compared with *C. elegans*.[80,81] Observational studies have shown that the clade V nematodes *Pellioditis* and *Halicephalobus*, and the clade IV pine-wood nematode *Bursaphelenchus*, undergo very similar embryogenesis, including the establishment of similar founder cells as in *C. elegans*.[82,83] Clade V *Rhabditis* species even require a P$_2$-EMS induction for endoderm to be made, similar to *C. elegans*.[84]

Despite the appearance of similar early embryonic development, however, underlying programs of cell-fate specification may be very different. In the clade IV nematode *Acrobeloides*, where a similar set of founder cells is also established, laser ablation of some early blastomeres results in compensation by the remaining cells, resulting in a nearly complete larva.[85] This is strikingly different from the mosaic nature of *C. elegans* and *Rhabditis* embryogenesis, in which lost cells are not replaced.[2,84]

Are the main players conserved?

One straightforward question to ask when comparing underlying developmental mechanisms is whether the same components even exist in a related species. Although identification of an interspecific ortholog does not guarantee that the gene performs the same function,[86] it can provide a starting point for further study. Two close relatives of *C. elegans* that are used for comparative molecular genetics are the hermaphroditic species *C. briggsae* and the closely related male–female species *C. remanei*.[79] *C. briggsae* and *C. remanei* are estimated to have diverged from *C. elegans* approximately 80–110 million years ago.[87,88] The recent availability of genome sequence information, and the ability to perform RNA interference in both species, should allow very rapid progress to be made with comparative developmental studies.[79,87,89]

Searches of the *C. briggsae* and *C. remanei* genome sequences have identified at least one clear ortholog of every *C. elegans* endomesoderm component shown in Table 1 (indexed in WormBase, http://www.wormbase.org, release WS160). The *end* and *med* genes appear to be undergoing frequent duplications: *end-3* is present as two nearly identical paralogs in *C. briggsae* (*Cb-end-3.1* and *Cb-end-3.2*), while the *med*s are represented by four paralogs in *C. briggsae* and seven in *C. remanei*.[90] *Cr-med-2* and *Cb-med-4* contain incomplete coding regions and may thus be nonfunctional,[90] similar to *elt-4*, a partial duplication of *elt-2* in *C. elegans*.[91] Gene duplications are somewhat rarer among the maternal components, as only *pos-1* appears to be present as two paralogous genes in *C. briggsae*.[92]

Many of the *med* and *end* orthologs have been tested for conserved function and expression by introducing them into *C. elegans* as transgenes. Two of the *Cb-med* and five of seven *Cr-med* genes can fully complement the embryonic lethal phenotype of *C. elegans med-1,2(-)* embryos,[90] and *Cb-end-1* and *Cb-end-3.1* can restore endoderm specification to *C. elegans end-1,3(-)* embryos.[30] Expression of reporters from the other species is also similar in *C. elegans*, suggesting that their *cis*-regulatory sites respond to similar activators.[30,90] Indeed, *Cr-med-1*::GFP expression in *C. elegans* was found to be dependent on *skn-1*, similar to endogenous *Ce-med-1,2*.[90] Therefore, among the *med*s and *end*s, the same regulatory hierarchy seems to be conserved within *Caenorhabditis*.

While more informative than expression data alone, cross-species rescue experiments still do not allow a firm conclusion about the function of an orthologous gene in its native context.[86] Here, classical genetics or reverse genetics approaches such as RNAi can be instructive. This has worked with the *end* genes in *C. briggsae*: RNAi of *Cb-end-1* or the *Cb-end-3.1,3.2* pair does not result in detectable phenotypes, but the triple *Cb-end-1,3.1,3.2(RNAi)* results in loss of endoderm in 96% of embryos, consistent with conservation of *end* function and redundancy.[30] However, RNAi knockdown of the *Cb-med* genes does not seem to produce a phenotype (Katy Lin and M.M., unpublished observations). To fully resolve function in related species, therefore, it may be necessary to isolate chromosomal mutants as has been done to study evolution of sex determination.[93]

High divergence of early zygotic regulators outside Caenorhabditis?

As a means by which to assess the extent of conservation outside of the rhabditids, the preliminary genome sequences of the vertebrate parasites *Haemonchus contortus* (clade V) and *Brugia malayi* (clade III) were searched for candidate orthologs of the *C. elegans* endomesoderm GRN components (Table 1). *H. contortus* diverged from *Caenorhabditis* approximately 400 million years ago,[94] which suggests that *B. malayi* is even more divergent.[78] While candidate homologs of most of the maternal and terminal differentiation genes seem to be present in *H. contortus* and *B. malayi*, many of the early zygotic regulators appear to be absent (Table 1). As these genome sequences are preliminary, the ability to identify candidate orthologs is necessarily affected by such factors as average intron size and sequence coverage. However, a similar difficulty in identifying sequence homologs for the early zygotic endomesoderm genes is encountered in the clade V nematode *Pristionchus pacificus*.[95] Thus, while the *med*s and *end*s are well-conserved within *Caenorhabditis*, these regulators might be highly diverged, if they are present at all, in more distant nematodes. Indeed, no MED-like GATA factors are known in any animal outside of *Caenorhabditis*, suggesting that they were a recent innovation.[32]

We might expect that higher constraints exist for early-acting maternal and late-acting organ identity genes as compared with blastomere identity genes. The maternal endomesoderm genes in *C. elegans* all have additional functions outside of the early endomesoderm GRN. For example, SKN-1 functions in the fully developed intestine in response to oxidative stress,[96] while POP-1 is important for many asymmetric cell divisions throughout development.[41] Late-acting regulators that activate tissue-specific 'gene batteries' are also expected to be well-conserved,[75] as an organ identity gene like PHA-4 regulates hundreds of targets.[54] In contrast to the multiple functions of the maternal regulators, and the complex network properties of tissue identity factors, early zygotic genes have comparatively simple functions: they regulate a small number of downstream targets and display a short window of expression in a small number of cells. Thus, more evolutionary flexibility may be found with these genes, as fewer steps may be required to transfer their functions to another regulator.

Based on studies of early blastomere specification in *Enoplus* (clade II), *Acrobeloides* (clade IV) and *C. elegans* (clade V), Schierenberg proposed that differences in embryonic cell specification mechanisms among distant nematodes might be maximal just prior to gastrulation,[97] the time window in which the *med*s and *end*s act. The strong conservation of the zygotic cell-fate-specification genes within *Caenorhabditis* might therefore reflect evolutionary constraints imposed by the rapid early embryonic development seen in this genus.

Conclusions

The *C. elegans* endomesoderm GRN comprises a relatively small number of factors organized into a hierarchy that links together maternal gene products, blastomere specification genes, tissue/organ identity genes and differentiation genes. While specific regulatory interactions are not conserved outside of nematodes, similar network motifs are found in diverse systems.

With most of the main players identified, interest now turns to using this GRN as a starting point for evolutionary studies among nematodes. Within *Caenorhabditis*, the genes that specify early blastomere identity in the endomesoderm GRN appear to be structurally and functionally conserved, although they appear to be undergoing frequent duplications. Outside the genus, these genes are the most divergent. It will be of interest to explore endomesoderm specification mechanisms outside of *Caenorhabditis* to assess how differences in the underlying GRNs might relate to developmental constraints imposed by the parasitic lifestyle of *H. contortus*, for example, or the rapid embryogenesis of *Caenorhabditis*. With new resources, such as forthcoming genome sequences from related nematodes and the ability to perform RNAi in these, elucidating GRNs in other species may progress relatively quickly. For example, if maternal and terminal regulators do turn out to be well-conserved outside of *Caenorhabditis*, it may be possible to use genomics approaches to build the homologous networks from the top-down or the bottom-up.

In the long term, an understanding of the endomesoderm GRNs in other nematodes will allow generalizations to be made about how the interaction between external forces influencing development and flexibility inherent in the genome drive evolutionary change in developmental GRNs.

References

1. Davidson EH, Rast JP, Oliveri P, Ransick A, Calestani C, et al. 2002. A provisional regulatory gene network for specification of endomesoderm in the sea urchin embryo. Dev Biol 246:162–190.
2. Sulston JE, Schierenberg E, White JG, Thomson JN. 1983. The embryonic cell lineage of the nematode Caenorhabditis elegans. Dev Biol 100:64–119.
3. Nance J. 2005. PAR proteins and the establishment of cell polarity during C. elegans development. Bioessays 27:126–135.
4. Goldstein B, Hird SN. 1996. Specification of the anteroposterior axis in Caenorhabditis elegans. Development 122:1467–1474.
5. Bowerman B, Shelton CA. 1999. Cell polarity in the early Caenorhabditis elegans embryo. Curr Opin Genet Dev 9:390–395.
6. Jenkins N, Saam JR, Mango SE. 2006. CYK-4/GAP Provides a Localized Cue to Initiate Anteroposterior Polarity upon Fertilization. Science.
7. Rodaway A, Patient R. 2001. Mesendoderm. an ancient germ layer? Cell 105:169–172.
8. 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 68:1061–1075.
9. Baugh LR, Hill AA, Claggett JM, Hill-Harfe K, Wen JC, et al. 2005. The homeodomain protein PAL-1 specifies a lineage-specific regulatory network in the C. elegans embryo. Development 132:1843–1854.
10. Seydoux G, Fire A. 1994. Soma-germline asymmetry in the distributions of embryonic RNAs in Caenorhabditis elegans. Development 120:2823–2834.
11. 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 74:443–452.
12. 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 87:217–226.
13. Priess JR, Schnabel H, Schnabel R. 1987. The glp-1 locus and cellular interactions in early C. elegans embryos. Cell 51:601–611.
14. Mango SE, Thorpe CJ, Martin PR, Chamberlain SH, Bowerman B. 1994. Two maternal genes, apx-1 and pie-1, are required to distinguish the fates of equivalent blastomeres in the early Caenorhabditis elegans embryo. Development 120:2305–2315.
15. 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 83:599–609.
16. Chen N, Greenwald I. 2004. The lateral signal for LIN-12/Notch in C. elegans vulval development comprises redundant secreted and transmembrane DSL proteins. Dev Cell 6:183–192.
17. Mello CC, Draper BW, Krause M, Weintraub H, Priess JR. 1992. The pie-1 and mex-1 genes and maternal control of blastomere identity in early C. elegans embryos. Cell 70:163–176.

18. Guedes S, Priess JR. 1997. The C. elegans MEX-1 protein is present in germline blastomeres and is a P granule component. Development 124:731–739.

19. Tenenhaus C, Schubert C, Seydoux G. 1998. Genetic requirements for PIE-1 localization and inhibition of gene expression in the embryonic germ lineage of Caenorhabditis elegans. Dev Biol 200: 212–224.

20. 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 7:475–485.

21. An JH, Vranas K, Lucke M, Inoue H, Hisamoto N, et al. 2005. Regulation of the Caenorhabditis elegans oxidative stress defense protein SKN-1 by glycogen synthase kinase-3. Proc Natl Acad Sci USA 102:16275–16280.

22. Gomes JE, Encalada SE, Swan KA, Shelton CA, Carter JC, et al. 2001. The maternal gene spn-4 encodes a predicted RRM protein required for mitotic spindle orientation and cell fate patterning in early C. elegans embryos. Development 128:4301–4314.

23. Tabara H, Hill RJ, Mello CC, Priess JR, Kohara Y. 1999. pos-1 encodes a cytoplasmic zinc-finger protein essential for germline specification in C. elegans. Development 126:1–11.

24. Hall TM. 2005. Multiple modes of RNA recognition by zinc finger proteins. Curr Opin Struct Biol 15:367–373.

25. Ogura K, Kishimoto N, Mitani S, Gengyo-Ando K, Kohara Y. 2003. Translational control of maternal glp-1 mRNA by POS-1 and its interacting protein SPN-4 in Caenorhabditis elegans. Development 130:2495–2503.

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

27. 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 171:545–555.

28. Maduro M, Broitman-Maduro G, Mengarelli I, Rothman J. 2006. Maternal deployment of the embryonic SKN-1→MED-1,2 cell specification pathway in C. elegans. Dev Biol accepted.

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

30. Maduro M, Hill RJ, Heid PJ, Newman-Smith ED, Zhu J, et al. 2005. Genetic redundancy in endoderm specification within the genus Caenorhabditis. Dev Biol 284:509–522.

31. 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. Dev Biol 248:128–142.

32. 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. Dev Cell 8:427–433.

33. Baugh LR, Hill AA, Slonim DK, Brown EL, Hunter CP. 2003. Composition and dynamics of the Caenorhabditis elegans early embryonic transcriptome. Development 130:889–900.

34. Lowry JA, Atchley WR. 2000. Molecular evolution of the GATA family of transcription factors: conservation within the DNA-binding domain. J Mol Evol 50:103–115.

35. Robertson SM, Shetty P, Lin R. 2004. Identification of lineage-specific zygotic transcripts in early Caenorhabditis elegans embryos. Dev Biol 276:493–507.

36. Broitman-Maduro G, Lin KT-H, Hung W, Maduro M. 2006. Specification of the C. elegans MS blastomere by the T-box factor TBX-35. Development 133:3097–3106.

37. Goldstein B. 1992. Induction of gut in Caenorhabditis elegans embryos. Nature 357:255–257.

38. Rocheleau CE, Downs WD, Lin R, Wittmann C, Bei Y, et al. 1997. Wnt signaling and an APC-related gene specify endoderm in early C. elegans embryos. Cell 90:707–716.

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

40. Bei Y, Hogan J, Berkowitz LA, Soto M, Rocheleau CE, et al. 2002. SRC-1 and Wnt signaling act together to specify endoderm and to control cleavage orientation in early C. elegans embryos. Dev Cell 3:113–125.

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

42. Korswagen HC. 2002. Canonical and non-canonical Wnt signaling pathways in Caenorhabditis elegans: variations on a common signaling theme. Bioessays 24:801–810.

43. Calvo D, Victor M, Gay F, Sui G, Luke MP, et al. 2001. A POP-1 repressor complex restricts inappropriate cell type-specific gene transcription during Caenorhabditis elegans embryogenesis. Embo J 20:7197–7208.

44. Rocheleau CE, Yasuda J, Shin TH, Lin R, Sawa H, et al. 1999. WRM-1 activates the LIT-1 protein kinase to transduce anterior/posterior polarity signals in C. elegans. Cell 97:717–726.

45. Meneghini MD, Ishitani T, Carter JC, Hisamoto N, Ninomiya-Tsuji J, et al. 1999. MAP kinase and Wnt pathways converge to downregulate an HMG-domain repressor in Caenorhabditis elegans. Nature 399:793–797.

46. 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. Dev Biol 285:584–592.

47. Labouesse M, Mango SE. 1999. Patterning the C. elegans embryo: moving beyond the cell lineage. Trends Genet 15:307–313.

48. Fukushige T, Hawkins MG, McGhee JD. 1998. The GATA-factor elt-2 is essential for formation of the Caenorhabditis elegans intestine. Dev Biol 198:286–302.

49. Maduro MF, Rothman JH. 2002. Making worm guts: the gene regulatory network of the Caenorhabditis elegans endoderm. Dev Biol 246:68–85.

50. Fukushige T, Goszczynski B, Yan J, McGhee JD. 2005. Transcriptional control and patterning of the pho-1 gene, an essential acid phosphatase expressed in the C. elegans intestine. Dev Biol 279:446–461.

51. Hawkins MG, McGhee JD. 1995. elt-2, a second GATA factor from the nematode Caenorhabditis elegans. J Biol Chem 270:14666–14671.

52. 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 133:287–295.

53. Zhu J, Fukushige T, McGhee JD, Rothman JH. 1998. Reprogramming of early embryonic blastomeres into endodermal progenitors by a Caenorhabditis elegans GATA factor. Genes Dev 12:3809–3814.

54. Gaudet J, Mango SE. 2002. Regulation of organogenesis by the Caenorhabditis elegans FoxA protein PHA-4. Science 295:821–825.

55. Vilimas T, Abraham A, Okkema PG. 2004. An early pharyngeal muscle enhancer from the Caenorhabditis elegans ceh-22 gene is targeted by the Forkhead factor PHA-4. Dev Biol 266:388–398.

56. Kalb JM, Lau KK, Goszczynski B, Fukushige T, Moons D, et al. 1998. pha-4 is Ce-fkh-1, a fork head/HNF-3alpha,beta,gamma homolog that functions in organogenesis of the C. elegans pharynx. Development 125: 2171–2180.

57. Roy Chowdhuri S, Crum T, Woollard A, Aslam S, Okkema PG. 2006. The T-box factor TBX-2 and the SUMO conjugating enzyme UBC-9 are required for ABa-derived pharyngeal muscle in C. elegans. Dev Biol.

58. Chen L, Krause M, Sepanski M, Fire A. 1994. The Caenorhabditis elegans MYOD homologue HLH-1 is essential for proper muscle function and complete morphogenesis. Development 120:1631–1641.

59. Fukushige T, Krause M. 2005. The myogenic potency of HLH-1 reveals wide-spread developmental plasticity in early C. elegans embryos. Development 132:1795–1805.

60. Mathies LD, Henderson ST, Kimble J. 2003. The C. elegans Hand gene controls embryogenesis and early gonadogenesis. Development 130:2881–2892.

61. Buckingham M. 2001. Skeletal muscle formation in vertebrates. Curr Opin Genet Dev 11:440–448.

62. Reim I, Frasch M. 2005. The Dorsocross T-box genes are key components of the regulatory network controlling early cardiogenesis in Drosophila. Development 132:4911–4925.

63. Reiter JF, Alexander J, Rodaway A, Yelon D, Patient R, et al. 1999. Gata5 is required for the development of the heart and endoderm in zebrafish. Genes Dev 13:2983–2995.

64. Good K, Ciosk R, Nance J, Neves A, Hill RJ, et al. 2004. The T-box transcription factors TBX-37 and TBX-38 link GLP-1/Notch signaling to

mesoderm induction in C. elegans embryos. Development 131:1967–1978.

65. Haun C, Alexander J, Stainier DY, Okkema PG. 1998. Rescue of Caenorhabditis elegans pharyngeal development by a vertebrate heart specification gene. Proc Natl Acad Sci USA 95:5072–5075.

66. Murakami R, Okumura T, Uchiyama H. 2005. GATA factors as key regulatory molecules in the development of Drosophila endoderm. Dev Growth Differ 47:581–589.

67. Shoichet SA, Malik TH, Rothman JH, Shivdasani RA. 2000. Action of the Caenorhabditis elegans GATA factor END-1 in Xenopus suggests that similar mechanisms initiate endoderm development in ecdysozoa and vertebrates. Proc Natl Acad Sci USA 97:4076–4081.

68. Rodaway A, Takeda H, Koshida S, Broadbent J, Price B, et al. 1999. Induction of the mesendoderm in the zebrafish germ ring by yolk cell-derived TGF-beta family signals and discrimination of mesoderm and endoderm by FGF. Development 126:3067–3078.

69. Erwin DH, Davidson EH. 2002. The last common bilaterian ancestor. Development 129:3021–3032.

70. Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, et al. 2002. Transcriptional regulatory networks in Saccharomyces cerevisiae. Science 298:799–804.

71. Swiers G, Patient R, Loose M. 2006. Genetic regulatory networks programming hematopoietic stem cells and erythroid lineage specification. Dev Biol.

72. Oliveri P, Davidson EH. 2004. Gene regulatory network controlling embryonic specification in the sea urchin. Curr Opin Genet Dev 14:351–360.

73. Chen JY, Bottjer DJ, Oliveri P, Dornbos SQ, Gao F, et al. 2004. Small bilaterian fossils from 40 to 55 million years before the cambrian. Science 305:218–222.

74. Lanjuin A, VanHoven MK, Bargmann CI, Thompson JK, Sengupta P. 2003. Otx/otd homeobox genes specify distinct sensory neuron identities in C. elegans. Dev Cell 5:621–633.

75. Davidson EH, Erwin DH. 2006. Gene regulatory networks and the evolution of animal body plans. Science 311:796–800.

76. Howard ML, Davidson EH. 2004. cis-Regulatory control circuits in development. Dev Biol 271:109–118.

77. Yuh CH, Brown CT, Livi CB, Rowen L, Clarke PJ, et al. 2002. Patchy interspecific sequence similarities efficiently identify positive cis-regulatory elements in the sea urchin. Dev Biol 246:148–161.

78. Blaxter ML, De Ley P, Garey JR, Liu LX, Scheldeman P, et al. 1998. A molecular evolutionary framework for the phylum Nematoda. Nature 392:71–75.

79. Mitreva M, Blaxter ML, Bird DM, McCarter JP. 2005. Comparative genomics of nematodes. Trends Genet 21:573–581.

80. Felix MA, Barriere A. 2005. Evolvability of cell specification mechanisms. J Exp Zoolog B Mol Dev Evol 304:536–547.

81. Sommer RJ. 2001. As good as they get: cells in nematode vulva development and evolution. Curr Opin Cell Biol 13:715–720.

82. Houthoofd W, Jacobsen K, Mertens C, Vangestel S, Coomans A, et al. 2003. Embryonic cell lineage of the marine nematode Pellioditis marina. Dev Biol 258:57–69.

83. Hasegawa K, Futai K, Miwa S, Miwa J. 2004. Early embryogenesis of the pinewood nematode Bursaphelenchus xylophilus. Dev Growth Differ 46:153–161.

84. Laugsch M, Schierenberg E. 2004. Differences in maternal supply and early development of closely related nematode species. Int J Dev Biol 48:655–662.

85. Wiegner O, Schierenberg E. 1999. Regulative development in a nematode embryo: a hierarchy of cell fate transformations. Dev Biol 215:1–12.

86. Haag ES, Pilgrim D. 2005. Harnessing Caenorhabditis Genomics for Evolutionary Developmental Biology. Curr Genomics 6:579–588.

87. Stein LD, Bao Z, Blasiar D, Blumenthal T, Brent MR, et al. 2003. The Genome Sequence of Caenorhabditis briggsae: A Platform for Comparative Genomics. PLoS Biol 1:E45.

88. Cho S, Jin SW, Cohen A, Ellis RE. 2004. A phylogeny of caenorhabditis reveals frequent loss of introns during nematode evolution. Genome Res 14:1207–1220.

89. Rudel D, Kimble J. 2001. Conservation of glp-1 regulation and function in nematodes. Genetics 157:639–654.

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

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

92. Konwerski J, Senchuk M, Petty E, Lahaie D, Schisa JA. 2005. Cloning and expression analysis of pos-1 in the nematodes Caenorhabditis briggsae and Caenorhabditis remanei. Dev Dyn 233:1006–1012.

93. Hill RC, de Carvalho CE, Salogiannis J, Schlager B, Pilgrim D, et al. 2006. Genetic flexibility in the convergent evolution of hermaphroditism in Caenorhabditis nematodes. Dev Cell 10:531–538.

94. Vanfleteren JR, Van de Peer Y, Blaxter ML, Tweedie SA, Trotman C, et al. 1994. Molecular genealogy of some nematode taxa as based on cytochrome c and globin amino acid sequences. Mol Phylogenet Evol 3:92–101.

95. Andres Pires da Silva, personal communication.

96. An JH, Blackwell TK. 2003. SKN-1 links C. elegans mesendodermal specification to a conserved oxidative stress response. Genes Dev 17:1882–1893.

97. Schierenberg E. 2001. Three sons of fortune: early embryogenesis, evolution and ecology of nematodes. Bioessays 23:841–847.

98. Park FD, Priess JR. 2003. Establishment of POP-1 asymmetry in early C. elegans embryos. Development 130:3547–3556.

99. Knight JK, Wood WB. 1998. Gastrulation initiation in Caenorhabditis elegans requires the function of gad-1, which encodes a protein with WD repeats. Dev Biol 198:253–265.

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

101. Marshall SD, McGhee JD. 2001. Coordination of ges-1 expression between the Caenorhabditis pharynx and intestine. Dev Biol 239:350–363.

102. Okkema PG, Fire A. 1994 . The Caenorhabditis elegans NK-2 class homeoprotein CEH-22 is involved in combinatorial activation of gene expression in pharyngeal muscle. Development 120:2175–2186.

103. Okkema PG, Harrison SW, Plunger V, Aryana A, Fire A. 1993. Sequence requirements for myosin gene expression and regulation in Caenorhabditis elegans. Genetics 135:385–404.

104. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410.

105. Couthier A, Smith J, McGarr P, Craig B, Gilleard JS. 2004. Ectopic expression of a Haemonchus contortus GATA transcription factor in Caenorhabditis elegans reveals conserved function in spite of extensive sequence divergence. Mol Biochem Parasitol 133:241–253.