Developmental Biology **246,** 68–85 (2002)
doi:10.1006/dbio.2002.0655, available online at http://www.idealibrary.com on IDEAL®

REVIEW

Making Worm Guts: The Gene Regulatory Network of the *Caenorhabditis elegans* Endoderm

Morris F. Maduro and Joel H. Rothman[1]

Department of MCD Biology and Neuroscience Research Institute, University of California, Santa Barbara, California 93106

The nematode *Caenorhabditis elegans* is a triploblastic ecdysozoan, which, although it contains too few cells during embryogenesis to create discernible germ "layers," deploys similar programs for germ layer differentiation used in animals with many more cells. The endoderm arises from a single progenitor, the E cell, and is selected from among three possible fates by a three-state combinatorial regulatory system involving intersecting cell-intrinsic and intercellular signals. The core gene regulatory cascade that drives endoderm development, extending from early maternal regulators to terminal differentiation genes, is characterized by activation of successive tiers of transcription factors, including a sequential cascade of redundant GATA transcription factors. Each tier is punctuated by a cell division, raising the possibility that intercession of one cell cycle round, or DNA replication, is required for activation of the next tier. The existence of each tier in the regulatory hierarchy is justified by the assignment of a unique task and each invariably performs at least two functions: to activate the regulators in the next tier and to perform one other activity distinct from that of the next tier. While the regulatory inputs that initiate endoderm development are highly divergent, they mobilize a gene regulatory network for endoderm development that appears to be common to all triploblastic metazoans. Genome-wide functional genomic approaches, including identification of >800 transcripts that exhibit the same regulatory patterns as a number of endoderm-specific genes, are contributing to elucidation of the complete endoderm gene regulatory network in *C. elegans*. Dissection of the architecture of the *C. elegans* endoderm network may provide insights into the evolutionary plasticity and origins of this germ layer.

Key Words: **gene network; endoderm; GATA factor; digestive tract.**

Introduction

In the process of developing from a relatively featureless zygote into a complex multicellular organism with diverse tissue types and functions, metazoan embryos have evolved strategies to successively restrict developmental potential and assign fates to subsets of cells. During early embryonic development of most metazoans, cells are organized into three germ layers, each endowed with the capacity to engender distinct tissue and organ types. Precursor cells in each germ layer must coordinate cell movements with adjoining germ layers during gastrulation, activate a germ layer-specific transcriptional cascade, and repress genes that function in other germ layers. Further, all of these events must be coordinated within the context of active cell division. The means by which distinct tissue layers are directed to undergo specific programs of morphogenesis and differentiation must therefore be precisely controlled, while at the same time remaining flexible to evolutionary change. A paradigm for how gene regulatory networks are coordinated to dictate a restricted program of differentiation can be obtained by elaborating the network controlling the development of an entire germ layer. Moreover, as the invention of the germ layers (first, ectoderm and endoderm of diploblasts, and later, addition of the mesoderm in triploblasts) marked major evolutionary transitions during metazoan evolution, an understanding of the gene regulatory networks underlying germ layer development will also contribute important insights into the events that created these large evolutionary steps.

Based on a few representative examples, it appears that

[1] To whom correspondence should be addressed. Fax: (805) 893-2005. E-mail: rothman@lifesci.ucsb.edu.

0012-1606/02 $35.00

the gene regulatory network controlling development of the innermost germ layer, the endoderm, is conserved across metazoan phylogeny. This is apparent even for an embryo in which too few cells are present at the time that the network becomes active for a "layer" to be evident. As recognized over 100 years ago, the endoderm germ "layer" in nematodes, members of the ecdysozoan clade of protostomes (Aguinaldo *et al.*, 1997), is established as a single cell, called E (Boveri, 1893, 1899). In the nematode *Caenorhabditis elegans*, this cell arises when there are only 7 cells in the embryo. The E cell undergoes no more than 5 rounds of division during embryogenesis, creating exclusively the 20 clonally derived cells of the juvenile intestine (Fig. 1) (Sulston *et al.*, 1983). The detailed morphogenetic events leading to assembly of the intestine into a coherent organ, and its anatomical structure, have been well described (Leung *et al.*, 1999). The regulatory program for endoderm-specific differentiation appears to be autonomous to the E cell: in the absence of all other embryonic cells, E can give rise to a full set of differentiated intestinal cells and even structural elements of a fully formed intestine (Priess and Thomson, 1987; Leung *et al.*, 1999).

In this review, we describe the emerging information regarding the gene regulatory network that directs development of the *C. elegans* endoderm, with emphasis on the cascade of early regulators. We describe the key regulatory factors constituting the backbone of the regulatory network that specifies the *C. elegans* endoderm and our initial understanding of the network through which they operate. The genome sequence (*C. elegans* Sequencing Consortium, 1998), a rapid reverse genetic method, RNA-mediated interference (RNAi; Fire *et al.*, 1998), procedures for detecting *in vivo* interactions between transcription factors and their targets in identified cells (e.g., Fukushige *et al.*, 1999), and the ready access to transcriptional profiling with DNA microarrays (e.g., Reinke *et al.*, 2000; Jiang *et al.*, 2000; Kim *et al.*, 2001), complement the developmental and genetic tools available for this organism and should make it possible to elaborate the entire regulatory network for endoderm-specific differentiation in this animal.

Specification of the Mesendoderm Progenitor, EMS, by a Maternal-to-Zygotic Switch

The events that establish the initial anteroposterior polarity of the *C. elegans* embryo and create differences between early descendant cells have been described in recent reviews (Bowerman, 1998; Bowerman and Shelton, 1999; Goldstein, 2000; Gotta and Ahringer, 2001).

The zygote divides asymmetrically into the two daughter cells, called AB and P_1, and embryonic polarity is manifested as the differential ability of these blastomeres to translate maternally provided mRNAs. P_1 divides to produce the mesendodermal precursor, EMS, and the germline/mesectodermal precursor, P_2 (Fig. 1A). The anterior daughter of EMS, called MS, gives rise to many mesodermal cell types, including body wall muscle and the posterior half of the feeding organ (the pharynx); the posterior daughter is the endodermal progenitor, E (Sulston *et al.*, 1983). In the past decade, the molecular mechanisms by which the identity of the EMS cell is specified and endoderm fate is subsequently restricted to the E blastomere have been elucidated (Table 1). Two maternal regulatory pathways specify the identity of EMS and its daughters. Both pathways function within EMS itself: the first activates the genes that specify both MS and E fates, while the second, which makes the E cell different from MS, is part of a reiterative switching system that directs daughters of asymmetric cell divisions to acquire different transcriptional states.

The first regulator of early blastomere fate in *C. elegans* to be identified was SKN-1 (Bowerman *et al.*, 1992), a composite bZIP/homeodomain transcription factor required maternally for EMS fates (Bowerman *et al.*, 1992; Blackwell *et al.*, 1994). While *skn-1* mRNA is contributed maternally and is found throughout all cells of the early embryo, SKN-1 protein is translated asymmetrically (Bowerman *et al.*, 1993; Seydoux and Fire, 1994), appearing at higher levels in P_1 descendants specifically. At the four-cell stage, maternal SKN-1 protein is present in the nuclei of both EMS and P_2. However, it is required for the identity of, and functions only in, the EMS blastomere. The level at which SKN-1 function is restricted exclusively to EMS was revealed when its zygotic targets, the *med-1* and *med-2* genes (for *m*es*e*ndoderm *d*etermining), a pair of unlinked though nearly identical target genes, were identified. In EMS, SKN-1 activates expression of the *med* genes, marking the switch from maternal to zygotic control in mesendoderm specification (Maduro *et al.*, 2001). The *med* genes encode GATA-type transcription factors, named for the degenerate consensus DNA binding site HGATAR (Lowry and Atchley, 2000) to which they bind. Regulation of *med-1* and *-2* by SKN-1 appears to be direct: the *med* promoters contain clusters of SKN-1 binding sites (Blackwell *et al.*, 1994) that bind SKN-1 protein *in vitro* and that are essential for reporter expression. At high levels, SKN-1 appears to be sufficient to activate *med* transcription: its widespread expression throughout the embryo results in ectopic activation of the *med* genes (Fig. 2), which in turn are able to convert non-EMS descendants into mesendoderm-generating cells.

The mechanism that prevents SKN-1-dependent activation of the *med* genes in P_2, the sister of EMS, was revealed when it was found that a maternally provided transcription factor, PIE-1, acts as a global repressor of transcription throughout the germline (P) lineage (Mello *et al.*, 1996; Seydoux *et al.*, 1996; Batchelder *et al.*, 1999). As such, PIE-1 blocks activation of *med-1/2* by SKN-1 in the P_2 lineage (Maduro *et al.*, 2001). While the *med* genes are first expressed exclusively in EMS and its early descendants, in *pie-1(−)* embryos, *med-1/2* are activated by SKN-1 in both EMS and, inappropriately, in P_2, causing both EMS and P_2 to adopt EMS-like fates (Mello *et al.*, 1992; Maduro *et al.*, 2001; Tenenhaus *et al.*, 2001).

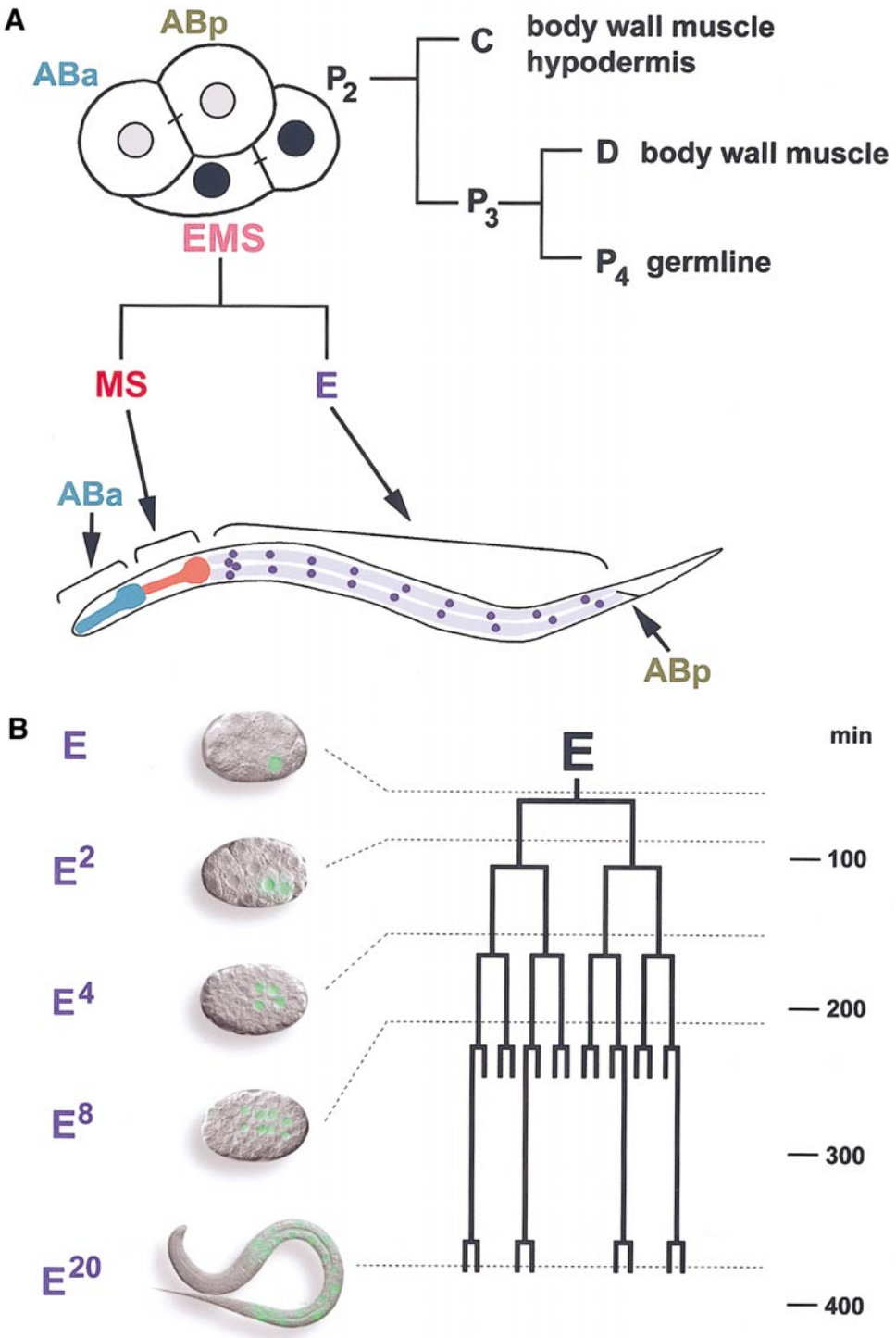

FIG. 1. Origin of the *C. elegans* digestive tract. (A) The embryonic fate map at the four-cell stage is diagrammed. The ventralmost cell, EMS, gives rise to the endoderm (E) precursor and a mesodermal precursor (MS), which produces primarily body wall muscle and the posterior half of the feeding organ, the pharynx. The anterior half of the pharynx and the entire rectum are produced by ABa and ABp, respectively. The sister of EMS, P_2, gives rise to body wall muscle, hypodermis, and the germline. (B) The *C. elegans* intestine is clonally derived from the E blastomere. The embryonic E lineage (adapted from Sulston *et al.*, 1983) is shown, along with a time scale marking embryonic development at 20°C. Horizontal lines indicate a cell division, while vertical lines indicate an undividing cell. Differential interference contrast (DIC) images overlaid with E-lineage-specific GFP fluorescence are shown for particular stages (indicated by the

TABLE 1
Genes Involved in Endoderm Formation

Gene	Product	EMS lineage phenotype[a]	Reference
Intermediate embryonic genes			
pos-1	CCCH finger	E, MS → C	Tabara *et al.*, 1999
spn-4	RNA recognition motif	E, MS → C	Gomes *et al.*, 2001
Restriction of endoderm to EMS			
pie-1	CCCH finger	P_2 → EMS	Mello *et al.*, 1992
sgg-1	GSK-3β kinase	C → EMS	Maduro *et al.*, 2001
EMS cell fate specification			
skn-1	bZIP/homeodomain TF	E, MS → C	Bowerman *et al.*, 1992
med-1/2	GATA-type TF	E, MS → C	Maduro *et al.*, 2001
end-1/3	GATA-type TF	E → C	Zhu *et al.*, 1997; M.F.M. and J.H.R., unpublished observations
E-MS polarity			
pop-1	TCF/LEF homolog	MS → E	Lin *et al.*, 1995a
wrm-1	β-catenin	E → MS	Rocheleau *et al.*, 1997
lit-1	Nemo-like kinase	E → MS	Rocheleau *et al.*, 1999
mom-2[b]	Wnt	E → MS	Thorpe *et al.*, 1997
Intestinal development			
elt-2	GATA-type TF	Loss of gut integrity	Fukushige *et al.*, 1998
elt-7	GATA-type TF	Unknown; redundant with *elt-2*	K. Strohmaier and J.H.R., unpublished observations
elt-4	GATA-type TF	Unknown	T. Fukushige and J. McGhee, personal communication

[a] Phenotypes due to loss of function.
[b] Among the "Mom" genes, we have only included *mom-2* by way of illustration. A full description of genes involved in the endoderm-including Wnt/MAPK pathway is found in Rocheleau *et al.* (1997), Thorpe *et al.* (1997), and Thorpe *et al.*, (2000).

In promoting EMS fate, it is apparent that MED-1/2 must also repress genes that direct the differentiation of other tissue types. In *skn-1* and *med-1/2* mutants, MS and E produce body wall muscle and hypodermis, tissue types that also arise from the somatic daughter of P_2, called C (Sulston *et al.*, 1983; Bowerman *et al.*, 1992; Maduro *et al.*, 2001). Specification of C fate requires the maternal CAUDAL-like transcription factor PAL-1, a protein found in all early P_1 descendants, including C, MS, and E (Hunter and Kenyon, 1996). The body wall muscle produced inappropriately by MS and E in *skn-1* and *med-1/2* mutant embryos also requires PAL-1 function (Hunter and Kenyon, 1996; Maduro *et al.*, 2001). The mesectodermal C-like program is therefore the default state for MS and E in the absence of mesendoderm-specifying information, owing to the presence of PAL-1. In the wild-type EMS lineage, PAL-1 C-promoting activity (and therefore acquisition of C fate) is overridden by SKN-1/MED-1/2. While it is not known at what level this inhibition of PAL-1-specific transcription functions, it seems likely that it is not the MED proteins, but rather their targets, that directly interfere with activation of PAL-1 targets, since elimination of the MED targets, *end-1* and *end-3*, also enables PAL-1 to activate C development in the E blastomere (see below).

Like EMS, the C cell contains high levels of both SKN-1 and PAL-1. Unlike EMS, however, C produces mesectoderm rather than mesendoderm. What distinguishes C from EMS? This mechanism operates at the level of *med* transcription: the *med* genes are not expressed in C irrespective of the presence of SKN-1. The repression of SKN-1-

number of E cells) on the left. For the top image, expression of a *med-1* reporter gene is shown (with the signal from the MS nucleus digitally erased; Maduro *et al.*, 2001). The image of the newly hatched larva at bottom shows expression of *elt-2*::GFP (with the signal from some nuclei digitally enhanced; Fukushige *et al.*, 1998); the remaining images are from an *end-3*::GFP strain (our unpublished observations). A *C. elegans* embryo is approximately 50 μm in length.

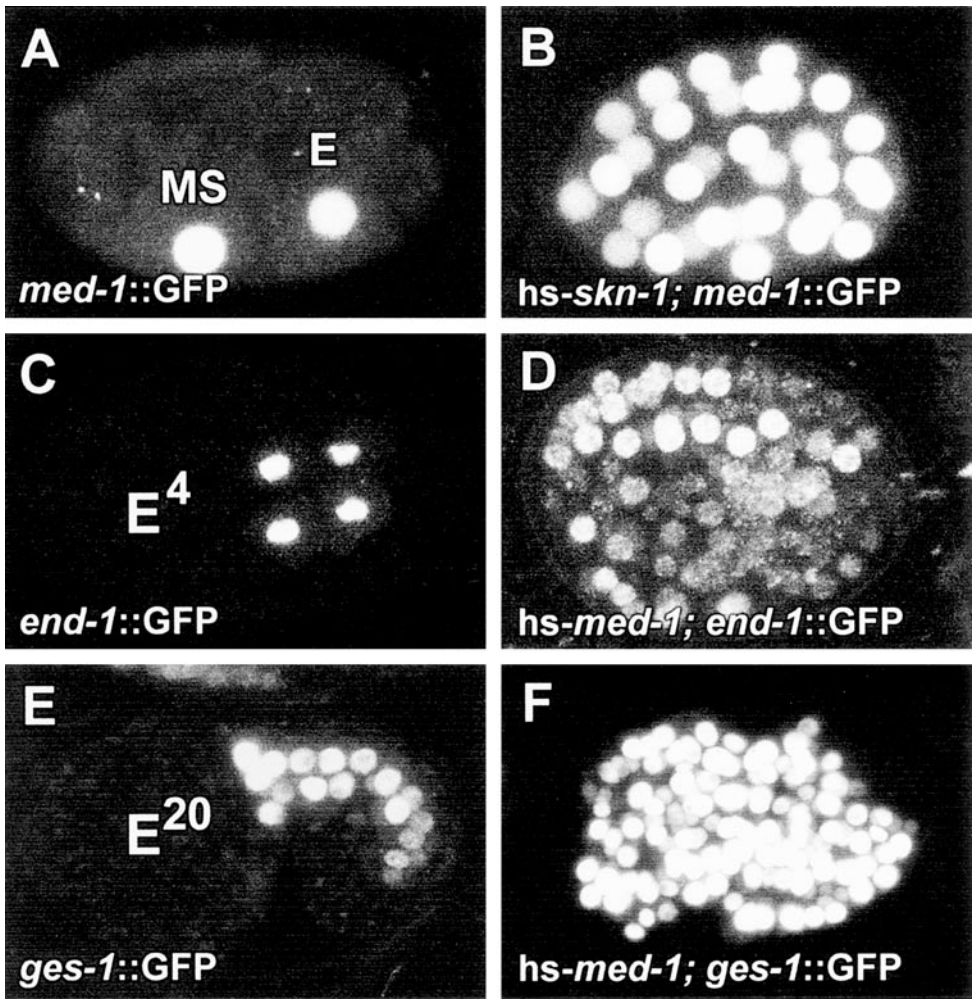

FIG. 2. Regulatory hierarchies demonstrated by ectopic expression experiments. Overexpression of endoderm regulators directs nonendodermal precursors to express downstream targets, demonstrating their potency as activators. Images are confocal micrographs showing expression of GFP reporters. (A) Normal expression of a *med-1* reporter at the eight-cell stage in MS and E. (B) Ectopic expression of *skn-1* driven by a heat-shock promoter leads to widespread *med* expression. (C) Expression of an *end-1* reporter at the E^4 stage. (D) Widespread expression of *end-1* when *med-1* is ectopically expressed. (E) Expression of a *ges-1* reporter in the embryonic intestine. (F) Overexpression of *med-1* in the early embryo leads to misexpression of *ges-1* in many lineages.

dependent activation of the *med* genes requires SGG-1, a GSK-3β homolog (Maduro *et al.,* 2001). In *sgg-1(−)* embryos, C inappropriately expresses the *med* genes and adopts an EMS-like fate. It is conceivable that this kinase might directly phosphorylate SKN-1, thereby abrogating its activation function; however, the well-described role for GSK-3β in the Wnt signaling pathway suggests an alternative mechanism that might involve Wnt signaling, which is required positively for E cell fate in the EMS lineage (see below).

med-1/2 are not the exclusive targets of *skn-1* in EMS. While both *skn-1(−)* and *med-1/2(−)* embryos fail to specify EMS as a mesendoderm progenitor, *skn-1* mutants are also defective in production of the secondary (induced) mesoderm engendered by ABa (Bowerman *et al.,* 1992). This effect results from the failure of MS to express an unidentified Delta-like ligand for the GLP-1 receptor that signals descendants of ABa to make mesoderm (Priess *et al.,* 1987; Bowerman *et al.,* 1992). Like the *med* genes, this ligand is apparently expressed in both E and MS (Lin *et al.,* 1995a). Thus, SKN-1 not only initiates mesendoderm specification in EMS by activating *med-1/2*, but also induces secondary mesoderm in another lineage through activation of another target. A number of genes encoding Delta-like ligands are present in the *C. elegans* genome, several of which contain SKN-1 consensus binding sites; these are strong candidates for other SKN-1 targets.

A Molecular Switch Distinguishes E from Its Sister MS

Early embryological experiments provided evidence that the fate of E is specified by cell-autonomous factors that are segregated through the early lineage (Laufer *et al.,* 1980; Edgar and McGhee, 1986). However, while E alone appears to contain the cell-intrinsic information necessary to produce differentiated intestine, the ability of EMS to engender an endoderm-producing E cell was shown to require a cell–cell interaction between it and P_2, which contacts EMS on its posterior side (Schierenberg, 1987; Goldstein, 1992). This signal induces an asymmetry in EMS that results in the adoption of endoderm fate by the daughter derived from the part of EMS that contacted the inducing cell (Goldstein, 1992). In the absence of this interaction, EMS divides symmetrically into two daughters that both exhibit an MS-like fate. The signal elicited by this interaction is mediated by intersecting Wnt and MAPK signaling pathways, whose components are maternally contributed (Thorpe *et al.,* 1997; Rocheleau *et al.,* 1997, 1999; Meneghini *et al.,* 1999; Shin *et al.,* 1999). Depletion of any of these components results in the production of two MS-like cells by EMS. Transduction of the endoderm-inducing Wnt/MAPK signals ultimately influences the state of maternally supplied POP-1, a member of the TCF/LEF class of HMG box proteins, which serve as the terminal transcription factors in Wnt signaling pathways (Brunner *et al.,* 1997; Korswagen and Clevers, 1999). Transduction of Wnt/MAPK signaling results in a change in nuclear POP-1 that is evident as a difference in levels: in MS, which does not receive the endoderm-inducing signal, POP-1 levels are high, while in an E cell that receives the signal, nuclear POP-1 is at low levels (Lin *et al.,* 1995a, 1998).

Elimination of maternal POP-1 reveals its most conspicuous role in the mesendoderm: that of a repressor of endoderm in MS (Lin *et al.,* 1995a). The Wnt/MAPK signal blocks the endoderm-repressing activity of POP-1 in the E cell; hence, this signaling system apparently induces endoderm by inactivating a repressor. Repression by POP-1 involves a complex containing the Groucho-like molecule UNC-37 and a histone deacetylase, HDA-1 (Calvo *et al.,* 2001). Specification of the endoderm, therefore, is directed by the combined action of a positive regulator, SKN-1 (through its zygotic targets, MED-1/2), and the switching mechanism provided by Wnt/MAPK signaling through POP-1. The combination of SKN-1 and POP-1 input, combined with the default activity of PAL-1, provides a mechanism by which, via these three transcriptional regulatory inputs, the E blastomere can adopt one of three fates (Fig. 3).

Recent findings indicate that POP-1 does not act exclusively in endoderm specification as a repressor. While mutants lacking either maternal SKN-1 or zygotic MED-1/2 appear to be completely deficient in specification of the MS blastomere, only ~50% of *med-1/2(RNAi)* and 80% of *skn-1(−)* embryos lack intestine (Bowerman *et al.,* 1992; Maduro *et al.,* 2001). Similarly, ≤60% of embryos lacking maternal MOM-2, the Wnt ligand, do not make intestine (Thorpe *et al.,* 1997; Rocheleau *et al.,* 1997). However, both *skn-1;mom-2* and *med-1/2;mom-2* embryos completely lack endoderm, suggesting the existence of a Wnt-dependent, and SKN-1/MED-1/2-independent, input that activates endoderm development (Rocheleau *et al.,* 1997; Maduro *et al.,* 2001). As TCF/LEFs function as activators downstream of Wnt signaling in other systems (reviewed in Korswagen and Clevers, 1999), one candidate for such a second positive activator of E fate is POP-1. Indeed, embryos lacking both SKN-1 and POP-1 show a much more penetrant loss of intestine, compared with *skn-1* mutants alone (J. Kasmir, J. Zhu, M. F. M., and J. H. R., unpublished observations). Moreover, a Lef-1-like consensus site is required for expression of a minimal E-specific promoter from the endoderm-promoting *end-1* gene (see below). Both observations suggest that, rather than blocking the repressive function of POP-1 in MS per se, Wnt/MAPK signaling in the E lineage converts POP-1 from a transcriptional repressor to a transcriptional activator. This toggling of POP-1 between repressing and activating states may occur throughout the entire development of the animal: POP-1 is a component of a molecular switch used reiteratively throughout *C. elegans* development to establish differences between sister cells that are born by A/P divisions (Lin *et al.,* 1998); moreover, a positive contribution of POP-1 has been implicated during larval development (Herman, 2001; Jiang and Sternberg, 1999).

Establishment of Endoderm by end-1 and end-3

How do the MED-1/2 activators and the Wnt/MAPK/POP-1 molecular switch collaborate to activate the endoderm-specific gene network? The notion that there must exist a zygotic gene responsive to the regulators that directs E cell identity was confirmed when deletions of an interval on chromosome *V* (the endoderm determining region, or "EDR") were identified that invariably eliminate endoderm. As in *skn-1* and *med-1/2* mutant embryos, these deletions cause E to adopt a C-like fate (Zhu *et al.,* 1997). Several screens for zygotic mutations that block endoderm formation failed to identify any penetrant gene-specific point mutations; the consequent inference that multiple genes in the EDR might act redundantly to specify endoderm was shown to be correct when each of two nearby genes in the EDR, *end-1* and *end-3,* was found to be individually capable of restoring endoderm formation in embryos lacking the EDR (Zhu *et al.,* 1997; our unpublished observations). Like *med-1/2,* both *end-1* and *-3* encode GATA factors. Transcripts from both genes are first detected in E shortly after its birth (Zhu *et al.,* 1997; our unpublished observations), indicating that the *end* genes are the earliest expressed genes known in the endoderm lineage.

end-1 and *end-3* have each been conserved (M.F.M. and J.H.R., unpublished observations) over the 20–40 million years since *C. elegans* and a close relative, *C. briggsae,* diverged (Kennedy *et al.,* 1993), indicating selective pressure for the presence of both genes. Indeed, while the *end*

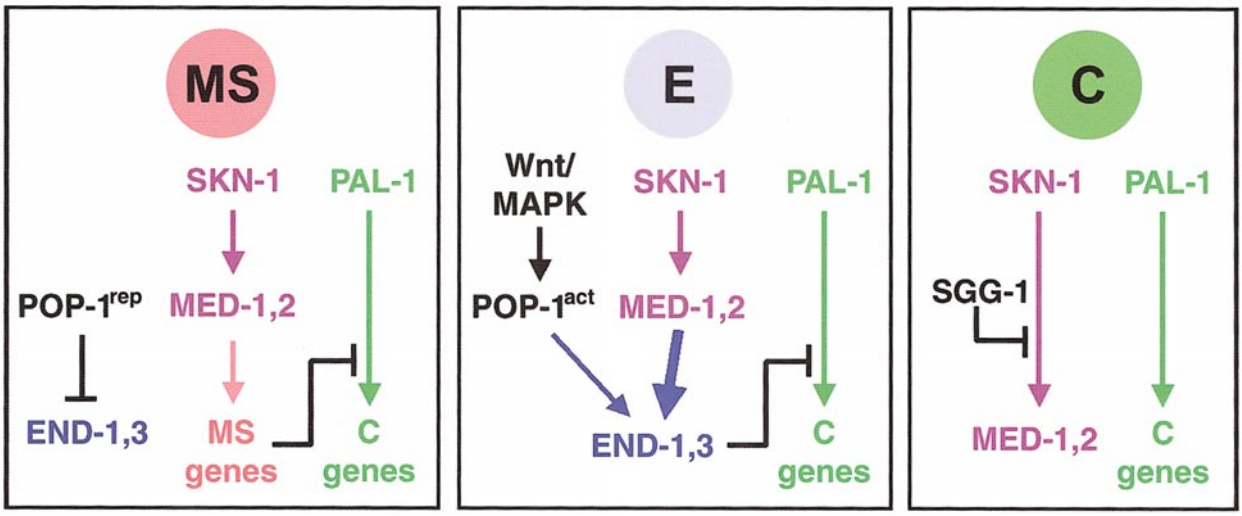

FIG. 3. Three-state model for specification of MS and E fates. The daughters of EMS can adopt one of three fates (MS, E, or C). In the absence of SKN-1, both MS and E adopt a C-like fate, owing to the presence of PAL-1 in these lineages. In the presence of SKN-1, high nuclear levels of POP-1 (POP-1^{rep}) direct MS fate, while the Wnt/MAPK-modified form of POP-1 (POP-1^{act}), at lower nuclear levels, directs endoderm fate by activating the *end-1/3* genes. The C-promoting activity of PAL-1 is blocked by SKN-1 targets. In the C blastomere, SGG-1 kinase prevents SKN-1 from activating the *med-1/2* genes in C, specifying an EMS-like fate. In the absence of *sgg-1* activity, *med-1/2* specify an E fate in Cp, the posterior daughter of C. In the absence of both SGG-1 and PAL-1, Cp adopts an E-like fate and Ca adopts an MS-like fate. The POP-1 switching system is presumed to direct an MS- vs E-like fate in the C daughters in such embryos.

genes overlap in function, they are not completely redundant for endoderm specification. The only zygotic point mutation identified that leads to (impenetrant) lack of intestine alters a residue in the zinc finger of the END-3 protein (M.F.M., R. Hill, J. Priess, J. Zhu, and J.H.R., unpublished results). Moreover, while simultaneous elimination of both gene functions leads to a large fraction of embryos without endoderm, a minor fraction depleted for *end-3* function alone also lack endoderm (E. Witze, M.F.M., and J.H.R., unpublished observations). One other gene in the EDR, *dpr-1* (formerly *end-2*), which encodes a nuclear receptor type transcription factor, can also rescue the endoderm differentiation defects of EDR mutants. However, RNAi of three genes (*end-1*, *dpr-1*, and *end-3*) does not produce significant enhancement compared with *end-1,3(RNAi)* alone, and further studies indicate that *dpr-1* performs a later role in endoderm development (E. Newman-Smith and J.H.R., unpublished observations).

end-1 and *end-3* are apparently direct targets of MED-1/2. Expression of *med-1* throughout early embryos is sufficient to activate expression of both (Maduro *et al.*, 2001) and both contain at least six GATA binding sites in their respective promoter regions, some of which are also present in the *C. briggsae* genes (Zhu *et al.*, 1997; our unpublished observations). Indeed, MED-1 interacts directly with both the *end-1* and *end-3* promoters *in vivo* based on the "nuclear spot assay" (M.F.M. and J.H.R., unpublished observations), which detects interactions of DNA binding proteins and their targets in individual cells of living animals (Carmi *et al.*, 1998; Fukushige *et al.*, 1999). Further, this assay demonstrated that MED-1 binds to the *end* genes in both the E *and* MS lineages, demonstrating that POP-1, which also binds to the *end* genes in the MS lineage, does not block activation of the *end* genes simply by precluding binding of the MED proteins to the *end* genes.

Other Maternal Pathways Provide Sequential Permissive States for SKN-1 and MED-1/2 Action, Respectively

High levels of SKN-1 or MED-1/2 are sufficient to drive expression of their downstream targets (*med-1/2* and *end-1/3*, respectively) in ectopic lineages, resulting in conversion of non-EMS blastomeres into mesendodermal precursors (Maduro *et al.*, 2001; our unpublished results). However, in the context of normal EMS development, other factors are required apparently to set the permissive state for target gene activation (Fig. 4). In embryos maternally depleted for the CCCH zinc finger protein POS-1, although SKN-1 is expressed at normal levels in EMS, a *med-1* reporter gene fails to be activated and E and MS adopt C-like fates, similar to *med-1/2(−)* embryos (Tabara *et al.*, 1999; Maduro *et al.*, 2001). Similarly, embryos lacking the maternal *spn-4* gene function show the same defect in E and MS specification (Gomes *et al.*, 2001); however, accumulation of SKN-1 protein and expression of a *med-1* reporter are unaffected in these mutants. Rather, *end-1* expression is abolished in *spn-4* mutants (Gomes *et al.*, 2001), suggesting

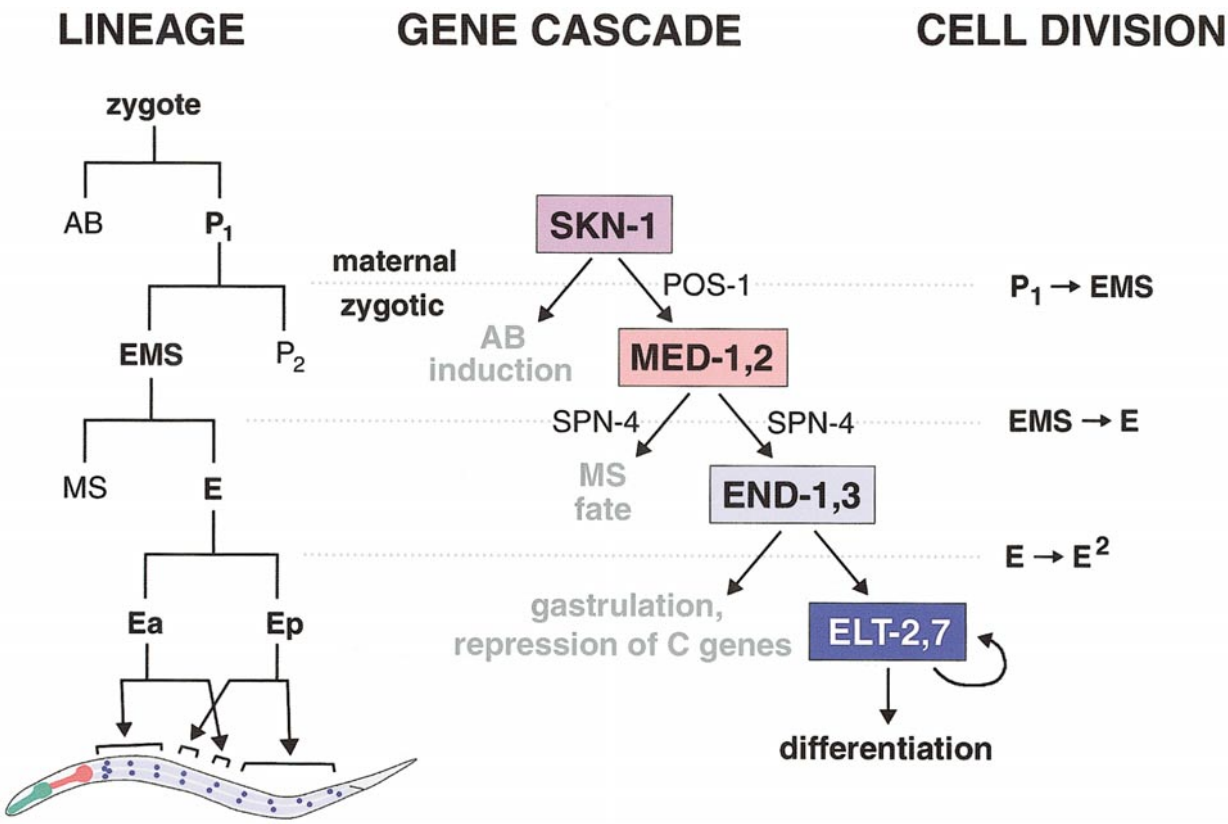

FIG. 4. The endoderm regulatory hierarchy. A gene cascade punctuated by cell divisions directs *C. elegans* endoderm differentiation. At each tier, a regulator activates another regulator following cell division, and also activates at least one other function, shown by divergent arrows. Downstream of SKN-1, the regulators are all GATA factors that function in redundant pairs.

that SPN-4 functions at the next tier in the E and MS gene regulatory hierarchy (Fig. 4). Both *pos-1* and *spn-4* mutant embryos display other phenotypes, consistent with other roles in the early embryo (Tabara *et al.*, 1999; Gomes *et al.*, 2001). POS-1 is primarily cytoplasmic and SPN-4 contains an RNA recognition motif, suggesting that the wild-type role of these proteins in *med* and *end* gene activation, respectively, is indirect and may involve regulation of RNA metabolism or expression. In constructing a coherent picture for how endoderm fate is assigned to E, it will be important to delineate the mechanisms by which POS-1 and SPN-4 establish permissive states that allow for the sequential activation of *med-1/2* and *end-1/3*, respectively.

Elaboration of Intestinal Fate: The end-1/3 Target Genes and GATAS Galore

As *end-1* and *-3* are the earliest known genes expressed specifically in the E lineage, they likely function at the top of a regulatory cascade that ultimately activates and maintains expression of terminal differentiation genes in the mature intestine. Consistent with this notion, overexpression of either *end-1* or *-3* in pregastrulation stage embryos is sufficient to promote endodermal differentiation from the descendants of nonendodermal precursors and to repress the differentiation of the cell types (i.e., mesoderm and ectoderm derivatives) normally made by those progenitors (Zhu *et al.*, 1997; our unpublished observations). However, while *end-1/3* expression appears to be temporally restricted to the early E lineage (no later than the E^8 stage), many terminal genes are not activated until much later in development. There must be other regulators, therefore, that function downstream of *end-1/3* to initiate and maintain the program of intestinal differentiation.

The *elt-2* gene, which encodes another GATA-type transcription factor, appears to be one such regulator of differentiation activated by END-1/3. ELT-2 was identified by its ability to bind to a pair of tandem GATA sites in the promoter of *ges-1*, a gene encoding a gut-specific esterase involved in digestion (Fukushige *et al.*, 1998; Kennedy *et al.*, 1993). Unlike the preceding chain of regulators, *skn-1*, *med-1/2*, or *end-1/3*, whose expression is detectable for only a few cell generations (Bowerman *et al.*, 1993; Maduro

et al., 2001; Zhu *et al.*, 1997; our unpublished observations), expression of *elt-2* begins in the immediate descendants of E and continues throughout the life of the animal (Fukushige *et al.*, 1998). The *elt-2* promoter itself contains GATA binding sites (Fukushige *et al.*, 1998), suggesting that expression of *elt-2* could be initiated by END-1/3. Indeed, ectopically expressed END-1 and -3 promote widespread expression of *elt-2* (Zhu *et al.*, 1997; our unpublished results) and GFP-tagged END-3 associates *in vivo* with the *elt-2* promoter in the E daughter nuclei during early interphase (our unpublished data).

Following activation of *elt-2* transcription by the ENDs, ELT-2 apparently maintains expression of its own structural gene. Ectopic *elt-2* itself can promote *elt-2* expression and widespread intestinal differentiation (Fukushige *et al.*, 1998) and GFP-tagged ELT-2 binds to the *elt-2* promoter *in vivo* (Fukushige *et al.*, 1999). Thus, expression of *elt-2* is apparently initiated by END-1,3; however, its continued expression is maintained thereafter by positive autoregulation. An *elt-2* knockout mutation results in a progressive loss of gut integrity, suggesting that, unlike the regulators that function upstream, *elt-2* expression is continuously required to maintain the differentiated state (Fukushige *et al.*, 1998).

Like the MED and END GATA factors, ELT-2 also shares a function with another factor. Although the GATA sites in the *ges-1* promoter are required for expression in the intestine (Aamodt *et al.*, 1991), the *elt-2* knockout mutation does not block *ges-1* expression, which suggested that another regulator acts in parallel with *elt-2* (Fukushige *et al.*, 1998). The ELT-7 GATA factor fulfilled this prediction (K. Strohmaier and J.H.R., unpublished observations). *elt-7* is similar to *elt-2* in its expression, ability to convert nonendodermal precursors into gut progenitor cells, and capacity to bind to its own gene. Depletion of *elt-7* activity by RNAi in an *elt-2(−)* strain results in a synergistic phenotype: though *ges-1* and many other gut-specific genes are still expressed in either single mutant, their expression is abolished when both gene functions are absent and an underdeveloped intestine is formed. Thus, like *med-1/2* and *end-1/3*, genetic redundancy operates at the next level of the endoderm gene regulatory network.

Finally, as if six GATA factors (MEDs, ENDs, and ELT-2/7) were not enough for endoderm development in this simple creature, a tiny (72-amino-acid) GATA factor encoded by an *elt-2*-adjacent gene, *elt-4*, is also expressed in the developing intestine (T. Fukushige and J. McGhee, personal communication). The ELT-4 zinc finger is nearly identical to that of ELT-2, suggesting that they may share some function. However, the later expression of ELT-4 and its inability to activate gut differentiation ectopically suggest that it may not be functionally redundant with ELT-2/7.

The transition from END-1/3 to ELT-2/7 (and possibly ELT-4) marks a conceptual shift from specification to differentiation of the endoderm. The ENDs provide the trigger for activation of the endoderm developmental program, and repress other differentiation programs; they then hand off the job to ELT-2/7, which, by an autoregulatory loop, appear to provide a "lock-down" system for maintenance of the differentiated state.

In addition to *elt-2/7*, there is evidence for additional target genes in the early E lineage acting immediately downstream of *end-1/3*. One hallmark of E specification is a delay in the cell cycle times of the E daughter cells, Ea and Ep, until such time as they have ingressed into the interior of the embryo, marking the onset of gastrulation (Sulston *et al.*, 1983). In the absence of *end-1/3*, the E daughter cells display a more rapid cell cycle time and divide on the ventral surface of the embryo (Zhu *et al.*, 1997; our unpublished results). As ELT-2/7 expression is detected after the E daughters have already ingressed, an additional END-1/3-dependent zygotic function, the identity of which is unknown, must exist to delay the Ea and Ep cell cycles.

Somewhat paradoxically, at least two maternal genes appear to contribute to the gastrulation function of *end-1/3*. Mutants lacking the function of GAD-1, a WD motif-containing protein, or EMB-5, a putative chromatin structure regulator similar to yeast SPT6, demonstrate a failure of the E daughter cells to gastrulate properly, although endoderm is specified correctly in most embryos (Knight and Wood, 1998; Nishiwaki *et al.*, 1993). The effect on gastrulation may be indirect, as *gad-1* and *emb-5* mutant embryos are pleiotropic for other developmental defects. Instead, the gastrulation defect may be attributable to their effect on the expression levels of *end-1*/3, as *skn-1* and *med-1/2* mutants that still produce endoderm often do not gastrulate correctly and there is evidence that the *gad-1* and *emb-5* mutations result in diminished expression levels of the *end* genes (E. Witze, M.F.M., I. Mengarelli, and J.H.R., unpublished observations). The study of mutants defective in gastrulation, therefore, may reveal new insights into the transcriptional activation of *end-1/3*.

Emergent Themes in Endoderm Specification

Our discussion has accounted for the core regulators known to direct endoderm specification and differentiation in *C. elegans*. These findings present a picture of the regulatory pathway for endoderm development that follows a familiar theme in development: cell fates are specified by stepwise restriction of developmental potential, followed by activation of instructive signals that direct terminal differentiation. We can explain the existence of the regulators in the *C. elegans* endoderm regulatory network based on two emergent principles as shown in Fig. 4: first, progression through successive tiers is punctuated by a cell division. Second, the regulators at each level not only activate the genes at the next tier in the cascade, but also carry out additional functions not encompassed by the next regulatory tier.

A direct correlation exists between cell division and successive levels of regulation downstream of SKN-1 (Fig. 4). The onset of *med-1,2* expression is detectable in EMS in

four-cell-stage embryos; expression peaks in the E and MS nuclei (Maduro *et al.*, 2001; our unpublished observation). The *end-1* and *-3* transcripts are first detectable in the E cell; expression peaks in the E daughters, Ea and Ep (Zhu *et al.*, 1997; our unpublished observation). Finally, *elt-2* and *-7* expression is first detected toward the end of the Ea/Ep cell cycles (Fukushige *et al.*, 1998; K. Strohmaier and J.H.R., unpublished observation).

Why should progression through these three levels in the gene regulatory cascade be correlated with the cell cycle? One possibility is that activation of genes at each successive tier, and establishment of a new transcriptional state, requires a complete cell cycle. This might be the case if DNA replication is a prerequisite for formation of a complex competent for transcription initiation, i.e., to allow for remodeling of chromatin and/or interaction of regulatory factors with their targets. In fact, the notion that sequential steps in the regulatory cascade are punctuated by one cell cycle was presaged many years ago when it was found that a round of DNA synthesis in the first cell cycle following birth of the E cell is essential for expression of at least one ELT-2 target, *ges-1* (Edgar and McGhee, 1988; Fukushige *et al.*, 1998). Alternatively, the correlation between transitions in the regulatory hierarchy and the occurrence of cell division may be coincidental and might simply reflect the fact that the rapid pace of cell division and sequential gene activation occur over similar time frames in the early *C. elegans* embryo, as can be addressed by experiments analogous to those described earlier (Edgar and McGhee, 1988).

Though we do not know whether cell division between tiers is an essential element in the deployment of the gene regulatory network, the *existence* of each tier is easily justified by considering the regulatory function at each step (Fig. 4). Initially, maternal factors set embryonic polarity and establish broad domains of specification, resulting in the segregation of functional SKN-1 activity to a single blastomere, EMS. This is achieved by differential translation of maternal *skn-1* mRNA in P_1 descendants and abrogation of its activity in the non-EMS lineage in which it is translated. SKN-1 then defines EMS fate broadly by activating expression of both the next tier of regulators, MED-1/2, and at least one other target gene that encodes the ligand for induction of secondary mesoderm in the AB lineage. The combined activities of MED-1/2 and Wnt/MAPK signaling through POP-1 lead to activation of *end-1/3* specifically in E. SKN-1 is an active transcription factor in EMS; if SKN-1 were to bypass *med-1/2* and activate *end-1/3* directly, *end-1/3* would be expressed in EMS, before POP-1 repression could occur. Hence, the intermediate regulators MED-1/2 provide a means of delaying *end-1/3* activation until the POP-1 repression/activation switch can function to distinguish the EMS daughters. Conversely, if SKN-1 activity was delayed until after the division of EMS, there might be insufficient time to activate expression of the signaling ligand that induces secondary mesoderm during a narrow window of time in the AB lineage. Hence, the need to integrate other SKN-1 functions and Wnt/MAPK input accounts for the existence of the *med* regulatory tier.

The division of function continues at the next tier in the hierarchy with the activity of END-1/3 in E. These regulators define the separation of endoderm from mesoderm, and set the state of the E lineage to commit to intestinal differentiation. In addition to activating the next tier of regulators, the autoregulatory *elt-2* and *elt-7* genes, END-1/3 block activation of other specification pathways (e.g., PAL-1-promoted C fate) and delay the Ea and Ep cell cycles, allowing gastrulation to occur properly, when there are only two E cells. Why is expression of END-1/3 transient, rather than maintained throughout intestinal development? We propose that sporadic low-level activation of *end-1/3*, which are potent activators of endoderm fate, in inappropriate lineages, coupled with their amplification through autoregulation, could result in ectopic mis-specification of endoderm. A two-tiered system may help to ensure fidelity: activation of *elt-2/7* by the combination of both END-1/3 and ELT-2/7 themselves may be required to raise the levels of ELT-2/7 to a critical threshold required for a self-sustaining autoactivation mechanism. Once ELT-2/7 autoactivation becomes fully functional, END-1/3 would become unnecessary and might even lead to detrimental overactivation of *elt-2/7*. Such a threshold mechanism is consistent with the observation that *end-1* reporter expression is detected at low levels in MS, which nonetheless never gives rise to endoderm. Therefore, we argue that the multiple roles of END-1/3 and the need for transience in their expression, preclude them from acting as direct regulators of terminal differentiation.

At the next tier, the *elt-2/7* genes are tissue-specific regulators that initiate and maintain intestinal differentiation. Through positive autoregulation, *elt-2/7* set a permanent state for the expression of terminal differentiation genes, ensuring continued function of the intestine throughout the life of the animal.

In total, therefore, the known regulatory cascade defines a logical paradigm for the way in which a specification/differentiation pathway is built, from the establishment of early embryonic polarity, through to the expression of terminal differentiation functions.

Expression of Terminal Differentiation Genes

The end point of endoderm development is, by definition, the expression of gene products that function in the differentiated intestine. Several such genes have been identified through genetic and biochemical approaches, and in a few cases, their expression and promoter requirements have been studied (Table 2). As mentioned earlier, the esterase-encoding *ges-1* gene is expressed when the developing gut consists of only four cells (Aamodt *et al.*, 1991; Kennedy *et al.*, 1993). This early expression of *ges-1* appears anomalous among genes expressed exclusively in the intestine. For example, the acid phosphatase-encoding *pho-1* gene is expressed beginning late in embryogenesis (Beh *et al.*, 1991; T.

TABLE 2
Examples of Terminal Differentiation Genes

Gene	Product	Expression	Reference
ges-1	Esterase	E^4 to adult	Kennedy *et al.*, 1993
pho-1	Acid phosphatase	Most intestinal cells, strongest from late embryogensis	Beh *et al.*, 1991; T. Fukushige and J. McGhee, personal communication
F10C1.7	Intermediate filament (MH33 antigen)	Early embryo to adult	T. Fukushige and J. McGhee, personal communication
asp-1	Aspartic protease	Late embryonic and early larval stages	Tcherepanova *et al.*, 2000
cpr-1	Cysteine protease	Postembryonic gut	Britton *et al.*, 1998
W03G9.4	Aminopeptidase P	Larvae and adults	Laurent *et al.*, 2001
vha-6	V-ATPase	Larvae and adults	Oka *et al.*, 2001
mtl-2	Metallothionein	Inducible in larvae and adults	Freedman *et al.*, 1999
vit-2	Vitellogenin	Adult intestine	Spieth *et al.*, 1988

Fukushige and J. McGhee, personal communication), while the *cpr-1* gene, which encodes a cysteine protease, is first expressed after hatching (Britton *et al.*, 1998). Mutation of two GATA consensus sites in the *cpr-1* promoter abolishes expression, implicating the GATA endoderm network in *cpr-1* regulation (Britton *et al.*, 1998); if ELT-2/7 or -4 are responsible for activating its expression, there is presumably an additional event that delays activation until larval development has begun. A similar situation exists with the metallothionein genes *mtl-1* and *mtl-2*, which are activated exclusively in the intestine by exposure to cadmium or heat stress (Freedman *et al.*, 1993). Expression of *mtl-1* and *mtl-2* reporters requires ELT-2 activity and GATA binding sites in the *mtl* promoters (Moilanen *et al.*, 1999).

Molecular approaches, such as *in situ* hybridization of expressed sequence tags (ESTs) from cDNA libraries and "promoter trapping" experiments, have also identified intestine-specific genes (e.g., Tabara *et al.*, 1996; Hope, 1991). More recently, transcriptional profiling using microarrays can, in a single experiment, reveal the relative expression of an entire suite of genes at a particular developmental stage or under a specific condition (e.g., Reinke *et al.*, 2000). Experiments using this approach have relied primarily on mutant backgrounds, developmental staging, or environmental manipulation to generate differences between populations of mRNAs. Kim *et al.* (2001) pooled the results of 553 such microarray experiments and grouped transcripts that are coregulated under a variety of experimental conditions. The data are represented as a three-dimensional "terrain map" in which groups of similarly regulated mRNAs are clustered, with gene density of each cluster represented by the height, thereby creating "mountains" of correlated gene expression patterns (Fig. 5A). The 8th largest cluster represents 803 transcripts, some of which are known to be specifically enriched in the intestine, and includes genes encoding digestive enzymes such

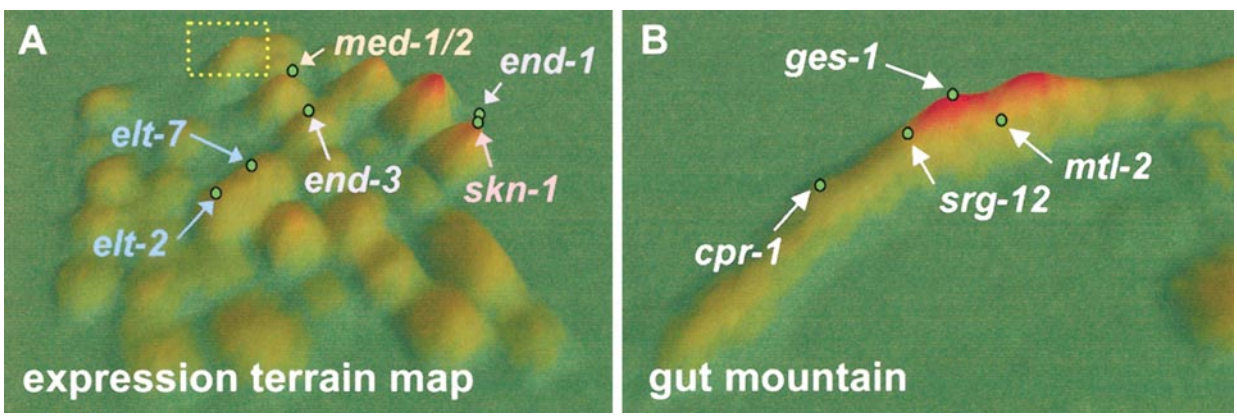

FIG. 5. Distribution of endoderm genes in the *C. elegans* gene expression map. (A) The endoderm-specifying regulators *skn-1*, *med-1/2*, *end-1/3*, and *elt-2/7* do not cluster in this view of the entire expression terrain map (Kim *et al.*, 2001). (B) An enlargement of the "gut mountain" (outlined by a box in A) reveals that many terminal differentiation genes (*ges-1*, *srg-12*, *mtl-2*, and *cpr-1*) show coordinate control. For both panels, VxInsight was used to produce terrain maps and locate individual genes, as described in Kim *et al.* (2001).

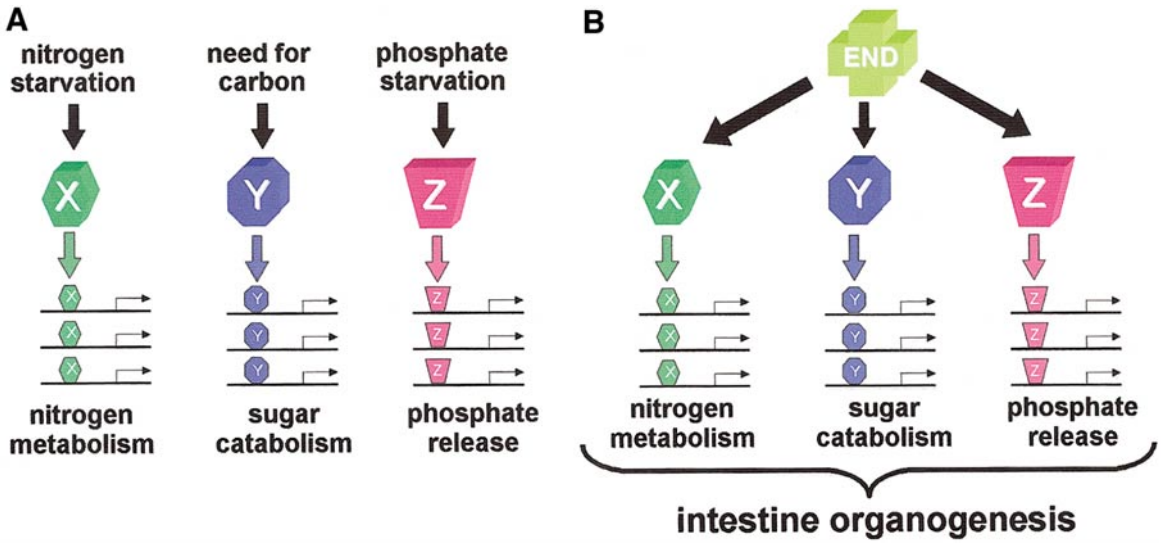

FIG. 6. A model for the hypothetical evolution of an endoderm gene regulatory network. We propose that an endoderm identity gene could arise when independent gene batteries preexisting in a primordial eukaryotic organism (X, Y, and Z in panel A) are placed under the control of a single regulator (END in B). The original regulators of the gene batteries would become intermediate regulators in the gene network and would be stably maintained in evolution, explaining in part the large number of downstream transcription factors present in the simple endoderm of *C. elegans*.

as proteases, carboxylesterases, and lipases (Fig. 5B). Of 8 genes known to be expressed in differentiated intestine, 5 were found in this cluster (Kim *et al.*, 2001). At least 15 of the 803 genes can be immediately classified as transcription factors by sequence homology, 9 of which are nuclear receptors. However, there are also some unexpected groupings of endoderm regulators: *elt-2* and *elt-7*, while clustered, are found in Mount 1, a group that includes muscle and neuronal genes, and while END-1 and -3 are in separate mountains, SKN-1 and END-1 map very closely in the same mountain. Hence, correlative transcriptional profiling is not likely to identify all genes whose expression is endoderm-specific. In addition, it is apparent that not all genes in the "gut mountain" are specific for the endoderm; for example, the *egl-5* Hox gene, which controls cell fates in the posterior of the worm, maps to the gut mountain even though its expression is not gut-specific (Kenyon *et al.*, 1997). Future microarray experiments that enrich for intestine-specific mRNAs should identify the complete set of endoderm-expressed genes and will clarify the relevance of correlated profiling analyses to the dissection of gene regulatory networks.

Other Transcriptional Regulators Expressed in the Intestine

Many other regulators have been identified that show expression in the E lineage; most of these are also expressed in other lineages as well. As the number is rapidly expanding and the role of these factors in endoderm development is unknown, we will not list them, but will describe a few examples here.

The best studied of these is PHA-4, an HNF3/forkhead-related transcription factor required for pharynx differentiation (Horner *et al.*, 1998; Kalb *et al.*, 1998). Although PHA-4 is expressed at high levels in the pharynx and rectum (where it programs organ-specific differentiation), it is also expressed at low levels in the intestine. Widespread overexpression of *elt-2* drives ectopic *pha-4* expression, and ELT-2 binds the *pha-4* promoter *in vitro*, consistent with direct regulation by ELT-2 (Kalb *et al.*, 1998).

Intriguingly, maternal genes that function in endoderm specification also have zygotic functions in the gut. Both *skn-1* and *pop-1* are zygotically expressed in the intestine (Bowerman *et al.*, 1993; Lin *et al.*, 1998). SKN-1 is apparently required to activate genes in response to heat or oxidative stress (K. Blackwell, personal communication), while the LIT-1/POP-1 pathway provides anteroposterior patterning information, similar to its role in other lineages (Schroeder and McGhee, 1998; Hermann *et al.*, 2000; Lin *et al.*, 1998).

An anticipated role for regulators expressed downstream of *elt-2/7* is the control of genes that function in intestinal morphogenesis. Indeed, the intestines of embryos lacking the function of both genes generally do not contain a lumen or other features of a differentiated gut (K. Strohmaier and J.H.R., unpublished observation). During morphogenesis of the developing gut tube, three pairs of cells in the anterior portion of the gut tube rotate relative to the other cells

(Sulston *et al.*, 1983; Leung *et al.*, 1999). This rotation was found to require the activity of LIN-12, a Notch-like receptor, and LAG-1, a transcription factor that is a Notch pathway effector (Christensen *et al.*, 1996; Hermann *et al.*, 2000). Although initially expressed throughout the gut primordium, LIN-12 is downregulated in the left half as the result of an interaction with descendants of MS that express the Delta-like ligand LAG-2 (Hermann *et al.*, 2000). LIN-12 asymmetry is correlated with the occurrence of intestinal twist later in development; it is not known whether LAG-1 is required specifically in the intestine, as it is expressed ubiquitously at the time it is required (Hermann *et al.*, 2000). These results demonstrate the contribution of extrinsic signals to gene regulation and intestinal development, and lay the groundwork for studies that correlate transcription factor activity to morphogenesis.

Several other transcription factors have been identified that are specifically expressed in the developing intestine; however, none is expressed as early as either the ENDs or ELT-2/7, suggesting that all may be subservient to these core regulators in the gene regulatory network. While the function of these transcription factors is obscure at present, the large number of endoderm-expressed genes will make it possible to dissect the regulatory interactions through which they act. Moreover, these factors are likely to reveal conserved elements in the metazoan endoderm gene regulatory network (see below).

Redundant GATA Factor "Pairs" in Germ Layer Specification

Of 11 GATA factors encoded in the *C. elegans* genome, 5 (END-1/3, ELT-2/7, and perhaps ELT-4) function specifically in endoderm development, 2 specify mesendoderm (MED-1/2), and the remaining 4 (ELT-1, -3, -5, -6) function in the ectoderm (Zhu *et al.*, 1997; Fukushige *et al.*, 1998; Maduro *et al.*, 2001; Gilleard and McGhee, 2001; Koh and Rothman, 2001; T. Fukushige and J. McGhee, personal communication; K. Strohmaier and J.H.R., unpublished results). It is intriguing that 8 of these (MED-1/2; END-1/3; ELT-2/7; ELT-5/6) constitute genetically redundant pairs. Of the remaining 3, only *elt-1* is convincingly not redundant with another gene (Page *et al.*, 1997): deletions of *elt-3* or *elt-4* do not result in any recognizable phenotype (Gilleard and McGhee, 2001; J. McGhee, personal communication). Curiously, only ELT-1 contains 2 GATA-type zinc fingers, a characteristic feature of the vertebrate GATA factors (Lowry and Atchley, 2000). The carboxyl zinc finger of vertebrate GATA factors is involved in binding to a cognate GATA site, and shares the most homology with the carboxyl finger of ELT-1 and the single fingers of the remaining *C. elegans* GATA factors (Lowry and Atchley, 2000; reviewed in Newton *et al.*, 2001). However, new data suggest that the amino finger, which has previously been described as being important for interactions with coregulators, can also interact with DNA (Newton *et al.*, 2001). One possibility, then, is that the single-finger GATA factor pairs achieve a higher promoter-binding specificity through heterodimeric interactions, perhaps necessitating the existence of pairs of partially redundant, single-finger GATA factors.

Endoderm Regulatory Networks in Other Nematodes

The elucidation of the regulatory network underlying *C. elegans* endoderm development provides a strong base from which to explore the mechanisms that drive evolutionary change within such networks. Comparisons of noncoding regions between *C. elegans* and *C. briggsae* have pointed to regions important for regulation (e.g., Kennedy *et al.*, 1993; Gower *et al.*, 2001; MacMorris *et al.*, 1994). Moreover, at least some components of endoderm specification appear to be conserved in *C. briggsae*. At the top of the cascade, *C. briggsae* contains a *skn-1* homolog (Kent and Zahler, 2000). A *C. briggsae end-1* homolog, which can rescue endoderm formation in *C. elegans* EDR-deficient mutants, more closely resembles *C. elegans end-1* than *C. elegans end-3*, indicating that the divergence of *end-1* and *end-3* occurred before the *briggsae/elegans* split (J. Kasmir, J. Zhu, M. F. Madura, and J. H. Rothman, unpublished observations). The *C. briggsae ges-1* gene is expressed in a similar manner when transgenes are introduced into either *C. briggsae* or *C. elegans*, implying conservation of the upstream regulatory network (Kennedy *et al.*, 1993). While the 5′ flanking sequences are not well conserved, several short regions show high conservation, including a 17-mer sequence that is completely conserved between the two species (Kennedy *et al.*, 1993). Intriguingly, deletion of a conserved promoter element, containing two GATA sites, in either species produces an expression component in the pharynx and tail, suggesting that a similar mechanism blocks *ges-1* activation outside the intestine (Egan *et al.*, 1995). As expected, such studies have identified more similarities than differences and one suspects that the *C. elegans* and *C. briggsae* endoderm regulatory networks will be largely identical. As such, *C. briggsae* may be more useful for identifying putative regulatory regions by sequence conservation, as a "proving ground" for models advanced by work in *C. elegans*, and a tool for identifying mechanisms that drive microevolutionary changes in the architecture of *cis*-acting regulatory elements. In some cases, the genetic redundancy within the core endoderm regulatory network means that such short time-scale changes can also be inferred from a single species. For example, although the *med* genes are nearly identical, the divergences are focused in noncoding regions and exclude known sites for interaction with transcriptional regulators, including SKN-1 and GATA factors.

Analyses in more deeply divergent nematodes have provided insights into developmental strategies that have been adopted in evolution to restrict endoderm to the appropriate blastomere during early embryogenesis. Early embryogenesis in *C. elegans* is highly mosaic, particularly with regard to generation of endoderm from the E cell. Even in the

marine nematode *Enoplus brevis*, which shows highly indeterminate development, E has the cell-intrinsic capacity to make intestine (Voronov and Panchin, 1998). However, unlike *C. elegans*, a related and morphologically similar nematode, *Acrobeloides nanus*, shows highly regulative development, which extends even to specification of the endoderm precursor (Wiegner and Schierenberg, 1998, 1999): endoderm fate can be assigned to any early blastomere cultured in isolation. Thus, unlike *C. elegans*, repressive intercellular interactions prevent the acquisition of endoderm fate in *A. nanus*, in which all early blastomeres are potentiated to make endoderm. With the core *C. elegans* endoderm specification components in hand, it will be of interest to ask at what level the regulatory mechanisms that restrict endoderm to the E cell lineage diverge between *A. nanus* and *C. elegans*.

Conservation of an Ancestral Endoderm Regulatory Networks

All metazoan phyla, with the exception of the sponges, contain an endoderm. As this germ layer was likely invented once during animal evolution, one supposes that at least some elements of the gene regulatory network regulating formation of the endoderm are pervasive across phylogeny. The scattered evidence available is consistent with this postulate. In particular, GATA transcription factors play a key role in endoderm development in many other systems. A striking parallel is seen with the *Drosophila* SERPENT GATA factor (Rehorn *et al.*, 1996): mutants lacking this gene fail to generate a midgut, the endodermal portion of the digestive tract, and cells that would become endodermal are apparently converted to ectoderm (Reuter, 1994). As in *C. elegans*, a second GATA factor, dGATAc, apparently acts downstream of this endoderm-specifying GATA factor (Lin *et al.*, 1995b). The GATA family of transcription factors in vertebrates consists of two groups of genes defined by sequence homology and expression pattern. GATA1, -2, and -3 function in hematopoiesis (reviewed in Orkin and Zon, 1997), while GATA4, -5, and -6 function primarily in cardiac and endoderm development (reviewed in Charron and Nemer, 1999). In the zebrafish, *faust*/GATA5 is required for the formation of the gut tube and other endodermal organs (Reiter *et al.*, 1999, 2001) and overexpression of GATA5 in *Xenopus* can respecify cells of mesodermal or ectodermal origin towards an endoderm fate (Weber *et al.*, 2000). In addition, the *C. elegans* END-1 GATA factor is capable of activating endoderm development in *Xenopus* animal caps, which would otherwise produce exclusively ectoderm (Shoichet *et al.*, 2000).

The expression of a number of other transcription factors in the *C. elegans* endoderm is also consistent with substantial conservation of the endoderm gene regulatory network. For example, members of the odd-skipped (K. Strohmaier and J.H.R., unpublished observations), HNF4 (K. Koh and J.H.R., unpublished observations), Sox (Hanna-Rose and Han, 1999), and HNF3 (Horner *et al.*, 1998; Kalb *et al.*, 1998) gene families are expressed in the *C. elegans* endoderm and in the digestive systems or developing endoderm of other animals ranging from *Drosophila* to vertebrates (Zhong *et al.*, 1993; Morrisey *et al.*, 1998; Stainier, 2002). It remains to be seen whether this similarity in expression is superficial or instead reflects genuine conservation of an endoderm regulatory network.

Finally, based on the similarities between *C. elegans* EMS specification and mesendoderm development in vertebrates, it has been proposed that the conjunction of endoderm and a subset of mesoderm as a mesendoderm layer is of ancient origin (Rodaway and Patient, 2001). Thus, while the inputs into the regulatory gene cascade are different, there is good evidence to suggest that they impinge upon a common, conserved pathway.

In comparing the known pathways for endoderm development in diverse animals, a general theme emerges: the regulatory inputs that initiate the program for endoderm are dramatically divergent, yet these divergent inputs may activate a well-conserved gene regulatory network that functions in all metazoans. For example, though endoderm differentiation in both *Drosophila* and *C. elegans* appears to be regulated by a cascade of GATA factors and other similar regulators, including HNF-4-like and Fkh-like factors, the early events that first select cells for the endoderm pathway are entirely distinct, involving the HKB Sp1/egr-like gap gene product (Bronner *et al.*, 1994) in the former and SKN-1 and Wnt/MAPK signaling in the latter (Bowerman *et al.*, 1993; Rocheleau *et al.*, 1997, 1999; Thorpe *et al.*, 1997).

If, as seems likely, there exists a conserved endoderm gene regulatory network across metazoans, in which not only specific components, but also the logic of cross-regulatory interactions and genetic circuitry are common, then it will be of great value to reveal the common parameters of the network. Elucidating the entire endoderm regulatory networks in disparate species will not only reveal limits on the degree of conservation and the rules for diversification of gene networks, but may also provide clues as to the steps by which germ layers were invented during the transition from a germ layer-less ancestral metazoan form (perhaps a sponge-like organism) to a diploblast, containing two germ layers. The plethora of transcription factors present in the endoderm, not only in, for example, sea urchins (e.g., Yuh *et al.*, 1998), in which the endoderm consists of many cells, but also in *C. elegans*, which contains one of the simplest endoderm layers of any animal may prove revealing as to the evolutionary steps leading to assembly of the network.

Why does the *C. elegans* endoderm, consisting of only 20 nearly identical cells at hatching, express so many transcription factors? Would it not be sufficient to maintain a simple, core regulatory cascade that activates the complete set of intestine-specific differentiation genes? A simplified model that could account for this apparent plethora of transcription factors in the endoderm proposes that the gene regulatory network might have been assembled during evolution from preexisting gene batteries present in a

primordial unicellular eukaryote (Fig. 6). As the endoderm in all animals gives rise to an intestine, one set of differentiation genes expressed in the endoderm of all animals would be those involved in digestion. In unicellular eukaryotes, there exist many independently regulated gene batteries, the members of which act in common metabolic processes (e.g., nitrogen or phosphate metabolism; Marzluf, 1997; Lenburg and O'Shea, 1996). These batteries are coupled together under the control of a single battery regulator, which itself is responsive to varying environmental conditions (e.g., GATA factors in nitrogen metabolism). Each battery regulator acts through recognition sites specific for the regulatory domains of all members of that particular gene battery. It would be possible, in principle, to create an organ in a multicellular creature that is constitutively dedicated to digestion by tying together the gene batteries under the control of a single, organ-specific regulator. Each of the individual gene batteries could become subordinated to the organ-specific regulator (e.g., an endoderm-promoting GATA factor) by a single step in which a binding site for this regulator (e.g., a GATA site) is created in the regulatory domain of the battery regulator. Although of relatively minor importance to the overall function of the organ, each subordinate regulator would then carry with it a set of subservient genes with their cognate recognition sites. These intermediaries, which link the global, organ-specific regulator to the gene batteries, would be maintained over long evolutionary periods, since a prohibitively large number of changes would be required to alter the recognition sites of the entire set of genes in each battery to one appropriate for the organ-specific regulator. If this model is correct, the gene networks in present-day metazoans might well be palimpsests that reveal underlying unicellular gene batteries subsumed within the larger regulatory networks of organs or germ layers. This proposal will be tested once the complete architecture of the gene regulatory network for endoderm is elaborated and individual circuits are compared to those in present-day unicellular eukaryotes.

Conclusion and Future Prospects

From the analysis of the core pathway for endoderm development in *C. elegans*, several principles have emerged that seem to guide the construction of the general transcriptional regulatory network for endoderm. It remains to be determined whether these principles actually reflect underlying mechanisms by which gene regulatory events are rapidly and sequentially deployed. As pathways for the specification of other early blastomeres and their descendant lineages in *C. elegans* are elucidated, it will become evident, for example, whether a cell division-linked gene cascade typifies such pathways in general.

With the identity of the main regulators in hand, the identification of ~800 genes that appear to be coregulated with gut-specific genes based on correlated transcriptional profiling and genome-wide methods for rapidly assessing gene function (Gonczy *et al.*, 2000; Fraser *et al.*, 2000), the *C. elegans* endoderm is likely to be among the first examples in which a transcriptional regulatory network for germ layer differentiation will be completely elucidated. The goal of such a task is to learn how such networks are assembled, why they are so complex, what features are common to all regulatory networks, how such networks might have evolved from simpler circuits, and to what extent the details of the network are conserved or subject to evolutionary plasticity.

ACKNOWLEDGMENTS

We are grateful to K. Blackwell, T. Fukushige, and J. McGhee for communicating unpublished results and to members of the Rothman laboratory for comments on the manuscript. M.F.M. was supported by a fellowship from the Natural Sciences and Engineering Research Council of Canada. This work was supported in part by grants from the NIH (HD37487) and the March of Dimes (to J.H.R.).

REFERENCES

Aamodt, E. J., Chung, M. A., and McGhee, J. D. (1991). Spatial control of gut-specific gene expression during *Caenorhabditis elegans* development. *Science* **252,** 579–582.

Aguinaldo, A. M., Turbeville, J. M., Linford, L. S., Rivera, M. C., Garey, J. R., Raff, R. A., and Lake, J. A. (1997). Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* **387,** 489–493.

Batchelder, C., Dunn, M. A., Choy, B., Suh, Y., Cassie, C., Shim, E. Y., Shin, T. H., Mello, C., Seydoux, G., and Blackwell, T. K. (1999). Transcriptional repression by the *Caenorhabditis elegans* germ-line protein PIE-1. *Genes Dev.* **13,** 202–212.

Beh, C. T., Ferrari, D. C., Chung, M. A., and McGhee, J. D. (1991). An acid phosphatase as a biochemical marker for intestinal development in the nematode *Caenorhabditis elegans*. *Dev. Biol.* **147,** 133–143.

Blackwell, T. K., Bowerman, B., Priess, J. R., and Weintraub, H. (1994). Formation of a monomeric DNA binding domain by Skn-1 bZIP and homeodomain elements. *Science* **266,** 621–628.

Bronner, G., Chu-LaGraff, Q., Doe, C. Q., Cohen, B., Weigel, D., Taubert, H., and Jackle, H. (1994). Sp1/egr-like zinc-finger protein required for endoderm specification and germ-layer formation in *Drosophila*. *Nature* **369**, 664–668.

Boveri, T. (1893). *Sitzungsber. Ges. Morphol. Physiol.* **8,** 114–125.

Boveri, T. (1899). *Festschrift Kupffer.* (G. Fischer, Jena), pp. 383–430.

Bowerman, B. (1998). Maternal control of pattern formation in early *Caenorhabditis elegans* embryos. *Curr. Top. Dev. Biol.* **39,** 73–117.

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

Bowerman, B., Draper, B. W., Mello, C. C., and Priess, J. R. (1993). The maternal gene *skn-1* encodes a protein that is distributed unequally in early *C. elegans* embryos. *Cell* **74,** 443–452.

Bowerman, B., and Shelton, C. (1999). Cell polarity in the early *Caenorhabditis elegans* embryo. *Curr. Opin. Genet. Dev.* **9,** 390–395.

Britton, C., McKerrow, J. H., and Johnstone, I. L. (1998). Regulation of the *Caenorhabditis elegans* gut cysteine protease gene *cpr-1*: Requirement for GATA motifs. *J. Mol. Biol.* **283,** 15–27.

Brunner, E., Peter, O., Schweizer, L., and Basler, K. (1997). pangolin encodes a Lef-1 homologue that acts downstream of Armadillo to transduce the Wingless signal in *Drosophila*. *Nature* **385**, 829–833.

C. elegans Sequencing Consortium. (1998). Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science* **282,** 2012–2018.

Calvo, D., Victor, M., Gay, F., Sui, G., Luke, M. P.-S., Dufurcq, P., Wen, G., Maduro, M., Rothman, J., and Shi, Y. (2001). A POP-1 repressor complex restricts inappropriate cell type-specific gene transcription during *C. elegans* embryogenesis. *EMBO J.* **20,** 7197–7208.

Carmi, I., Kopczynski, J. B., and Meyer, B. J. (1998). The nuclear hormone receptor SEX-1 is an X-chromosome signal that determines nematode sex. *Nature* **396,** 168–173.

Charron, F., and Nemer, M. (1999). GATA transcription factors and cardiac development. *Semin. Cell Dev. Biol.* **10,** 85–91.

Christensen, S., Kodoyianni, V., Bosenberg, M., Friedman, L., and Kimble, J. (1996). *lag-1*, a gene required for *lin-12* and *glp-1* signaling in *Caenorhabditis elegans*, is homologous to human CBF1 and Drosophila Su(H). *Development* **122,** 1373–1383.

Edgar, L., and McGhee, J. D. (1986). Embryonic expression of a gut-specific esterase in *Caenorhabditis elegans*. *Dev. Biol.* **114**, 109–118.

Edgar, L., and McGhee, J. (1988). DNA synthesis and the control of embryonic gene expression in *C. elegans*. *Cell* **53**, 589–599.

Egan, C. R., Chung, M. A., Allen, F. L., Heschl, M. F., Van Buskirk, C. L., and McGhee, J. D. (1995). A gut-to-pharynx/tail switch in embryonic expression of the *Caenorhabditis elegans ges-1* gene centers on two GATA sequences. *Dev. Biol.* **170,** 397–419.

Ferreira, H. B., Zhang, Y., Zhao, C., and Emmons, S. W. (1999). Patterning of *Caenorhabditis elegans* posterior structures by the Abdominal-B homolog *egl-5*. *Dev. Biol.* **207,** 215–228.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E, and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811.

Fraser, A. G., Kamath, R. S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M., and Ahringer, J. (2000). Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. *Nature* **408,** 325–330.

Freedman, J. H., Slice, L. W., Dixon, D., Fire, A., and Rubin, C. S. (1993). The novel metallothionein genes of *Caenorhabditis elegans*. Structural organization and inducible, cell-specific expression. *J. Biol. Chem.* **268,** 2554–2564.

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

Fukushige, T., Hendzel, M. J., Bazett-Jones, D. P., and McGhee, J. D. (1999). Direct visualization of the *elt-2* gut-specific GATA factor binding to a target promoter inside the living *Caenorhabditis elegans* embryo. *Proc. Natl. Acad. Sci. USA* **96,** 11883–11888.

Gilleard, J. S., and McGhee, J. D. (2001). Activation of hypodermal differentiation in the *Caenorhabditis elegans* embryo by GATA transcription factors ELT-1 and ELT-3. *Mol. Cell. Biol.* **21,** 2533–2544.

Goldstein, B. (1992). Induction of gut in *Caenorhabditis elegans* embryos. *Nature* **357,** 255–257.

Goldstein, B. (2000). Embryonic polarity: A role for microtubules. *Curr. Biol.* **10,** 820–822.

Gomes, J. E., Encalada, S. E., Swan, K. A., Shelton, C. A., Carter, J. C., and Bowerman, B. (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.

Gonczy, P., Echeverri, G., Oegema, K., Coulson, A., Jones, S. J., Copley, R. R., Duperon, J., Oegema, J., Brehm, M., Cassin, E., Hannak, E., Kirkham, M., Pichler, S., Flohrs, K,. Goessen, A., Leidel, S., Alleaume, A. M., Martin, C., Ozlu, N., Bork, P., and Hyman, A. A. (2000). Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. *Nature* **408**, 331–336.

Gotta, M., and Ahringer, J. (2001). Axis determination in *C. elegans*: Initiating and transducing polarity. *Curr. Opin. Genet. Dev.* **11**, 367–373.

Gower, N. J., Temple, G. R., Schein, J. E., Marra, M., Walker, D. S., and Baylis, H. A. (2001). Dissection of the promoter region of the inositol 1,4,5-trisphosphate receptor gene, *itr-1*, in *C. elegans*: A molecular basis for cell-specific expression of IP3R isoforms. *J. Mol. Biol.* **306,** 145–157.

Hanna-Rose, W, and Han, M. (1999). COG-2, a Sox domain protein necessary for establishing a functional vulval-uterine connection in *Caenorhabditis elegans*. *Development* **126**, 169–179.

Herman, M. (2001). *C. elegans* POP-1/TCF functions in a canonical Wnt pathway that controls cell migration and in a noncanonical Wnt pathway that controls cell polarity. *Development* **128,** 581–590.

Hermann, G. J., Leung, B., and Priess, J. R. (2000). Left–right asymmetry in *C. elegans* intestine organogenesis involves a LIN-12/Notch signaling pathway. *Development* **127**, 3429–3440.

Hope, I. A. (1991). "Promoter trapping" in *Caenorhabditis elegans*. *Development* **113**, 399–408.

Horner, M. A., Quintin, S., Domeier, M. E., Kimble, J., Labouesse, M., and Mango, S. E. (1998). *pha-4*, an HNF-3 homolog, specifies pharyngeal organ identity in *Caenorhabditis elegans*. *Genes Dev.* **12**, 1947–1952.

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

Jiang, M., Ryu, J., Kiraly, M., Duke, K., Reinke, V., and Kim, S. K. (2000). Genome-wide analysis of developmental and sex-regulated gene expression profiles in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* **98**, 218–223.

Jiang, L. I., and Sternberg, P. (1999). An HMG1-like protein facilitates Wnt signaling in *Caenorhabditis elegans*. *Genes Dev.* **13**, 877–889.

Kalb, J. M., Lau, K. K., Goszczynski, B., Fukushige, T., Moons, D., Okkema, P. G., and McGhee, J. D. (1998). *pha-4* is *Ce-fkh-1*, a fork head/HNF-3α,β,γ homolog that functions in organogenesis of the *C. elegans* pharynx. *Development* **125,** 2171–2180.

Kennedy, B. P, Aamodt, E. J., Allen, F. L., Chung, M. A., Heschl, M. F., and McGhee, J. D. (1993). The gut esterase gene (*ges-1*) from the nematodes *Caenorhabditis elegans* and *Caenorhabditis briggsae*. *J. Mol. Biol.* **229,** 890–908.

Kent, W. J., and Zahler, A. M. (2000). Conservation, regulation, synteny, and introns in a large-scale *C. briggsae–C. elegans* genomic alignment. *Genome Res.* **10,** 1115–1125.

Kenyon, C. J., Austin, J., Costa, M., Cowing, D. W., Harris, J. M., Honigberg, L., Hunter, C. P., Maloof, J. M., Muller-Immergluck, M. M., Salser, S. J., Waring, D. A., Wang, B. B., and Wrischnik,

L. A. (1997). The dance of the Hox genes: patterning the anteroposterior body axis of *Caenorhabditis elegans*. *Cold Spring Harbor Symp. Quant. Biol.* **62,** 293–305.

Kim, S. K., Lund, J., Kiraly, M., Duke, K., Jiang, M., Stuart, J. M., Eizinger, A., Wylie, B. N., and Davidson, G. S. (2001). A gene expression map for *Caenorhabditis elegans*. *Science* **293,** 2087–2092.

Knight, J. K., and Wood, W. B. (1998). Gastrulation initiation in *Caenorhabditis elegans* requires the function of *gad-1*, which encodes a protein with WD repeats. *Dev. Biol.* **198**, 253–265.

Koh, K., and Rothman, J. H. (2001). ELT-5 and ELT-6 are required continuously to regulate epidermal seam cell differentiation and cell fusion in *C. elegans*. *Development* **128**, 2867–2880.

Korswagen, H. C., and Clevers, H. C. (1999). Activation and repression of wingless/Wnt target genes by the TCF/LEF-1 family of transcription factors. *Cold Spring Harbor Symp. Quant. Biol.* **64,** 141–147.

Laufer, J. S., Bazzicalupo, P., and Wood, W. B. (1980). Segregation of developmental potential in early embryos of *Caenorhabditis elegans*. *Cell* **19,** 569–577.

Laurent, V., Brooks, D. R., Coates, D., and Isaac, R. E. (2001). Functional expression and characterization of the cytoplasmic aminopeptidase P of *Caenorhabditis elegans*. *Eur. J. Biochem.* **268,** 5430–5438.

Lenburg, M. E., and O'Shea, E. K. (1996). Signaling phosphate starvation. *Trends Biochem. Sci.* **21,** 383–387.

Leung, B., Hermann, G. J., and Priess, J. R. (1999). Organogenesis of the *Caenorhabditis elegans* intestine. *Dev. Biol.* **216,** 114–134.

Lin, R., Hill, R. J., and Priess, J. R. (1998). POP-1 and anterior–posterior fate decisions in *C. elegans* embryos. *Cell* **92**, 229–239.

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

Lin, W. H., Huang, L. H., Yeh, J. Y., Hoheisel, J., Lehrach, H., Sun, Y. H., and Tsai, S. F. (1995b). Expression of a *Drosophila* GATA transcription factor in multiple tissues in the developing embryos. Identification of homozygous lethal mutants with P-element insertion at the promoter region. *J. Biol. Chem.* **270,** 25150–25158.

Lowry, J. A., and Atchley, W. R. (2000). Molecular evolution of the GATA family of transcription factors: Conservation within the DNA-binding domain. *J. Mol. Evol.* **50,** 103–115.

MacMorris, M., Spieth, J., Madej, C., Lea, K., and Blumenthal, T. (1994). Analysis of the VPE sequences in the *Caenorhabditis elegans vit-2* promoter with extrachromosomal tandem array-containing transgenic strains. *Mol. Cell. Biol.* **14,** 484–491.

Maduro, M., Meneghini, M., Bowerman, B., Broitman-Maduro, G., and Rothman, J. H. (2001). Restriction of mesendoderm to a single blastomere by the combined action of SKN-1 and a GSK-3β homolog is mediated by MED-1 and -2 in *C. elegans*. *Mol. Cell* **7**, 475–485.

Marzluf, G. A. (1997). Genetic regulation of nitrogen metabolism in the fungi. *Microbiol. Mol. Biol. Rev.* **61,** 17–32.

Mello, C. C., Draper, B. W., Krause, M., Weintraub, H., and Priess, J. R. (1992). The *pie-1* and *mex-1* genes and maternal control of blastomere identity in early *C. elegans* embryos. *Cell* **70,** 163–176.

Mello, C. C., Schubert, C., Draper, B., Zhang, W., Lobel, R., and Priess, J. R. (1996). The PIE-1 protein and germline specification in *C. elegans* embryos. *Nature* **382**, 710–712.

Meneghini, M. D., Ishitani, T., Carter, J. C., Hisamoto, N., Ninomiya-Tsuji, J., Thorpe, C. J., Hamill, D. R., Matsumoto, K., and Bowerman, B. (1999). MAP kinase and Wnt pathways converge to downregulate an HMG-domain repressor in *Caenorhabditis elegans*. *Nature* **399**, 793–797.

Moilanen, L. H., Fukushige, T., and Freedman, J. H. (1999). Regulation of metallothionein gene transcription. Identification of upstream regulatory elements and transcription factors responsible for cell-specific expression of the metallothionein genes from *Caenorhabditis elegans*. *J. Biol. Chem.* **274,** 29655–29665.

Morrisey, E. E., Tang, Z., Sigrist, K., Lu, M. M., Jiang, F., Ip, I. S., and Parmacek, M. S. (1998). GATA6 regulates HNF4 and is required for differentiation of the visceral endoderm in the mouse embryo. *Genes Dev.* **12,** 3579–3590.

Newton, A., Mackay, J., and Crossley, M. (2001). The N-terminal zinc finger of the erythroid transcription factor GATA-1 binds GATC motifs in DNA. *J. Biol. Chem.* **276,** 35794–35801.

Nishiwaki, K., Sano, T., and Miwa, J. (1993). *emb-5*, a gene required for the correct timing of gut precursor cell division during gastrulation in *Caenorhabditis elegans*, encodes a protein similar to the yeast nuclear protein SPT6. *Mol. Gen. Genet.* **239**, 313–322.

Oka, T., Toyomura, T., Honjo, K., Wada, Y., and Futai, M. (2001). Four subunit a isoforms of *Caenorhabditis elegans* vacuolar H+-ATPase. Cell-specific expression during development. *J. Biol. Chem.* **276,** 33079–33085.

Orkin, S. H., and Zon, L. I. (1997). Genetics of erythropoiesis: Induced mutations in mice and zebrafish. *Annu. Rev. Genet.* **31,** 33–60.

Page, B. D., Zhang, W., Steward, K., Blumenthal, T., and Priess, J. R. (1997). ELT-1, a GATA-like transcription factor, is required for epidermal cell fates in *Caenorhabditis elegans* embryos. *Genes Dev.* **11**, 1651–1661.

Priess, J. R., and Thomson, J. N. (1987). Cellular interactions in early *C. elegans* embryos. *Cell* **48,** 241–250.

Priess, J. R., Schnabel, H., and Schnabel, R. (1987). The *glp-1* locus and cellular interactions in the early *C. elegans* embryo. *Cell* **51,** 601–611.

Rehorn, K. P., Thelen, H., Michelson, A. M., and Reuter, R. (1996). A molecular aspect of hematopoiesis and endoderm development common to vertebrates and *Drosophila*. *Development* **122,** 4023–4031.

Reinke, V., Smith, H. E., Nance, J., Wang, J., Van Doren, C., Begley, R., Jones, S. J., Davis, E. B., Scherer, S., Ward, S., and Kim, S. K. (2000). A global profile of germline gene expression in *C. elegans*. *Mol. Cell* **6**, 605–616.

Reiter, J. F., Alexander, J., Rodaway, A., Yelon, D., Patient, R., Holder, N., and Stainier, D. Y. R. (1999). Gata5 is required for the development of the heart and endoderm in zebrafish. *Genes Dev.* **13,** 2983–2995.

Reiter, J. F., Kikuchi, Y., and Stainier, D. Y. (2001). Multiple roles for Gata5 in zebrafish endoderm formation. *Development* **128,** 125–135.

Reuter, R. (1994). The gene serpent has homeotic properties and specifies endoderm versus ectoderm within the *Drosophila* gut. *Development* **120,** 1123–1135.

Rocheleau, C. E., Downs, W. D., Lin, R., Wittmann, C., Bei, Y., Cha, Y. H., Ali, M., Priess, J. R., and Mello, C. C. (1997). Wnt signaling and an APC-related gene specify endoderm in early *C. elegans* embryos. *Cell* **90,** 707–716.

Rocheleau, C. E., Yasuda, J., Shin, T. H., Lin, R., Sawa, H., Okano, H., Priess, J. R., Davis, R. J., and Mello C. C. (1999). WRM-1 activates the LIT-1 protein kinase to transduce anterior/posterior polarity signals in *C. elegans*. *Cell* **97**, 717–726.

Rodaway, A., and Patient, R. (2001). Mesendoderm: An ancient germ layer? *Cell* **105**, 169–172.

Schierenberg, E. (1987). Reversal of cellular polarity and early cell–cell interaction in the embryos of *Caenorhabditis elegans*. *Dev. Biol.* **122,** 452–463.

Schroeder, D. F., and McGhee, J. (1998). Anterior–posterior patterning within the *Caenorhabditis elegans* endoderm. *Development* **125,** 4877–4887.

Seydoux, G., and Fire, A. (1994). Soma-germline asymmetry in the distributions of embryonic RNAs in *Caenorhabditis elegans*. *Development* **120**, 2823–2834.

Seydoux, G., Mello, C. C., Pettitt, J., Wood, W. B., Priess, J. R., and Fire, A. (1996). Repression of gene expression in the embryonic germ lineage of *C. elegans*. *Nature* **382,** 713–716.

Shin, T. H., Yasuda, J., Rocheleau, C. E., Lin, R., Soto, M., Bei, Y., Davis, R. J., and Mello, C. C. (1999). MOM-4, a MAP kinase kinase kinase-related protein, activates WRM-1/LIT-1 kinase to transduce anterior/posterior polarity signals in *C. elegans*. *Mol. Cell* **4**, 475–480.

Shoichet, S. A., Malik, T. H., Rothman, J. H., and Shivdasani, R. A. (2000). Action of the *Caenorhabditis elegans* GATA factor END-1 in *Xenopus* suggests that similar mechanisms initiate endoderm development in ecdysozoa and vertebrates. *Proc. Natl. Acad. Sci. USA* **97,** 4076–4081.

Spieth, J., MacMorris, M., Broverman, S., Greenspoon, S., and Blumenthal, T. (1988). Regulated expression of a vitellogenin fusion gene in transgenic nematodes. *Dev. Biol.* **130,** 285–293.

Stainer, D. Y. R. (2002). A glimpse into the molecular entrails of endoderm formation. *Genes Dev.* **16,** 893–907.

Sulston, J. E., Schierenberg, E., White, J. G., and Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100,** 64–119.

Tabara, H., Hill, R. J., Mello, C. C., Priess, J. R., and Kohara, Y. (1999). *pos-1* encodes a cytoplasmic zinc-finger protein essential for germline specification in *C. elegans*. *Development* **126**, 1–11.

Tabara, H., Motohashi, T., and Kohara, Y. (1996). A multi-well version of *in situ* hybridization on whole mount embryos of *Caenorhabditis elegans*. *Nucleic Acids Res.* **24,** 2119–2124.

Tcherepanova, I., Bhattacharyya, L., Rubin, C. S., and Freedman, J. H. (2000). Aspartic proteases from the nematode *Caenorhabditis elegans*. Structural organization and developmental and cell-specific expression of *asp-1*. *J. Biol. Chem.* **275,** 26359–26369.

Tenenhaus, C., Subramaniam, K., Dunn, M. A., and Seydoux, G. (2001). PIE-1 is a bifunctional protein that regulates maternal and zygotic gene expression in the embryonic germ line of *Caenorhabditis elegans*. *Genes Dev.* **15,** 1031–1040.

Thorpe, C. J., Schlesinger, A., Carter, J. C., and Bowerman, B. (1997). Wnt signaling polarizes an early *C. elegans* blastomere to distinguish endoderm from mesoderm. *Cell* **90,** 695–705.

Thorpe, C. J., Schlesinger, A., and Bowerman, B. (2000). Wnt signalling in *Caenorhabditis elegans*: Regulating repressors and polarizing the cytoskeleton. *Trends Cell Biol.* **10,** 10–17.

Voronov, D. A., and Panchin, Y. V. (1998). Cell lineage in marine nematode *Enoplus brevis*. *Development* **125,** 143–150.

Weber, H., Symes, C. E., Walmsley, M. E., Rodaway, A. R. F., and Patient, R. K. (2000). A role for GATA5 in *Xenopus* endoderm specification. *Development* **127,** 4345–4360.

Wiegner, O., and Schierenberg, E. (1998). Specification of gut cell fate differs significantly between the nematodes *Acrobeloides nanus* and *Caenorhabditis elegans*. *Dev. Biol.* **204,** 3–14.

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

Yuh, C. H., Bolouri, H., and Davidson, E. H. (1998). Genomic *cis*-regulatory logic: Experimental and computational analysis of a sea urchin gene. *Science* **279,** 1896–1902.

Zhong, W., Sladek, F. M., and Darnell, J. E. (1993). The expression pattern of a *Drosophila* homolog to the mouse transcription factor HNF-4 suggests a determinative role in gut formation. *EMBO J.* **12,** 537–544.

Zhu, J., Hill, R. J., Heid, P. J., Fukuyama, M., Sugimoto, A., Priess, J. R., and Rothman, J. H. (1997). *end-1* encodes an apparent GATA factor that specifies the endoderm precursor in *Caenorhabditis elegans* embryos. *Genes Dev.* **11**, 2883–2896.

Received for publication January 7, 2002
Revised March 13, 2002
Accepted March 14, 2002