

# Cell Fate Specification in the *C. elegans* Embryo

Morris F. Maduro\*

Cell specification requires that particular subsets of cells adopt unique expression patterns that ultimately define the fates of their descendants. In *C. elegans*, cell fate specification involves the combinatorial action of multiple signals that produce activation of a small number of “blastomere specification” factors. These initiate expression of gene regulatory networks that drive development forward, leading to activation of “tissue specification” factors. In this review, the *C. elegans* embryo is considered as a model system for studies of cell specification. The techniques used to study cell fate in this species, and the themes that have emerged, are described. *Developmental Dynamics* 239:1315–1329, 2010. © 2010 Wiley-Liss, Inc.

**Key words:** *C. elegans*; cell fate specification; embryo; model systems; gene regulation

Accepted 15 December 2009

## CELL SPECIFICATION AND ITS SIGNIFICANCE

During embryonic development in metazoans, embryonic cells must ultimately assign fates to individual cells and tissues. When a cell has been specified for a particular fate, it is capable of generating those tissues if cultured away from the remainder of the embryo (Wolpert et al., 2007). As all embryonic cells are derived from the zygote by mitosis, cells must acquire differences in gene expression over time. In most animal embryos this is achieved through a combination of inherited maternal determinants, cell–cell interactions, positional information and factors segregated within a lineage from a mother cell to its descendants. In modern developmental biology, the description of gene activities that result in specification constitutes a gene regulatory network, or GRN (Davidson and Erwin, 2006).

Although animal embryos have evolved different ways of specifying very early embryonic cells, the properties of GRNs are similar across many systems (Davidson, 1991; Davidson and Levine, 2008). Hence, the study of cell specification in model systems can be used to exploit the advantages of that system to reveal new mechanistic insights into general GRN properties. In addition, this knowledge might ultimately result in the ability to control cell fates experimentally with only minimal information about the underlying GRNs.

## USE OF *C. ELEGANS* TO STUDY CELL SPECIFICATION

### The *C. elegans* Embryo

The nematode *C. elegans* exhibits a short generation time of 3 days. Embryogenesis occurs in approximately 14 hr at room temperature,

and takes place entirely within an ovoid, chitinous eggshell approximately 50  $\mu\text{m}$  long (Fig. 1). By the end of embryogenesis, the larva has 558 cells, is some 250  $\mu\text{m}$  long and will undergo a series of molts before forming the adult (Brenner, 1974; Sulston et al., 1983). *C. elegans* demonstrates eutely, displaying a constant number of cells from animal to animal, the result of a nearly invariant cell lineage that was described more than 25 years ago (Sulston et al., 1983). All blastomeres at any particular time have a unique history that can be described in terms of the series of cell divisions that gave rise to them from the zygote. The cell lineage, therefore, gives studies of development a highly detailed frame of reference that places every embryonic cell in a defined spatiotemporal context.

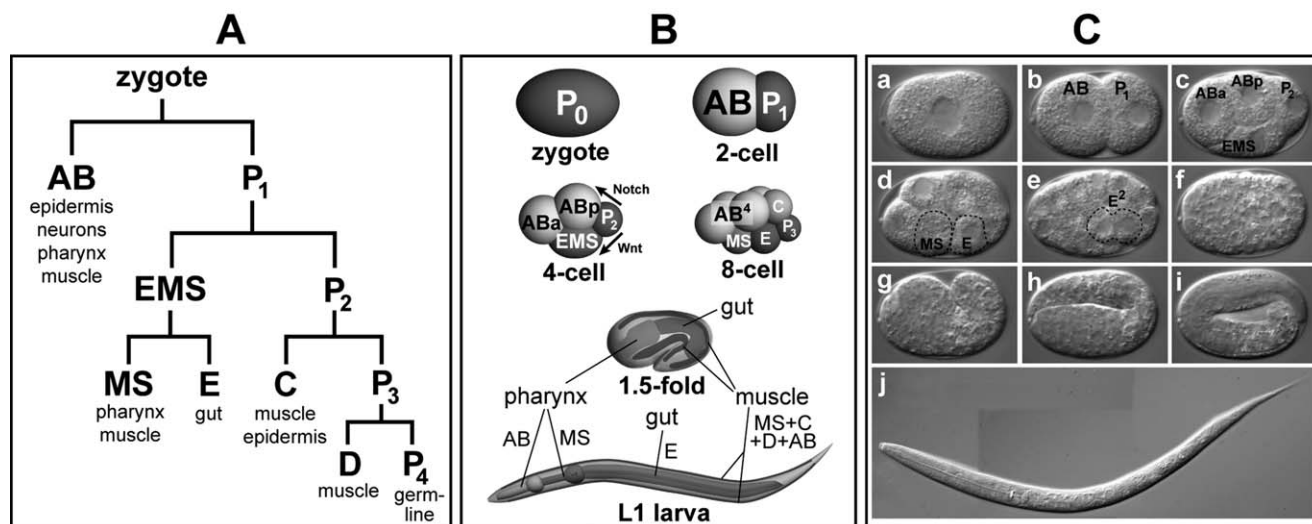
Studies of cell fate specification have focused primarily on two classes of specification event (Table 1;

Department of Biology, University of California, Riverside, Riverside, California  
Grant sponsor: NSF; Grant number: IOS#0643325.

\*Correspondence to: Morris F. Maduro, 2121A Genomics Building, Department of Biology, University of California, Riverside, Riverside, CA 92521. E-mail: mmaduro@ucr.edu

DOI 10.1002/dvdy.22233

Published online 27 January 2010 in Wiley InterScience (www.interscience.wiley.com).



**Fig. 1.** Partial *C. elegans* lineage, diagram and images showing stages of embryogenesis. **A:** The zygote undergoes a set of asymmetric cleavages to generate the six founder cells, AB, MS, E, C, D, and P<sub>4</sub> (Sulston et al., 1983). Anterior cells are to the left. **B:** Diagram of early and mid-embryogenesis stages showing arrangement of blastomeres (first four diagrams) and location of pharynx, muscle, and gut at the 1.5-fold stage and L1 larva. The P<sub>2</sub> inductions of ABp (mediated by Notch/GLP-1 in ABp and Delta/APX-1 in P<sub>2</sub>) and EMS (Wnt/MAPK/Src pathway, abbreviated Wnt for simplicity) are indicated (Priess et al., 1987; Moskowitz et al., 1994; Rocheleau et al., 1997; Thorpe et al., 1997; Rocheleau et al., 1999; Bei et al., 2002). **C:** Stages of embryogenesis shown by differential interference contrast (DIC) micrographs. a–i: Stills from time lapse series (courtesy Bob Goldstein, used with permission). (a) Zygote. (b) Two-cell stage. (c) Four-cell stage. (d) Eight-cell stage. (e) Gastrula showing E daughter cells (E<sup>2</sup>). (f) One-fold stage before start of elongation. (g) The 1.5-fold stage. (h) The two-fold stage. (i) The three-fold elongation. (j) Larva (different embryo than in time lapse images) shortly after hatching. Images in a–f are shown with dorsal at top and anterior to the left; in g and h dorsal is to the bottom. A *C. elegans* embryo is approximately 50  $\mu\text{m}$  along its long axis, and a larva is approximately 250  $\mu\text{m}$  long.

**TABLE 1. Cell Specification Genes in the *C. elegans* Embryo**

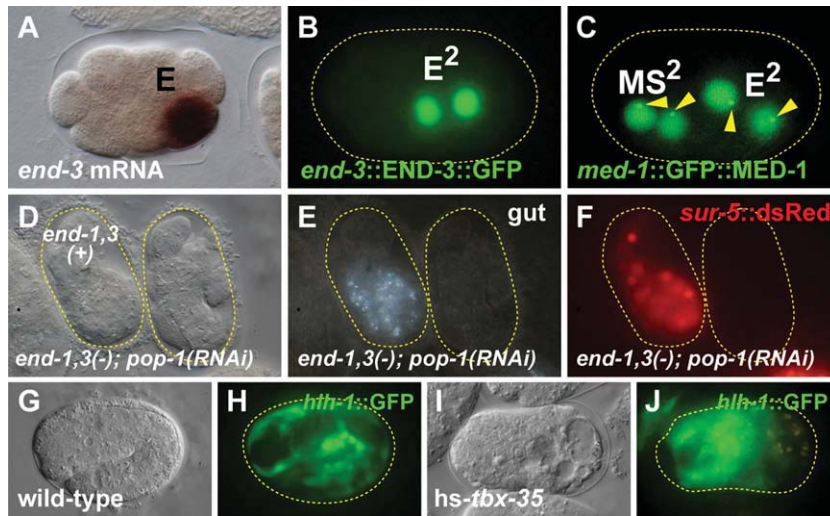
Protein/product	Blastomeres and/or tissues specified	Reference
<i>Blastomere Identity - Maternal</i>		
SKN-1/bZIP-HD	MS, E	(Bowerman et al., 1992)
PAL-1/Caudal	C,D, E (weak contribution)	(Hunter and Kenyon, 1996; Baugh et al., 2005; Maduro et al., 2005b)
PIE-1/CCCH Zinc Finger	P lineage (germline)	(Mello et al., 1992)
<i>Blastomere Identity - Zygotic</i>		
MED-1,2/GATA <sup>a</sup>	MS, E	(Maduro et al., 2001)
END-1,3/GATA	E	(Zhu et al., 1997; Maduro et al., 2005a)
TBX-35/T-box	MS	(Broitman-Maduro et al., 2006)
CEH-51/NK-2 HD	MS	(Broitman-Maduro et al., 2009)
PAL-1/Caudal	C,D	(Hunter and Kenyon, 1996; Baugh et al., 2005)
<i>Tissue Identity</i>		
ELT-2/GATA	Intestine (E)	(Fukushige et al., 1998)
PHA-4/FoxA	Pharynx (AB, MS)	(Horner et al., 1998; Kalb et al., 1998)
HLH-1/MyoD, UNC-120/MRF, HND-1/Hand	Body muscle (AB, MS, C, D)	(Fukushige et al., 2006)
ELT-1/GATA	Epidermis (AB, C)	(Page et al., 1997)
ELT-3/GATA	Epidermis (AB, C)	(Gilleard et al., 1999)

<sup>a</sup>There is some evidence of a weak maternal contribution of the *meds* to E specification, but this has been disputed (Captan et al., 2007; Maduro et al., 2007).

Labouesse and Mango, 1999). The first are the events that specify early progenitors (Priess et al., 1987; Bowerman et al., 1992; Mello et al., 1992; Lin et al., 1995; Hunter and

Kenyon, 1996; Zhu et al., 1997; Maduro et al., 2001, 2005a; Good et al., 2004; Broitman-Maduro et al., 2006). After fertilization, the zygote undergoes a set of holoblastic clea-

vages to generate the six founder cells AB, MS, E, C, D, and P<sub>4</sub>, distinguished by the synchrony of cell cycle timing of their immediate descendants and the similar suite of fates



**Fig. 2.** Analysis of gene expression and perturbations of normal development. **A:** In situ hybridization of a normal embryo with antisense *end-3* probe showing expression in the gut precursor E (Maduro et al., 2007). **B:** Nuclear expression of a translational *end-3::END-3::GFP* fusion in the two daughters of E (Maduro et al., 2005a). **C:** A translational *med-1::GFP::MED-1* fusion is localized to the nuclei of the MS and E daughters and further enriched in small puncta that correspond to an extrachromosomal array carrying the *end-3* promoter (Maduro et al., 2002). **D–F:** *end-1(ok558) end-3(ok1448); pop-1(RNAi)* embryos. The embryo on the left is rescued by an extrachromosomal array that carries wild-type *end-1* and *end-3* and a ubiquitously expressed *sur-5::dsRed* reporter transgene (Yochem et al., 1998). **D:** A differential interference contrast (DIC) image showing terminal phenotype of *pop-1(RNAi)* (left embryo) and *end-1,3; pop-1(RNAi)* (right embryo). **E:** Excess gut, visualized by polarized light, is seen in the left embryo due to endoderm made by both MS and E (Owraghi et al., 2010). In the right-hand embryo, no gut is made due to loss of *end-1* and *end-3*, which are essential for specification of the endoderm precursor E. Both embryos arrest without significant embryo elongation because knockdown of *pop-1* blocks morphogenesis. **F:** Expression of a *sur-5* reporter marks the presence of the rescuing *end-1,3(+)* transgene in the embryo on the left. **G:** Wild-type embryo at the 1.75-fold stage. **H:** Expression of *hjh-1::GFP* in embryo in G. **I:** Terminal phenotype of an embryo in which TBX-35 was expressed before gastrulation. **J:** Expression of *hjh-1::GFP* occurs throughout the embryo (Broitman-Maduro et al., 2006).

produced among descendants of each founder (Fig. 1A) (Sulston et al., 1983). Perturbations that alter these early fates typically cause a transformation of one or more of the founder cells into one of the other founder cells, resulting in the loss or reduction of one tissue type and the excess of another (e.g., Fig. 2D–F). The second type of specification event studied includes those that specify tissue types, irrespective of cell lineage (Page et al., 1997; Fukushige et al., 1998, 2006; Horner et al., 1998; Kalb et al., 1998). These factors are activated among descendants of multiple founder cells, with intestine and germline being the exceptions as they are derived from only E and P4, respectively. Loss of tissue identity factors causes either another tissue type to be made instead (Page et al., 1997; Horner et al., 1998; Fukushige et al., 2006) or a loss of tissue integrity (Fukushige et al., 1998). Recent

work has begun to elucidate connections between the two types of specification events, with the surprising result that tissue identity factors can be directly activated by blastomere factors long before the differentiated tissues are made (Baugh et al., 2005; Lei et al., 2009).

Development of *C. elegans* is mostly autonomous, though cell–cell interactions induce some cells to adopt fates by breaking inherent symmetries, examples of which will be given later. Development is also mosaic: ablation of a cell, or extrusion of cells from the embryo, does not result in compensation by the remaining cells. As a side note, these characteristics are not typical of all nematodes. In *Acroboloides nanus*, embryonic development can show regulation, as ablation of the AB blastomere can result in compensation by the remaining cells (Wiegner and Schierenberg, 1999). As well, in the basal nematode *Enoplus brevis*, no

early specified blastomeres are apparent (Voronov and Panchin, 1998).

The study of cell specification in *C. elegans* has been greatly facilitated by the rich tools and resources available in this system. This includes the ability to do forward and reverse genetics, make transgenics, manipulate blastomeres, visualize all life stages, and make use of the known cell lineage and its completely sequenced genome. What follows are examples of the techniques that have been used for the study of embryonic cell specification mechanisms in this species, followed by a description of specification paradigms that have been found. Finally, a perspectives section contemplates the future of the field.

## TECHNIQUES IN THE *C. elegans* EMBRYO

### Following Cell Lineages

Studying embryos at the level of the cell lineage has been possible because of the transparency of *C. elegans* at all stages. During characterization of wild-type development, lineages were originally followed with repeated observations by differential interference microscopy, aided by some video recordings for early stages (Sulston et al., 1983). Subsequently, time-lapse four-dimensional (4D) microscopy has allowed partial lineages to be constructed retrospectively in single embryos until muscle movements commence at approximately ~400 min after first cleavage, at the 1.5-fold stage (Schnabel et al., 1997; Chisholm and Hardin, 2005). Early technologies used analog laserdisc video of differential interference contrast microscopy (Hird and White, 1993; Schnabel et al., 1997). More recently, computer-assisted fluorescence time-lapse microscopy using transgene reporters has been developed (discussed below) (Murray et al., 2006). In both cases, lineage reconstruction can be facilitated by computer software such as Simi BioCell or StarryNite (Schnabel et al., 1997; Bao et al., 2006; Murray et al., 2006).

Studying cell fate at this level can be important for understanding the nature of changes in cell fate specification that result from experimental perturbation. As mentioned above,

blastomeres that have undergone a change in fate continue to divide but ultimately give rise to different tissues. In the process, the transformed lineage may adopt a cell division pattern characteristic of a new lineage. For example, an apparently complete transformation is visible in the case of deletions that remove *end-1* and *end-3*, a pair of linked GATA factor genes that together specify the endoderm fate (Zhu et al., 1998; Maduro et al., 2005a). In embryos lacking function of both genes, E adopts the fate of C, producing body muscle and epidermis, with a transformation of its lineage to one that is C-like. However, such changed fates do not always show a clear lineage transformation. In *mex-1* mutant embryos, the AB granddaughters adopt MS-like fates, but the transformations in the lineage itself are not as clear (Mello et al., 1992; Schnabel et al., 1996). Presumably, incomplete transformations in mutant backgrounds result from competition among multiple factors in the same lineage, as well as a change in cell–cell interactions when a blastomere adopts a fate outside of its normal context.

Because of the mosaic nature of *C. elegans* embryonic development, in practice it is not necessary to evaluate the cell division pattern of an entire lineage to identify a cell fate transformation. Ectopic activation of early GFP reporters that anticipate the descendants made in a particular lineage can be taken as supportive evidence of a transformation in fate. At the end of embryogenesis, the location and number of ectopically produced cells can be used to suggest a cell fate transformation through the absence of a particular tissue and the excess of another. This is usually done through the light microscope, but transmission electron microscopy has also been used to characterize ectopically mis-specified tissues (Wolf et al., 1983; Bowerman et al., 1992). A more confident assignment of ectopic tissues to an early progenitor can be confirmed by using a laser microbeam to ablate all other blastomeres, as was done to characterize transformed E and MS blastomeres in *end-1,3(-)*, *med-1,2(RNAi)*, and *tbx-35(-)*; *ceh-51(-)* embryos (Zhu et al., 1997; Maduro et al., 2001; Broitman-Maduro et al., 2009). Cell morphology, antibody

staining or green fluorescent protein (GFP) reporters can then be used to identify the descendants produced in the resultant partial embryo.

### Examining the Influence of Cell–Cell Interactions

One of the more elegant techniques in the *C. elegans* embryo is the study of contextual effects on specification by direct manipulation of blastomeres participating in a cell–cell interaction. The study of context-specific effects on cell fate can be achieved without breaching the eggshell, by putting pressure from the outside to nudge cells, as was used to characterize interactions that make ABp different from ABa, and in the study of left–right asymmetry (Priess and Thomson, 1987; Wood, 1991). A laser microbeam can be used to ablate particular cells and prevent cell–cell interactions, as was done to characterize specification within the AB lineage (Gendreau et al., 1994) or to block generation of AB- or MS-derived pharynx (Hunter and Kenyon, 1996). To isolate or block specific cell–cell interactions, cells can be extruded through holes in the eggshell to reduce the number of possible cell contacts in the remaining cells (Priess and Thomson, 1987). Alternatively, particular cell–cell interactions can be reconstructed outside the eggshell, as was done to characterize gut specification (Goldstein, 1992). Cell–cell contacts up to the 150-cell stage were recently characterized systematically in intact embryos (Hench et al., 2009). In that study, approximately 1,500 cell–cell contacts lasting 2.5 min each were identified, and the cell interactions were modeled by computer. This resource, called Ce2008, will likely be very useful for future studies on cell interactions (Hench et al., 2009).

### Forward and Reverse Genetics

In addition to embryo manipulations, cell–cell interactions can also be studied genetically. *C. elegans* possess a hermaphrodite mode of reproduction, which has greatly facilitated the recovery of mutations in developmentally important genes. The first regulators to be identified that affected

early embryonic cell fate specification events were recovered in forward mutagenic screens that looked for maternal-effect embryonic lethal phenotypes exhibiting a deficit of tissues of one type and an excess of another. Chemical or transposon-mediated mutagenesis screens, followed by mapping and molecular cloning, have proven to be good ways of identifying such regulators (Priess et al., 1987; Kempfues et al., 1988). Such screens recovered *glp-1*, which encodes a Notch-like receptor important for asymmetries in the AB lineage; *apx-1*, which encodes a Delta-like ligand that mediates a GLP-1-dependent cell–cell interaction between P<sub>2</sub> and ABp that alters the fate of ABp (Mango et al., 1994); a maternal-specific allele of *pop-1*, which encodes the *C. elegans* TCF-like Wnt effector, important for MS and E fates (Lin et al., 1995; Shetty et al., 2005); and *skn-1*, important for specification of EMS fate (Priess et al., 1987; Bowerman et al., 1992). Similar zygotic screens identified the two endoderm specification genes *end-1* and *end-3*, by large deletions that removed both genes (Zhu et al., 1997; Maduro et al., 2005a), the epidermal specification gene *elt-1* (Page et al., 1997), and *tbx-37* and *tbx-38*, which participate in specification of some ABa descendants (Good et al., 2004).

Factors not accessible by forward mutagenesis, for example due to redundancy or requirement of genes in other processes, can be discovered by other approaches. Regulators can be identified by their expression in an early blastomere lineage. *pal-1*, which encodes a Caudal-like homeoprotein, was first identified by zygotic mutants and later found to play a central role in specification of C and D, the somatic descendants of P<sub>2</sub>, following the discovery that PAL-1 protein is found in the early P<sub>2</sub> and EMS lineages (Hunter and Kenyon, 1996; Hunter et al., 1999). Similarly, *tbx-35* was identified as a candidate MS factor because of its expression in the early MS lineage (Robertson et al., 2004; Broitman-Maduro et al., 2005).

The discovery of RNA interference (Fire et al., 1998), coupled with the completion of the *C. elegans* genome sequence (Consortium, 1998), led to the use of reverse genetics as a

method to identify factors important for cell specification. Components of the Wnt/MAPK signaling pathway that mediate the P<sub>2</sub>-EMS interaction were identified by both genetics and RNAi (Rocheleau et al., 1997; Thorpe et al., 1997). A systematic search of GATA-like transcription factors encoded in the *C. elegans* genome led to the identification of the MED-1,2 divergent GATA factors, which specify MS and participate in specification of E as first characterized by RNAi (Maduro et al., 2001; Goszczynski and McGhee, 2005). More recently, a genome-wide RNAi screen identified *ceh-51*, a gene which functions in MS specification (Broitman-Maduro et al., 2009). Although originally thought to act in muscle specification because of its larval arrest phenotype, the expression of *ceh-51* was found to occur in the early MS lineage, directly downstream of the MS specification factor TBX-35 (Broitman-Maduro et al., 2006, 2009).

As another reverse genetics approach, candidate genes may be selectively targeted for recovery of chromosomal mutations. Gene knockouts made following UV/TMP mutagenesis (Gengyo-Ando and Mitani, 2000) are available from the National Bioresource Project (Mitani laboratory, Japan) and the *Caenorhabditis* Genetics Center (CGC, University of Minnesota, MN). Mutants at the CGC are made by the *C. elegans* Gene Knockout Consortium (Vancouver, British Columbia, Canada, and Oklahoma City, OK; <http://celeganskoconsortium.omrf.org>). Random insertions of transposons, such as Tc1 and Mos1, have also been generated (Zwaal et al., 1993; Granger et al., 2004). Alleles generated in these screens have been used to study chromosomal mutants for MS and E specification factors (Broitman-Maduro et al., 2006, 2009; Maduro et al., 2007). For example, after the T-box factor gene *tbx-35* was identified as a candidate for MS specification because of its early MS lineage expression, a deletion allele, created by reverse genetics, was used to characterize its loss-of-function phenotype (Robertson et al., 2004; Broitman-Maduro et al., 2005, 2006). The chromosomal mutant proved invaluable for identification of *tbx-35* as a critical MS regulator, as RNA interference against *tbx-*

*35* was found to be almost completely ineffective (Broitman-Maduro et al., 2006). RNAi against the related genes *tbx-37* and *tbx-38*, which function together in specification of ABa descendants, was also found to be ineffective (Good et al., 2004), suggesting that early zygotic regulator genes may be generally resistant to RNAi.

### Transgene Reporters

Before the establishment of techniques to make *C. elegans* transgenic, identification of cells or tissue types was achieved primarily by their appearance in the light microscope or by immunohistochemistry (Duerr, 2006; Yochem, 2006). RNA in situ hybridization has also been used to localize transcripts (e.g., Fig. 2A; Seydoux and Fire, 1995; Kohara, 2001; Coroian et al., 2005).

Use of reporter transgenes, by far the most frequently used method to assess gene expression and to mark tissues in *C. elegans*, became standardized through the use of gonadal microinjection, first with lacZ as a reporter, and subsequently GFP and its variants (Fire et al., 1990; Mello et al., 1991; Chalfie et al., 1994). Transgene arrays made this way are generally useful for zygotic genes, but not for maternal genes which frequently undergo silencing due to the repetitive nature of the arrays (Kelly et al., 1997). Problems with expression of maternally expressed transgenes and mosaic expression of zygotic transgenes have been overcome by the use of complex arrays (Kelly et al., 1997), microparticle bombardment (Praitis et al., 2001), or transposon-mediated chromosomal integration (Frokjaer-Jensen et al., 2008). The latter approach, which relies on transgene-dependent repair of a chromosomal Mos1 excision event, is advantageous as it results in a known location of the inserted transgene, favors low copy number insertions, and requires no additional technology to implement in a lab that already has standard microinjection working (Frokjaer-Jensen et al., 2008).

Determining which cells in the early embryo express a given gene is usually a matter of recognizing the expressing cells by their morphology and context (e.g., *end-3* reporter

expression in Fig. 2B, *hll-1* reporter expression in Fig. 2H). For more complex patterns, cell identification can be facilitated by comparison with 4D time lapse recordings (Hope, 1994). Alternatively, mutant backgrounds that transform specific lineages can be used to confirm that expression patterns change appropriately (Maduro et al., 2001). More recently, semi-automated 4D fluorescence microscopy has been used to follow expression of embryonic transgenes simultaneously with the lineage (Murray et al., 2008). In these experiments, the expression of a cell-specific reporter is viewed in a separate fluorescent channel from a ubiquitously expressed nuclear reporter, with the pattern of cell divisions informing the cell identities. Recently, a computational approach has been developed to analyze expression of genes in L1 larvae (Long et al., 2009). In this approach, cells are identified by their positions in the larva, rather than their lineage. Using this method, single-cell resolution expression patterns of 93 reporter genes were used to deduce the points in the embryonic cell lineage that particular tissues were specified by asymmetric cell divisions (Liu et al., 2009). The coupling of fluorescent in vivo reporters with computer-aided image analysis is clearly emerging as a powerful approach for the rapid characterization of factors that are involved in specification of cells or tissues.

Aside from the use of transgene reporters to identify the cells in which a factor is expressed, there are three other applications of transgenes that are worth mentioning. First, factors that are hypothesized to act as “master genes” can be tested outside their normal context through overexpression driven by a heat shock promoter (Stringham et al., 1992; Horner et al., 1998; Zhu et al., 1998; Maduro et al., 2001; Broitman-Maduro et al., 2006; Fukushige et al., 2006). Such constructs permit the conditional activation of the factor of interest in essentially all somatic cells in the embryo (e.g., *hs-tbx-35* embryo in Fig. 2I,J; Stringham et al., 1992). Second, translational GFP fusions can be used to visualize subcellular localization of tagged proteins (e.g., GFP-tagged END-3 in Fig. 2B). This has been

done for proteins that show differential subcellular localization or stability in different cells, such as the  $\beta$ -catenin/asymmetry pathway components TCF/POP-1 (Maduro et al., 2002) and the divergent  $\beta$ -catenin SYS-1 (Huang et al., 2007; Phillips et al., 2007). A third transgene-based technique, the “nuclear spot assay”, takes advantage of the repetitive nature of multicopy transgene arrays. In this approach, a GFP-tagged transcription factor can be localized to the subnuclear region occupied by target arrays that contain its binding sites. For example, this approach was used to show interaction of the endoderm factor ELT-2 with its own promoter (Fukushige et al., 1999) and interaction of POP-1 and MED-1 with the promoters of *end-1* and *end-3* (e.g., GFP::MED-1 subnuclear spots in Fig. 2C; Maduro et al., 2002). In principle, this assay could be used to identify and map transcription factor-DNA interactions in an *in vivo* context, with no prior knowledge of the binding sites.

### Genome-Wide Approaches

Studies of *C. elegans* benefit from a complete genome sequence with thorough annotation by means of the centralized resource WormBase (<http://www.wormbase.org>) (Consortium, 1998; Harris et al., 2010). The genome sequence has greatly facilitated molecular identification of mutated genes and has also been used for multiple other applications. Owing to the ability of *C. elegans* to knock down endogenous gene expression by RNAi through ingested dsRNA (Timmons and Fire, 1998), libraries of bacteria have been constructed that express dsRNA targeted to each gene, permitting genome-wide RNAi-based screens by bacterial feeding (Kamath and Ahringer, 2003; Johnson et al., 2005). A comprehensive list of the entire suite of  $\sim 1,000$  *C. elegans* transcription factors has been compiled, and this information has been used to generate GFP reporters for many of these (Reece-Hoyes et al., 2005, 2007). Yeast two-hybrid studies have been used to identify protein-protein interactions (Boxem et al., 2008). The results of genome-wide screens, expression pattern studies and pro-

tein interactions become available on WormBase soon after publication.

The genome sequence has been invaluable for whole-genome studies of gene expression. An early use of the sequence was in the generation of microarrays, which can be used to identify genes expressed in various tissues such as the germline (Reinke et al., 2000). Microarrays were used to identify transcripts expressed in a small number of precisely staged early embryos (Baugh et al., 2003). To work with such small samples, protocols were developed to linearly amplify transcripts from limiting amounts of RNA (Eberwine, 1996; Baugh et al., 2003). These approaches were used to identify genes that participate in C specification, by isolating mRNA from intact embryos in mutant backgrounds that either created extra C cells or abolished their specification (Baugh et al., 2005). These experiments elucidated a gene network for PAL-1-dependent cell specification of the early C lineage (Baugh et al., 2005). In a similar strategy designed to identify genes that function in a specific tissue, microarrays were used to identify 240 transcripts enriched in the pharynx in similarly made pharynx(+) and pharynx(-) embryos (Gaudet and Mango, 2002).

Other genome-wide techniques are aimed at identifying *cis*-regulatory sites of interest. For example, all promoters, comprising the *C. elegans* “promoterome,” have been cloned and used to generate reporter GFP fusions of a subset of these (Dupuy et al., 2004, 2007). A subset of protein-DNA interactions have been identified by yeast one-hybrid analysis using the promoterome (Deplancke et al., 2006). A recent study reported the use of chromatin immunoprecipitation as a method for verifying direct interaction between PAL-1 and the muscle regulator MyoD/*hlh-1* (Lei et al., 2009). A large, multi-investigator project, mod-ENCODE (model organism ENCyclopedia Of DNA Elements), is working to systematically document all transcription factor-DNA interactions in *C. elegans* and *Drosophila* (Celniker et al., 2009). Information from all of these genome-level projects will likely lead to identification of new genes and protein-DNA interactions that are important for cell specification.

Other approaches can be used to identify lineage-specific transcripts. These include using tissue-specific expression of epitope-tagged Poly-A binding protein to immunoprecipitate transcripts, coupled with microarray analysis or serial analysis of gene expression (SAGE) tagging (Roy et al., 2002; Pauli et al., 2006; McGhee et al., 2009). Newer massively parallel sequencing methods, such as RNA-seq, permit efficient identification of full-length transcripts by coupling the technology with approaches that enrich for mRNAs present in particular lineages or cells (Hillier et al., 2009; McGhee et al., 2009; Wang et al., 2009). The field is just beginning to explore the possibilities with this new technology.

### Biochemical Methods

Biochemical methods have been used to characterize protein-DNA or protein-protein interactions important for cell specification in *C. elegans*. As examples, DNA-binding properties of TCF/POP-1 and interactions of POP-1 with the Nemo-like kinase LIT-1 and the divergent  $\beta$ -catenin WRM-1 were studied in cultured cells (Shin et al., 1999; Korswagen et al., 2000); co-immunoprecipitations from *C. elegans* embryos were used to detect interaction between WRM-1 and LIT-1 (Rocheleau et al., 1999); gel shift analysis and DNaseI footprinting studies have been reported for E and MS specification factors in *C. elegans* (Stroher et al., 1994; Hawkins and McGhee, 1995; Broitman-Maduro et al., 2005, 2009); mass spectrometry was used to identify proteins that can bind to the promoter of *end-1* (Witze et al., 2009); and an NMR solution structure has been solved for the MED-1 DNA-binding domain (Lowry et al., 2009). Therefore, although it can be difficult in practice to obtain biochemical quantities of *C. elegans* embryos (particularly those of a specific stage), this has not been a severe limitation to doing biochemical experiments.

### Comparative Studies With Other Nematodes

Approximately a dozen species have been identified in *Caenorhabditis* (Kiontke et al., 2004). Embryogenesis

is found to be very similar between *C. briggsae* and *C. elegans* (Zhao et al., 2008). However, there are RNAi-induced phenotype differences between these two species with respect to the orthologues of *skn-1* and *pop-1* (Lin et al., 2009), suggesting that the underlying gene networks have undergone some changes. Hence, comparative studies of embryo cell specification among *Caenorhabditis* species may be a way to probe the flexibility of specification mechanisms that produce the same output, similar to what has been done postembryonically in the vulval lineages (Felix, 2007). Genome sequences are available for different nematode species, and comparisons among noncoding regions have generated cisRED, a compendium of conserved putative cis-regulatory sites (Sleumer et al., 2009). RNAi by bacterial feeding is also possible in *C. briggsae* (Winston et al., 2007); this technique was used in examining knockdown of *skn-1* and *pop-1* (Lin et al., 2009), and it should be possible to perform genome-wide RNAi screens to identify cell fate genes in *C. briggsae*. A recent comparative study described differences in very early embryonic development, before cell specification events take place, among multiple nematode species in different genera (Brauchle et al., 2009). Comparisons of cell specification events can therefore be useful both from an evolutionary standpoint, and to inform studies of *C. elegans* development.

The *C. elegans* system clearly offers a rich set of tools for the study of cell fate specification pathways, and newer approaches that exploit massively parallel sequencing and computer-aided analysis of expression patterns will almost certainly revolutionize the next phase of work in this field. Many themes of development have already emerged from cell specification studies in *C. elegans*, as described below.

## THEMES OF CELL FATE SPECIFICATION IN THE *C. elegans* EMBRYO

*C. elegans* development proceeds through a lineage specification phase and an organ/tissue specification

phase (Labouesse and Mango, 1999). This suggests that there should be two broad classes of cell specification genes: Those that act in the early embryo in individual blastomeres to specify lineage progenitors, irrespective of the tissue types made, and those that act later, in multiple lineages, to specify the progenitors of particular tissues. Studies over the past two decades have identified regulators of both types, and as mentioned earlier, recent work has begun to connect them together. The results of some of these studies will be described in brief below to illustrate examples of developmental themes.

### Asymmetric Maternal Determinants

Asymmetrically localized maternal determinants are important to blastomere specification in *C. elegans*. The anterior–posterior axis of the embryo is set by the point of sperm entry (Gotta and Ahringer, 2001), and a network of gene products, including products of the *par* genes, links this information to the asymmetric expression of maternal determinants (Bowerman, 1998; Gonczy and Rose, 2005). The blastomere-specific activity of cell specification factors occurs by a combination of specific translation of maternal mRNAs, differential activity and differential stability of factors. For example, the maternally contributed EMS specification factor SKN-1, which encodes a homeodomain/bZIP transcription factor, is enriched in P<sub>1</sub> and its daughters P<sub>2</sub> and EMS (Bowerman et al., 1992, 1993; Seydoux and Fire, 1994). SKN-1 is also present in the daughters of P<sub>2</sub> (P<sub>3</sub> and C) and EMS (MS and E; Fig. 1). The CCCH zinc finger proteins PIE-1, MEX-1, and OMA-1 are involved in restricting SKN-1 activity to the EMS lineage. Loss of function of *pie-1* or *mex-1*, or a gain of function mutation in *oma-1*, results in ectopic appearance of mesoderm and/or endoderm, the tissues that are specified by SKN-1 (Mello et al., 1992; Lin, 2003; Shirayama et al., 2006).

### Cell Induction Events

The lineage gives the appearance of an autonomous system of fate deter-

mination. However, several cell induction events are required to specify particular fates, and hence patterns of gene expression, in the early embryo (Gonczy and Rose, 2005). Two of these involve the four-cell stage blastomere P<sub>2</sub>, which provides an inductive signal to both ABp and EMS, the two cells it contacts (Fig. 1B). In the first interaction, the sister cells ABa and ABp are initially equivalent, but the interaction of ABp with P<sub>2</sub> breaks this equivalence (Priess et al., 1987; Priess and Thomson, 1987; Moskowitz et al., 1994). This induction is dependent on the Notch signaling pathway: the AB cells express maternal GLP-1/Notch protein, and P<sub>2</sub> expresses APX-1/Delta (Evans et al., 1994; Mango et al., 1994; Mickey et al., 1996). A key result of the P<sub>2</sub>-ABp interaction is the repression of *tbx-37* and *tbx-38* in the ABp lineage, so that their activation is restricted to the ABa lineage (Good et al., 2004). GLP-1/Notch signaling is involved in other cell inductions in the early embryo, including an interaction between MS and ABa descendants that induces formation of the anterior portion of the pharynx (Priess et al., 1987; Moskowitz et al., 1994).

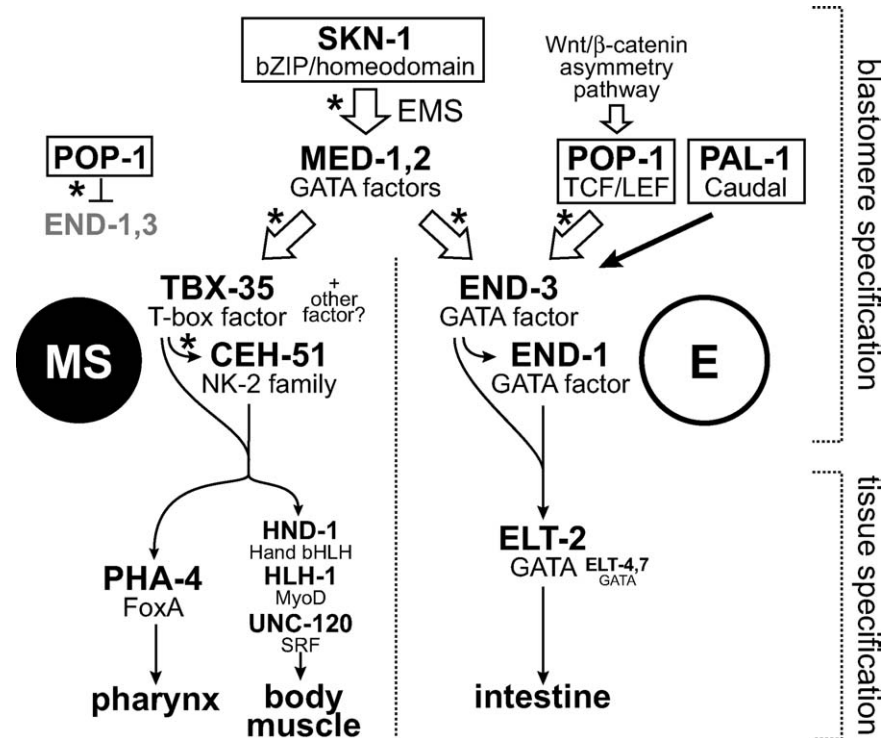
In the second P<sub>2</sub>-dependent cell–cell interaction, P<sub>2</sub> polarizes its sister cell EMS, such that the EMS spindle rotates to become aligned with the anterior–posterior axis, and the side of EMS in contact with P<sub>2</sub> will generate the endoderm precursor E when EMS divides (Goldstein, 1992; Schlesinger et al., 1999). The default state of the daughters of EMS is the MS fate, producing mesoderm (body muscles, pharynx, and other cells). Overlapping Wnt/MAPK/Src signals are used in the P<sub>2</sub>-EMS interaction (Rocheleau et al., 1997, 1999; Thorpe et al., 1997, 2000; Schlesinger et al., 1999; Bei et al., 2002). Loss of function of these components leads to an E to MS transformation. The nuclear effector TCF/POP-1 is required primarily for repression of endoderm fate in MS, as loss of *pop-1* function leads to the reverse transformation, MS to E (Lin et al., 1995). Though initially characterized in the MS-E decision, the Wnt/β-catenin asymmetry pathway was found to be involved in several asymmetric cell divisions, both in the embryo and in many postembryonic

cell divisions (Kaletta et al., 1997; Lin et al., 1998; Mizumoto and Sawa, 2007; Phillips et al., 2007). Recurrent “Wnt asymmetry” has been found in a polychaete annelid, suggesting this mechanism for making sister cells different may be widespread in protozoans (Schneider and Bowerman, 2007).

Although the known cell–cell interactions involve cells in direct contact, a recent study suggests that a relay mechanism, operating over many cell diameters, orients cell divisions along the anteroposterior axis (Bischoff and Schnabel, 2006). As Wnt/ $\beta$ -catenin asymmetry operates in cell divisions that occur along this axis, such a global mechanism would play an important role in orienting dividing cells to access this system for the correct establishment of asymmetric fates.

### Autonomous Lineage Specification

Lineage specification factors are active very early in embryogenesis and are expressed only transiently, passing control of lineage identity to another tier of regulators. The specification of MS and E is an example of a combinatorial control mechanism in which two maternal pathways converge to result in lineage-specific activation of blastomere-specific factors (Fig. 3). Maternal SKN-1 specifies the descendants of EMS as endomesodermal, and the input of the Wnt/ $\beta$ -catenin asymmetry pathway, through the nuclear effector POP-1, makes the descendants of EMS different from each other (Maduro and Rothman, 2002). A pair of zygotic target genes of SKN-1, *med-1* and *med-2*, are activated in EMS in response to SKN-1 (Maduro et al., 2001). MED-1,2, divergent members of the GATA transcription factor family, are required to specify MS, and they also contribute to specification of E in parallel with SKN-1, Wnt-signaled POP-1, and a small but detectable input from PAL-1 (Maduro et al., 2001, 2005b; Goszczynski and McGhee, 2005; Shetty et al., 2005; Broitman-Maduro et al., 2009). The result of these inputs is the activation of T-box factor gene *tbx-35* in MS, and the endoderm GATA factor genes *end-1* and *end-3* in E (Zhu et al., 1997; Maduro et al., 2005a; Broitman-



**Fig. 3.** The abbreviated endomesoderm gene regulatory network as an example of a *C. elegans* gene regulatory network for cell specification. Maternal factors (boxed) act in combinations to activate zygotic blastomere specification factors, which in turn leads to activation of tissue specification factors. SKN-1 activates zygotic *med-1,2*, which in MS activates *tbx-35* (Maduro et al., 2001; Broitman-Maduro et al., 2006). TBX-35 and another unknown factor activate *ceh-51* (Broitman-Maduro et al., 2009). The combination of TBX-35 and CEH-51 activities specifies MS identity. Farther downstream, the pharynx organ identity factor PHA-4, and the muscle identity network consisting of Hand/HND-1, MyoD/HLH-1, and SRF/UNC-120 are activated (Horner et al., 1998; Kalb et al., 1998; Fukushige et al., 2006). In MS, POP-1 represses the endoderm specification genes *end-1* and *end-3* (Lin et al., 1995; Maduro et al., 2002; Shetty et al., 2005). In E, POP-1 is modified by the Wnt/ $\beta$ -catenin asymmetry pathway so that it now becomes an activator in conjunction with the  $\beta$ -catenin SYS-1 (Maduro et al., 2005b; Shetty et al., 2005; Huang et al., 2007; Phillips et al., 2007). PAL-1, SKN-1 and the MEDs also contribute to activation of *end-1* and *end-3*, and there is evidence that END-3 also contributes to *end-1* activation (Maduro et al., 2001, 2005b, 2007). Downstream of E specification, the GATA factor ELT-2 (with possible contributions by ELT-4 and ELT-7) specifies the intestinal fate (McGhee et al., 2009). The blastomere specification factors are expressed only transiently, for as little as one or two cell cycles, while the tissue identity factors are expressed throughout the lifespan of the animal. Asterisks denote regulatory connections which are known to be direct. For simplicity, the model presented here leaves out other tissues made by MS and other factors that restrict SKN-1 activity to EMS; for additional information, the reader is referred to more comprehensive versions of this network (Maduro, 2009; Owrighi et al., 2010).

Maduro et al., 2006). In MS, POP-1 (in its default, unsignaled state) blocks endoderm specification by directly repressing *end-1* and *end-3* (Maduro et al., 2002, 2005b, 2007; Shetty et al., 2005). In E, this repression is relieved by the Wnt-dependent modification of POP-1, which lowers its nuclear levels and allows it to form a bipartite activator with the divergent  $\beta$ -catenin SYS-1 (Huang et al., 2007; Phillips et al., 2007). The major role of POP-1 in MS appears to be repression of *end-1* and *end-3*, as triple mutant *pop-1; end-1 end-3*

embryos appear to specify MS (Owrighi et al., 2010). Once activated, the zygotic MS and E specification factors specify these cells autonomously. Consistent with this, when overexpressed outside of the EMS context, ectopically expressed TBX-35 and END-1/3 are able to reprogram early blastomeres into mesodermal or endodermal precursors (Fig. 2I,J; Zhu et al., 1998; Maduro et al., 2005a; Broitman-Maduro et al., 2006).

In a second example of blastomere specification, the C cell is specified by maternal PAL-1, which activates



zygotic *pal-1* in the C lineage and activates a suite of genes to specify C-derived tissues (Hunter and Kenyon, 1996; Baugh et al., 2005). In the C specification network, zygotic PAL-1 activates unique sets of transcription factor targets to specify epidermal or muscle progenitors. These include a pair of redundant T-box factors, *tbx-8* and *tbx-9*, and the GATA factor gene *elt-1* to specify ectoderm, and the muscle specification network, consisting of Hand/HND-1, SRF/UNC-120, and MyoD/HLH-1 (Page et al., 1997; Andachi, 2004; Baugh et al., 2005; Fukushige et al., 2006), discussed below. The activation of the PAL-1 targets was observed to occur sequentially in temporal phases (Baugh et al., 2005); indeed, the zygotic transcriptome of the embryo appears to change dynamically in such temporal waves, suggesting that tiers of regulatory cascades, operating at each cell cycle, are a feature of most zygotic cell specification events (Baugh et al., 2003).

### Tissue Specification Factors

The second major class of cell specification factors determines the identity of individual organs and/or tissues. Unlike lineage identity factors, which are expressed transiently, tissue identity factors are expressed both in progenitors and through adulthood. Endoderm (or gut) identity is specified by the GATA factor ELT-2, immediately downstream of the E identity factors END-1 and END-3 (Fukushige et al., 1998; Maduro et al., 2005a; McGhee et al., 2009). ELT-2 is at the top of an endoderm gene network that includes thousands of genes (Pauli et al., 2006; McGhee et al., 2007). In the AB and C lineages, the GATA factor ELT-1 specifies epidermal identity, perhaps in conjunction with the related factor ELT-3 (Page et al., 1997; Gilleard and McGhee, 2001). Body muscle cells, made primarily in the C, D and MS lineages, are specified by the combined activity of Hand/HND-1, MyoD/HLH-1, and SRF/UNC-120 (Baugh and Hunter, 2006; Fukushige et al., 2006). The myogenic potential of HLH-1 is determined in part by activity of POP-1, showing that tissue identity factors also work in combinatorial interactions similar

to blastomere identity factors (Fukushige and Krause, 2005). In the MS and AB lineages, pharynx identity is determined by the FoxA/PHA-4 factor, which activates hundreds of genes in the developing pharynx (Horner et al., 1998; Kalb et al., 1998; Gaudet and Mango, 2002).

Within tissues, additional gene networks must be activated that specify unique domains within the territory. In the pharynx, multiple types of tissues are specified during organogenesis, as the pharynx consists of neurons, epithelial and muscle cells, among others (Mango, 2007). PHA-4 specifies these cells in combination with other factors. To specify the muscle cells that are made in the pharynx, PHA-4 works with the CEH-22 homeoprotein to activate the pharynx-specific myosin gene *myo-2* (Okkema and Fire, 1994; Vilimas et al., 2004). In the gland cells, a combinatorial code of factors, including PHA-4, activates the gland-specific gene *hlh-6* (Raharjo and Gaudet, 2007). Hence, within a single organ or tissue, additional specification networks specify particular cell types. Such tissue factors may also maintain the differentiation state of the tissue, and even activate different genes in response to the environment after development (Fukushige et al., 1998; McGhee et al., 2007).

### Redundancy

One theme that has emerged during the study of cell specification is that multiple gene products, with overlapping functions, can participate in the same event. The pairs of regulators MED-1,2, END-1,3, and TBX-37,38 are good examples of redundant function in development of the MS, E and ABa lineages, respectively (Maduro et al., 2001, 2005a; Good et al., 2004). While these paralogous pairs of factors are likely the result of gene duplications, there are other examples of overlapping function that result from convergent activities of unrelated factors. In MS specification, TBX-35 and the homeoprotein CEH-51 make similar contributions, and *ceh-51* is also a direct target of *tbx-35* (Broitman-Maduro et al., 2009). In C specification, multiple factors function in parallel downstream of PAL-1 (Baugh

et al., 2005). The P<sub>2</sub>-EMS interaction involves overlapping Wnt/MAPK/Src pathways (Rocheleau et al., 1997; Thorpe et al., 1997; Meneghini et al., 1999; Bei et al., 2002). Zygotic activation of endoderm specification in E is the result of input from MED-1,2, SKN-1, Wnt-signaled POP-1 and PAL-1 (Maduro et al., 2005b, 2007; Huang et al., 2007; Phillips et al., 2007). In these examples involving E, a fully penetrant loss of endoderm specification is generally apparent only when multiple factors are knocked down simultaneously (Bei et al., 2002; Maduro et al., 2005a, 2007). In tissue specification, muscle identity results from the convergent activities of HND-1, UNC-120, and HLH-1, as only knockdown or mutation of all three together can abolish specification of muscle (Fukushige et al., 2006).

It is not clear why this type of redundancy exists. It may be adaptive to have multiple genes specify lineages to ensure developmental robustness under different environmental conditions, or perhaps different gene batteries were recruited into common lineages over time. For redundancy involving multiple paralogs of an ancestral gene, duplications may be stabilized by some other selective force, such as acquisition of new functions or mutual loss of subsets of the original function or expression (Force et al., 1999). For example, *med-1*-like genes have undergone extensive duplications in other *Caenorhabditis* species (Coroian et al., 2005; Lowry et al., 2009), while *end-3* has some expression and functional differences with *end-1* (Maduro et al., 2005a).

Redundancy is also evident in embryo morphogenesis. Many cell specification factors were identified by mutagenic screens that presupposed embryonic lethality as a consequence of blastomere mis-specification. However, there is at least one example of a cell specification mutant that displays both embryonic and larval lethality: Double mutants of *end-1* and *end-3*, while lacking all endoderm, can still undergo a fairly normal morphogenesis and generate elongated larvae that can hatch (Owraghi et al., 2010). Hence, it seems likely that multiple, parallel pathways contribute to the

morphogenesis of the embryo besides those that specify blastomere identity, suggesting that other cell specification factors could be identified as those for which a mutant phenotype results in larval lethality or even viability.

### The Use of Developmental Regulators in Other Contexts

As in many other systems, early *C. elegans* developmental regulators have been found to have unrelated functions that occur after their roles in specification. This is especially true of the maternal factors. The Wnt/ $\beta$ -catenin asymmetry pathway and POP-1 are used in multiple asymmetric divisions, as mentioned earlier (Kaletta et al., 1997; Lin et al., 1998; Mizumoto and Sawa, 2007). The EMS specification factor SKN-1 has a more ancient role in the differentiated intestine in response to stress (An and Blackwell, 2003; An et al., 2005). The C specification factor PAL-1 has later roles in development of the male tail (Hunter et al., 1999; Edgar et al., 2001). In contrast, additional roles for the zygotic specification factors *med-1, 2, end-1, 3, tbx-35, tbx-37, 38,* and *ceh-51* have yet to be defined. This may be related to how recently these genes were recruited into cell specification. A genome comparison of *C. elegans* with nematodes outside the *Caenorhabditis* genus found that the early zygotic regulators were the least likely to be recognizable in other genomes, while the maternal and tissue specification factors were more conserved (Maduro, 2006). Hence, these genes may not have yet acquired other functions, or else their later roles are subtle and have escaped identification.

### Conservation of Mechanisms Across Species

Although *C. elegans* development is considered to be derived relative to other nematodes, as a result of optimization for rapid embryogenesis (Schierenberg, 2001), there are multiple examples of the similar use of factors for similar specification events, suggesting that core features have been conserved through evolution. Endomesoderm in vertebrates is speci-

fied by GATA factors, similar to the specification of MS/E by the MED-1,2 GATA factors (Reiter et al., 1999; Maduro et al., 2001; Patient and McGhee, 2002). In *Drosophila*, the GATA factor Serpent specifies endoderm, similar to the END-1,3 and ELT-2 GATA factors in *C. elegans* (Reuter, 1994; Fukushige et al., 1998; Maduro et al., 2005a). T-box and NK-2 homeodomain factors (similar to the MS specification factors TBX-35 and CEH-51, respectively) also act in mesoderm development in *Drosophila* and vertebrates (Lyons et al., 1995; Horb and Thomsen, 1999; Reim and Frasch, 2005). Body muscle specification by MyoD/HLH-1 and Hand/HND-1 is similar to the use of the bHLH factors MyoD and Hand in specification of different muscle types in vertebrates (Srivastava, 1999; Tapscott, 2005). As well, the regulatory network formed by the endomesoderm and C lineage specification genes has features similar to those in other systems, such as feed forward loops, autoregulation, and multiple input motifs (Baugh et al., 2005; Maduro, 2006; Swiers et al., 2006). The *C. elegans* system, therefore, provides a way to learn about processes of specification that are common to all metazoans.

## PERSPECTIVES

### Building a Complete Network

The goal of embryonic cell specification studies in *C. elegans*, following the lead of work in the sea urchin *Strongylocentrotus purpuratus* (Davidson et al., 2002) and other fields, is the generation of a comprehensive gene regulatory network that connects all transcription factor and cell-cell signaling events from zygote to larva. A comprehensive network would itself be an immense achievement, and it would lead to other applications, such as mathematical modeling, and there may also be new paradigms of cell specification to be discovered. To date, there is a fair bit of information about key regulators that are important for both blastomere specification and tissue specification, but there is a knowledge gap in the linkage between these, and the connection of these factors with the generation of the complex lineage pat-

terns and cellular morphogenesis within each tissue type. How do cells adopt reproducible positions in the embryo? How are cell movements in morphogenesis tied to cell specification? Do these later lineage specification events use similar mechanisms as those that happen early?

Construction of an exhaustive gene regulatory network requires spatiotemporal profiling of all transcription factors, perturbation of these, and characterization of the effects that these have on expression of all other regulators (Materna and Oliveri, 2008; Li and Davidson, 2009). Thus far, building the *C. elegans* embryo specification network has relied primarily on mutations identified through open-ended screening, or by identification of candidate genes based on expression patterns. Mutational screens have been biased toward factors that play more essential roles. Phenotypes of genes acting later in the cell lineage may exhibit only larval lethality, or even be viable, because they may affect only a small number of cells. As well, specification events that involve parallel functions of redundant genes, or events that involve factors that have both early and late functions, might be missed entirely without conditional alleles. Hence, a candidate gene approach that starts with known expression patterns may ultimately be the best way to build a more comprehensive network. Newer technologies aimed at identifying promoter-DNA interactions, and the expression of all transcription factors, will likely provide a data set that can be mined for cell specification gene candidates. Without a full perturbation analysis, however, these studies are still only a starting point. One potential technical limitation to a large-scale molecular characterization of expression is the difficulty of obtaining large numbers of synchronized embryos; newer technologies are emerging that allow characterization of gene expression profiles on small samples with high sensitivity (Geiss et al., 2008). Given the rate of technological advances, it is likely that a description of transcription at the level of single cells, describing all blastomeres in the *C. elegans* embryo, will soon be possible at modest cost.

## Knowledge Gaps: What Is Left to Find?

Aside from the completion of the gene network, there continue to be questions about cell fate specification in *C. elegans*. First, there is no known set of regulators that dictates neuronal identity, analogous to the regulators identified for muscle, endoderm, epidermis and pharynx. RNAi knock-down of the histone acetyltransferase gene *cbp-1* results in a failure of the embryo to generate the aforementioned tissue types, but neurons continue to be made, suggesting that neuron fate may be a default condition when no other fate has been specified (Shi and Mello, 1998; Hobert, 2005). Many genes have been identified that are required for specification of a subset of neurons, or even just asymmetric divisions that lead to neurons; however, it may be that a complex set of regulators works together and has simply not yet been identified, analogous to the muscle specification network (Finney and Ruvkun, 1990; Hallam et al., 2000; Frank et al., 2003; Fukushige et al., 2006). Just as a neuronal identity factor is not apparent, it is similarly not clear what factor(s) may specify AB fate, as blastomere identity factors are known for all the other founder cells. It may be that because AB generates the majority of neurons, there is no "AB factor" to be found, only factors that are required to break symmetrical divisions through cell-cell interactions (Sulston et al., 1983; Moskowitz et al., 1994; Good et al., 2004).

Whereas much is known about cell specification in the *C. elegans* embryo, some questions are wide open for study. Do miRNAs or other noncoding RNAs play roles in lineage specification? How is the cell specification network robust to environmental changes? And given that there are other ways that nematodes have evolved early embryo specification mechanisms (Voronov and Panchin, 1998; Goldstein, 2001; Schulze and Schierenberg, 2009), what is inherently special about the *C. elegans* gene network, and what is flexible? Future work, with new tools and questions, will no doubt make many more contributions to this area.

## ACKNOWLEDGMENTS

I thank Bob Goldstein for permission to use video stills from one of the time-lapse movies made in his laboratory for Figure 1. Work in the Maduro lab is currently funded by NSF.

## REFERENCES

- An JH, Blackwell TK. 2003. SKN-1 links *C. elegans* mesendodermal specification to a conserved oxidative stress response. *Genes Dev* 17:1882–1893.
- An JH, Vranas K, Lucke M, Inoue H, Hisamoto N, Matsumoto K, Blackwell TK. 2005. Regulation of the *Caenorhabditis elegans* oxidative stress defense protein SKN-1 by glycogen synthase kinase-3. *Proc Natl Acad Sci U S A* 102:16275–16280.
- Andachi Y. 2004. *Caenorhabditis elegans* T-box genes *tbx-9* and *tbx-8* are required for formation of hypodermis and body-wall muscle in embryogenesis. *Genes Cells* 9:331–344.
- Bao Z, Murray JI, Boyle T, Ooi SL, Sandel MJ, Waterston RH. 2006. Automated cell lineage tracing in *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A* 103:2707–2712.
- Baugh LR, Hunter CP. 2006. MyoD, modularity, and myogenesis: conservation of regulators and redundancy in *C. elegans*. *Genes Dev* 20:3342–3346.
- 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.
- Baugh LR, Hill AA, Claggett JM, Hill-Harfe K, Wen JC, Slonim DK, Brown EL, Hunter CP. 2005. The homeodomain protein PAL-1 specifies a lineage-specific regulatory network in the *C. elegans* embryo. *Development* 132:1843–1854.
- Bei Y, Hogan J, Berkowitz LA, Soto M, Rocheleau CE, Pang KM, Collins J, Mello CC. 2002. SRC-1 and Wnt signaling act together to specify endoderm and to control cleavage orientation in early *C. elegans* embryos. *Dev Cell* 3:113–125.
- Bischoff M, Schnabel R. 2006. A posterior centre establishes and maintains polarity of the *Caenorhabditis elegans* embryo by a Wnt-dependent relay mechanism. *PLoS Biol* 4:e396.
- Bowerman B. 1998. Maternal control of pattern formation in early *Caenorhabditis elegans* embryos. *Curr Top Dev Biol* 39:73–117.
- 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.
- 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.
- Boxem M, Maliga Z, Klitgord N, Li N, Lemmens I, Mana M, de Lichtervelde L, Mul JD, van de Peut D, Devos M, Simonis N, Yildirim MA, Cokol M, Kao HL, de Smet AS, Wang H, Schlaitz AL, Hao T, Milstein S, Fan C, Tipsword M, Drew K, Galli M, Rhrissorrakrai K, Drechsel D, Koller D, Roth FP, Iakoucheva LM, Dunker AK, Bonneau R, Gunsalus KC, Hill DE, Piano F, Tavernier J, van den Heuvel S, Hyman AA, Vidal M. 2008. A protein domain-based interactome network for *C. elegans* early embryogenesis. *Cell* 134:534–545.
- Brauchle M, Kiontke K, MacMenamin P, Fitch DH, Piano F. 2009. Evolution of early embryogenesis in rhabditid nematodes. *Dev Biol* 335:253–262.
- Brenner S. 1974. The genetics of *Caenorhabditis elegans*. *Genetics* 77:71–94.
- 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.
- 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.
- Broitman-Maduro G, Owraghi M, Hung W, Kuntz S, Sternberg PW, Maduro M. 2009. The NK-2 class homeodomain factor CEH-51 and the T-box factor TBX-35 have overlapping function in *C. elegans* mesoderm development. *Development* 176:2735–2746.
- Captan VV, Goszczynski B, McGhee JD. 2007. Neither maternal nor zygotic *med-1/med-2* genes play a major role in specifying the *Caenorhabditis elegans* endoderm. *Genetics* 175:969–974.
- Celniker SE, Dillon LA, Gerstein MB, Gunsalus KC, Henikoff S, Karpen GH, Kellis M, Lai EC, Lieb JD, MacAlpine DM, Micklem G, Piano F, Snyder M, Stein L, White KP, Waterston RH. 2009. Unlocking the secrets of the genome. *Nature* 459:927–930.
- Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC. 1994. Green fluorescent protein as a marker for gene expression. *Science* 263:802–805.
- Chisholm AD, Hardin J. 2005. Epidermal morphogenesis. *WormBook*: 1–22.
- Consortium CeS. 1998. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 282:2012–2018.
- 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.
- Davidson EH. 1991. Spatial mechanisms of gene regulation in metazoan embryos. *Development* 113:1–26.
- Davidson EH, Erwin DH. 2006. Gene regulatory networks and the evolution of animal body plans. *Science* 311:796–800.
- Davidson EH, Levine MS. 2008. Properties of developmental gene regulatory

- networks. *Proc Natl Acad Sci U S A* 105:20063–20066.
- Davidson EH, Rast JP, Oliveri P, Ransick A, Caletani C, Yuh CH, Minokawa T, Amore G, Hinman V, Arenas-Mena C, Otim O, Brown CT, Livi CB, Lee PY, Revilla R, Schilstra MJ, Clarke PJ, Rust AG, Pan Z, Arnone MI, Rowen L, Cameron RA, McClay DR, Hood L, Bolouri H. 2002. A provisional regulatory gene network for specification of endomesoderm in the sea urchin embryo. *Dev Biol* 246:162–190.
- Deplancke B, Mukhopadhyay A, Ao W, Elewa AM, Grove CA, Martinez NJ, Sequerra R, Doucette-Stamm L, Reece-Hoyes JS, Hope IA, Tissenbaum HA, Mango SE, Walhout AJ. 2006. A gene-centered *C. elegans* protein-DNA interaction network. *Cell* 125:1193–1205.
- Duerr JS. 2006. Immunohistochemistry. *WormBook*: 1–61.
- Dupuy D, Li QR, Deplancke B, Boxem M, Hao T, Lamesch P, Sequerra R, Bosak S, Doucette-Stamm L, Hope IA, Hill DE, Walhout AJ, Vidal M. 2004. A first version of the *Caenorhabditis elegans* Promoterome. *Genome Res* 14: 2169–2175.
- Dupuy D, Bertin N, Hidalgo CA, Venkatesan K, Tu D, Lee D, Rosenberg J, Svrzika N, Blanc A, Carnec A, Carvunis AR, Pulak R, Shingles J, Reece-Hoyes J, Hunt-Newbury R, Viveiros R, Mohler WA, Tasan M, Roth FP, Le Peuch C, Hope IA, Johnsen R, Moerman DG, Barabasi AL, Baillie D, Vidal M. 2007. Genome-scale analysis of in vivo spatiotemporal promoter activity in *Caenorhabditis elegans*. *Nat Biotechnol* 25: 663–668.
- Eberwine J. 1996. Amplification of mRNA populations using aRNA generated from immobilized oligo(dT)-T7 primed cDNA. *Biotechniques* 20:584–591.
- Edgar LG, Carr S, Wang H, Wood WB. 2001. Zygotic expression of the caudal homolog *pal-1* is required for posterior patterning in *Caenorhabditis elegans* embryogenesis. *Dev Biol* 229:71–88.
- Evans TC, Crittenden SL, Kodoyianni V, Kimble J. 1994. Translational control of maternal *glp-1* mRNA establishes an asymmetry in the *C. elegans* embryo. *Cell* 77:183–194.
- Felix MA. 2007. Cryptic quantitative evolution of the vulva intercellular signaling network in *Caenorhabditis*. *Curr Biol* 17:103–114.
- Finney M, Ruvkun G. 1990. The *unc-86* gene product couples cell lineage and cell identity in *C. elegans*. *Cell* 63: 895–905.
- Fire A, Harrison SW, Dixon D. 1990. A modular set of lacZ fusion vectors for studying gene expression in *Caenorhabditis elegans*. *Gene* 93:189–198.
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391:806–811.
- Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. 1999. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151:1531–1545.
- Frank CA, Baum PD, Garriga G. 2003. HLH-14 is a *C. elegans* achaete-scute protein that promotes neurogenesis through asymmetric cell division. *Development* 130:6507–6518.
- Frokjaer-Jensen C, Davis MW, Hopkins CE, Newman BJ, Thummel JM, Olesen SP, Grunnet M, Jorgensen EM. 2008. Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nat Genet* 40: 1375–1383.
- Fukushige T, Krause M. 2005. The myogenic potency of HLH-1 reveals widespread developmental plasticity in early *C. elegans* embryos. *Development* 132: 1795–1805.
- 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.
- Fukushige T, Hendzel MJ, Bazett-Jones DP, McGhee JD. 1999. Direct visualization of the *elt-2* gut-specific GATA factor binding to a target promoter inside the living *Caenorhabditis elegans* embryo. *Proc Natl Acad Sci U S A* 96: 11883–11888.
- Fukushige T, Brodigan TM, Schriefer LA, Waterston RH, Krause M. 2006. Defining the transcriptional redundancy of early bodywall muscle development in *C. elegans*: evidence for a unified theory of animal muscle development. *Genes Dev* 20:3395–3406.
- Gaudet J, Mango SE. 2002. Regulation of organogenesis by the *Caenorhabditis elegans* FoxA protein PHA-4. *Science* 295:821–825.
- Geiss GK, Bumgarner RE, Birditt B, Dahl T, Dowidar N, Dunaway DL, Fell HP, Ferree S, George RD, Grogan T, James JJ, Maysuria M, Mitton JD, Oliveri P, Osborn JL, Peng T, Ratcliffe AL, Webster PJ, Davidson EH, Hood L, Dimitrov K. 2008. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol* 26:317–325.
- Gendreau SB, Moskowitz IP, Terns RM, Rothman JH. 1994. The potential to differentiate epidermis is unequally distributed in the AB lineage during early embryonic development in *C. elegans*. *Dev Biol* 166:770–781.
- Gengyo-Ando K, Mitani S. 2000. Characterization of mutations induced by ethyl methanesulfonate, UV, and trimethylpsoralen in the nematode *Caenorhabditis elegans*. *Biochem Biophys Res Commun* 269:64–69.
- Gilleard JS, McGhee JD. 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.
- Gilleard JS, Shafi Y, Barry JD, McGhee JD. 1999. *ELT-3*: a *Caenorhabditis elegans* GATA factor expressed in the embryonic epidermis during morphogenesis. *Dev Biol* 208:265–280.
- Goldstein B. 1992. Induction of gut in *Caenorhabditis elegans* embryos. *Nature* 357:255–257.
- Goldstein B. 2001. On the evolution of early development in the Nematoda. *Philos Trans R Soc Lond B Biol Sci* 356: 1521–1531.
- Gonczy P, Rose LS. 2005. Asymmetric cell division and axis formation in the embryo. *WormBook*: 1–20.
- Good K, Ciosk R, Nance J, Neves A, Hill RJ, Priess JR. 2004. The T-box transcription factors *TBX-37* and *TBX-38* link *GLP-1/Notch* signaling to mesoderm induction in *C. elegans* embryos. *Development* 131:1967–1978.
- Goszczynski B, McGhee JD. 2005. Reevaluation of the role of the *med-1* and *med-2* genes in specifying the *C. elegans* endoderm. *Genetics* 171:545–555.
- Gotta M, Ahninger J. 2001. Axis determination in *C. elegans*: initiating and transducing polarity. *Curr Opin Genet Dev* 11:367–373.
- Granger L, Martin E, Segalat L. 2004. Mos as a tool for genome-wide insertional mutagenesis in *Caenorhabditis elegans*: results of a pilot study. *Nucleic Acids Res* 32:e117.
- Hallam S, Singer E, Waring D, Jin Y. 2000. The *C. elegans* *NeuroD* homolog *cnd-1* functions in multiple aspects of motor neuron fate specification. *Development* 127:4239–4252.
- Harris TW, Antoshechkin I, Bieri T, Blasiar D, Chan J, Chen WJ, De La Cruz N, Davis P, Duesbury M, Fang R, Fernandes J, Han M, Kishore R, Lee R, Muller HM, Nakamura C, Ozersky P, Petcherski A, Rangarajan A, Rogers A, Schindelman G, Schwarz EM, Tuli MA, Van Auken K, Wang D, Wang X, Williams G, Yook K, Durbin R, Stein LD, Spieth J, Sternberg PW. 2010. *WormBase*: a comprehensive resource for nematode research. *Nucleic Acids Res* 38:D463–D467.
- Hawkins MG, McGhee JD. 1995. *elt-2*, a second GATA factor from the nematode *Caenorhabditis elegans*. *J Biol Chem* 270:14666–14671.
- Hench J, Henriksson J, Luppert M, Burglin TR. 2009. Spatio-temporal reference model of *Caenorhabditis elegans* embryogenesis with cell contact maps. *Dev Biol* 333:1–13.
- Hillier LW, Reinke V, Green P, Hirst M, Marra MA, Waterston RH. 2009. Massively parallel sequencing of the polyadenylated transcriptome of *C. elegans*. *Genome Res* 19:657–666.
- Hird SN, White JG. 1993. Cortical and cytoplasmic flow polarity in early embryonic cells of *Caenorhabditis elegans*. *J Cell Biol* 121:1343–1355.
- Hobert O. 2005. Specification of the nervous system. *WormBook*: 1–19.
- Hope IA. 1994. *PES-1* is expressed during early embryogenesis in *Caenorhabditis elegans* and has homology to the fork head family of transcription factors. *Development* 120:505–514.
- Horb ME, Thomsen GH. 1999. *Tbx5* is essential for heart development. *Development* 126:1739–1751.

- Horner MA, Quintin S, Domeier ME, Kimble J, Labouesse M, Mango SE. 1998. pha-4, an HNF-3 homolog, specifies pharyngeal organ identity in *Caenorhabditis elegans*. *Genes Dev* 12:1947–1952.
- Huang S, Shetty P, Robertson SM, Lin R. 2007. Binary cell fate specification during *C. elegans* embryogenesis driven by reiterated reciprocal asymmetry of TCF POP-1 and its coactivator beta-catenin SYS-1. *Development* 134:2685–2695.
- 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.
- Hunter CP, Harris JM, Maloof JN, Kenyon C. 1999. Hox gene expression in a single *Caenorhabditis elegans* cell is regulated by a caudal homolog and intercellular signals that inhibit wnt signaling. *Development* 126:805–814.
- Johnson NM, Behm CA, Trowell SC. 2005. Heritable and inducible gene knockdown in *C. elegans* using Wormgate and the ORFeome. *Gene* 359:26–34.
- Kalb JM, Lau KK, Goszczynski B, Fukushige T, Moons D, Okkema PG, McGhee JD. 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.
- Kaletta T, Schnabel H, Schnabel R. 1997. Binary specification of the embryonic lineage in *Caenorhabditis elegans*. *Nature* 390:294–298.
- Kamath RS, Ahringer J. 2003. Genome-wide RNAi screening in *Caenorhabditis elegans*. *Methods* 30:313–321.
- Kelly WG, Xu S, Montgomery MK, Fire A. 1997. Distinct requirements for somatic and germline expression of a generally expressed *Caenorhabditis elegans* gene. *Genetics* 146:227–238.
- Kemphues KJ, Priess JR, Morton DG, Cheng NS. 1988. Identification of genes required for cytoplasmic localization in early *C. elegans* embryos. *Cell* 52:311–320.
- Kiontke K, Gavin NP, Raynes Y, Roehrig C, Piano F, Fitch DH. 2004. *Caenorhabditis* phylogeny predicts convergence of hermaphroditism and extensive intron loss. *Proc Natl Acad Sci U S A* 101:9003–9008.
- Kohara Y. 2001. [Systematic analysis of gene expression of the *C. elegans* genome]. *Tanpakushitsu Kakusan Koso* 46:2425–2431.
- Korswagen HC, Herman MA, Clevers HC. 2000. Distinct beta-catenins mediate adhesion and signalling functions in *C. elegans*. *Nature* 406:527–532.
- Labouesse M, Mango SE. 1999. Patterning the *C. elegans* embryo: moving beyond the cell lineage. *Trends Genet* 15:307–313.
- Labouesse M, Hartweg E, Horvitz HR. 1996. The *Caenorhabditis elegans* LIN-26 protein is required to specify and/or maintain all non-neuronal ectodermal cell fates. *Development* 122:2579–2588.
- Lei H, Liu J, Fukushige T, Fire A, Krause M. 2009. Caudal-like PAL-1 directly activates the bodywall muscle module regulator hlh-1 in *C. elegans* to initiate the embryonic muscle gene regulatory network. *Development* 136:1241–1249.
- Li E, Davidson EH. 2009. Building developmental gene regulatory networks. *Birth Defects Res C Embryo Today* 87:123–130.
- Lin R. 2003. A gain-of-function mutation in oma-1, a *C. elegans* gene required for oocyte maturation, results in delayed degradation of maternal proteins and embryonic lethality. *Dev Biol* 258:226–239.
- Lin R, Thompson S, Priess JR. 1995. pop-1 encodes an HMG box protein required for the specification of a mesoderm precursor in early *C. elegans* embryos. *Cell* 83:599–609.
- Lin R, Hill RJ, Priess JR. 1998. POP-1 and anterior-posterior fate decisions in *C. elegans* embryos. *Cell* 92:229–239.
- Lin KT, Broitman-Maduro G, Hung WW, Cervantes S, Maduro MF. 2009. Knockdown of SKN-1 and the Wnt effector TCF/POP-1 reveals differences in endomesoderm specification in *C. briggsae* as compared with *C. elegans*. *Dev Biol* 325:296–306.
- Liu X, Long F, Peng H, Aerni SJ, Jiang M, Sanchez-Blanco A, Murray JI, Preston E, Mericle B, Batzoglu S, Myers EW, Kim SK. 2009. Analysis of cell fate from single-cell gene expression profiles in *C. elegans*. *Cell* 139:623–633.
- Long F, Peng H, Liu X, Kim SK, Myers E. 2009. A 3D digital atlas of *C. elegans* and its application to single-cell analyses. *Nat Methods* 6:667–672.
- Lowry JA, Gamsjaeger R, Thong SY, Hung W, Kwan AH, Broitman-Maduro G, Matthews JM, Maduro M, Mackay JP. 2009. Structural analysis of MED-1 reveals unexpected diversity in the mechanism of DNA recognition by GATA-type zinc finger domains. *J Biol Chem* 284:5827–5835.
- Lyons I, Parsons LM, Hartley L, Li R, Andrews JE, Robb L, Harvey RP. 1995. Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeo box gene *Nkx2-5*. *Genes Dev* 9:1654–1666.
- Maduro MF. 2006. Endomesoderm specification in *Caenorhabditis elegans* and other nematodes. *Bioessays* 28:1010–1022.
- Maduro MF. 2009. Structure and evolution of the *C. elegans* embryonic endomesoderm network. *Biochim Biophys Acta* 1789:250–260.
- Maduro MF, Rothman JH. 2002. Making worm guts: the gene regulatory network of the *Caenorhabditis elegans* endoderm. *Dev Biol* 246:68–85.
- 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.
- 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.
- Maduro M, Hill RJ, Heid PJ, Newman-Smith ED, Zhu J, Priess J, Rothman J. 2005a. Genetic redundancy in endoderm specification within the genus *Caenorhabditis*. *Dev Biol* 284:509–522.
- Maduro MF, Kasmir JJ, Zhu J, Rothman JH. 2005b. The Wnt effector POP-1 and the PAL-1/Caudal homeoprotein collaborate with SKN-1 to activate *C. elegans* endoderm development. *Dev Biol* 285:510–523.
- Maduro MF, Broitman-Maduro G, Mengarelli I, Rothman JH. 2007. Maternal deployment of the embryonic SKN-1->MED-1,2 cell specification pathway in *C. elegans*. *Dev Biol* 301:590–601.
- Mango SE. 2007. The *C. elegans* pharynx: a model for organogenesis. *WormBook*:1–26.
- 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.
- Materna SC, Oliveri P. 2008. A protocol for unraveling gene regulatory networks. *Nat Protoc* 3:1876–1887.
- McGhee JD, Sleumer MC, Bilenky M, Wong K, McKay SJ, Goszczynski B, Tian H, Krich ND, Khattra J, Holt RA, Baillie DL, Kohara Y, Marra MA, Jones SJ, Moerman DG, Robertson AG. 2007. The ELT-2 GATA-factor and the global regulation of transcription in the *C. elegans* intestine. *Dev Biol* 302:627–645.
- McGhee JD, Fukushige T, Krause MW, Minnema SE, Goszczynski B, Gaudet J, Kohara Y, Bossinger O, Zhao Y, Khattra J, Hirst M, Jones SJ, Marra MA, Ruzanov P, Warner A, Zapf R, Moerman DG, Kalb JM. 2009. ELT-2 is the predominant transcription factor controlling differentiation and function of the *C. elegans* intestine, from embryo to adult. *Dev Biol* 327:551–565.
- Mello CC, Kramer JM, Stinchcomb D, Ambros V. 1991. Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J* 10:3959–3970.
- 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.
- Meneghini MD, Ishitani T, Carter JC, Hisamoto N, Ninomiya-Tsuji J, Thorpe CJ, Hamill DR, Matsumoto K, Bowerman B. 1999. MAP kinase and Wnt pathways converge to downregulate an HMG-domain repressor in *Caenorhabditis elegans*. *Nature* 399:793–797.
- Mickey KM, Mello CC, Montgomery MK, Fire A, Priess JR. 1996. An inductive interaction in 4-cell stage *C. elegans*

- embryos involves APX-1 expression in the signalling cell. *Development* 122: 1791–1798.
- Mizumoto K, Sawa H. 2007. Two betas or not two betas: regulation of asymmetric division by beta-catenin. *Trends Cell Biol* 17:465–473.
- Moskowitz IP, Gendreau SB, Rothman JH. 1994. Combinatorial specification of blastomere identity by *glp-1*-dependent cellular interactions in the nematode *Caenorhabditis elegans*. *Development* 120:3325–3338.
- Murray JI, Bao Z, Boyle TJ, Waterston RH. 2006. The lineaging of fluorescently-labeled *Caenorhabditis elegans* embryos with StarryNite and AceTree. *Nat Protoc* 1:1468–1476.
- Murray JI, Bao Z, Boyle TJ, Boeck ME, Mericle BL, Nicholas TJ, Zhao Z, Sandel MJ, Waterston RH. 2008. Automated analysis of embryonic gene expression with cellular resolution in *C. elegans*. *Nat Methods* 5:703–709.
- 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.
- Owraghi M, Broitman-Maduro G, Luu T, Roberson H, Maduro MF. 2010. Roles of the Wnt effector POP-1/TCF in the *C. elegans* endomesoderm specification gene network. *Dev Biol* [Epub ahead of print].
- Page BD, Zhang W, Steward K, Blumenthal T, Priess JR. 1997. ELT-1, a GATA-like transcription factor, is required for epidermal cell fates in *Caenorhabditis elegans* embryos. *Genes Dev* 11: 1651–1661.
- Patient RK, McGhee JD. 2002. The GATA family (vertebrates and invertebrates). *Curr Opin Genet Dev* 12:416–422.
- 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.
- Phillips BT, Kidd AR, 3rd, King R, Hardin J, Kimble J. 2007. Reciprocal asymmetry of SYS-1/beta-catenin and POP-1/TCF controls asymmetric divisions in *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A* 104:3231–3236.
- Praitis V, Casey E, Collar D, Austin J. 2001. Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. *Genetics* 157:1217–1226.
- Priess JR, Thomson JN. 1987. Cellular interactions in early *C. elegans* embryos. *Cell* 48:241–250.
- Priess JR, Schnabel H, Schnabel R. 1987. The *glp-1* locus and cellular interactions in early *C. elegans* embryos. *Cell* 51:601–611.
- Raharjo I, Gaudet J. 2007. Gland-specific expression of *C. elegans* *hlh-6* requires the combinatorial action of three distinct promoter elements. *Dev Biol* 302: 295–308.
- Reece-Hoyes JS, Deplancke B, Shingles J, Grove CA, Hope IA, Walhout AJ. 2005. A compendium of *Caenorhabditis elegans* regulatory transcription factors: a resource for mapping transcription regulatory networks. *Genome Biol* 6:R110.
- Reece-Hoyes JS, Shingles J, Dupuy D, Grove CA, Walhout AJ, Vidal M, Hope IA. 2007. Insight into transcription factor gene duplication from *Caenorhabditis elegans* Promoterome-driven expression patterns. *BMC Genomics* 8: 27.
- 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.
- Reinke V, Smith HE, Nance J, Wang J, Van Doren C, Begley R, Jones SJ, Davis EB, Scherer S, Ward S, Kim SK. 2000. A global profile of germline gene expression in *C. elegans*. *Mol Cell* 6: 605–616.
- Reiter JF, Alexander J, Rodaway A, Yelon D, Patient R, Holder N, Stainier DY. 1999. Gata5 is required for the development of the heart and endoderm in zebrafish. *Genes Dev* 13:2983–2995.
- Reuter R. 1994. The gene serpent has homeotic properties and specifies endoderm versus ectoderm within the *Drosophila* gut. *Development* 120: 1123–1135.
- Robertson SM, Shetty P, Lin R. 2004. Identification of lineage-specific zygotic transcripts in early *Caenorhabditis elegans* embryos. *Dev Biol* 276:493–507.
- Rocheleau CE, Downs WD, Lin R, Wittmann C, Bei Y, Cha YH, Ali M, Priess JR, Mello CC. 1997. Wnt signaling and an APC-related gene specify endoderm in early *C. elegans* embryos. *Cell* 90: 707–716.
- Rocheleau CE, Yasuda J, Shin TH, Lin R, Sawa H, Okano H, Priess JR, Davis RJ, Mello CC. 1999. WRM-1 activates the LIT-1 protein kinase to transduce anterior/posterior polarity signals in *C. elegans*. *Cell* 97:717–726.
- Roy PJ, Stuart JM, Lund J, Kim SK. 2002. Chromosomal clustering of muscle-expressed genes in *Caenorhabditis elegans*. *Nature* 418:975–979.
- Schierenberg E. 2001. Three sons of fortune: early embryogenesis, evolution and ecology of nematodes. *Bioessays* 23: 841–847.
- Schlesinger A, Shelton CA, Maloof JN, Meneghini M, Bowerman B. 1999. Wnt pathway components orient a mitotic spindle in the early *Caenorhabditis elegans* embryo without requiring gene transcription in the responding cell. *Genes Dev* 13:2028–2038.
- Schnabel R, Weigner C, Hutter H, Feichtinger R, Schnabel H. 1996. *mex-1* and the general partitioning of cell fate in the early *C. elegans* embryo. *Mech Dev* 54:133–147.
- Schnabel R, Hutter H, Moerman D, Schnabel H. 1997. Assessing normal embryogenesis in *Caenorhabditis elegans* using a 4D microscope: variability of development and regional specification. *Dev Biol* 184:234–265.
- Schneider SQ, Bowerman B. 2007. beta-Catenin asymmetries after all animal/vegetal-oriented cell divisions in *Platynereis dumerilii* embryos mediate binary cell-fate specification. *Dev Cell* 13:73–86.
- Schulze J, Schierenberg E. 2009. Embryogenesis of *Romanomermis culicivorax*: an alternative way to construct a nematode. *Dev Biol* 334:10–21.
- Seydoux G, Fire A. 1994. Soma-germline asymmetry in the distributions of embryonic RNAs in *Caenorhabditis elegans*. *Development* 120:2823–2834.
- Seydoux G, Fire A. 1995. Whole-mount in situ hybridization for the detection of RNA in *Caenorhabditis elegans* embryos. *Methods Cell Biol* 48:323–337.
- 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.
- Shi Y, Mello C. 1998. A CBP/p300 homolog specifies multiple differentiation pathways in *Caenorhabditis elegans*. *Genes Dev* 12:943–955.
- Shin TH, Yasuda J, Rocheleau CE, Lin R, Soto M, Bei Y, Davis RJ, Mello CC. 1999. MOM-4, a MAP kinase kinase-related protein, activates WRM-1/LIT-1 kinase to transduce anterior/posterior polarity signals in *C. elegans*. *Mol Cell* 4:275–280.
- Shirayama M, Soto MC, Ishidate T, Kim S, Nakamura K, Bei Y, van den Heuvel S, Mello CC. 2006. The Conserved Kinases CDK-1, GSK-3, KIN-19, and MBK-2 Promote OMA-1 Destruction to Regulate the Oocyte-to-Embryo Transition in *C. elegans*. *Curr Biol* 16:47–55.
- Sleumer MC, Bilenky M, He A, Robertson G, Thiessen N, Jones SJ. 2009. *Caenorhabditis elegans* cisRED: a catalogue of conserved genomic elements. *Nucleic Acids Res* 37:1323–1334.
- Srivastava D. 1999. HAND proteins: molecular mediators of cardiac development and congenital heart disease. *Trends Cardiovasc Med* 9:11–18.
- Stringham EG, Dixon DK, Jones D, Candido EP. 1992. Temporal and spatial expression patterns of the small heat shock (*hsp16*) genes in transgenic *Caenorhabditis elegans*. *Mol Biol Cell* 3: 221–233.
- Stroeher VL, Kennedy BP, Millen KJ, Schroeder DF, Hawkins MG, Goszczynski B, McGhee JD. 1994. DNA-protein interactions in the *Caenorhabditis elegans* embryo: oocyte and embryonic factors that bind to the promoter of the gut-specific *ges-1* gene. *Dev Biol* 163: 367–380.
- Sulston JF, Schierenberg E, White JG, Thomson JN. 1983. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev Biol* 100:64–119.
- Swiers G, Patient R, Loose M. 2006. Genetic regulatory networks programming hematopoietic stem cells and erythroid lineage specification. *Dev Biol* 294:525–540.
- Tapscott SJ. 2005. The circuitry of a master switch: MyoD and the regulation of

- skeletal muscle gene transcription. *Development* 132:2685–2695.
- 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.
- Thorpe CJ, Schlesinger A, Bowerman B. 2000. Wnt signalling in *Caenorhabditis elegans*: regulating repressors and polarizing the cytoskeleton. *Trends Cell Biol* 10:10–17.
- Timmons L, Fire A. 1998. Specific interference by ingested dsRNA. *Nature* 395:854.
- 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.
- Voronov DA, Panchin YV. 1998. Cell lineage in marine nematode *Enoplos brevis*. *Development* 125:143–150.
- Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10:57–63.
- Wiegner O, Schierenberg E. 1999. Regulative development in a nematode embryo: a hierarchy of cell fate transformations. *Dev Biol* 215:1–12.
- Winston WM, Sutherlin M, Wright AJ, Feinberg EH, Hunter CP. 2007. *Caenorhabditis elegans* SID-2 is required for environmental RNA interference. *Proc Natl Acad Sci U S A* 104:10565–10570.
- Witze ES, Field ED, Hunt DF, Rothman JH. 2009. *C. elegans* *pur alpha*, an activator of *end-1*, synergizes with the Wnt pathway to specify endoderm. *Dev Biol* 327:12–23.
- Wolf N, Priess J, Hirsh D. 1983. Segregation of germline granules in early embryos of *Caenorhabditis elegans*: an electron microscopic analysis. *J Embryol Exp Morphol* 73:297–306.
- Wolpert L, Smith J, Jessell T, Lawrence P. 2007. *Principles of development*. Oxford: Oxford University Press.
- Wood WB. 1991. Evidence from reversal of handedness in *C. elegans* embryos for early cell interactions determining cell fates. *Nature* 349:536–538.
- Yochem JK. 2006. Nomarski images for learning the anatomy, with tips for mosaic analysis. *WormBook*: 1–47.
- Yochem J, Gu T, Han M. 1998. A new marker for mosaic analysis in *Caenorhabditis elegans* indicates a fusion between *hyp6* and *hyp7*, two major components of the hypodermis. *Genetics* 149:1323–1334.
- Zhao Z, Boyle TJ, Bao Z, Murray JI, Mericle B, Waterston RH. 2008. Comparative analysis of embryonic cell lineage between *Caenorhabditis briggsae* and *Caenorhabditis elegans*. *Dev Biol* 314:93–99.
- Zhu J, Hill RJ, Heid PJ, Fukuyama M, Sugimoto A, Priess JR, Rothman JH. 1997. *end-1* encodes an apparent GATA factor that specifies the endoderm precursor in *Caenorhabditis elegans* embryos. *Genes Dev* 11:2883–2896.
- 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.
- Zwaal RR, Broeks A, van Meurs J, Groenen JT, Plasterk RH. 1993. Target-selected gene inactivation in *Caenorhabditis elegans* by using a frozen transposon insertion mutant bank. *Proc Natl Acad Sci U S A* 90:7431–7435.