

Environmentally Induced Foregut Remodeling by PHA-4/FoxA and DAF-12/NHR

Wanyuan Ao,^{1*} Jeb Gaudet,^{1*†} W. James Kent,²
Srikanth Muttumu,¹ Susan E. Mango^{1‡}

Growth and development of the *Caenorhabditis elegans* foregut (pharynx) depends on coordinated gene expression, mediated by pharynx defective (PHA)-4/FoxA in combination with additional, largely unidentified transcription factors. Here, we used whole genome analysis to establish clusters of genes expressed in different pharyngeal cell types. We created an expectation maximization algorithm to identify cis-regulatory elements that activate expression within the pharyngeal gene clusters. One of these elements mediates the response to environmental conditions within pharyngeal muscles and is recognized by the nuclear hormone receptor (NHR) DAF-12. Our data suggest that PHA-4 and DAF-12 endow the pharynx with transcriptional plasticity to respond to diverse developmental and physiological cues. Our combination of bioinformatics and in vivo analysis has provided a powerful means for genome-wide investigation of transcriptional control.

A critical question in developmental biology is how complex programs of gene expression are orchestrated by a class of regulators known as selector genes. Selector genes code for transcription factors that autonomously govern the fates of groups of cells related to each other by virtue of their

cell type, position, or affiliation to an organ (*J*). For example, the FoxA transcription factor PHA-4 dictates the identity of cells within the *Caenorhabditis elegans* pharynx. Embryos that lack *pha-4* fail to generate pharyngeal cells, and conversely ectopic *pha-4* is sufficient to drive nonpharyngeal cells toward a pharyngeal fate (2, 3). PHA-4 also functions during postembryonic development, because reduction of *pha-4* activity at birth is lethal (4). These dramatic phenotypes reflect the global regulatory role of PHA-4 within pharyngeal cells. Many, perhaps all, genes selectively expressed within the pharynx are activated directly by PHA-4, including genes expressed in different pharyngeal cell types and at different developmental stages (4).

This strategy raises the question of how a single transcription factor can mediate diverse transcriptional outcomes within different cellular contexts. Combinatorial regulation by multiple transcription factors is one likely mechanism. However, few transcription factors have been discovered that could act in conjunction with PHA-4 (5, 6). Here, we explore the cis-regulatory logic for expression within pharyngeal cells, with a particular focus on growth control in response to food availability.

Gene expression profiles for pharyngeal cell types. To analyze expression profiles within subsets of pharyngeal cells, we subdivided 339 candidate pharyngeal genes previously identified by microarray analysis (Fig. 1, fig. S1, and table S1) (4, 7). We took advantage of RNA in situ patterns (8), reporter expression, and placement on the *C. elegans* Topological Expression Map (9) (Topo Map) to generate five clusters of genes with shared expression profiles. On the Topo Map, genes are positioned according to their correlated expression across 553 independent microarray experiments to generate “mountains” that reflect expression within a common cell type or cellular process. The clusters of pharyngeal genes encompassed 65% of the microarray positives and defined genes expressed in pharyngeal muscles (Ph-M), pharyngeal glands (Ph-G), pharyngeal marginal cells (Ph-MC), epithelia (Epi), and pharyngeal muscle and marginal cells combined (Ph-MMC) (fig. S1 and table S1). We validated these assignments with green fluorescent protein (GFP) reporters for previously uncharacterized members of each cluster (Fig. 1).

The compendium of pharyngeal gene clusters provided a tool to explore the regulatory circuitry for cell type-specific

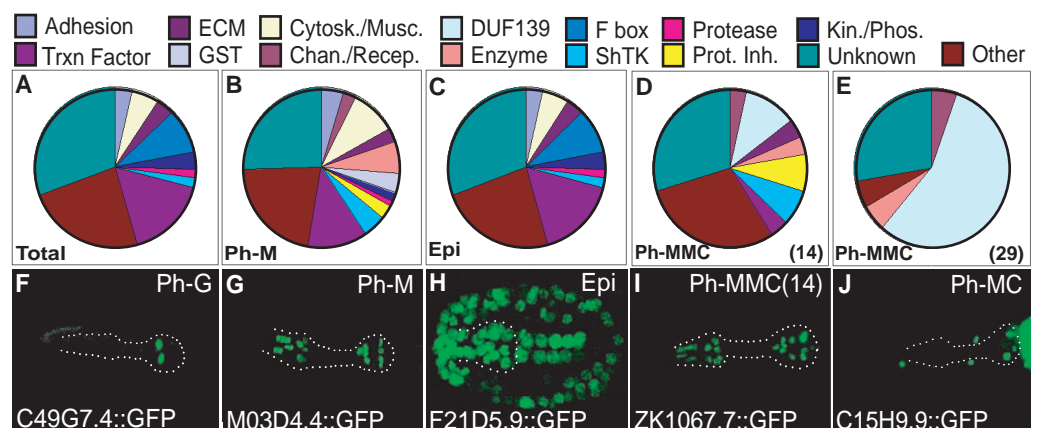
¹Huntsman Cancer Institute, University of Utah, 2000 Circle of Hope, Salt Lake City, UT 84112, USA. ²Genome Bioinformatics Group, University of California, Santa Cruz, CA 95064, USA.

*These authors contributed equally to this work.

†Present address: Genes and Development Research Group, University of Calgary, Calgary, Alberta AB T2N 4N1, Canada.

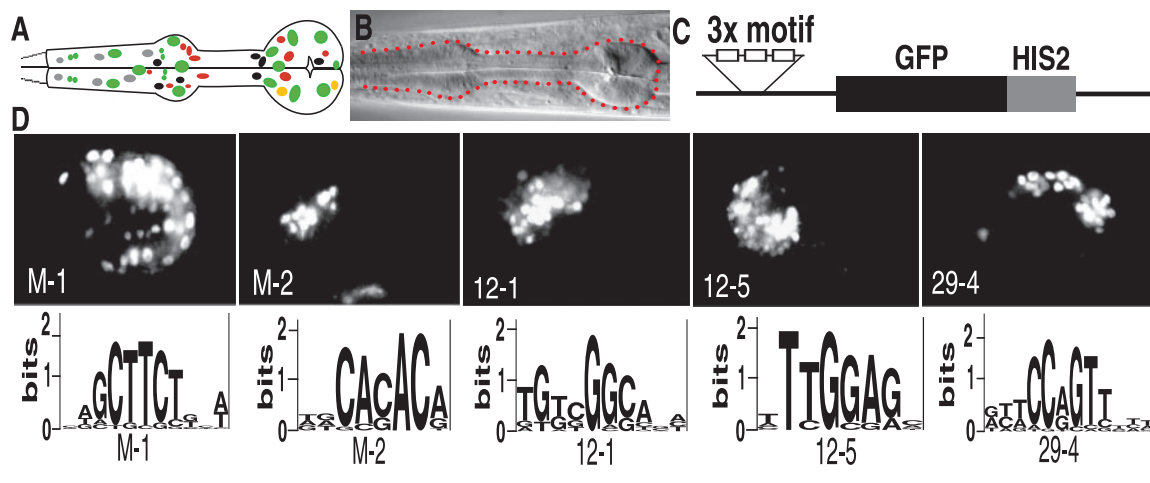
‡To whom correspondence should be addressed. E-mail: susan.mango@hci.utah.edu

Fig. 1. Five pharyngeal clusters. Pie charts of protein categories for all candidate pharyngeal genes (A) and individual clusters (B to E) and individual clusters (B to E). GST, glutathione S-transferase; ECM, extracellular matrix protein; Prot. Inh., protease inhibitor; ShTK, ShTK domain-containing protein; Chan./Recep., channel molecule or receptor; Cytosk./Musc., cytoskeleton or muscle protein; Enzyme, metabolism enzyme; DUF139, domain of unknown function 139 domain; Trxn Factor, transcription factor; and Kin./Phos., kinase or phosphatase. (F to J) Previously uncharacterized genes were expressed as predicted. C49G7.4 in Ph-G, confirmed by co-staining with a gland-specific antibody, J126 (19), M03D4.4 in Ph-M, F21D5.9 in embryonic Epi, ZK1067.7 in late muscles and marginal cells from Topo Mounts 14 or 29 (Ph-MMC) (9), and C15H9.9 in Ph-MC. Dotted lines indicate the pharynx [(F) to (J)].



C49G7.4 in Ph-G, confirmed by co-staining with a gland-specific antibody, J126 (19), M03D4.4 in Ph-M, F21D5.9 in embryonic Epi, ZK1067.7 in late muscles and marginal cells from Topo Mounts 14 or 29 (Ph-MMC) (9), and C15H9.9 in Ph-MC. Dotted lines indicate the pharynx [(F) to (J)].

Fig. 2. Predicted cis-regulatory motifs function in vivo. (A) Diagram of pharyngeal nuclei (green, myoepithelia; red, neurons; gray, pharyngeal epithelia cells; orange, gland cells; and black, marginal cells). (B) Head of an adult worm featuring the bilobed pharynx (dotted). (C and D) Five of seven motifs displayed enhancer activity in embryos when placed upstream of a minimal *pes-10* reporter (C) (4), which alone had no activity (7).



expression. We created the Improbizer algorithm (10) to identify potential cis-regulatory sites enriched in the pharyngeal gene clusters (11). Improbizer searches for motifs in DNA or RNA sequences that occur with unexpected frequency by using a variation of the expectation maximization algorithm (12). As a negative control, we examined groups of genes (DNA synthesis or a set of 194 randomly selected genes) that were not specifically expressed in the pharynx and removed motifs common to both the pharyngeal and negative control sets. This analysis identified seven candidate regulatory elements (Fig. 2, fig. S2, and tables S4 and S5). To test whether these elements were biologically active, we introduced three copies of each motif upstream of a minimal *pes-10* promoter (13) and examined them for their ability to activate expression.

Five of seven candidate motifs displayed enhancer activity, indicating that our bioinformatics approach efficiently identified cis-regulatory elements that functioned in vivo. Four were active within restricted domains, whereas one, M-1, behaved ubiquitously throughout the embryo (Fig. 2D). Motif M-2, identified from Ph-M, was the most specific and activated expression within pharyngeal muscles. GFP expression initiated during mid-embryogenesis (bean stage) (14), peaked during pharyngeal morphogenesis (two- to threefold stages), and decreased dramatically after hatching. Motif 29-4, identified from Ph-MMC, was active in anterior embryonic cells, including those of the pharynx. Reporters were expressed in the head and broadly throughout the pharynx beginning at mid-embryogenesis (1.5- to 2-fold stage). Robust expression was maintained until hatching, after which only a few cells continued to express. Two regulatory motifs, 12-1 and 12-5, were associated with the Epi cluster and activated expression in

Table 1. Pharyngeal muscle gene expression is reduced when M-2 or *daf-12* is mutated under growth conditions. GFP expression was determined in independent transgenic lines (*n*), with >20 animals examined per line. "mut" indicates mutant M-2 sites. Inactivation of M-2, *daf-12*, or both lead to an equivalent reduction in GFP expression.

Reporter	Strain	Strong (<i>n</i>)	Reduced (<i>n</i>)	Dramatically reduced (<i>n</i>)	None (<i>n</i>)
<i>myo-2P::GFP</i>	Wild type	4	2	3	1
<i>myo-2(mut)P::GFP</i>	Wild type	1	4	8	4
<i>myo-2P::GFP</i>	<i>daf-12 (m20)</i>	0	1	3	2
<i>myo-2(mut)P::GFP</i>	<i>daf-12 (m20)</i>	0	0	3	3
<i>ceh-22P::GFP</i>	Wild type	2	6	1	2
<i>ceh-22(mut)P::GFP</i>	Wild type	1	1	6	4
<i>ceh-22P::GFP</i>	<i>daf-12 (m20)</i>	0	4	8	2
<i>ceh-22(mut)P::GFP</i>	<i>daf-12 (m20)</i>	0	2	8	2

ectoderm (pharynx, neurons, and epidermis). For example, reporters for 12-5 expressed GFP in the pharynx (71%); pharynx and neurons (24%); or pharynx, neurons, and epidermis (5%). Both 12-1 and 12-5 were active in bean-stage embryos. Reporters for 12-1 maintained GFP expression to adulthood, whereas 12-5-dependent expression peaked at the two- to threefold stage and subsequently declined. These five cis-regulatory elements suggest a simple code for expression within pharyngeal epithelia. We propose that PHA-4 sites contribute to organ-wide activation throughout the pharynx whereas additional elements impart positional (29-4) or cell-type (M-2, 12-1, and 12-5) information.

The M-2 motif. We examined the activity of the cis-regulatory elements within natural pharyngeal promoters to address their importance for endogenous genes. We focused on the M-2 motif because it was highly selective for pharyngeal muscle. To identify relevant instances of the motif, we chose four genes that were expressed in pharyngeal muscles and carried a perfect match to the M-2 consensus in *C. elegans* and the related species *C. briggsae* (15): the heavy chain myosin genes F45G2.2 and *myo-2* (16), the Nkx2.5 homeobox homolog *ceh-22*

(17, 18), and the predicted zinc finger M03D4.4.

The M-2 motif was required in natural promoters for robust expression. Whereas most transgenic animals expressed GFP strongly in pharyngeal muscles throughout life, expression was reduced dramatically at all stages when M-2 was mutated (Fig. 3D and Table 1) (19). Fewer animals had GFP, and expression was weaker for those that did. The presence of additional regulatory elements accounts for the residual expression observed for promoters bearing M-2 mutations (4, 17, 20).

A current challenge for computational searches for cis-regulatory sites is to identify the trans-acting factors and upstream regulatory pathways that impinge upon these sites. Motifs 12-1 and 12-5 resemble known binding sites for the novel factors PEB-1 and CEB/P α , respectively (21, 22). We considered that T-box proteins might recognize the M-2 element (23); however, no effect on M-2-dependent expression was observed when eight predicted T-box genes were inactivated (table S6). We used a yeast one-hybrid screen (24) and obtained a single candidate, isolated 12 times, that recognized six copies of an intact M-2 motif but not a mutated version (Fig. 3B). This candidate was *daf-12*, which encodes

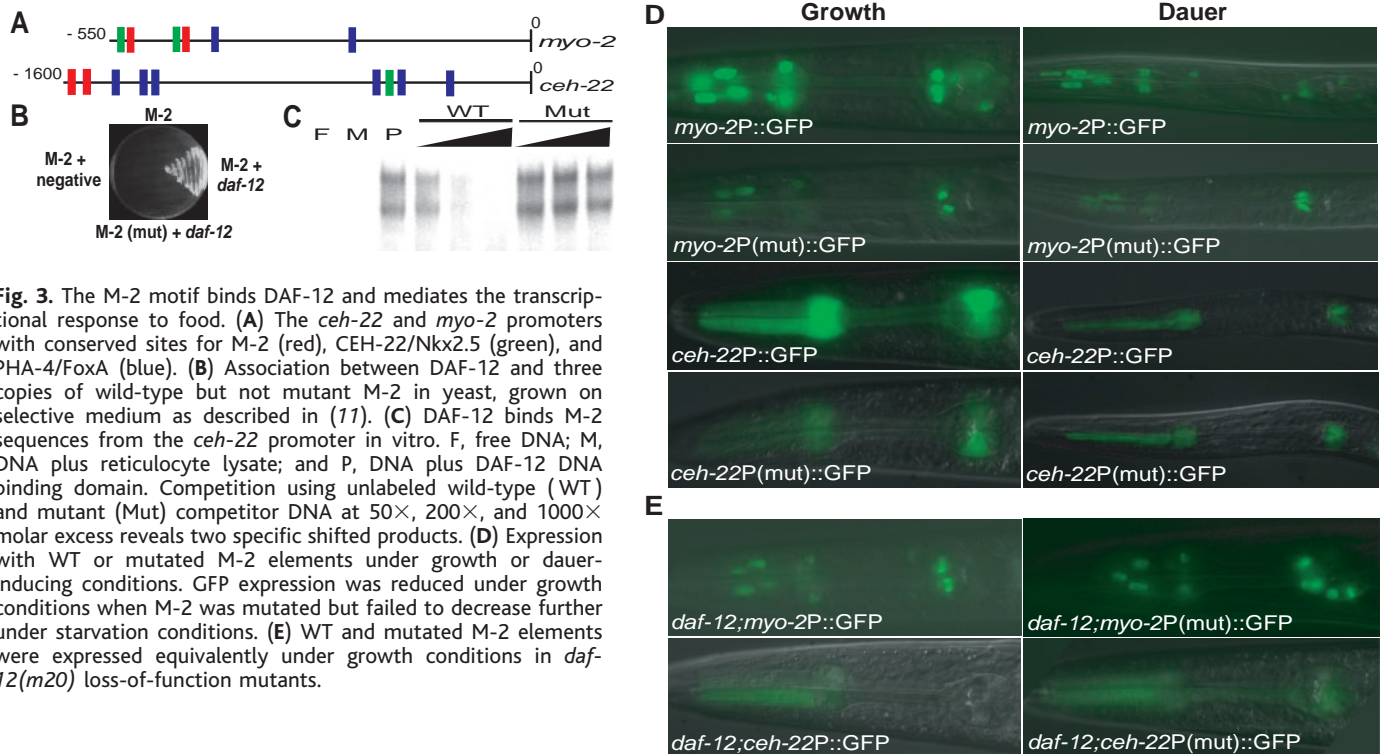


Fig. 3. The M-2 motif binds DAF-12 and mediates the transcriptional response to food. (A) The *ceh-22* and *myo-2* promoters with conserved sites for M-2 (red), CEH-22/Nkx2.5 (green), and PHA-4/FoxA (blue). (B) Association between DAF-12 and three copies of wild-type but not mutant M-2 in yeast, grown on selective medium as described in (11). (C) DAF-12 binds M-2 sequences from the *ceh-22* promoter in vitro. F, free DNA; M, DNA plus reticulocyte lysate; and P, DNA plus DAF-12 DNA binding domain. Competition using unlabeled wild-type (WT) and mutant (Mut) competitor DNA at 50 \times , 200 \times , and 1000 \times molar excess reveals two specific shifted products. (D) Expression with WT or mutated M-2 elements under growth or dauer-inducing conditions. GFP expression was reduced under growth conditions when M-2 was mutated but failed to decrease further under starvation conditions. (E) WT and mutated M-2 elements were expressed equivalently under growth conditions in *daf-12(m20)* loss-of-function mutants.

Table 2. M-2 is required for recovery of pharyngeal muscle gene expression during dauer exit. Entry and exit into the dauer stage were induced by removal or readdition of food, respectively. GFP expression was scored at 0, 3, or 8 hours after food restoration. Entries indicate the number of worms examined.

Reporter	0 hours recovery	3 hours	8 hours
<i>myo-2P::GFP</i>	weak:16 strong: 4	weak:16 strong: 4	weak: 2 strong:18
<i>myo-2(mut)P::GFP</i>	weak:20 strong: 0	weak:20 strong: 0	weak:20 strong: 0
<i>ceh-22P::GFP</i>	weak:14 strong: 6	weak:14 strong: 6	weak: 4 strong:16
<i>ceh-22(mut)P::GFP</i>	weak:20 strong: 0	weak:20 strong: 0	weak:20 strong: 0

a nuclear hormone receptor similar to the pregnane X and vitamin D receptors (25, 26). Electrophoretic mobility shift assays with the *ceh-22* promoter revealed that DAF-12 bound M-2-containing sequences in vitro and that binding was specific (Fig. 3C). These findings demonstrate that the M-2 element is recognized by DAF-12 in vitro.

To ascertain the importance of DAF-12 for M-2 activity in vivo, we examined *myo-2P::GFP* and *ceh-22P::GFP* in *daf-12(m20)* loss-of-function mutants (27). Mutant animals expressed GFP at a reduced level: Both the number of cells and GFP intensity were lower, similar to the effects of M-2 mutations (Fig. 3D and Table 1). Animals carrying double mutations in *daf-12* and M-2 resembled those with either single mutation, as expected if DAF-12

binds M-2 (Fig. 3E and Table 1). We conclude that DAF-12 is required for *myo-2* and *ceh-22* expression and functions through the M-2 element.

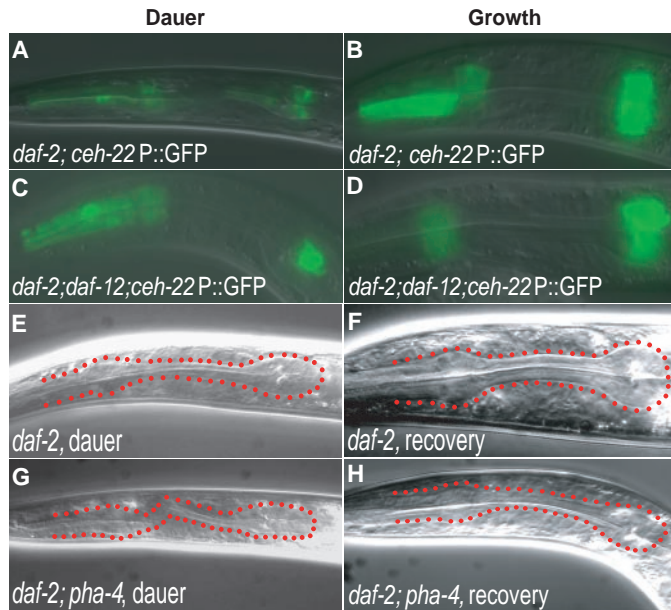
M-2 and modulation of expression by food availability. DAF-12 is a critical regulator of developmental progression and response to environmental cues (28, 29). When conditions are harsh, nematode larvae enter a long-lived state of diapause called the dauer stage. Dauer larvae exhibit multiple adaptations to stress, including remodeling of the pharynx to become thinner radially and cessation of feeding. The choice between growth and dauer development depends on insulin, transforming growth factor- β , and heterotrimeric guanine nucleotide-binding protein-coupled receptor signaling pathways, which converge on a handful of transcription factors, including DAF-12 (25, 26, 30). A key

question is how DAF-12 executes its biological functions to mediate these developmental events.

Given the role of DAF-12, we tested whether the M-2 motif modulates expression in dauer larvae. Wild-type *myo-2P::GFP* and *ceh-22P::GFP* were each expressed strongly in pharyngeal cells and down-regulated in dauer animals (Fig. 3D and Table 1) (31). Once food was restored, GFP gradually returned to its previous robust level as animals exited the dauer program (Table 2). By contrast, animals bearing a mutated M-2 motif failed to decrease expression in dauer larvae or reactivate expression upon feeding (Fig. 3D and Table 2). These effects were specific because transcriptional or translational PHA-4::GFP constructs, which lack an M-2 element, were not modulated in dauer larvae (19). We conclude that regulation of *myo-2* and *ceh-22* during dauer development depends critically on the M-2 motif.

Reduced *myo-2* and *ceh-22* expression in dauer larvae could reflect repression by DAF-12 or loss of activation. To distinguish between these possibilities, we examined CEH-22::GFP in *daf-2(insulin receptor)*; *daf-12* double mutants, which form dauer larvae despite reduced *daf-12* (30). We reasoned that GFP expression would decrease normally in *daf-12* dauer larvae if down-regulation reflected loss of activation but would increase relative to the wild type if down-regulation required DAF-12 repression. We found that CEH-22::GFP was derepressed in *daf-2*; *daf-12* double mutants and pharyngeal morphology resem-

Fig. 4. (A to D) DAF-12 is required to repress pharyngeal muscle expression in dauer larvae. Reduction of *ceh-22P::GFP* in dauer larvae [(A) and (B)] was lost in *daf-12* mutants [(C) and (D)]. (E to H) PHA-4 is required for dauer exit but not entry. *daf-2(ts); pha-4(RNAi)* animals were shifted to a nonpermissive temperature to induce dauer entry and pharyngeal remodeling [(E) and (G)]. When transferred to permissive temperature, *daf-2(ts); pha-4(RNAi)* animals failed to exit the dauer stage and resume growth [(F) and (H)].



bled that of growth conditions (Fig. 4). These data suggest that DAF-12 functions as a repressor in dauer larvae and that repression is essential for pharyngeal remodeling.

We identified pharyngeal genes with one or more predicted M-2 motifs and surveyed those genes for regulation during dauer development. A total of 94 pharyngeal genes carried a conserved M-2 element, of which 26 were down-regulated in dauer animals (28%) (31) (table S7). Conversely, 191 genes lacked a conserved M-2 element, of which 25 were modulated in dauers (13%). The enrichment of the M-2 element in dauer-regulated genes supports a role for this motif in response to environmental conditions. The majority of dauer-regulated genes were associated with pharyngeal muscle clusters (22 of 26, or 85%) (table S7) and included genes important for feeding, growth, and metabolism. For example, *inx-6* is required for electrical coupling between pharyngeal muscles during feeding (32), *nrs-2* and *cyp-17* are predicted to contribute to protein synthesis or folding, and *ech-4* is likely involved in fatty acid metabolism (33). These target genes, in addition to *ceh-22*, *F45G2.2*, and *myo-2*, are attractive candidates to modulate feeding and growth of the pharynx in response to nutritional cues.

Exit from the dauer program, but not entry, was critically dependent on PHA-4/FoxA. Dauer animals were induced by placing *daf-2(ts)* larvae at a nonpermissive temperature. Animals with *pha-4* amounts reduced by RNA interference (RNAi) [*pha-4(RNAi)*] could form dauer larvae as efficiently as *pha-4(+)* animals (Fig. 4). Furthermore, destruction of two predicted Fox binding sites within the *myo-2* promoter did not disrupt dauer-dependent repression of *myo-2P::GFP* (fig. S4). We conclude that *pha-4* is not required for the dauer program of

development. However, *pha-4* was essential for exit from the dauer program and hypertrophy of the pharynx under growth conditions. When *daf-2(ts)* dauer animals were shifted to permissive temperature for 32 hours, only 13% of *pha-4(RNAi)* larvae recovered and resumed growth ($n = 32$), compared to 92% of *pha-4(+)* animals ($n = 37$). The *pha-4(RNAi)* larvae remained small with partially remodeled pharynges and undeveloped gonads (Fig. 4). *pha-4(RNAi)* dauer larvae were viable for extended periods and resumed growth when *pha-4* activity was restored with food (19). Thus, the failure of *pha-4(RNAi)* larvae to reinitiate growth did not reflect a general unhealthiness of these animals. We conclude that the developmental selector gene *pha-4* is redeployed postembryonically to promote larval growth, pharyngeal hypertrophy, and sexual maturation in response to food.

We computationally identified five cis-regulatory sequences and tested their function in vivo. One of these elements mediates the response to environmental cues and is recognized by DAF-12. We speculate that when conditions are favorable, DAF-12, bound to an unidentified ligand, and PHA-4 function as transcriptional activators for genes involved in feeding, hypertrophy, and metabolism. When conditions are harsh, loss of the DAF-12 ligand converts DAF-12 into a transcriptional repressor and thereby triggers dauer development, pharyngeal remodeling, and cessation of feeding. This model provides an explanation for why loss of *daf-12* activity blocks entry into dauer development whereas mutations predicted to interfere with production of a DAF-12 ligand lead to constitutive dauer formation (34–37). These studies reveal how the transcriptional program can respond to nutrition to coordinate organ growth and activity.

References and Notes

- R. S. Mann, S. B. Carroll, *Curr. Opin. Genet. Dev.* **12**, 592 (2002).
- S. E. Mango, E. J. Lambie, J. Kimble, *Development* **120**, 3019 (1994).
- M. A. Horner *et al.*, *Genes Dev.* **12**, 1947 (1998).
- J. Gaudet, S. E. Mango, *Science* **295**, 821 (2002).
- P. G. Okkema, A. Fire, *Development* **120**, 2175 (1994).
- J. D. Thatcher, A. P. Fernandez, L. Beaster-Jones, C. Haun, P. G. Okkema, *Dev. Biol.* **229**, 480 (2001).
- J. Gaudet, S. Muttumu, M. Horner, S. E. Mango, *PLoS Biol.*, in press.
- More information is available at <http://nematode.lab.nig.ac.jp/db/index.html>.
- S. K. Kim *et al.*, *Science* **293**, 2087 (2001).
- The Improbizer algorithm is available online at www.cse.ucsc.edu/~kent/improbizer/improbizer.html.
- Materials and methods are available as supporting material on Science Online.
- T. L. Bailey, C. Elkan, *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2**, 28 (1994).
- A. Fire *et al.*, in *GFP Strategies and Applications*, M. Chalfie, S. Kain, Eds. (Wiley, New York, 1998), pp. 153–168.
- J. E. Sulston, E. Schierenberg, J. G. White, J. N. Thomson, *Dev. Biol.* **100**, 64 (1983).
- L. D. Stein *et al.*, *PLoS Biol.* **1**, E45 (2003).
- D. M. Miller, F. E. Stockdale, J. Karn, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 2305 (1986).
- P. G. Okkema, E. Ha, C. Haun, W. Chen, A. Fire, *Development* **124**, 3965 (1997).
- T. Vilimas, A. Abraham, P. G. Okkema, *Dev. Biol.* **266**, 388 (2004).
- W. Ao, J. Gaudet, W. J. Kent, S. Muttumu, S. E. Mango, unpublished data.
- J. D. Thatcher, C. Haun, P. G. Okkema, *Development* **126**, 97 (1999).
- L. Beaster-Jones, P. G. Okkema, *J. Mol. Biol.* **339**, 695 (2004).
- More information is available at www.gene-regulation.com.
- A. Kispert, B. G. Herrmann, *EMBO J.* **12**, 3211 (1993).
- M. M. Wang, R. R. Reed, *Nature* **364**, 121 (1993).
- A. Antebi, W. H. Yeh, D. Tait, E. M. Hedgecock, D. L. Riddle, *Genes Dev.* **14**, 1512 (2000).
- M. I. Snow, P. L. Larsen, *Biochim. Biophys. Acta* **1494**, 104 (2000).
- D. L. Riddle, M. M. Swanson, P. S. Albert, *Nature* **290**, 668 (1981).
- M. Tatar, A. Bartke, A. Antebi, *Science* **299**, 1346 (2003).
- C. E. Finch, G. Ruvkun, *Annu. Rev. Genomics Hum. Genet.* **2**, 435 (2001).
- P. L. Larsen, P. S. Albert, D. L. Riddle, *Genetics* **139**, 1567 (1995).
- J. Wang, S. K. Kim, *Development* **130**, 1621 (2003).
- S. Li, J. A. Dent, R. Roy, *Mol. Biol. Cell* **14**, 2630 (2003).
- More information is available at www.wormbase.org.
- P. S. Albert, D. L. Riddle, *Dev. Biol.* **126**, 270 (1988).
- B. Gerisch, C. Weitzel, C. Kober-Eisermann, V. Rotters, A. Antebi, *Dev. Cell* **1**, 841 (2001).
- K. Jia, P. S. Albert, D. L. Riddle, *Development* **129**, 221 (2002).
- A. Antebi, J. G. Culotti, E. M. Hedgecock, *Development* **125**, 1191 (1998).
- We thank A. Fire, Y. Kohara, and J. Kiefer for plasmids; R. Barstead for cDNA libraries; K. Yamamoto, H. Mak, and G. Ruvkun for communicating results before publication; and S. Dames for technical assistance. Some nematode strains were provided by the C. elegans Genetics Center, funded by NIH National Center for Research Resources. This work was supported by NIH R01-GM56264 to S.E.M. and Canadian Institutes of Health Research funding to J.G. S.E.M. is an associate investigator of the Huntsman Cancer Institute. Molecular interaction data have been deposited in the Biomolecular Interaction Network Database with the accession code 151736. DNA sequencing was supported by the NIH (2P303A42014).

Supporting Online Material

www.sciencemag.org/cgi/content/full/305/5691/1743/DC1
Materials and Methods
Figs. S1 to S4
Tables S1 to S8
References

1 July 2004; accepted 4 August 2004