Development 130, 1621-1634

doi:10.1242/dev.00363

Global analysis of dauer gene expression in *Caenorhabditis elegans*

John Wang[1] and Stuart K. Kim[1,2,*]

[1]Department of Developmental Biology, Stanford University, Stanford, CA 94305, USA
[2]Department of Genetics, Stanford University, Stanford, CA 94305, USA
*Author for correspondence (e-mail: kim@cmgm.stanford.edu)

Accepted 30 December 2002

SUMMARY

The dauer is a developmental stage in *C. elegans* that exhibits increased longevity, stress resistance, nictation and altered metabolism compared with normal worms. We have used DNA microarrays to profile gene expression differences during the transition from the dauer state to the non-dauer state and after feeding of starved L1 animals, and have identified 1984 genes that show significant expression changes. This analysis includes genes that encode transcription factors and components of signaling pathways that could regulate the entry to and exit from the dauer state, and genes that encode components of metabolic pathways important for dauer survival and longevity. Homologs of *C. elegans* dauer-enriched genes may be involved in the disease process in parasitic nematodes.

Supplemental data available on-line

Key words: Dauer, Microarray, Timecourse, *C. elegans*

INTRODUCTION

When conditions are not favorable, *Caenorhabditis elegans* can develop into dauer larvae to maximize survival (Cassada and Russell, 1975). In this facultative stage, worms become developmentally arrested, non-feeding, stress resistant and long lived (Cassada and Russell, 1975; Klass and Hirsh, 1976). Morphological remodeling and behavior are specialized for dispersal, while metabolism is altered to use internal energy stores. Low amounts of food, high population density and temperature induce first larval stage (L1) worms to become dauers (Golden and Riddle, 1984).

Genetic and molecular analyses (Riddle and Albert, 1997) have revealed that dauer-inducing pheromone is detected by sensory neurons that subsequently regulate larval development through the TGFβ (Ren et al., 1996) and cGMP (Birnby et al., 2000) signaling pathways. An insulin-like (Kimura et al., 1997) signaling pathway acts in parallel to control dauer entry in a pheromone-independent mechanism (Gottlieb and Ruvkun, 1994). Neuroendocrine cells probably integrate the manifold inputs from the insulin-like and TGFβ signaling pathways and then distribute the decision to become dauer to target tissues, in a *daf-12* nuclear hormone receptor-dependent manner (Antebi et al., 2000).

Dauer recovery is initiated once environmental conditions become favorable, particularly a high food to pheromone ratio and low temperature (Golden and Riddle, 1984). Upon transfer to favorable conditions, the first visible change is a change in surface lipophilicity (30 minutes), and commitment to dauer recovery occurs after approximately 50-60 minutes (Golden and Riddle, 1984; Proudfoot et al., 1993). Pharyngeal pumping, movement and increased body volume occur within 3 hours at 25°C (Cassada and Russell, 1975). There is a shift in internal pH from about pH 7.3 to about pH 6.3 prior to feeding (Wadsworth and Riddle, 1988). The recovering dauer molts into a post-dauer L4 at ~10 hours at 25ºC (Cassada and Russell, 1975). There are distinct sets of temporally regulated genes expressed upon exit from the dauer stage (Dalley and Golomb, 1992); expression of the Hsp70 and polyubiquitin genes peaks 75 minutes before the pharynx begins to pump and diminishes within 4 hours. SAGE analysis of dauer animals identified 358 genes that are dauer specific (Jones et al., 2001).

Dauer larvae have been described as ‘non-aging’, as post-dauer lifespan is not affected by the duration that an animal spends in the dauer state (Klass and Hirsh, 1976). Dauers can survive five times the normal lifespan (at least 70 days). Dauer longevity might be mediated in part by increased expression of stress resistance genes, decreased metabolic rates and by the insulin-like signaling pathway (Dorman et al., 1995; Kenyon et al., 1993; Tissenbaum and Ruvkun, 1998).

C. elegans dauer larvae have been considered analogous to the infectious larvae of parasitic nematodes because of morphological, behavioral and physiological similarities (Bird et al., 1999; Blaxter and Bird, 1997; Burglin et al., 1998). Additionally, muscarinic-receptor agonists induce recovery from the dauer stage in both *C. elegans* and the parasitic nematode, *Ancylostoma caninum*, indicating conservation at the neuronal level for these two species (Tissenbaum et al., 2000). Although the dauer stage is facultative in *C. elegans*, it is often obligatory in other species. As parasitic nematodes can infect humans and agricultural crops, a fundamental understanding of the mechanisms that underlie the dauer state in *C. elegans* might illuminate methods for control of nematodes that are pests (Aboobaker and Blaxter, 2000;

Blaxter and Bird, 1997). Analysis of dauer recovery in *C. elegans* could help define a conserved developmental transition for all nematodes.

In this paper, we have profiled gene expression patterns of the *C. elegans* dauer and associated dauer recovery process by using high-density DNA microarrays (Schena et al., 1995). Microarray analysis permits parallel and unbiased identification of genes that might have multiple functions, genetic redundancy or subtle phenotypes. We hypothesized that some of the transcriptional changes might simply be a response to the introduction of food rather than to development events per se, so we compared gene expression changes during dauer exit to those following exit from starvation. These experiments define a molecular profile for the dauer state involving 1984 dauer-regulated and 446 feeding-regulated genes.

MATERIALS AND METHODS

Strain

The strain used was *C. elegans* variety Briston strain N2.

Dauer growth and purification

Dauers were isolated as previously described with minor modifications (Epstein and Shakes, 1995). Wild-type (N2) *C. elegans* were grown on egg white plates for 8-12 days at room temperature (~22.5°C) until dauers had become plentiful. Worms were first isolated by washing with 0.1 M NaCl and sucrose floatation. Next, dauers were purified by treating the worms with 1% SDS for 1 hour, sucrose floatation and pelleting through 15% Ficoll. Purification was slightly variable, and in the worst case, estimated to be about 1:1 dauer to carcass by mass and 1000:1 dauers to non-dauers. Approximately 0.5 to 1 ml of this dauer sample was kept for the 0 hour time point, and a similar amount of worms were inoculated onto *E. coli* (OP50) seeded 15 cm plates for each time point. Concentrated *E. coli* paste derived from a fermentor was added to the plates at 0, 3 and 8 hours to prevent starvation. Each timecourse was repeated four times with the exception of the 4 and 7 hour time points, which only have three repetitions each.

During the course of our data analysis, it became clear that the carcasses contributed RNA. Carcass-derived RNA appeared dauer enriched as this RNA gradually degraded during the dauer exit timecourse. To filter out genes that change due simply to RNA degradation in carcasses, we analyzed four populations of pure dauers. We purified dauers through a nylon mesh with 15 μm pores for 20 minutes. In the worst case, the ratio of dauer to carcass was estimated to be about 33:1 and dauer to non-dauer worms was 1000:1. RNA was purified from the pure dauer populations before (0 hour) and after (12 hour) feeding with *E. coli*.

Isolation of L1 larvae

L1 larvae were isolated as previously described with minor modifications (Epstein and Shakes, 1995). Briefly, embryos were prepared by hypochlorite treatment, and were hatched in 0.1M NaCl overnight. L1 larvae were purified by sucrose floatation to remove dead carcasses, and allowed to recover in 0.1 M NaCl overnight. The starved L1 larvae were then plated onto 15 cm plates seeded with OP50. More *E. coli* was added at 3 and 8 hours to prevent starvation.

RNA preparation and microarray hybridization

The reference RNA used for all experiments is from a mixed staged population of N2 worms grown at room temperature (~22.5°C). Worms were harvested by washing the plates twice with 0.1 M NaCl at room temperature and then suspending in six volumes of Trizol. Total RNA and polyA(+) RNA were isolated as previously described (Reinke et al., 2000). Reference and experimental cDNA probes were labeled with Cy3 and Cy5, respectively, from 5-10 μg of polyA(+) RNA as previous described (DeRisi et al., 1997). Reference and experimental probes were purified with a Qiagen PCR purification kit and 28 μl subsequently hybridized (labeled probe, 8.3 mM Tris, 2×SSC, 0.17% SDS and 0.67 μg yeast tRNA) to near full-genome *C. elegans* DNA microarrays (Jiang et al., 2001).

Data analysis

Scanning was carried out with an Axon 3000 scanner. Data was acquired and quantitated with GenePix software. The raw data were uploaded into the Stanford Microarray Database. Normalized data [$\log_2$ (experiment/reference)] were downloaded using the following filter criteria: flag=0, failed=0, spot size >13 pixels, and red or green intensity >1.5-fold of the background intensity.

To identify genes that change in expression during the dauer exit timecourse, a standard one-way ANOVA ($P<0.001$) was applied to each gene. There were 2650 genes that passed these criteria, including genes that were derived from contamination from carcasses. Contaminating genes would appear downregulated in the dauer exit timecourse but not downregulated when comparing the 0 hour time point with the 12 hour time point using pure dauers. Contaminating genes would also be more enriched in the samples containing carcasses in comparison to the samples with pure dauers (at 0 hours). Specifically, we looked at the downregulated genes and removed genes whose expression was greater in the samples containing carcasses than in the pure dauer sample at 0 hours (one-tailed Students *t*-test, $P<0.05$) and that did not show significant downregulation in dauer exit using pure dauers (0 hours versus 12 hours, one-tailed Student's *t*-test, $P>0.001$). In this fashion, we removed 220 genes that were likely to be due to contamination from carcasses. The remaining 2430 genes are differentially regulated during the dauer exit timecourse and are reported in the Results.

For these 2430 genes, we employed a two-way mixed-model ANOVA as implemented by SAS to identify 1984 genes with different kinetics between the dauer exit and L1 starvation timecourses ($P<0.05$) (Romagnolo et al., 2002). We used a self-organizing map to cluster the 1984 genes of the dauer-specific data set into 45 nodes using 15 million iterations with Cluster software (Eisen et al., 1998). We used hierarchical clustering followed by visual inspection to subdivide the 45 nodes into the five groups described in the Results. For all other analyses, we used hierarchical clustering to organize individual genes. To compare whether gene classes were over-represented with respect to subsets of the data, we used the hypergeometric probability (J. Lund, data not shown). Treeview software was used to facilitate visualization of the microarray expression data (Eisen et al., 1998). When reporting the magnitude and confidence of dauer-enrichment in Figs 4, 5, we used the 12 horr/0 hour time point ratio and one-tailed Student's *t*-test *P* values derived from comparing the pure dauer 0 hour time point with the pure dauer 12 hour time point. Most gene annotations are from Proteome but some are manually annotated (e.g. different isoforms of *daf-16*). Gene lists are from Proteome, Kim et al. (Kim et al., 2001) or manually compiled.

RESULTS

In order to determine gene expression changes during the dauer recovery process, we performed DNA microarray experiments examining the transition from dauers into normal development over a 12 hour timecourse. As some of the gene expression changes might be related to the introduction of food rather than be specific to dauer exit, we compared the dauer exit timecourse to a timecourse after feeding of starved normal

worms. Although the dauer stage is the only developmentally regulated stage of arrest, the L1 larval stage will arrest in the absence of food (Epstein and Shakes, 1995). The arrested L1 larvae do not display dauer-like properties (e.g. prolonged life, stress resistance and altered morphology) but will continue development with the addition of food (Epstein and Shakes, 1995). Genes behaving similarly after feeding of starved L1 animals and dauer recovery might define common feeding programs. Genes that are regulated in the dauer but not in the starvation timecourse would be dauer recovery-specific.

We wanted to generate a high temporal resolution of the dauer exit process, so dauer animals were harvested at approximately 1 hour intervals over 12 hours after feeding. In order to engage statistical analysis, each timecourse was repeated four times. The L1 starvation timecourse involved a similar timecourse after feeding of starved L1 worms. From each experimental sample, we purified polyA(+) RNA, synthesized Cy3-labeled cDNA and compared it with Cy5-labeled reference cDNA synthesized from mixed-staged wild-type hermaphrodite polyA(+) RNA. We hybridized both probes to DNA microarrays contained 17,088 genes that correspond to 88% of the known predicted genes (Jiang et al., 2001) (M. Kiraly, personal communication). For each hybridization, we calculated the $\log_2$ of the expression level of the experimental sample relative to the expression level of the reference, and then averaged the results for the three or four replicates of each time point. Because all of the samples were compared with the same reference RNA, we could compare the expression levels of one time point to any other in either timecourse. We plotted how each gene behaved in the dauer recovery and L1 starvation timecourses relative to the starting time point (0 hours). The complete data for the dauer recovery and L1 starvation timecourses as well as an application to view expression profiles for individual genes can be viewed at http://cmgm.stanford.edu/~kimlab/dauer/

Statistical analysis

To identify genes that were significantly induced or repressed during the dauer exit timecourse, we first used one-way ANOVA analysis to identify 2430 genes that change expression in the dauer exit timecourse (P<0.001; Materials and Methods) (see supplemental Table S3 at http://dev.biologists.org/supplemental/ and at http://cmgm.stanford.edu/~kimlab/dauer/ExtraData.htm). The large number of genes identified is not unexpected as dauer larvae are highly differentiated and proceed through a dramatic change during exit. Some of the 2430 genes are related to the introduction of food while others might be specific to dauer exit. We distinguished between these two categories by comparing the expression profile of the dauer exit and L1 starvation timecourses using two-way mixed-model ANOVA (P<0.05). This analysis identified 1984 genes that were considered to be differentially expressed between the two timecourses (see supplemental Table S4 at http://dev.biologists.org/supplemental/ and at http://cmgm.stanford.edu/~kimlab/dauer/ExtraData.htm).

Feeding response

Enriched gene groups

The feeding response is defined by a set of 446 genes that change in the dauer exit timecourse and whose expression kinetics in the L1 starvation timecourse are not significantly different from the dauer exit timecourse based on two-way mixed-model ANOVA (P>0.05, Fig. 1). We compared the feeding response genes to groups of previously defined functional classes to determine which classes were statistically over-represented. We calculated a representation factor, which is the fold-enrichment of a particular class of genes over the number expected as a result of random chance given the number of genes in each group and the genome size. A gene class that is highly enriched suggests that its biological function plays an important role during recovery from starvation.

We find that the list of feeding-regulated genes are enriched for those involved in general metabolism and growth processes (Table 1). For example, mitochondrial genes are over-represented in the feeding response. There are 179 genes that encode proteins that are localized to the mitochondria in the *C. elegans* genome. Based on the 17,088 genes on the microarray, 4.5 genes would be expected to be included in the set of 446 genes with common feeding response kinetics by random chance. Instead, 19 mitochondrial genes are present in the feeding response gene set, corresponding to a 4.2-fold

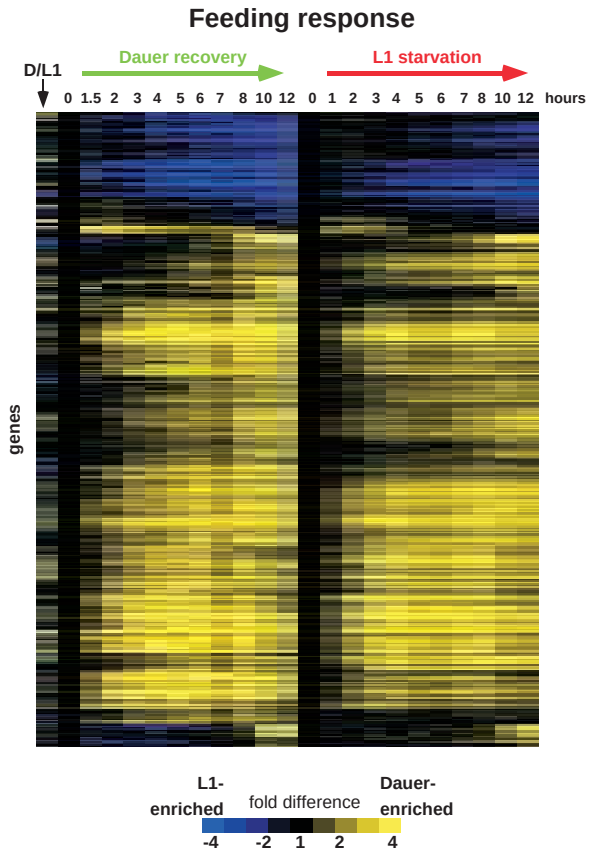

Fig. 1. Expression patterns of genes in feeding response. Hierarchical clustering was used to display the ratio of dauer to L1 at 0 hours (D/L1) and the expression ratios of the 446 commonly regulated genes during the dauer exit and L1 starvation timecourses. Rows represent different genes and columns represent different times during the timecourse. Scale shows level of expression. D, Dauer; L1, L1 larvae. Full data for this figure are available at http://dev.biologists.org/supplemental/ and at http://cmgm.stanford.edu/~kimlab/dauer/

Table 1. Feeding response genes

A Functional classes

Gene class*	Total on array	Fold enrichment in feeding response†	Number in feeding response	Ref.‡
Total number of gene	17761		446	
Nuclear: nucleolus	23	13.7	8	
tRNA synthetase genes	30	9.2	7	
Translation factors	43	9.2	10	
RNA processing/modification	71	6.7	12	
Endoplasmic reticulum	63	4.4	7	
Mitochondrial	179	4.2	19	
Protein expression	390	3.9	39	Kim et al. (2001)
RNA binding	151	3.4	13	
Germ line-enriched	508	2.5	32	Reinke et al. (2000)
Biosynthesis	478	2.4	29	Kim et al. (2001)
Muscle expressed genes	1329	2.1	70	Roy (2002)
Worm-yeast orthologs	1921	1.9	94	Kim et al. (2001)

B Gene expression mountains

Gene class*	Total on array	Fold enrichment in feeding response†	Number in feeding response	Comments§
Genome	17761		446	
Mount 41	7	11.3	2	Protein expression
Mount 40	8	14.8	3	Protein expression
Mount 20	160	11.4	46	Germline, biosynthesis, protein expression, heat shock
Mount 18	190	4.6	22	Germline, oocyte, biosynthesis, protein synthesis
Mount 05	978	2.3	56	
Mount 02	1465	2.2	80	Germline, oocyte, operons

*Over-represented functional classes in the feeding response gene groups. For clarity, gene groups with $P>0.001$ are not listed. Gene lists are shown in supplemental Table S7.

†All representation factors are significant at $P<0.001$. The probability was determined using the normal approximation of the hypergeometric probability distribution.

‡Functional classes are from Proteome except where noted.

§Functional groups over-represented in the mountains (Kim et al., 2001).

enrichment that is statistically significant (hypergeometric, $P<0.001$). Thus, we infer that mitochondrial function is important in the feeding response gene set, as would be expected considering that energy generation is needed for animals transitioning from an inactive to an active state.

By a similar analysis we find that genes encoding nucleolar proteins (13-fold, $P<0.001$), proteins in the endoplasmic reticulum (4.4-fold, $P<0.001$), translation factors (9.2-fold, $P<0.001$) and tRNA synthetases (9.2-fold, $P<0.001$) are over-represented (Table 1). Each of these gene groups is involved in protein expression, reflecting increased protein synthesis after feeding. Other over-represented categories are: genes involved in RNA binding, genes involved in RNA processing and modification, genes involved in biosynthetic pathways, germline-enriched genes, muscle-enriched genes, and worm-yeast orthologs genes (Table 1). The worm-yeast orthologs are a set of orthologous genes between worm and yeast that might define core biological functions, so the enrichment of worm-yeast orthologs in the set of feeding response genes is consistent with increased general metabolism and biosynthesis (Chervitz et al., 1998).

Enriched gene expression mountains

Kim et al. have organized genes into 44 groups based on correlated expression across 553 diverse experiments (Kim et al., 2001). The algorithm used to generate the gene groups produced a 3D output for visual representation (a gene expression topomap) in which the gene groups appear as mountains of co-expressed genes (Kim et al., 2001). Six gene-expression mountains are significantly over-represented in the feeding response (Table 1). Five (mounts 2, 18, 20, 40 and 41) are enriched for genes involved in protein expression or synthesis (Table 1) (Kim et al., 2001). Three of these (mounts 2, 18 and 20) are also enriched for biosynthesis genes. These results provide further evidence that genes involved in protein expression and biosynthesis are involved in the feeding response. The sixth mountain (mount 5) is not enriched for any previously defined gene classes.

The gene expression mountains can also be used to partition the list of feeding response genes into smaller clusters; each cluster contains genes that are co-expressed not only during feeding but in a diverse set of microarray experiments used to generate the expression topomap. Genes within a mountain exhibit tight co-regulation, and tight clustering of genes on the topomap is evidence that these genes may function together. The list of feeding response genes includes those whose function are known (e.g. because they encode proteins similar to proteins of known function) and also many genes with unknown function. We can use the expression topomap to infer the function of these unknown genes based on co-expression with genes of known function. For example, the feeding response genes whose functions are currently unknown and that are in mounts 2, 18, 20, 40 and 41 are likely involved in biosynthesis or protein expression. Similarly, genes with

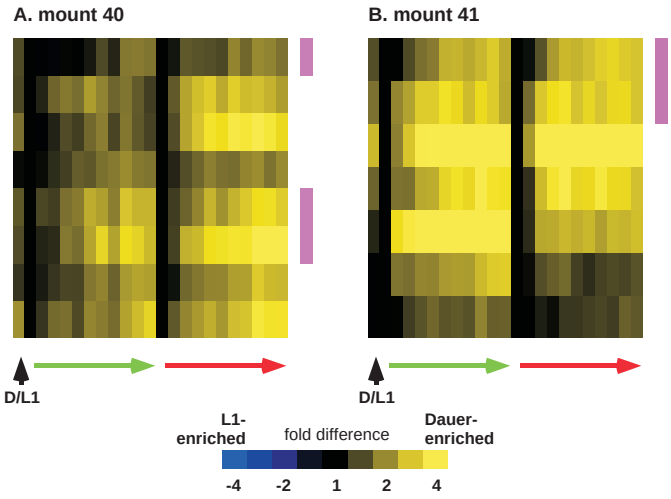

Fig. 2. Expression patterns of genes in mounts 40 and 41. Hierarchical clustering was used to display the ratio of dauer to L1 at 0 hours and the expression ratios of the commonly regulated genes during the dauer exit and L1 starvation timecourses for the (A) eight genes in mount 40 and (B) seven genes in mount 41. The genes in mounts 40 and 41 have homogeneous expression patterns. The feeding response genes that are in common with the genes in mounts 40 and 41 are indicated with pink bars to the right of the clusters. Rows represent different genes and columns represent different times during the timecourse. Green and red arrows are the dauer exit and L1 starvation timecourses, respectively. Scale shows level of expression. D, Dauer; L1, L1 larvae. Full data for this figure are available at http://dev.biologists.org/supplemental/ and at http://cmgm.stanford.edu/~kimlab/dauer/

unknown function in other mountains could match the ascribed function of that mountain.

As genes within a mountain exhibit tight co-regulation across many experiments, the gene expression mountains can be used to find additional genes that may be involved in the feeding response. For example, we plotted all of the genes in mounts 40 and 41 in the dauer and L1 starvation timecourses and found that almost all are upregulated in both timecourses (Fig. 2). These two mountains include not only the five genes that were found by these microarray experiments but also an additional ten genes that appear to be upregulated in both timecourses. Hence, these ten genes are likely to be feeding response genes; they are co-expressed in these gene expression mountains but were probably missed because they did not meet the stringent statistical criteria used by the ANOVA analysis. These results show that feeding response is another characteristic that underlies the expression profiles of the genes in mounts 40 and 41.

In contrast to mounts 40 and 41, the genes in mount 5 do not show a uniform response during the dauer and L1 starvation timecourses. Mount 5 is over-represented in the feeding response, but is not enriched in any of our previously defined gene classes (Table 1) (Kim et al., 2001). We plotted the 978 mount 5 genes on the two timecourses and found that expression of only about 40% of the genes was upregulated similarly during both timecourses (see supplemental Fig. S1 at http://dev.biologists.org/supplemental/ and at http://cmgm.stanford.edu/~kimlab/dauer/). Another 40% of the genes were upregulated only during the dauer recovery timecourse and not during the L1 starvation timecourse. Thus, although 56 mount 5 genes are defined by our microarray experiments to be feeding response genes (Table 1), the remaining may not necessarily be feeding response genes. Consequently, feeding response is not a new biological annotation for mount 5. The genes in mounts 2, 18 and 20 also exhibit heterogeneous expression patterns in the dauer exit and L1 starvation timecourses, indicating that feeding response is not a common biological characteristic for genes in these mountains (see Fig. 6C and data not shown).

These results exemplify how specific microarray experiments, such as those described in this paper, and global clustering using a compendium of microarray experiments, such as the expression topomap, are complementary. Genes defined by statistical criteria from an individual microarray experiment (e.g. the experiments described in this paper) can be partitioned into more refined functional groups using gene expression mountains. In cases where mountains exhibit homogenous expression patterns, the genes resident to that mountain but that lie outside of the statistical criteria for the individual experiment can be implicated in the biological process (e.g. feeding response). In these cases, the expression topomap mountains can be annotated as being enriched for the feeding response gene class. In cases in which only some of the genes in a specific gene expression mountain are feeding regulated, the current set of experiments help to subdivide that mountain.

Patterns of expression of dauer-specific genes

We used a self-organizing map (SOM) to cluster the 1984 dauer-specific genes into 45 groups, termed nodes, based on their patterns of expression during the dauer exit and L1 starvation timecourses [Fig. 3 (Kohonen, 1997); Cluster Software (Eisen et al., 1998)]. Each row in Fig. 3 depicts the normalized expression profile of the genes contained within that node.

By focusing only on the expression patterns during the dauer exit timecourse, we divided the dauer exit-specific genes into five groups: dauer-enriched, transient, early, climbing and late (Fig. 3). The expression of the genes in the dauer-enriched (540 genes) group decline in the dauer exit timecourse, showing either immediate or delayed kinetics. These genes are presumably involved in dauer-specific processes, such as stress resistance or prolonged life. Four groups of genes are induced as the dauer recovers: transient (195 genes, induced earliest and transiently), early (538 genes, induced to a steady state level), climbing (386 genes, induced and increasing continually) and late (325 genes, induced late) (Fig. 3).

As in the case for feeding response, we compared each of the groups defined in these microarray experiments to groups of previously-defined functional classes to determine if some are significantly over-represented. We highlight some biological functions for each of the five expression categories. The complete set of over-enriched categories along with their relative enrichment are listed in Table 2.

Dauer enriched

Dauer-enriched genes are downregulated upon dauer recovery, showing either early or delayed kinetics. The set of dauer-enriched genes probably define those that confer dauer-specific properties, such as stress resistance and longevity. Previous

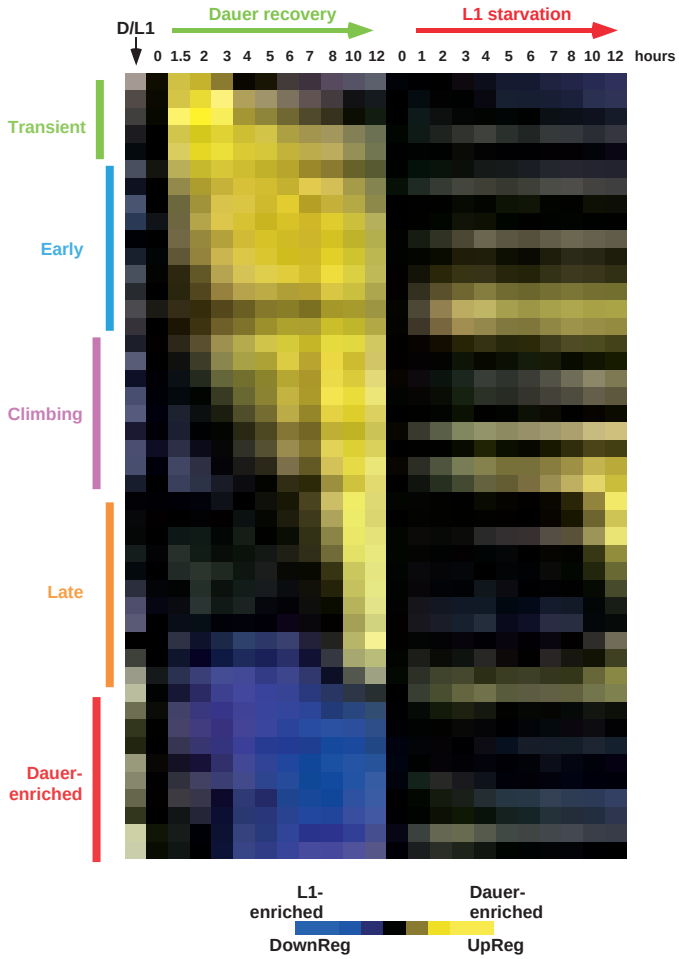

Fig. 3. A gene expression profile during the dauer exit and L1 starvation timecourses. A self-organizing map was used to cluster the 1984 dauer-specific genes with common normalized expression patterns into 45 nodes using 1.5×10^7 iterations. Each row represents the average expression pattern of the genes in the node (14-100 genes, average of 45). The columns represent different times during the timecourse except for the first column (D/L1) which is the ratio of dauer to L1 at 0 hours. The nodes have also been aggregated into five groups (vertical axis) based on the dauer exit timecourse expression patterns as labeled. These groups are partitioned into many nodes because the expression patterns in the L1 starvation timecourse and relative dauer-enrichment are different. Because normalized data are shown, there are no units on the scale bar. D, Dauer; L1, L1 larvae. The complete data are available at http://dev.biologists.org/supplemental/ and at http://cmgm.stanford.edu/~kimlab/dauer/

work has shown that *daf-21*, *hsp-20*, *sod-3* and *ctl-1* (which encode HSP90, HSP12.6, Mn-superoxide dismutase and cytosolic catalase, respectively) are involved in stress resistance and are enriched in dauer larvae (Birnby et al., 2000; Honda and Honda, 1999; Jones et al., 2001; Larsen, 1993; Taub et al., 1999). Our microarray data confirmed these results showing that *daf-21* (2.9 fold, $P<0.002$), *hsp-20* (125-fold, $P<0.001$), *ctl-1* (5.9-fold, $P<0.001$) and *sod-3* (4.5-fold, $P<0.021$) are significantly enriched in the dauer relative to the non-dauer (12 hours after feeding).

We observed that genes encoding cytochrome P450 enzymes and UDP-glucuronosyltransferases are over-represented 3.2-fold ($P<0.001$) and 4.9-fold ($P<0.001$), respectively, in the dauer-enriched expression class. Cytochrome P450 enzymes are mono-oxygenases that metabolize many endogenous and exogenous lipophilic compounds including steroidal hormones, xenobiotics and fatty acids (Mansuy, 1998). UDP-glucuronosyltransferases attach sugar residues to lipophilic molecules that facilitate their export from a cell (Tukey and Strassburg, 2000). These enzymes might inactivate external contamination or internally generated toxins, and thus provide stress resistance and perhaps prolonged lifespan for dauers. In addition to stress and toxin resistance, both cytochrome P450 enzymes and UDP-glucuronosyltransferases could have a role in the metabolism of a putative dauer entrance or exit hormone(s). Cytochrome P450 enzymes could participate in synthesizing, modifying or degrading a putative dauer hormone(s), while UDP-glucuronosyltransferases could regulate ligand activity by decorating a putative hormone(s) with sugar moieties.

Dauer metabolism overview

Normal worms ingest bacteria for energy and biosynthesis components. Carbohydrates and some derivatives of nucleic acids and amino acids are metabolized through the glycolytic pathway into acetyl-coA (via pyruvate), whereas lipids are metabolized by fatty acid β-oxidation into acetyl-coA (Stryer, 1995). Acetyl-coA is subsequently metabolized by the tricarboxylic acid cycle (TCA, citric acid cycle) and oxidative phosphorylation to generate ATP (Stryer, 1995). Thus, the glycolytic, TCA, oxidative phosphorylation and fatty acid β-oxidation pathways are all active in normal worms. With respect to biosynthesis, many building blocks can be obtained by ingestion. For example, nucleic acids and amino acids can be directly recycled for RNA and protein synthesis.

By contrast, the dauer is non-feeding and must rely on internal reservoirs for energy and biosynthesis. Fat is the major source for energy and biosynthetic precursors in dauers, and accumulates in the intestine and hypodermis during dauer formation. Fat is predominantly found in the form of triacylglycerides, consisting of three fatty acids attached to a glycerol backbone. As in normal worms, energy generation involves metabolizing acetyl-coA (derived from fatty acids) via the TCA cycle and oxidative phosphorylation. However, in dauers, fat must also provide biosynthetic precursors. This can be achieved in part through fatty acid β-oxidation, the glyoxylate cycle and gluconeogenesis (Stryer, 1995).

Dauer-enriched genes confirm metabolic predictions

The microarray data provide supporting evidence that all three of these metabolic pathways are active in the dauers. First, we find that genes encoding fatty acid β-oxidation enzymes are dauer enriched (data not shown). Previous results had shown that the specific activities of the enzymes involved in fatty acid β-oxidation are high in dauers, although at levels lower than adults, indicating that lipids are an important energy reserve for dauers (O'Riordan and Burnell, 1990). Our results support the importance of fat reservoirs in dauers, and furthermore, identify the genes that encode the dauer-enriched β-oxidation isoenzymes (data not shown). Because glycerol is the backbone for triacylglyceride fats, metabolism of fat for energy also implies concurrent metabolism of glycerol for energy. Glycerol can be used for energy generation after conversion into glyceraldehyde-3-phosphate, a metabolic intermediate of

glycolysis, by two enzymes: glycerol kinase and glycerol-3-phosphate dehydrogenase. The microarray results show that R11F4.1 (3.9-fold, $P<0.001$), which encodes glycerol kinase, and K11H3.1 (2.8-fold, $P<0.001$) and F47G4.3 (3.6-fold, $P<0.002$), each of which encodes glycerol-3-phosphate kinase, are dauer enriched. The upregulation of both enzymatic activities for metabolizing glycerol in dauers, as inferred from transcriptional upregulation, further supports the importance of fat reserves in dauers.

Second, the microarray data indicate that the glyoxylate cycle is active in dauers. We see that expression of C05E4.9, which encodes the *C. elegans* bi-functional glyoxylate enzyme, is higher in dauers relative to the end of the timecourse (13-fold; $P<0.001$). Furthermore, F54H12.1 (aconitase), F48E8.3

Table 2. Dauer recovery

A Functional classes

Gene class*	Total on array	Dauer-enriched†	Transient	Early	Climbing	Late
Genome	17761	540	195	538	386	325
Glycolysis	21				6.5 (3)	
TCA	30				9.1 (6)	
Muscle expressed genes	1329				2.6 (76)	
Hedgehog-like	46					12 (10)
Patched-like/NPC	22					17 (7)
Cytochrome P450 enzymes	81	3.2 (8)	5.6 (5)			
UDP-Glucoronosyltransferase	73	4.9 (11)	4.9 (4)			
Nuclear hormone receptors	270		3.3 (10)			
Peroxisome localization	18		15.0 (3)			
Collagen	179				2.8 (11)	12 (41)
Worm-yeast orthologs	1921				1.7 (72)	
Lectins	220			2.8 (19)		
Lipid metabolism genes	303	2.9 (27)				
Cell adhesion genes	49					13.3 (12)
Cell structure genes	219					3.0 (12)
Mitochondria localization	179				4.6 (18)	
Biosynthesis	478			2.1 (31)	5.1 (53)	
Basement membrane	14				13.0 (4)	
Fatty acid oxidation	48				4.8 (5)	
Other phosphatases	31			5.3 (5)		
Carbohydrate metabolism	121				4.9 (13)	
Amino acid metabolism	104			5.0 (16)	7.0 (16)	
Energy generation genes	117				6.2 (16)	
Translation factors	41			4.8 (6)		
Transporters	407		2.9 (13)	2.9 (36)		
Small molecule transport	561	2.2 (37)				
Passive channels	203	3.2 (20)				
Protein degradation	250			2.2 (17)		
Protease	186			3.3 (19)		

B Gene expression mountains

Gene class*	Total on array	Dauer-enriched†	Transient	Early	Climbing	Late
Genome	17761	540	195	538	386	325
Mount 01	1818		1.7 (34)	1.5 (85)		
Mount 06	909	4.5 (126)	2.8 (28)			
Mount 08	803	2.2 (55)			2.6 (45)	
Mount 14	353				6.3 (49)	26 (169)
Mount 15	247	8.6 (65)				
Mount 16	230				5.6 (28)	11.5 (49)
Mount 19	189			8.8 (51)	5.8 (24)	
Mount 20	160			3.9 (19)	5.1 (18)	
Mount 21	154	2.8 (13)	16.0 (27)	4.0 (19)		
Mount 22	151	2.8 (13)				
Mount 23	143				3.8 (12)	
Mount 24	133		10.0 (15)	7.9 (32)		
Mount 27	87			11.0 (30)		
Mount 29	40					12 (9)
Mount 31	25			9.2 (7)		
Mount 41	7			19.0 (4)		

*Over-represented functional classes in the dauer recovery subgroups. For clarity, gene groups with $P>0.001$ are not listed. Gene lists are shown in supplemental Table S7.

†The representation factor is shown and the number of overlapping genes between the two groups is in parentheses. All representation factors are significant at $P<0.001$. The probability was determined using the normal approximation of the hypergeometric probability distribution.

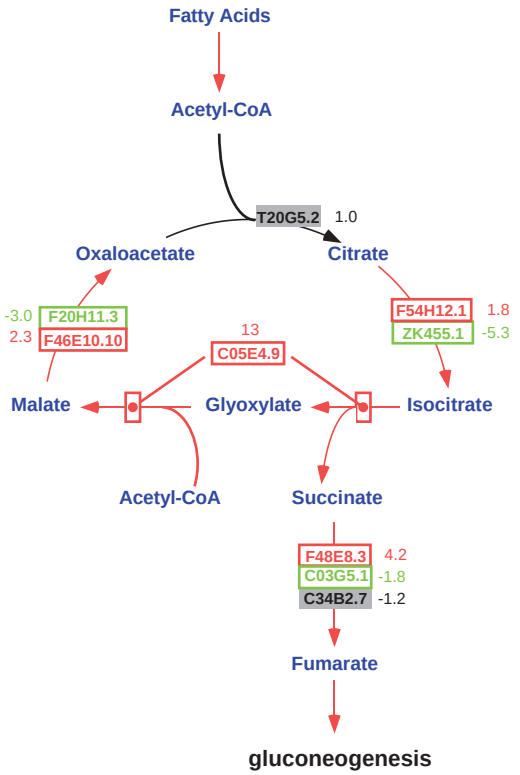

Fig. 4. Dauer expression of glyoxylate pathway genes. The glyoxylate pathway genes show increased expression in dauers. Red lettering and boxes refer to genes that are dauer enriched (P<0.01) when comparing dauers at 0 hours with dauers at 12 hours. The corresponding red arrows imply increased levels of enzymatic activity that has not been experimentally verified. Green lettering and boxes refer to genes and enzymatic steps that are downregulated (P<0.01) when comparing dauer at 0 hours to dauer at 12 hours. Black lettering in gray boxes indicates no significant difference in expression. The magnitude of induction or repression for each gene is indicated. Some key metabolic intermediates are shown.

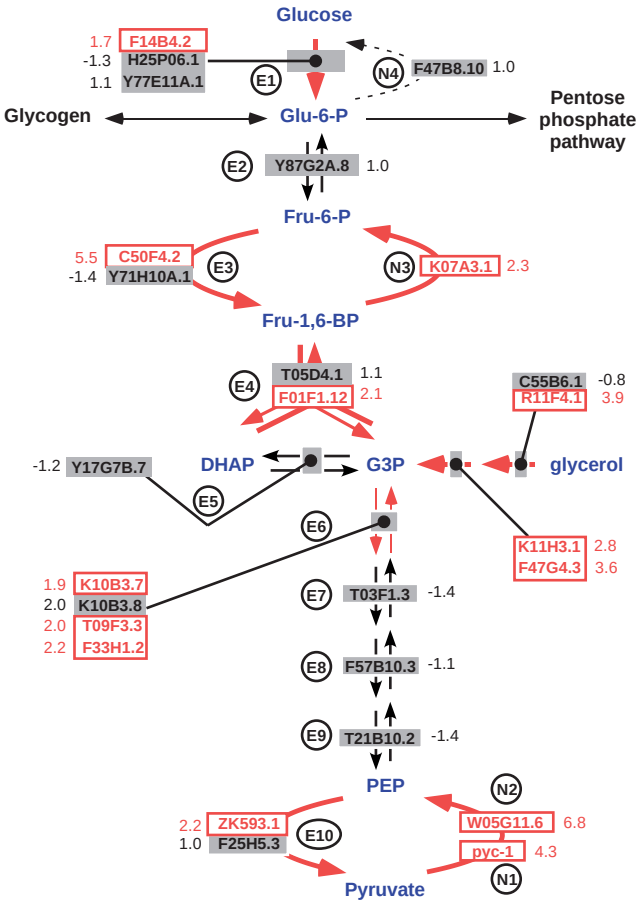

Fig. 5. Dauer expression of glycolysis and gluconeogenesis genes. The glycolytic and gluconeogenic pathways are upregulated in dauers, as inferred from changes in gene expression. None of the *C. elegans* genes encoding the enzymes that catalyze each step in these two pathways has been explicitly defined, but the ORFs illustrated show significant homology to the respective enzymes. Red lettering and boxes refer to genes that are dauer enriched (P<0.01) when comparing pure dauers at 0 hours to 12 hours. Correspondingly, red arrows indicate inferred high levels of enzyme activity in dauers. Black lettering in gray boxes indicates no significant difference in expression. The magnitude of induction or repression for each gene is indicated. Lines that end in black circles connect genes with the enzymatic step. Some key metabolic intermediates are shown. Glucose-6-phosphatase activity (N4) is depicted as a broken line because there is no clear homolog that encodes that activity. F47B8.10 encodes a glucose-6-phosphate translocase, which functions in a complex with the phosphatase in vertebrates. Glycolysis enzymes: E1, hexokinase; E2, glucose-6-phosphate isomerase; E3, phosphofructokinase; E4, fructose-bisphosphate aldolase; E5, triose phosphate isomerase; E6, glyceraldehyde-3-phosphate dehydrogenase; E7, phosphoglycerate kinase; E8, phosphoglycerate mutase; E9, enolase; E10, pyruvate kinase. Gluconeogenic enzymes: N1, pyruvate carboxylase; N2, phosphenolpyruvate carboxykinase; N3, fructose-1,6-bisphosphatase; N4, glucose-6-phosphatase. Glu-6-P, glucose-6-phosphate; Fru-6-P, fructose-6-phosphate; Fru-1,6-BP, fructose-1,6-bisphosphate; DHAP, dihydroxyacetone phosphate; G3P, glyceraldehyde-3-phosphate; PEP, phosphoenolpyruvate.

(succinate dehydrogenase) and F46E10.10 (malate dehydrogenase) are more abundantly expressed in the dauer relative to the end of the dauer exit timecourse (Fig. 4; Table 3). Thus, we observe that the genes encoding four of the five enzymatic activities of the glyoxylate cycle are enriched in dauers suggesting that there is a greater capacity to metabolize fatty acids via this pathway. Furthermore, as there is more than one isoenzyme for aconitase and succinate dehydrogenase, we may have identified the genes that encode the glyoxylate cycle-specific isoenzymes.

Third, the gluconeogenic pathway appears to be active in dauers as three out of the four key gluconeogenic bypass enzymes are dauer-enriched (Fig. 5): *pyc-1* (pyruvate carboxylase, 4.3-fold, P<0.001), W05G11.6 (phosphoenolpyruvate carboxykinase, 6.8-fold, P<0.001) and K07A3.1 (fructose-1,6-bisphosphatase, FBP, 2.3-fold, P<0.002). The final step that generates glucose from glucose-6-phosphate requires both a glucose-6-phosphate translocase and a glucose-6-phosphatase. There is not an apparent glucose-6-phosphatase encoded by the *C. elegans* genome, indicating either that *C. elegans* cannot generate glucose from glucose-6-phosphate or an unknown gene has glucose-6-phosphatase activity. The reactions catalyzed by these four enzymes are essentially irreversible, and, thus, increased expression of three out of four enzymes indicates increased flux through the

Table 3. Expression ratios of shared TCA/glyoxylate cycle genes

Gene	Dauer-enrichment*	*P*	Predicted enzyme activity
F54H12.1	1.9	<0.010	Aconitase
ZK455.1	–5.3	<0.001	Aconitase
C03G5.1	–1.8	<0.006	Succinate dehydrogenase/fumarate reductase
C34B2.7	–1.2	<0.192	Succinate dehydrogenase/fumarate reductase
F48E8.3	4.2	<0.001	Cytoplasmic succinate dehydrogenase/fumarate reductase
F20H11.3	–3.0	<0.001	Mitochondrial malate dehydrogenase
F46E10.10	2.8	<0.002	Cytosolic malate dehydrogenase

*Average expression ratio of pure dauers at 0 hours versus pure dauers at 12 hours.

gluconeogenesis pathway to generate glucose-6-phosphate in dauers. Despite the absence of an identified glucose-6-phosphatase in the genome, glucose may still be generated in dauers as we observe increased expression of key glycolytic enzymes (below).

There are three essentially irreversible enzymatic steps in glycolysis (those catalyzed by hexokinase, phosphofructokinase and pyruvate kinase), and the microarray data show that the genes that encode all three of these steps are enriched in dauers. Specifically, F14B4.2 (hexokinase, 1.7-fold, $P<0.003$), C50F4.2 (phosphofructokinase, 5.5-fold, $P<0.001$) and ZK593.1 (pyruvate kinase, 2.2-fold, $P<0.001$) are enriched in dauers relative to 12 hours after dauer exit (Fig. 5). Increased expression of phosphofructokinase gene expression in dauers is consistent with previous results (O'Riordan and Burnell, 1989). These data indicate that flux through glycolysis may be higher in dauers relative to the end of the dauer exit timecourse. The two major sources of glucose for glycolysis are extracellular or derived from the breakdown of glycogen. Because the genes that encode enzymes of glycogen metabolism are not upregulated in dauers (data not shown), extracellularly derived glucose may be the primary substrate for glycolysis

Glycolysis and gluconeogenesis are competing processes that are both shown to be enriched in dauers in the dauer exit timecourse microarray experiments. As co-expression of both metabolic pathways would generate a futile cycle, it is possible that cells expressing glycolytic enzyme activity are different from those that express the gluconeogenic enzymes. For example, fatty acids are stored in the intestine and epidermis of dauers (Ogg et al., 1997). Glucose is generated from fatty acids via gluconeogenesis in these tissues and may be transported to other tissues, such as neurons and muscles, and serve as an energy source. Tissues receiving glucose from the intestine and hypodermis would be active in glycolysis for energy production. Alternatively, increased mRNA levels might not reflect enzymatic activity as some glycolytic and gluconeogenic enzymes are under allosteric regulation (Stryer, 1995). For example, the gluconeogenic enzyme FBP activity is under tight metabolic allosteric control (Stryer, 1995). We found that K07A3.1, which encodes FBP, is 2.3-fold dauer-enriched but previously, FBP enzyme activity had been shown to be low in dauers relative to adults (O'Riordan and Burnell, 1989).

Gene expression mountains that are over-represented in the dauer-enriched gene class

Five mountains are over-represented by the dauer-enriched gene class upon comparison to the expression topomap: mounts 6, 8, 15, 21 and 22 (Table 2). Neuronal genes are over-represented in mount 6 (Kim et al., 2001), suggesting that other genes in this mountain may also have neuronal function. Dauers display different behaviors than non-dauers, such as seeking different temperatures, crawling up objects, projecting their heads into the air and waving it back and forth, perhaps to permit attachment to passing insects and allow transportation to more fertile soil (Riddle, 1988). The dauer-enriched genes in mount 6 could be expressed in neurons and act to modify these dauer-specific behaviors.

We examined the expression profile of all the genes in each of the over-represented mountains in the dauer exit timecourse. The overlap between the dauer-enriched gene class and mount 15 is particularly interesting. Mount 15 contains a total of 247 genes, of which 65 genes were selected from the microarray experiments as dauer-enriched using stringent statistical criteria ($P<0.001$). We found that essentially all of the genes in mount 15 are dauer enriched (Fig. 6B), suggesting that in addition to the 65 genes previously selected from the dauer exit timecourse, the other 182 genes in mount 15 are likely to be dauer enriched, although at a level below the one used in our stringent selection.

Genes in mount 15 are both aging and dauer regulated, as they are expressed at higher levels in old versus young adults and in dauers versus non-dauer worms (Lund et al., 2002). The dauer animal is considered essentially to be non-aging, as they do not appear to senesce (Klass and Hirsh, 1976). Hence, mount 15 may identify a subset of aging and dauer-regulated genes involved in a common mechanism to prolong life and delay senescence.

Most of the genes in mount 22 appear to be dauer enriched. However, this result is probably an artifact of adult carcasses contaminating the dauer populations used in the timecourse experiments, as the genes in mount 22 do not appear to be dauer enriched using a pure population of dauers (see Materials and Methods; see supplemental tables at http://dev.biologists.org/supplemental/ and at http://cmgm.stanford.edu/~kimlab/dauer/).

Transient

The expression of transient genes peaks around two hours after introduction of food and then declines (Fig. 3). The expression of these genes peak after change in surface lipophilicity (about 30 minutes) and commitment to exit (about 1 hour), but before many morphological and behavioral changes, such as pharyngeal pumping (about 3 hours) and increase in the diameter of the worm (Cassada and Russell, 1975; Golden and

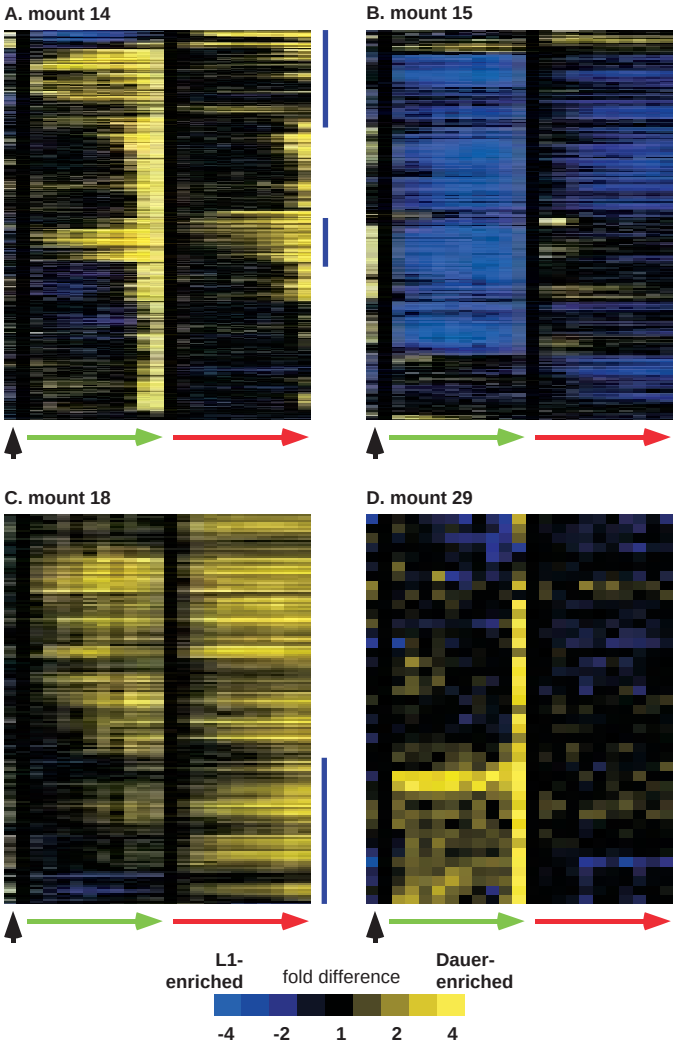

Fig. 6. Expression patterns of genes in four gene expression mountains. Hierarchical clustering was used to display the expression ratio of dauers to L1 worms at 0 hours and the expression ratios during the dauer exit and L1 starvation timecourses for genes in four mountains. (A) Mount 14, 353 genes. Blue bars highlight genes that are not induced late in the dauer exit timecourse. (B) Mount 15, 247 genes. (C) Mount 18, 190 genes. Blue bar highlights genes that do not have a dauer exit timecourse expression pattern that is similar to the L1 starvation timecourse expression pattern. (D) mount 29, 40 genes. Almost all genes in mount 29 are induced at the dauer-L4 (or L3-L4) molt. Rows represent different genes and columns represent different times during the timecourse, except the first column [which is the ratio of dauer to L1 at 0 hours (arrowhead)]. Green and red arrows are the dauer exit and L1 starvation timecourses, respectively. Scale shows level of expression. Full data for this figure are available at http://dev.biologists.org/supplemental/ and at http://cmgm.stanford.edu/~kimlab/dauer/

Riddle, 1984; Proudfoot et al., 1993). These genes could be an early transcriptional response that sets up the dauer recovery process. In order for dauer recovery to be successful, all of the cells in an animal must exit the dauer stage synchronously. One plausible mechanism to coordinate dauer exit would be via distribution of a hormone(s) throughout the entire body. We see four gene classes that may regulate the synthesis and distribution of a hormone important for dauer recovery.

First, genes encoding cytochrome P450 enzymes are over-represented in the transient class of dauer exit genes (Table 2). There are 81 cytochrome P450 genes in the genome, and eight of these are transient genes in the dauer exit timecourse (5.6-fold over-representation, $P<0.001$). Cytochrome P450 enzymes are mono-oxygenases that metabolize many endogenous and exogenous lipophilic compounds including steroidal hormones, xenobiotics and fatty acids (Mansuy, 1998). The cytochrome P450 enzymes and perhaps other biosynthetic enzymes could synthesize a dauer exit hormone(s). *daf-9* encodes a cytochrome P450 involved in regulating dauer development (Antebi et al., 2000; Gerisch et al., 2001; Jia et al., 2002). The microarray experiments show that *daf-9* expression increases early and then remains high during the dauer exit timecourse.

Second, genes encoding UDP-glucuronosyltransferases are over-represented in the transient class of dauer exit genes. There are 73 UDP-glucuronosyltransferase genes in the genome, and four of these are transient genes in the dauer exit timecourse (4.9-fold over-representation, $P<0.001$). UDP-glucuronosyltransferases facilitate the export of lipophilic molecules by attaching sugar residues (Tukey and Strassburg, 2000). UDP-glucuronosyltransferases could function within the signaling cell by conjugating a sugar residue to a putative dauer exit hormone, thereby facilitating its export and distribution to the rest of the body. Alternatively, they might function within the receiving cell as an initial cellular response to attenuate the hormonal signal.

Third, genes encoding transporters are over-represented in the transient class of dauer exit genes. There are 407 transporter genes in the genome, and 13 of these are transient genes in the dauer exit timecourse (2.9-fold over-representation, $P<0.001$). Transporters facilitate the movement of molecules across cell membranes. Of the genes encoding transporters in the transient gene class, four encode members similar to the multidrug resistance protein family. The multidrug resistance proteins are a family of transporters that are often elevated in expression in cells resistant to toxic compounds (e.g. multidrug resistant cancer cell lines) (Dean et al., 2001). The physiological role of some multidrug resistance proteins include transporting steroidal and lipophilic signaling molecules (Dean et al., 2001). One of the transporters, *pgp-1*, is implicated in arsenite and cadmium resistance, while the specific functions of the other three transporters (F14D7.6, T21E8.1 and T21E8.2) are unknown and could be to export a dauer-exit hormonal molecule (Broeks et al., 1996).

Fourth, genes encoding nuclear hormone receptors are over-represented in the transient class of dauer exit genes. There are 270 nuclear hormone receptor genes in the genome, and 10 of these are transient genes in the dauer exit timecourse (3.3-fold over-representation, $P<0.001$). Nuclear hormone receptors are ligand-activated transcription factors that are involved in many diverse process such as metabolic regulation, sexual differentiation and embryonic development. The large number of induced nuclear hormone receptors present in the transient class of dauer exit genes might reflect redundant functions, cell- or tissue-specific expression, a specific function for each nuclear hormone receptor or a combination of these possibilities.

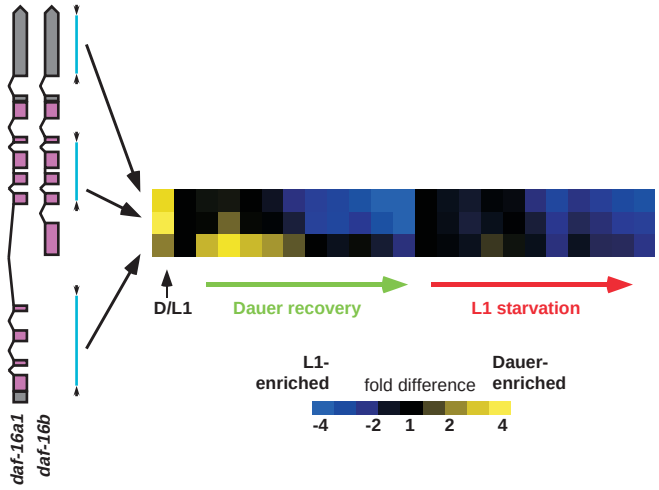

Fig. 7. Differential expression of *daf-16* isoforms. Intron and exon structures of the *daf-16a* and *daf-16b* isoforms are depicted. Pink boxes correspond to coding sequences and gray boxes are 5′ or 3′ untranslated regions. Light blue lines with inward facing arrows represent predicted amplified PCR products that are on the microarray. The corresponding expression profile for each PCR product is depicted. Columns represent different times during the timecourse except the first column (D/L1) which is the ratio of dauer to L1 at 0 hours. Scale shows level of expression.

daf-16 encodes a Forkhead transcription factor and is a member of the transient gene class (Ogg et al., 1997). *daf-16* promotes dauer formation and functions downstream of the *daf-2*/insulin-like receptor signal transduction pathway that regulates dauer development and longevity (Dorman et al., 1995; Tissenbaum and Ruvkun, 1998). *daf-16* has at least three alternatively spliced forms designated *daf-16a1*, *daf-16a2* and *daf-16b*. The Forkhead DNA binding domains of DAF-16a and DAF-16b are highly related but distinct.

DAF-16B but not DAF-16A is expressed in the pharynx and is required for pharynx remodeling during dauer formation (Henderson and Johnson, 2001; Lee et al., 2001; Lin et al., 2001; Ogg et al., 1997). Our microarray data show different expression kinetics for the different *daf-16* isoforms; specifically, *daf-16a* is transiently induced, whereas *daf-16b* decreases expression during the dauer-exit timecourse. There are three spots on the microarray that correspond to different regions of *daf-16*: R13H8.2, R13H8.1 and DAF-16#2. R13H8.2 hybridizes to *daf-16a* but not *daf-16b* and exhibits transient kinetics in the dauer exit timecourse (Fig. 7). R13H8.1 and DAF-16#2 hybridize to both *daf-16* isoforms and are downregulated during the dauer exit timecourse. These results suggest that *daf-16a* may have both dauer formation and dauer recovery functions, whereas *daf-16b* might be important for the maintenance of some aspect of the dauer state (e.g. the pharynx).

We compared the genes showing transient expression in the dauer exit timecourse to the mountains in the gene expression topomap, and found that four mountains are over-represented: mounts 1, 6, 21 and 24 (Table 2). This comparison partitioned the genes in the transient gene class into more refined compartments. For example, partitioning the nuclear hormone receptors contained in the transient class into different gene expression mountains suggests possible functions. Specifically, mount 6 is enriched for neuronal genes, suggesting that the two transient class nuclear hormone receptor genes contained in this mountain (F07C3.10 and F16B4.12) might function in neurons during dauer recovery. Mount 1 is enriched for muscle and neuronal genes, suggesting that the three transient nuclear hormone receptor genes in mount 1 (*nhr-45*, *nhr-46* and *nhr-63*) might function in these tissues. Mount 21 is enriched in lipid metabolism genes, so the transient class nuclear hormone receptor gene contained in this mountain (R11G11.2) might be associated with that process during dauer recovery.

Early

Early genes are induced early in the dauer exit timecourse, and then remain induced (Fig. 3). Genes that are expressed early in dauer exit could be required for resumption into an active state from a quiescent one. For example, normal worms eat and digest food to generate energy for sustenance and growth. Proteases and acid phosphatases (in the other phosphatases gene group) are involved with digestion and are over-represented 3.3- and 5.3-fold in the early gene class (Table 2). There is also an increase in sugar and amino acid transporters (2.9-fold over-represented, Table 2) that may function in food uptake.

Genes that encode six ABC transporters/multidrug resistance family proteins are found in the early gene class: *mrp-2*, *pgp-3*, C44B7.8, T28C12.2, W04C9.1 and ZK484.2. One gene (C44B7.8) encodes a peroxisomal ABC transporter that specifically transports long-chain fatty acids across the peroxisomal membrane for fatty acid β-oxidation. Two (*mrp-2* and *pgp-3*) are involved in toxic metal ions or cholchine/chloroquine resistance (Broeks et al., 1996; Broeks et al., 1995). The precise function of the remaining three genes is unknown. Worms ingest potential toxins as they exit the dauer state. These three ABC transporter/multidrug resistance family proteins could export a variety of these toxic molecules, and thereby confer resistance to them.

Climbing

Climbing genes have continually increasing expression during the dauer exit timecourse (Fig. 3). The transient and early genes that encode regulatory proteins such as transcription factors might control, directly or indirectly, the expression of genes in the climbing class.

Energy is generated from food through the glycolytic, tricarboxylic acid cycle (TCA), fatty acid oxidation and oxidative phosphorylation pathways. We observed that genes involved in these pathways are over-represented in the climbing gene class (Table 2): glycolysis (6.5-fold), TCA (9.1-fold), fatty acid oxidation (4.8-fold) and mitochondrial (4.6-fold) gene classes. The mitochondrial gene class includes the oxidative phosphorylation enzymes that are important for ATP generation. All of these biological gene groups would also be expected to be required during the L1 starvation timecourse. In general, these biological gene groups are regulated in both the dauer exit and L1 starvation timecourses, but to a different extent or with different kinetics.

Late

The late genes are induced just before the dauer-to-L4 molt in the dauer exit timecourse. Collagens are expressed in

preparation for a molt in order to synthesize cuticle for the next larval stage. We find collagen genes to be over-represented 12-fold in the genes showing late kinetics during dauer exit (Table 2). Other genes associated with cuticle or collagen processing are also expressed at this time. For example, M153.1 encodes a pyrroline-5-carboxylate reductase (which catalyzes the final step in proline biosynthesis) and is expressed at a late time during dauer exit. This is consistent with increased collagen biosynthesis as proline is abundant in collagens.

The *C. elegans* genome does not encode a homolog to *Drosophila hedgehog* (*hh*). However, it does encode 46 genes that are distantly related to *hh* called, *warthog* (*wrt-1* to *wrt-10*), *groundhog* (*grd-1* to *grd-14*) and *ground-like* (28 *grl* genes) (Aspock et al., 1999) (see supplemental Fig. S3 at http://dev.biologists.org/supplemental/ and at http://cmgm.stanford.edu/~kimlab/dauer/). Ten *wrt*, *grd* or *grl* genes are present in the late gene set, corresponding to a 12-fold enrichment that is statistically significant (hypergeometric, $P<0.001$). In flies and vertebrates, members of the *hh* family encode secreted signaling molecules important for anteroposterior patterning and cellular differentiation events (Hammerschmidt et al., 1997), although in *C. elegans* the function of *wrt*, *grd* and *grl* genes is unknown. Our data suggest that some of the *wrt*, *grd* and *grl* genes may function during the dauer-to-L4 molt or during molting in general (see supplemental Fig. S3 at http://dev.biologists.org/supplemental/ and at http://cmgm.stanford.edu/~kimlab/dauer/).

hh signaling leads to increased *ptc* transcription (Marigo et al., 1996). Seven Patched-related/Niemann-Pick type C (NPC)-related genes, corresponding to 17-fold over-representation (hypergeometric, $P<0.001$), are found in the late group. These data support the possibility that the *C. elegans* Patched-related/NPC-related genes might be transcriptionally activated by *wrt*, *grd* and *grl* signaling. Alternatively, human NPC1 is implicated in retrograde sterol transport from lysosomes (Neufeld et al., 1999), so increased expression of these seven genes may reflect an increased need to mobilize cholesterol or its derivatives at this time.

Three mountains on the expression topomap are over-represented among genes showing late kinetics during dauer exit: mounts 14, 16 and 29. Mount 16 is enriched in muscle genes and collagen (Kim et al., 2001). There are 49 genes in the late class that are also part of mount 16. Of these, 31 are collagen genes. The remaining 18 genes might function in association with collagens or in muscle. No common biological function was previously noted for genes in mount 29. We plotted all the genes in mount 29 in the dauer exit and L1 starvation timecourses, and found that nearly all were induced at the dauer-L4 (or L3-L4) molt (Fig. 6D). This result suggests that genes in mount 29 may be involved in these molts. The genes in mount 14 show a heterogeneous expression response in the dauer exit and L1 starvation timecourses (Fig. 6A).

DISCUSSION

Overview

We have used DNA microarrays to profile gene expression in the dauer larvae and after feeding of starved worms. We used statistical criteria to determine which genes are dauer exit specific and which are common to both dauer exit and L1-feeding. For the dauer exit-specific genes, we have further subdivided those genes into five clusters with similar kinetic profiles. Our experiments provide a global molecular description of the dauer exit process.

Comparison to SAGE analysis

The dauer larvae transcriptional profile has also been analyzed using serial analysis of gene expression (SAGE) (Jones et al., 2001). By comparing the transcript tag abundances found in dauer with that of a mixed-staged population, Jones et al. identified a set of 358 dauer-specific tags ($P<0.05$). Of these, 305 are present on the DNA microarrays used in the experiments from this paper, and 43 (14%) overlap with the 540 dauer-enriched genes defined by our criteria. Hence, there is a poor overlap between the SAGE data and dauer-enriched genes identified by our microarray data. One possibility for the poor agreement is that the SAGE experiment selected genes with a lower stringency than the DNA microarray experiments ($P<0.05$ in the SAGE dataset versus $P<0.001$ in the microarray dataset). A more stringent criteria of $P<0.001$ for the SAGE dataset requires at least eight dauer-specific tags and corresponds to 53 dauer-specific genes. However, only 20 out of the 53 genes selected by SAGE (38%) are also selected by the microarray analysis. Second, the source of dauer larvae is different in the two experiments; the SAGE experiments prepared dauers from liquid cultures, whereas the microarray experiments prepared dauers from agar plates. Third, the experimental design and analysis were different; the SAGE experiments defined a set of dauer-specific genes by comparing dauers to mixed-staged animals, while the microarray experiments defined a set of dauer-enriched genes using a dauer exit timecourse. Finally, the SAGE analysis was performed using only one biological sample, whereas we repeated our experiments using eleven time points from four independent timecourses. Thus, some of the dauer-specific genes identified by the SAGE experiment may not be reproducible in other samples of dauers.

Evaluating gene expression changes

In addition to functioning in the dauer itself, the dauer-enriched genes may be expressed in dauers in preparation for exit to normal development. As the primary survival strategy of *C. elegans* is to propagate successfully as often as possible, exiting the dauer state rapidly and efficiently may be important. The pre-dauer animal may express a stored mRNA pool in preparation for dauer recovery. Consistent with this, transcriptional inhibition does not prevent the initiation of dauer recovery (Reape and Burnell, 1991).

Global transcription is depressed in the dauer larvae but increases dramatically upon exit (Dalley and Golomb, 1992). The microarray experiments assay changes in mRNA levels for one gene relative to other genes in the sample, and does not measure changes in absolute mRNA levels. Hence, genes that show increased expression in the microarray analysis are those that induce expression even more than would be expected due to the global increase in expression during dauer exit. Conversely, genes that show decreased expression in the microarray analysis do so relative to the general increase exhibited by the rest of the genome. These genes may maintain steady state levels or even show small expression increases that are smaller than the general level of increase in mRNA accumulation.

We would like to thank the members of the Kim laboratory for useful advice and comments. This work was supported by grants from the National Institute of General Medical Sciences and from the National Center for Research Resources.

REFERENCES

Aboobaker, A. A. and Blaxter, M. L. (2000). Medical significance of Caenorhabditis elegans. *Ann. Med.* **32**, 23-30.

Antebi, A., Yeh, W. H., Tait, D., Hedgecock, E. M. and Riddle, D. L. (2000). daf-12 encodes a nuclear receptor that regulates the dauer diapause and developmental age in C. elegans. *Genes Dev.* **14**, 1512-1527.

Aspock, G., Kagoshima, H., Niklaus, G. and Burglin, T. R. (1999). Caenorhabditis elegans has scores of hedgehog-related genes: sequence and expression analysis. *Genome Res.* **9**, 909-923.

Bird, D. M., Opperman, C. H., Jones, S. J. and Baillie, D. L. (1999). THE CAENORHABDITIS ELEGANS GENOME: A guide in the post genomics age. *Annu. Rev. Phytopathol.* **37**, 247-265.

Birnby, D. A., Link, E. M., Vowels, J. J., Tian, H., Colacurcio, P. L. and Thomas, J. H. (2000). A transmembrane guanylyl cyclase (DAF-11) and Hsp90 (DAF-21) regulate a common set of chemosensory behaviors in caenorhabditis elegans. *Genetics* **155**, 85-104.

Blaxter, M. and Bird, D. (1997). Parasitic nematodes. In C. elegans *II*, pp. 851-878. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Broeks, A., Janssen, H. W., Calafat, J. and Plasterk, R. H. (1995). A P-glycoprotein protects Caenorhabditis elegans against natural toxins. *EMBO J.* **14**, 1858-1866.

Broeks, A., Gerrard, B., Allikmets, R., Dean, M. and Plasterk, R. H. (1996). Homologues of the human multidrug resistance genes MRP and MDR contribute to heavy metal resistance in the soil nematode Caenorhabditis elegans. *EMBO J.* **15**, 6132-6143.

Burglin, T. R., Lobos, E. and Blaxter, M. L. (1998). Caenorhabditis elegans as a model for parasitic nematodes. *Int. J. Parasitol.* **28**, 395-411.

Cassada, R. C. and Russell, R. L. (1975). The Dauerlarva, a post-embryonic developmental variant of the nematode Caenorhabditis elegans. *Dev. Biol.* **46**, 326-342.

Chervitz, S. A., Aravind, L., Sherlock, G., Ball, C. A., Koonin, E. V., Dwight, S. S., Harris, M. A., Dolinski, K., Mohr, S., Smith, T. et al. (1998). Comparison of the complete protein sets of worm and yeast: orthology and divergence. *Science* **282**, 2022-2028.

Dalley, B. K. and Golomb, M. (1992). Gene expression in the Caenorhabditis elegans dauer larva: developmental regulation of Hsp90 and other genes. *Dev. Biol.* **151**, 80-90.

Dean, M., Hamon, Y. and Chimini, G. (2001). The human ATP-binding cassette (ABC) transporter superfamily. *J. Lipid Res.* **42**, 1007-1017.

DeRisi, J. L., Iyer, V. R. and Brown, P. O. (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278**, 680-686.

Dorman, J. B., Albinder, B., Shroyer, T. and Kenyon, C. (1995). The age-1 and daf-2 genes function in a common pathway to control the lifespan of Caenorhabditis elegans. *Genetics* **141**, 1399-1406.

Eisen, M. B., Spellman, P. T., Brown, P. O. and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **95**, 14863-14868.

Epstein, H. F. and Shakes, D. C. (eds) (1995). Caenorhabditis elegans*: Modern Biological Analysis of an Organism* (*Methods in Cell Biology*). London: Academic Press.

Gerisch, B., Weitzel, C., Kober-Eisermann, C., Rottiers, V. and Antebi, A. (2001). A hormonal signaling pathway influencing C. elegans metabolism, reproductive development, and life span. *Dev. Cell* **1**, 841-851.

Golden, J. W. and Riddle, D. L. (1984). The Caenorhabditis elegans dauer larva: developmental effects of pheromone, food, and temperature. *Dev. Biol.* **102**, 368-378.

Gottlieb, S. and Ruvkun, G. (1994). daf-2, daf-16 and daf-23: genetically interacting genes controlling Dauer formation in Caenorhabditis elegans. *Genetics* **137**, 107-120.

Hammerschmidt, M., Brook, A. and McMahon, A. P. (1997). The world according to hedgehog. *Trends Genet.* **13**, 14-21.

Henderson, S. T. and Johnson, T. E. (2001). daf-16 integrates developmental and environmental inputs to mediate aging in the nematode Caenorhabditis elegans. *Curr. Biol.* **11**, 1975-1980.

Honda, Y. and Honda, S. (1999). The daf-2 gene network for longevity regulates oxidative stress resistance and Mn-superoxide dismutase gene expression in Caenorhabditis elegans. *FASEB J.* **13**, 1385-1393.

Jia, K., Albert, P. S. and Riddle, D. L. (2002). DAF-9, a cytochrome P450 regulating C. elegans larval development and adult longevity. *Development* **129**, 221-231.

Jiang, M., Ryu, J., Kiraly, M., Duke, K., Reinke, V. and Kim, S. K. (2001). Genome-wide analysis of developmental and sex-regulated gene expression profiles in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. USA* **98**, 218-223.

Jones, S. J., Riddle, D. L., Pouzyrev, A. T., Velculescu, V. E., Hillier, L., Eddy, S. R., Stricklin, S. L., Baillie, D. L., Waterston, R. and Marra, M. A. (2001). Changes in gene expression associated with developmental arrest and longevity in Caenorhabditis elegans. *Genome Res* **11**, 1346-1352.

Kenyon, C., Chang, J., Gensch, E., Rudner, A. and Tabtiang, R. (1993). A C. elegans mutant that lives twice as long as wild type. *Nature* **366**, 461-464.

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

Kimura, K. D., Tissenbaum, H. A., Liu, Y. and Ruvkun, G. (1997). daf-2, an insulin receptor-like gene that regulates longevity and diapause in Caenorhabditis elegans. *Science* **277**, 942-946.

Klass, M. and Hirsh, D. (1976). Non-ageing developmental variant of Caenorhabditis elegans. *Nature* **260**, 523-525.

Kohonen, T. (1997). *Self-Organizing Maps*. New York: Springer.

Larsen, P. L. (1993). Aging and resistance to oxidative damage in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. USA* **90**, 8905-8909.

Lee, R. Y., Hench, J. and Ruvkun, G. (2001). Regulation of C. elegans DAF-16 and its human ortholog FKHRL1 by the daf-2 insulin-like signaling pathway. *Curr. Biol.* **11**, 1950-1957.

Lin, K., Hsin, H., Libina, N. and Kenyon, C. (2001). Regulation of the Caenorhabditis elegans longevity protein DAF-16 by insulin/IGF-1 and germline signaling. *Nat. Genet.* **28**, 139-145.

Lund, J., Tedesco, P., Duke, K., Wang, J., Kim, S. and Johnson, T. (2002). Transcriptional profile of aging in C. elegans. *Curr. Biol.* **12**, 1566-1573.

Mansuy, D. (1998). The great diversity of reactions catalyzed by cytochromes P450. *Comp. Biochem. Physiol. C Pharmacol. Toxicol. Endocrinol.* **121**, 5-14.

Marigo, V., Davey, R. A., Zuo, Y., Cunningham, J. M. and Tabin, C. J. (1996). Biochemical evidence that patched is the Hedgehog receptor. *Nature* **384**, 176-179.

Neufeld, E. B., Wastney, M., Patel, S., Suresh, S., Cooney, A. M., Dwyer, N. K., Roff, C. F., Ohno, K., Morris, J. A., Carstea, E. D. et al. (1999). The Niemann-Pick C1 protein resides in a vesicular compartment linked to retrograde transport of multiple lysosomal cargo. *J. Biol. Chem.* **274**, 9627-9635.

Ogg, S., Paradis, S., Gottlieb, S., Patterson, G. I., Lee, L., Tissenbaum, H. A. and Ruvkun, G. (1997). The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in C. elegans. *Nature* **389**, 994-999.

O'Riordan, V. B. and Burnell, A. M. (1989). Intermediary metabolism in the dauer larva of the nematode Caenorhabditis elegans – I Glycolysis, gluconeogenesis, oxidative phosphorylation and the tricarboxylic acid cycle. *Comp. Biochem. Physiol. B* **92B**, 233-238.

O'Riordan, V. B. and Burnell, A. M. (1990). Intermediary metabolism in the dauer larva of the nematode Caenorhabditis elegans – II The glyoxylate cycle and fatty-acid oxidation. *Comp. Biochem. Physiol. B* **95B**, 125-130.

Proudfoot, L., Kusel, J. R., Smith, H. V., Harnett, W., Worms, M. J. and Kennedy, M. W. (1993). Rapid changes in the surface of parasitic nematodes during transition from pre- to post-parasitic forms. *Parasitology* **107**, 107-117.

Reape, T. J. and Burnell, A. M. (1991). Dauer larva recovery in the nematode Caenorhabditis elegans – I The effects of mRNA synthesis inhibitors on recovery growth and pharyngeal pumping. *Comp. Biochem. Physiol. B* **98B**, 239-243.

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

Ren, P., Lim, C. S., Johnsen, R., Albert, P. S., Pilgrim, D. and Riddle, D. L. (1996). Control of C. elegans larval development by neuronal expression of a TGF-beta homolog. *Science* **274**, 1389-1391.

Riddle, D. L. (1988). The dauer larva. In *The Nematode* Caenorhabditis elegans, pp. 393-412. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Riddle, D. L. and Albert, P. S. (1997). Genetic and environmental regulation of dauer larva development. In C. elegans *II*: Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Romagnolo, B., Jiang, M., Kiraly, M., Breton, C., Begley, R., Wang, J., Lund, J. and Kim, S. K. (2002). Downstream targets of let-60 Ras in Caenorhabditis elegans. *Dev. Biol.* **247**, 127-136.

Roy, P. J., Stuart, J. M., Lund, J. and Kim, S. K. (2002). Chrosomal clustering of muscle-expressed genes in Caenorhabditis elegans. *Nature* **418**, 975-979.

Schena, M., Shalon, D., Davis, R. W. and Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467-470.

Stryer, L. (1995). *Biochemistry*. New York: W. H. Freeman and Company.

Taub, J., Lau, J. F., Ma, C., Hahn, J. H., Hoque, R., Rothblatt, J. and Chalfie, M. (1999). A cytosolic catalase is needed to extend adult lifespan in C. elegans daf-C and clk-1 mutants. *Nature* **399**, 162-166.

Tissenbaum, H. A., Hawdon, J., Perregaux, M., Hotez, P., Guarente, L. and Ruvkun, G. (2000). A common muscarinic pathway for diapause recovery in the distantly related nematode species Caenorhabditis elegans and Ancylostoma caninum. *Proc. Natl. Acad. Sci. USA* **97**, 460-465.

Tissenbaum, H. A. and Ruvkun, G. (1998). An insulin-like signaling pathway affects both longevity and reproduction in Caenorhabditis elegans. *Genetics* **148**, 703-717.

Tukey, R. H. and Strassburg, C. P. (2000). Human UDP-glucuronosyltransferases: metabolism, expression, and disease. *Annu. Rev. Pharmacol. Toxicol.* **40**, 581-616.

Wadsworth, W. G. and Riddle, D. L. (1988). Acidic intracellular pH shift during Caenorhabditis elegans larval development. *Proc. Natl. Acad. Sci. USA* **85**, 8435-8438.