Changes in Gene Expression Associated with Developmental Arrest and Longevity in Caenorhabditis elegans

Steven J.M. Jones, Donald L. Riddle, Anatoli T. Pouzyrev, Victor E. Velculescu, LaDeana Hillier, Sean R. Eddy, Shawn L. Stricklin, David L. Baillie, Robert Waterston and Marco A. Marra

References	Article cited in: **http://www.genome.org/cgi/content/full/11/8/1346#otherarticles**
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here**

Notes

To subscribe to *Genome Research* go to:
http://www.genome.org/subscriptions/

Changes in Gene Expression Associated with Developmental Arrest and Longevity in *Caenorhabditis elegans*

Steven J.M. Jones,[1,7] Donald L. Riddle,[2] Anatoli T. Pouzyrev,[2] Victor E. Velculescu,[3] LaDeana Hillier,[4] Sean R. Eddy,[5] Shawn L. Stricklin,[5] David L. Baillie,[6] Robert Waterston,[4] and Marco A. Marra[1]

[1]*Genome Sequence Centre, British Columbia Cancer Research Centre, Vancouver, British Columbia V5Z 4E6, Canada;* [2]*Molecular Biology Program and Division of Biological Sciences, University of Missouri, Columbia, Missouri 65211, USA;* [3]*Johns Hopkins Oncology Center, Baltimore, Maryland 21231, USA;* [4]*The Genome Sequencing Center, Washington University School of Medicine, St. Louis, Missouri 63108, USA;* [5]*Howard Hughes Medical Institute, Department of Genetics, Washington University School of Medicine, St. Louis Missouri 63110, USA;* [6]*Institute of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada*

Gene expression in a developmentally arrested, long-lived dauer population of *Caenorhabditis elegans* was compared with a nondauer (mixed-stage) population by using serial analysis of gene expression (SAGE). Dauer (152,314) and nondauer (148,324) SAGE tags identified 11,130 of the predicted 19,100 *C. elegans* genes. Genes implicated previously in longevity were expressed abundantly in the dauer library, and new genes potentially important in dauer biology were discovered. Two thousand six hundred eighteen genes were detected only in the nondauer population, whereas 2016 genes were detected only in the dauer, showing that dauer larvae show a surprisingly complex gene expression profile. Evidence for differentially expressed gene transcript isoforms was obtained for 162 genes. H1 histones were differentially expressed, raising the possibility of alternative chromatin packaging. The most abundant tag from dauer larvae (20-fold more abundant than in the nondauer profile) corresponds to a new, unpredicted gene we have named *tts–1* (transcribed telomere-like sequence), which may interact with telomeres or telomere-associated proteins. Abundant antisense mitochondrial transcripts (2% of all tags), suggest the existence of an antisense-mediated regulatory mechanism in *C. elegans* mitochondria. In addition to providing a robust tool for gene expression studies, the SAGE approach already has provided the advantage of new gene/transcript discovery in a metazoan.

The nematode *Caenorhabditis elegans* can enter the dauer diapause stage under conditions of high population density and limited food (Golden and Riddle 1984). In the soil, dauer larvae disperse to a fresh environment and resume development to the reproductive adult. Molecular genetic analyses have shown that the nervous system responds to environmental stimuli to regulate larval development via transforming growth factor (TGF)–β (Ren et al. 1996), cyclic guanosine monophosphate (cGMP) (Birnby et al. 2000), and insulin-like (Kimura et al. 1997) signaling pathways, which, in turn, act on a nuclear receptor (Antebi et al. 2000) to control dauer versus nondauer morphogenesis (Riddle and Albert 1997). Dauer larvae are long lived and resistant to environmental stress (Klass and Hirsh 1976). Mutations that reduce insulin-like signaling promote both dauer diapause and adult longevity (Kimura et al. 1997; Gems et al. 1998). As a step toward understanding the unique physiology of the long-lived dauer state, we compared the overall gene expression profiles of dauer larvae and nondauer, growing (mixed-stage) populations. Serial analysis of gene expression (SAGE) (Velculescu et al. 1995) was used to comprehensively survey the transcriptomes of the two populations and to search for previously unidentified transcripts. SAGE allows the detailed profiling of mRNA populations through the isolation of unique sequence tags from individual transcripts, typically 14 nucleotides in length. Detection and enumeration of the tags is conducted by the concatenation of the tags and the subsequent sequencing of concatemer clones. In the case of *C. elegans*, the identification of the cellular transcripts from the SAGE tag is aided greatly through the availability of a sequenced and annotated genome.

RESULTS AND DISCUSSION

Considering the arrested state of development, we suspected that the dauer transcriptome would not be complex. Transcription rates in the dauer stage previously were estimated to be approximately six- to seven-fold reduced relative to growing larvae (Dalley and Golomb 1992). Surprisingly, SAGE tags for 8449 transcripts were detected in the dauer stage. In total, SAGE tags were identified for 11,130 (58.8%) of the 18,931 *C. elegans* transcripts predicted to contain the *Nla*III (CATG) restriction site required for detection by using the SAGE technology. Of these, 29.2% of the transcripts were identified by

tags observed only once (singletons). Two thousand sixteen (18.1%) of the detected transcripts were found only in the dauer stage and 2681 (24.1%) only in the mixed stages, although many of these transcripts were detected at low levels. Dauer (358) and nondauer (533) -specific transcripts were detected at significance of $P \leq 0.05$. The dauer-specific genes (many of which encode novel proteins) are candidates for specifying functions associated with efficient life maintenance and longevity. The overall abundance of SAGE tags is shown in Figure 1. Of the transcripts detected, 5765 (51%) previously had been confirmed by expressed sequence tag (EST) data (Kohara 1996).

Thirty-three thousand forty unique tag sequences (tag species) representing 129,813 observed tags could not be unambiguously assigned to a predicted transcript. Of these, 1397 tag species matched either ESTs (sense strand) or unfinished sections of the *C. elegans* genomic sequence yet to be annotated. There were 23,628 (71.5%) singletons that lacked validating sequence data confirming sequence integrity. Unassigned tag species fell into three classes. (1) Two thousand seven hundred thirteen matched more than one gene and therefore lacked an unambiguous gene assignment. (2) Sixteen thousand three hundred seventy-nine could not be matched with either *C. elegans* genomic or cDNA sequence. These tags arose presumably because of sequencing errors, strain differences, or tags overlapping undiscovered splice junctions or poly(A) tails. (3) Twelve thousand five hundred fifty-one tag species (66.4% of these were singletons) matched genomic DNA sequence but could not be assigned to a predicted transcript. A subset of these tags derive from transcripts that currently are not predicted within the genomic annotation or where the gene structure is mispredicted. Of the 12,551 unmapped tag species, 6489 (63.4% of which were singletons) matched a transcript in the antisense orientation (Fig. 2). It is likely that most of the observed antisense tags are caused by mispriming of the oligo(dT) to internal poly(A) stretches after first-strand synthesis during SAGE library construction.

Mitochondrial transcripts were found to be highly represented in both SAGE libraries profiled. The five mitochondrial protein coding transcripts detectable by the SAGE methodology (cytochrome C oxidase subunits I, II, and III and NADH-ubiquinone oxidoreductase subunits 4 and 5) contributed 4247 observed tags (2.8%) in the dauer population and 2150 tags (1.4%) in the mixed stages. However, 3924 tags (2.6%) from the dauer library and 2302 (1.6%) from the mixed-stage profile could only be assigned to the antisense strand of mitochondrial genes (Table 1). The relative abundance of antisense and sense mitochondrial tags indicates that the generation of antisense tags from mitochondrial transcripts possibly is not an artifact caused by mispriming (Fig. 2). This observation is supported by the previous detec-
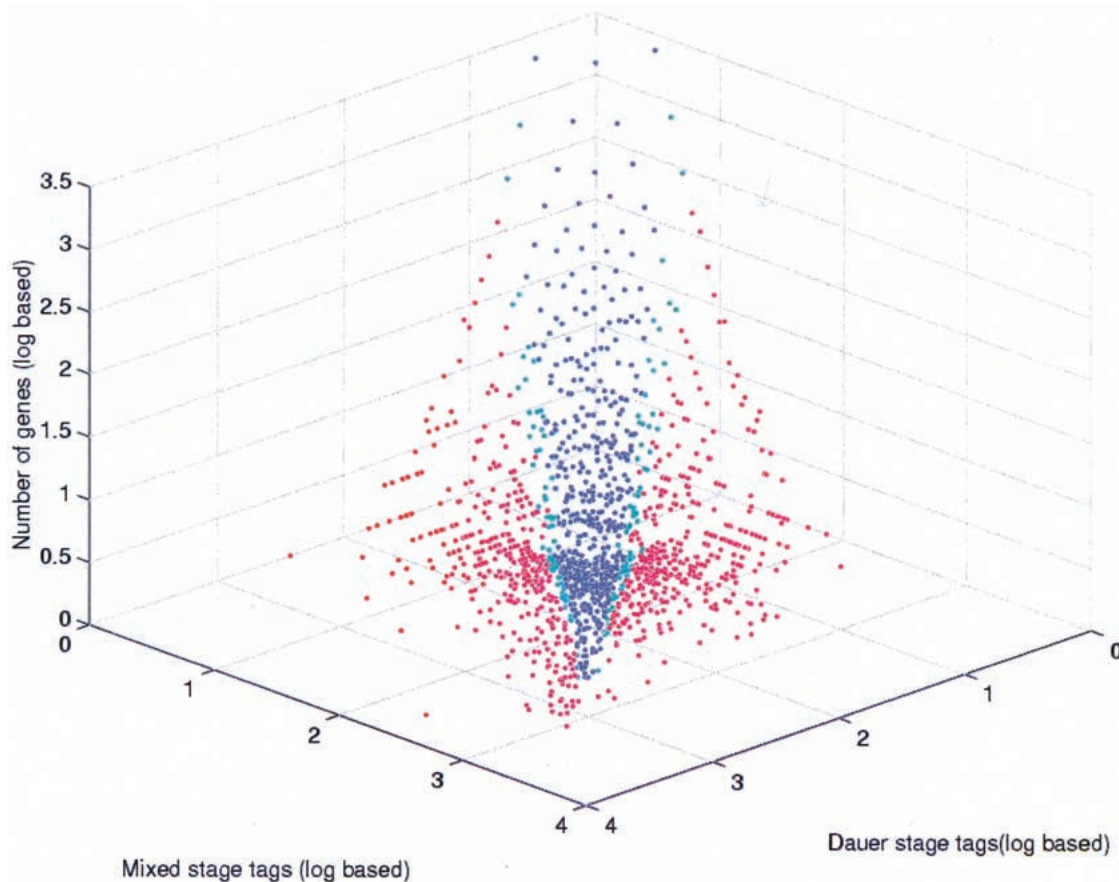

Figure 1 Expression profile comparing relative expression in dauer and mixed stages. Singleton tags were excluded. (blue) Tags for which no significant expression difference is observed; (green) significance is 95%–99%; (red) significance >99% confidence, as determined by the G test (Sokal and Rohlf 1995). The z-axis represents the number of transcripts with a specific mixed/dauer tag ratio.
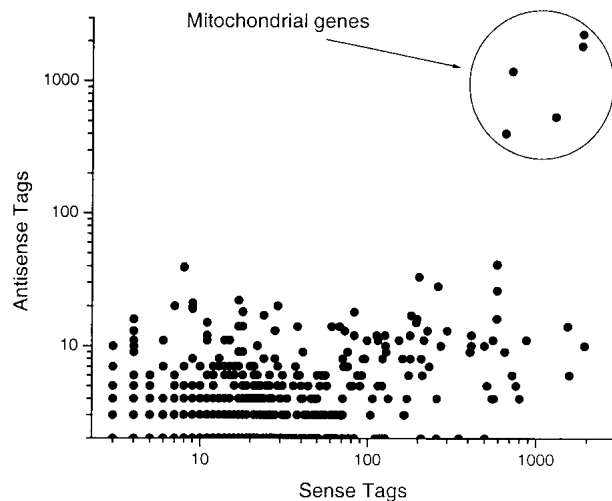
Figure 2 Comparison of sense and antisense tag frequencies. From 6489 tag sequences matching only an antisense transcript, 549 transcripts were studied in which both the antisense and sense SAGE tags could be unambiguously correlated. Only in 56 transcripts are the antisense transcripts in a higher abundance than the sense tags.

tion of stable polyadenylated antisense mRNA for both NADH-ubiquinone oxidoreductase in the rat (Tullo et al. 1994) and for cytochrome C oxidase in a human cell line (Shirafuji et al. 1997). As no proteins with RNAase activity currently are known to be imported into the *C. elegans* mitochondria (Costanzo et al. 2000), antisense RNA may serve to negatively regulate mitochondrial translation.

Genes previously implicated in longevity showed increased representation in the dauer expression profile (Table 2) including two superoxide dismutase genes (C08A9.1/*sod-3* and F55H2.1/*sod-4*) and two glutathione peroxidase genes adjacent to each other in the genome (C11E4.1 and C11E4.2). The tRNA dimethylallyltransferase (DMAPP transferase) *gro-1* gene (Heikimi et al. 1997), mutations in which increase longevity in *C. elegans*, was not detected among dauer-derived tags. DMAPP transferase is required for correct activation of tRNA molecules whose anticodons begin with U, and therefore its reduced expression in the dauer form should decrease the rate of protein synthesis (Dihanich et al. 1987). A subsequent increase in levels of DMAPP transferase might allow for the rapid resumption of protein synthesis on exit from the dauer stage. The expression of a single poly(A)-binding (PABP) like protein, elevated levels of which are known to decrease message turnover (Stambuk and Moon 1992), has been shown previously by subtractive cDNA library approach to be increased in dauer (Cherkasova et al. 2000). Two *C. elegans* genes, Y106G6H.2 and F18H3.3, show the greatest degree of similarity to human PABP. Y106G6H.2 (dauer 217; mixed 80) and both alternative transcripts of F18H3.3 (F18H3.3a dauer 12, mixed 5; F18H3.3b dauer 15, mixed 4) show increased levels of expression in the dauer stage. In this instance, the SAGE methodology discriminated between two paralogous genes and detected alternative transcripts.

The most abundant tag in the dauer stage (dauer, 4329 tags; mixed, 215 tags) derives from a transcript lacking a long open reading frame or protein similarity (dbEST nos. 1280045 and 1300904; GenBank accession number U41749, positions 17209–17831), which we have named *tts-1* (transcribed telomere-like sequence). This transcript is enigmatic, but it shares

characteristics with known telomerase RNAs, from ciliate through human (Chen et al. 2000). These include (1) a one-and-a-half-length repeat of the *C. elegans* telomeric template sequence 60 nucleotides from the 5′ end; (2) a 20-nucleotide region of base pair complementarity upstream and downstream of the telomeric template sequence; and (3) a canonical pseudoknot predicted within the region between the telomeric template and the 3′ end of the predicted 20 nucleotide helix. However, other characteristics of *tts-1* differ from telomerase RNA genes. Alternative splicing was detected between the two ESTs representing this locus (Kohara 1996), resulting in the possible presence of an intron (52 nucleotides) within the central region of the transcript. Although the 3′ end of the *tts-1* transcript can be folded in a manner consistent with the three terminal helices seen in the vertebrate model of telomerase RNA (and a degenerate H box sequence lies in the proper place), no matching ACA box can be found three nucleotides in from the 3′ end.

We would not expect high levels of telomerase activity in the dauer where the germ line remains undeveloped and hence predict no increased requirement for the telomerase template RNA. Accordingly, six tags were observed for the putative *C. elegans* telomerase reverse transcription protein DY3.4 (SWISS-PROT accession O45321) in the mixed stages and none in the dauer. This seems inconsistent with *tts-1* being a telomerase component, but it remains possible that *tts-1* has a function involving binding to the telomeres or telomere-associated proteins. It is possible that *tts-1* may have a chromosome-protective function, especially in the dauer stage.

A second transcribed telomerase-like sequence, *tts-2* (dauer 5; mixed 10), was identified through sequence similarity to *tts-1* (GenBank accession number Z48795 17488–18489). The *tts-2* transcript also contains degenerate telomeric repeat sequences but lacks any of the structural signatures of other telomerase RNA subunits. The two cDNA clones deriving from *tts-2* also showed alternative splicing with one of the transcripts excluding the telomeric repeat sequence.

The most abundant dauer and mixed-stage–specific transcripts are listed in Table 3. Except where noted, these tags were not detected in the other library. The most highly expressed dauer-specific transcript is F38E11.2/*hsp-12.6*, encoding a small heat shock protein of the α-crystallin family. Small heat shock proteins have been implicated in oxidative and mechanical stress responses in many organisms, but in *C. elegans hsp12.6* is not induced by biological stress through exposure to heat or chemical agents (Leroux et al. 1997). It has been detected at high levels in synchronized populations arrested at the L1 stage by starvation. Therefore, *hsp12.6* may be specifically induced during times of developmental arrest. Other heat shock genes such as *Hsp90* (dauer 578: mixed 172) and Hsp70 (dauer 166: mixed 172) are also expressed within the dauer stage, although not in a dauer-speific manner, consistent with previous studies (Dalley and Golomb 1992; Cherkasova et al. 2000). The two highly abundant, dauer-specific G-protein coupled receptors (Table 3) are candidates for chemoreceptors involved in triggering dauer exit. It is striking that 15 of the 20 most abundant dauer-specific proteins are novel, lacking similarities with known proteins indicative of putative functions.

Chromosomes may be packaged differently in the dauer stage. Tags for histone H1-like genes C18G1.5 and M163.3/*his-24* were prominent in the mixed-stage profile. Histone H1-like genes C30G7.1 and F22F1.1 showed increased dauer expression whilst C18G1.5 was reduced (Fig. 3A). The 40-fold

Table 1. SAGE Tags Corresponding to Five Mitochondrial Sense and Antisense Transcripts

Gene	Dauer sense	Dauer antisense	Mixed sense	Mixed antisense
NADH-ubiquinone oxido-reductase chain 4	346	245	314	157
NADH-ubiquinone oxido-reductase chain 5	365	540	356	640
Cytochrome C oxidase subunit I	899	328	398	211
Cytochrome C oxidase subunit II	1222	1099	626	742
Cytochrome C oxidase subunit III	1415	1712	456	552

enrichment of tags for the histone H1 variant C30G7.1 (only 33% and 31% identical to the histones encoded by M163.3 and C18G1.5, respectively) suggests that dauer chromatin is altered, possibly to reduce overall transcription rate or assume a structure less susceptible to damage. Two different M163.3/*his-24* transcripts were detected, with the longer transcript more highly represented in the mixed-stage profile than in the dauer (dauer 10; mixed 31). This is consistent with a previous observation that a longer 1.3-kb transcript for M163.3 is present within male germ cells (Sanicola et al. 1990), a tissue represented only in the mixed-stage library. The D2096.8 nucleosome assembly protein transcripts were also present in two different forms, one of which displayed a 10-fold increase in tag abundance in the dauer profile. The differential mRNA forms may represent splice variants or alternative transcriptional termination.

Other genes more highly represented in the dauer SAGE library include the C47D12.8 DNA repair endonuclease, further indicating possible increased levels of proteins involved in maintaining DNA integrity during diapause (Fig. 3B). Tags for the F56C9.1 protein phosphatase I (PP1C) γ protein, also more prominent in the dauer profile, has been postulated to negatively regulate of entry into mitosis (Doonan and Morris 1989; Ohkura et al. 1989). A high level of activated PP1C protein would be antagonistic to the role of the DAF-2 insulin receptor family protein, which is a negative regulator of genes involved in oxidative stress (Honda and Honda 1999) and dauer entry. The *spo-11* gene (Dernburg et al. 1998), the *gld-1* gene (Jan et al. 1999), and the gene for caveolin (Scheel et al. 1999), all involved in meiotic and germ-line function, were not observed within the dauer stage, where the germ line does not develop.

To explore the fecundity of the SAGE approach to detect novel expressed sequences, the 10 most abundant tag species that could not be correlated to the sense strand of a predicted transcript were examined. Three were antisense tag species

Table 2. Expression of Genes Implicated Previously in Lifespan and Dauer

Gene	Dauer tags	Mixed tags	Significance[a]
C15F1.b (*sod-1*) superoxide dismutase (Giglio et al. 1994)	60	59	0.96
F55H2.1 (*sod-4*) superoxide dismutase (Fujii et al. 1998)	26	5	0
C08A9.1 (*sod-3*) Manganese superoxide dismutase (Honda and Honda 1999)	18	0	0
ZK430.1 superoxide dismutase	0	0	NA
T24H7.1 prohibitin-like (Coates et al. 1997)	3	1	0.39
ZC395.2 (*clk-1*) (Lakowski and Hekimi 1996)	8	10	0.60
C12D8.10a/b (*akt-1*) protein kinase (Paradis and Ruvkun 1998)	28	6	0
B0334.8 (*daf-23,age-1*) phosphatidylinositol-3-OH kinase (Larsen et al. 1995)	6	3	0.33
C11E4.1 glutathione peroxidase (Vanfleteren 1993)	68	15	0
C11E4.2 glutathione peroxidase (Vanfleteren 1993)	17	2	0
R03G5.5 glutathione peroxidase (Vanfleteren 1993)	2	0	0.09
F26E4.12 glutathione peroxidase (Vanfleteren 1993)	2	1	0.57
Y54G11A.6 catalase (*ctl-1*) (Vanfleteren 1993; Taub et al. 1999)	15	7	0.10
Y54G11A.5a/b peroxisomal catalase (Vanfleteren 1993)	11	7	0.37
ZC395.6 (*gro-1*) (Hodgkin and Doniach 1997)	0	4	0.02
R11A8.4 (*sir-2*) Yeast regulatory protein SIR2-like (Tissenbaum and Guarente 2001)	2	3	0.63
H42K12.1 (*pdk-1*) Akt/PKB kinase like (Paradis et al. 1999)	0	0	NA
Y55D5A.b (*daf-2*) Insulin receptor-like (Kimura et al. 1997)	1	1	0.98
T28B8.2 (*ins-1*) Insulin-like (Pierce et al. 2001)	6	1	0.05
R13H8.1 (*daf-16*) HNF-3/forkhead family member (Lin et al. 1997)	NA[b]		
R10H10.2 (*spe-26*) (Varkey et al. 1995)	0	0	NA
C08H9.5 (*cold-1*) Protein tyrosine kinase (Rikke et al. 2000)	4	1	0.17
F29C4.1 (*daf-1*) Protein kinase (Georgi et al. 1990)	3	3	0.97
F25E2.5 (*daf-3*) Smad protein (Patterson et al. 1997)	3	5	0.45
C05D2.1a (*daf-4*) Receptor protein kinase (Estevez et al. 1993)	1	3	0.29
B0412.2 (*daf-7*) TGF-β (Ren et al. 1996)	2	4	0.39
B0240.3 (*daf-11*) transmembrane guanylyl cyclase (Birnby et al. 2000)	0	6	0
T07A9.6 (*daf-18*) PTEN homolog (Ogg and Ruvkun 1998)	1	8	0.01
F33H1.1 (*daf-19*) RFX-type transcription factor (Swoboda et al. 2000)	NA[b]		

[a]Significance indicates the probability of the null hypothesis that no difference exists between the expression in both dauer and mixed stage populations, determined by the G test (Sokal and Rohlf 1995).
[b]Indicates genes for which SAGE tags could not unambiguously identify the transcript profile.

Table 3. The Twenty Most Abundant Dauer and Mixed-Stage-Specific SAGE Tags that Have Been Correlated to a Transcript

Dauer

Gene	Function	Dauer tags
F38E11.2 (hsp-12.6)	Alpha-B-crystallin/hsp20	263
R02C2.3	G-protein coupled receptor	62
T26C11.2	Unknown function	53
W07A12.7	G-protein coupled receptor	46
F36D3.9*	Cysteine protease	46
T25D10.3	Unknown function	33
Y39H10A.i	Unknown function	33
T09F5.8	Unknown function	30
K01G12.2	Unknown function	26
W01A11.7	Unknown function	26
T21D12.11	Unknown function	26
T05B4.12	Unknown function	25
R02E12.6	Unknown function	23
C18A11.2	Unknown function	21
F39H2.1*	Putative DNA repair protein	21
F20A1.6	Unknown function	21
F17C11.4	Unknown function	21
F14D7.5	Unknown function	20
K09H9.5	Unknown function	20
T16A1.4	Unknown function	20

Mixed Stages

Gene	Function	Mixed tags
F59D8.b	Vitellogenin	51
R09B5.3	Unknown function	34
R09B5.9	Unknown function	32
C46A5.3/col-14	Collagen	31
Y2H9A.3	Collagen	31
C50B6.4	Collagen	31
F53F1.4	Cuticlin 2 precursor	28
F57H12.3	Unknown function	25
F08G5.4	Collagen	25
C45B2.1	Unknown function	24
T23G11.2	Acetyltransferase	23
F28D1.5	Unknown function	22
T23G11.3/gld-1	Germline differentiation	21
W05F2.3	Unknown function	20
W05F2.1	Unknown function	20
M195.1	Collagen	20
F09G8.6	Collagen	20
C33G8.2	Actin-cytoskeleton associated	19
W03C9.7/mex-1	Zn Finger/germline differentiaton	19
Y18D10A.11	Unknown function	18

*These transcripts possess a second tag which is present in the mixed-stage profile (F36D3.9 mixed 1, dauer 1; F39H2.1 mixed 4, dauer 0).

base variant of the highly expressed F15A2.1 collagen gene. One tag species correlated with the mitochondrial small ribosomal RNA gene (dauer 102; mixed 6; Okimoto et al. 1992), and the remaining species, CATGCGCTAAAAAA, is most likely the result of the CATG restriction site being too close to the polyadenylation site for an unambiguous gene assignment. Therefore, we can account for highly expressed unmapped tags as deriving from the transcriptome.

We investigated whether SAGE data also could be used to detect alternative splicing. Different (15,596) tag species could be correlated to the 10,581 transcripts unambiguously assigned to SAGE tags (excluding the 549 transcripts from 326 genes known to have alternatively spliced transcripts (Durbin and Thierry-Mieg 1991). Two thousand six hundred sixty-five transcripts had two tag species present, 721 had three, 194 had four, and 76 transcripts had five or more tag species present (if singleton tags are ignored these numbers become 1161, 159, 74, and 24, respectively). Many of these tag vari-
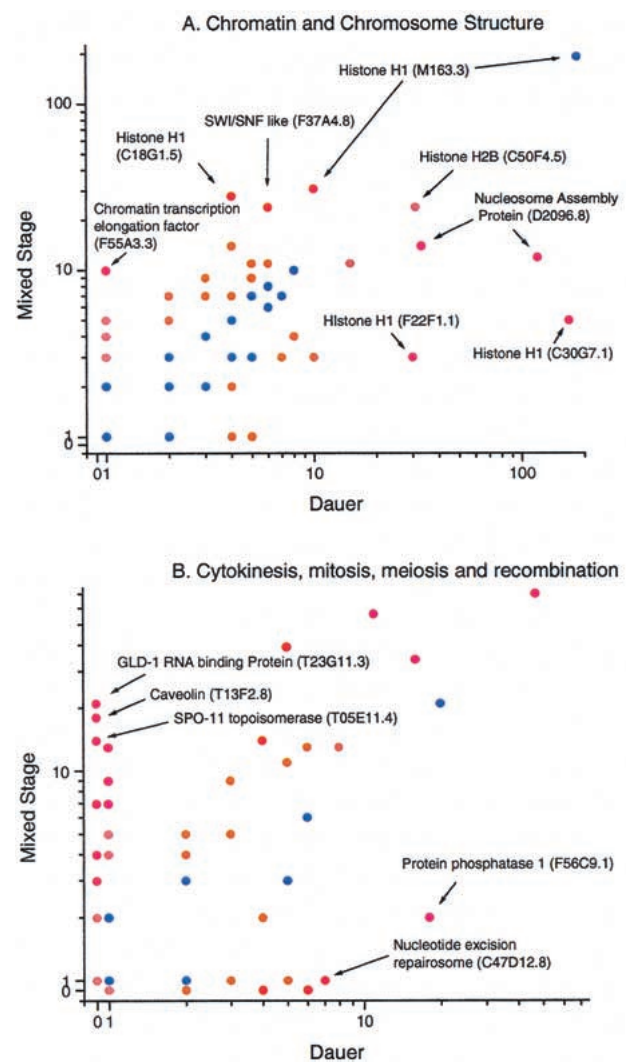

Figure 3 Relative abundance of dauer and mixed-stage tags for genes from specific cellular processes (Costanzo et al. 2000). (blue) No significant expression difference between the two sets. (orange) Significance of $P \leq 0.05$; (red) significance of $P \leq 0.01$, as determined by the G test (Sokal and Rohlf 1995).

from mitochondrial genes. Three other tag species were from genes represented in EST data sets but not currently in the annotated genomic data: the Ribosomal L27A protein (dauer 505; mixed 305), a nuclear cytochrome C oxidase (dauer 120; mixed 143), and a gene of unknown function (dauer 56; mixed 71). Two tag species were single-base variants of abundant tags, possibly representing sequence heterogeneity in the N2 populations. One species, CATGCGACTTCTGA (dauer 94; mixed 4), possessed a single mismatch with the *tts-1* tag, and its relative abundance in the two profiles was consistent with that observed for *tts-1*. Another correlated with a single-

ants may be resulting from incomplete *Nla*III digestion during SAGE library construction. However, 153 genes were candidates for possessing alternate transcript structures as these showed an altered ratio of different tag species between the mixed and dauer populations, as determined by the Fisher exact test ($P \leq 0.05$). In addition, nine of the genes known to have alternatively spliced transcripts had altered transcript representation between the dauer and mixed-stage profiles, for a total of 162 such genes.

Our findings represent the first large-scale investigation into gene expression during *C. elegans* diapause. The mixed-stage profile provides a reference point for future study of specific life stages or of mutant strains. Furthermore, our analysis has revealed some surprising and previously unsuspected differences leading to new hypotheses about the nature of the long-lived dauer form. The most abundant dauer-specific transcript shows strong similarity to the vertebrate α-B-crystalline/hsp20, which is predicted to function as a stabilizer of protein structure and cellular integrity (Leroux et al. 1997). At least three transcripts possibly affecting the structure or stability of chromatin (*tts-1*, a variant histone H1 gene and an isoform of a nucleosome assembly protein) are much more prominent in the dauer than in the nondauer expression profile, suggesting that chromatin structure may be altered to improve stability, to reduce overall gene expression, or possibly both. Finally, antisense RNA corresponding to mitochondrial gene transcripts is present in both populations and may play a role in preventing continuous translation of mitochondrial transcripts.

METHODS

Nematode Growth

Wild-type (N2) *C. elegans* were grown asynchronously in three 250-mL liquid cultures containing 4% *Escherichia coli* χ1666 in S medium, harvested by centrifugation, then allowed to settle for 30 min at 25°C to digest any *E. coli* present in the gut (Epstein Henry et al. 1995). The pooled, mixed-stage populations contained stages in the estimated ratio 20 L1 : 2 L2 : 1 L3 : 1 L4 : 1 adult. N2 dauer larvae were purified by the sucrose flotation method (Epstein Henry et al. 1995) from two starved liquid cultures grown at 25°C. The final preparation of dauer larvae did not contain more than one animal of any other stage per 500 dauer larvae. Settled worms were quick-frozen in small pellets by dropping into liquid nitrogen and stored at −80°C.

RNA Preparation

Frozen animals were crushed in liquid nitrogen using a mortar and pestle for 5–10 min. Total RNA was isolated by the guanidinium isothiocyanate : phenol method (Chomczynski and Sacchi 1987). The yields of total RNA from 5 mL of settled dauer larvae and from 3.5 mL of settled mixed-stage populations were ~12 and 20 mg, respectively.

Generation and Analysis of SAGE Tags

SAGE libraries were produced as previously described (Velculescu et al. 1995) (detailed SAGE protocol available at www.sagenet.org), and the resultant clones were used to generate 15,371 successful sequencing reads performed on ABI 377 automated DNA sequencers. Quality assessment and clipping of the reads was performed by using PHRED (Ewing et al. 1998) and vector_clip (Staden et al. 2000). Sage tags derived from the linker sequence used in SAGE library construction were removed. Conceptual transcript sequences were predicted from the *C. elegans* genome sequence, obtained from *C. el-*

egans ACEDB version WS9 and the mitochondrial genome (GenBank accession number X54252). Where possible, 5′ and 3′ untranslated region (UTR) sequences were derived from the coordinates of similarity to *C. elegans* EST sequences. In the absence of EST data, up to 270 bases were added to the 5′ end of genes and up to 460 bases added to the 3′ end of genes. No estimated UTR sequences extended into other gene predictions. Lengths added were sufficient to encompass the UTRs of 99% of the genes studied, as estimated from EST data. Conceptual SAGE tags were generated for all transcript sequences by determining the presence of the *Nla*III restriction site and the subsequent 10 3′ base pairs. Observed tags were correlated to conceptual transcript sequences. Where a tag matched two or more transcripts, the tags corresponding to the most 3′ CATG site were used to resolve the ambiguity, where possible. Estimated UTR sequences were not used to map tags to antisense transcripts. The SAGE data and the gene correlations are available from http://elegans.bcgsc.bc.ca/SAGE.

When comparing the abundance of mRNA species corresponding to SAGE tags, we used the overall expression profile reflected in each library. For example, the *tts-1* tag species in the dauer library was present 4329 times among the total of 152,314 tags (2.8%). In the mixed-stage library, this tag was present 215 times among a total of 148,324 (0.14%). Hence, when the two profiles are compared, this tag is 20 times more abundant in the dauer than in the mixed-stage profile. This is not to say that there is necessarily 20 times more of this RNA per cell in the dauer.

REFERENCES

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.

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.

Chen, J.L., Blasco, M.A., and Greider, C.W. 2000. Secondary structure of vertebrate telomerase RNA. *Cell* **100:** 503–514.

Cherkasova, V., Ayyadevara, S., Egilmez, N., and Reis, R.S. 2000. Diverse *Caenorhabditis elegans* genes that are upregulated in dauer larvae also show elevated transcript levels in long-lived, aged, or starved adults. *J. Mol. Biol.* **300:** 433–448.

Chomczynski, P. and Sacchi, N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162:** 156–159.

Coates, P.J., Jamieson, D.J., Smart, K., Prescott, A.R., and Hall, P.A. 1997. The prohibitin family of mitochondrial proteins regulate replicative lifespan. *Curr. Biol.* **7:** 607–610.

Costanzo, M.C., Hogan, J.D., Cusick, M.E., Davis, B.P., Fancher,

A.M., Hodges, P.E., Kondu, P., Lengieza, C., Lew-Smith, J.E., Lingner, C., et al. 2000. The yeast proteome database (YPD) and *Caenorhabditis elegans* proteome database (WormPD): Comprehensive resources for the organization and comparison of model organism protein information. *Nucleic Acids Res.* **28:** 73–76.

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.

Dernburg, A.F., McDonald, K., Moulder, G., Barstead, R., Dresser, M., and Villeneuve, A.M. 1998. Meiotic recombination in *C. elegans* initiates by a conserved mechanism and is dispensable for homologous chromosome synapsis. *Cell* **94:** 387–398.

Dihanich, M.E., Najarian, D., Clark, R., Gillman, E.C., Martin, N.C., and Hopper, A.K. 1987. Isolation and characterization of MOD5, a gene required for isopentenylation of cytoplasmic and mitochondrial tRNAs of *Saccharomyces cerevisiae* [published erratum appears in *Mol. Cell. Biol.* 1987;7:2035]. *Mol. Cell. Biol.* **7:** 177–184.

Doonan, J.H. and Morris, N.R. 1989. The bimG gene of *Aspergillus nidulans*, required for completion of anaphase, encodes a homolog of mammalian phosphoprotein phosphatase 1. *Cell* **57:** 987–996.

Durbin, R. and Thierry-Mieg, J. 1991. A. *C. elegans* Database. Documentation, code, and data available at anonymous FTP server, ftp.sanger.ac.uk.

Epstein Henry, F. and Shakes, D.C. American Society for Cell, Biology 1995. *Caenorhabditis elegans*: Modern biological analysis of an organism. Academic Press, San Diego, CA and London.

Estevez, M., Attisano, L., Wrana, J.L., Albert, P.S., Massague, J., and Riddle, D.L. 1993. The daf-4 gene encodes a bone morphogenetic protein receptor controlling *C. elegans* dauer larva development. *Nature* **365:** 644–649.

Ewing, B., Hillier, L., Wendl, M.C., and Green, P. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8:** 175–185.

Fujii, M., Ishii, N., Joguchi, A., Yasuda, K., and Ayusawa, D. 1998. A novel superoxide dismutase gene encoding membrane-bound and extracellular isoforms by alternative splicing in *Caenorhabditis elegans*. *DNA Res.* **5:** 25–30.

Gems, D., Sutton, A.J., Sundermeyer, M.L., Albert, P.S., King, K.V., Edgley, M.L., Larsen, P.L., and Riddle, D.L. 1998. Two pleiotropic classes of daf-2 mutation affect larval arrest, adult behavior, reproduction and longevity in *Caenorhabditis elegans*. *Genetics* **150:** 129–155.

Georgi, L.L., Albert, P.S., and Riddle, D.L. 1990. daf-1, a *C. elegans* gene controlling dauer larva development, encodes a novel receptor protein kinase. *Cell* **61:** 635–645.

Giglio, A.M., Hunter, T., Bannister, J.V., Bannister, W.H., and Hunter, G.J. 1994. The copper/zinc superoxide dismutase gene of *Caenorhabditis elegans*. *Biochem Mol. Biol. Int.* **33:** 41–44.

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.

Heikimi, S., Lemieux, J., Lakowski, B., and Barnes, T. 1997. The *C. elegans* clock gene gro-1. In Canadian Patent 2210251.

Hodgkin, J. and Doniach, T. 1997. Natural variation and copulatory plug formation in *Caenorhabditis elegans*. *Genetics* **146:** 149–-164.

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.

Jan, E., Motzny, C.K., Graves, L.E., and Goodwin, E.B. 1999. The STAR protein, GLD-1, is a translational regulator of sexual identity in *Caenorhabditis elegans*. *EMBO J.* **18:** 258–269.

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.

Kohara, Y. 1996. Large scale analysis of *C. elegans* cDNA. *Tanpakushitsu Kakusan Koso* **41:** 715–720.

Lakowski, B. and Hekimi, S. 1996. Determination of life-span in *Caenorhabditis elegans* by four clock genes. *Science* **272:** 1010–1013.

Larsen, P.L., Albert, P.S., and Riddle, D.L. 1995. Genes that regulate both development and longevity in *Caenorhabditis elegans*. *Genetics* **139:** 1567–1583.

Leroux, M.R., Ma, B.J., Batelier, G., Melki, R., and Candido, E.P. 1997. Unique structural features of a novel class of small heat shock proteins. *J. Biol. Chem.* **272:** 12847–12853.

Lin, K., Dorman, J.B., Rodan, A., and Kenyon, C. 1997. daf-16: An HNF-3/forkhead family member that can function to double the life-span of *Caenorhabditis elegans*. *Science* **278:** 1319–1322.

Ogg, S. and Ruvkun, G. 1998. The *C. elegans* PTEN homolog, DAF-18, acts in the insulin receptor-like metabolic signaling pathway. *Mol. Cell* **2:** 887–893.

Ohkura, H., Kinoshita, N., Miyatani, S., Toda, T., and Yanagida, M. 1989. The fission yeast dis2+ gene required for chromosome disjoining encodes one of two putative type 1 protein phosphatases. *Cell* **57:** 997–1007.

Okimoto, R., Macfarlane, J.L., Clary, D.O., and Wolstenholme, D.R. 1992. The mitochondrial genomes of two nematodes, *Caenorhabditis elegans* and *Ascaris suum*. *Genetics* **130:** 471–498.

Paradis, S. and Ruvkun, G. 1998. *Caenorhabditis elegans* Akt/PKB transduces insulin receptor-like signals from AGE-1 PI3 kinase to the DAF-16 transcription factor. *Genes & Dev.* **12:** 2488–2498.

Paradis, S., Ailion, M., Toker, A., Thomas, J.H., and Ruvkun, G. 1999. A PDK1 homolog is necessary and sufficient to transduce AGE-1 PI3 kinase signals that regulate diapause in *Caenorhabditis elegans*. *Genes & Dev.* **13:** 1438–1452.

Patterson, G.I., Koweek, A., Wong, A., Liu, Y., and Ruvkun, G. 1997. The DAF-3 Smad protein antagonizes TGF-β-related receptor signaling in the *Caenorhabditis elegans* dauer pathway. *Genes & Dev.* **11:** 2679–2690.

Pierce, S.B., Costa, M., Wisotzkey, R., Devadhar, S., Homburger, S.A., Buchman, A.R., Ferguson, K.C., Heller, J., Platt, D.M., Pasquinelli, A.A.,et al. 2001. Regulation of DAF-2 receptor signaling by human insulin and ins-1, a member of the unusually large and diverse *C. elegans* insulin gene family. *Genes & Dev.* **15:** 672–686.

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-β homolog. *Science* **274:** 1389–1391.

Riddle, D.L. and Albert, P.S. 1997. C. elegans *II*. In C. elegans *II*. pp. 739–768. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Rikke, B.A., Murakami, S., and Johnson, T.E. 2000. Paralogy and orthology of tyrosine kinases that can extend the life span of *Caenorhabditis elegans*. *Mol. Biol. Evol.* **17:** 671–683.

Sanicola, M., Ward, S., Childs, G., and Emmons, S.W. 1990. Identification of a *Caenorhabditis elegans* histone H1 gene family. Characterization of a family member containing an intron and encoding a poly(A)+ mRNA. *J. Mol. Biol.* **212:** 259–268.

Scheel, J., Srinivasan, J., Honnert, U., Henske, A., and Kurzchalia, T.V. 1999. Involvement of caveolin-1 in meiotic cell-cycle progression in *Caenorhabditis elegans*. *Nat. Cell Biol.* **1:** 127–129.

Shirafuji, N., Takahashi, S., Matsuda, S., and Asano, S. 1997. Mitochondrial antisense RNA for cytochrome C oxidase (MARCO) can induce morphologic changes and cell death in human hematopoietic cell lines. *Blood* **90:** 4567–4577.

Sokal, R.R. and Rohlf, F.J. 1995. Biometry: The principles and practice of statistics in biological research. Freeman, New York.

Staden, R., Beal, K.F., and Bonfield, J.K. 2000. The Staden package, 1998. *Methods Mol. Biol.* **132:** 115–130.

Stambuk, R.A. and Moon, R.T. 1992. Purification and characterization of recombinant Xenopus poly(A)(+)-binding protein expressed in a baculovirus system. *Biochem. J.* **287:** 761–766.

Swoboda, P., Adler, H.T., and Thomas, J.H. 2000. The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in *C. elegans*. *Mol. Cell* **5:** 411–421.

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. and Guarente, L. 2001. Increased dosage of a sir-2 gene extends lifespan in *Caenorhabditis elegans*. *Nature* **410:** 227–230.

Tullo, A., Tanzariello, F., D'Erchia, A.M., Nardelli, M., Papeo, P.A., Sbisa, E., and Saccone, C. 1994. Transcription of rat mitochondrial NADH-dehydrogenase subunits. Presence of antisense and precursor RNA species. *FEBS Lett.* **354:** 30–36.

Vanfleteren, J.R. 1993. Oxidative stress and ageing in *Caenorhabditis elegans*. *Biochem. J.* **292:** 605–608.

Varkey, J.P., Muhlrad, P.J., Minniti, A.N., Do, B., and Ward, S. 1995. The *Caenorhabditis elegans* spe-26 gene is necessary to form spermatids and encodes a protein similar to the actin-associated proteins kelch and scruin. *Genes & Dev.* **9:** 1074–1086.

Velculescu, V.E., Zhang, L., Vogelstein, B., and Kinzler, K.W. 1995. Serial analysis of gene expression. *Science* **270:** 484–487.