Mol Gen Genet (1995) 247:367–378

ORIGINAL PAPER

Denise V. Clark · Dinar S. Suleman
Karen A. Beckenbach · Erin J. Gilchrist
David L. Baillie

Molecular cloning and characterization of the *dpy-20* gene of *Caenorhabditis elegans*

Received: 17 January 1994 / Accepted: 1 December 1994

Abstract We describe the molecular analysis of the *dpy-20* gene in *Caenorhabditis elegans*. Isolation of genomic sequences was facilitated by the availability of a mutation that resulted from insertion of a Tc1 transposable element into the *dpy-20* gene. The Tc1 insertion site in the *m474::Tc1* allele was identified and was found to lie within the coding region of *dpy-20*. Three revertants (two wild-type and one partial revertant) resulted from the excision of this Tc1 element. Genomic *dpy-20* clones were isolated from a library of wild-type DNA and were found to lie just to the left of the *unc-22* locus on the physical map, compatible with the position of *dpy-20* on the genetic map. Cosmid DNA containing the *dpy-20* gene was successfully used to rescue the mutant phenotype of animals homozygous for another *dpy-20* allele, *e1282*ts. Sequence analysis of the putative *dpy-20* homologue in *Caenorhabditis briggsae* was performed to confirm identification of the coding regions of the *C. elegans* gene and to identify conserved regulatory regions. Sequence analysis of *dpy-20* revealed that it was not similar to other genes encoding known cuticle components such as collagen or cuticulin. The *dpy-20* gene product, therefore, identifies a previously unknown type of protein that may be directly or indirectly involved in cuticle function. Northern blot analysis showed that *dpy-20* is expressed predominantly in the second larval stage and that the mRNA is not at all abundant. Data from temperature shift studies using the temperature-sensitive allele *e1282*ts showed that the sensitive period also occurs at approximately the second larval stage. Therefore, expression of *dpy-20* mRNA and function of the DPY-20 protein are closely linked temporally.

Key words *Caenorhabditis elegans* · *dpy-20* · Cuticle *Caenorhabditis briggsae*

Communicated by D. J. Finnegan

D. V. Clark[1] · D. S. Suleman · K. A. Beckenbach
E. J. Gilchrist · D. L. Baillie (✉)
Institute of Molecular Biology and Biochemistry,
Department of Biological Sciences, Simon Fraser University,
Burnaby, B.C., Canada V5A 1S6

Present address:
[1] Department of Biology,
University of New Brunswick (Bag Service 45111),
Fredericton, N.B., Canada E3B 6E1

Introduction

In the nematode *Caenorhabditis elegans* there are many genes that have mutant phenotypes affecting body morphology. Mutations in the *dpy-20* (*dumpy*) gene on chromosome IV cause a short body phenotype. There are at least 29 other *dpy* genes known which, when mutated, cause homozygous individuals to exhibit a shorter body length than wild-type animals. Many of these genes are involved in the formation or maintenance of the cuticle. The cuticle is an extracellular structure that protects the animal from environmental stresses, and, as in most nematodes, acts as an exoskeleton to control body shape in *C. elegans*. During *C. elegans* development from embryo to adult, there are four larval stages (L1, L2, L3 and L4) separated by four molts. The old cuticle is completely shed at each molt and a new, biochemically distinct one is synthesized (Cox et al. 1981b). The body shape of the nematode is first established during embryogenesis, primarily as a result of the contraction of circumferentially oriented microfilaments in the hypodermal cells that form the outer layer of the embryo (Priess and Hirsh 1986). Following elongation, and prior to hatching, the hypodermal cells secrete the L1 cuticle. The different cuticles of the various stages are then responsible for maintaining the elongated shape of the nematode throughout its life cycle.

The cuticle consists of several layers of cross-linked collagenous and non-collagenous proteins (Cox et al. 1981a). The amounts and types of collagens and other proteins vary with each developmental stage (Cox et al. 1981b, c; Cox and Hirsh 1985). In addition, the ultrastructure of the cuticle varies between some larval stages, or the dauer stage (an alternate to the L3 stage)

and the adult stage (Cox et al. 1981c). The stage-specific pattern of collagens and other gene products suggests that cuticle formation occurs through the progressive activation of different sets of genes during development (Cox and Hirsh 1985). Several genes have been identified that are known to be involved in the formation or function of the cuticle in *C. elegans*. Some genetically identified loci, such as *sqt-1* (*squat*), *sqt-3*, and *rol-6* (*roller*), encode collagen genes, as do the *dpy* genes *dpy-2, dpy-7, dpy-10,* and *dpy-13* (for review see Kramer 1994). Others, such as *bli-4* (*blister*), encode enzymes that are thought to be responsible for processing structural cuticle components (K. Peters and A. Rose, personal communication). There are several other genetically defined genes that affect body shape and may be involved in cuticle function whose roles are not yet understood at the molecular level. Animals homozygous for mutations in the *dpy-5* gene, for example, have a short body phenotype and cuticle defects (Ouazana et al. 1985), yet this gene does not encode a collagen (Babity 1993). In fact, the protein product of this locus does not resemble any known proteins, indicating that there are probably still other classes of cuticle components that have not yet been identified. Several genes that encode components of the cuticle have been cloned but not identified mutationally. Most of these are collagen genes (for review see Kramer 1994), but one non-collagenous *cut*icul*in* gene has also been identified (*cut-1*; Sebastiano et al. 1991). The functions of many other genes with morphological mutant phenotypes may also be cuticle related.

The investigation of the *dpy-20* gene at the molecular level was undertaken partly to determine if it encoded a protein that was involved in the formation or maintenance of the cuticle. Evidence that *dpy-20* coded for a cuticle component came from observations that Dpy-20 mutants exhibited aberrant cuticle structure on examination by electron microscopy (M. Kusch, unpublished results, cited in Cox et al. 1985). The *col*lagen gene, *col-5*, had been placed close to *dpy-20* on the genetic map (Cox et al. 1985) yet had not been identified mutationally. Because collagens are important components of the cuticle, it was possible that the gene identified mutationally as *dpy-20* was in fact the gene that had been molecularly identified as *col-5*. The determination of the sequence of *dpy-20* was undertaken, in part, to resolve the question of whether these two genes represented a single locus.

Another reason that the molecular nature of *dpy-20* was of interest was the fact that animals homozygous for the various mutant alleles exhibited a wide range of phenotypes, from weak dumpy (e.g. *e1282*ts) to severely dumpy (e.g. *m474::Tc1*). The phenotype of null, or loss-of-function, mutations had not been determined but the fact that most alleles were almost inviable at 15°C suggested that the null phenotype might be lethal. The identification of the DPY-20 protein product was likely to be of general developmental interest if the gene was essential. We reasoned that if the *dpy-20* gene product displayed similarity to a known polypeptide, this might allow us to determine if it was likely to be essential, or how it was likely to function during development.

The third reason for cloning *dpy-20* was that it lay in a region of chromosome IV that had been extensively studied both genetically (Rogalski et al. 1982; Rogalski and Baillie 1985; Clark et al. 1988; Clark and Baillie 1992) and molecularly (Baillie et al. 1985, Prasad and Baillie 1988) with respect to mutations in essential genes and the detection of putative open reading frames. The cloning of *dpy-20* would provide another link between the physical map (Coulson et al. 1988) and the genetic map in the *unc-22* region of the chromosome, which would benefit our overall understanding of genome organization and function.

In this paper we describe the cloning of the *dpy-20* gene in *C. elegans* using the techniques of transposon tagging and transformation rescue. We present sequence analysis of a *dpy-20* cDNA clone, *dpy-20* genomic clones, and a clone that represents part of the *Caenorhabditis briggsae dpy-20*-like gene. We also investigate *dpy-20* expression using Northern analysis and the genetic examination of a temperature-sensitive allele *dpy-20(e1282)*ts.

Materials and methods

Nematode strains and culture conditions

The nomenclature in this paper follows the recommendations of Horvitz et al. (1979). Nematodes were maintained on petri plates containing nematode growth medium streaked with *Escherichia coli* strain OP50 (Brenner 1974). Unless otherwise indicated, nematodes were grown at 20°C.

Two strains were kindly provided by P. Albert and D. Riddle (University of Missouri, Columbia, Mo.). One of these strains, DR1040, was severely Dpy in phenotype and had the genotype (*mut*ator) *mut-6(st702) dpy-20(m474::Tc1)*. The other strain, DR1058, was wild type in phenotype and had the genotype *mut-6 dpy-20(m474m495)*wt. Two additional DR1040 revertant strains were used in this study: BC3212, which is wild type in phenotype; and BC3226, which is mildly Dpy and, therefore, only partially revertant.

The following mutations were used in this study: *dpy-20(e1282)*ts, *dpy-20(e1362), dpy-20(e1415), unc-5(e152)* and *unc-26(e345)*. These were obtained from the MRC Laboratory, Cambridge, U.K. Strains carrying *dpy-20(cn142), dpy-20(cn322), dpy-20(cn326),* and *dpy-20(e2017)*am were obtained from the Caenorhabditis Genetics Center at the University of Missouri, Columbia, Mo., USA. *unc-22(s7)* was isolated at Simon Fraser University (Moerman and Baillie 1979).

Strain construction

The mutator locus *mut-6* lies one map unit to the right of *dpy-13* on chromosome IV and is associated with mobility of the transposable element Tc1 (Mori et al. 1988). The strains DR1040, DR1058, BC3212, and BC3226 contained *mut-6(st702)* and a number of Tc1 elements not found in the Bristol (N2) strain. The *mut-6* locus was removed from the four strains to reduce the probability of mutational events. Hermaphrodites of the genotype *mut-6(st702)+dpy-20(x)+/+unc-5(e152)+unc-22(s7)* were constructed and allowed to self-fertilize. Their progeny were screened for Unc-5 recombinants that twitched in nicotine (*+unc-5dpy-20(x)+/+unc-5+unc-22*). The Unc-5 recombinant hermaphrodites were used to establish new strains that were homozygous for

Table 1 Derivatives of DR1040 [*mut-6(st702) dpy-20(m474::Tc1)IV*]

Strain	Subsequent derivatives	Genetic manipulation	Genotype	Phenotype
DR1040			*mut-6 dpy-20(m474::Tc1)*	Severe Dpy-20
	DR1058	Revertant	*mut-6 dpy-20(m474m495)*wt	Wild type
	BC3212	Revertant	*mut-6 dpy-20(m474s1711)*wt	Wild type
	BC3226	Revertant	*mut-6 dpy-20(m474s2033)*	Regular Dpy-20
	BC3325	Revertant	*mut-6 dpy-20(m474s1774)*wt	Wild type
	BC3326	Revertant	*mut-6 dpy-20(m474s1776)*wt	Wild type
	BC3321	Remove *mut-6*	*unc-5 dpy-20(m474::Tc1)*	Dpy-20 Unc-5
	BC3103	Remove *unc-5*	*dpy-20(m474::Tc1)*	Severe Dpy-20
	BC3484	Outcross 6×	*dpy-20(m474::Tc1)*	Severe Dpy-20
DR1058			*mut-6 dpy-20(m474m495)*wt	Wild type
	BC3323	Remove *mut-6*	*unc-5 dpy-20(m474m495)*wt	Unc-5
	BC3302	Remove *unc-5*	*dpy-20(m474m495)*wt	Wild type
	BC3486	Outcross 6×	*dpy-20(m474m495)*wt	Wild type
BC3212			*mut-6 dpy-20(m474s1711)*wt	Wild type
	BC3324	Remove *mut-6*	*unc-5 dpy-20(m474s1711)*wt	Unc-5
	BC3304	Remove *unc-5*	*dpy-20(m474s1711)*wt	Wild type
	BC3487	Outcross 5×	*dpy-20(m474s1711)*wt	Wild type
BC3226			*mut-6 dpy-20(m474s2033)*	Regular Dpy-20
	BC3322	Remove *mut-6*	*unc-5 dpy-20(m474s2033)*	Dpy-20 Unc-5
	BC3110	Remove *unc-5*	*dpy-20(m474s2033)*	Regular Dpy-20
	BC3485	Outcross 6×	*dpy-20(m474s2033)*	Regular Dpy-20

the recombinant chromosome but had presumably lost any closely linked Tc1 polymorphisms to the left of *dpy-20*. Then closely linked Tc1 polymorphisms to the right of *dpy-20* were removed by selecting Unc-26 recombinants from the self-progeny of hermaphrodites of genotype *unc-5(e152) dpy-20(x)++/++unc-22(s7)unc-26(e345)*. The flanking markers, *unc-5* and *unc-26*, were then crossed off of the *dpy-20* chromosome so that the resultant strains were simply *dpy-20(x)* in genotype. Hermaphrodites of each of these strains were then mated to males and Dpy-20 progeny selected in the next generation. This procedure was repeated six times. This ensured that approximately 98% of the original, mutator strain genome was replaced with N2 DNA except in the region immediately surrounding *dpy-20* (Baillie et al. 1985). The strains that were derived from BC3212 and DR1058 (see Table 1) were crossed to *dpy-20(e1282)*ts animals with each round of outcrossing to keep track of the revertant allele. Strains that were derived from DR1040 and BC3226 (see Table 1) were crossed to animals carrying a wild-type chromosome IV with each round of outcrossing. The strains retained from these constructions are listed in Table 1.

*dpy-20(e1282)*ts temperature shift experiment

Hermaphrodites homozygous for *dpy-20(e1282)*ts were allowed to lay eggs for 1.5 h at either 15°C or 25°C. Five parents per plate yielded about 20 eggs in this time period, or more at the higher temperature. The time at which the parents were removed was taken as the 0 h time point. Individual plates were shifted up or down between 15°C and 25°C at 4 h intervals for the first 32 h, and then every 8 h subsequently, until the worms had reached the adult stage. The times of hatching and of the L4 molt were noted to normalize the growth rates at the two temperatures. The effect of the temperature shifts on the phenotype of the adult hermaphrodites was assayed by measuring the body length of the worms during the time after they began to lay eggs but before these eggs hatched (a period of 10 h at 25°C). The body length of each worm was measured after it was heat-killed by placing a soldering iron tip in the agar close to the worm. Measurements were made with an ocular micrometer at 50× magnification.

The eggs for the temperature "pulse" experiment were synchronized in the same manner as for the temperature shift experiment. Dpy-20 individuals raised at 15°C were shifted to 25°C for 8 h pulses at 0, 24, 48, 72, and 96 h after the end of egg laying. Individuals raised at 25°C were shifted to 15°C for 24 h pulses at 0, 8, 16, 24 and 32 h after egg laying. The "pulses" are approximately equivalent in length developmentally if the growth rates at 15°C and 25°C are normalized. The adult body length was measured as described above.

Nucleic acid techniques

Plasmid DNA was prepared using the alkaline lysis procedure outlined by Maniatis et al. (1982) or the Miniprep Kit Plus (Pharmacia). Plasmid DNA was purified for manual sequencing using Sephacryl S-400 Spun columns (Pharmacia). No further purification of DNA was required for automated sequencing. Genomic DNA was extracted from nematodes using the protocol described by Emmons et al. (1979). Restriction enzymes were obtained from Bethesda Research Laboratories (BRL) or from Pharmacia, and were used as recommended by the manufacturer. DNA was transferred to nitrocellulose (Schleicher and Schuell) using the bi-directional transfer method described by Smith and Summers (1980).

DNA fragments used for making probes were extracted from agarose gels either by electroelution in dialysis tubing as described by Maniatis et al. (1982), by using an Elutrap apparatus (Schleicher and Schuell), or by centrifugation through siliconized glass wool in a microcentrifuge at 6000 rpm for 10 min (Heery et al. 1990). The vectors pBluescript (Stratagene) and pVZ1 (Henikoff and Eghtedarzadeh 1987) were used for subcloning. Ligations were carried out as described in Snutch (1984). Probes were labeled with [^{32}P]dATP (Amersham) using the oligolabelling technique of Feinberg and Vogelstein (1983). Random hexamer primers were obtained from Pharmacia. The probe used for screening the BC3484 genomic DNA library (described below) was labeled with [^{32}P]dATP using the nick-translation technique of Rigby et al. (1977). Prior to hybridization with probes, DNA filters were soaked in hybridization solution (5 X SSPE, 0.3% SDS and 2.5 X Denhardt's solution; 1 X Denhardt's is 0.02% BSA, 0.02% Ficoll, 0.02% polyvinyl pyrrolidone) at 68°C for at least 1 h. Fresh hybridization solution containing denatured probe was added, and hybridization was carried out for at least 16 h. Filters were then exposed to either Kodak X-OMAT K or X-OMAT AR preflashed film at –80°C.

Northern analysis

For Northern analysis, total RNA from synchronized populations of nematodes at one of the four stages of development (L1, L2, L3/L4, or adult) was isolated by T. Snutch (Chirgwin et al. 1979; Snutch 1984). Total and poly(A)$^+$ RNA from a mixed population of worms was isolated by Shiv Prasad using the same technique (personal communication). Aliquots of approximately 10 μg of total RNA from mixed stages, 10 μg of total RNA from specific developmental stages and 1 μg of poly(A)$^+$ RNA were subjected to denaturation, electrophoresis, and transfer to Genescreen (Dupont) membranes as described by Prasad and Baillie (1989). After electrophoresis, the presence of intact RNA was verified by ethidium bromide staining of the gel (not shown). Hybridization was carried out as described above at 68°C. The probe was washed from the filter at 68°C with 1 X SSPE and 0.2% SDS.

Construction and screening of a partial genomic library of mutant DNA

Approximately 20 μg of genomic DNA from the mutant strain BC3484 was digested to completion with *Eco*RI and run on a 0.75% low melting point agarose gel. Fragments between 6 and 10 kb were isolated. *Eco*RI-digested phage λZAP DNA was then treated with alkaline phosphatase and about 1 μg of λZAP DNA and 250 ng of genomic DNA were pooled and ligated, and the phage packaged as directed in the Stratagene protocol. The BC3484 library was screened with the 0.9 kb *Eco*RI-*Hin*dIII fragment of Tc1 (pCes100; Baillie et al. 1985). After three rounds of screening, pBluescript plasmids were isolated from the phage using the Stratagene protocol.

Screening of *C. elegans* cDNA and genomic libraries

Approximately 1×10^4 phage from a *C. elegans* cDNA library in λZAP (Barstead and Waterston 1989) were screened with the 1.4 kb *Eco*RI-*Eco*RV fragment from clone pCes105. This fragment consists of DNA that flanks the Tc1 element of *dpy-20(m474::Tc1)* (see Results). After three rounds of screening, plasmids were isolated from the phage as directed by Stratagene. Subsequently, 6×10^5 phage of the Barstead cDNA library, and 5×10^4 of an L2 stage-specific cDNA library in λgt10 (J. Arhinger and J. Kimble, personal communication) were screened for additional cDNAs.

A *C. elegans* wild-type genomic DNA library (T. Snutch 1984) was also screened with the 1.4 kb fragment of pCes105. After three rounds of screening, the isolated phage clones were collected from plate lysates, and isolated by precipitation in 1 M NaCl, 10% polyethylene glycol and ultracentrifugation in 0.75 g/ml CsCl at 45|000 rpm overnight (Maniatis et al. 1982). DNA was extracted from the phage using the formamide method of Davis et al. (1980).

Screening of a *C. briggsae* genomic DNA library

Approximately 5×10^4 phage of a Charon 4 library of wild-type *C. briggsae* genomic DNA, constructed by T. Snutch (1984), were screened with the *C. elegans dpy-20* cDNA and the 6.0 kb *Xba*I subclone of the cosmid C35H3 (pCes177; M. Han and P. Sternberg, personal communication) that was used to rescue *dpy-20(1282)*ts. After three rounds of screening (65°C hybridization), phage clones were collected, purified in CsCl, and extracted using the formamide method of Davis et al. (1980).

DNA sequence analysis

Manual sequencing was performed using the Sequenase Version 2.0 DNA sequencing kit (United States Biochemical), with a DMSO (dimethyl sulfoxide) modification developed by T. Snutch (personal communication). Automated sequencing was performed using the Applied Biosystems (ABI) dye primer and dideoxy dye terminator sequencing methods devised for the Model 373A sequencing apparatus.

Genomic sequence of the *dpy-20* gene was derived from the 4.9 kb *Eco*RI and 2.7 kb *Eco*RI fragments to which the cDNA hybridized (see Results) and the 6.0 kb *Xba*I fragment of cosmid C35H3 (pCes177). Deletion subclones of pCes177 were constructed using the Erase-a-Base system (Promega). Custom oligonucleotide primers and commercially available T3 and T7, together with M13 reverse and –21M13 (forward), dye primers were used for sequencing. Oligonucleotide primers 18–20 bp in length were designed using the program OLIGO (Rychlik and Rhoads 1989) and synthesized using the ABI Model 391 oligonucleotide synthesizer.

Nucleotide sequences were aligned, formatted, and translated using the Eyeball Sequence Editor (ESEE) program (Cabot and Beckenbach 1989). Amino acid sequence predicted using ESEE was used to search the latest available releases of the SwissProt, PIR, GenPept, and GPUpdate databases using the BLAST algorithm (Altschul et al. 1990). Hydropathy plots of the protein sequence were generated using TGREASE (Pearson and Lipman 1988). The inferred amino acid sequence was further analyzed using the PCGENE suite of programs (Bairoch 1990) specifically PROSITE, PSIGNAL, PESTFIND, AACOMP, and CHARGPRO.

Germ-line transformation

Cosmid DNA used for germ-line transformation was prepared using the protocol supplied by J. Sulston (personal communication). DNA from overnight liquid cultures was extracted using the alkaline lysis technique described by Maniatis et al. (1982), except that the phenol and chloroform extractions were omitted and the TE buffer did not contain ribonuclease A. To 50 μl of DNA solution, 50 μl of 4.4 M LiCl was added, mixed, and left at 4°C for 4 h. After a 5-min spin in a microcentrifuge, 90 μl of supernatant was kept. The DNA was then precipitated with 95% ethanol, washed with 70% ethanol, and dispersed slowly in 100 μl of 0.1 M potassium acetate (pH 7.4). The DNA was then precipitated with 200 μl 95% ethanol at –20°C, was washed with 70% ethanol, vacuum dried and resuspended in 10 mM KPO_4 (pH 7.5). The appropriate amount of DNA was then mixed with injection solution (see Fire 1986) to a final concentration of 200–400 ng/μl. The injection technique was essentially as described by Fire (1986) except that the recovery temperature was 20°C.

Results

Dpy-20 mutant phenotypes

The strain DR1040 has a strong dumpy phenotype (Fig. 1D) and has the genotype *mut-6(st702)dpy-20(m474::Tc1)*. The *mut-6* mutator locus (Mori et al. 1988) was active in this strain and resulted in Tc1 transposition. Screening for revertants resulted in the isolation of one individual with a less severe dumpy phenotype (strain BC3226; Fig. 1E) and four wild-type revertant hermaphrodites (strains BC3212, BC3325, BC3326 and DR1058).

Animals homozygous for various *dpy-20* alleles were examined using Nomarski microscopy in order to understand the effects of disrupting the normal function of *dpy-20*. The *e1282*ts allele is temperature sensitive and *dpy-20(e1282)*ts individuals were raised at 15°C, 20°C and 25°C and examined for phenotypic differences (Fig. 1A, B and C, respectively). The alleles *cn142*,

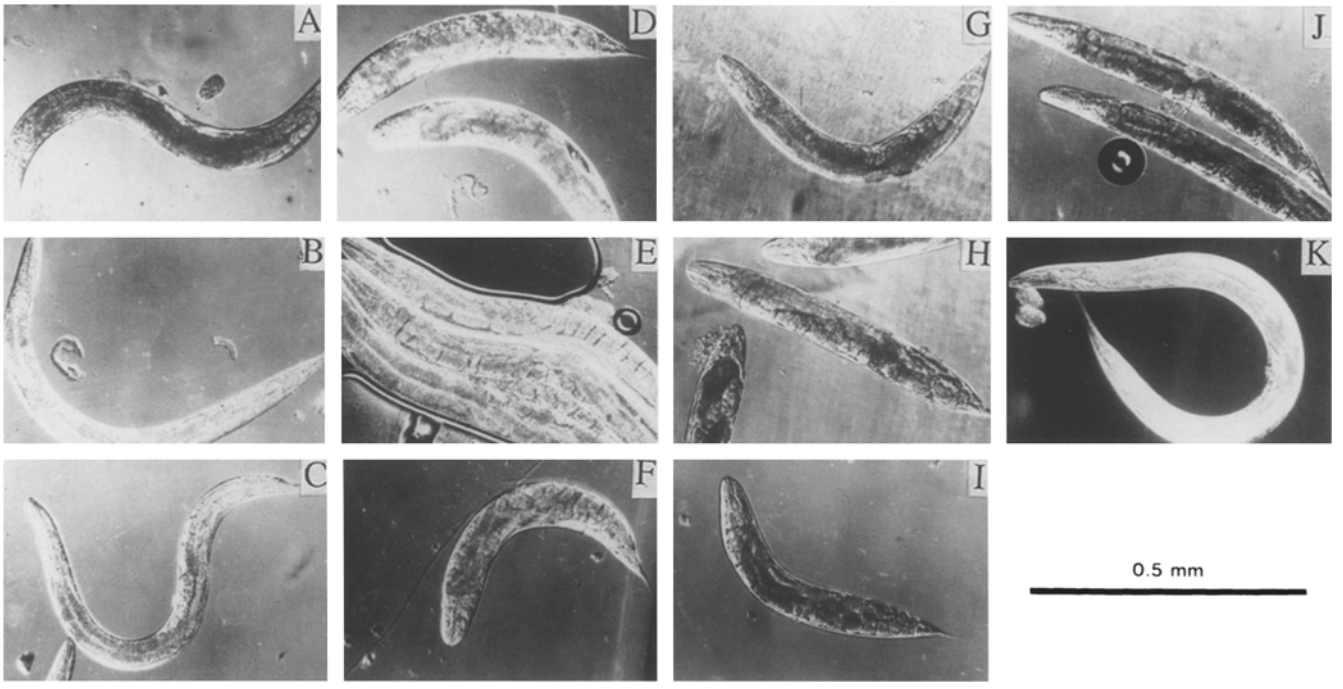


Fig. 1A–K Photographs of Dpy-20 mutants taken with Nomarski optics. Genotypes are as follows: **A** *dpy-20(e1282)*ts 25°C. **B** *dpy-20(e1282)*ts 20°C. **C** *dpy-20(e1282)*ts 15°C. **D** *dpy-20(m474::Tc1)* 20°C. **E** *dpy-20(m474s2033)* 20°C. **F** *dpy-20(cn142)* 20°C. **G** *dpy-20(cn322)* 20°C. **H** *dpy-20(e1362)* 20°C. **I** *dpy-20(e192)* 20°C. **J** *dpy-20(e2017)*am 20°C. **K** Wild-type control (N2) 20°C

cn322 (Fig. 1F, G) and *cn326* (not shown) were isolated in a general screen for dumpy mutants by Hosono et al. (1982). The *cn142* and *m474::Tc1* (Fig. 1D) alleles exhibited a "rounded head" phenotype, as well as the shorter body length typical of other *dpy-20* alleles. The phenotype of animals homozygous for the amber mutation *e2017*am (Fig. 1J) was less severely Dpy than that of animals carrying the *m474::Tc1* or *cn142* alleles. The remaining three alleles, *e1362, e192* (Fig. 1H and I, respectively) and *e1422* (not shown), all conferred strong Dpy phenotypes upon homozygotes. The allele *cn142* exhibited cold-sensitive lethality at 13°C. The parents of homozygous *cn142* animals laid a few eggs that hatched but the progeny did not develop beyond the L1 or L2 stage. The *e1362* allele also exhibited cold-sensitive lethality at 13°C, but this aspect of the phenotype was not fully penetrant in animals homozygous for this allele.

*dpy-20(e1282)*ts temperature shift experiments

Figure 2 summarizes the data gathered from the temperature shift (Fig. 2A) and temperature pulse (Fig. 2B) experiments. The plots for the down-shift and up-shift experiments in Fig. 2A indicate that the temperature-sensitive period was at 14–24 h after egg laying at 25°C, i.e., during the L1 and L2 stages (Wood et al. 1980). The body length of *dpy-20(e1282)*ts homozygotes grown at 15°C was 1.05±0.01 mm and at 25°C was 0.92 ±0.02 mm (95% confidence intervals). A temperature pulse experiment was conducted (Fig. 2B) to confirm that the *dpy-20(e1282)*ts mutants were sensitive to the restrictive temperature at the L1–L2 stage and that the phenotypes seen in the temperature shift experiment were not just a function of the length of time spent at a given temperature. The data indicated a temperature-sensitive period at approximately 16 h (at 25°C), which was consistent with the data from the temperature shift experiment.

A Tc1-containing fragment correlated with the Dpy-20 mutant phenotype

The BC3484 strain was homozygous for the *m474::Tc1* allele which carries a Tc1 insertion in the *dpy-20* gene. BC3485 animals were homozygous for the revertant allele *m474s2033*, that resulted in a less severe Dpy phenotype than *m474::Tc1*. BC3486 and BC3487 individuals exhibited a wild-type, revertant phenotype. It was reasonable to assume that these revertant strains had lost the Tc1 insertion that was the cause of the Dpy phenotype of BC3484 animals. Therefore, genomic DNA from the strains BC3484, BC3485, BC3486 and BC3487 was digested with *Eco*RI and probed with a 0.9 kb XhoI-*Eco*RV fragment of Tc1 in order to confirm the loss of the Tc1 element in the revertants. The Tc1 probe identified only a single fragment that correlated with the appearance and disappearance of the severe *m474::Tc1* Dpy phenotype (data not shown). A fragment of 6.5 kb was detected in BC3484 DNA but not in DNA extracted from the revertant strains (including the partial revertant), or from wild-type N2 animals.

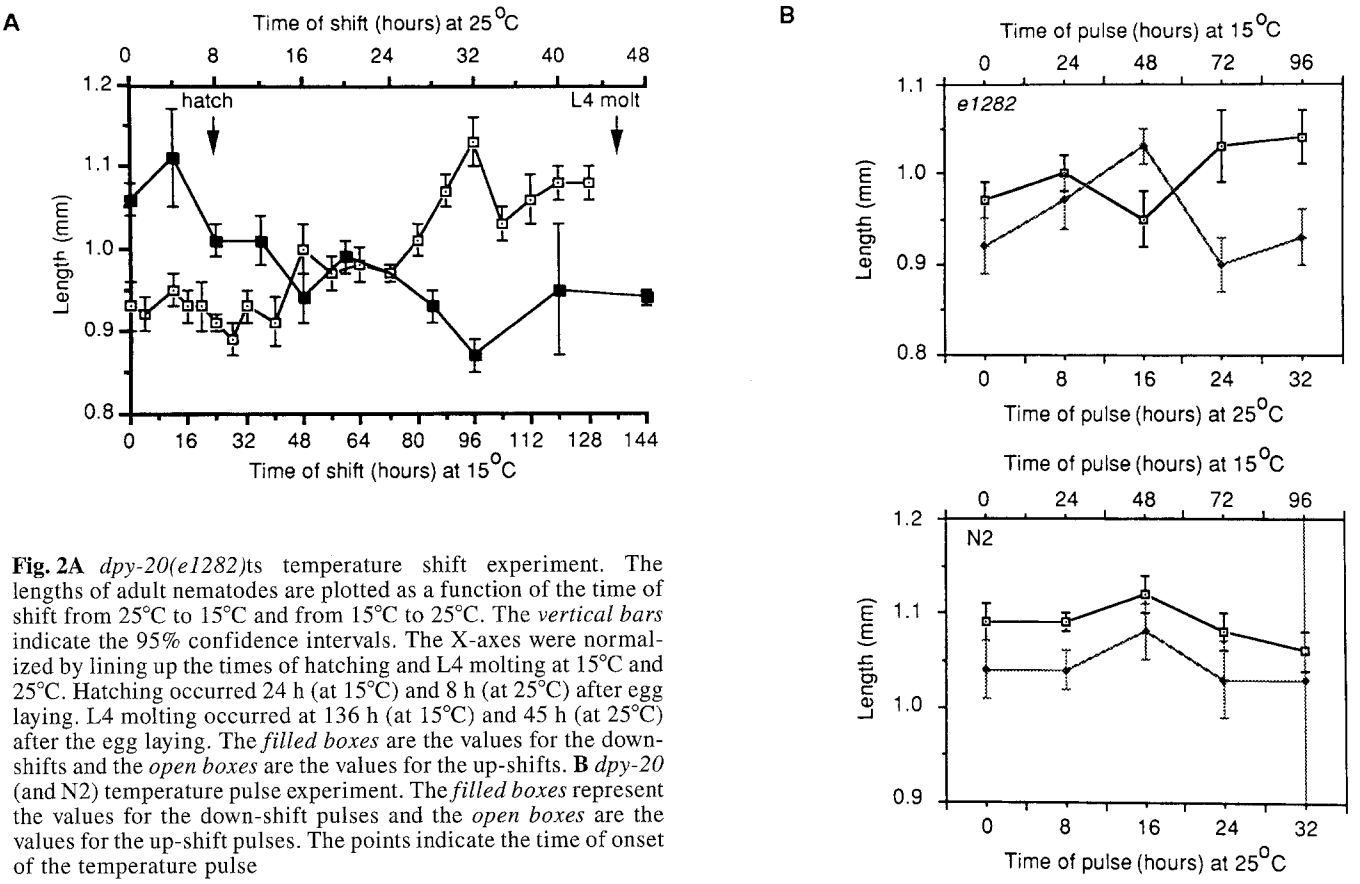


Fig. 2A *dpy-20(e1282)*ts temperature shift experiment. The lengths of adult nematodes are plotted as a function of the time of shift from 25°C to 15°C and from 15°C to 25°C. The *vertical bars* indicate the 95% confidence intervals. The X-axes were normalized by lining up the times of hatching and L4 molting at 15°C and 25°C. Hatching occurred 24 h (at 15°C) and 8 h (at 25°C) after egg laying. L4 molting occurred at 136 h (at 15°C) and 45 h (at 25°C) after the egg laying. The *filled boxes* are the values for the down-shifts and the *open boxes* are the values for the up-shifts. **B** *dpy-20* (and N2) temperature pulse experiment. The *filled boxes* represent the values for the down-shift pulses and the *open boxes* are the values for the up-shift pulses. The points indicate the time of onset of the temperature pulse

Isolation and subcloning of DNA flanking Tc1 insertion

A partial library of BC3484 genomic DNA was constructed as described in Materials and methods. After screening with the Tc1 probe described above, three clones were purified and the pBluescript plasmids were excised from the λZAP vector. An *Eco*RI insert of 6.5 kb was isolated (clone pCes103), which was of the correct size to represent the fragment that had been correlated with the *dpy-20(m474::Tc1)* phenotype. A second clone containing an insert of 6.4 kb was also isolated, but this Tc1-containing fragment was present in the mutant and revertant strains as well as in the wild-type N2 strain. No further work was done with this latter clone since it was not correlated with the mutant phenotype of interest.

The DNA flanking the Tc1 insertion in pCes103 was subcloned by digestion with *Bal*I which cut in the inverted repeats that are found at either end of the Tc1 element. The DNA was then re-ligated to produce pCes105, which contained a 4.9 kb insert with a synthetic *Bal*I site. To confirm that the DNA in pCes105 did indeed flank the Tc1 insertion site associated with the *dpy-20* mutant phenotype, a 1.4 kb fragment of pCes105 was used to probe a Southern blot of *Eco*RI-digested genomic DNA (see Fig. 3). There was a 1.6 kb difference between the sizes of the DNA fragments detected by this probe in the mutant and the wild-type or revertant strains – exactly the size of the Tc1 element. Hence, the absence or presence of this element is correlated with the absence or presence of the original mutant phenotype. Because the partially revertant strain BC3485 carried the allele *m474s2033*, and yet apparently had the same size fragment as the wild-type and the complete revertants (see Fig. 3), most of the Tc1 element must have been excised to generate the revertant *s2033* allele. The milder Dpy phenotype exhibited by BC3485 animals, therefore, presumably resulted from imprecise excision of Tc1, that resulted in a lesion that was not detectable at the level of resolution of our autoradiogram.

The sequence flanking the insertion of the *m474::Tc1* mutation was sequenced and found to lie within the *dpy-20* coding element, very near the 3′ end of the gene (Fig. 4). Since only the last eleven amino acids of the DPY-20 protein would be affected by this mutation, the severity of the phenotype of homozygous *dpy-20(m474::Tc1)* animals suggests that this region must be very important for its function.

Isolation and characterization of cDNA

Approximately 1×10^4 phage of the *C. elegans* cDNA library in λZAP (Barstead and Waterston 1989) were screened with the probe that was used to identify the Tc1 insertion site in *dpy-20(m474::Tc1)*. One positive clone

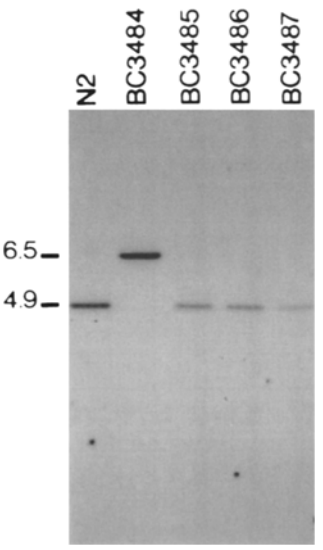


Fig. 3 Southern blot of *Eco*RI-digested wild-type (N2), BC3484, BC3485, BC3486 and BC3487 genomic DNA probed with the 1.4 kb *Eco*RI-*Eco*RV fragment from pCes105. The 6.5 kb band represents the Tc1-containing fragment in BC3484. The 4.9 kb band in all the other strains indicates the lack of the Tc1 element

was isolated and the pBluescript plasmid was designated pCes108. Subsequent screens of 6×10^5 phage of the Barstead cDNA library and 5×10^4 phage of the Arhinger/Kimble cDNA library (personal communication) using pCes108 detected no other positive colonies, suggesting that the *dpy-20* mRNA is not at all abundant. The L2 stage-specific library was used because the temperature sensitivity experiment had indicated that the DPY-20 protein is required at the L2 larval stage.

To determine the relative abundance of the *dpy-20* transcript in the larval stages and in the adult, the *dpy-20* cDNA was hybridized to a developmental Northern blot (see Fig. 5). A single transcript appeared to be expressed most abundantly in the L2 stage. Under fairly high stringency conditions (68°C), no other transcripts were apparent, and the single band shown in Fig. 5 was detected only after 10–12 days of exposure to film, supporting the hypothesis that the *dpy-20* transcript is present in very low abundance.

To determine which genomic fragments hybridized to the cDNA, Southern blots of *Eco*RI-digested DNA from N2 animals, and from three Charon 4 phage clones (containing wild-type, genomic DNA) were probed with the whole cDNA. The cDNA hybridized to the same sized fragments in both the Charon 4, cloned DNA, and the

Fig. 4 Genomic sequence of *dpy-20* is shown *above* the inferred amino acid sequence. *Underlined* sequence is that of the cDNA. Positions of intron-exon boundaries are indicated by an *arrow* above the sequence. The Tc1 insertion site is indicated by an *asterisk* (↓*)

```
ATGGAAGGGCATAGTAATACTTCTGTTATCCTGCATACTTATCAACAACAACAACCACAACAACTTATTTCTTCAATTATTCCTCAAACATTGTCAGATG  100
 M  E  G  H  S  N  T  S  V  I  L  H  T  Y  Q  Q  Q  Q  P  Q  Q  L  I  S  S  I  I  P  Q  T  L  S  D   33

GGTACGGTATTTCTGCACCTATATCACAAGCCGATTGTTCTGGAAATTCGAATTTTGTAAACACGAATCCCTGGAATAACACGACTTTTCAAGCATATCA  200
 G  Y  G  I  S  A  P  I  S  Q  A  D  C  S  G  N  S  N  F  V  N  T  N  P  W  N  N  T  T  F  Q  A  Y  Q   67

                  1↓
AACAGATGCTCAATTCACTAATGAG ATTTCCGTATTGAACTATCCAAATGTATACGATGATTTTTCAAAGGAATCCTCAAATGGAATATTATCCGACCCA  300
  T  D  A  Q  F  T  N  E    I  S  V  L  N  Y  P  N  V  Y  D  D  F  S  K  E  S  S  N  G  I  L  S  D  P  100

                                                  2↓
TCACTTCATGGATCACACAGAACCAGTTCGACTTCGGATGTTGGATCATCGGTCGATTGCAGT ATTTCCCCGGAGCCAATCCTAATGCCTTCAGTCAAAC  400
 S  L  H  G  S  H  R  T  S  S  T  S  D  V  G  S  S  V  D  C  S    I  S  P  E  P  I  L  M  P  S  V  K   133

                                                                3↓
GGGGTAGACCTACAGAGAATCCATGTTGGGCTTATTTCCACAGAATTGATGATCAGCTTGTAAAATGTCGGCTATGTACAAAG GTAGTAAGATCGGCGTG  500
 R  G  R  P  T  E  N  P  C  W  A  Y  F  H  R  I  D  D  Q  L  V  K  C  R  L  C  T  K    V  V  R  S  A  C  167

                                                                                  4↓
TGCAACAAATATGACAAAACATCTCGAAAGACATCACGCAGACGATTATCAAAAAGTGACAGGACAATTGAAACTATTCAGAATGAATGATATTGG AATT  600
  A  T  N  M  T  K  H  L  E  R  H  H  A  D  D  Y  Q  K  V  T  G  Q  L  K  L  F  R  M  N  D  I  G    I  200

CGAAGTAAAATGCATTATGAAGTGGCGGACAATCCTTTATCAACTCTCCCAGTTGTTACAAACTCTTACATTTTTACCAAAAATGGAACCAATGGACACA  700
 R  S  K  M  H  Y  E  V  A  D  N  P  L  S  T  L  P  V  V  T  N  S  Y  I  F  T  K  N  G  T  N  G  H   233

TTCGATCCCTCTCAATATTATGGAATGCAACAAGACCCAACATTAATCAACAGCAATTTATTCAATTCGAGCAACCTCAAGCTTCAATTGATCCTCAGAC  800
 I  R  S  L  S  I  L  W  N  A  T  R  P  N  I  N  Q  Q  Q  F  I  Q  F  E  Q  P  Q  A  S  I  D  P  Q  T  267

ATGGCCTCTAACCCATTTTTGGCAGAACACTGCTCAAACAAATATAATGTCTCATTTGGGAGAAGCTTCAACTTCTACATTAGGATCTTCTGTAATTCAA  900
  W  P  L  T  H  F  W  Q  N  T  A  Q  T  N  I  M  S  H  L  G  E  A  S  T  S  T  L  G  S  S  V  I  Q  300

GAAGCTCATAAAATGAATCCAGATTCAAATGACAAGAGTTTTCAACTTGATTTGGAGCAACAGAATTGTGTAATTATTGAGCAAAAGTTAGATGATAAAC 1000
 E  A  H  K  M  N  P  D  S  N  D  K  S  F  Q  L  D  L  E  Q  Q  N  C  V  I  I  E  Q  K  L  D  D  K   333

                                                                                  ↓*
CACGTGTCACGAAGCCAACTACAAAGCCGTATCAAAAGAGAAATA GGTCAGTTTTTAGTCAATTTGGAAGTTTATTTTGA                    1080
 P  R  V  T  K  P  T  T  K  P  Y  Q  K  R  N    R  S  V  F  S  Q  F  G  S  L  F  *                    359
```

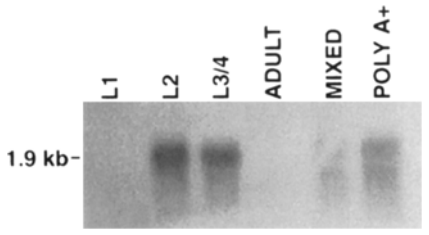


Fig. 5 Northern blot analysis of *dpy-20* transcripts during development using cDNA as a probe. The RNA in lanes L1, L2, L3/L4, adult and mixed is total RNA. From 10 to 12 μg of total RNA was loaded per lane. About 1 μg of polyA$^+$ RNA was loaded. Sizes indicated are based on positions of the 28S and 18S RNA bands as indicated by ethidium bromide staining of total RNA

wild-type, genomic DNA: the 4.9 kb *Eco*RI fragment and an additional 2.7 kb *Eco*RI fragment (data not shown).

Isolation and characterization of genomic clones

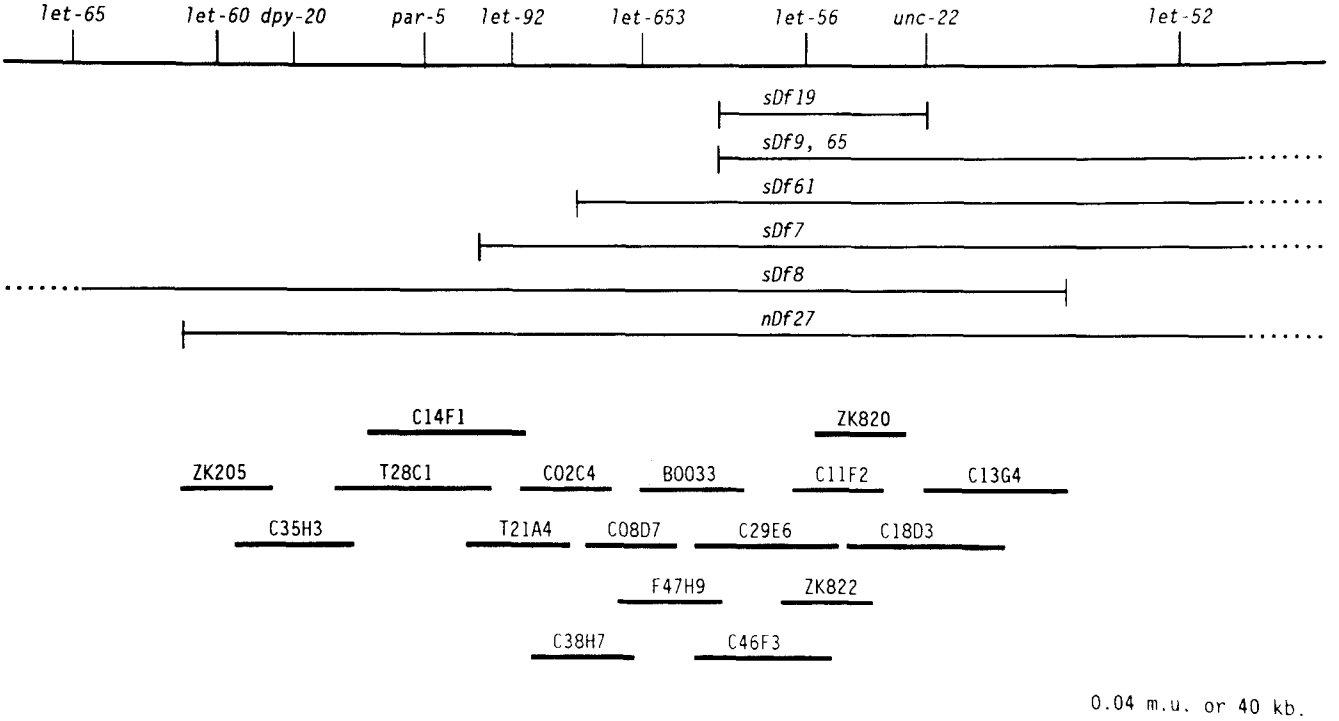
A *C. elegans* genomic DNA library in Charon 4 was screened for two reasons. Firstly, to generate a genomic clone for fingerprinting analysis at the MRC, Cambridge, which would identify the location of *dpy-20* on the physical map (Coulson et al. 1988). Secondly, to facilitate sequencing and characterization of the DNA from wild-type animals in the region around the Tc1 insertion site of *m474::Tc1*. Three overlapping λ phage were isolated (λ301, λ302, and λ303) using the 1.4 kb fragment of pCes105 (see above) as a probe. λ302 was sent to A. Coulson and J. Sulston (MRC, Cambridge, England) for placement on the physical map and it was found to lie on cosmid C35H3, about 200 kb to the left of *unc-22* (Fig. 6).

Approximately 5×10^4 phage of a *C. briggsae* genomic DNA library were screened with both the *dpy-20* cDNA and pCes177 (a 6.0 kb *Xba*I subclone of C35H3). Only a single positive *C. briggsae* clone was detected. This phage clone was digested with *Eco*RI and probed with the *C. elegans* cDNA, which detected a 1.6 kb fragment. The pCes177 probe detected the same fragment, along with a 3.9 kb fragment that probably represented extraneous DNA. The 1.6 kb fragment that was detected by both probes was sequenced to confirm the conserved starting position of the *dpy-20* gene. Sequence analysis of the *C. briggsae* gene that was similar to *dpy-20* also allowed the identification of putative upstream regulatory regions.

Germ-line rescue of *dpy-20(e1282)*ts

To provide further evidence that *dpy-20* had indeed been cloned and correctly placed on the physical map, DNA from the cosmid C35H3 was injected into the germ-line of three *dpy-20(e1282)*ts hermaphrodites that were subsequently placed on individual plates and allowed to produce self-fertilized progeny. One of the injected Dpy hermaphrodites produced a single wild-type F_1 off-

Fig. 6 Genetic map of the *dpy-20–unc-22* region with the genetically mapped deficiency breakpoints. All the genetic loci on this map have been ordered with respect to each other using deficiency mapping and three-factor crosses. Below the genetic map is a representation of the physical map, showing some of the cosmid clones



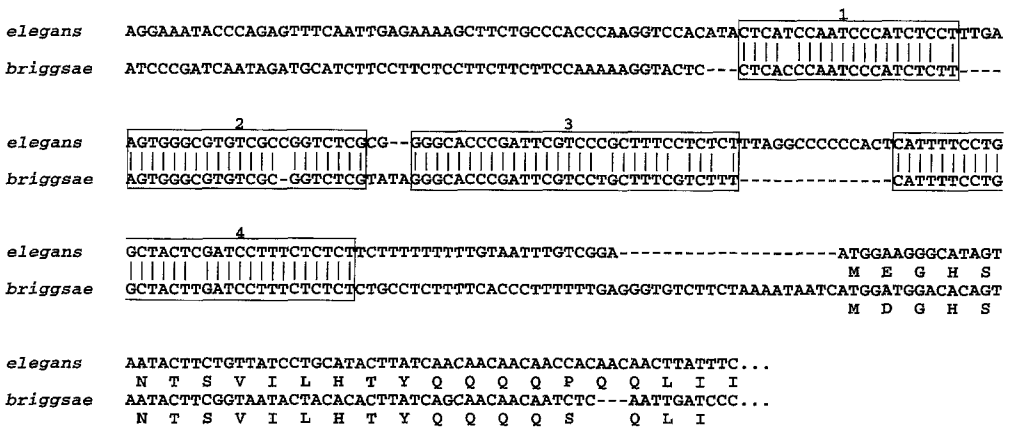


Fig. 7 *C. elegans* and *C. briggsae dpy-20* DNA sequence comparison

spring that, in turn, produced both wild-type and Dpy F_2 progeny. This indicated that the injected cosmid DNA was capable of rescuing the Dpy phenotype of *e1282*ts. The C35H3 cosmid DNA had obviously been incorporated into the germ-line of the transformed F_1 animal since it was inherited by some of the F_2 progeny, but it cannot have been stably maintained in these germ cells since some of the offspring still exhibited the Dpy phenotype of the *e1282*ts mutation.

A second hermaphrodite that had been injected with C35H3 DNA had two wild-type F_1 progeny. These F_1 animals, however, produced only Dpy F_2 animals, indicating that the transformation rescue had been transient and that the injected DNA had not been incorporated into the germ line of the parent nematode. The other injected hermaphrodite produced only Dpy progeny and was not examined further. The plasmid pCes177, containing only a portion of the C35H3 cosmid, was subsequently shown to be sufficient to rescue the Dpy phenotype of *dpy-20(e1282)*ts (M. Han and P. Sternberg, personal communication).

dpy-20 encodes a novel protein

The nucleotide sequence of the *dpy-20* coding region was determined from the cDNA and genomic DNA clones. It was found to be 1080 bp in length, encoding 359 amino acids. The cDNA sequence was used to help determine the intron-exon boundaries. There are four introns of 693 bp, 293 bp, 52 bp, and 211 bp and their positions are shown in Fig. 4. The deduced DPY-20 amino acid sequence was used to search peptide databases using the BLAST search algorithm (see Materials and methods). No similarity to any existing protein sequence was found. Specifically, the structure of dpy-20 did not resemble the collagen structure typical of most other sequenced *dpy* gene products. Two regions of a putative gene that has been sequenced by the *C. elegans* sequencing consortium (Sulston et al. 1994) did exhibit similarity to *dpy-20,* however. The putative new gene lies on the overlapping cosmids C15H7 and F58A4, on chromosome III and has the potential to encode a 874-residue protein product. Amino acids 129–186 of DPY-20 exhibit 43% identity and 60% similarity to amino acids 263–298 of the gene identified by the sequencing project, and 47% identity and 64% similarity to amino acids 621–656 of this gene.

Using the PROSITE and AACOMP programs in PCGENE, the inferred amino acid sequence of DPY-20 was analyzed. The protein was shown to contain 11.4% serine and 8.0% threonine, indicative of possible phosphorylation sites, and 7.5% asparagine residues, which may represent glycosylation sites. A hydrophobicity plot was generated using TGREASE which revealed that the DPY-20 protein is hydrophilic. Using the CHARGPRO program of PCGENE, DPY-20 was found to have an isoelectric point of 6.71. A search for PEST sequences was performed using the program PESTFIND in PCGENE. DPY-20 had no PEST regions within the inferred amino acid sequence. A search for secretory signal sequences was also performed using the program PSIGNAL in PCGENE. DPY-20 was not found to have any secretory signal sequences within its coding region.

Comparison of *C. elegans* and *C. briggsae* sequences

The *C. elegans dpy-20* gene sequence was aligned with the sequence of 500 bp near the 5′ end of the *C. briggsae dpy-20*-like gene (Fig. 7). This confirmed the position of the putative start codon of *dpy-20* since the sequences of the 60 nucleotides of the *C. elegans* and *C. briggsae* genes following this methionine were 75% identical. The amino acid sequences were 80% identical and, when conserved changes were considered, the protein products were found to be 100% conserved in the first 20 amino acids. Upstream of the initiating methionine, the nucleotide sequence similarity between *C. elegans* and *C. briggsae* decreased to 55% over approximately 300 nucleotides. There were four blocks of strong homology in the 5′ flanking sequence, however, having 90%, 83%, 95%, and 96% sequence similarity, respectively. These blocks of sequence homology probably represent regulatory elements of the *dpy-20* gene.

Discussion

Dpy-20 animals have aberrant cuticle structure when examined using electron microscopy (M. Kusch, unpublished results, cited in Cox et al. 1985). The *dpy-20* gene, therefore, presumably plays some role in cuticle development. The cuticle is a collagenous exoskeleton synthesized and secreted by the hypodermal cells during each of the four postembryonic molts of the *C. elegans* life cycle (Cox et al. 1981a). Comparison of the predicted DPY-20 protein with known peptide sequences using the BLAST search algorithm (see Materials and methods) revealed no similarity to collagen genes that are known to encode the predominant structural proteins of the cuticle. DPY-20 does not appear to contain secretory signal sequences, indicating that it is probably not an extracellular protein (von Heijne 1986). Its function may, therefore, not be structural. Northern and cDNA analysis indicate that the *dpy-20* transcript is present in very low abundance, suggesting that it is an enzyme or regulatory protein that is involved in the synthesis, processing, or export of structural proteins of the cuticle.

Lewis et al. (1980) proposed that some Dpy phenotypes may be due simply to hypercontraction of the body wall muscles that are attached to the external cuticle, but these mutants would also be expected to exhibit impaired motility. Movement in Dpy-20 animals is normal, so the mutations in this gene do not appear to affect the muscle cells of the body wall. At least six other *dpy* genes have been shown to be involved in X chromosome dosage compensation mechanisms. In *C. elegans*, hermaphrodites carry two X chromosomes, while males carry only one; hence, dosage compensation is used to equalize the levels of X-linked gene products in the two sexes. The cause of the Dpy phenotype of animals homozygous for these dosage compensation mutations is not understood, but in all cases it is sensitive to the X chromosome content of the animal's genome (Meyer and Casson 1986; Hodgkin 1987). These genes also affect the phenotype exhibited by animals carrying mutations on the X chromosome. *dpy-20* does not appear to be involved in dosage compensation since *dpy-20* mutations have no effect on the abnormal cell *lin*eage mutation *lin-15*, which is on the X chromosome. Four other *dpy* genes that are known to be involved in dosage compensation all affect the phenotype of animals carrying hypomorphic *lin-15* alleles (Meneely and Wood 1987).

Comparison of *dpy-20*-like DNA sequences in *C. elegans* and *C. briggsae* was undertaken because it had been previously shown that conserved sequences between these two species primarily define coding regions of their genomes (Snutch 1984; Prasad and Baillie 1989). *C. elegans* and *C. briggsae* are two closely related nematode species that are morphologically similar and cross fertile. Evolutionarily, the two species are tens of millions of years apart. Non-coding regions of *C. elegans* and *C. briggsae* DNA fail to hybridize to one another even under very low stringency conditions. Comparison of similar sequences between the two species can, therefore, be a useful method for detecting potential coding regions of genes (Snutch 1984; Heine and Blumenthal 1986; Prasad and Baillie 1989; Heschl and Baillie 1990). A sequence comparison between the *C. elegans dpy-20* gene and a similar *C. briggsae* gene was used in this study to reliably identify the initiating methionine of the *dpy-20* transcript. The nucleotide sequence similarity upstream of the methionine is low (55%) relative to that downstream of the methionine codon (75%; Fig. 7).

Comparisons of upstream sequences between homologous *C. elegans* and *C. briggsae* DNA have also been used to identify potential regulatory and structural elements. Heschl and Baillie (1990) compared the 5′ flanking DNA sequences of homologous, grp78-like heat shock genes in these two *Caenorhabditis* species. They demonstrated that regions of sequence identity in the upstream DNA sequence corresponded to previously identified viral enhancer sequences, a heat shock regulatory element, and other regulatory regions of the mammalian grp78 and grp94 genes. Comparisons of the upstream sequences of *dpy-20* and its putative *C. briggsae* homologue revealed four islands of strong homology (showing 83%–96% similarity; see Fig. 7) suggesting that these regions may be regulatory regions although their function is unclear at this point.

Only a single band was detected when *C. elegans* genomic DNA was probed with the *dpy-20* cDNA, even at moderately low stringency, suggesting that *dpy-20* is not a member of a multi-gene family. The *C. elegans* genome sequencing consortium (Wilson et al. 1994) has, however, detected a putative coding region on chromosome III that contains two domains that exhibit similarity to one region within *dpy-20* (data not shown). The similarity between this sequence and *dpy-20* does not extend beyond these regions, so there is no obvious functional relationship between these two loci.

An analysis of the *dpy-20(e1282)*ts mutation (Wood 1988) was undertaken in order to determine the developmental time of function of the DPY-20 gene product. Temperature shift and temperature pulse experiments showed that the mutant Dpy phenotype is exhibited if *dpy-20(e1282)*ts homozygotes are exposed to the restrictive temperature during the L1 and L2 larval stages. The *dpy-20* transcript is also most abundant in L2 total RNA, although transcripts were detected in L3/L4 RNA. The peak of expression of the *dpy-20* transcript at the L2 stage coincides with the temperature-sensitive period of the *e1282*ts allele. This observation indicates that the DPY-20 protein is required during the period at which *dpy-20* mRNA is detectable, at the L2 stage of larval development.

The phenotype of animals homozygous for the *dpy-20(e2017)*am mutation has been shown to be suppressible by mutations in $tRNA^{Trp}$ genes that allow these tRNAs to recognize the amber codon, UGA (Hodgkin 1985). Five different amber suppressors, have been shown to suppress the *e2017*am allele as well as several

amber mutations in other genes (Hodgkin 1985; Kondo et al. 1988). The degree of suppression by one or two copies of each suppressor varies depending upon the gene that is being suppressed. It has been suggested that the suppressor $tRNA^{Trp}$ genes vary in their spatial and/or temporal patterns of expression, resulting in variation in their ability to suppress amber alleles of different genes (Kondo et al. 1988). Of the four genes that have been studied (*unc-13, dpy-20, unc-52 and tra-3*), the pattern of suppression of the *dpy-20(e2017)*am mutation by different amber suppressors most closely resembles that of *unc-52(e669)*am. The *unc-52* gene product is made in the muscle cells of the body wall but is exported to the basement membrane that lies between these cells and the hypodermis (Rogalski et al. 1993). The similarity in the patterns of suppression of *dpy-20* and *unc-52* suggests that either DPY-20 may a component of the basement membrane, or that these two loci may have similar temporal patterns of expression. The latter hypothesis is not likely since the *unc-52* gene product is known to be required during embryogenesis (Rogalski et al. 1994), well before *dpy-20* transcription is detected. The fact that the two type IV collagen genes, (abnormal *emb*ryogenesis) *emb-9* (or *clb-1*), and (*let*hal) *let-2* (or *clb-2*), are known to be components of the basement membrane in *C. elegans* (Guo and Kramer 1989) supports the first hypothesis. Most mutations in these genes result in embryonic lethality, but their function has also been shown to be essential for larval development and for adult fertility (Meneely and Herman, 1979; Wood et al. 1980; Isnenghi et al. 1983).

A genetic analysis of some *dpy-20* mutants had been previously carried out by Hosono et al. (1982). They isolated a *dpy-20* mutant allele, *cn142,* and screened for suppressors of this allele by EMS treatment of *cn142* homozygotes. They observed that *cn142* homozygotes had both a shortened-body and a rounded-head phenotype. In screening for suppression of these mutant characteristics, they isolated a less severely Dpy animal that did not exhibit the rounded head phenotype (strain TN6). This suppressor phenotype appeared to be tightly linked to *dpy-20* and was dominant to *cn142.* They suggested, therefore, that the rounded head phenotype might be a manifestation of the severity of the *cn142* allele and that the suppression of the severe Dpy phenotype had resulted from an intragenic reversion event. The TN6 case parallels the difference in phenotype between animals homozygous for the *m474::Tc1* allele and those homozygous for *m474s2033,* the partially revertant allele isolated from the *mut-6* mutator strain (see Fig. 1; Table 1). Sequencing of the DNA from this region in the *dpy-20(m474s2033)* revertant strain is needed to confirm that the excision event that resulted in loss of the Tc1 element was imprecise and thus gave rise to a phenotype that is only partially revertant. Similarly, sequencing of DNA from the *dpy-20(cn142)* strain and the partially revertant strain (TN6) is necessary to confirm the hypothesis that the phenotypes of animals homozygous for the mutant or partial revertant alleles are due to different types of physical lesions in the *dpy-20* gene. Correlation of the molecular defects in *dpy-20* with the rounded-head phenotype might also be possible when the sequences of these alleles are known, and might indicate whether this is a phenotype typical of generally severe alleles of *dpy-20* or whether the 3′ region of the gene that is affected by *m474::Tc1* is responsible for this dysfunction.

In summary, this paper describes the molecular cloning and characterization of the *dpy-20* gene of *C. elegans.* We have shown that the function of the DPY-20 protein and steady-state levels of *dpy-20* mRNA are temporally related around the L2 larval stage. We have also established that, in spite of its effect on the cuticle, *dpy-20* is not a collagen gene. Because collagens are the major known components of the *C. elegans* cuticle, *dpy-20* appears to encode a novel type of protein that is involved in cuticle function and is essential for normal morphological development.

Acknowledgements We thank M.A. Marra, A. Rose, and J.E. Schein for critical review of the manuscript. We also wish to thank M.A.M. for his technical expertise and assistance. This work was supported by an Arthur Minden Muscular Dystrophy Association of Canada predoctoral fellowship to D.V.C., and by National Science and Engineering Research Council of Canada, and Muscular Dystrophy Association of Canada operating grants to D.L.B.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Babity JM (1993) A low level Tc1 transposition and Tc1 induced mutations in *dpy-5.* Ph.D. Thesis, University of British Columbia, Vancouver, B.C., Canada

Baillie DL, Beckenbach KA, Rose AM (1985) Cloning within the *unc-43* to *unc-31* interval (linkage group IV) of the *Caenorhabditis elegans* genome using Tc1 linkage selection. Can J Genet Cytol 27:457–466

Bairoch A (1990) PROSITE: a dictionary of protein sites and patterns. University of Geneva, fifth release

Barstead RJ, Waterston RH (1989) The basal component of the nematode dense body is vinculin. J Biol Chem 264:10177–10185

Brenner S (1974) The genetics of *Caenorhabditis elegans.* Genetics 77:71–94

Cabot EL, Beckenbach AT (1989) Simultaneous editing of multiple nucleic acid and protein sequences with ESEE. Comp Appl Biosci 5:233–234

Chirgwin JM, Przybyla AE, MacDonald RJ, Rutter WJ (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18:5294–5299

Clark DV, Baillie DL (1992) Genetic analysis and complementation by germ-line transformation of lethal mutations in the *unc-22(IV)* region of *Caenorhabditis elegans.* Mol Gen Genet 232:97–105

Clark DV, Rogalski TM, Donati LM, Baillie DL (1988) The *unc-22(IV)* region of *Caenorhabditis elegans*: genetic analysis of lethal mutations. Genetics 119:345–353

Coulson A, Waterston R, Kiff J, Sulston J, Kohara Y (1988) Genome linking with yeast artificial chromosomes. Nature 335:184–186

Cox GN, Hirsh D (1985) Stage-specific patterns of collagen gene expression during development of *Caenorhabditis elegans.* Mol Cell Biol 5:363–372

Cox GN, Kusch M, DeNevi K, Edgar RS (1981a) Temporal regulation of cuticle synthesis during development of *Caenorhabditis elegans*. Dev Biol 84: 277–285

Cox GN, Kusch M, Edgar RS (1981b) Cuticle of *Caenorhabditis elegans*: its isolation and partial characterization. J Cell Biol 90: 7–17

Cox GN, Staprans S, Edgar RS (1981c) The cuticle of *Caenorhabditis elegans* II. Stage-specific changes in ultrastructure and protein composition during postembryonic development. Dev Biol 86: 456–470

Cox GN, Carr S, Kramer JM, Hirsh D (1985) Genetic mapping of *Caenorhabditis elegans* collagen genes using DNA polymorphisms as phenotypic markers. Genetics 109: 513-528

Davis RW, Botstein D, Roth JR (1980) A manual for genetic engineering: advanced bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Emmons SW, Klass MR, Hirsh D (1979) Analysis of the constancy of DNA sequences during development and evolution of the nematode *Caenorhabditis elegans*. Proc Natl Acad Sci USA 76: 1333–1337

Feinberg A, Vogelstein B (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132: 6–13

Fire A (1986) Integrative transformation of *Caenorhabditis elegans*. EMBO J 5: 2673-2680

Guo X, Kramer JM (1989) The two *Caenorhabditis elegans* basement membrane (type IV) collagen genes are located on separate chromosomes. J Biol Chem 264: 17575–17582

Heery DM, Gannon F, Powell R (1990) A simple method for subcloning DNA fragments from gel slices. Trends Genet 6: 173

Heine U, Blumenthal T (1986) Characterization of regions of the *Caenorhabditis elegans* X chromosome containing vitellogenin genes. J Mol Biol 188: 301–312

Henikoff S, Eghtedarzadeh K (1987) Conserved arrangement of nested genes at the *Drosophila Gart* locus. Genetics 117: 711–725

Heschl MFP, Baillie DL (1990) Functional elements and domains inferred from sequence comparisons of a heat shock gene in two nematodes. J Mol Evol 31: 3–9

Hodgkin J (1985) Novel nematode amber suppressors. Genetics 111: 287–310

Hodgkin J (1987) Sex determination and dosage compensation in *Caenorhabditis elegans*. Annu Rev Genet 21: 133–154

Horvitz HR, Brenner S, Hodgkin J, Herman RK (1979) A uniform genetic nomenclature for the nematode *Caenorhabditis elegans*. Mol Gen Genet 175: 129–133

Hosono R, Hirahara K, Kuno S, Kurihara T (1982) Mutants of *Caenorhabditis elegans* with dumpy and rounded head phenotype. J Exp Zool 224: 135–144

Isnenghi E, Cassada R, Smith K, Denich K, Radnia K, von Ehrenstein G (1983) Maternal effects and temperature-sensitive period of mutations affecting embryogenesis in *Caenorhabditis elegans*. Dev Biol 98: 465–480

Kondo K, Hodgkin J, Waterston RH (1988) Differential expression of five $tRNA^{trp}$(AUG) amber suppressors in *Caenorhabditis elegans*. Mol Cell Biol 8: 3627–3635

Kramer JM (1994) Structures and functions of collagens in *Caenorhabditis elegans*. FASEB J 8: 329–336

Lewis JA, Wu C-H, Berg H, Levine JH (1980) The genetics of levamisole resistance in the nematode *Caenorhabditis elegans*. Genetics 95: 905–928

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Meneely PM, Herman RK (1979) Suppression and function of X-linked lethal and sterile mutations in *Caenorhabditis elegans*. Genetics 97: 65–84

Meneely PM, Wood WB (1987) Genetic analysis of X-chromosome dosage compensation in *Caenorhabditis elegans*. Genetics 117: 25–41

Meyer BJ, Casson LP (1986) *Caenorhabditis elegans* compensates for the difference in X chromosome dosage between the sexes by regulation of transcript levels. Cell 47: 871–881

Moerman DG, Baillie DL (1979) Genetic organization in *Caenorhabditis elegans*: fine structure analysis of the *unc-22* gene. Genetics 91: 95–103

Mori I, Moerman DG, Waterston RH (1988) Analysis of a mutator activity necessary for germline transposition and excision of Tc1 transposable elements in *Caenorhabditis elegans*. Genetics 120: 397–407

Ouazana R, Garrone R, Godet J (1985) Characterization of morphological and biochemical defects in the cuticle of a dumpy mutant of *Caenorhabditis elegans*. Comp Biochem Physiol B 80: 481–484

Pearson WR, Lipman DJ (1988) Improved tools for biological sequence analysis. Proc Natl Acad Sci USA 85: 2444–2448

Prasad SS, Baillie DL (1989) Evolutionarily conserved coding sequences in the *dpy-20–unc-22* region of *Caenorhabditis elegans*. Genomics 5: 185–198

Priess JR, Hirsh D (1986) *Caenorhabditis elegans* morphogenesis: the role of the cytoskeleton in elongation of the embryo. Dev Biol 117: 156–173

Rigby PW, Dieckmann, Rhodes C, Berg P (1977) Labelling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase. J Mol Biol 113: 237-251

Rogalski TM, Baillie DL (1985) Genetic organization of the *unc-22 IV* gene and the adjacent region in *Caenorhabditis elegans*. Mol Gen Genet 201: 409–414

Rogalski TM, Gilchrist EJ, Mullen GP, Moerman DG (1995) Mutations in the *unc-52* gene responsible for body wall muscle defects in adult *Caenorhabdits elegans* are located in alternatively spliced exons. Genetics 139: 159–169

Rogalski TM, Moerman DM, Baillie DL (1982) Essential genes and deficiencies in the *unc-22 IV* region of *Caenorhabitis elegans*. Genetics 102: 725–736

Rogalski TM, Williams BD, Mullen GP, Moerman DG (1993) Products of the *unc-52* gene in *Caenorhabditis elegans* are homologous to the core protein of the mammalian basement membrane heparan sulfate proteoglycan. Genes Dev 7: 1471–1484

Rychlik W, Rhoads RE (1989) A computer programme for choosing optimal oligonucleotides for filter hybridization, sequencing, and *in vitro* amplification of DNA. Nucleic Acids Res 17: 8543–8551

Sebastiano M, Lassandro F, Bazzicalupo P (1991) *cut-1*, a *Caenorhabditis elegans* gene coding for a dauer-specific noncollagenous component of the cuticle. Dev Biol 146: 519–530

Smith GE, Summers MO (1980) The bidirectional tranfer of DNA and RNA to nitrocellulose or DBM paper. Anal Biochem 109: 123–129

Snutch TP (1984) A molecular and genetic analysis of the heat shock response of *Caenorhabditis elegans*. Ph.D. Thesis, Simon Fraser University, Burnaby, B.C., Canada

Von Heijne G (1986) A new method for predicting signal sequence cleavage sites. Nucleic Acids Res 14: 4683–4690

Wilson R et al. (1994) 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. Nature 368: 32–38

Wood WB (1988) The nematode *Caenorhabditis elegans*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Wood WB, Hecht R, Carr S, Vanderslice R, Wolf N, Hirsh D (1980) Parental effects and phenotypic characterization of mutations that affect early development in *Caenorhabditis elegans*. Dev Biol 74: 446–469