The use of deficiencies to determine essential gene content in the *let-56–unc-22* region of *Caenorhabditis elegans*

JACQUIE E. SCHEIN[1] AND MARCO A. MARRA

Institute of Molecular Biology and Biochemistry, Department of Biological Sciences, Simon Fraser University, Burnaby, BC V5A 1S6, Canada

GUY M. BENIAN

Department of Pathology, Emory University, Atlanta, GA 30322, U.S.A.

CHRIS FIELDS

The Institute for Genomic Research, 932 Clopper Road, Gaithersburg, MD 20782, U.S.A.

AND

DAVID L. BAILLIE

Institute of Molecular Biology and Biochemistry, Department of Biological Sciences, Simon Fraser University, Burnaby BC V5A 1S6, Canada

Corresponding Editor: J. Bell

Received April 22, 1993

Accepted September 7, 1993

SCHEIN, J.E., MARRA, M.A., BENIAN, G.M., FIELDS, C., and BAILLIE, D.L. 1993. The use of deficiencies to determine essential gene content in the *let-56–unc-22* region of *Caenorhabditis elegans*. Genome, **36**: 1148-1156.

We have investigated the possibility of using the polymerase chain reaction to detect deletions of coding elements in the *unc-22–let-56* interval on chromosome IV in the nematode *Caenorhabditis elegans*. Our analysis of approximately 13 kb of genomic sequence immediately to the left of the *unc-22* gene resulted in the identification of four possible genes. Partial cDNAs have been identified for three of them. To determine whether any of these coding elements are essential for development, we required a method for the induction and selection of mutations in these elements. Our approach was to identify a set of formaldehyde and gamma radiation induced *unc-22* mutations that mapped to the *unc-22–let-56* region, and then employ polymerase chain reaction methodology to identify deficiencies that affected one or more of the four identified coding elements. Two small deficiencies were identified in this manner. Characterization of these deficiencies shows that there are no coding elements between *unc-22* and *let-56* (the nearest mutationally identified gene to the left of *unc-22*), which are required in development under laboratory conditions. We conclude that the polymerase chain reaction is a practical tool for the detection of deletions of coding elements identified in this region, and that characterization of such deficiencies provides a method for assessing whether or not these elements are required for development.

Key words: *Caenorhabditis elegans*, deficiencies, coding elements, *unc-22*.

SCHEIN, J.E., MARRA, M.A., BENIAN, G.M., FIELDS, C., et BAILLIE, D.L. 1993. The use of deficiencies to determine essential gene content in the *let-56–unc-22* region of *Caenorhabditis elegans*. Genome, **36** : 1148-1156.

La possibilité d'utiliser la réaction en chaîne de la polymérase pour détecter les délétions des éléments codants dans l'intervalle *unc-22–let-56* sur le chromosome IV de la nématode *Caenorhabditis elegans* a fait l'objet de la présente recherche. L'analyse d'environ 13 kb de séquences génomiques situées immédiatement à gauche du gène *unc-22* a permis d'identifier quatre gènes possibles; une partie des ADNc a été identifiée pour trois d'entre eux. Pour déterminer si certains de ces éléments codants sont essentiels au développement, il a fallu trouver une méthode pour induire et sélectionner des mutations dans ces éléments. L'approche choisie a été d'identifier un groupe de mutations induites dans l'*unc-22*, cartographié dans la région *unc-22–let-56*, par du formaldéhyde et des radiations gamma, puis d'utiliser la méthodologie de la réaction de la polymérase en chaîne pour identifier les déficiences affectant l'un ou l'autre des quatre éléments codants identifiés. Deux petites déficiences ont été identifiées par cette méthode. La caractérisation de ces déficiences montre qu'il n'y a pas d'éléments codants entre *unc-22* et *let-56* (gène mutationnel identifié comme le plus près à gauche de l'*unc-22*) qui soient requis pour le développement dans des conditions de laboratoire. La conclusion dégagée est que la réaction de la polymérase en chaîne est un outil pratique pour la détection de délétions d'éléments codants identifiés dans cette région et que la caractérisation de telles déficiences fournit un moyen d'établir si ces éléments sont requis ou non pour le développement.

Mots clés : *Caenorhabditis elegans*, déficiences, éléments codants, *unc-22*.

[Traduit par la rédaction]

Introduction

The *unc-22* region of chromosome IV in *Caenorhabditis elegans* has been extensively studied in our laboratory, with particular emphasis on the region delimited by the breakpoints of the deficiency *sDf2* (Rogalski et al. 1982; Rogalski and Baillie 1985; Clark et al. 1988; Clark 1990; Clark and Baillie 1992). This 2 map unit region is estimated to be 70% saturated for mutations in essential genes (Clark and Baillie 1992). A total of 36 essential genes have been identified and these fall into 14 zones, which are defined by deficiency breakpoints (Clark and Baillie 1992).

Correlation of the genetic and physical maps in this region has been achieved by germline transformation using cosmid clones. Using this technique, four genes in the 0.2 map unit

[1]Corresponding author.

Printed in Canada / Imprimé au Canada

interval bounded on the left by *dpy-20* and on the right by *unc-22* (see Fig. 1) have been placed on the physical map: *dpy-20* (Clark 1990), *let-653*, *let-56* (Clark and Baillie 1992), and *par-5* (D. Shakes, personal communication).

Several coding regions in the *dpy-20–unc-22* interval were identified by Prasad and Baillie (1989), who subcloned restriction fragments from four cosmids in the *dpy-20–unc-22* cosmid contig, and, using interspecies cross-hybridization to *Caenorhabditis briggsae* genomic DNA, identified nine putative coding regions. Seven of these nine restriction fragments also hybridized to RNA on *C. elegans* Northern blots. The four cosmids analyzed represent approximately half the DNA between *dpy-20* and *unc-22* (Prasad and Baillie 1989). If we extrapolate from these results, there should be at least 14 coding elements in the *dpy-20–unc-22* region. However, only six genes have been identified genetically in this interval, including *dpy-20* and *unc-22* (see Fig. 1). It therefore appears that the *dpy-20–unc-22* interval contains additional genes that have yet to be identified genetically.

This study concentrates on a small (approximately 0.05 map unit) interval bounded on the left by *let-56* and on the right by *unc-22*. This small interval has been well characterized at the molecular level. In addition to the cosmid rescue of *let-56* (Clark and Baillie 1992) and the detection of at least three coding regions between *let-56* and *unc-22* (Prasad and Baillie 1989), approximately 13 kb of genomic sequence 3′ (to the left) of *unc-22* was generated during the sequencing of the *unc-22* gene (Benian et al. 1989, 1993). Our analysis of this 13 kb of sequence identifies four potential coding elements. By recovery of partial cDNA clones, we have demonstrated that at least three of these genes are transcribed. Although no genes have been identified genetically between *let-56* and *unc-22* in screens designed to detect recessive lethal mutations in the *sDf2* region (Rogalski et al. 1982; Rogalski and Baillie 1985; Clark et al. 1988; Clark and Baillie 1992; M. Marra, J. Schein and D. Baillie, manuscript in preparation), it is possible that some of the identified coding elements are essential genes. This possibility is likely given that *let-56* and *unc-22* are separated by approximately one cosmid length of DNA, and germline transformation data suggest an average of one essential gene per cosmid in the *dpy-20–unc-22* interval (Clark and Baillie 1992). To determine whether any of the four coding elements identified immediately 3′ of the *unc-22* gene encode genes essential for development requires a method for the induction and isolation of mutations in these elements. We reasoned that characterization of individuals carrying deletions of these coding elements would likely provide insight as to whether the coding elements are required for development. The goal of this study, as a result of a suggestion by R.H. Waterston, was to determine whether the polymerase chain reaction (PCR) (Saiki et al. 1988) could be employed to identify *unc-22* deficiencies that deleted one or more of these coding elements. To screen for deletions of this region, we made use of the fact that mutations in *unc-22* alone, or deficiencies that delete *unc-22* and neighboring DNA, are easily recovered since individuals carrying such mutations express a visible, conditionally dominant Unc-22 phenotype (Moerman and Baillie 1979, 1981). We examined a set of strains exhibiting the dominant Unc-22 phenotype obtained after treatment with mutagens known to cause chromosome breakage. We designed PCR primer pairs specific for each of the four coding elements immediately 3′ of *unc-22*, and, together with genetic analysis, used these primers to identify *unc-22* mutations that were deficiencies with breakpoints in the *let-56–unc-22* region. Once such deficiencies were identified, the viability of deficiency homozygotes could be correlated with the coding elements deleted. Our results suggest this approach to be a practical strategy for examining whether the coding elements in the *let-56–unc-22* interval encode genes essential for development.

Materials and methods

The genetic nomenclature follows the recommendations of Horvitz et al. (1979).

Nematode strains and culture conditions

Nematodes were maintained on Petri plates containing nematode growth medium (NGM) streaked with *Escherichia coli* OP50 (Brenner 1974). The wild-type strain N2 (var. Bristol) and a strain carrying the *unc-31(e169)* mutation were obtained from the stock collection at the Medical Research Council, Cambridge, England. The remaining LGIV mutations used in this study were isolated at Simon Fraser University: *unc-22(s7)* (Moerman and Baillie 1979), *let-52(s42)*, *let-56(s173)* (Rogalski et al. 1982), and *let-653(s1733)* (Clark and Baillie 1992).

All *unc-22* mutations have a recessive "twitcher" phenotype with the exception of *sDf19* (Rogalski and Baillie 1985), which has a dominant twitcher phenotype. Mutations in *unc-22* also have a conditionally dominant phenotype, such that individuals heterozygous for an *unc-22* mutation twitch in a 1% nicotine solution (Sigma), but they are otherwise phenotypically wild type (Moerman and Baillie 1979). Homozygous wild-type individuals are contracted and paralyzed in nicotine.

Isolation of deficiencies of unc-22

Twenty-seven of the 58 putative *unc-22* deficiencies analyzed were isolated in this study using essentially the same protocol described by Moerman and Baillie (1981). However, only 0.1% formaldehyde was used, as recommended by Johnsen and Baillie (1988). Five mutagenized N2 hermaphrodites were placed on each of 95 60-mm Petri plates and allowed to lay eggs for 24 h. These animals were then transferred to fresh plates and allowed to lay eggs for an additional 24 h, after which they were removed. The F_1 generation was screened in 1% nicotine for the presence of twitchers. Only one twitcher was kept from each plate if more than one was found. The self-progeny of each F_1 twitcher selected were screened for the presence of Unc-22's. Individual F_1's that gave no fertile Unc-22's were suspected of carrying a deficiency affecting *unc-22* and a neighboring essential gene or genes. These strains were maintained every generation by selecting phenotypically wild-type individuals that twitched in nicotine. Lines that gave viable, fertile Unc-22's were also retained and maintained as homozygous twitchers.

Thirty-one of the 58 putative *unc-22* deficiencies analyzed in this study were isolated previously in our lab in a number of screens. All 31 strains are viable when homozygous for the *unc-22* mutation. Twenty-four of these 31 were isolated in screens similar to that described above using either 0.07 or 0.1% formaldehyde (Moerman and Baillie 1981). Six were isolated from screens using either 1500 R gamma radiation (D.L. Baillie and D.V. Clark, unpublished results) or 2500 R gamma radiation (D.G. Moerman, unpublished results), while one was isolated incidentally following mutagenesis of a *dpy-5–unc-13* strain with 0.05% formaldehyde (A.M. Rose, unpublished results).

Complementation tests

All complementation tests were performed at 20°C. Complementation tests with lethal mutations were performed as follows. Hermaphrodites of the genotype *sDfx*/+ were mated to either *let-y unc-22 unc-31*/+ + + males or *let-y unc-22*/+ + males. The outcrossed progeny were screened for the presence of

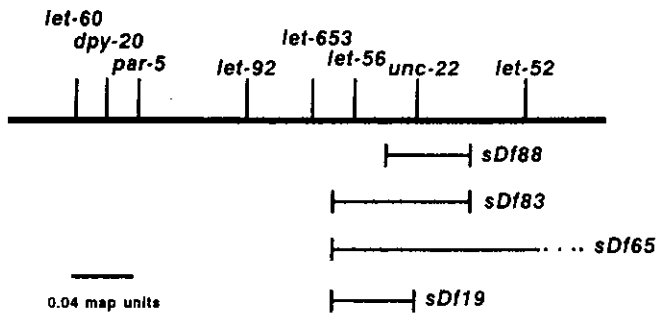


FIG. 1. A genetic map of the *let-60–let-52* region on chromosome IV. The positions of genes in the region are shown above the heavy horizontal line. The genetic breakpoints of deficiencies in the region that are discussed in this study are shown below the line.

fertile Unc-22 hermaphrodites. If no fertile Unc-22 hermaphrodites were found in the outcrossed progeny, then the lethal mutation was considered to be uncovered by the deficiency.

Construction of sDf83 *homozygous animals bearing* sEx25

Hermaphrodites of the genotype *sDf83/+* were mated with wild-type males. Male progeny (of the genotype *sDf83/+*) resulting from this cross were selected in nicotine and mated with hermaphrodites carrying *sEx25* (BC4537), an extrachromosomal array containing plasmids that represent the rightmost 12 kb of the cosmid C11F2. *sEx25* also contains the plasmid pRf4. pRf4 contains the dominant Roller allele *rol-6(su1006)* (Kramer et al. 1990), so that an extrachromosomal array containing pRf4 confers a dominant Roller phenotype on individuals carrying the array. *sEx25* has been shown to rescue *let-56(s173)* (S. Jones and M. Marra, unpublished results). Mated hermaphrodites were set individually and allowed to lay eggs. The progeny of this cross were screened for the presence of rolling animals that twitched in nicotine. Progeny exhibiting this phenotype were of the genotype *sDf83/+; sEx25*. These animals were picked individually to separate plates and allowed to lay eggs. Their progeny were screened for the presence of constitutive twitcher animals that rolled. These animals are of the genotype *sDf83/sDf83; sEx25*. When animals exhibiting this phenotype were found they were removed singly to individual plates for observation.

PCR mapping

Oligonucleotide primers 18 bp in length were designed using the program OLIGO (Rychlik and Rhoads 1989) from the sequence 3′ of *unc-22* (Benian et al. 1989, 1993). The primer pairs were designed from sequences in the putative 5′ region of each of the coding elements identified. The direction of transcription of each element is shown in Fig. 2. The Na^+/H^+ antiporter-like gene primers (DB19 and MAM2) and DOPA decarboxylase-like gene primers (R1 and R8) were designed from cDNA sequences (Marra et al. 1993). Primers KT01 and KT03 are specific for a region approximately 2 kb upstream of *unc-22*. Control primer pairs specific for sequences on chromosome V (either 265F and 265R, or KO8-3 and KO8-7) were kindly provided by K.S. McKim. PCR reactions were performed following a whole worm protocol devised by Barstead and Waterston (1991) using reagents supplied in the GeneAmp PCR kit (Perkin Elmer Cetus), and either *Taq* DNA polymerase (Perkin Elmer Cetus) or *Pfu* DNA polymerase (Stratagene). The PCR products were resolved on agarose gels. The sequences of primers specific for the four coding elements identified immediately 3′ of *unc-22* are (5′ to 3′). Glucose transporter-like primers: Lunc1-CCGTTTGAGAGTTGAGCG and Lunc2-CGACGAGACAGCCACAAC. Proline transporter-like primers: Mys1-AAGTTCGTAAAGATGCCC and Mys2-AGATTTGGGCATTTTTCG. Dup primers: Dup1-ATTCACAAAACACATCCC and Dup2-GTTATGTGACCGATGAGC. C6 primers: C61-GCCAGCAACAAACCGAAG and C62-AAGAACAGCAGACAACGC.

Detection of cDNAs using PCR

Phage particles from 20 mL of a lambda ZAP (Stratagene) cDNA library (Barstead and Waterston 1989) (2×10^{10} pfu) were isolated and polyethylene glycol (PEG) (Sigma) precipitated following the protocol described by Sambrook et al. (1989). The precipitate was brought up in 1 mL of SM buffer (1 L of SM buffer contains 5.8 g NaCl, 2 g $MgSO_4$, 50 mL 1 M Tris–HCl (pH 7.5), and 5 mL 2% gelatin solution) and, following PEG extraction, was digested with Proteinase K (Sigma) for 5 min at 65°C. This phage DNA solution was diluted 10-fold and 1 μL was used in PCR reactions with a primer pair specific for either the 5′ region of the proline transporter-like gene (primers Mys1 and Mys3), the 5′ region of the Dup sequence (primers Dup3 and Dup4), or a primer pair encompassing sequence between the 5′ regions of the two genes (primers Mys1 and Dup2). Figure 2 shows the positions and transcriptional polarity of these genes. A primer pair (R3 and R11) designed from the cDNA sequence of the DOPA decarboxylase-like gene (Marra et al. 1993) was used as a control. PCR reactions were carried out as described above. The sequences of the Dup and proline transporter-like gene primers used are (5′ to 3′) as follows. Proline transporter-like gene: Mys1 (see above) and Mys3-AACAACTCTGCCCATTCG. Dup gene: Dup2 (see above), Dup3-GCTCATCGGTCACATAAC and Dup4-TAATCTTTTCACTGCTCC.

Database searches

Amino acid sequences were searched against entries in the SwissProt, PIR, GenPept, and GPUpdate peptide sequence databases using the BLAST algorithm (Altschul et al. 1990). The computation was performed at the National Center for Biotechnology Information (NCBI) using the BLAST network service.

Results

Isolation of unc-22 *deficiencies*

A total of 27 wild-type F_1 individuals that twitched in nicotine were isolated from approximately 190 000 wild-type chromosomes treated with 0.1% formaldehyde. Each was presumed to be heterozygous for a recessive *unc-22* mutation. Of these, 17 gave fertile Unc-22 progeny, while 10 gave no fertile Unc-22 progeny. Strains were established for each of the 27 mutations (see Material and methods).

Complementation tests

The mutations carried by the 10 strains that failed to produce fertile Unc-22 progeny were presumed to be deficiencies of *unc-22* and a neighboring essential gene(s). These mutations were tested for complementation with *let-653*, which lies immediately to the left of *let-56* (Fig. 1). Nine of the 10 mutations failed to complement *let-653* and thus were likely deficiencies of *unc-22*, which extended at least as far as *let-653*. These nine deficiencies were too large for our purposes and were not analyzed further in this study.

The single mutation that complemented *let-653* was further tested and found to fail to complement *let-56*, which lies between *let-653* and *unc-22*. This result demonstrated that the mutation was likely a deficiency affecting both *let-56* and *unc-22*. We designated this mutation *sDf83*. The phenotype of *sDf83* homozygous individuals is similar to that observed for *let-56* homozygous individuals. Individuals homozygous for *sDf83* are very slow growing animals with an Unc-22 phenotype. Two weeks after being laid as eggs, these animals reach an early adult stage, based on their length as compared with the *unc-22* growth curve (Rogalski et al. 1982)

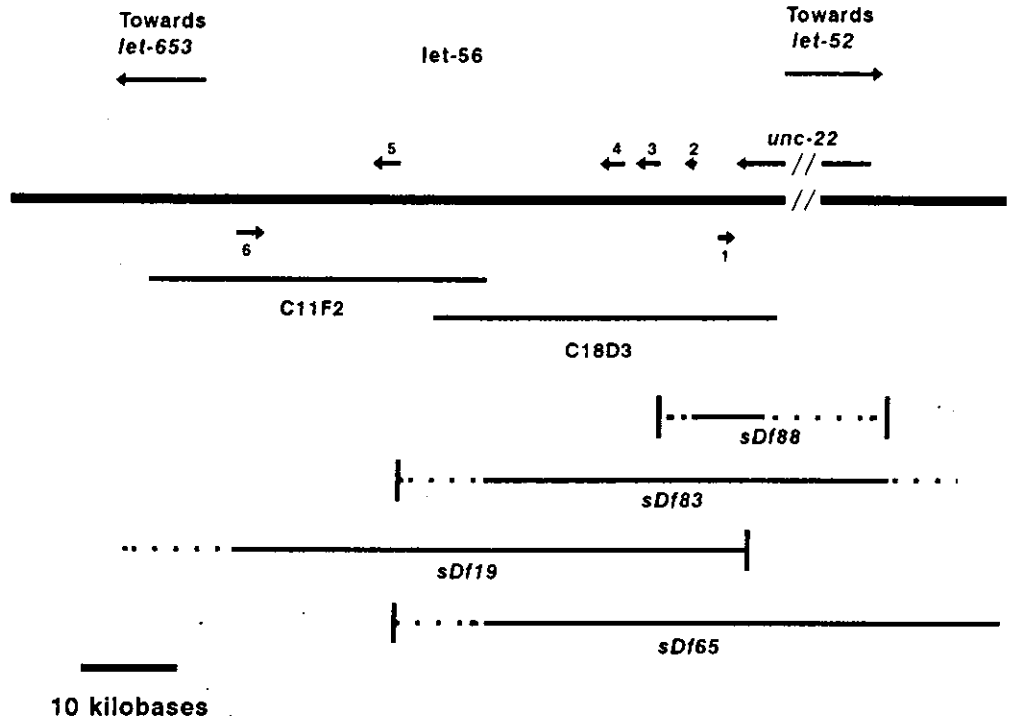


FIG. 2. A physical map showing the positions of the coding elements in the *let-56–unc-22* region (numbered arrows) and the inferred molecular breakpoints of the deficiencies discussed in this study. Arrows indicate directions of transcription. The break in the *unc-22* gene indicates that approximately 28 kb are not shown. The approximate position of *let-56* is indicated above the heavy horizontal line. The positions of cosmids C11F2 and C18D3 and the deficiencies are shown below the heavy horizontal line. Dotted lines indicate that the exact position of the deficiency breakpoint is unknown. The identity of the numbered coding elements is as follows: 1, glucose transporter-like gene; 2, proline transporter-like gene; 3, Dup; 4, C6; 5, DOPA decarboxylase-like gene; 6, Na^+/H^+ antiporter-like gene.

and examination of gonadal development (data not shown). We have never observed *sDf83* homozygous individuals to be fertile, although occasionally some appear to carry an oocyte. To determine the position of the right breakpoint of *sDf83*, a complementation test was performed with *let-52*, the nearest gene to the right of *unc-22* on the genetic map for which multiple presumed point mutations are available. *sDf83* was found to complement *let-52*. *sDf83* therefore has its left breakpoint between *let-653* and *let-56* and its right breakpoint between *unc-22* and *let-52* (Fig. 1). The strain bearing *sDf83* was retained for further analysis.

Analysis of sequence 3′ of unc-22

The *unc-22* sequence (Benian et al. 1989, 1993) includes approximately 13 kb of sequence 3′ of the *unc-22* gene. Analysis of this 13-kb sequence for open reading frames and for typical nematode introns (consensus donor and acceptor sites, high AT content, and small size (Blumenthal and Thomas 1988)), identified four possible genes (indicated as arrows in Fig. 2). The first of these genes, the glucose transporter-like gene, is oriented with opposite transcriptional polarity to *unc-22* and extends from a putative initiator methionine at nucleotide 43 928 to nucleotide 42 190 in the sequence (nucleotide positions correspond to those in the updated sequence, GenBank Accession No. L10351, Benian et al. 1993). It specifies a protein that has similarity to a family of glucose transporters, including a human insulin-responsive glucose transporter (Fukumoto et al. 1989) (Fig. 3A). We detect a 2.4-kb message on blots of total RNA, and a partial 870-bp cDNA shows that its polyadenylated message overlaps the 944 nucleotide 3′ untranslated sequence of the *unc-22* message by 78 nucleotides, on the opposite strand (data not shown).

The second gene, the proline transporter-like gene, has the same transcriptional polarity as *unc-22*. A putative initiator methionine lies at nucleotide 46 267. It is unclear where the 3′ end of this gene lies, but it is likely to extend at least as far as nucleotide 47 476. A search of peptide sequence databases (see Materials and methods) showed that a 327 amino acid protein predicted from this sequence has similarity to a family of neurotransmitter transporters. The best match was with a high affinity L-proline transporter (rat) (Fig. 3B). The sequence similarity between this *C. elegans* sequence and the neurotransmitter transporter family was also noted by Liu et al. (1992). A partial, approximately 700-bp cDNA corresponding to this gene was PCR amplified from a cDNA library using primers specific for the putative 5′ region of the gene (see Materials and methods). When this cDNA was labeled and used to probe Northern blots of total RNA, an approximately 4.2-kb message was detected (data not shown).

The third putative gene Dup (for duplicated), has a potential initiator methionine at nucleotide 49 875, and the putative 3′ end lies at nucleotide 52 239. Dup does not show similarity to any sequences in the peptide databases searched. The possibility exists that Dup is part of the proline transporter-like gene, given the size of the mRNA detected for the proline transporter-like gene, and the proximity of the putative 5′ end of Dup and the putative 3′ end of the proline transporter-like gene (Fig. 2). Attempts to PCR amplify a

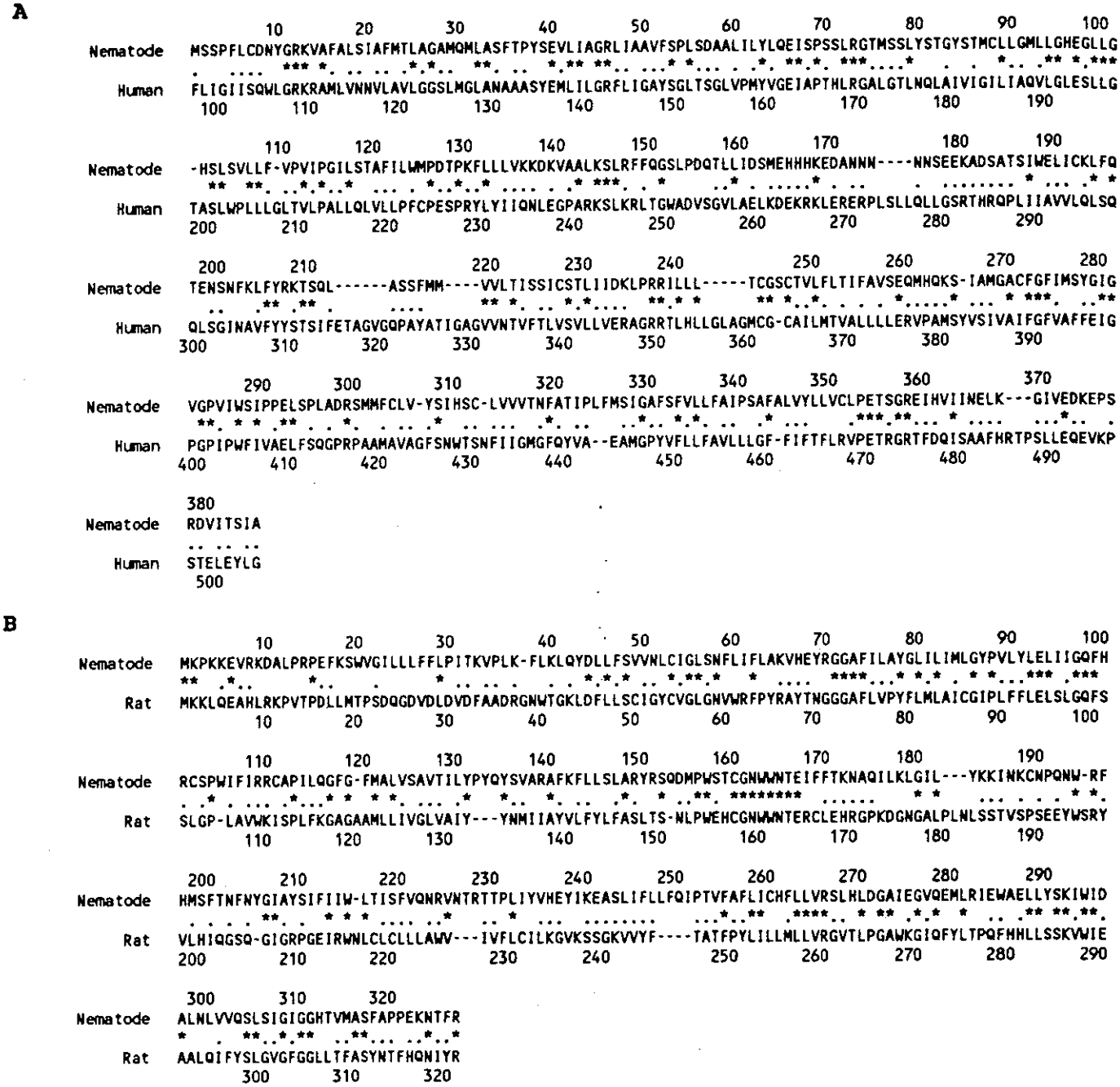


FIG. 3. FASTA amino acid sequence comparisons (Pearson and Lipman 1988). Asterisks indicate identical amino acids and single dots indicate conservative amino acid changes. (A) Comparison of the *C. elegans* glucose transporter-like gene (Nematode) and the 509 amino acid human GLUT4 insulin-responsive glucose transporter (Fukumoto et al. 1989). The *C. elegans* amino acid sequence is derived from a combination of genomic and cDNA sequence data. (B) Comparison of the *C. elegans* L-proline transporter-like gene (Nematode) and the 637 amino acid rat high affinity L-proline transporter (Fremeau et al. 1992).

cDNA using either Dup-specific primers, or a primer pair encompassing both Dup and proline transporter-like gene sequences (see Materials and methods), were unsuccessful.

The fourth gene, called C6, has a potential initiator methionine at nucleotide 53 438 and extends past the end of the *unc-22* sequenced region. The sequence of a partial cDNA, isolated previously (S. Prasad, unpublished results), was found to overlap with the last approximately 600 bp of the *unc-22* region sequence. A 2.4-kb message is detected on Northern blots of total RNA (Prasad and Baillie 1989). We believe C6 mRNA to be moderately abundant, as we found it represented twice in the 1517 cDNAs sequenced by Waterston et al. (1992) from a partially normalized cDNA library (clones cm12a11 and cm01b8). C6 does not show similarity to any sequences in the peptide databases searched.

We designed oligonucleotide primer pairs specific for a region near the 5′ end of each element and used PCR to determine if any of these four coding elements are deleted in the viable Unc-22 strains.

Identification of unc-22 *deficiencies using PCR*

If one or more of the coding elements immediately to the left (3′) of *unc-22* encode nonessential genes, then an individual homozygous for a deficiency that deletes *unc-22* and

only these nonessential genes, would likely be viable and phenotypically Unc-22. For this reason, the homozygous viable Unc-22 mutations were analyzed at the molecular level.

The first pair of primers tested were those specific for the 5′ region of the glucose transporter-like gene. A deficiency of *unc-22* that extends 3′ of *unc-22* and deletes the glucose transporter-like gene and one or both of the priming sites would be identified by failure of DNA from an individual homozygous for the deficiency to produce a PCR product with this primer pair. In this manner, deficiencies that break 3′ of *unc-22* could be identified. We tested 17 viable Unc-22 strains isolated in this study, 31 viable Unc-22 strains isolated previously in our laboratory (see Materials and methods), and *sDf83* homozygous larvae with the glucose transporter-like gene primers. If *sDf83* is in fact a deficiency of *unc-22* that extends as far left as *let-56*, then individuals homozygous for *sDf83* would not be expected to produce a PCR product with the glucose transporter-like primers. A primer pair specific for sequences on chromosome V (provided by K.S. McKim) was used as an internal control in the PCR reactions. As predicted, *sDf83* homozygotes failed to produce the expected PCR product of 510 bp with the glucose transporter-like gene primers (data not shown). In addition, one of the viable Unc-22 strains isolated in this study also failed to produce the expected product (Fig. 4A), indicating that this strain is homozygous for a deficiency of *unc-22* that deletes both *unc-22* and the glucose transporter-like gene. This deficiency has been designated *sDf88*.

Characterization of sDf83 *and* sDf88

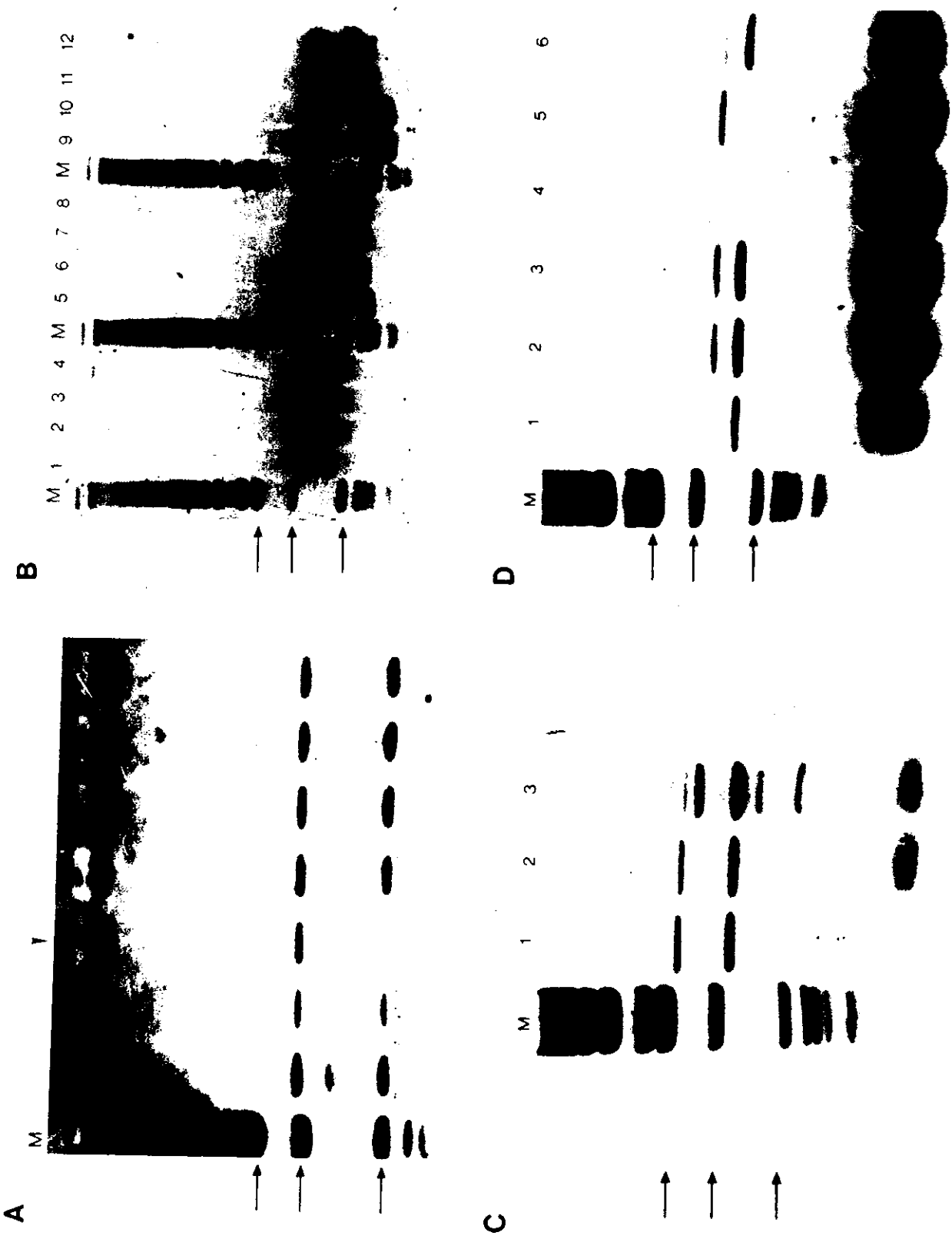
To determine the extent of the deficiencies *sDf83* and *sDf88*, we tested them using PCR with primer pairs specific for the proline transporter-like gene, Dup and C6. As expected, DNA from individuals homozygous for *sDf83* failed to produce PCR products with all of these primer pairs (Fig. 4B, lanes 9–12, and data not shown), indicating that, in addition to the glucose transporter-like gene, the proline transporter-like gene, Dup and C6 are also deleted by *sDf83*. DNA from individuals homozygous for *sDf88* failed to produce the expected product with the primer pair for the proline transporter-like gene (565 bp) (Fig. 4B, lanes 5–8) but did produce the expected products with the Dup primers (680 bp) (Fig. 4B, lanes 1–4) and C6 primers (513 bp) (data not shown), indicating that the left breakpoint of *sDf88* lies between the 5′ end of the proline transporter-like gene and the 5′ end of Dup (Fig. 2). *sDf88* therefore deletes two of the four coding elements immediately 3′ of *unc-22*: the glucose transporter-like gene and the proline transporter-like gene. *sDf88* homozygous individuals are viable and fertile, suggesting that these two coding elements are not essential for development under laboratory conditions, either singly or in combination with each other.

To map the left breakpoint of *sDf83* more precisely, we required primers specific for the *let-56* region. Although the *let-56* gene has not been cloned, it has been mapped to cosmid C11F2 by injection rescue (Clark and Baillie 1992) and localized to an approximately 10-kb region at the right end of the cosmid (Marra et al. 1993). Two genes have been identified and sequenced on C11F2. One of these has sequence similarity to a human growth factor activatable Na^+/H^+ antiporter gene and the other has sequence similarity to DOPA decarboxylase genes from a number of different organisms, including human, rat, and *Drosophila* (Marra et al. 1993). The location of these genes on cosmid C11F2 is shown in Fig. 2. Since *sDf83* fails to complement *let-56*, it must extend into the region of the genome represented in cosmid C11F2 and may affect one or both of the DOPA decarboxylase-like and Na^+/H^+ antiporter-like genes. We tested this using PCR with primer pairs specific for each of these genes. The expected PCR products were produced with both the Na^+/H^+ antiporter-like gene primers (data not shown) and the DOPA decarboxylase-like gene primers (1.4 kb) (Fig. 4C) with *sDf83* homozygotes. Thus it appears that neither the DOPA decarboxylase-like gene nor the Na^+/H^+ antiporter-like gene are deleted by *sDf83*. The left breakpoint of *sDf83* must therefore lie between the site of the DOPA decarboxylase-like gene primers used and the site corresponding to the right end of cosmid C11F2, a distance of approximately 10 kb (Fig. 2). Since *sDf83* fails to complement *let-56*, the *let-56* gene must also lie in this 10-kb interval. This finding supports results obtained with another *unc-22* deficiency, *sDf65* (Clark and Baillie 1992), which also breaks between *let-56* and *let-653*. *sDf65* also produces the expected PCR product with DOPA decarboxylase-like gene specific primers (Marra et al. 1993).

To determine if either *sDf83* or *sDf88* extend to the right of *unc-22*, we used a primer pair (KT01 and KT03) specific for a region approximately 2 kb upstream of the *unc-22* gene. When tested with this primer pair, *sDf88* homozygous individuals produced the expected PCR product of 674 bp (data not shown), while *sDf83* homozygous individuals did not (Fig. 4D). Since *sDf88* does not delete these primer sites, the right breakpoint of *sDf88* either lies within the *unc-22* gene or between the 5′ end of *unc-22* and these primer sites. *sDf83* deletes one or both of these upstream primer sites and therefore extends at least 2 kb 5′ of the *unc-22* gene but, as a consequence of the complementation results discussed above, does not extend as far as *let-52* (Fig. 2). The observation that the phenotype of *sDf83* homozygous individuals resembles that of *let-56* homozygous individuals suggests that there are no genes between *unc-22* and *let-56* which are required in development before the requirement for *let-56*. To test this, we examined the phenotype of *sDf83* homozygous individuals carrying an extrachromosomal array bearing the rightmost 12 kb of cosmid C11F2 (see Materials and methods). This extrachromosomal array has been shown to rescue individuals homozygous for *let-56(s173)* (S. Jones and M. Marra, unpublished results). *sDf83* homozygotes bearing this array are viable and fertile, demonstrating that the interval between *let-56* and *unc-22* does not contain any genes required in development.

Discussion

This study has focused on a small (approximately 0.05 map unit) interval between *let-56* and *unc-22*, in the *sDf2* region of chromosome IV. We identified four coding elements in the 13 kb of sequence immediately 3′ of *unc-22* and established that at least three of the four are transcribed. We have investigated whether deletions of these elements could be identified by PCR from a set of *unc-22* mutations, reasoning that such deficiencies could be used to determine if any of these coding elements encode genes essential for development. Of a total of 58 *unc-22* mutations examined, two (*sDf83* and *sDf88*) were found to be deficiencies that break in the *let-56–unc-22* region.


A
M
Y
B
M 1 2 3 4 M 5 6 7 8 M 9 10 11 12
C
M 1 2 3
D
M 1 2 3 4 5 6

Characterization of *sDf88* indicates that the two coding elements immediately to the left of *unc-22* (the glucose transporter-like and proline transporter-like genes) are not essential for development under laboratory conditions. However, we cannot rule out the possibility that these elements provide functions required under other conditions. The glucose transporter is most likely a member of a multicopy gene family, as one of the cDNAs sequenced by Waterston et al. (1992) (clone cm06d2) was also identified as a glucose transporter-like gene. A search of ACEDB (a *C. elegans* database, R. Durbin and J. Thierry-Mieg, personal communication) determined that cm06d2 maps to the right of *unc-22* on chromosome IV, between *unc-31* and *unc-30*. Subjecting *sDf88* homozygous individuals to other than normal laboratory conditions (such as alternative food sources or environmental stress) will likely be necessary to determine if the lack of the glucose transporter-like gene results in an observable phenotype. Likewise, although the lack of the L-proline transporter-like gene appears to have no readily discernable effect on *sDf88* homozygous individuals, this is likely due to the fact that under laboratory conditions most neurons are not essential in hermaphrodites (Chalfie and White 1988). Individuals lacking this neurotransmitter transporter may express one or more of a number of more subtle phenotypes associated with nervous system defects, such as uncoordinated movement and chemotactic/thermotactic abnormalities. Such phenotypes may, however, be masked by the Unc-22 phenotype of *sDf88* homozygous individuals.

Our observation that the early adult blocking stage of *sDf83* homozygotes appeared to be the same as that of both *let-56* homozygotes and *sDf83/let-56* trans-heterozygotes prompted us to investigate the phenotype of *sDf83* homozygous individuals carrying an extrachromosomal array that rescues the *let-56* mutation. These animals were viable and fertile. Our interpretation of this data is that there are no essential genes in the *sDf83* interval that are required in development under laboratory conditions. If *sDf83* does in fact delete essential genes other than *let-56*, such genes must be represented more than once in the genome to account for our results.

Comparison of the developmental blocking stage of *sDf83* homozygous individuals with that of individuals homozygous for other *unc-22* deficiencies provides additional information on the essential gene content in the immediate vicinity of *let-56*. Most deficiency homozygotes arrest development at the time when the earliest acting gene that is deleted is required. There is another small deficiency in this region, *sDf19* (Rogalski et al. 1982), which has a dominant Unc-22 phenotype and is lethal when homozygous. The right breakpoint of *sDf19* was predicted (Rogalski and Baillie 1985) and subsequently shown (Benian et al. 1989) to lie in the 3′ end of the *unc-22* gene, and the left breakpoint lies between *let-56* and *let-653* (Rogalski and Baillie 1985) (Fig. 1). *sDf19* has been shown by PCR to delete both the DOPA decarboxylase-like gene and the Na^+/H^+ antiporter-like gene (Marra et al. 1993). *sDf19* homozygotes usually arrest as eggs, but occasionally hatch prior to arrest. The arrest stage of *sDf19* homozygotes is thus earlier than either *let-56* homozygotes or *sDf83* homozygotes. This suggests that there are one or more essential genes that lie between *let-56* and *let-653*, which are deleted by *sDf19* but not *sDf83*. The Na^+/H^+ antiporter-like gene and the DOPA decarboxylase-like gene may therefore be essential genes that have yet to be identified genetically.

In summary, we have found PCR to be a useful tool for the detection of deletions of coding elements in the *let-56–unc-22* region. Using this method, we identified two *unc-22* deficiencies that also delete coding elements to the left of the *unc-22* gene. Characterization of *sDf83* and *sDf88* has shown that there are no genes between *let-56* and *unc-22* that are required in development under normal laboratory conditions. In addition, characterization of *sDf83* has confirmed the placement of *let-56* by Marra et al. (1993) to within 10 kb of the right end of cosmid C11F2, demonstrating that *let-56* cannot be the Na^+/H^- antiporter-like gene. Comparison of the blocking stages of *sDf83* and *sDf19*, which share the same left breakpoint position on the genetic map (between

FIG. 4. PCR analysis of *sDf83* and *sDf88*. PCR products were electrophoresed on 1% agarose gels and visualized with ethidium bromide. Lanes containing DNA size markers are labeled M. Arrows indicate marker bands of sizes 1.6 kb, 1.0 kb, and 500 bp, from top to bottom, respectively. (A) Products of PCR reactions using the glucose transporter-like gene primers on DNA from viable Unc-22 strains isolated in this study. The lane marked with an arrowhead indicates the PCR reaction performed with DNA from individuals homozygous for *sDf88*. The band migrating at approximately 1 kb in all reactions corresponds to the product produced by control primers specific for chromosome V. The band migrating at 510 bp in all lanes except the *sDf88* lane corresponds to the product produced by the glucose transporter-like gene primers. (B) Products of PCR reactions using primers for the Dup gene (lanes 1–4), proline transporter-like gene (lanes 5–8), and C6 gene (lanes 9–12). Reactions were run in duplicate. Reactions in lanes 3, 4, 7, 8, 11, and 12 were performed with N2 (wildtype) DNA. Reactions in lanes 1, 2, 5, and 6 were performed with DNA from individuals homozygous for *sDf88*. Reactions in lanes 9 and 10 were performed with DNA from larvae homozygous for *sDf83*. Lane 10 is a failed reaction.The upper, approximately 900-bp band in all reactions corresponds to the product produced by the chromosome V control primers. In lanes 1–4, the band migrating at 680 bp corresponds to the product produced by the Dup gene primers. In lanes 7 and 8, the band migrating at 565 bp corresponds to the product produced by the proline transporter-like gene primers. Note that this product is missing in lanes 5 and 6. The band migrating at 513 bp in lanes 11 and 12 corresponds to the product produced by the C6 gene primers. Note that this 513-bp product is missing in lane 9. (C) Products of PCR reactions using the DOPA decarboxylase-like gene primers. Reaction in lane 1 was performed with N2 DNA. Reaction in lane 2 was performed with DNA from individuals homozygous for *sDf88*. Reaction in lane 3 was performed with DNA from individuals homozygous for *sDf83*. The band migrating at approximately 1.4 kb in all lanes corresponds to the product produced by the DOPA decarboxylase-like gene primers. The band migrating at approximately 900 bp in all lanes corresponds to the product produced by chromosome V control primers. (D) Products of PCR reactions using primers KT01 and KT03, specific for a region approximately 2 kb upstream of the *unc-22* gene. Reactions in lanes 1–3 were performed with DNA from viable Unc-22 strains isolated in this study. Lane 4 is a failed reaction. Reaction in lane 5 was performed with DNA from *sDf83* homozygous individuals. Reaction in lane 6 was performed with N2 DNA. The band migrating at approximately 900 bp corresponds to the product produced by chromosome V control primers. The band migrating at 674 bp corresponds to the product produced by primers KT01 and KT03. Note that the 674-bp band is absent in lane 5.

let-56 and *let-653*) but differ in their molecular breakpoints, has also shown that there is a gene required early in development which lies between *let-56* and *let-653*. This approach to the detection of deletions of coding elements may also be applicable to other regions of the genome for which there is sequence data available and an adequate system for the induction and selection of deficiencies. With the sequence data available from the *C. elegans* genome sequencing project, which is now well underway (Sulston et al. 1992), this method should provide a tool with which to assay for essential gene content in sequenced regions of the genome.

Acknowledgments

The authors thank Brad Barbazuk for help with the mutagenesis screen, Steven Jones for help with the database searches, and M.R. Rogalski for technical assistance. We would also like to thank Raja Rosenbluth, Ann Rose, and Sharon Gorski for their helpful discussions throughout this work. This research was supported in part by a Natural Sciences and Engineering Research Council of Canada operating grant to D.L.B. The work of D.L.B., J.E.S., and M.A.M. was supported in part by Department of Health and Human Sciences grant No. RR07127. This work was also supported by a National Institutes of Health grant to G.M.B., a U.S. Department of Energy grant to C.F., and a Natural Sciences and Engineering Research Council of Canada Postgraduate Scholarship to M.A.M. G.M.B. is an Established Investigator for the American Heart Association.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. J. Mol. Biol. **215**: 403–410.

Barstead, R.J., and Waterston, R.H. 1989. The basal component of the nematode dense-body is vinculin. J. Biol. Chem. **264**: 10177–10185.

Barstead, R.J., and Waterston, R.H. 1991. Vinculin is essential for muscle function in the nematode. J. Cell Biol. **114**: 715–724.

Benian, G.M., Kiff, J.E., Neckelmann, N., Moerman, D.G., and Waterston, R.H. 1989. Sequence of an unusually large protein implicated in regulation of myosin activity in *C. elegans*. Nature (London), **342**: 45–50.

Benian, G.M., L'Hernault, S.W., and Morris, M.E. 1993. Additional sequence complexity in the muscle gene, *unc-22*, and it's encoded protein, twitchin, of *Caenorhabditis elegans*. Genetics, **134**: 1097–1104.

Blumenthal, T., and Thomas, J. 1988. Cis and trans mRNA splicing in *C. elegans*. Trends Genet. **4**: 305–308.

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

Chalfie, M., and White, J. 1988. The nervous system. *In* The nematode *Caenorhabditis elegans*. *Edited by* W.B. Wood. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. pp. 337–391.

Clark, D.V. 1990. The *unc-22(IV)* region of *Caenorhabditis elegans*: genetic analysis and molecular mapping. Ph.D. thesis, Simon Fraser University, Burnaby, B.C., Canada.

Clark, D.V., and Baillie, D.L. 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, D.V., Rogalski, T.M., Donati, L.M., and Baillie, D.L. 1988. The *unc-22(IV)* region of *Caenorhabditis elegans*: genetic analysis of lethal mutations. Genetics, **119**: 345–353.

Fremeau, R.T. Jr., Caron, M.G., and Blakely, R.D. 1992. Molecular cloning and expression of a high affinity L-proline transporter expressed in putative glutamatergic pathways of rat brain. Neuron, **8**: 915–926.

Fukumoto, H., Kayano, T., Buse, J.B., Edwards, Y., Pilch, P.F., Bell, G.I., and Seino, S. 1989. Cloning and characterization of the major insulin-responsive glucose transporter expressed in human skeletal muscle and other insulin-responsive tissues. J. Biol. Chem. **264**: 7776–7779.

Horvitz, H.R., Brenner, S., Hodgkin, J., and Herman, R.K. 1979. A uniform genetic nomenclature for the nematode *Caenorhabditis elegans*. Mol. Gen. Genet. **175**: 129–133.

Johnsen, R.C., and Baillie, D.L. 1988. Formaldehyde mutagenesis of the *eT1* balanced region in *Caenorhabditis elegans*: dose response curve and the analysis of mutational events. Mutat. Res. **201**: 137–147.

Kramer, J.M., French, R.P., Park, E.C., and Johnson, J.J. 1990. The *Caenorhabditis elegans rol-6* gene, which interacts with the *sqt-1* collagen gene to determine organismal morphology, encodes a collagen. Mol. Cell. Biol. **10**: 2081–2089.

Liu, Q.-R., Mandiyan, S., Nelson, H., and Nelson, N. 1992. A family of genes encoding neurotransmitter transporters. Proc. Natl. Acad. Sci. U.S.A. **89**: 6639–6643.

Marra, M.A., Prasad, S.S., and Baillie, D.L. 1993. Molecular analysis of two genes between *let-653* and *let-56* in the *unc-22(IV)* region of *C. elegans*. Mol. Gen. Genet. **236**: 289–298.

Moerman, D.G., and Baillie, D.L. 1979. Genetic organization in *Caenorhabditis elegans*: fine structure analysis of the *unc-22* gene. Genetics, **91**: 95–103.

Moerman, D.G., and Baillie, D.L. 1981. Formaldehyde mutagenesis in the nematode *Caenorhabditis elegans*. Mutat. Res. **80**: 273–279.

Pearson, W.R., and Lipman, D.J. 1988. Improved tools for biological sequence analysis. Proc. Natl. Acad. Sci. U.S.A. **85**: 2444–2448.

Prasad, S.S., and Baillie, D.L. 1989. Evolutionarily conserved coding sequences in the *dpy-20–unc-22* region of *Caenorhabditis elegans*. Genomics, **5**: 185–198.

Rogalski, T.M., and Baillie, D.L. 1985. Genetic organization of the *unc-22 IV* gene and the adjacent region in *Caenorhabditis elegans*. Mol. Gen. Genet. **201**: 409–414.

Rogalski, T.M., Moerman, D.G., and Baillie, D.L. 1982. Essential genes and deficiencies in the *unc-22 IV* region of *Caenorhabditis elegans*. Genetics, **102**: 725–736.

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

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science (Washington, D.C.), **239**: 487–491.

Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sulston, J., Du, Z., Thomas, K., Wilson, R., Hillier, L., Staden, R., Halloran, N., Green, P., Thierry-Mieg, J., Qiu, L., Dear, S., Coulson, A., Craxton, M., Durbin, R., Berks, M., Metzstein, M., Hawkins, T., Ainscough, R., and Waterston, R. 1992. The *C. elegans* genome sequencing project: a beginning. Nature (London), **356**: 37–41.

Waterston, R., Martin, C., Craxton, M., Huynh, C., Coulson, A., Hillier, L., Durbin, R., Green, P., Shownkeen, R., Halloran, N., Metzstein, M., Hawkins, T., Wilson, R., Berks, M., Du, Z., Thomas, K., Thierry-Mieg, J., and Sulston, J. 1992. A survey of expressed genes in *Caenorhabditis elegans*. Nature Genet. **1**: 114–123.