Molecular cloning of *lin-29*, a heterochronic gene required for the differentiation of hypodermal cells and the cessation of molting in *C.elegans*

A.Papp+, A.E.Rougvie and V.Ambros*
Department of Cellular and Developmental Biology, Harvard University Biological Laboratories, 16 Divinity Ave, Cambridge, MA 02138, USA

Received August 24, 1990; Revised and Accepted November 30, 1990

ABSTRACT

The *lin-29* gene product of *C.elegans* activates a temporal developmental switch for hypodermal cells. Loss-of-function *lin-29* mutations result in worms that fail to execute a stage-specific pattern of hypodermal differentiation that includes exist from the cell cycle, repression of larval cuticle genes, activation of adult cuticle genes, and the cessation of molting. Combined genetic and physical mapping of restriction fragment length polymorphisms (RFLPs) was used to identify the *lin-29* locus. A probe from the insertion site of a Tc1 (*maP1*), closely linked and to the left of *lin-29* on the genetic map, was used to identify a large set of overlapping cosmid, lambda and yeast artificial chromosome (YAC) clones assembled as part of the *C.elegans* physical mapping project. Radiolabeled DNA from one YAC clone identified two distinct allele-specific alterations that cosegregated with the *lin-29* mutant phenotype in *lin-29* intragenic recombinants. *lin-29* sequences were severely under-represented in all cosmid and lambda libraries tested, but were readily cloned in a YAC vector, suggesting that the *lin-29* region contains sequences incompatible with standard prokaryotic cloning techniques.

INTRODUCTION

Genes that control the timing of cellular division and differentiation in *C.elegans* have been identified by mutations that cause alterations in the stage-specificity of developmental events (1). Four of these 'heterochronic' genes, *lin-4, lin-14, lin-28* and *lin-29* have been organized into a phenotypic hierarchy based on their relative pleiotropies (1) and into a regulatory hierarchy based on their epistasis relationships (2). *lin-29*, the most downstream of these genes, is required for lateral hypodermal cells to execute a particular larval to adult switch (L/A switch) in cell fate at the final molt (2). The timing of *lin-29* activation, and hence the timing of the L/A switch, is proposed to be regulated by the upstream genes, *lin-4, lin-14* and *lin-28*.

The L/A switch involves the terminal differentiation of the lateral hypodermal ('seam') cells, which contribute to the synthesis of the worm's cuticle (5). During wild-type *C.elegans* development, the seam cells divide concomitantly with the first three larval molts (L1, L2, and L3) in a stem cell-like lineage (6). At each division, the anterior daughter joins a hypodermal syncytium and synthesizes the next larval cuticle, while the posterior daughter remains a stem cell, with the potential to divide at the next molt. However, just prior to the final (L4) molt, seam cells cease cell division and fuse together to form a new syncytium and synthesize a morphologically and biochemically distinct adult cuticle (6, 7). Thus, the L/A switch involves a diverse set of cellular differentiation processes, including cell cycle exit, cell fusion, stage-specific changes in cuticle gene expression and morphogenesis.

Animals homozygous for *lin-29* loss-of-function mutations fail to execute the L/A switch. Although the development of all non-hypodermal tissues appears to proceed normally, animals lacking *lin-29* function undergo supernumerary larval stages. At the L4 molt, and at subsequent supernumerary molts, the seam cells divide, synthesize larval cuticle rather than adult cuticle, and do not fuse together. The *lin-29* gene product(s) may regulate the coordinated expression of processes associated with the larval to adult switch, including the control of cell division, cell fusion and the expression of stage-specific cuticle genes, including collagens. In as much as *lin-29* activity is required by cells for a switch from the proliferative to terminally differentiated state, *lin-29* can be included in the general class of tumor suppressor genes (8). To initiate molecular studies of how *lin-29* regulates these various aspects of the terminal differentiation of seam cells, and how *lin-29* activity is regulated temporally and anatomically, we have cloned the *lin-29* locus.

Since we could not presuppose any particular biochemical properties for products of *lin-29*, the gene was cloned solely on the basis of its genetic properties. Furthermore, since no transposon-induced alleles of *lin-29* were identified in screens that would have been expected to recover *lin-29* mutations (S. Kim, personal communication; A. Papp and V. Ambros,

* To whom correspondence should be addressed
+Present address: 12012 Goshen Avenue, Suite 103, Brentwood, CA 90049, USA

unpublished), we elected to use techniques similar to those used for cloning *Drosophila* homeotic genes and loci corresponding to human genetic diseases: the identification of linked Restriction Fragment Length Polymorphisms (RFLPs) combined with chromosomal walking (9, 10). Parallel RFLP mapping of strain-specific insertions of the transposon Tc1 (11) was used to identify and genetically map RFLPs linked to *lin-29*. A probe from the insertion site of a Tc1 to the left of *lin-29* on the genetic map was used to initiate a chromosome walk that identified and extended a large set of overlapping cloned sequences (a 'contig') assembled as part of the *C.elegans* physical mapping project (12, 13). Sequences cloned in lambda, cosmid and Yeast Artificial Chromosome (YAC) vectors (14) were used to search for DNA alterations associated with *lin-29* mutations. Some sequences within the contig were represented only in YAC libraries. Radiolabelled DNA from one such YAC clone and a lambda clone from this walk detected two distinct allele-specific aberrations of a 7.3 kb *Eco*RI fragment in genomic DNA of two of the five different *lin-29* alleles. These two DNA aberrations were unambiguously mapped to the *lin-29* locus by examining their segregation in *lin-29* intragenic recombinants, indicating that this 7.3 kb fragment contains at least part of *lin-29*. Our observation that sequences in the vicinity of *lin-29* were severely under-represented in cosmid and lambda libraries suggests that this region contains sequences incompatible with standard prokaryotic cloning techniques.

MATERIALS AND METHODS

Strains and alleles

A strain (VT455) congenic to N2 with Bergerac sequences in the *lin-29* region of LGII was a gift from Joan Park and H. Robert Horvitz. *C.elegans* var. Bergerac, strain RW7000, was obtained from the *C.elegans* Genetics Stock Center. All mutants used in this study were derived from *C.elegans* var. Bristol, strain N2, and were obtained from the Cambridge collection (15) unless otherwise noted. *mnCI* is a balancer for the *lin-29* region of LGII which carries recessive mutations *dpy-10(e128) unc-52(e444)* (16). *lin-29* alleles are described by Ambros and Horvitz (1), except for *n1368* and *n1440*, which were induced by Ethylmethanesulfonate (EMS) and gamma-rays, respectively, and provided by H.R.Horvitz.

Cosmid and lambda libraries

pHC79 (17) and pJB8 (12, 18) cosmid libraries in host HB101, each containing approximately 10 genome-equivalents of cloned DNA, were analysed here. Of these two libraries, only the pJB8 library yielded clones (e.g., C25F2, C32C11) near the *lin-29* locus. A lorist 2 (19) library yielded the *lin-29*-proximal cosmid clones F28H4, T15G9 and W03C9. Phage libraries of wild type DNA were in vector lambda-2001 (20) and were plated on rec+ hosts Q358, Q359 or K802 (21) or the rec$^-$host CES200 (*hsdR recB21 recC22 sbcB15*) (22).

Parallel RFLP mapping

Parallel mapping of multiple strain-specific Tc1 insertion sites ('Parallel RFLP mapping') was performed as described elsewhere (11). Hermaphrodites of strain VT455 (chromosomes denoted *lin-29*[Berg +]) or *C.elegans* var. Bergerac (strain RW7000) were mated with *sqt-1 (sc13) lin-29(n333)/mnCI* males and separately with *lin-29(n546) rol-1(e91)/mnCI* males. From these crosses, hermaphrodites of genotypes, *sqt-1(sc13) lin-29(n333)/lin-29*[Berg +] and *lin-29(n546) rol-1(e91)/lin-29*[Berg +] were obtained and cultured individually. Sqt-non-Lin and Rol-non-Lin recombinants were isolated from among the self progeny of these hermaphrodites, and homozygous recombinant strains were derived by cloning self-progeny in the subsequent generation.

Genomic DNA from strains homozygous for recombinant chromosomes or parental chromosomes was prepared by the proteinase K method (23), digested with restriction enzymes, separated by agarose gel electrophoresis, transferred to filters, and hybridized with a cloned Tc1 probe as described below.

Transfers and hybridizations

After agarose gel electrophoresis, DNA was transferred to Zetabind (AMF Cuno, Meriden, CT) or Nytran (Schleicher & Schuell) as per manufacturer's protocols.

Cloning sequences containing a Bergerac-specific Tc1 element linked to *lin-29*

A 17 kb *Eco*RI/*Bam*HI fragment containing the *maP1* Bergerac-specific Tc1 insertion was isolated from an EMBL4 library of size-selected, *Eco*RI/*Bam*HI digested VT455 DNA. To obtain a substantial quantity of ligatable fragments, *Eco*RI/*Bam*HI digested VT455 DNA was size separated on a 10−40% sucrose gradient (21). DNA from the highest molecular weight fractions was diluted 1:1 with H_2O and ethanol precipitated. This DNA was combined at an equimolar ratio with a mixture of *Eco*RI digested EMBL4 DNA and *Bam*HI digested EMBL4 DNA. Ligation was carried out at ~1 mg/ml, overnight, at 20°C. The ligation was packaged in vitro (Gigapack, Stratagene) and amplified on *E.coli* strain Q359. The resulting phage stock was plated on K802, a non-selective host that gives larger and faster growing plaques than Q359. Plaques were transferred to Nitrocellulose (Schleicher & Schuell) 82 mm circular filters (21), and hybridized with a Tc1-containing plasmid probe (pCeT2, obtained from H.R. Horvitz, originally constructed by S. Emmons). Two positive plaques were picked from the ~9600 that were plated, and both contained the desired 17 kb *Eco*RI/*Bam*HI fragment. This fragment was excised and recloned into pUC18 to yield clone pAPP-HB4.

Cloning the ends of inserts in Yeast Artificial Chromosomes

Yeast artificial chromosome clones containing large segments of *C.elegans* genomic DNA were maintained as described elsewhere (14). YAC clones are linear, with vector sequences and yeast telomeres on each end of the insert DNA. To identify sequences at the ends of the *C.elegans* DNA in a YAC clone, we employed two procedures, each designed to recover a small segment of insert DNA immediately adjacent to vector sequences.

To recover the end of the YAC clone Y17F2 containing the bacterial origin or replication and the gene for ampicillin resistance (14), genomic DNA from the YAC-containing yeast strain was digested to completion with *Xho*I. This generated a fragment extending from the unique *Xho*I site between the telomere and amp-ori, to the first *Xho*I site in the insert. The enzyme was then heat-inactivated at 70°C for 10 minutes after addition of an equal volume of 20 mM EDTA, and the DNA was diluted to <50 μg/ml in ligation buffer, ligated, transformed into *E.coli* and ampicillin-resistant colonies were selected.

To clone the end of Y17F2 adjacent to the other telomere, yeast genomic DNA from a strain carrying Y17F2 was cut separately with the restriction enzymes *Hin*dIII, *Xho*I, *Nsi*I, *Eco*RV, *Pst*I, and *Sal*I and Southern blots were probed with vector sequences from this end of pYAC4 (eg. *Sal*I/*Eco*RI fragment). A 6.4 kb *Xho*I restriction fragment detected in this way was chosen for

cloning, since it was expected to contain 2.9 kb of vector sequence and the adjacent 3.5 kb of insert DNA. Genomic DNA from a yeast strain containing Y17F2 was digested with *Xho*I, ligated to *Xho*I-digested lambda 2001, packaged, and the resulting phage library screened with the *Sal*I/*Eco*RI fragment of pYAC4. A clone containing the desired end fragment was thus identified and confirmed by restriction enzyme analysis. This clone contained the expected 3.5 kb *Eco*RI/*Xho*I fragment of *C.elegans* DNA and it was subcloned into pBluescript (Stratagene).

Labeling YACs for use as probes

Hybridization probes were made from entire Yeast Artificial Chromosomes by electrophoretically separating the yeast chromosomes as previously described (25), and electrophoresing the YAC into a trough filled with 1% low melting-point agarose. The agarose was then melted at 65°C, 8 μl diluted 1:1 with H_2O and labeled using a Random Primed DNA Labeling Kit (Boehringer Mannheim) with 3000 Ci/mmol [α^{32}P]dATP (NEN), in a 30 μl reaction using kit instructions.

Identifying lambda clones of sequences linked to *lin-29*

A *lin-29(n836)* genomic *Eco*RI library was constructed in Lambda-Zap II, packaged with Gigapack, as per Stratagene's protocols, but plated on *E.coli* strain CES200 *(hsdR recB21 recC22 sbcB15)* (22). Plaques were screened with partially purified YAC clone Y11E9 DNA probe (see above), and positive clones were analyzed by restriction digests to identify those containing the 9.2 kb *Eco*RI fragment associated with the *n836* mutation (see Results section). To identify overlapping phage clones of DNA contained within YAC clones, genomic libraries of partially Sau3A-digested N2 DNA were constructed in lambda 2001, plated on *E.coli* strain K802 or CES200, and plaque lifts were probed with gel-isolated YAC DNA.

lin-29 intragenic recombinants

Hermaphrodites of the genotype *unc-4(e120) lin-29(n1440)/ sqt-1*(sc13) *lin-29(n836)* or *unc-4(e120) lin-29(n1440)/ lin-29(n836)*, were placed individually on plates and their F1 and F2 progeny were screened for wild type recombinants. The identification of wild type recombinants was facilitated by the fact that *lin-29(n836/n1440)* hermaphrodites display very highly penetrant morphological and egg-laying defective phenotypes and thus rare *lin-29(n836/+)* or *lin-29(n1440/+)* recombinant animals were easily identified among *lin-29/lin-29* mutants by their wild type morphology and normal deposition of eggs on the surface of agar plates. Strains homozygous for recombinant chromosomes were derived by cloning self-progeny from the subsequent generation.

RESULTS

Identification and cloning of a Tc1-containing RFLP, *maP1*, closely linked to *lin-29*

The method of parallel RFLP mapping (11) was used to identify and genetically map polymorphisms caused by strain-specific Tc1 insertions. This method exploits the fact that the *lin-29* mutants were isolated in *C.elegans* var. Bristol which has about 30 stable copies of the Tc1 transposon, while another strain, var. Bergerac, has about 300 Tc1 copies (26). Thus, given the genome size of approximately 100,000 kb (13), and assuming a random distribution of Tc1 elements, Tc1 insertions unique to the Bergerac strain are expected to lie within approximately 100 kb of any given point in the genome.

To identify Tc1 elements linked to *lin-29* we used a strain (VT455) congenic to Bristol for all linkage groups, except with Bergerac sequences in the *lin-29* region (J. Park, H.R. Horvitz, pers. comm.). When DNA from VT455 was digested simultaneously with *Eco*RI and *Bam*HI (neither of which cuts

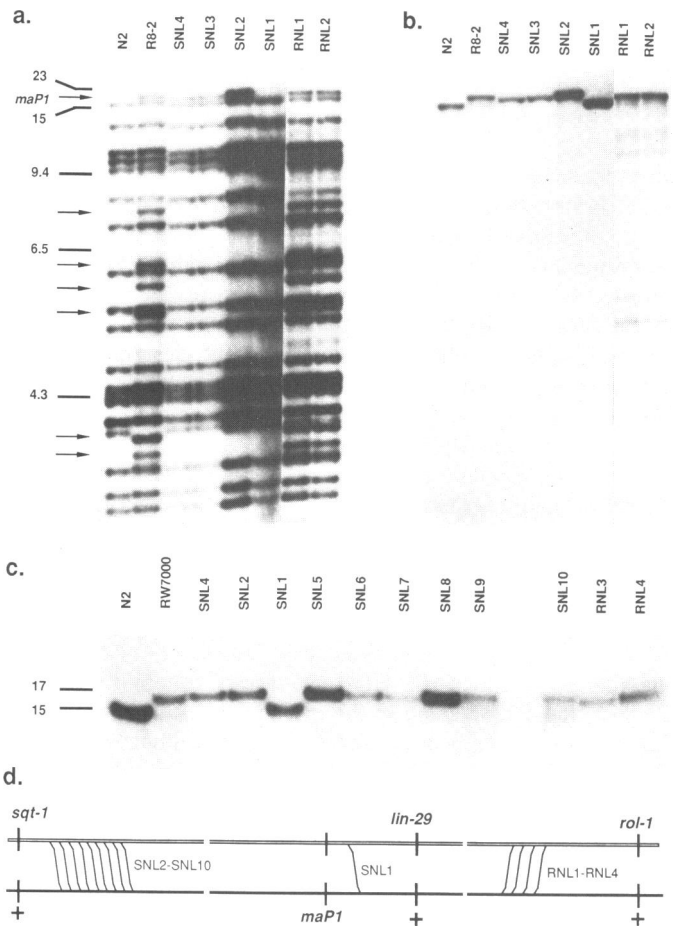

Figure 1. a) Parallel RFLP mapping of Bergerac Tc1 insertions linked to *lin-29*. Genomic DNAs were digested with *Eco*RI + *Bam*HI and hybridized with a Tc1 probe. DNA from *lin-29*[Berg +] strain VT455 contains at least 7 extra Tc1-containing fragments (indicated by arrows) compared to Bristol N2. One 17 kb Bergerac-specific Tc1-hybridizing band, designated *maP1*, is present in Sqt-non-Lin recombinants (SNL) while the other six extra bands are absent. Thus, these other Tc1 elements map closely linked to or left of *sqt-1* and were not of further interest. A single recombinant, SNL1, retains the Bristol allele of *maP1*, as evidenced by absence of the 17 kb hybridization, and two Rol-non-Lin recombinants, RNL1 and RNL2, retain the Bergerac allele of *maP1*. Thus, *maP1* maps between *sqt-1* and *lin-29*, relatively closely linked to *lin-29*. Size standards are in kilobases. b) The 17 kb *Eco*RI/*Bam*HI fragment containing *maP1* was cloned as described in Materials and Methods. Sequences flanking the Tc1 insertion were used to probe the same filter shown in panel 1a after removal of the previous probe. Hybridization is to the 17 kb Bergerac allele of *maP1* (a 17 kb fragment in VT455 and recombinants that retained Bergerac *maP1*, see above) or an approximately 15.4 kb Bristol allele of *maP1* (in N2 and SNL1; compare with panel 1a). c) The recombinants used in parts a and b and additional recombinants, isolated as described in Materials and Methods, were digested with *Eco*RI and *Bam*HI, blotted and probed as in part 1b. The SNL1−SNL4 recombinants were selected between Bristol and VT455 chromosomes in the *sqt-1−lin-29* interval; the SNL5−SNL10 recombinants were selected between Bristol and Bergerac RW7000 chromosomes in the same interval. The RNL1−RNL4 recombinants were selected between Bristol and VT455 chromosomes in the *rol-1−lin-29* interval. d) Partial genetic map of the *lin-29* region of LGII showing the positions of *sqt-1, maP1, lin-29* and *rol-1*. Shown schematically are the relative positions of the recombination events generating the recombinant chromosomes analysed in parts a, b and c. Assuming that the recombination events were evenly distributed between *sqt-1* and *lin-29, maP1* maps approximately 0.2% left of *lin-29*.

within Tc1 elements) and then analyzed by Southern hybridization using Tc1 sequences as a probe, 7 or 8 extra Tc1-containing *Eco*RI/*Bam*HI restriction fragments were observed (Fig. 1a, arrows). These extra bands represent Bergerac-specific Tc1 insertions that persisted through the backcrossing procedure used to generate the VT455 congenic strain. Since VT455 resulted from backcrosses designed to retain Bergerac sequences linked to the genetic region containing *lin-29*, these extra Tc1-containing fragments were likely to be genetically linked to *lin-29*. To identify which of these Tc1 elements were closest to *lin-29*, Bergerac-specific fragments were genetically mapped by selecting for recombination between the VT455 chromosome and Bristol chromosomes carrying the flanking marker *sqt-1*, which maps to the left of *lin-29*, or *rol-1*, which maps to the right (Fig. 1d). Three of four recombinants in the *sqt-1 − lin-29* interval and two recombinants in the *rol-1 − lin-29* interval, all of which were selected as retaining the Bergerac allele of *lin-29*, retained a 17 kb Bergerac-specific Tc1-containing fragment (Fig. 1a, '*maP1*'). Thus, this polymorphic restriction fragment, designated *maP1*, appeared to lie between *sqt-1* and *lin-29*. All of the other Bergerac-specific Tc1 elements in VT455 mapped left of *maP1* and were less closely linked to *lin-29*.

To obtain genomic sequences adjacent to *maP1*, the 17 kb *Eco*RI/*Bam*HI fragment containing the *maP1* Bergerac Tc1 was cloned from a library of VT455 genomic DNA digested with *Eco*RI and *Bam*HI (see Materials and Methods section). Clones containing the 17 kb *Eco*RI/*Bam*HI fragment were identified by hybridization to a Tc1 probe and restriction analysis. The *Eco*RI/*Bam*HI insert fragment was subcloned into pUC18, yielding the clone pAPP-HB4. *C.elegans* genomic sequences flanking the Tc1 element in this plasmid hybridized to a 17 kb *Eco*RI/*Bam*HI fragment in VT455 and in recombinants carrying *maP1*, and to a fragment 1.6 kb smaller (the size of a Tc1) in N2 and in recombinants lacking *maP1* (compare Fig. 1a and 1b). Thus, the cloned fragment was confirmed to contain *maP1*.

To further map the location of *maP1* with respect to *lin-29* sequences, additional recombinants were isolated by selecting for recombination between a Bristol chromosome and a chromosome derived from Bergerac strain RW7000. Recombinants were isolated between *lin-29* and either *sqt-1* or *rol-1* as described above and in Materials and Methods. These recombinants were selected such that they retained the Bergerac allele of *lin-29*$^+$. Sequences flanking the *maP1* Tc1 were used to probe Southern blots of genomic DNA isolated from Bristol, Bergerac (RW7000) and the recombinant strains. Six of six Sqt-nonLin and two of two Rol-nonLin recombinants contained the Bergerac allele of *maP1* (Fig. 1c). The combined recombination data from the experiments described in Figs 1a and 1c, (nine of ten Sqt-nonLin and four of four Rol-nonLin recombinants contained the Bergerac allele of *maP1*), indicate a map position of *maP1* between *lin-29* and *sqt-1*, approximately 0.2% recombination left of *lin- 29* (Fig. 1d).

Identification of a genetically oriented contig containing *maP1*

The pAPP-HB4 clone, containing the *maP1* 17 kb *Eco*RI/*Bam*HI fragment, was fingerprinted by A. Coulson and J. Sulston as part of the *C.elegans* physical mapping project (12) and placed on the far right end of a contig of approximately 9000 kb of overlapping cosmid and YAC clones (Fig. 2). This contig contains several clones corresponding to genetically mapped loci on LGII, including *sqt-1* (27), establishing the orientation of this contig with respect to the genetic map.

Since the *maP1*-containing clone, pAPP-HB4, matched by fingerprint a cosmid (C25F2) at the far right end of an existing contig, and since *lin-29* mapped genetically right of *maP1*, an effort was undertaken to extend the contig rightward from C25F2. However, using short, single-copy, high specific-activity probes from the rightmost end of C25F2, and analyzing more than 20 genome-equivalents of phage and cosmid clones, no significant extensions were found past the C25F2 endpoint, although a large

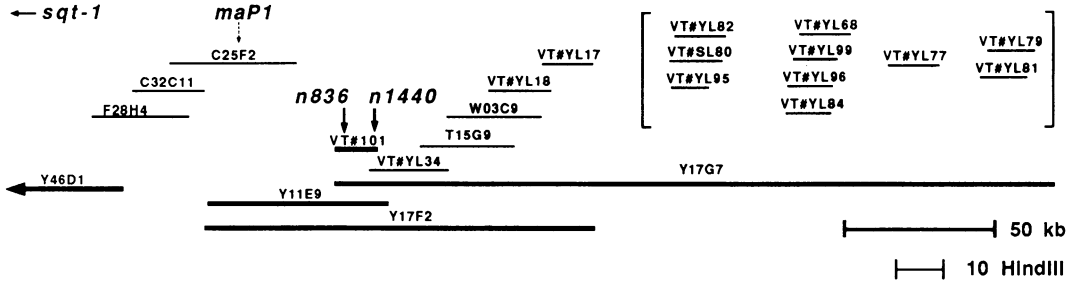

Figure 2. A portion of the physical map of *C.elegans* LGII. The map is drawn in the same left to right orientation as the genetic map; the positions of genetically mapped loci *sqt-1* and *maP1* are indicated. Yeast Artificial Chromosome (YAC) clones are named with a Y prefix. All other clones are cosmid or phage clones. Lambda clones with the prefix 'VT # YL' were from *Sau*3A partial genomic libraries in lambda 2001, donated by A. Coulson. The four small phage contigs indicated by the brackets did not overlap previously isolated clones other than Y17G7 and thus could not be ordered. Overlaps between cosmid clones were identified as part of the *C.elegans* physical mapping project (12, A. Coulson and J. Sulston, pers. comm.), and linkage of certain cosmids to YACs was determined by hybridization, as described elsewhere (13) and in Materials and Methods. Clones indicated by thick lines were mapped with respect to other clones by hybridization, and their lengths in kilobase pairs (kb) were estimated using agarose gel electrophoresis. Clones whose extent of overlap was determined by fingerprinting as part of the *C.elegans* physical mapping project are shown by thin lines, with lengths proportional to number of *Hin*dIII sites (12). Shown here is the right end of a contig of approximately 9,000 kb (A. Coulson and J. Sulston, pers. comm.). The set of clones from VT # YL34 through VT # YL17 was linked to the rest of the contig and oriented based on the following hybridization data: Y11E9 and Y17F2 hybridize to C25F2, but not C32C11, indicating that these two YACs extend from C25F2 to the right. Y17F2 hybridized to cosmids T15G9 and W03C9 and to phage VT # YL34. Probing genomic phage libraries with Y17F2 identified, VT # YL17, VT # YL18, and VT # YL34, as well as other phage clones overlapping C25F2, T15G9, and W03C9. Y17G7 did not hybridize to C25F2 but did hybridize to T15G9 and contains all the sequences of VT # YL34 and VT # 101. *C.elegans* DNA sequences flanking the two yeast telomeres of Y17F2 were cloned (see Materials and Methods) and probed to phage and cosmid clones. One end of the Y17F2 insert hybridized to C25F2 as expected, and the other end hybridized to phage VT # YL17, indicating that VT # YL17 is the furthest from C25F2 among the clones covered by Y17F2 (approximately 150 kb from C25F2). Phage VT # YL34 hybridized to part of Y11E9, indicating a partial overlap, and placing VT # YL34 approximately 30 kb from C25F2. All clones were isolated from wild type (N2) libraries with the exception of VT # 101 which was isolated from a library of *lin-29(n836)* DNA.

number of new clones were identified that extended leftward, back into the contig.

In contrast to the results of probing *E.coli*-based genomic libraries, several genomic clones maintained as Yeast Artificial Chromosomes (YACs) were detected by colony hybridization using a C25F2 probe and were found to extend to the right of C25F2 (Fig. 2). DNA prepared from these YACs was used to probe filters containing physically mapped *C.elegans* cosmid clones by J. Sulston and A. Coulson as part of the *C.elegans* genomic physical mapping project. Two of these YACs, Y11E9 (~70 kb) and Y17F2 (~170 kb) overlapped cosmid C25F2, but not C32C11, indicating that these clones extend 'rightward' from C25F2 toward *lin-29* (J. Sulston and A. Coulson, pers. comm.). Genomic DNA was isolated from yeast strains containing these putative *lin-29*-linked YAC clones and was compared with *C.elegans* genomic DNA by Southern hybridization to restriction digests, using the putative overlapping cosmids as probes. All YAC clones showed patterns of hybridization identical to those of the worm DNA, except for end junction-fragments (data not shown), indicating that these genomic clones likely represent genuine genomic extensions from C25F2. A subset of these YAC clones, including Y17F2 but not Y11E9, also identified two new, extensively overlapping cosmids, T15G9 and W03C9 (J. Sulston and A. Coulson, pers. comm.; see Fig. 2).

In order to identify additional *E.coli*-based clones containing sequences in Y17F2, gel-purified Y17F2 DNA was radiolabeled and used as a probe to a freshly packaged (un- amplified) genomic phage library. Phage clones detected by hybridization were plaque-purified and fingerprinted by J. Sulston and A. Coulson (pers. comm.). The resulting phage clones produced contiguous extensions of ~10 kb leftward from T15G9 (VT # YL34) and ~20 kb rightward from W03C9 (VT # YL17) (Fig. 2), as well as clones overlapping previously identified sequences in C25F2, T15G9, and W03C9 (data not shown). This second contig, containing VT # YL34, W03C9, T15G9 and VT # YL17, was oriented with respect to the first contig based on hybridization between VT # YL34 and Y11E9, and between an end-clone of Y17F2 and VT # YL17 (data not shown, see Materials and Methods). Sequences in VT # YL17 wee used to identify additional overlapping YACs by colony hybridization. One of these, Y17G7 (~250 kb), hybridized to T15G9, but not C25F2, indicating that it extended 100–150 kb farther 'right' than Y17F2. Using Y17G7 DNA as a probe to a genomic phage library, four small contigs of overlapping lambda clones were identified. These small contigs did not overlap with one another, or any previously identified clones and thus could not be ordered or oriented. Representative phage clones, designated VT # YL68, VT # YL77, VT # YL81, and VT # YL82 were confirmed to be fully contained within Y17G7 by comparing hybridization patterns to Southern blots of Y17G7 and Bristol N2 DNA and hence appear to be authentically linked to the *lin-29* contig (data not shown).

The above genetic and physical mapping experiments, combined with results from the *C.elegans* genomic physical mapping project, placed the probable physical position of *lin-29* near the right end of the contig of approximately 9000 kb. Our recombination map data indicate a position of *lin-29* approximately 0.2% recombination to the right of *maP1*. Based on the estimated *C.elegans* genome size of approximately 100,000 kb and an estimate of total recombinational distance for all linkage groups, the average ratio of physical distance to genetic distance in *C.elegans* is approximately 300 kb/cM (26), with a maximum value in certain regions of approximately 1500 kb/cM (26, 28). Thus, *lin-29* is likely to be about 50 kb (and probably no more than 300 kb) to the right of *maP1*. The contig identified by *maP1* flanking sequences contained approximately 250–300 kb of cloned sequences to the right of *maP1*, suggesting that this contig includes the *lin-29* gene. It is noteworthy that significant portions of this region of the contig expected to contain *lin-29*, specifically the region between C25F2 and VT # YL34 and regions between the four small contigs under Y17G7, were represented only by YAC clones.

Identification of two *lin-29* allele-specific RFLPs

To detect *lin-29* allele-specific alterations in restriction fragments, phage, cosmid, and YAC clones spanning the region from C25F2 through VT # YL17 (Fig. 2) were used to probe restriction digests of genomic DNA from wild type and four *lin-29* mutant strains. Using Alternating Field Gel-isolated Y11E9 DNA as a probe, an allele-specific polymorphism was readily detected in *Eco*RI digests of *lin-29(n836)* DNA. Two wild-type *Eco*RI fragments

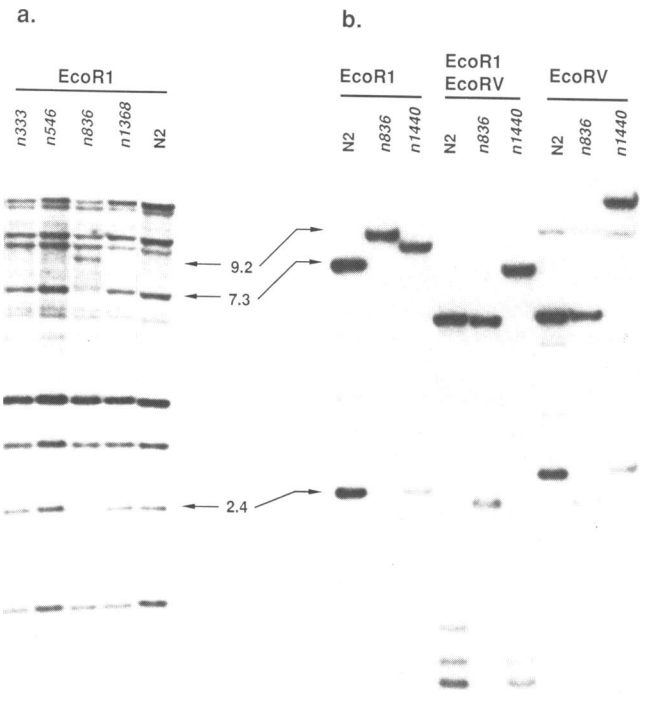

Figure 3. Allele-specific aberrations associated with *lin-29(n836)* and *lin-29(n1440)*. a) Gel-isolated Y11E9 DNA was radiolabeled and used as a probe to *Eco*RI digests of genomic DNA from wild type and *lin-29* mutants. Lanes 1–5: *lin-29(n333)*, *lin-29(546)*, *lin-29(n836)*, and *lin-29(n1368)*, N2 Bristol. *lin-29(n1440)* DNA was not analysed in this experiment. In *lin-29(n836)* DNA, fragments of 2.4 kb and 7.3 kb apparently fuse to form an approximately 9.2 kb fragment. Background bands are likely due to cross-hybridization between the Y11E9 probe and repeated sequences in *C.elegans* DNA. b) Analysis of *lin-29(n836)* and *lin-29(n1440)* allele-specific polymorphisms. The novel 9.2 kb *Eco*RI fragment detected in part a) was cloned from *lin-29(n836)* genomic DNA (yielding clone VT # 101; see Figs. 2 and 4) and used to probe genomic digests of N2 Bristol (the strain from which all *lin-29* mutants were isolated), *lin-29(n836)*, and *lin-29(n1440)*. The ethylmethanesulfonate (EMS)-induced allele *n836* results in the fusion of the 7.3 kb and 2.4 kb fragments into one of 9.2 kb, and the fusion of 1.50 kb and 1.15 kb *Eco*RI/*Eco*RV fragments to produce one of 2.15 kb. The 1.3 kb *Eco*RI/*Eco*RV band (the middle band of the triplet near the bottom) is detectable in the *n836 Eco*RI/*Eco*RV lane, but is not apparent in this exposure because the *n836* lane is relatively underloaded. A 0.7 kb band is not shown (see map, Fig. 4). *lin-29(n1440)* DNA shows an altered restriction pattern consistent with a chromosomal aberration, an insertion or inversion break point close to the 7.3 kb *Eco*RI fragment's terminal *Eco*RI and *Eco*RV sites.

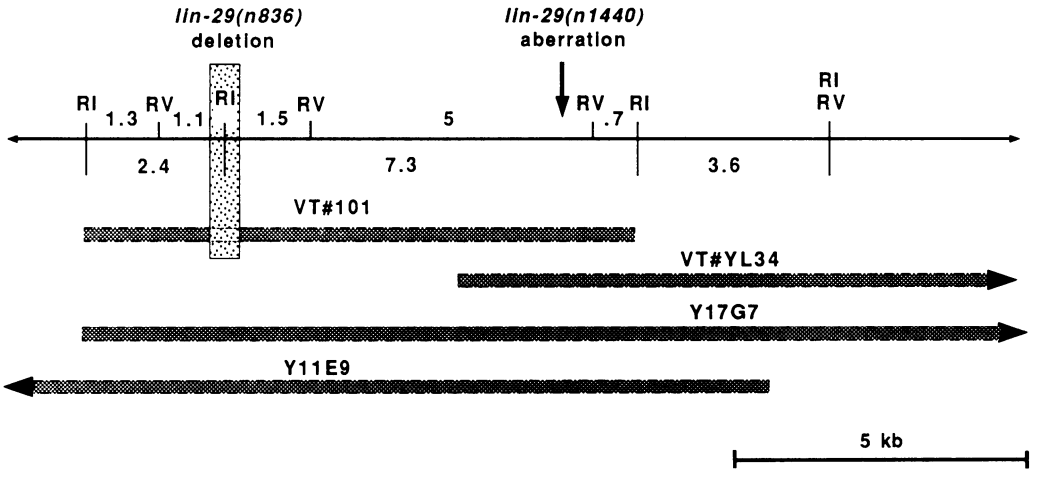

Figure 4. Partial restriction map of the *lin-29* region containing two *lin-29* allele-specific aberrations detected by Y11E9, based on the data in Fig. 3. These aberrations, and the corresponding 7.3 kb and 2.4 kb wild type *Eco*RI fragments are also detected by phage VT # 101 probe. The sizes, in kilobase pairs, of *Eco*RI and *Eco*RI/*Eco*RV digestion products are shown below and above the line, respectively. The precise end points of Y11E9 and VT # YL34 are not known; these clones are shown to end within the last genomic restriction fragment that they hybridize to. VT # YL34 hybridizes to the 7.3 kb but not the 2.4 kb *Eco*RI fragment. Y11E9 hybridizes to the 7.3 kb, and 3.6 kb *Eco*RI fragments that are also detected by VT # YL34, but does not detect other VT # YL34 *Eco*RI fragments. Y17G7 contains all of the sequences in VT # 101, but may extend somewhat to the left of the indicated end point. The ethylmethanesulfonate (EMS)-induced allele *n836* removes an *Eco*RI site, resulting in the fusion of the 7.3 kb and 2.4 kb fragments into one of 9.2 kb, and the fusion of 1.5 kb and 1.1 kb *Eco*RI/*Eco*RV fragments to produce one of 2.1 kb. This implies that the *n836* mutation deletes approximately 500 bp (indicated by the stippled box) surrounding the *Eco*RI site (see Fig. 3). The aberration associated with the *n1440* lesion is consistent with an insertion or rearrangement affecting the opposite end of the 7.3 kb *Eco*RI fragment from the one adjacent to the 2.4 kb fragment (Fig. 3). The fragment above the dotted lines represents an insertion consistent with the data regarding *lin-29(n1440)*. The 8000 bp region in this proposed insertion represents a minimum value, which is correct if the insertion contains no *Eco*RV sites. Otherwise, the inserted DNA could be any size larger than 8000 bp.

(7.3 kb and 2.4 kb) appear to be fused by the deletion of an *Eco*RI site and approximately 500 bp of sequences flanking it, producing a single new *Eco*RI fragment of approximately 9.2 kb (Fig. 3a, lane 3).

To obtain a less complex probe that detects the *n836* aberration, the 9.2 kb *Eco*RI fusion fragment from *n836* was cloned from a genomic *Eco*RI library of *n836* DNA. This subcloned fragment (clone VT # 101) specifically hybridizes to the *n836* RFLP described above and also detects an RFLP associated with *lin-29(n1440)* DNA. In *n1440* DNA, the 7.3 kb *Eco*RI fragment is increased in size to 8.3 kb while the 2.4 kb *Eco*RI fragment appears unaltered (Fig. 3b, lane 3). The restriction fragment alterations detected by the VT # 101 probe in various other digests of *n836* DNA are consistent with the above interpretation of the *n836* mutation as a deletion of approximately 500 bp that deletes the *Eco*RI site between the 7.3 and 2.4 kb fragments (Figs. 3, 4). The *n1440* mutation appears to involve a rearrangement (possibly an insertion) of DNA in the same 7.3 kb *Eco*RI fragment affected by *n836* (Figs. 3, 4).

Intragenic recombinants establish the identity of *n836* and the *n1440* RFLPs

To confirm that the two distinct aberrations of the 7.3 kb *Eco*RI fragment in *lin-29(n836)* and *lin-29(n1440)* DNAs correspond to the mutant alleles themselves, we selected phenotypically for intragenic recombinants in which a recombination event occurred between the two mutations. Four independent wild type recombinants were isolated from among approximately 40,000 progeny of *n836/n1440* hermaphrodites, as described in Materials and Methods. From each of these four recombinants, strains homozygous for the respective intragenic recombinant chromosome were derived and their DNAs were digested with *Eco*RI and analyzed by Southern blotting. All four recombinants were found to contain neither of the two allele-specific aberrations and only the wild type 7.3 and 2.4 kb *Eco*RI fragments (Fig. 4, lanes 3, 6, 9 and 12). Thus, in all four cases recombination occurred between the two DNA aberrations and hence within the 7.3 kb *Eco*RI fragment. Therefore, at least part of the *lin-29* gene must lie within this 7.3 kb *Eco*RI fragment. All four recombinant chromosomes retained an *unc-4(e120)* flanking marker, which was originally in cis to *n1440*, indicating that *n836* maps genetically to the left of *n1440*. This genetic order is consistent with the physical order derived from the hybridization of VT # YL34 probe to the 7.3 kb *Eco*RI fragment but not the 2.3 kb *Eco*RI fragment (data not shown; see summary in Fig. 4).

DISCUSSION

Localization of *lin-29* sequences to a 7.3 kb *Eco*RI fragment

Here we report the physical localization of the heterochronic gene *lin-29* to a small region of cloned *C.elegans* genomic DNA. These findings will serve as a foundation for the identification and sequence analysis of the *lin-29* gene product(s). Since the stage-specific activation of *lin-29* by upstream regulators triggers a developmental switch in the differentiation of lateral hypodermal cells (2), and since *lin-29* may also play roles in the differentiation of other cells (29), the structural and functional analysis of the *lin-29* gene products may illuminate mechanisms underlying the temporal regulation of cellular differentiation.

Using the technique of parallel RFLP mapping (11), DNA sequences in the vicinity of *lin-29* were identified and cloned. These *lin-29*-linked sequences identified an approximately 9,000 kb contig that was previously identified and genetically oriented with respect to the LGII genetic map by the *C.elegans* physical mapping project (A. Coulson and J. Sulston, pers. comm). Since this contig ended at a point genetically left of *lin-29*, and hence

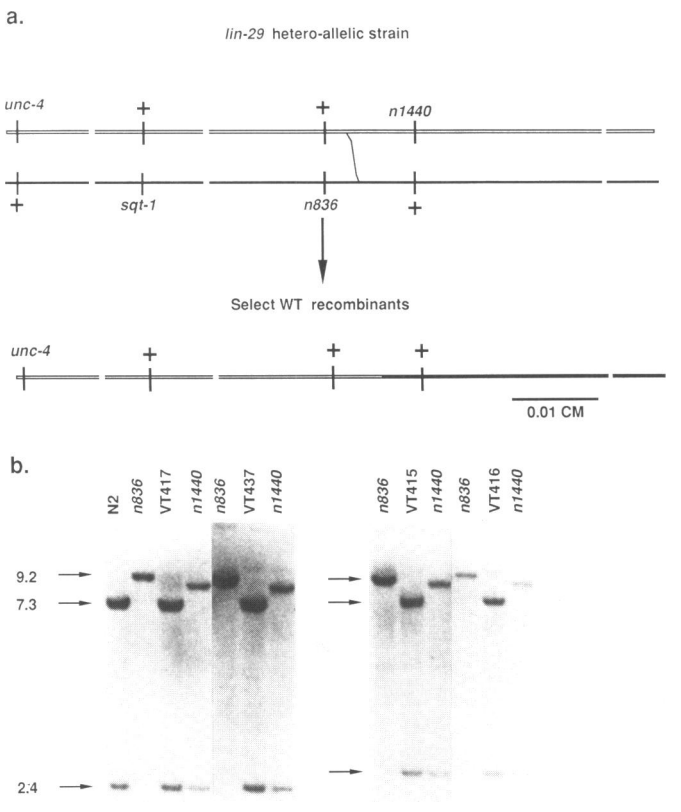

Figure 5. The segregation of *lin-29(n836)* and *lin-29(n1440)* allele-specific aberrations in intragenic recombinants. a) Schematic diagram showing the generation of *lin-29* intragenic recombinants from *lin-29(n836)/ lin-29(n1440)* heteroallelic parents. An *unc-4(e120)* marker was in cis to *n1440*. In some experiments, a *sqt-1* marker was in cis to *n836*. Strains were constructed as described in Materials and Methods. Four wild type recombinants, which would necessarily have resulted from recombination between the two alleles, were recovered from separate cultures. b) DNA was prepared from homozygous recombinant strains, and from strains homozygous for the parental *n836* and *n1440* chromosomes, digested with *Eco*RI and subjected to Southern blot analysis using VT # 101 probe as described in Materials and Methods. The arrows indicate the positions of the wild type 7.3 kb and 2.4 kb *Eco*RI fragments. The first lane is N2 DNA. Thereafter, each set of three lanes contains one recombinant strain (VT415, VT416, VT417 or VT437), and two strains homozygous for the parental *n546* and *n1440* chromosomes recovered from the same culture as the corresponding recombinant. Lanes 1−7 and lanes 8−13 were from separate gels.

did not at the outset include *lin-29*, genomic walking techniques were used to identify Yeast Artificial Chromosome (YAC), lambda and cosmid clones that extended this contig to the right. The *lin-29* gene was located in this new, extended region of the contig by using a radiolabeled YAC clone as hybridization probe to identify two allele-specific RFLPs. The fact that two independent *lin-29* alleles, the EMS induced *n836* mutation and the gamma-ray induced *n1440* mutation, cause distinct aberrations of the same *Eco*RI fragment strongly suggests that this fragment contains *lin-29* sequences. The physical position of this *Eco*RI fragment maps approximately 50−60 kb to the right of *maP1*, a position agreeing well with the predicted location of *lin-29* based on its genetic position and recombination distance from *maP1*.

Four independent intragenic recombinants were recovered by selecting, phenotypically, for recombination between the *n836* and *n1440* mutations. All four recombination events were shown by DNA analysis to have occurred between the two aberrations of the *Eco*RI fragment, confirming the physical correspondence of the *n836* and *n1440* mutations to their respective DNA aberrations, and hence confirming the location of *lin-29* sequences to this *Eco*RI fragment. Further, the segregation of a flanking marker, *unc-4(e120)*, in these intragenic recombinants indicated a genetic order of *n836* and *n1440* which is consistent with the physical order of the corresponding DNA alterations. These intragenic recombinants thus provide unambiguous proof that at least part of the *lin-29* gene is located within the 7.3 kb *Eco*RI fragment.

Sequences in the *lin-29* region are cloned in a YAC vector but resist cloning in cosmid and phage vectors

Significant portions of the contig in the region of the *C.elegans* genome expected to contain *lin-29* were represented only by YAC clones. Two approaches were used to obtain additional lambda and cosmid clones. First, after probing an estimated 20 or more genome-equivalents of phage and cosmid librairies with end probes from existing cosmid contigs, no new clones were found extending significantly to the right of C25F2. This suggests that sequences contiguous to the right end of C25F2 are significantly underrepresented in the phage and cosmid libraries examined. Second, gel-purified Y11E9 DNA, Y17F2 DNA and Y17G7 DNA were radiolabeled and used to probe unamplified genomic phage libraries of wild type *C.elegans* DNA. The results of these YAC-to-phage library hybridization experiments demonstrate the existence of several stretches of *C.elegans* DNA sequences, specifically the region between C25F2 and VT # YL34 and regions between the four small contigs under Y17G7, which are severely underrepresented in genomic *E.coli*-based libraries. These sequences were absent in spite of the use of a recBC host. Based on estimated sizes for Y17F2 and Y17G7 of approximately 170 kb and 250 kb, respectively (as determined by Alternating Field Gel electrophoresis, data not shown), it appears that of the approximately 400 kb of YAC sequences right of *maP1* on this contig, approximately 120 kb of sequences are not represented in the phage and cosmid libraries that we examined.

lin-29 seems to lie in a genomic region readily clonable using YAC vectors but which is significantly under-represented in cosmid and phage libraries fingerprinted as part of the *C.elegans* genomic mapping project. The under-representation of certain regions in cosmid libraries is a feature of the *C.elegans* physical map that accounts, in part, for the separation of the map into contigs after exhaustive fingerprinting of random clones (13). Connections between contigs have been efficiently achieved by identifying YAC clones that span contig gaps (13). Thus, the expectation that YAC libraries of eukaryotic sequences would be more representative than lambda and cosmid libraries (14) is borne out by the *C.elegans* physical mapping project in general, and by this analysis of the *lin-29* region in particular. The precise reasons why certain sequences are easily cloned in YAC vectors, but not in cosmid or phage vectors, is not known.

Further structural and functional characterization of *lin-29*'s role in hypodermal cell differentiation will require plasmid and phage subclones suitable for DNA transformation by microinjection (30, 31) and for sequence analysis. In this manuscript, we have described manipulations of YAC clones that have facilitated the analysis of *lin-29*-linked sequences. For example, labeled YAC DNA was used to probe genomic Southern blots and plaque lifts of phage libraries. Also, procedures for the cloning of sequences adjacent to the ends of YAC clones were employed to determine the limits of YAC clones with respect to other cloned sequences. Gel purified YAC DNA clones containing *lin-29* may be a convenient source of

DNA for subcloning *lin-29* sequences in prokaryotic or yeast plasmid vectors. Y17F2 may also be suitable for rescue of *lin-29* mutations by microinjection since a YAC clone covering *par-2* has been successfully used to rescue a *par-2* mutation by DNA transformation (D. Levitan, C. Mello and D. Stinchcomb, pers. comm.) and a YAC clone of the *tra-3* gene has been found to rescue a *tra-3* mutation (T. Barnes, pers. comm.).

ACKNOWLEDGEMENTS

Thanks to Alan Coulson and John Sulston for analyzing YAC, lambda and cosmid clones to identify overlaps with previously characterized clones, and also for their gift of gel purified YAC DNA and *C.elegans* genomic phage libraries. We are also grateful to Joan Park for providing the VT455 chromosome, to Diane Parry and H.R. Horvitz for providing *lin-29(n1440)* and to Robert Waterson for providing the gridded YAC library. Special thanks to Craig Mello, Diane Levitan and members of the laboratory for helpful discussions, and to Hugo Giambarella for technical assistance. This work was supported by Public Health Service research grant GM34028 and American Cancer Society Grant NP479A (to VA), and by the Damon Runyon/Walter Winchell Cancer Fund (to AER). Some of the strains used in this work were provided by the Caenorhabditis Genetics Center, which is supported by Contact N01RR-4-2111 between the NIH Division of Research Resources and Curators of the University of Missouri.

REFERENCES

1. Ambros,V. and Horvitz,H.R. (1984) *Science* **226**, 409–416.
2. Ambros,V. (1989) *Cell* **57**, 49–57.
3. Ambros,V. and Horvitz,H.R. (1987) *Genes Dev.* **1**, 398–414.
4. Ruvkun,G. and Guisto,J. (1989) *Nature* **338**, 313.
5. Singh,R. and Sulston,J.E. (1978) *Nematologica* **24**, 63–71.
6. Sulston,J. and Horvitz,H.R. (1977) *Dev. Biol.* **56**, 110–156.
7. Cox,G.N. and Hirsh,D. (1985) *Mol. Cell. Biol.* **5**, 363–372.
8. Sager,R. (1989) *Science* **246**, 1406–1412.
9. Bender,W., Spierer,P. and Hogness,D. (1983) *J. Mol. Biol.* **168**, 17–34.
10. Koenig,M., Hoffman,E., Bertelson,C., Monoco,A., Feener,C. and Kunkel,L. (1987) *Cell.* **50**, 509–517.
11. Ruvkun,G., Ambos,V., Coulson,A., Waterston,R., Sulston,J. and Horvitz,H.R. (1989) *Genetics* **121**, 501–516.
12. Coulson,A., Sulston,J., Brenner,S. and Karn,J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7821–7825.
13. Coulson,A., Watertson,R., Kiff,J., Sulston,J. and Korana,Y. (1988) *Nature* **335**, 184–186.
14. Burke,D., Carle,G. and Olson,M. (1987) *Science* **236**, 806–812.
15. Brenner,S. (1974) *Genetics* **77**, 71–94.
16. Herman,R.K. (1978) *Genetics* **88**, 49–65.
17. Hohn,B. and Collins,J. (1980) *Gene* **11**, 291–298.
18. Ish-Horowicz,D. and Burke,J. (1981) *Nucl. Acids Res.* **9**, 2989–2998.
19. Gibson,T., Coulson,A., Sulston,J. and Little,P. (1987) *Gene* **53**, 275–281.
20. Karn,J., Matthews,H., Gait,M. and Brenner,S. (1984) *Gene* **32**, 217–224.
21. Maniatis,T., Fritsch,E. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor University Press, Cold Spring Harbor.
22. Wertman,K., Wyman,A. and Botstein,D. (1986) *Gene* **49**, 253–262.
23. Wood,W.B. (1988) *The Nematode Caenorhabditis elegans*. (ed) Cold Spring Harbor Monograph Series **17**. Cold Spring Harbor University Press, Cold Spring Harbor.
24. Southern,E. (1975) *J. Mol. Biol.* **98**, 503.
25. Carle,G. and Olsen,M. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3756–3760.
26. Emmons,S. (1988) In *The Nematode Caenorhabditis elegans*. CSH Press, Cold Spring Harbor, NY.
27. Kramer,J., Johnson,J., Edgar,R., Basch,C. and Roberts,S. (1988) *Cell*, **55**, 555–565.
28. Greenwald,I., Coulson,A., Sulston,J. and Priess,J. (1987) *Nucl. Acids Res.* **15**, 2295–2307.
29. Liu,Z. and Ambros,V. (1988) *Genes and Development* **3**, 2039–2049.
30. Stinchcomb,D.T., Shaw,J.E., Carr,S.H. and Hirsh,D. (1985) *Mol. Cell. Biol.* **5**, 3484–3496.
31. Fire,A. (1986) *EMBO J.* **5**, 2673–2680.