Molecular Genetics of the *Caenorhabditis elegans* Heterochronic Gene *lin-14*

Gary Ruvkun,*†[1] Victor Ambros,†‡ Alan Coulson,§ Robert Waterston,‖ John Sulston§ and H. Robert Horvitz†

**Department of Molecular Biology, Massachusetts General Hospital and Department of Genetics, Harvard Medical School, Boston, Massachusetts 02114; ‡Department of Cellular and Developmental Biology, Harvard University, Cambridge, Massachusetts 02138; §Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 2QH, England; ‖Department of Genetics, Washington University School of Medicine, St. Louis, Missouri 63110; and †Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

Manuscript received September 20, 1988
Accepted for publication November 30, 1988

ABSTRACT

We describe a general strategy for the genetic mapping in parallel of multiple restriction fragment length polymorphism (RFLP) loci. This approach allows the systematic identification for cloning of physical genetic loci within about 100 kb of any gene in *Caenorhabditis elegans*. We have used this strategy of parallel RFLP mapping to clone the heterochronic gene *lin-14*, which controls the timing and sequence of many *C. elegans* postembryonic developmental events. We found that of about 400 polymorphic loci in the *C. elegans* genome associated with the Tc1 family of repetitive elements, six are within 0.3 map unit of *lin-14*. The three closest *lin-14*-linked Tc1-containing restriction fragments were cloned and used to identify by hybridization an 830-kb region of contiguous cloned DNA fragments assembled from cosmid and yeast artificial chromosome libraries. A *lin-14* intragenic recombinant that separated a previously cryptic *lin-14* semidominant mutation from a *cis*-acting *lin-14* suppressor mutation was used to map the location of the *lin-14* gene to a 25-kb region of this 830-kb contig. DNA probes from this region detected *lin-14* allele-specific DNA alterations and a *lin-14* mRNA. Two *lin-14* semi-dominant alleles, which cause temporally inappropriate *lin-14* gene activity and lead to the reiterated expression of specific early developmental events, were shown to delete sequences from the *lin-14* gene and mRNA. These deletions may define *cis*-acting sequences responsible for the temporal regulation of *lin-14*.

A hierarchy of interacting control genes specifies the diversity and configuration of cell types that arise during the development of multicellular organisms. Control genes of this type that coordinately regulate the temporal and spatial pattern of cell divisions and differentiations during the development of the nematode *Caenorhabditis elegans* have been identified by lineage (*lin*) mutations that alter the normally invariant cell lineage of this animal (HORVITZ and SULSTON 1980; STERNBERG and HORVITZ 1984). Many of these cell lineage mutations cause particular cells or groups of related cells to execute patterns of cell lineage normally executed by other cells. Mutations in the heterochronic genes, including *lin-4*, *lin-14*, *lin-28* and *lin-29*, coordinately affect diverse postembryonic cell lineages and tissues, causing particular cells from these lineages to express fates normally expressed by cells found earlier or later in the same lineages (AMBROS and HORVITZ 1984). These mutations are heterochronic in that they cause a change in the developmental stage at which particular cell types and differentiated structures are generated. This genetically induced change in the relative timing of various developmental events is similar to the heterochrony observed in the phylogenetic variation that exists among related species (AMBROS and HORVITZ 1984; GOULD 1977). Mutation in heterochronic genes that control temporal patterning may be the underlying cause of this phylogenetic variation in developmental timing.

The analysis of mutations in the *lin-14* gene has indicated that this heterochronic gene plays a central role in the global control of the temporal pattern of the *C. elegans* postembryonic cell lineages (AMBROS and HORVITZ 1984, 1987). These studies have suggested that postembryonic blast cells from a variety of tissues have alternative potential fates: one that normally occurs early in the postembryonic lineage of that blast cell, an early fate, and one that normally occurs late in that same lineage, a late fate. The level of *lin-14* gene activity selects between these two fates: a high level of *lin-14* activity causes particular blast cells to execute their early fates, whereas a low level of *lin-14* activity causes these cells to execute their late fates. This observation suggests that during normal development the activity of the *lin-14* gene is progressively reduced, and that this reduction causes multipotential blast cells to execute late cell fates

[1] To whom correspondence should be addressed.

instead of early cell fates. Mutations in the *lin-14* gene alter temporal patterns of cell lineage during development by altering this temporal pattern of *lin-14* gene activity: recessive alleles cause reduced gene activity at early stages and lead to the precocious expression of late cell fates, and semidominant alleles cause inappropriately elevated gene activity at later stages and lead to the reiterated expression of early cell fates.

The particular early or late cell fate specified by the level of *lin-14* gene activity is distinct for many of the postembryonic cell lineages affected by *lin-14* mutations, although the fates inappropriately executed in these mutants are always fates normally executed by a closely related descendent or ancestor cell (AMBROS and HORVITZ 1984). The *lin-14* gene product may function to convey general temporal information to these cell lineages. The specific response made by each cell must be caused by unique properties of that cell that either modify the *lin-14* signal or interpret it differently.

To understand both how *lin-14* gene activity is temporally regulated and how the level of *lin-14* gene activity causes each cell to execute its particular early or late fate, we have initiated molecular studies of the *lin-14* gene. Here we report the molecular cloning of DNA sequences spanning the *lin-14* locus. We used a general strategy we call parallel restriction fragment length polymorphism (RFLP) mapping to rapidly and systematically identify RFLP loci closely linked to *lin-14*. These loci were then cloned and used to identify an 830 kb contiguous region (contig) of overlapping cloned DNA fragments from the collection of contigs now being assembled in a project to physically map the entire *C. elegans* genome (COULSON *et al.* 1986, 1988). The *lin-14* gene was genetically mapped within this 830-kb contig using a *lin-14* intragenic recombinant. We find that two *lin-14* semidominant alleles, which cause elevated levels of *lin-14* activity late during development, are associated with the deletion of sequences from the *lin-14* gene and mRNA. These sequences may normally negatively regulate *lin-14* gene activity.

MATERIALS AND METHODS

Construction of the *lin-14(+ Berg)* congenic strain: The *lin-14* gene is *X*-linked and *C. elegans* males are *XO*. Nonconditional *lin-14* mutant males are incapable of mating because of cell lineage defects in the generation of the male tail (AMBROS and HORVITZ 1987). Therefore, a strain carrying the temperature-sensitive *lin-14(n1790ts)* mutation was used to transfer the Bristol *X* chromosome into the Bergerac strain RW7000 (MOERMAN and WATERSTON 1984). Bristol *him-5(e1467); lin-14(n179ts Bris)/0* males grown at 15° (therefore phenotypically non-Lin-14) (AMBROS and HORVITZ 1987) were crossed to strain Bergerac *lin-14(+ Berg)* hermaphrodites. Heterozygous *lin-14(n179ts Bris)/lin-14(+ Berg)* hermaphrodites were then mated at 25° to strain Bristol *him-5(e1467); lin-14(n179ts Bris)/O* males (which had been grown at 15°), for a total of 10 crosses of the Bergerac *lin-14(+ Berg)* region into the Bristol strain. After the last cross a *lin-14(n179ts Bris)/lin-14(+ Berg)* hermaphrodite was allowed to self-fertilize and non-Lin-14 hermaphrodites were picked at successive generations until a strain was isolated that no longer segregated Lin-14 hermaphrodites and therefore was homozygous for the *lin-14(+ Berg)* chromosome.

Genetic mapping of Tc1-dimorphic loci: The *(nP1 nP2...nP32dpy-6(+ Berg) lin-14(+ Berg)) sma-5(+ Berg)) X* chromosome was crossed into two different doubly marked Bristol strains containing a recessive *lin-14(−)* allele linked in *cis* to the closest convenient visible genetic markers mapping to the left and right of *lin-14*: *dpy-6(e14 Bris) lin-14(n355n679 Bris) X* and *lin-14(n179 Bris) sma-5(n678 Bris) X*. The *lin14(n355n679)* allele is temperature sensitive. The *dpy-6* and *lin-14* genes are separated by about 7 map units (M. STERN and C. NUSSBAUM, personal communication) and *lin-14* and *sma-5* are separated by 1.5 map units (see Figure 2B). The *dpy-6(e14 Bris) lin-14(n355n679)/(nP1 nP2...nP32 dpy-6(+ Berg) lin-14(+ Berg)* strain was constructed in the following manner. Wild-type N2 males were mated to *(nP1 nP2...nP32 dpy-6(+ Berg) lin-14(+ Berg))* hermaphrodites at 20°. The resulting *(nP1 nP2...nP32 dpy-6(+ Berg) lin-14(+ Berg))/O* males were mated at 25° to *dpy-6(e14) lin-14(n355n679)* hermaphrodites previously grown at 20°, at which temperature they are capable of mating. Non-Lin-14 non-Dpy-6 hermaphrodite progeny that segregated 25% Dpy-6 Lin-14 progeny were isolated and their progeny were screened for recombinants. Similarly, *(nP1 nP2...nP32 lin-14(+ Berg) sma-5(+ Berg))/O* males were mated to *lin-14(n179 Bris) sma-5(n678 Bris)* hermaphrodites to construct the heterozygote *lin-14(n179 Bris) sma-5(n678 Bris)/(nP1 nP2...nP32 lin-14(+ Berg) sma-5(+ Berg))*.

The Lin-14 non-Dpy-6 recombinants were initially picked as *dpy-6(+ Berg) lin-14(n355n679 Bris)/dpy-6(e14 Bris) lin-14(n355n679 Bris)* animals. The Lin-14 non-Sma-5 recombinants were initially picked as *lin-14(n179 Bris) sma-5(+ Berg)/lin-14(n179 Bris) sma-5(n678 Bris)* animals. Lin-14 non-Dpy-6 animals and Lin-14 non-Sma-5 animals were individually isolated and allowed to self-fertilize for successive generations until strains that no longer segregated Lin-14 Dpy-6 progeny and Lin-14 Sma-5 progeny, respectively, were established. These strains were therefore homozygous for each of the recombinant chromosomes and used as sources of DNA.

The *dpy-6(e14 Bris) lin-14(+ Berg) lin-2(+ Bris) unc-9(e101 Bris)* strain was constructed by the following procedure: a Dpy-6 non-Lin-14 animal was isolated from the progeny of a heterozygous *dpy-6(e14 Bris) lin-14(n355n679 Bris)/lin-14(+ Berg)* hermaphrodite to generate a *dpy-6(e14 Bris) lin-14(+ Berg)* recombinant chromosome. The *tra-1* mutation (HODGKIN and BRENNER 1977) was used to generate *XX* males so that this recombinant chromosome could be transferred to other strains by male mating. In this way, a heterozygote of genotype *tra-1(e1099 Bris)/+; dpy-6(e14 Bris) lin-14(+ Berg)/lin-14(n179 Bris) lin-2(e1309 Bris)* was constructed. A Lin-2 non-Lin-14 recombinant was picked from the progeny from this strain and shown to segregate *dpy-6(e14 Bris) lin-14(+ Berg) lin-2(e1309 Bris)* homozygotes. A heterozygote of genotype *tra-1(e1099 Bris)/+; dpy-6(e14 Bris) lin-14(+ Berg) lin-2(e1309 Bris)/dpy-7(e1324 Bris) unc-9(e101 Bris)* was constructed, Unc-9 non-Dpy-7 recombinants were picked, and strains containing homozygous *dpy-6(e14 Bris) lin-14(+ Berg* or *Bris) unc-9(e101 Bris)* chromosomes were isolated. DNA samples from six such recombinants were analyzed by Southern blotting using a ^{32}P-Tc1

DNA probe. One of the six *dpy-6(e14 Bris) lin-14(+ Berg) unc-9(e101 Bris)* recombinant strains retained the *lin-14*-linked Tc1-containing *Eco*RI fragment loci *nP1, nP3, nP8, nP11, nP13, nP21* plus the flanking Tc1 loci *nP15* and *nP23* to the left and *nP18* to the right. This strain was used as the source of the *lin-14(+ Berg)* region and flanking Tc1-containing *Eco*RI fragment loci for the intragenic recombination experiment. One of five other recombinant strains (shown as #2 in Figure 3) apparently resulted from a fortuitous recombination within the cluster of *lin-14*-linked Tc1-containing *Eco*RI fragments: this strain retained only *nP1, nP11,* and *nP21* from this cluster and lost *nP3, nP8,* and *nP13.*

The *lin-14(n536sd n540)/dpy-6(e14 Bris) lin-14(+ Berg) unc-9(e101 Bris)* strain was constructed as follows: Males of genotype *tra-1(e1099 Bris); dpy-6(e14 Bris) lin-14(+ Berg) unc-9(e101 Bris)/+++* were constructed and mated to *lin-14(n536sd n540)/szT1 lin-14(+ Bris)* hermaphrodites. Individual cross-progeny were isolated and their progeny isolated at successive generations until a strain that segregated ¼ Dpy-6 Unc-9, ¼ Lin-14, and ½ wild-type progeny was identified.

Cloning Tc1-dimorphic loci: The Tc1 DNA probe for Southern blots was plasmid pCe2001, which contains a Bergerac Tc1 element (EMMONS *et al.* 1983). This DNA probe was oligo-labeled (FEINBERG and VOGELSTEIN 1983) and used to probe Southern blots of the *Eco*RI-digested DNAs isolated from the recombinant strains, separated electrophoretically on 0.8% agarose gels, and transferred to Gene Screen (New England Nuclear) or Biodyne (ICN).

DNA was isolated from the *dpy-6(e14 Bris) (nP18 nP23) (nP1 nP11 nP21) (nP8 nP13) lin-14(+ Berg) nP3 nP18 unc-9(e101 Bris)* strain as described (EMMONS, KLASS and HIRSH 1979), digested with *Eco*RI and electrophoretically separated on a preparative 0.7% agarose gel (MANIATIS, FRITSCH and SAMBROOK 1982). Size fractions (3.1, 4.0 and 5.4 kb, respectively) expected to contain Tc1-containing *Eco*RI fragments corresponding to *nP3, nP8,* and *nP13* were electroeluted, ligated to λgt7 vector DNA (DAVIS, BOTSTEIN and ROTH 1980), and this ligation mix was packaged into lambda particles *in vitro* (Pharmacia). Phage were plated onto *Escherichia coli* strain C600 and plaques were absorbed to nitrocellulose filters and hybridized to ^{32}P-labeled Tc1 DNA. Positive-hybridizing clones were shown to contain the correct Tc1-containing *Eco*RI fragment as follows. Unique sequence DNA adjacent to the Tc1 insertion site was hybridized to Southern blots of *Eco*RI-digested DNAs isolated from the recombinant Lin-14 non-Dpy-6 and Lin-14 non-Sma-5 strains described in the text. The cloned restriction fragments hybridized to the Bristol non-Tc1-containing alleles of these RFLPs in all of these recombinants, and to the 1.6-kb larger Bergerac alleles in the *lin-14(+ Berg)* strain (data not shown), showing that these fragments map to the *lin-14* region. Unique sequence restriction fragments from each of these clones was used to isolate cosmids by colony hybridization.

Detection of non-Tc1-associated RFLPs: Bristol/Bergerac RFLPs not associated with Tc1 were found by digesting DNAs from Bergerac strain RW7000 and Bristol strain N2 with restriction enzymes *Eco*RI, or *Bgl*II, or *Hin*dIII, or *Xho*I, and analyzing the digests by Southern blotting using ^{32}P-labeled cosmid probes. Any difference between the Bristol and Bergerac DNAs in the size of a hybridization band detected by a particular cosmid probe defined a new RFLP. The presence of the Bergerac or Bristol allele of this RFLP was then assessed in the *lin-14* intragenic recombinant strain by digesting DNA from this strain with the restriction enzyme with which the Bristol/Bergerac RFLP was originally detected and analyzing this DNA on Southern blots. RFLP *nP34* was detected by *Xho*I digestion of the genomic DNAs and hybridization to cosmid KKH9; the Bergerac allele is 7 kb and the Bristol allele is 9 kb. RFLP *nP33* was detected by digestion of genomic DNAs with *Eco*RI and hybridization to ^{32}P-labeled cosmids EEG4 or PPE4; the Bergerac allele is 1.5 kb and the Bristol allele is 1.2 kb. RFLP *nP35* was detected by digestion of genomic DNAs with *Eco*RI and hybridization to cosmid HHG9; the Bergerac allele is 4.5 kb and the Bristol allele is 4.9 kb.

The DNAs isolated from strains containing the following *lin-14* alleles were digested with *Eco*RI, *Hin*dIII, *Bgl*II, and *Xho*I and subjected to Southern blot analysis with probes from the cosmids EEG4, PPE4, and KKH9 (see Figure 5) as described in the text: MT1143 = *lin-14(n530)*, MT1144 = *lin-14(n355n531)*, MT1146 = *lin-14(n355n533)*, MT1147 = *lin-14(n355n534)*, MT1149 = *lin-14(n536)*, MT1153 = *lin-14(n536n540)*, MT1842 = *lin-14(n536n838)*, MT1849 = *lin-14(n536n837)/szT1*, MT2000 = *lin-14(n536n839)/szT1*, MT1150 = *lin-14(n536n537)*, MT1151 = *lin-14(n536n538)*, MT1152 = *lin-14(n536n539)*, MT355 = *lin-14(n355)*, MT1534 = *lin-14(n355)/lin-14(n355n726)*, MT925 = *lin-14(n355n407)*, MT1851 = *lin-14(n727)*, MT1388 = *lin-14(n355n679)*, MT1846 = *lin-14(n355n840)*, MT1397 = *lin-14(n179)*, MT1848 = *lin-14(n360)*. Using cosmid probes HHG9, C15G3 and C03B2 from the *nP3* region, cosmid probes C02H8 and C12B3 from the *nP13* region, and cosmid probe C10B6 from the *nP8* region (data not shown), no changes in hybridization pattern were observed on Southern blots of *Eco*RI-digested or *Hin*dIII-digested DNAs isolated from the strains containing *lin-14* alleles. Also, no changes from wild-type were detected in DNA isolated from the following non-*lin-14* strains, digested with *Bgl*II or *Hin*dIII, and subjected to Southern blot analysis using ^{32}P-labeled cosmid EEG4 probe: wild-type N2, TR287 = *unc-54(r241)*, CB190 = *unc-54(e190)*, CB2384 = *unc-54(e1660)*, MT177 = *lin-12(n177)*, MT302 = *lin-12(n302)*, MT2067 = *unc-86(n994)*, CB1416 = *unc-86(e1416)*, MT2206 = *dpy-19(e1259) sup-5(e1464)*, MT2208 = *lon-1(e185) sup-5(e1464) sma-2(e502)*.

Detection of *lin-14* mRNA: Poly(A)$^+$ RNA was prepared from cultures grown on agarose plates or in liquid (SULSTON and BRENNER 1974; COX *et al.* 1985) using the guanidinium isothiocyanate and oligo dT column protocols as described (MANIATIS, FRITSCH and SAMBROOK 1982). Northern blots were done using Gene Screen (New England Nuclear) or Biodyne (ICN) matrix. Double-stranded DNA probes from restriction fragments excised from agarose gels were prepared using the oligo-labeling method (FEINBERG and VOGELSTEIN 1983). Single-stranded DNA templates for hybridization probes were prepared as follows. Cells bearing the 7.6-kb and 15-kb *Bgl*II restriction fragments subcloned in either orientation in Bluescribe were superinfected with M13 phage and the resulting phage particles were purified as specified by the vendor (Stratagene Cloning Systems, San Diego, California). This single-stranded DNA served as a template for a universal hybridization primer (New England Biolabs) from which ^{32}P-dATP was incorporated using Klenow DNA polymerase I. Hybridizations and washes were done as described (MANIATIS, FRITSCH and SAMBROOK 1982).

RESULTS

Parallel restriction fragment length polymorphism mapping: The *lin-14* gene was defined by the

anatomical and cell lineage defects that result from mutations in this gene (AMBROS and HORVITZ 1984). Although we knew the genetic location of *lin-14* on the *X* chromosome, we had no hint of the biochemical identity of its gene product. We chose to clone *lin-14* by cloning genetically linked restriction fragments identified by RFLP mapping (BOTSTEIN *et al.* 1980), which, to then reach the *lin-14* gene itself, were used to identify overlapping larger contiguous regions (contigs) of cloned DNA assembled as part of the incipient *C. elegans* physical genome map (COULSON *et al.* 1986, 1988). However, in contrast to, for example, human RFLP mapping, in which the segregation behavior in pedigrees of individual RFLP probes are monitored serially until RFLP loci closely linked to the gene of interest are identified, the parallel RFLP mapping protocol we developed could monitor in parallel using one DNA probe on one pedigree, the genetic linkage of the 32 RFLP loci most closely linked to *lin-14* out of about 400 such loci in the *C. elegans* genome. In this way, we could rapidly and systematically identify and clone RFLP loci mapping very close to *lin-14*.

Two interbreeding strains of *C. elegans*, strain Bergerac and strain Bristol, contain a high level of restriction fragment length polymorphism (EMMONS *et al.* 1983; COX *et al.* 1985). In addition to RFLPs caused by random DNA sequence variation, a major difference between the Bristol and Bergerac strains is that in Bristol there are 30 copies of the 1.6-kb element Tc1, while in Bergerac there are 300–500 copies (EMMONS *et al.* 1983; FINNEY 1987). Thus, we assumed that randomly scattered throughout the genome there are approximately 400 loci that are "Tc1 dimorphic" between Bristol and Bergerac. These loci can be detected by Southern blot analysis using a Tc1 DNA probe; a restriction fragment encompassing each such locus in Bergerac will hybridize to a Tc1 DNA probe, while the corresponding Bristol fragment will not. The genetic linkage of each of these dimorphic loci with respect to other visible genetic markers and to each other can be monitored by Southern blot analysis of progeny from a Bristol/Bergerac cross using a Tc1 DNA probe; each Tc1-containing restriction fragment is detected as a hybridization band of a characteristic size and segregates as a unique genetic locus in genetic mapping experiments.

The *C. elegans* genome contains approximately 8×10^4 kb of DNA (SULSTON and BRENNER 1974). Assuming a random distribution of the 400 Tc1 elements in the Bergerac strain, there should be one dimorphic Bergerac Tc1-associated RFLP locus approximately every 200 kb. Assuming that genes are also distributed randomly within this dimorphic Tc1 locus linkage map, and using a weighted Poisson distribution, we calculated that the *lin-14* locus (or any other genetic locus in *C. elegans*) should map within about 100 kb of its closest dimorphic Tc1-containing restriction fragment locus. Thus, to clone *lin-14*, we identified the Bergerac-specific Tc1-containing restriction fragment loci most closely linked to and flanking *lin-14*, and using Tc1 DNA as a probe, cloned those restriction fragments. These cloned DNA fragments were derived from chromosomal locations genetically, and thus physically, linked to *lin-14* and were used as probes to isolate cosmid clones and larger multiclone "contigs" to reach the *lin-14* locus.

Construction of a *lin-14*(+ *Bergerac*) strain congenic with the Bristol strain: Because we needed to genetically map only those Tc1-dimorphic loci in the *lin-14* region of the Bergerac *X* chromosome (and not all 400 of the Tc1-dimorphic loci scattered throughout the genome), we constructed a strain that contained the *lin-14*(+) gene of Bergerac and its genetically linked Tc1-dimorphic loci within about 10 map units on either side of *lin-14*(+ *Berg*), but contained the Bristol genome for essentially all genetic regions not closely linked to *lin-14* (Fig. 1A). We denote this Bergerac chromosome region *lin-14*(+ *Berg*). Similarly, we denote the same chromosomal region derived from the strain Bristol as *lin-14*(+ *Bris*).

As shown in Figure 1A, this *lin-14*(+ *Berg*) chromosome was constructed by repeatedly crossing animals carrying the Bergerac chromosome containing the *lin-14*(+ *Berg*) allele with Bristol animals containing a recessive *lin-14*(−) allele, selecting for the Bergerac *lin-14*(+ *Berg*) allele derived from the original Bergerac parent in the resulting progeny. In this way, Tc1-dimorphic loci on chromosomes other than the *X* and on regions of the *X* chromosome not closely linked to the selected Bergerac *lin-14*(+ *Berg*) region were replaced by Bristol chromosomal regions containing the Bristol alleles of these loci.

DNA was prepared from the *lin-14*(+ *Berg*) strain congenic with the Bristol strain, digested with *Eco*RI, which does not cut within the Tc1 element (ROSENZWEIG, LIAO and HIRSH 1983), and analyzed by Southern blotting with ^{32}P-Tc1 DNA probe. As shown in Figure 2A, the congenic *lin-14*(+ *Berg*) strain (lane 6) had acquired many Tc1-hybridizing *Eco*RI bands that were not present in the parent Bristol strain (lane 7). These Bergerac-derived Tc1-containing *Eco*RI fragments (and the corresponding Bristol-derived *Eco*RI fragments, which do not contain Tc1) defined 32 genetic loci presumably genetically linked to *lin-14* on the *X* chromosome. These dimorphic Bergerac Tc1-containing *Eco*RI fragment loci in the congenic strain are denoted *nP1*, *nP2*,...,*nP32* in order of increasing size. The Bristol alleles of these *Eco*RI fragments that do not contain Tc1 we denote *nP1*(*Bris*) to *nP32*(*Bris*).

Genetic mapping of the *lin-14*-linked Tc1-containing loci using 34-factor crosses: We performed two

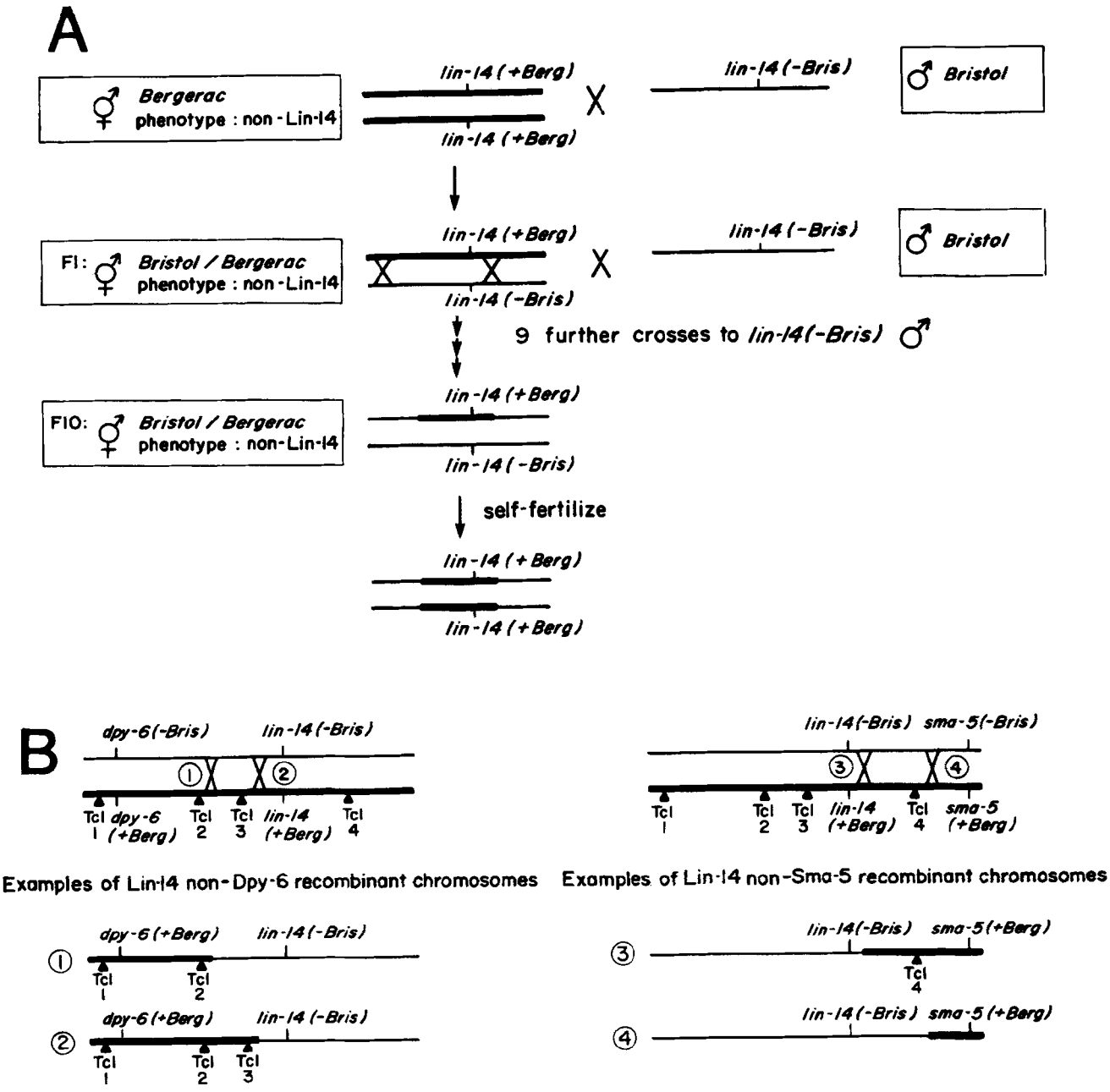
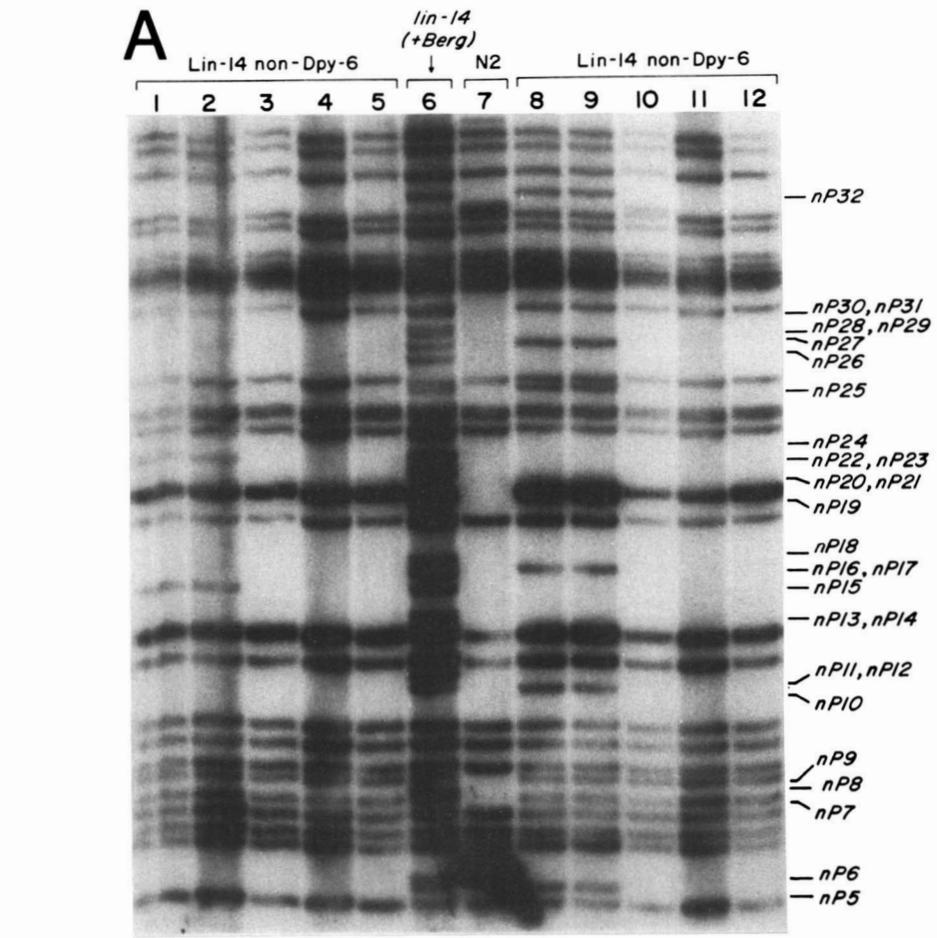
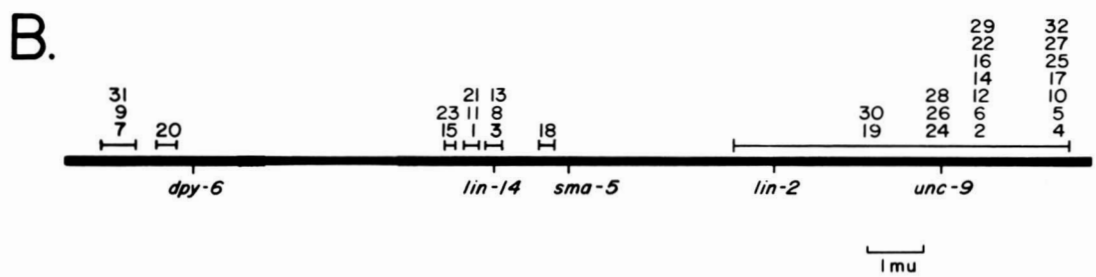
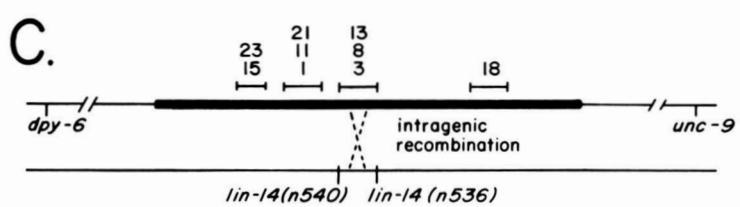
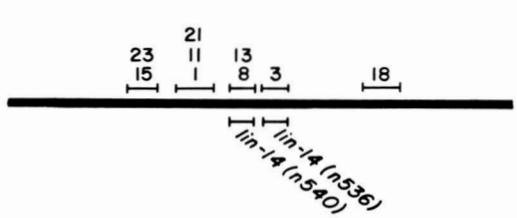


FIGURE 1.—Parallel RFLP mapping using a repetitive DNA probe. A, Construction of a *lin-14*(+ *Bergerac*) strain congenic with the Bristol strain. The *lin-14*(+ *Berg*) strain was constructed by serially crossing ten times a Bristol strain carrying the *lin-14*(*n179 Bris*) *X* chromosome with a strain carrying an *X* chromosome bearing the *lin-14*(+ *Berg*) genetic region (see MATERIALS AND METHODS). While progeny retaining the *lin-14*(+ *Berg*) region were selected after each cross, all other genetic regions were allowed to segregate randomly. Thus after 10 rounds of such crosses, the Bergerac Tc1-dimorphic loci unlinked to *lin-14* would have a probability of 0.5 per generation after the F_1, or $1/500$, of remaining in the strain. Recombination events separating *X*-linked Bergerac Tc1-dimorphic loci genetically mapping more than 10 map units from the *lin-14*(+ *Berg*) gene would each have a 10% probability of occurring per generation after the F_1; after 10 generations, only those Bergerac Tc1-dimorphic loci within 10 map units both left and right of the Bergerac *lin-14*(+ *Berg*) allele are expected to be maintained. B, Parallel RFLP mapping of the dimorphic Tc1-containing *Eco*RI fragments near *lin-14*: an example of how Tc1-dimorphic loci can be genetically mapped. The *dpy-6*(*e14 Bris*) *lin-14*(*n355n679r Bris*)/*lin-14*(+ *Berg*) strain shown in the left panel segregates ¾ wild-type and ¼ Dpy-6 Lin-14 progeny, except for the approximately 7% of the progeny that are Lin-14 non-Dpy-6 or Dpy-6 non-Lin-14 due to recombination between the *dpy-6*(*e14*) and *lin-14*(*n355n679r*) loci. Similarly, the *lin-14*(*n179ts*) *sma-5*(*n678*)/*lin-14*(+ *Berg*) strain shown in the right panel segregates ¾ wild-type and ¼ Lin-14 Sma-5 progeny, except for the approximately 1.5% recombinant Lin-14 non-Sma-5 or Sma-5 non-Lin-14 progeny. The recombination event that generated each Lin-14 non-Dpy-6 or Lin-14 non-Sma-5 strain could have occurred anywhere in the 7-map unit *dpy-6* to *lin-14* or 1.5-map unit *lin-14* to *sma-5* interval, respectively. As shown, the Lin-14 non-Dpy-6 strain #1 resulting from a recombination event at crossover point 1 would retain Tcls 1 and 2 and lose Tcls 3 and 4. The Lin-14 non-Dpy-6 strain #2 resulting from a recombination event at crossover point 2 would retain Tcls 1, 2, and 3 and lose Tcl 4. The presence or absence of these Tcl-containing *Eco*RI fragments is scored by probing a Southern blot containing DNA isolated from the recombinant strains with ^{32}P Tcl DNA probe. Coupled with the observation that no Lin-14 non-Sma-5 recombinants retain Tcls 1, 2, and 3, these two Lin-14 non-Dpy-6 recombinants would map Tcl 3 to between *dpy-6* and *lin-14*, and Tcls 1 and 2 near or to the left of *dpy-6*. The order of Tcls 1 and 2 could be determined by an additional recombination event between *dpy-6* and Tcl 2. The observation that only Lin-14 non-Sma-5 recombinant #3, and not Lin-14 non-Sma-5 recombinant #4, and no Lin-14 non-Dpy-6 recombinants retain Tcl 4 would map Tcl 4 to the *lin-14* to *sma-5* interval.


A
lin-14
(+Berg)
Lin-14 non-Dpy-6
N2
Lin-14 non-Dpy-6
1 2 3 4 5 6 7 8 9 10 11 12
nP32
nP30, nP31
nP28, nP29
nP27
nP26
nP25
nP24
nP22, nP23
nP20, nP21
nP19
nP18
nP16, nP17
nP15
nP13, nP14
nP11, nP12
nP10
nP9
nP8
nP7
nP6
nP5


B.
31 9 7
20
21 13
23 11 8
15 1 3
18
29 22 16 14 12 6 2
32 27 25 17 10 5 4
30 19
28 26 24
dpy-6
lin-14
sma-5
lin-2
unc-9
1 mu


C.
23 15
21 11 1
13 8 3
18
dpy-6
intragenic recombination
unc-9
lin-14(n540)
lin-14 (n536)


23 15
21 11 1
13 8 3
18
lin-14 (n536)
lin-14 (n540)

sets of 34-factor crosses to genetically map the 32 dimorphic Tc1-containing loci relative to each other, to *lin-14*, and to the flanking visible genetic markers *dpy-6* and *sma-5*. The segregation behavior of the 32 Tc1-dimorphic loci detected on Southern blots using ^{32}P-labeled Tc1 DNA probe was assessed in strains in which recombination events had been selected in genetic intervals to the left and right of *lin-14* (Figure 1B). These Tc1-dimorphic loci behave as unselected genetic markers, the segregation of which can be scored in these recombinants using a DNA probe. These segregation data are used to generate a genetic map analogously to mapping unselected markers in any three-factor cross (BRENNER 1974).

We collected 25 independent Lin-14 non-Dpy-6 recombinant progeny from a *dpy-6(e14 Bris) lin-14(n355n679 Bris)/dpy-6(+ Berg) lin-14(+ Berg)* parent strain and nine independent Lin-14 non-Sma-5 recombinant progeny from a *lin-14(n179 Bris) sma-5(n678 Bris)/lin-14(+ Berg) sma-5(+ Berg)* parent strain. DNA was prepared from each of the 34 homozygous recombinant strains, digested with *Eco*RI, and probed with ^{32}P-Tc1 DNA in a Southern blot analysis (Figure 2A). By analyzing the segregation of the Tc1-containing *Eco*RI fragment loci with respect to each

FIGURE 2.—A, Mapping Tc1-containing *Eco*RI fragment loci by 34-factor crosses: Twenty-five independent strains with recombination points between *dpy-6* and *lin-14* were isolated from a *dpy-6(e14 Bris) lin-14(n355n679r Bris)/(nP1 nP2...nP32 dpy-6(+ Berg) lin-14(+ Berg))* strain. Nine independent strains with recombination points between *lin-14* and *sma-5* were isolated from a *lin-14(n179ts) sma-5(n678)/(nP1 nP2...nP32,lin-14(+ Berg) sma-5(+ Berg))* strain. DNA was isolated from each strain homozygous for these recombinant chromosomes, digested with *Eco*RI, and analyzed by Southern blotting using ^{32}P-labeled Tc1 DNA probe. Shown here are the Tc1-containing *Eco*RI fragment loci *nP5* to *nP32* as seen in 10 of the 34 recombinant strains analyzed, as well as Tc1-containing *Eco*RI fragments from wild-type Bristol strain N2 and the congenic *lin-14(+ Berg)* strain. In lanes 1–5 are DNAs from independent Lin-14 non-Dpy-6 recombinant strains; lane 6: *lin-14(+ Berg)* congenic strain; lane 7: wild-type Bristol strain N2; lanes 8–12: independent Lin-14 non-Dpy-6 recombinant strains. The Tc1-containing *Eco*RI fragments present in strain Bristol are not named. Those Tc1-containing *Eco*RI fragments present in the *lin-14(+ Berg)* strain are denoted by *nP1* to *nP32* in order of increasing size as labeled on the right of the figure. In DNA isolated from Bristol strain N2 (lane 7), two of the 30 Bristol-specific Tc1 loci which also happen to be dimorphic with Bergerac can be seen at positions just below the Bergerac Tc1-containing loci *nP32* and *nP26*. B, Genetic map of Tc1-containing *Eco*RI fragment loci on the *X* chromosome near *lin-14* as derived from intergenic recombinants. The presence or absence of each Tc1-containing *Eco*RI fragment locus was scored in each recombinant strain and map positions were assigned. Those Tc1-dimorphic loci that always segregated together in these recombinants are shown as vertical clusters above the genetic map. Figure 3 shows a summary of the Tc1 genotypes of each recombinant strain from which this map was generated. Because the Lin-14 non-Dpy-6 and Lin-14 non-Sma-5 recombinant strains were selected for the *lin-14(− Bris)* chromosomal region, the Tc1-containing Bergerac alleles of the dimorphic loci most closely linked to the *lin-14* gene were not present in any of these recombinants. Nine Bergerac Tc1-dimorphic loci appeared in few or no recombinants in either genetic interval (*nP1, nP3, nP8, nP11, nP13, nP15, nP18, nP21, nP23*) and thus mapped closest to *lin-14*. The *nP3, nP8*, and *nP13* cluster and the *nP1, nP11*, and *nP21* cluster of Tc1-dimorphic loci were not separated from *lin-14* by these recombinants. The Tc1-containing *Eco*RI fragment loci retained in the Lin-14 non-Dpy-5 recombinant strains (genotype of visible genetic markers: *dpy-6(+ Berg) lin-14(n355n679 Bris)*), but not in the Lin-14 non-Sma-5 recombinant strains were positioned on the map near or to the left of *dpy-6(+ Berg)* on the *X* chromosome. Similarly, the Tc1-containing *Eco*RI fragment loci retained in the Lin-14 non-Sma-5 recombinant strains (genotype of visible genetic markers: *lin-14(n179 Bris) sma-5(+ Berg)*), but not in the Lin-14 non-Dpy-6 recombinant strains (except in the case of unselected second recombination events described below) were positioned on the map near or to the right of Bergerac *sma-5(+ Berg)*. Four Tc1-dimorphic loci (*nP7, nP9, nP20, nP31*) mapped to the left of (or some of them possibly to the right and very near to) *dpy-6(+ Berg)* and 19 loci (*nP2, nP4, nP5, nP6, nP10, nP12, nP14, nP16, nP17, nP19, nP22, nP24, nP25, nP26, nP27, nP28, nP29, nP30, nP32*) mapped to the right of (or some of them possibly to the left and very near to) *sma-5(+ Berg)*. The 19 loci to the right of *sma-5* were subdivided into four clusters, and the four loci to the left of *dpy-6* were subdivided into two clusters (Figures 2 and 3) by unselected second recombination events that caused the simultaneous appearance or disappearance of these sets of loci. One unselected recombination event separated *nP20* and the set *nP7, nP9*, and *nP31* into two clusters of loci as shown. The 19 Tc1-containing *Eco*RI fragment loci mapped to the right of *sma-5* as shown by Figure 3 strain 1. The locations of *lin-2* and *unc-9* relative to these 19 Tc1-containing *Eco*RI fragment loci were not determined, as indicated by the bracket surrounding these Tc1 loci. For the Tc1 loci between *dpy-6* and *sma-5*, the map distance of each Tc1 locus (or linked cluster of loci) from visible genetic markers and from each other was estimated from the proportion of the independent recombinants in the genetic interval that acquired each dimorphic Tc1 locus, as in any three-factor genetic mapping. For example, three of the 25 Lin-14 non-Dpy-6 recombinants acquired both Tc1-containing *Eco*RI fragment loci *nP15* and *nP23*, allowing them to be mapped as a cluster 3/25 × 7 map units or about 0.8 map unit to the left of *lin-14*. Outside of the *dpy-6* to *sma-5* interval, relative map distances were estimated by the frequency of second unselected recombination events. Because all of these unselected recombinations occurred in strains selected to have a recombination event in the *dpy-6* to *sma-5* region, interference could affect the recombination frequencies we have measured. Nonetheless, the order of these markers established by these data is unambiguous. C, Fine-structure genetic mapping of the *lin-14*-linked Tc1-containing *Eco*RI fragment loci using a *lin-14* intragenic recombinant: a strain was constructed that contained on one chromosome the *lin-14(+ Berg)* region of Bergerac with the flanking Tc1-dimorphic and visible genetic loci: *dpy-6(e14 Bris) (nP15 nP23)(nP1 nP11 nP21)(nP3 nP8 nP13 lin-14(+ Berg)) nP18 unc-9(e101 Bris)* (the Bergerac region is indicated with a thick line), and on the other chromosome a doubly mutant *lin-14* gene from Bristol: *lin-14(n540 n536)* (shown with a thin line below). Tc1-dimorphic loci listed above the bracketed intervals on the map and within parentheses as described here were genetically unordered relative to each other before this experiment. One recombination event between *n536* and *n540* in *trans* to the *lin-14(+ Berg)* chromosome was detected as an animal with a dominant Lin-14 phenotype. This *lin-14* intragenic recombinant strain retained the Bergerac alleles of the *lin-14*-linked Tc1-containing *Eco*RI fragment loci *nP1, nP8, nP11, nP13, nP15, nP21*, and *nP23*, and the *dpy-6(e14 Bris)* marker but had segregated away the Tc1-containing Bergerac allele of the *nP3* and *nP18* loci and the *unc-9(e101 Bris)* marker. These data mapped *nP3* and *n536* to the right of this intragenic recombination event, although their order was not defined from these data. The data also mapped *nP8, nP13*, and *n540* to the left of the intragenic recombination event and to the left of *nP3* and *n536*. The order of *nP8, nP13* and *n540* could not be determined from these data.

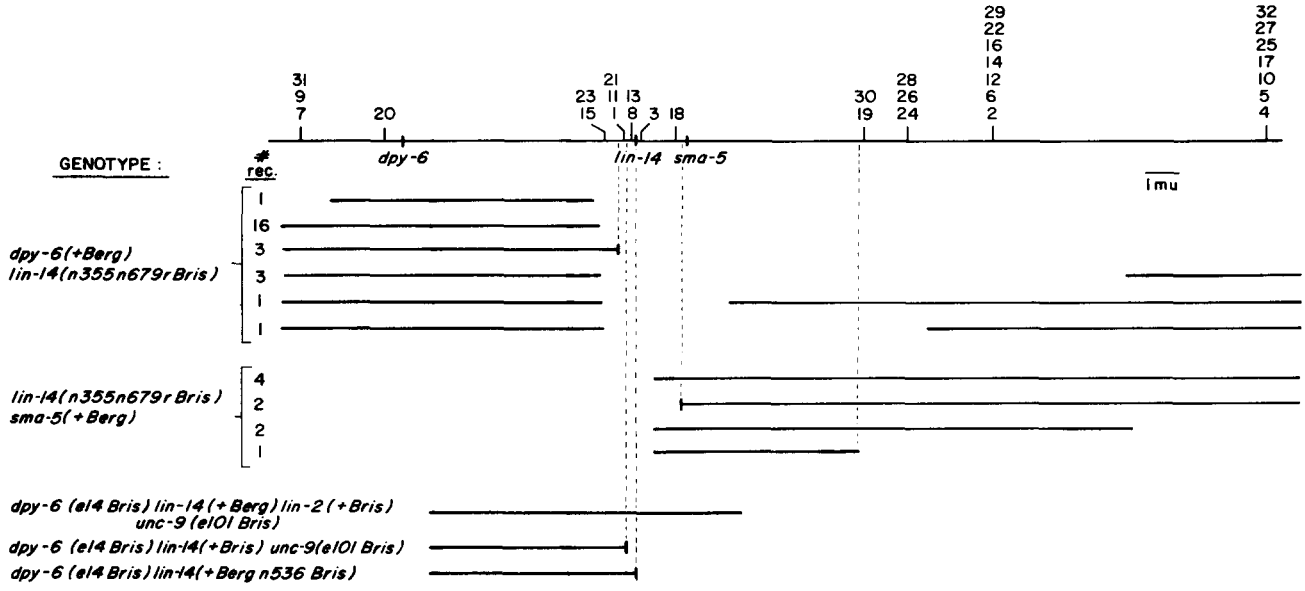


FIGURE 3.—Genetic map of the dimorphic Tcl loci near *lin-14* showing the recombinant strains used to construct the map. The genotypes of each of the recombinant strains as determined by both the visible genetic markers and the Tcl-containing *Eco*RI fragment loci they carry is summarized below the map. Each line below the genetic map indicates that a recombinant strain(s) carried the corresponding Bergerac Tcl-containing alleles of each locus where a line is shown and the Bristol alleles where no line is shown. Lin-14 non-Dpy-6 recombinants retain the Bergerac region to the left and Lin-14 non-Sma-5 recombinants retain the Bergerac region to the right of *lin-14*. The locations of unselected second recombination events that occurred outside the *dpy-6* to *sma-5* interval are indicated by the disconnected second line corresponding to the additional Bergerac alleles of the Tcl-dimorphic loci that appeared in those strains shown. Certain genotypes were independently generated multiple times and are indistinguishable based on their visible and Tcl-containing *Eco*RI fragment loci phenotypes. The numbers of these independently derived identical recombinant strains are recorded in the column marked "# rec." Each set of Tcl-containing *Eco*RI fragment loci that coordinately appeared and disappeared in recombinants, that is were not separated from each other by any of the 47 recombination events used in this mapping, are shown on the map as an unseparated cluster. The visible and Tcl genotypes of three other recombinant strains used in this study are indicated below the genetic map. The *dpy-6(e14 Bris) lin-14(+ Berg) lin-2(+ Bris) unc-9(e101 Bris)* strain #1 was used as the source of *lin-14*-linked Tcl-containing *Eco*RI fragments in the intragenic recombination experiment described in Figure 5. The *dpy-6(e14 Bris) (nP15 nP23)(nP1 nP11 nP21) lin-14(+ Bris) unc-9(e101 Bris)* strain #2 allowed the separation of the *nP1*, *nP11*, and *nP21* cluster of Tcl-dimorphic loci from the *nP8*, *nP13*, and *nP3* cluster of *lin-14*-linked Tcl-dimorphic loci. The *dpy-6(e14 Bris) lin-14(+ Berg n536sd Bris)* strain #3 was isolated in the intragenic recombination experiment described in Figures 2 and 5.

other and to *lin-14*, *dpy-6*, and *sma-5* (Figure 3) we generated the map shown in Figure 2B.

A total of nine Tcl-dimorphic loci mapped between *dpy-6* and *sma-5* and thus close to *lin-14* on the Bergerac chromosome. Two of these loci (*nP15* and *nP23*) were present in some but not all Lin-14 non-Dpy-6 recombinants and in no Lin-14 non-Sma-5 recombinants, indicating a map position between *dpy-6* and *lin-14*, and one locus (*nP18*) was present in some but not all Lin-14 non-Sma-5 recombinants and no Lin-14 non-Dpy-6 recombinants, indicating a map position between *lin-14* and *sma-5*.

The six remaining Tcl-dimorphic loci (*nP1*, *nP3*, *nP8*, *nP11*, *nP13*, *nP21*) mapping in the *dpy-6 sma-5* interval were completely linked to *lin-14* at the level of resolution of this experiment. None of the 34 recombinants we collected in the 8.5-map unit *dpy-6 sma-5* interval separated those six Tcl-containing *Eco*RI fragment loci from the *lin-14*(+ *Berg*) locus or from each other. Thus, these six Tcl-dimorphic loci must be linked to *lin-14* and to each other within about 0.3 map unit. Thus in one pedigree, using one DNA probe, parallel RFLP mapping identified the six Tcl-dimorphic loci, out of 400 in the genome, most closely linked to *lin-14*.

This cluster of six linked Tcl-dimorphic loci was further divided into two clusters each containing three Tcl-dimorphic loci (*nP1*, *nP11*, *nP21* to the left and *nP3*, *nP8*, and *nP13* to the right) by the isolation of a fortuitous recombination event within the cluster of *lin-14*-linked Tcl-dimorphic loci (Figure 3, recombinant #2). However, because no *lin-14* mutant allele was present in the parent chromosomes of this recombinant, the recombination event did not map these loci relative to the *lin-14* gene.

Fine-structure RFLP mapping of the *lin-14* gene using a *lin-14* intragenic recombinant: To establish more precisely the location of the *lin-14* gene relative to the six Tcl-dimorphic loci most closely linked to *lin-14*, we performed a genetic screen to detect a *lin-14* intragenic recombination event between a Bristol chromosome carrying two *lin-14* mutations and a chromosome carrying the *lin-14*(+ *Berg*) gene and flanking Tcl-dimorphic loci. This recombination event created a hybrid Bristol/Bergerac *lin-14* gene in which the recombination point, and thus the *lin-14*

gene itself, could be mapped relative to Bristol/Bergerac RFLP loci.

The Bristol chromosome used in this experiment contained both a semidominant *lin-14* mutation (*n536sd*) and a loss-of-function *lin-14* allele (*n540*) that acts in *cis* to suppress *n536sd* (AMBROS and HORVITZ 1987). This doubly mutant *lin-14(n536sd n540)* chromosome was generated by mutating the semidominant retarded *lin-14(n536sd)* allele to a recessive *lin-14* allele; the new allele failed to complement other recessive *lin-14* alleles and resulted in the same precocious phenotype (AMBROS and HORVITZ 1987). This suppressor mutation, *lin-14(n540)*, is recessive: *lin-14(n536sd)/lin-14(n536sd n540)* animals display the semidominant Lin-14 retarded phenotype, whereas *lin-14(n536sd n540)/lin-14(n536sd n540)* animals display the recessive Lin-14 precocious phenotype (AMBROS and HORVITZ 1987). Thus, among the progeny of a *lin-14(+ Berg)/lin-14(n536sd n540 Bris)* strain, a *lin-14* intragenic recombination event separating the semidominant *n536sd* mutation from its recessive suppressor mutation *n540* would be detected by the appearance of a rare animal displaying the *n536sd* retarded phenotype. The semidominant nature of the *n536sd* mutation allowed such a recombinant to be detected in the first generation after the recombination event.

The source of Bergerac alleles of RFLP loci to be followed in the cross was a derivative of the congenic *lin-14(+ Berg)* chromosome, *dpy-6(e14 Bris)* (*nP15 nP23*)(*nP1, nP3, nP8, nP11, nP13, nP21, lin-14*(+ *Berg*))(*nP18*) *unc-9(e101 Bris)*, carrying all of the Bergerac Tc1-containing *Eco*RI fragment loci mapping within about 8 map units of *lin-14* plus the flanking visible genetic markers *dpy-6* and *unc-9* (Figure 3).

One recombinant animal of genotype *dpy-6(e14) lin-14(+ Berg n536sd)/lin-14(n536sd n540)* was detected after screening 10^4 to 10^5 *dpy-6(e14 Bris)*(*nP15 nP23*)(*nP1, nP3, nP8, nP11, nP13, nP21, lin-14*(+ *Berg*))(*nP18*) *unc-9(e101 Bris*))*/lin-14(n536sd n540)* progeny. DNA was prepared from the recombinant homozygous *dpy-6(e14) lin-14(n536sd)* strain and probed with ^{32}P-labeled Tc1 DNA to ascertain which of the six *lin-14*-linked Tc1-containing Bergerac loci were retained in the strain. This strain retained five of the six Tc1-containing *Eco*RI fragment loci mapping closest to *lin-14*; it was missing Tc1-containing *Eco*RI fragment *nP3*, and therefore contained the Bristol allele of this *Eco*RI fragment *nP3(Bris)*. In addition, the strain was missing the Tc1-containing Bergerac allele of *Eco*RI fragment *nP18* that we had mapped using the Lin-14 non-Sma-5 recombinants to the right of *lin-14*, while it retained Tc1-containing *Eco*RI fragments *nP1, nP8, nP11, nP13, nP15, nP21*, and *nP23*. This experiment mapped the *lin-14* recessive allele *n540* and the *lin-14*-linked Tc1-containing *Eco*RI fragments *nP1, nP8, nP11, nP13*, and *nP21* to the left of the *lin-14* semidominant mutation *n536sd*, and the *lin-14*-linked Tc1-containing *Eco*RI fragment *nP3* to the right of (or near and to the left of) the *lin-14* semidominant mutation *n536sd* (Figure 2C).

Thus, the *lin-14* intragenic recombinant mapping data further divided the closest *lin-14*-linked Tc1-dimorphic cluster, *nP3, nP8*, and *nP13*, with Tc1-dimorphic locus *nP3* mapping closest to the right of the *lin-14* intragenic recombination point and with one of the Tc1-dimorphic loci *nP8* or *nP13* mapping closest to the left (Figure 3). These data suggested that the *lin-14* locus could be cloned by isolating overlapping clones spanning these three Tc1-dimorphic loci and precisely mapping within this cloned region the location of the *lin-14* intragenic recombination point. This recombination point, marking the location of the *lin-14* gene, could be mapped by detecting non-Tc1-associated RFLP loci that flank the recombination point based on their segregation behavior in the *lin-14* intragenic recombinant strain.

Cloning the *lin-14*-linked Tc1-containing *Eco*RI fragments and identifying a *lin-14* contig: We cloned the *nP3, nP8*, and *nP13* Tc1-containing *Eco*RI fragments most closely linked to *lin-14* by screening with a Tc1 DNA probe size-selected *Eco*RI fragment libraries made from the *lin-14(+ Berg)* strain. Unique DNA flanking the Tc1 insertion sites was purified, ^{32}P-labeled, and used as probe to isolate larger overlapping genomic clones from *C. elegans* strain Bristol cosmid libraries (G. BENIAN, personal communication; COULSON *et al.* 1986). These cosmid clones were identified by their "fingerprints" and assigned to a large contig containing overlapping cosmid and yeast artificial chromosome clones (COULSON *et al.* 1986, 1988; BURKE, CARLE and OLSON 1986).

In this way, the cosmid clones corresponding to the genetic loci *nP3, nP8*, and *nP13* were placed on a single 830-kb contig, shown in Figure 4. The genetic mapping of *nP13* and *nP8* to the left and *nP3* to the right orients the physical map relative to the genetic map: *nP13 nP8 nP3*. The physical mapping orders the loci by placing *nP8* in the middle. The physical mapping confirms and extends the genetic mapping of the Tc1-dimorphic loci and yields the physical distances between them, as *nP8* is about 300 kb from both *nP13* and *nP3*. Because the *lin-14* intragenic recombination event separated the *nP13* and *nP8* loci cluster from the *nP3* locus, at least part of the *lin-14* gene must lie in the approximately 300 kb of cloned DNA between *nP8* and *nP3*.

Detection and mapping of non-Tc1-associated Bristol/Bergerac RFLPs flanking *lin-14*: The *lin-14* intragenic recombinant strain contained Bergerac DNA sequences surrounding the closest Tc1-dimorphic locus *nP8* on the left, and Bristol DNA

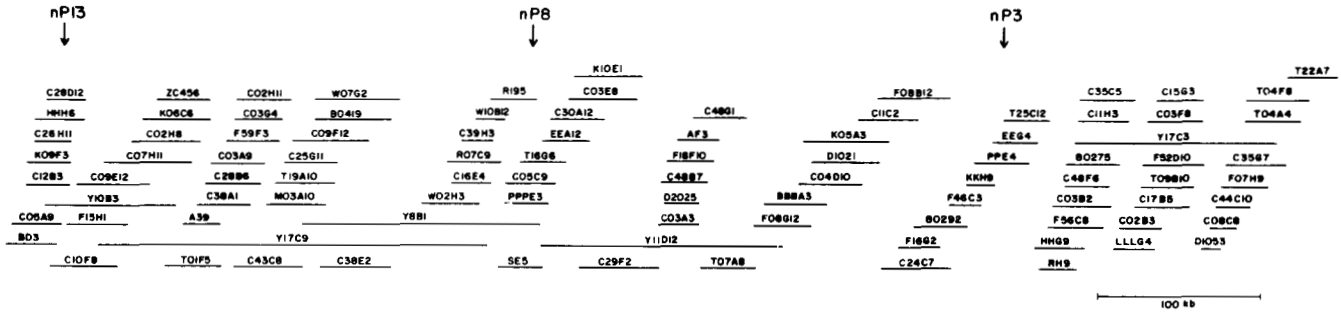


FIGURE 4.—Overlapping cosmid clone and yeast artificial chromosome clone contig from the *lin-14*-linked Tc1-dimorphic loci. Shown are the DNA clones isolated near the *lin-14*-linked Tc1-dimorphic loci *nP3*, *nP8*, and *nP13*. Each cosmid clone is indicated by a horizontal line with the alphanumeric clone name above the line. Yeast artificial chromosome (YAC) clone names begin with a *Y* and are longer. Overlaps between cosmids were determined by "fingerprinting" as described (COULSON *et al.* 1986) and in some cases by DNA hybridization as well. Overlaps between cosmid clones and YACs were determined by DNA hybridization. While the methods by which the set of clones has been assembled do not allow precise calculation of physical distances between clones, these were estimated based on known sizes of YACs and estimated sizes of cosmids. This 830-kb contiguous stretch of cloned DNA shows that *nP8* is about 300 kb to the left of *nP3*. Thus the site of the *lin-14* intragenic recombination point and at least part of the *lin-14* gene must lie in this 300 kb between *nP8* and *nP3*. The *nP13* locus maps about 300 kb to the left of *nP8* on this contig.

sequences surrounding the closest Tc1-dimorphic locus *nP3(Bris)* on the right. At some point between *nP8* and *nP3*, the *lin-14* intragenic recombinant *X* chromosome must shift from Bristol-derived DNA sequences to Bergerac-derived DNA sequences. This Bristol-Bergerac transition point marks the location of the *lin-14* gene and therefore was mapped precisely. To find this point, cosmids from the region between *nP8* and *nP3* were used to search for Bristol/Bergerac non-Tc1-associated RFLPs, and the presence or absence of these RFLPs in the intragenic recombinant strain was assessed. The intragenic recombinant strain was expected to carry the Bergerac allele of any RFLP detected with a clone from the contig mapping to the left of the *lin-14* intragenic recombination point and the Bristol allele of any RFLP detected with a clone mapping to the right.

Cosmids from both sides of *nP3* in this contig were probed to Southern blots of restriction digests of Bristol and Bergerac DNAs, and RFLPs were detected: cosmid KKH9 to the left detects an *Xho*I site Bristol/Bergerac RFLP *nP34*, cosmid HHG9 to the right detects an *Eco*RI site Bristol/Bergerac RFLP *nP35*, and cosmid EEG4 from the region between cosmids KKH9 and HHG9 detects a Bristol/Bergerac *Eco*RI RFLP *nP33* (as well as the Tc1-associated RFLP *nP3*). The intragenic recombinant strain contained the Bergerac alleles of RFLP loci *nP34* and *nP33* to the left, and the Bristol alleles of RFLP loci *nP3* and *nP35* to the right (Figure 5). The *nP33* and *nP3* RFLP loci thus flank the *lin-14* intragenic recombination point. Because these RFLP loci correspond to sequences both located on cosmid EEG4, this cosmid must contain the site of the *lin-14* intragenic recombination event and at least part of the *lin-14* gene. The restriction map of this cosmid was determined, and the *nP33* RFLP was found to be separated by about 27 kb from the *nP3* RFLP (Figure 7).

Detection of *lin-14* allele-specific DNA alterations: We used cosmid EEG4, as well as flanking cosmids from the region, to probe Southern blots of DNAs from 20 strains containing independently isolated *lin-14* alleles and 10 non-*lin-14* strains. We detected DNA alterations associated with both of the existing dominant *lin-14* mutations, *n536*sd and *n355*sd, and with two of the 18 *lin-14* recessive mutations tested, *n360* and *n407* (Figures 6 and 7).

The only two recessive alleles isolated after γ-ray mutagenesis, *n360* and *n407*, alter the same 2.3-kb *Eco*RI fragment and the same 5.2-kb *Hin*dIII fragment located about 18 kb to the left of *nP3* (Figures 6 and 7). Hybridization of ^{32}P-labeled 2.3-kb and 2.0-kb *Eco*RI fragments derived from cosmid EEG4 to *Hin*dIII-digested DNAs from these strains results in disappearance of the normally 5.2-kb *Hin*dIII hybridization band in *n407* strains (Figures 6 and 7) and replacement of the 5.2-kb *Hin*dIII band with a 12-kb *Hin*dIII hybridization band in *n360* strains (data not shown). Hybridization of the same probe to *Eco*RI-digested DNA isolated from a strain containing *lin-14(n360)* results in no 2.3-kb *Eco*RI fragment hybridization band (data not shown). Using the same probe, a weak 2.8-kb *Eco*RI hybridization band replacing the normally 2.3-kb *Eco*RI hybridization band can be observed in DNA isolated from a strain containing *lin-14(n407)* (data not shown). All other *Hin*dIII or *Eco*RI hybridization bands detected with ^{32}P-labeled cosmid EEG4 probe are normal in these strains. These data suggest that the *n360* and *n407* mutations are both associated with rearrangements with an end point in the 2.3-kb *Eco*RI fragment.

Using cosmid KKH9 as a probe, another allele-specific DNA alteration associated with the *lin-14* recessive allele *n360* was observed about 30 kb to the

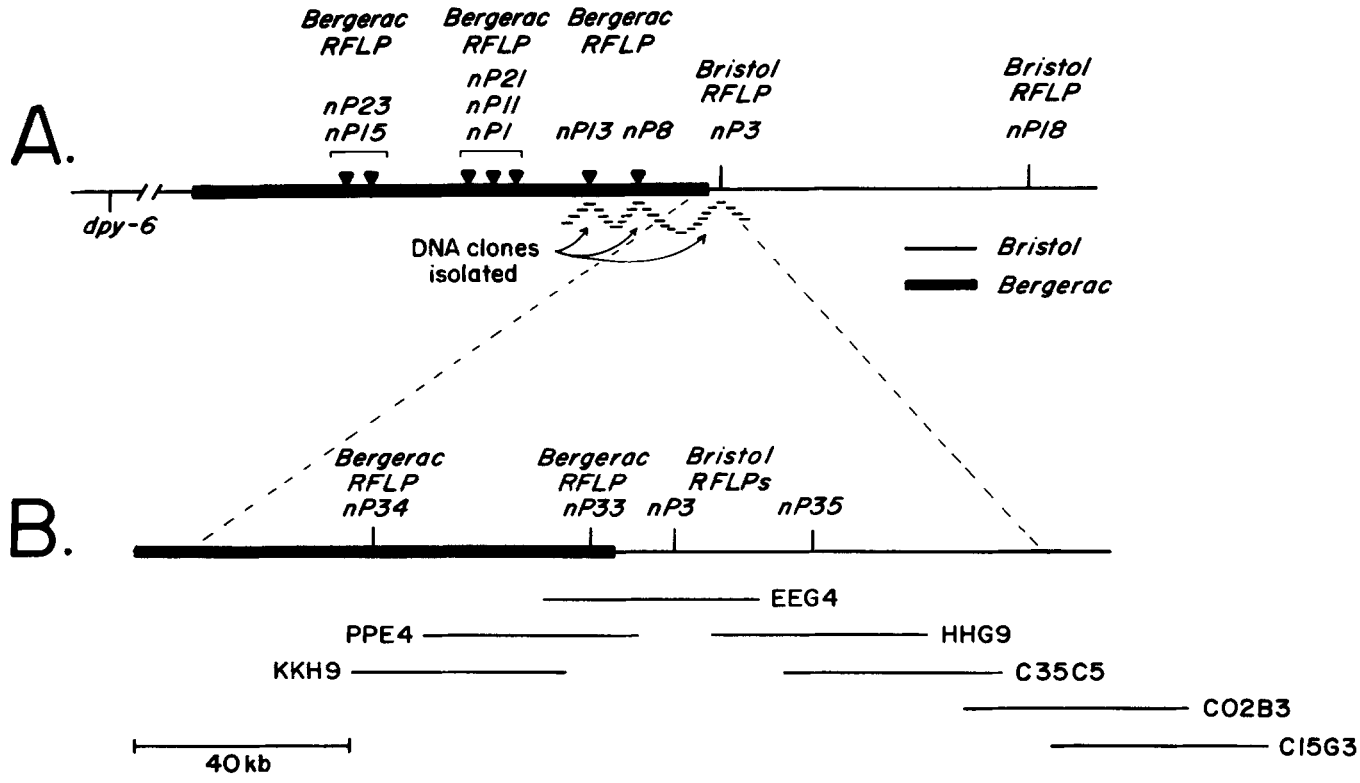


FIGURE 5.—A, Fine-structure physical genetic mapping of the *lin-14*-linked RFLP loci using the *lin-14* intragenic recombinant chromosome. A map of the *lin-14* region of the *lin-14* intragenic recombinant strain is shown above. Chromosomal regions derived from Bergerac are shown in thick lines and those derived from Bristol are shown in thin lines. The *dpy-6(e14 Bris)* mutation and flanking Bristol genetic regions shown to the left of the Bergerac region on this chromosome were derived from the original chromosome in the parental strain from which the *lin-14* intragenic recombinant was isolated. The intragenic recombinant strain contains the Bergerac alleles of the Tc1-dimorphic loci *nP1*, *nP8*, *nP11*, *nP13*, *nP15*, *nP21*, *nP23* and the Bristol alleles of Tc1-dimorphic loci *nP3* and *nP18*. RFLP *nP18* maps between *lin-14* and *sma-5*, about 1.3 map units to the right of *lin-14* (Figure 2). Thus *nP3* must be the closest *lin-14*-linked Tc1-containing *Eco*RI fragment locus on the right of the *lin-14* intragenic recombination point. The other two *lin-14*-linked Tc1-containing *Eco*RI fragment loci *nP8* and *nP13* mapped to the left of the intragenic recombination point. B, Physical genetic map of the *lin-14* region near the *lin-14* intragenic recombination point. The physical genetic map of the region around *nP3* is shown in an expanded view below. Cosmid EEG4 was isolated using the cloned *nP3(Bris)* *Eco*RI fragment as a probe and assigned to the contig as described in the text. Bristol/Bergerac RFLPs were detected using ^{32}P-labeled cosmids KKH9, PPE4, EEG4, and HHG9 as probes to Southern blots of DNAs isolated from strain Bristol N2 and strain Bergerac RW7000 and digested with various restriction enzymes. The intragenic recombinant strain contained the Bergerac alleles of RFLP *nP34*, detected by KKH9 and RFLP *nP33*, detected by PPE4 and EEG4, and the Bristol alleles of RFLPs *nP3*, detected by EEG4, and *nP35* detected by HHG9. These data mapped the *lin-14* intragenic recombination point to a 27-kb region cloned on cosmid EEG4.

left of the 2.3-kb *Eco*RI fragment *lin-14(n360)* DNA alteration (data not shown). It is possible that these two *n360*-associated changes are inversion endpoints. Although either or both of these mutational changes could be the cause of the *lin-14* phenotype, we suspect that the *lin-14(n360)*-associated DNA alteration in the 2.3-kb *Eco*RI fragment detected by cosmid EEG4 is the cause of the *lin-14* phenotype, since the recessive *lin-14* allele *n407* maps to the same *Eco*RI fragment.

The DNA alterations associated with the two semi-dominant alleles map about 15 kb to the right of the recessive alleles and alter the same 3.8-kb *Eco*RI (right) fragment (Figures 6 and 7). The *n536*sd mutation, isolated after EMS mutagenesis (AMBROS and HORVITZ 1984), is associated with a simple deletion of about 600 bp from the 3.8- kb *Eco*RI (right) fragment. Digestion of DNA isolated from strains bearing the *n536*sd mutation with the restriction enzymes *Bgl*II, *Hin*dIII, *Eco*RI, or *Xho*I results in a fragment about 600 bp shorter than that from wild type on Southern blot analysis using ^{32}P-labeled 3.8-kb *Eco*RI (right) fragment probe (Figures 6 and 7, data not shown). The *n355*sd allele, isolated after γ-ray mutagenesis, is associated with a rearrangement of the 3.8-kb *Eco*RI (right) fragment. Using ^{32}P-labeled 3.8-kb *Eco*RI (right) fragment probe, DNA isolated from strains containing *n355* and analyzed on Southern blots yielded the following DNA alterations: digestion with *Hin*dIII results in a 9-kb rather than 6.2-kb band (Figure 6), digestion with *Eco*RI results in a 10-kb rather than 3.8-kb band (data not shown), and digestion with *Bgl*II results in a 7.3-kb rather than 7.6-kb band (data not shown). These data suggest that the *n355* mutation is associated with a rearrangement affecting the 3.8-kb (right) *Eco*RI fragment. All other adjacent hybridization bands using ^{32}P-labeled cosmid

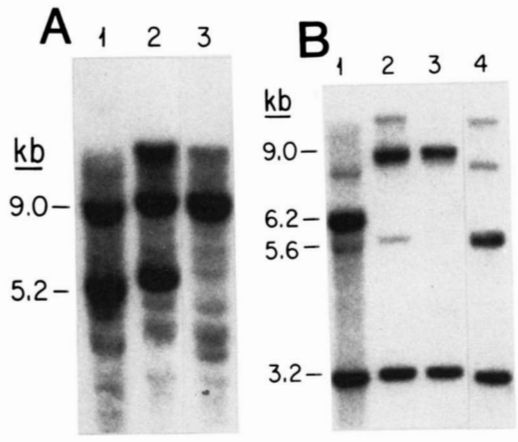


FIGURE 6.—DNA alterations associated with *lin-14* alleles. Shown are Southern blots of *Hind*III-digested DNAs isolated from the wild type (N2) and three mutant strains. In panel A this blot is probed with ^{32}P-labeled 2.0 and 2.3 kb *Eco*RI fragments from the 5′ end of the *lin-14* gene. In panel B, the blot is probed with ^{32}P-labeled 1.9-kb and 3.8-kb (right) *Eco*RI fragments from the 3′ end of the gene. Lane 1 = wild-type N2, Lane 2 = *n355*, lane 3 = *n355n407*, lane 4 = *n536*. Only the *n355n407* strain (lane 3) is altered from wild type as seen using the 5′ probe in A; the normal 5.2-kb *Hind*III fragment is not visible. Because the *n355* strain is like wild type in this region (lane 2), the change observed in *n355n407* must be due to the *n407* mutation. As seen using the 3′ probe in B, strains containing the *n355* mutation (lane 2 and 3), replace the normal 6.2-kb *Hind*III hybridization band with one at 9 kb. The 6.2-kb *Hind*III fragment is changed to 5.6 kb in the *n536* strain (lane 4). The adjacent 3.2 kb *Hind*III band is normal in all the strains shown. The weaker unmarked bands on this blot are due to slight contamination of *C. elegans* DNA with *E. coli* DNA that hybridizes to sequences in the Bluescribe cloning vector.

EEG4 probe were normal in DNA isolated from strains containing *n536*sd or *n355*sd.

Using cosmid EEG4 as a probe, no changes in the restriction map of the *lin-14* region were detected in any of 10 *lin-14*(+) strains examined. Therefore the changes in this region we see in the *lin-14* mutants are not simply due to a highly polymorphic or mutable region.

The clustering of four *lin-14* allele-specific DNA alterations over an 18-kb region argues that these are caused by the corresponding *lin-14* mutations. Three mapping criteria further correlate the physical genetic alterations we detect with the genetic locations of the four *lin-14* mutations: (1) The *lin-14* intragenic recombinant genetically maps the *lin-14*(*n536sd*) mutation to the right of this recombination point. The 600-bp deletion associated with *lin-14*(*n536sd*) maps physically to the right of the *lin-14* intragenic recombination point. (2) The only other semidominant *lin-14* mutation, *n355sd*, is associated with a rearrangement that maps to the same 3.8-kb *Eco*RI fragment as *n536sd*. (3) The *lin-14* intragenic recombinant strain genetically maps the *lin-14* recessive allele *n540* to the left of semidominant *lin-14* allele *n536sd*. While no DNA alteration associated with *n540* has been detected, this mutation is allelic with the recessive *lin-14*(*n360*) mutation, which is associated with a DNA alteration that physically maps to the 2.3-kb *Eco*RI fragment to the left of *lin-14*(*n536sd*).

The outer boundaries of the *lin-14* gene have not been determined by our RFLP mapping.

Detection of *lin-14* transcripts: The cosmid EEG4 was used to probe Northern blots of RNA isolated from wild-type *C. elegans* strain Bristol N2. Two major mRNA species, one of about 7 kb (data not shown) and one of 3.5 kb were detected using this probe (Figure 8). Using single-stranded DNA probes carrying restriction fragments subcloned from cosmid EEG4, we found that the 3.5-kb mRNA is detected by subclones of the 2.3-kb, 1.9-kb, and 3.8-kb (right) *Eco*RI fragments and is transcribed left to right on the genetic map. These genomic regions span about 18 kb (Figure 7), suggesting that one or more introns are present in this region. No other probes from the 45-kb region shown in Figure 7 detected the 3.5-kb transcript (data not shown).

The fact that the 2.3-kb or 3.8-kb (right) *Eco*RI fragments that hybridize to the 3.5-kb mRNA are altered in strains bearing the *lin-14* mutations in *n360*, *n407*, *n355*sd and *n536*sd argues that this mRNA is a *lin-14* mRNA. In addition, the *lin-14* intragenic recombination point between *n536*sd and *n540* maps to a region (between *nP33*, just to the left of the 2.3-kb *Eco*RI, and the 600-bp deletion associated with *n536*sd in the 3.8-kb (right) *Eco*RI fragment on the right) that is nearly congruent with the locations of the exons of the 3.5-kb mRNA (Figure 7).

The 7-kb mRNA is transcribed left to right on the genetic map and is detected only by subclones of the 15-kb *Bgl*II fragment to the right of *nP3* (Figure 7). This transcript therefore maps to the right of the *lin-14* intragenic recombination point and to the right of all four physically mapped *lin-14* mutations.

***lin-14* transcripts in *lin-14* mutant strains:** Changes in size of the normally 3.5-kb wild-type *lin-14* transcript detected on Northern blots using ^{32}P-labeled restriction fragment probes isolated from cosmid EEG4 were observed in RNAs isolated from two *lin-14* mutant strains. Using ^{32}P-labeled 3.8-kb *Eco*RI (right), 1.9-kb *Eco*RI, or 2.3-kb *Eco*RI fragment probes, we found the 3.5-kb mRNA decreased in size to about 3.0 kb in all strains containing the *lin-14* semidominant allele *n536*sd (Figures 7 and 8). All strains containing the *lin-14* semidominant allele *n355*sd replace the 3.5-kb mRNA with two mRNAs of 3.0 kb and 2.2 kb, with the 2.2-kb mRNA about four times as intensely hybridizing to these probes as the 3.0-kb mRNA (Figures 7 and 8). These shorter 3.5-kb-related mRNAs must have deletions of RNA sequences normally present in the *lin-14* mRNA (Figure 7). Because these mRNAs from the semidominant *lin-14* mutants hybridize to the same three DNA

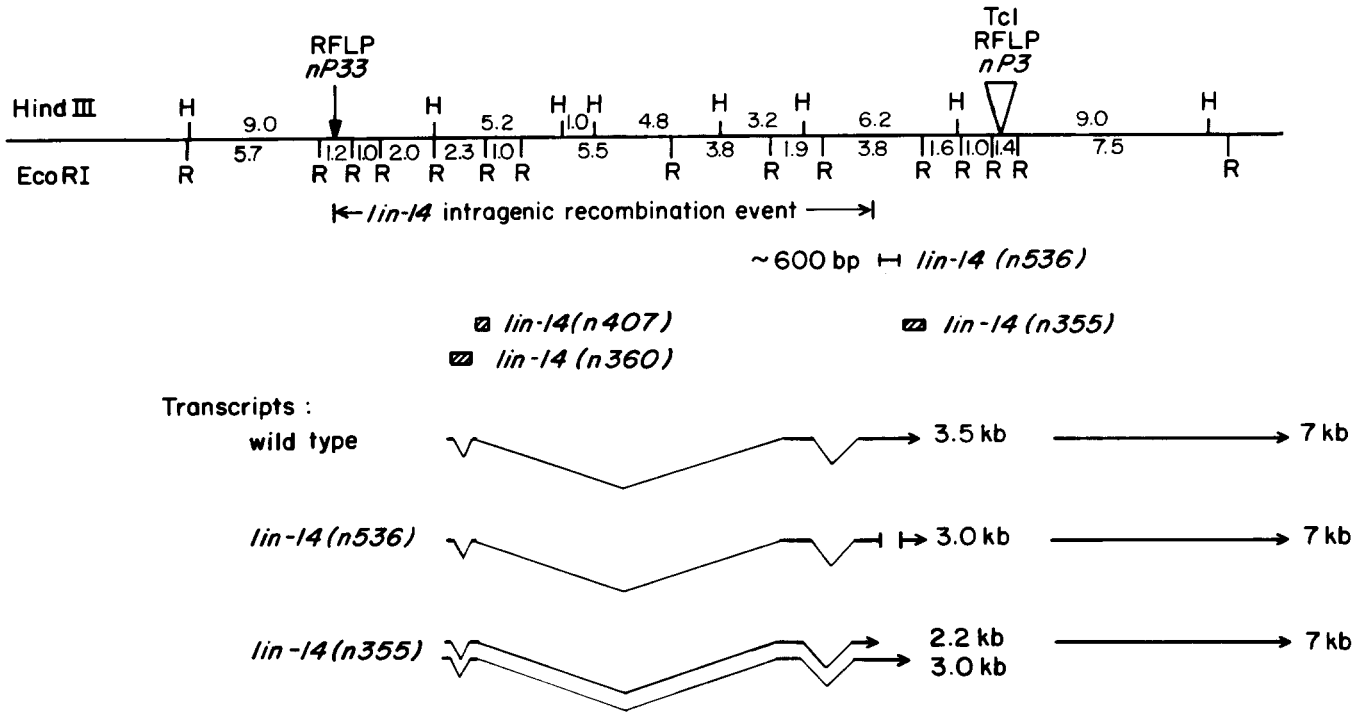


FIGURE 7.—Physical genetic map of *lin-14* region showing locations of mutations and transcripts in the wild type and *lin-14* dominant mutants. Genetically mapped Bristol/Bergerac RFLPs *nP33* and *nP3* are shown above the line. The locations of Bristol/Bergerac RFLPs *nP33* and *nP3*, and *lin-14* allele-specific DNA alterations *n360*, *n407*, *n355*, and *n536* were mapped by probing with ^{32}P-cosmid EEG4 Southern blots of genomic DNAs from strains containing these *lin-14* alleles or *lin-14*(+ *Berg*) to detect which restriction fragments were altered from the wild-type Bristol pattern. These restriction fragments were located on the restriction maps of cosmids in the region. The mutations detected in *lin-14*(*n360*) strains and *lin-14*(*n355n407*) strains are rearrangements that both map to the same 2.3-kb *Eco*RI fragment but most likely do not overlap. The *lin-14*(*n536*sd) and *lin-14*(*n355*sd) mutations are both associated with DNA alterations that map to the same 3.8-kb *Eco*RI (right) fragment. The *n536*sd mutation is a simple deletion of 600 bp. The *n355*sd mutation is a rearrangement. Transcripts were detected using ^{32}P-labeled cosmid EEG4 as a probe to Northern blots of poly(A)$^+$-selected mRNAs from the wild-type N2 strain and the two semidominant mutants strains shown. The genomic regions from which these transcripts are derived were determined by hybridizing *Eco*RI fragment and *Bgl*II fragment subclones from cosmid EEG4 to Northern blots of these same RNAs. All cosmid EEG4 *Bgl*II restriction fragments (data not shown) and *Eco*RI fragments 2.3 kb, 3.8 kb (left), 1.9 kb, and 3.8 kb (right) were tested. Only those restriction fragments corresponding to exonic sequences on the mRNAs as shown in the figure were observed to hybridize to the mRNAs. This procedure would miss micro-exons. The presence of introns between the 1.9-kb and 3.8-kb (right) *Eco*RI fragments, between the 2.3-kb and 1.9-kb *Eco*RI fragments, and within the 2.3-kb *Eco*RI fragment were inferred from hybridization of *lin-14* cDNA clones to genomic clones and partial DNA sequences of *lin-14* cDNA clones (G. RUVKUN, J. GIUSTO, J. GATTO and T. BÜRGLIN, unpublished observations). The locations of the 5′ end, splice sites, and 3′ end of the *lin-14* mRNAs shown are approximate. The direction of transcription of both the 7-kb and 3.5-kb-related transcripts was determined using subclones containing the 5.0-kb *Bgl*II, 7.6-kb *Bgl*II, and 15-kb *Bgl*II fragments inserted in both orientations into the single-stranded cloning vector Bluescribe. Only one orientation of each subclone hybridized to these mRNAs from the *lin-14* region. The location of the 600-bp deletion in the *lin-14* mRNA associated with *n536* maps to exons derived from the 3.8-kb *Eco*RI (right) fragment but has not been more precisely mapped within this fragment. The *n355*-associated rearrangement of the 3.8-kb *Eco*RI fragment results in two shorter *lin-14* mRNAs of 3.0 and 2.2 kb. Both of these mRNAs contain exonic sequences derived from the 3.8-kb (right), 1.9-kb, and 2.3-kb *Eco*RI fragments and do not inappropriately hybridize to other restriction fragments from the *lin-14* region. We assign the region deleted in these shorter *lin-14* mRNAs to the 3.8-kb *Eco*RI (right) fragment because the *n355*sd mutation affects only this exon-containing *Eco*RI fragment. We have not determined the effects of the *n360* or *n407* mutations on the 3.5-kb *lin-14* transcript.

probes as the wild-type *lin-14* mRNA, they, at least grossly, contain similar sequences 5′ to the dominant mutations (data not shown). We assume that the *lin-14*(*n536*) mutant mRNA has a simple 600-bp deletion because it is approximately 600 bases shorter than the wild-type *lin-14* mRNA and the genomic 3.8-kb *Eco*RI (right) fragment is also reduced in size by about 600 bp in this mutant (Figure 6). The *n355sd* rearrangement mutation results in two shorter mRNAs of 3.0 kb and 2.2 kb. This mutation could add or expose variable transcriptional splice sites, terminator sites, or polyadenylation sites to the *lin-14* mRNA. While we can be sure that both shorter *lin-14* mRNAs from these semidominant alleles are missing some sequences present on the wild-type *lin-14* mRNA, we do not know if these mutant mRNAs contain any additional sequences. To examine this possibility, we are currently determining the DNA sequences of *lin-14* cDNA clones isolated from both the wild type and *lin-14* semidominant mutants.

The 3.5-kb related mRNA is shorter in strains carrying either *n536sd* or *n355sd* regardless of whether there is an additional *cis*-acting recessive suppressor mutation that reverts the dominant mutation,

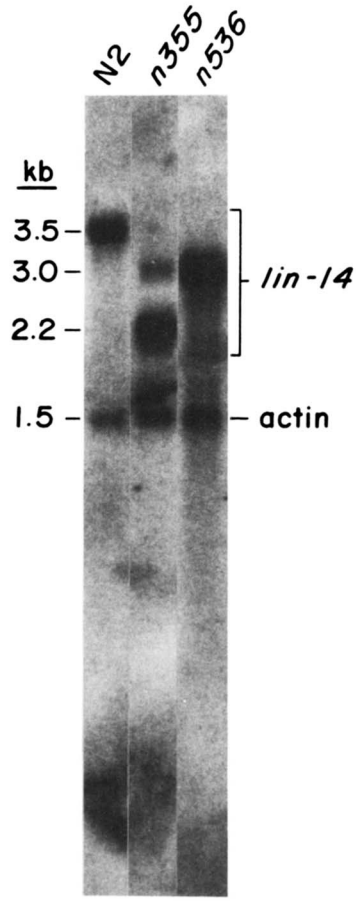


FIGURE 8.—Transcription of *lin-14* in the wild-type and *lin-14* mutants. Northern blot analysis of *lin-14* mRNA. The 3.8-kb *Eco*RI (right) fragment to which the semidominant mutations *n355* and *n536* map was used as a probe and found to hybridize to a 3.5-kb mRNA isolated from wild-type strain N2, to a 3.0-kb mRNA isolated from strains containing the *lin-14*(*n536*) mutation, and to 3.0- and 2.2-kb mRNAs isolated from strains containing the *lin-14*(*n355*) mutation. Two micrograms of poly(A)$^+$ mRNA isolated from each strain were separated by electrophoresis in 1.2% formaldehyde agarose gels. The amounts of total mRNA in each lane were shown to be equivalent using ^{32}P-labeled actin (FILES, CARR and HIRSH 1983) probe that hybridizes to a 1.5-kb mRNA. The poly(A)$^+$ RNAs shown are: lane 1: wild type (N2) mixed stages; lane 2: MT355 = *lin-14*(*n355*sd) mixed stages, lane 3: MT1149 = *lin-14*(*n536*sd) mixed stages.

such as in *n540 n536sd* strains (data not shown). This observation argues that the reduction in the size of this transcript in these mutants is not an indirect consequence of their retarded or precocious phenotypes, but rather is a direct consequence of the *n536sd* or *n355sd* mutations. The 7-kb mRNA is unchanged in these strains (data not shown).

The fact that both dominant *lin-14* mutations affect the 3.5-kb mRNA further supports the argument that the 3.5-kb mRNA is a *lin-14* transcript.

DISCUSSION

The nature of the *lin-14* semidominant mutations: The two semidominant *lin-14* mutations are DNA deletions or rearrangements of the 3′ region of the *lin-14* gene. These mutations appear to decrease the size of the normally 3.5-kb *lin-14* transcript by eliminating 3′ sequences from this mRNA. These semidominant mutations have been shown genetically to cause an increase of *lin-14* gene activity (AMBROS and HORVITZ 1987). Assuming that the semidominant *lin-14* mutations add no functionally relevant sequences to the *lin-14* gene and mRNA, the sequences deleted from the *lin-14* gene in these mutants must encode a *cis*-acting element that negatively regulates *lin-14* gene activity during normal development.

Various models can explain how deletions within the *lin-14* gene and mRNA could cause an increase in *lin-14* gene activity: (1) at the DNA level, the deleted sequences could encode a transcriptional repressor binding site or a negative enhancer (JOHNSON and HERSKOWITZ 1985), and affect the level of *lin-14* transcription; (2) at the RNA level, the deleted sequences could define a site that is recognized by a specific or nonspecific RNAse (COLE 1986) that degrades the *lin-14* mRNA, or a site that is recognized by an enzyme that regulates *lin-14* mRNA processing (BOGGS *et al.* 1987) or translation (DESCHAMPS *et al.* 1985) as development proceeds; or (3) at the protein level, the deleted sequences could encode a domain of the *lin-14* protein responsible for its normal instability (DUNCAN 1986), or for some form of allosteric negative regulation of *lin-14* activity (JOVE and HANAFUSA 1987). These models are currently being tested by examining the levels of *lin-14* mRNA and protein in wild-type and *lin-14* mutant strains and by determining the DNA sequences of *lin-14* cDNA clones isolated from the wild-type and *lin-14* semidominant mutants.

The utility of parallel-RFLP mapping in cloning *C. elegans* genes: Our use of a highly repetitive element for the genetic mapping and cloning of *lin-14* is generally applicable to other genes in *C. elegans*. Assuming a uniform distribution of genes and Tc1-dimorphic loci, we calculate that on average there should be one such locus every 200 kb or within about 100 kb of any *C. elegans* gene. In the *lin-14* region, we have found that Tc1-dimorphic loci are separated by about 300 kb, and that one of these loci is located within 5 kb of the *lin-14* gene. A study of the *lin-12* region showed that two flanking Tc1-dimorphic loci are separated by about 450 kb; the *lin-12* gene is within 75 kb of one of them (GREENWALD *et al.* 1987). Thus, based both upon these examples and our calculation, it appears likely that any gene of interest will be within about 200 kb of its nearest Tc1-dimorphic locus. As in the case of *lin-14*, such closely linking Tc1-dimorphic loci can be cloned and used to identify an overlapping multiclone contig assembled in the course of the *C. elegans* genome mapping project

(COULSON *et al.* 1986, 1988). Currently, this project has progressed to the point that at least 90% of the genome is represented by 247 contigs of average size 368 kb and ranging in size from 40 to 5100 kb (A. COULSON, J. SULSTON, R. WATERSTON, Y. KOHARA, D. ALBERTSON and R. FISHPOOL, unpublished results). Clones from these contigs can then be used, as we have done, to map more precisely the gene location using other RFLP loci and selected genetic recombinants. Given that the average separation of Tc1-dimorphic loci and the average contig size are now about equal (and the average contig size is continuing to increase as the project progresses), there is a high probability that the gene of interest will be located on the contig identified with the cloned Tc1-dimorphic locus, obviating the need for chromosome walking (BENDER, SPIERER and HOGNESS 1983). Indeed, genetic mapping of linked Tc1-dimorphic loci and the *C. elegans* physical genetic map have already been used together to clone the *C. elegans* cell lineage gene *unc-86* (CHALFIE, HORVITZ and SULSTON 1981; FINNEY, RUVKUN and HORVITZ 1988) and the cell death gene *ced-3* (ELLIS and HORVITZ 1986; J. YUAN and R. HORVITZ, unpublished results).

While the 247 contigs currently assembled represent a midpoint in the compilation of a complete *C. elegans* genome map, our results show that these contigs are already extremely useful in yielding the "medium scale" (100–1000 kb) data about genome organization necessary to map and clone genes using any RFLP mapping approach. A partially complete physical genetic map at an analogous point in any genome mapping project would have similar utility.

Intragenic recombinants and the etiology of sporadic dominant mutations in human genetic disease: The unveiling of a *lin-14* dominant mutation from its *cis*-acting suppressor mutation by the *lin-14* intragenic recombination event is an example of a general mechanism by which apparently spontaneous germ line or somatic dominant mutations, for example in human genetic disease or oncogenesis, could appear. Thus, some of these mutants might arise not via a new mutation but rather by a similar recombinational revelation of an extant but cryptic dominant mutation. Unlike newly induced mutations, such recombinationally revealed dominant mutations would always be carried on a chromosome that bore evidence of a recombination event within the dominantly mutant gene. The presence of such a recombinant chromosome in either mutant progeny or tumors could be detected using RFLP probes. The crossover point within the dominantly mutant gene, and thus the gene, could then be mapped precisely, as we have done for *lin-14*, if polymorphic chromosomes were present in the parental generation. Such precise physical genetic mapping is a necessary prerequisite for cloning such genetically defined loci by an RFLP mapping/chromosome walking strategy. Unlike more traditional two-factor and three-factor RFLP data (BOTSTEIN *et al.* 1980), which yield statistical measures of co-segregation of a given RFLP locus with a gene of interest, and therefore require large pedigrees for fine structure genetic mapping, one intragenic recombination event can define the location of at least part of the gene as precisely as the density of detectable RFLPs can define the location of the intragenic recombination event. Thus, large pedigrees are not necessary for the precise mapping and eventual cloning of such genes.

We thank D. ALBERTSON and R. FISHPOOL for assistance with the contig mapping, GUY BENIAN for his cosmid library, and CORI BARGMANN, ROGER BRENT, THOMAS BÜRGLIN, MIKE FINNEY, JIM GUSELLA, STUART KIM, DAVID PAGE and NATASHA STALLER for helpful discussions. This work was supported by the Harvard Society of Fellows and Hoechst, AG (G.R.); U.S. Public Health Service research grant GM24663 and GM24943 and Research Career Development Award HD00369 (H.R.H.); NIH postdoctoral fellowship F32 GM08642 and American Cancer Society fellowship NYPF-1726 (V.A.).

LITERATURE CITED

AMBROS, V., and H. R. HORVITZ, 1984 Heterochronic mutants of the nematode *Caenorhabditis elegans*. Science **226:** 409–416.

AMBROS, V., and H. R. HORVITZ, 1987 The *lin-14* locus of *Caenorhabditis elegans* controls the time of expression of specific postembryonic developmental events. Genes Dev. **1:** 398–414.

BENDER, W., P. SPIERER and D. S. HOGNESS, 1983 Gene isolation by chromosome walking. J. Mol. Biol. **168:** 17–34.

BOGGS, R. T., P. GREEN, S. IDRISS, J. M. BELOTE and M. MCKEOWN, 1987 Regulation of sexual differentiation in *Drosophila melanogaster* via alternative splicing of RNA from the *transformer* gene. Cell **48:** 489–499.

BOTSTEIN, D., R. L. WHITE, M. SKOLNICK and R. W. DAVIS, 1980 Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. **32:** 314–331.

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

BURKE, D. T., G. F. CARLE and M. V. OLSON, 1987 Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science **236:** 806–812.

CHALFIE, M., H. R. HORVITZ and J. E. SULSTON, 1981 Mutations that lead to reiterations in the cell lineage of *Caenorhabditis elegans*. Cell **24:** 59–69.

COLE, M. D., 1986 The *myc* oncogene: its role in transformation and differentiation. Annu. Rev. Genet. **20:** 361–384.

COULSON, A., J. E. SULSTON, S. BRENNER and J. KARN, 1986 Towards a physical genetic map of the genome of the nematode *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA **83:** 7821–7825.

COULSON, A., R. WATERSTON, J. KIFF, J. SULSTON and Y. KOHARA, 1988 Genome linking with yeast artificial chromosomes. Nature **335:** 184–186.

COX, G. N., S. CARR, J. M. KRAMER and D. HIRSH, 1985 Genetic mapping of *Caenorhabditis elegans* collagen genes using DNA polymorphisms as phenotypic markers. Genetics **109:** 513–528.

DAVIS, R. W., D. BOTSTEIN and J. R. ROTH, 1980 *Advanced Bacterial Genetics*. Cold Spring Harbor Press, Cold Spring Harbor, N.Y.

DESCHAMPS, J., *et al.*, 1985 Proto-oncogene *fos* is expressed during development, differentiation, and growth. Cold Spring Harbor Symp. Quant. Biol. **50:** 733–745.

DUNCAN, I. M., 1986 Control of bithorax complex functions by segmentation gene *fushi tarazu* of *Drosophila melanogaster*. Cell **47:** 297–309.

ELLIS, H. M., and H. R. HORVITZ, 1986 Genetic control of programmed cell death in the nematode *Caenorhabditis elegans*. Cell **44:** 817–829.

EMMONS, S. W., M. R. KLASS and D. HIRSH, 1979 Analysis of the constancy of DNA sequences during development and evolution of the nematode *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA **76:** 1333–1337.

EMMONS, S. W., L. YESNER, K. RUAN and D. KATZENBERG, 1983 Evidence for a transposon in *Caenorhabditis elegans*. Cell **32:** 55–65.

FEINBERG, A. P., and B. VOGELSTEIN, 1983 A technique for radiolabeling DNA restriction fragments to high specific activity. Anal. Biochem. **132:** 6–13.

FILES, J. G., S. CARR and D. HIRSH, 1983 Actin gene family of *Caenorhabditis elegans*. J. Mol. Biol. **164:** 355–375.

FINNEY, M., 1987 The genetics and molecular biology of *unc-86*, a *Caenorhabditis elegans* cell lineage gene. Ph.D. thesis, MIT, Cambridge, Mass.

FINNEY, M., G. B. RUVKUN and H. R. HORVITZ, 1988 The *Caenorhabditis elegans* cell lineage and differentiation gene *unc-86* encodes a protein with a homeodomain and extended similarity to transcription factors. Cell **55:** 757–769.

GOULD, S. J., 1977 *Ontogeny and Phylogeny*. Belknap Press, Cambridge, Mass.

GREENWALD, I., A. COULSON, J. SULSTON and J. PRIESS, 1987 Correlation of the physical and genetic map in the *lin-12* region of *Caenorhabditis elegans*. Nucleic Acids Res. **15:** 2295–2307.

HODGKIN, J., and S. BRENNER, 1977 Mutations causing transformation of sexual phenotype in the nematode *Caenorhabditis elegans*. Genetics **86:** 275–287.

HORVITZ, H. R., and J. E. SULSTON, 1980 Isolation and genetic characterization of cell lineage mutants of the nematode *Caenorhabditis elegans*. Genetics **96:** 435–454.

JOHNSON, A. D., and I. HERSKOWITZ, 1985 A repressor (*Matα2* product) and its operator control expression of a set of cell-type specific genes in yeast. Cell **42:** 237–247.

JOVE, R., and H. HANAFUSA, 1987 Cell transformation by the viral *src* oncogene. Annu. Rev. Cell Biol. **3:** 31–56.

MANIATIS, T., E. F. FRITSCH and J. SAMBROOK, 1982 *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

MOERMAN, D. G., and R. H. WATERSTON, 1984 Spontaneous unstable *unc-22* IV mutations in *C. elegans* var. Bergerac. Genetics **108:** 859–877.

ROSENZWEIG, B., L. LIAO and D. HIRSH, 1983 Sequence of the *Caenorhabditis elegans* transposable element Tc1. Nucleic Acids Res. **11:** 4201–4209.

STERNBERG, P. W., and H. R. HORVITZ, 1984 The genetic control of cell lineage during nematode development. Annu. Rev. Genet. **18:** 489–524.

SULSTON, J. E., and S. BRENNER, 1974 The DNA of *Caenorhabditis elegans*. Genetics **77:** 95–104.

Communicating editor: R. K. HERMAN