Macrorestriction Analysis of *Caenorhabditis elegans* Genomic DNA

Heidi Browning,[1] Laura Berkowitz, Cynthia Madej,[2] Janet E. Paulsen,[3] Miriam E. Zolan and Susan Strome

Department of Biology and Institute for Molecular and Cellular Biology, Indiana University, Bloomington, Indiana 47405

Manuscript received March 28, 1996
Accepted for publication June 22, 1996

ABSTRACT

The usefulness of genomic physical maps is greatly enhanced by linkage of the physical map with the genetic map. We describe a "macrorestriction mapping" procedure for *Caenorhabditis elegans* that we have applied to this endeavor. High molecular weight, genomic DNA is digested with infrequently cutting restriction enzymes and size-fractionated by pulsed field gel electrophoresis. Southern blots of the gels are probed with clones from the *C. elegans* physical map. This procedure allows the construction of restriction maps covering several hundred kilobases and the detection of polymorphic restriction fragments using probes that map several hundred kilobases away. We describe several applications of this technique. (1) We determined that the amount of DNA in a previously uncloned region is <220 kb. (2) We mapped the *mes-1* gene to a cosmid, by detecting polymorphic restriction fragments associated with a deletion allele of the gene. The 25-kb deletion was initially detected using as a probe sequences located ~400 kb away from the gene. (3) We mapped the molecular endpoint of the deficiency *hDf6*, and determined that three spontaneously derived duplications in the *unc-38-dpy-5* region have very complex molecular structures, containing internal rearrangements and deletions.

THE molecular analysis of many organisms has been greatly facilitated by the construction of physical maps of their genomes (for example, OLSON *et al.* 1986, KOHARA *et al.* 1987; HARTL *et al.* 1992; COHEN *et al.* 1993; MORTIMER *et al.* 1992; HOHEISEL *et al.* 1993; MIZUKAMI *et al.* 1993). The *Caenorhabditis elegans* physical map consists of contiguous regions of YAC and cosmid clones referred to as "contigs" (COULSON *et al.* 1986, 1988, 1991, 1995). In regions where there is good linkage between the physical and genetic maps, it can be relatively straightforward to molecularly identify genetically defined loci. In addition, the physical map is being used for the massive effort to sequence the *C. elegans* genome (SULSTON *et al.* 1992; WILSON *et al.* 1994).

The initial phase of physical map construction involved cloning *C. elegans* DNA into cosmid vectors (COULSON *et al.* 1986). Many regions remained uncloned, resulting in a large number of unconnected contigs. To overcome the biases and limitations of cosmid cloning, the second phase of physical map construction relied on the use of YAC vectors (COULSON *et al.* 1988, 1991). This combined approach has been very successful: >95% of the *C. elegans* genome has been cloned, and there are currently only seven gaps in clone coverage (COULSON *et al.* 1995; HODGKIN *et al.* 1995). The final goal of the project is to close these gaps and reduce the number of contigs to six, the haploid number of chromosomes.

We have adapted the macrorestriction mapping technique (LAWRENCE *et al.* 1987; SMITH *et al.* 1987; FAN *et al.* 1988; LINK and OLSON 1991) to an analysis of the *C. elegans* physical map. High molecular weight DNA is isolated from *C. elegans* larvae and digested with restriction enzymes that have 8-bp recognition sequences. The digested DNA is size-fractionated by pulsed field gel electrophoresis (PFGE), and cosmids from the physical map are hybridized to Southern blots of these gels. The power of the approach is that the restriction fragments are very large (generally hundreds of kilobases), and thus large regions of the genome can be analyzed.

In one application of this technique, we demonstrate its usefulness for eventually closing gaps in clone coverage. By constructing a macrorestriction map of a gap-containing region of linkage group II, we were able to provide an estimate of the distance between two neighboring contigs and identify restriction fragments that contain the intervening DNA.

A second use of the technique is to locate genes on the physical map. A mutant allele caused by a sizeable deletion or DNA rearrangement can result in alteration of the size of a macrorestriction fragment. Such a polymorphic fragment can be detected using probe sequences that map far from the actual polymorphism. Indeed, we detected a deletion allele of the *mes-1* gene from ~400 kb away and subsequently mapped the deletion to a cosmid.

Corresponding author: Susan Strome, Department of Biology, Indiana University, Bloomington, IN 47405.
E-mail: sstrome@bio.indiana.edu

[1] *Present address:* Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO 80309.

[2] *Present address:* Department of Molecular Genetics, University of Cincinnati Medical Center, Cincinnati, OH 45267.

[3] *Present address:* Genetics Institute, Cambridge, MA 02140.

In a third application of the technique, we characterized chromosome duplications and deficiencies in an effort to more closely align the genetic and physical maps. A large collection of duplications and deficiencies exists for the *C. elegans* genome. Because *C. elegans* chromosomes are holocentric, with centromeric function dispersed along the length of the chromosome (ALBERTSON and THOMSON 1982), duplications and deficiencies are stably transmitted from generation to generation. The genetic coverages of many duplications and deficiencies have been mapped (for example HERMAN *et al.* 1976, 1979; MENEELY and HERMAN 1979; SIGURDSON *et al.* 1984; MCKIM and ROSE 1990, 1994). Determining the molecular boundaries of genetically characterized duplications and deficiencies would serve to more closely link the genetic and physical maps. Several strategies have been employed previously to accomplish this task. *In situ* hybridization to meiotic chromosomes has been used to molecularly map chromosome rearrangements (ALBERTSON 1993). Quantitative Southern hybridization analysis of DNA from animals heterozygous for a deficiency (TAX *et al.* 1994) and PCR analysis of animals homozygous for a deficiency (HENDERSON *et al.* 1994) have been used to determine whether specific regions of the genome are missing from deficiency-bearing chromosomes. In addition, deficiency endpoints have been molecularly mapped by restriction fragment length polymorphism (RFLP) analysis (KRAMER *et al.* 1988, 1990; VON MENDE *et al.* 1988). In this last approach, because only a small portion of the genome is sampled with each Southern hybridization analysis, detection of a duplication- or deficiency-specific RFLP requires that the probe be at or very near the breakpoint. We have used the macrorestriction mapping technique to detect duplication- and deficiency-specific RFLPs using as probes DNA sequences that map hundreds of kilobases away from the duplication or deficiency endpoints. We describe molecular characterization of three spontaneously derived duplications and two deficiencies induced by gamma radiation. The numbers and sizes of the RFLPs created by the duplications suggest that their structures are extremely complex, containing internal rearrangements and deletions. The two deficiencies analyzed appear to be less complex.

MATERIALS AND METHODS

Maintenance of strains: Strains were maintained as described by BRENNER (1974). The *C. elegans* strain N2 was the wild-type parent of all mutant strains. The following genes and mutations were used:

LGI: *dpy-5(e61), dpy-14(e188), unc-31(e450)*

LGX: *mes-1(bn24), mes-1(bn74)*

The following chromosomal rearrangements were used: *hDf6 (I), hDp31 (I;f), hDp24 (I;f), hDp76 (I;f),* and *hDp29 (I;f).* Some strains were provided by the Caenorhabditis Genetics Center. The duplication- and deficiency-bearing strains were provided by K. MCKIM, R. ROSENBLUTH, A. HOWELL, and A. ROSE.

Maintenance of deletion strains: *mes-1* mutant animals bearing the *bn74*-associated deletion were maintained as homozygotes, since the sterile phenotype is incompletely expressed and fertile progeny are produced at every generation (CAPOWSKI *et al.* 1991; STROME *et al.* 1995). The *hDf6*-bearing chromosome was maintained in the presence of a duplication (*hDp31*) that has wild-type DNA for the region deleted by *hDf6* (*i.e., Df/Df/Dp*).

Separation of duplication- and non-duplication-bearing animals: For the molecular analysis of duplications, it is preferable to compare duplication-bearing animals with non-duplication-bearing animals in the same genetic background. This eliminates from consideration any RFLPs that result from strain variability rather than from the duplication. Duplications are lost at variable frequencies from generation to generation. It was possible to compare animals bearing the *hDp24, hDp76,* or *hDp29* duplication with non-duplication-bearing animals in the same genetic background because the normal chromosomes carry mutations in *dpy-5* and *dpy-14* while the duplications carry wild-type copies of these genes. Therefore, duplication-bearing animals are phenotypically wild-type and non-duplication-bearing animals are Dpy-5 Dpy-14. Dpy-5 Dpy-14 animals are very short and crawl slowly and therefore were easily separated from wild-type animals by placing mixed populations of worms in the center of 10-cm agar plates spread with *Escherichia coli.* The phenotypically wild-type animals rapidly crawled from the center after which each of the centers was cut out of the plates with a sterile scalpel. The worms from each area were washed off the agar with M9 buffer (BRENNER 1974) and collected by centrifugation. This process was repeated as necessary to collect relatively pure populations of non-duplication-bearing animals and to enrich for duplication-bearing animals.

For duplication strains, a small aliquot of each larval sample that was to be embedded in agarose (see next section) was transferred to an agar plate spread with *E. coli.* The animals were allowed to grow to adulthood, and the percentages of duplication-bearing (non-Dpy) and non-duplication-bearing (Dpy) animals were determined. In the duplication-bearing populations, 53–67% of the worms carried the duplication. In the non-duplication-bearing populations, 92–100% of the worms lacked the duplication.

Preparation of high molecular weight, genomic DNA: High molecular weight, genomic DNA was prepared from young larvae. Young larvae were used because they lack the tough, protease-resistant eggshell that surrounds embryos, their cuticles may be more susceptible to protease digestion than the cuticles of older worms, and R. WATERSTON successfully obtained high molecular weight DNA by protease lysis of L1s (personal communication). Embryos prepared by hypochlorite treatment of gravid adult hermaphrodites (JOHNSON and HIRSH 1979) were allowed to hatch in M9 buffer at 16°. Larvae were separated from debris and unhatched embryos by filtering through a 20-μM mesh. The filtrate was centrifuged at 400–450 × *g* for 2 min, and the pelleted larvae were rinsed once in double distilled H_2O followed by a second centrifugation. The worm pellet was diluted 1:10 with ddH_2O. An equal volume of 2% low melt preparative grade agarose (made in H_2O) was added to the worm suspension, and the mixture was dispensed into BioRad plug molds and allowed to solidify at 4°. The solidified plugs were digested in lysis buffer (1 mg/ml proteinase K, 0.1 M EDTA pH 8.0, 0.01 M Tris·Cl pH 8.0, 1% sarkosyl) at 50° for ~24 hr. To remove the proteinase K, the plugs were washed at room temperature in several changes of storage buffer (0.1 M EDTA pH 8.0, 0.01 M Tris·Cl pH 8.0), including one wash of 6 hr to overnight. The plugs were stored at 4° in storage buffer. We found that plugs could be stored for ≥1 yr.

Restriction digestion of agarose-embedded DNA: The agarose-embedded DNA was digested using restriction enzymes with 8-bp recognition sequences, under conditions rec-

ommended by New England Biolabs. The plugs were cut into pieces slightly smaller than the well of the gel on which they were to be size-fractionated, equilibrated in 10 mM Tris·Cl pH 8.0, 5 mM EDTA pH 8.0 for 15 min on ice, and then equilibrated with the recommended restriction enzyme buffer for 15 min on ice. Fresh restriction buffer and enzyme were added, and the plugs were digested overnight at the temperature recommended by the manufacturer. For digestions with two restriction enzymes, the plugs were first digested with one enzyme, and then the procedure was repeated for the second enzyme, beginning with the Tris·Cl-EDTA equilibration.

Pulsed-field gel electrophoresis: Restriction-digested DNA was size-fractionated by PFGE on a BioRad CHEF (clamped homogeneous electrophoresis field) DRII apparatus. The digested DNA plugs and agarose-embedded molecular weight markers (BioRad) were equilibrated in 0.5× TBE for >20 min. The plugs were loaded into the wells of a 1% LE agarose 0.5× TBE gel by placing them flush against the front side of the well with a sterile scalpel. The plugs were sealed in place with 1% LE agarose. Gels were run in 0.5× TBE at ~15°. We used the BioRad Autoalgorithm Program to determine optimal running parameters for separating restriction fragments of various sizes. The running parameters for the gels shown in Figures 1, 5, 6, and 7 were initial pulse time 0.2 sec, final pulse time 93.7 sec at 200 V for 18.5 hr or initial pulse time 1 sec, final pulse time 93.7 sec at 200 V for 23.2, 31, or 36 hr. The running parameters for the gels shown in Figure 2 were initial pulse time 0.2 sec, final pulse time 8.6 sec at 200 V for 27.5 hr.

The sizes of the chromosomal restriction fragments and RFLPs were calculated based on migration distances compared with yeast chromosomes (BioRad) and a lambda ladder marker (BioRad). There was ≤10% variation in size calculated for some fragments in independent experiments.

Southern blot analysis of CHEF gels: Southern transfers were performed essentially as described in AUSUBEL *et al.* (1988) except that depurination was in 0.25 N HCl for 10 min. The filters were initially prehybridized in 1× SSC, 0.5% SDS at 65° for 1 hr and then prehybridized in 4× SSC, 1% SDS, 0.5% nonfat dry milk at 65° for 4–24 hr. The second prehybridization solution was replaced with fresh prehybridization solution plus denatured probe. Hybridization proceeded for 36–48 hr. The filters were washed two times for 5 min at 65° in 2× SSC, 0.5% SDS followed by two times for 1 hr at 65° in 0.2× SSC, 0.1% SDS.

Preparation of cosmid probes for Southern hybridization analysis: Cosmids were prepared by the alkaline lysis procedure essentially as described in SAMBROOK *et al.* (1989). Two methods were used to prepare cosmids for radiolabeling: cosmids were purified by conventional agarose gel electrophoresis, followed by isolation from the gel using the Bio101 Gene Clean Kit or cosmids were partially digested with *Eco*RI, extracted with phenol/chloroform, and ethanol precipitated. Purified cosmids were radiolabeled with α-^{32}P-dCTP using the Boehringer Mannheim Random Primed Labeling Kit.

Conventional preparation of genomic DNA and Southern blot analysis: Genomic DNA for conventional RFLP analysis was prepared as described in SPIETH *et al.* (1988). Southern blot transfer and hybridizations were essentially as described in AUSUBEL *et al.* (1988). After 14–48 hr hybridization, the filters were washed as described above for CHEF Southerns.

RESULTS

Macrorestriction analysis to estimate the sizes of gaps in clone coverage: *lin-29* and *unc-52* map to the right arm of linkage group II, approximately 19 map units (m.u.) apart (see a *C. elegans* data base or ACeDB; R. DURBIN and J. THIERRY-MIEG, personal communication). Both genes have been cloned and assigned to separate contigs, referred to here as the *lin-29* contig and *unc-52* contig (Figure 1). The physical distance between the contigs (*i.e.*, the amount of uncloned DNA between the contigs) was not known. To estimate this distance, we investigated whether the endmost cosmids on each contig hybridize to a common restriction fragment(s). The cosmids were radiolabeled and hybridized

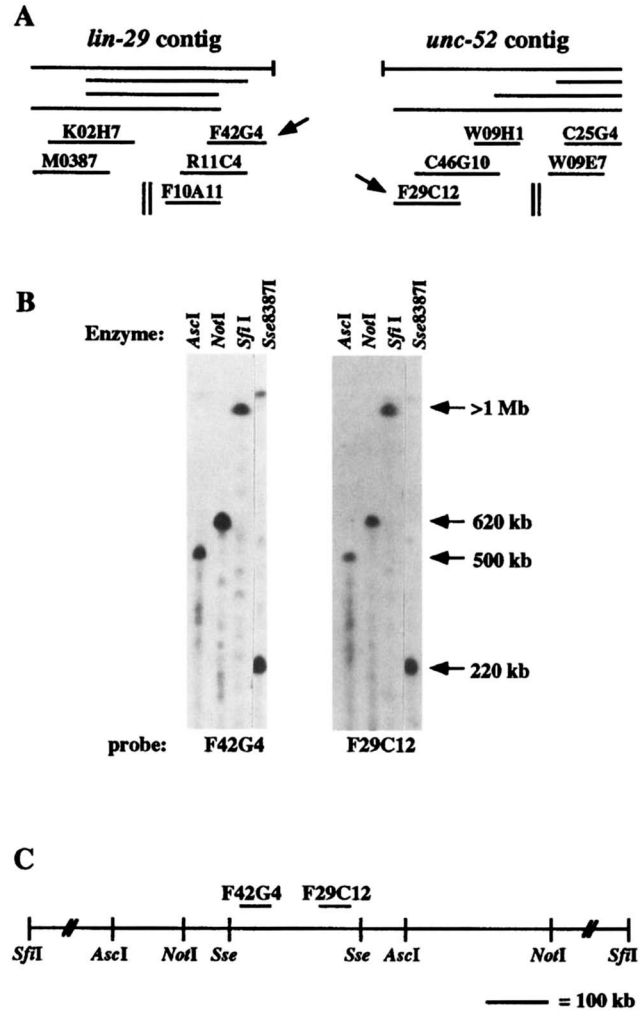

FIGURE 1.—Estimation of the distance between the *lin-29* and *unc-52* contigs. (A) Physical map of the *lin-29-unc-52* region. Cosmid and YAC clones on the end of each contig are shown. Cosmids (labeled lines) and YAC clones (unlabeled lines) are diagrammed as in ACeDB (a *C. elegans* data base; R. DURBIN and J. THIERRY-MIEG, personal communication). A single YAC may span the gap between the two contigs, but this is listed as "highly unconfirmed". On each contig, double vertical lines indicate gaps in cosmid coverage. Arrows indicate the cosmids that were used for macrorestriction analysis. (B) Macrorestriction Southern hybridization analysis of the *lin-29-unc-52* region. Genomic DNA from wild-type animals was digested with *Asc*I, *Not*I, *Sfi*I, and *Sse*8387I, size-fractionated by PFGE, transferred to Hybond, and probed with a 5.7-kb *Eco*RI fragment from cosmid F42G4 and a 1.9-kb *Eco*RI fragment from cosmid F29C12. The sizes of the fragments are indicated. (C) Macrorestriction map of the *lin-29-unc-52* region. The map is drawn based on the single digest results shown in B and the results of double digests (data not shown).

to Southern blots of wild-type *C. elegans* genomic DNA that had been digested with restriction enzymes with 8-bp recognition sequences and size-fractionated by PFGE (see MATERIALS AND METHODS for details of DNA preparation and pulsed-field gel parameters).

The cosmids F42G4 from the *lin-29* contig and F29C12 from the *unc-52* contig hybridized to many chromosomal bands generated by each restriction digestion (data not shown), making it impossible to construct a map. We assumed that repeated sequences in the cosmids are responsible for this complexity, and so we used restriction fragments from the cosmids as probes in the analysis. For each cosmid, several restriction fragments were tested to determine which fragments do not contain repeat sequences. A 5.7-kb *Eco*RI fragment from F42G4 appears to contain only unique sequences and hybridizes to a 500-kb *Asc*I chromosomal fragment, a 620-kb *Not*I chromosomal fragment, a >1-Mb *Sfi*I chromosomal fragment, and a 220-kb *Sse8387*I chromosomal fragment (Figure 1). Similarly, a 1.9-kb *Eco*RI restriction fragment from F29C12 appears to contain only unique sequences and hybridizes to the same size fragments as the F42G4 fragment (Figure 1). Based on the sizes of the chromosomal fragments, both cosmids hybridize to the same restriction fragments and thus are within 220 kb (the size of the *Sse8387*I fragment) of each other.

Macro RFLP analysis to locate genes: Molecularly mapping an RFLP created by a deletion allele of *mes-1* was the first step in physically locating and cloning the *mes-1* gene. *mes-1* maps 0.2 m.u. left of *sma-5*, between *egl-15* and *sma-5* (Figure 2A). There are currently 10 alleles of *mes-1*, four of which were induced by gamma radiation (CAPOWSKI *et al.* 1991; STROME *et al.* 1995). With the goal of identifying a *mes-1* allele-associated polymorphism, DNA from homozygous mutant *mes-1* animals was analyzed by macro RFLP analysis using as probes cosmids from the *egl-15-sma-5* region of the physical map. DNA from *mes-1(bn74)* animals exhibited an allele-associated polymorphism initially detected in DNA digested with *Not*I and *Asc*I and probed with cosmid T04A4 (data not shown; see Figure 2A for map). The polymorphism is the result of deletion of the *Not*I site and was detected by T04A4 from a distance of 400 kb from the actual deletion. The polymorphism was delimited to the region of the genome contained in the cosmid C38D5 (Figure 2B). Both C34E11 and C38D5 detect a polymorphic *Sma*I fragment in DNA from *bn74* animals: a 170-kb *Sma*I chromosomal fragment is reduced to 140 kb in DNA from *bn74* animals. The *bn74* deletion appears to be entirely contained within the 170-kb *Sma*I fragment, as the neighboring *Sma*I fragments in DNA from *bn74* animals are the same size as those in DNA from wild-type animals: to the left of the polymorphic fragment, C34E11 detects 10- and 200-kb fragments in DNA from both wild-type and *bn74* animals, and to the right T08D10 detects an 80-kb fragment in both DNAs (Figure 2, A and B). The relatively weak hybridization of C38D5 to the polymorphic 140-kb *Sma*I fragment in *bn74* DNA suggested that much of the DNA contained in C38D5 is deleted. Indeed, Southern analysis using more frequently cutting enzymes revealed that a 25-kb deletion removed DNA corresponding to most of the right half of C38D5 (Figure 2, C and D). *Bam*HI fragments of 14.8, 9.4, 1.8, and 0.5 kb present in wild-type DNA are deleted from *bn74* DNA. The presence in *bn74* DNA of the 24- and 1.9-kb fragments on the left and of some sequence on the right end of C38D5 is consistent with the deletion being entirely contained within C38D5. The evidence that the 25-kb deletion is responsible for the Mes-1 phenotype is that a 17-kb restriction fragment entirely contained within the 25-kb region deleted in *bn74* DNA is capable of transformation rescue of *mes-1* mutant strains (L. BERKOWITZ and S. STROME, unpublished results).

Macro RFLP analysis to map the endpoints of chromosomal deficiencies: Defining the endpoint of a large deficiency with a breakpoint near *mes-3* provided a physical boundary for the *mes-3*-containing region and facilitated cloning *mes-3*. *mes-3* maps 0.2 m.u. to the right of *unc-38* (CAPOWSKI *et al.* 1991). The deficiency *hDf6* fails to complement *unc-38* but does complement *mes-3* (Figure 3A) (HOWELL and ROSE 1990; E. CAPOWSKI, unpublished result). Thus, to provide a leftward boundary for *mes-3* on the physical map, the righthand endpoint of *hDf6* was molecularly mapped. The data are summarized in the map in Figure 3B. By macro RFLP analysis, the cosmid F14A9 detects a 340- and a 170-kb *Not*I fragment and a 300-kb *Sfi*I fragment in DNA from both wild-type and deficiency-bearing animals. In addition, this cosmid detects a 390-kb *Not*I fragment and a 350-kb *Sfi*I fragment only in DNA isolated from the deficiency-bearing strain. The cosmid C48A4 does not detect RFLPs from deficiency DNA. Therefore, the molecular endpoint of *hDf6* is in F14A9 or between C48A4 and F14A9. The cosmids C30F8 and F56A3 do not detect RFLPs, suggesting that this region of the deficiency chromosome is devoid of gross rearrangements. Consistent with our mapping of the endpoint of *hDp6*, *mes-3* has been molecularly identified and is contained within a cosmid just to the right of F14A9 (PAULSEN *et al.* 1995).

Macro RFLP analysis to assess the structure and integrity of chromosomal duplications: Many duplications have been generated by gamma radiation. Some of these have undergone spontaneous shortening. Several spontaneously shortened duplications have genetically defined endpoints between *unc-38* and *dpy-5* (Figure 4) (MCKIM and ROSE 1990, 1994). Complementation analysis was used to determine how far leftward the duplications extend: the farthest extending duplication should complement the most genes and the shortest duplication should complement the fewest genes. In addition, the order of the genes shown in Figure 4 is based on duplication coverage: failure of complementation or complementation by only one duplication would place a gene toward the left, while complementa-

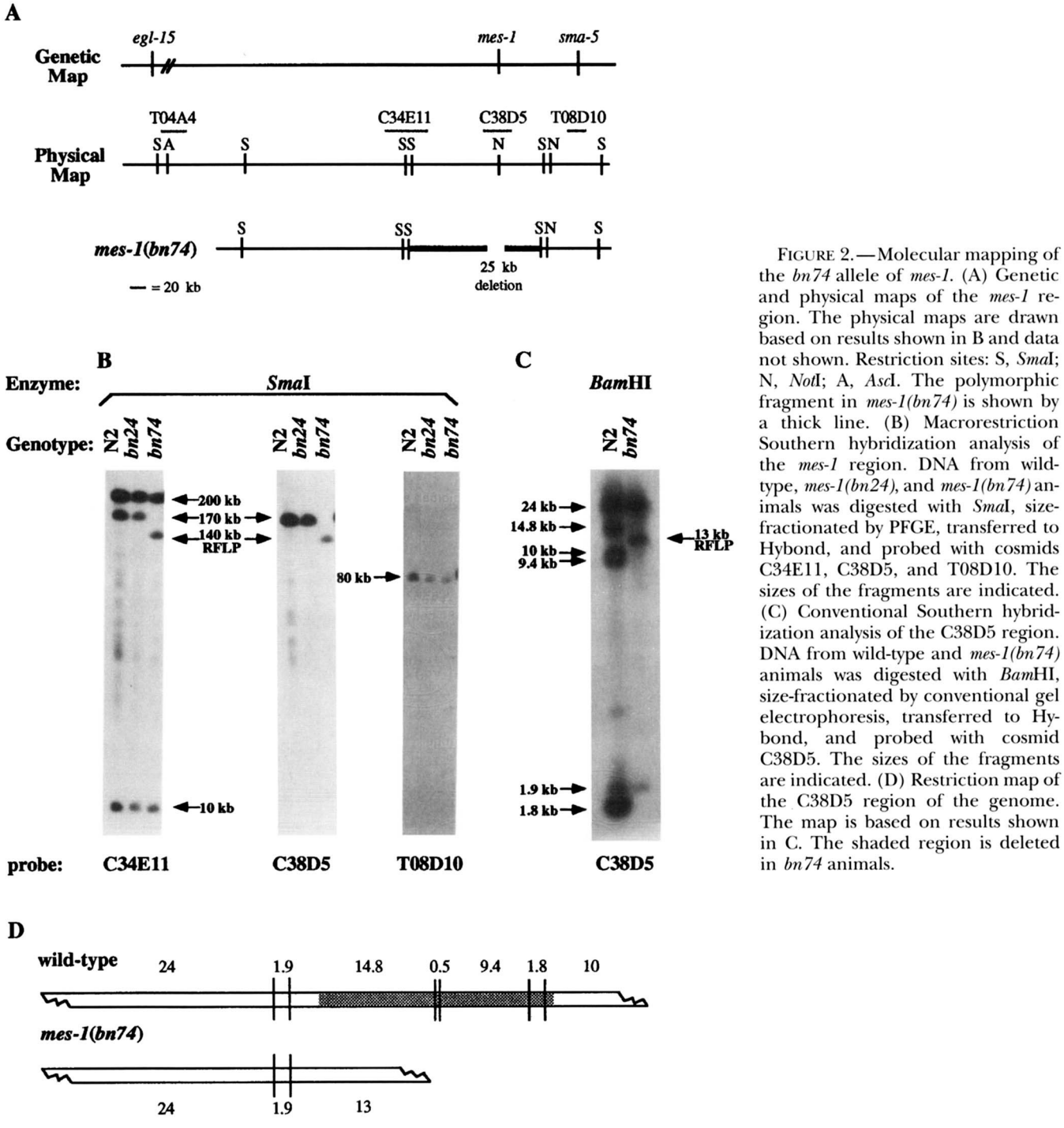

FIGURE 2.—Molecular mapping of the *bn74* allele of *mes-1*. (A) Genetic and physical maps of the *mes-1* region. The physical maps are drawn based on results shown in B and data not shown. Restriction sites: S, *Sma*I; N, *Not*I; A, *Asc*I. The polymorphic fragment in *mes-1(bn74)* is shown by a thick line. (B) Macrorestriction Southern hybridization analysis of the *mes-1* region. DNA from wild-type, *mes-1(bn24)*, and *mes-1(bn74)* animals was digested with *Sma*I, size-fractionated by PFGE, transferred to Hybond, and probed with cosmids C34E11, C38D5, and T08D10. The sizes of the fragments are indicated. (C) Conventional Southern hybridization analysis of the C38D5 region. DNA from wild-type and *mes-1(bn74)* animals was digested with *Bam*HI, size-fractionated by conventional gel electrophoresis, transferred to Hybond, and probed with cosmid C38D5. The sizes of the fragments are indicated. (D) Restriction map of the C38D5 region of the genome. The map is based on results shown in C. The shaded region is deleted in *bn74* animals.

tion by multiple duplications would place a gene toward the right (McKIM and ROSE 1990, 1994). In the process of molecularly identifying *spe-11*, we analyzed the three duplications shown in Figure 4 with the intention of placing molecular boundaries on the physical map for *spe-11*. The physical map in this region is characterized by two cosmid-rich areas bridged by YAC clones (selected cosmids are shown in Figures 5–7). The gap in cosmid coverage extends from C30F8 to F56A3.

Molecular structure of hDp24: Macro RFLP analysis of *hDp24* indicates that this duplication has a large deletion as well as internal rearrangements. F56A3 detects a 340-kb *Not*I chromosomal fragment and a duplication-specific 1080-kb *Not*I RFLP, while B0342 detects an 820-kb *Not*I chromosomal fragment and a duplication-specific 1130-kb *Not*I RFLP (Figure 5A). (Although these RFLPs appear to be different sizes, the resolution of the gel from which the sizes were determined was not sufficient to definitively determine whether these RFLPs are or are not the same size.) The intervening cosmids, C24C1, B0261, C50D8, K11B12, and F27E10, detect only an 820-kb chromosomal *Not*I fragment (Figure 5A; C24C1 and K11B12 shown), suggesting that most of this region is deleted from the duplication.

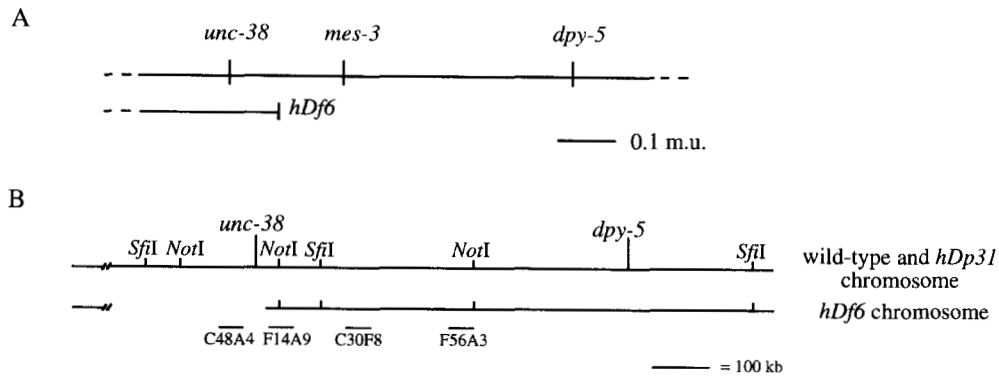

FIGURE 3.—Genetic and physical maps of the right-hand endpoint of *hDf6*. (A) Genetic map. *mes-3* maps 0.2 m.u. to the right of *unc-38* and between *unc-38* and *dpy-5*. *hDf6* does not complement *unc-38* but does complement *mes-3* and *dpy-5*. (B) Macrorestriction map of wild-type and *hDp31* DNA and of *hDf6* DNA. DNA from wild-type animals and from animals homozygous for *hDf6* and carrying the *hDp31* duplication [genotype: *hDf6 dpy-5(e61) unc-31(e450)/ hDf6 dpy-5(e61) unc-31(e450)/ hDp31 (I;f)*] was digested with *Not*I and *Sfi*I and probed with the cosmids shown below the map. Only F14A9 detects deficiency-specific RFLPs. The positions of *unc-38* and *dpy-5* on the restriction map are estimated from their positions on the physical map.

However, conventional RFLP analysis using enzymes with 6-bp recognition sequences demonstrates that the missing segment is not a simple deletion of the entire intervening region. Both F56A3 and B0342 detect duplication-specific *Xba*I and *Xho*I RFLPs (Figure 5B). Because the cosmids do not detect the same size RFLPs, there must be some DNA present between them. This DNA could be part of the original genomic region or an insertion of DNA from another part of the genome. C30F8 does not detect a *Not*I or *Sfi*I RFLP (Figure 5A, *Not*I shown), suggesting that the duplication ends between C30F8 and F56A3. Alternatively, *hDp24* may extend to the left of C30F8 but be missing DNA detected by C30F8. Figure 5C is a diagrammatic representation of *hDp24* based on this analysis.

Molecular structure of hDp76: Macro RFLP analysis indicates that *hDp76* is rearranged and contains a deletion. Both C24C1 and B0342 detect an 820-kb *Not*I and a 750-kb *Sfi*I chromosomal fragment. However, the two cosmids detect different size duplication-specific *Not*I and *Sfi*I RFLPs: C24C1 detects a 610-kb *Not*I RFLP and an 870-kb *Sfi*I RFLP, while B0342 detects a 380-kb *Not*I RFLP and a 500-kb *Sfi*I RFLP (Figure 6A). The most likely explanation for the absence of a *Not*I site and an *Sfi*I site between these cosmids is that novel DNA has been introduced between these two cosmids in the duplication, either by an inversion or an insertion. No RFLPs are detected by F27E10 using several 8-bp enzymes and enzyme combinations: *Asc*I, *Not*I, *Asc*I/*Not*I, *Asc*I/*Sfi*I, and *Not*I/*Sfi*I (data not shown). This indicates that the duplication has a deletion that removes the DNA contained in F27E10. Although F56A3 detects duplication-specific *Not*I and *Sfi*I RFLPs, C30F8 does not detect an RFLP with either of these enzymes (Figure 6A). This suggests that the duplication either ends in the gap in cosmid coverage between C30F8 and F56A3 or is deleted for but extends beyond C30F8. Figure 6B is a diagrammatic representation of *hDp76* based on this analysis.

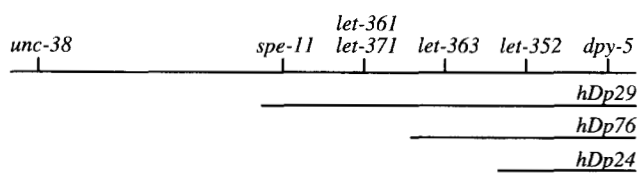

FIGURE 4.—Genetic coverage of spontaneously derived duplications in the *unc-38-dpy-5* region. Duplications are drawn as covering those genes that they complement. The order of the genes between *unc-38* and *dpy-5* is based on duplication coverage. (map adapted from MCKIM and ROSE 1994).

Molecular structure of hDp29: Macro RFLP analysis of *hDp29* demonstrates that it is also rearranged. F56A3 detects a duplication-specific 190-kb *Not*I RFLP, while C24C1 and B0342 detect a duplication-specific *Not*I RFLP migrating at limiting mobility (Figure 7A; F56A3 and C24C1 shown). At least part of the larger *Not*I RFLP is not internally rearranged; although C24C1 detects a duplication-specific *Not*I RFLP and an *Sfi*I RFLP, it does not detect an RFLP when the duplication DNA is double digested with *Not*I and *Sfi*I (Figure 7A). Similarly, B0342 does not detect an RFLP when the duplication DNA is digested with *Asc*I or double digested with *Not*I and *Sfi*I (data not shown). In addition, while C50D8 detects an *Asc*I RFLP and is thus clearly contained within the duplication, it does not detect a *Not*I/*Asc*I RFLP (data not shown). These results indicate that a large portion of the duplication is devoid of gross rearrangements and deletions (Figure 7B). C30F8 does not detect a *Not*I or *Sfi*I RFLP (data not shown), suggesting that the duplication ends in the gap in cosmid coverage. However, the physical end of the duplication is more complex than a simple termination of the DNA sequences from this region. F56A3 detects a 190-kb *Not*I RFLP and a 530-kb *Asc*I RFLP (Figure 7A), a difference of 340 kb. If the duplication simply ended between F56A3 and C30F8, one would expect the *Asc*I and *Not*I polymorphisms to differ by 140 kb (the distance between the middle *Not*I and *Asc*I sites). Therefore, *hDp29* probably has additional, noncontiguous sequences at its genetically defined left end. Figure 7B is a diagrammatic representation of *hDp29* based on this analysis.

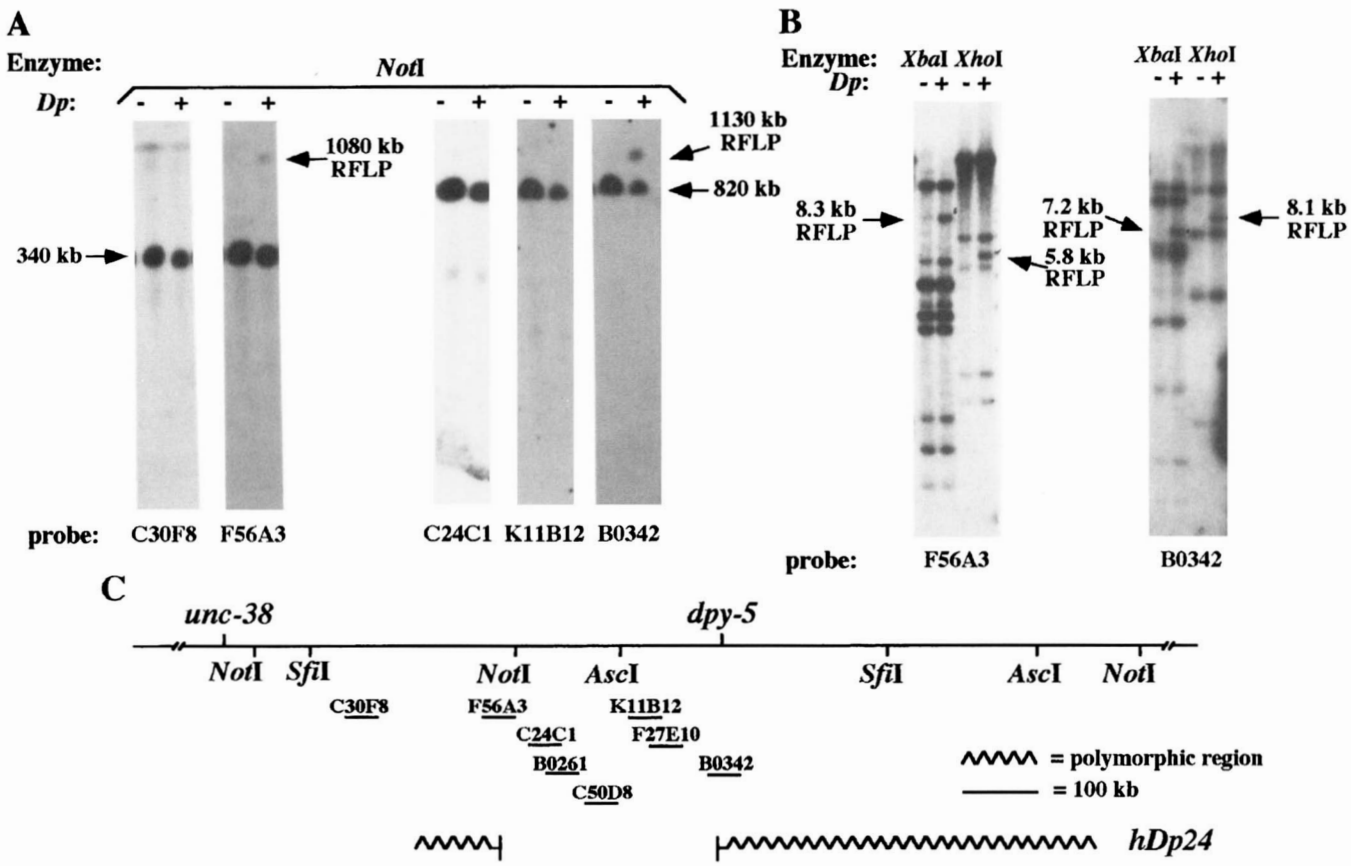

FIGURE 5.—RFLP analysis of *hDp24*. Dp + indicates DNA from strains carrying the duplication, Dp − indicates DNA from the same strain without the duplication (see MATERIALS AND METHODS for details). (A) Macro RFLP analysis. DNA from *dpy-5(e61) dpy-14(e188)* worms and *dpy-5(e61) dpy-14(e188)/hDp24 (I;f)* worms was digested with *Not*I, size-fractionated by PFGE, transferred to Hybond, and probed with cosmids C30F8, F56A3, C24C1, K11B12, and B0342. The sizes of the fragments are indicated. (B) Conventional RFLP analysis. DNA from *dpy-5(e61) dpy-14(e188)* worms and *dpy-5(e61) dpy-14(e188)/hDp24 (I;f)* worms was digested with *Xba*I or *Xho*I, size-fractionated by conventional gel electrophoresis, transferred to Hybond, and probed with cosmids F56A3 and B0342. (C) Diagrammatic representation of *hDp24* based on RFLP analysis. The region of *hDp24* analyzed contains RFLPs, designated as "polymorphic region", and is deleted for sequences between F56A3 and B0342.

DISCUSSION

Applications of macrorestriction mapping to assembly of a physical map for *C. elegans*: We have developed a macrorestriction mapping technique for *C. elegans* DNA and demonstrated how it can be used to facilitate assembly of the complete *C. elegans* physical map. There are currently seven gaps in coverage of the genome by cosmid and YAC clones (COULSON *et al.* 1995). A first step toward filling in these gaps is estimating the amount of DNA that remains to be cloned. We successfully analyzed one gap on the right end of linkage group II and demonstrated that the two contigs in the region are separated by <220 kb. This gap may exist because parts of this 220 kb of DNA are difficult or impossible to clone in standard cosmid and YAC vectors and hosts (HODGKIN *et al.* 1995). Some clones from the gap region may already exist but may not yet be assigned to a contig or may be part of contigs not yet assigned to a linkage group. Macrorestriction analysis using unassigned clones as probes could indicate if some of these clones represent DNA from the 220-kb region.

Macrorestriction mapping could also be applied to the verification of assignments of clones to contigs and linkage groups. Some clusters of cosmid clones are linked by a single YAC. Such linkages are tenuous, because YAC clones can be chimeric and contain inserts from noncontiguous parts of the *C. elegans* genome. Fortunately, coligation events are thought to be relatively rare in the *C. elegans* YAC library (COULSON *et al.* 1991). In addition, the assignments of cosmids to the edges of contigs are occasionally uncertain (A. COULSON, personal communication). Macrorestriction analysis of these regions could be used to confirm both of these types of linkages.

The usefulness of macro RFLP analysis for cloning genes: We defined the molecular location of the *mes-1* gene by macro RFLP analysis of the *bn74* allele of *mes-1*, which was induced by gamma radiation and is a 25-kb deletion. Allele-associated DNA polymorphisms that result in alteration of the size of a macrorestriction fragment can be detected using a probe to any portion of that fragment. Thus, it is possible to detect allele-associated polymorphisms using as a probe sequences that map hundreds of kilobases away from the actual polymorphic site. Through multiple probings with

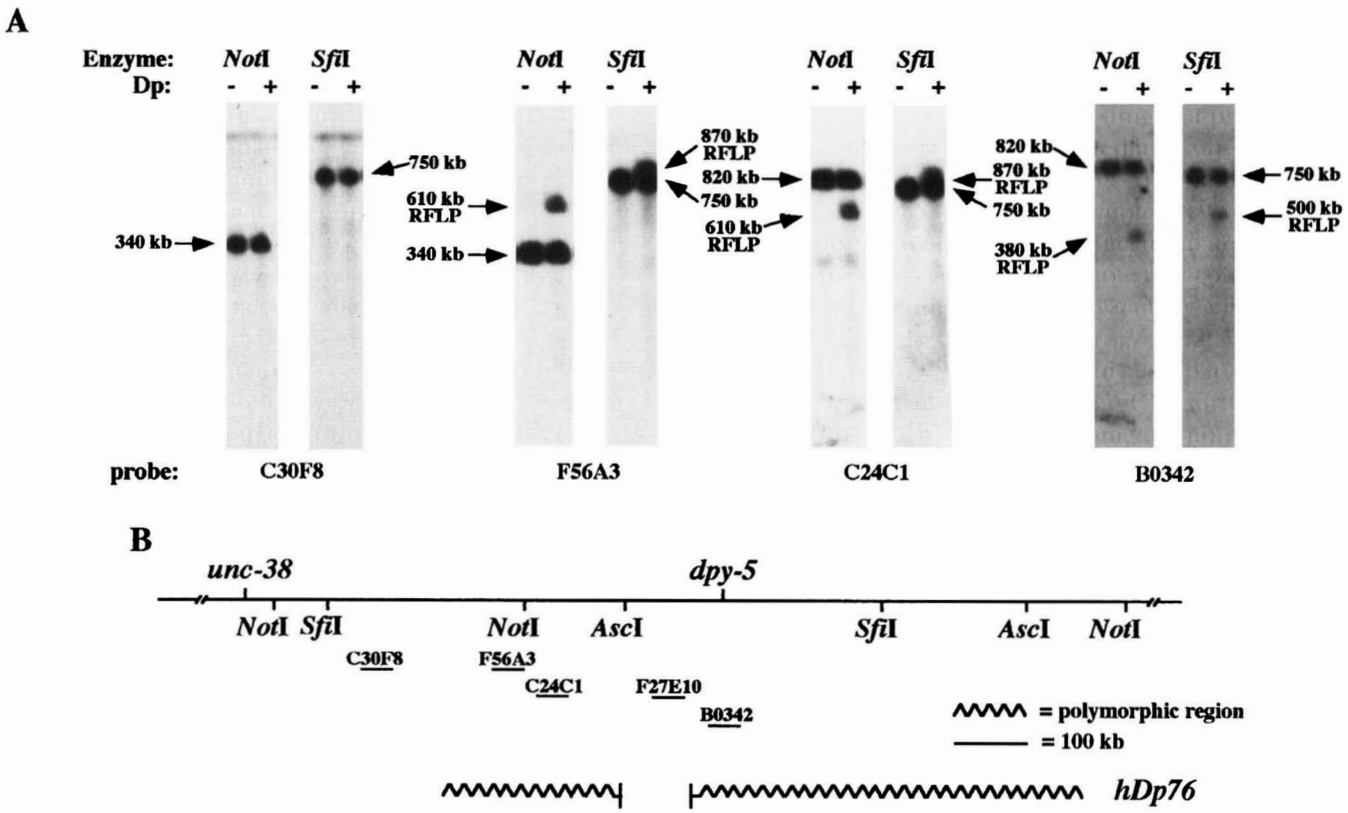

FIGURE 6.—RFLP analysis of *hDp76*. Dp + indicates DNA from strains carrying the duplication, Dp − indicates DNA from the same strain without the duplication (see MATERIALS AND METHODS for details). (A) Macro RFLP analysis. DNA from *dpy-5(e61) dpy-14(e188)* worms and *dpy-5(e61) dpy-14(e188)/ hDp76 (I;f)* worms was digested with *Not*I or *Sfi*I, size-fractionated by PFGE, transferred to Hybond, and probed with cosmids C30F8, F56A3, C24C1, and B0342. The sizes of the fragments are indicated. (B) Diagrammatic representation of *hDp76* based on RFLP analysis. The region of *hDp76* analyzed contains RFLPs ("polymorphic region") and is deleted for the sequences contained in F27E10.

clones from the physical map, one can rapidly scan a relatively large region of the genome for allele-associated polymorphisms in search of the gene of interest.

Once allele-associated polymorphisms are mapped, at least two approaches can be used to determine whether they are responsible for the mutant phenotype. DNA from the region can be tested for the ability to rescue the mutant phenotype in transgenic animals (FIRE and WATERSTON 1989) or anti-sense RNA produced against transcribed genes in the region can be tested for the ability to "phenocopy" the mutant phenotype of the gene (GUO and KEMPHUES 1995).

Macro RFLP analysis of deficiencies and duplications: The usefulness of the physical map relies upon a close correlation between the physical and the genetic maps. Linkages between the two maps are established by identifying DNA clones that contain genetically defined genes, by identifying RFLPs that can be mapped relative to genes and by mapping the endpoints of genetically defined duplications and deficiencies. Macrorestriction mapping is one of several approaches (see Introduction) that can be used to determine the molecular boundaries of duplications and deficiencies.

The three spontaneous duplications we analyzed, *hDp24*, *hDp76*, and *hDp29*, appear to be extremely complex. The duplications were derived from the large, free, radiation-induced duplication *sDp2* (MCKIM and ROSE 1990): exposure of *sDp2*-bearing worms to gamma radiation resulted in several shortened versions of *sDp2*, including *hDp7*, *hDp20*, and *hDp5*, which underwent further spontaneous shortening to generate *hDp24*, *hDp76*, and *hDp29*, respectively. MCKIM and ROSE (1994) analyzed the genetic coverage of a large number of the spontaneous duplications derived from *sDp2*. They also found that several have a complex structure; the parental duplication appears to have lost sequences from both ends and sometimes from an internal position as well. MCKIM and ROSE propose that some duplications exist as ring chromosomes and may contain inversions of some sequences relative to the linear map. Other examples of the spontaneous breakdown of free duplications have been documented in mosaic analysis experiments (HERMAN 1984; VILLENEUVE and MEYER 1990). At least some cases of the spontaneous breakdown of *mnDp3* appear to be due to formation of a deficiency in the duplication (HERMAN 1984).

Based on genetic coverage, *hDp24*, *hDp76*, and *hDp29* appear to have arisen from their parental duplications by a simple deletion of sequences from one end (MCKIM and ROSE 1994). However, our molecular analysis has demonstrated that the structures of the duplications are very complex. The number and sizes of RFLPs

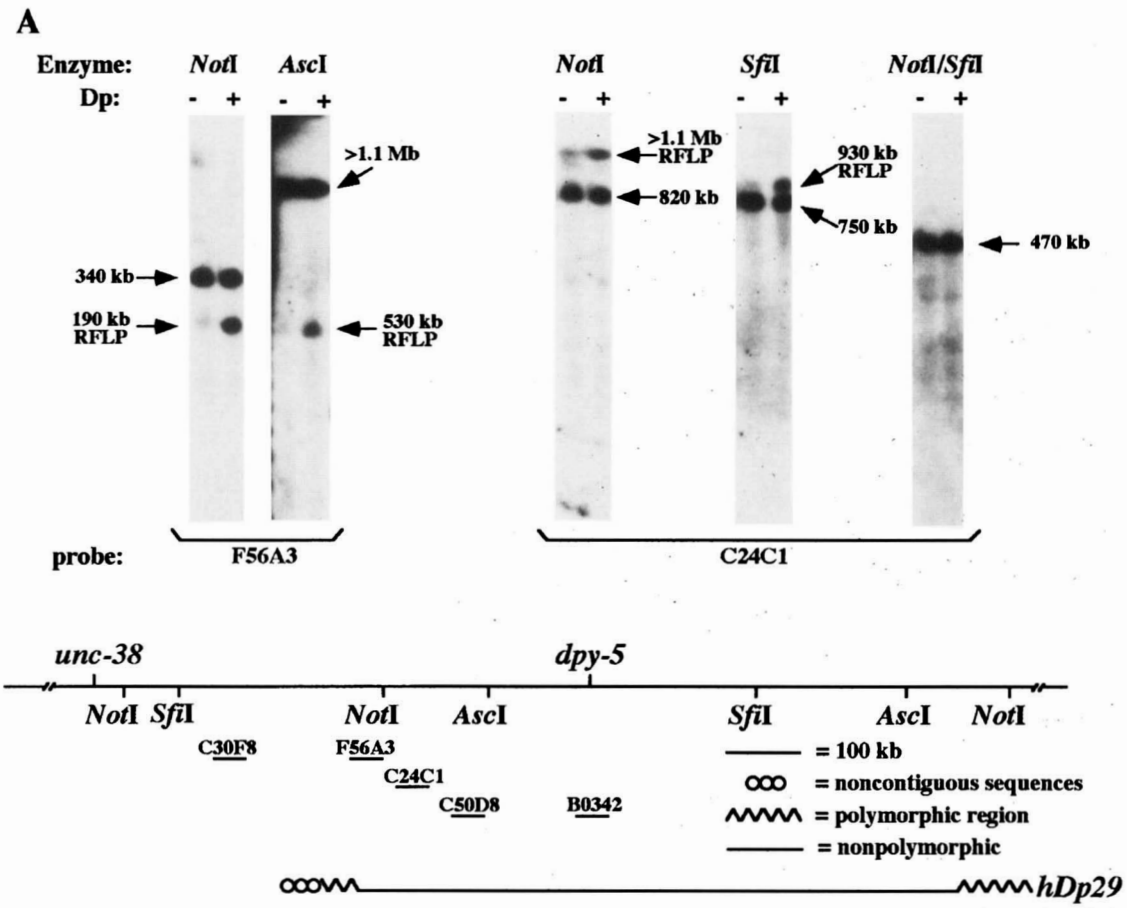

FIGURE 7.—RFLP analysis of *hDp29*. Dp + indicates DNA from strains carrying the duplication, Dp − indicates DNA from the same strain without the duplication (see MATERIALS AND METHODS for details). (A) Macro RFLP analysis. DNA from *dpy-5(e61) dpy-14(e188)* worms and *dpy-5(e61) dpy-14(e188)/ hDp29 (I;f)* worms was digested with *Not*I, *Asc*I, or *Sfi*I or double digested with *Not*I and *Sfi*I, size-fractionated by PFGE, transferred to Hybond, and probed with cosmids F56A3 and C24C1. The sizes of the fragments are indicated. (B) Diagrammatic representation of *hDp29* based on RFLP analysis. On a gross level, *hDp29* appears to be devoid of gross rearrangements ("nonpolymorphic") between the *Not*I site left of *dpy-5* and the *Asc*I site right of *dpy-5*, but contains RFLPs ("polymorphic region") outside of this region. This duplication also appears to contain noncontiguous sequences beyond its genetically defined end.

associated with the three duplications do not permit the construction of simple molecular maps for any of the three duplications, but instead suggest that they are the products of multiple rearrangements, including some deletions. These results suggest that the spontaneous shortening of duplications involves multiple breakage and ligation events. These events may involve inversion followed by deletion, as suggested by MCKIM and ROSE (1994) to explain the structure of other spontaneous duplications. Alternatively, the original duplication DNA may have broken into numerous pieces, some of which randomly ligated back together to form the new "shortened" duplication. In either case, although the cosmids used for RFLP analysis are from neighboring regions in the chromosomal DNA, they may be hybridizing to non-neighboring regions in the duplications.

In addition to internal rearrangement(s), it appears that *hDp29* also has additional noncontiguous sequences beyond its genetically defined left end. Because *hDp24* and *hDp76* are extensively rearranged, it is not possible to determine if they too have additional noncontiguous sequences. There are several possible explanations for *hDp29*'s extra sequences, including the following: (1) The duplication could be circular, in which case the additional sequences are from the other genetically defined end of the duplication. (2) This region of the duplication could be part of an inversion, leading to the juxtaposition of sequences that come from nonadjacent regions of the chromosome. (3) There could be telomeric sequences attached to the end of the duplication. (4) The duplication could be deleted for sequences that hybridize to the left-most cosmid used as a probe, but in fact extend further leftward.

Due to the complexity of the spontaneously derived duplications analyzed, they are of limited use for linking the genetic and physical maps. Furthermore, because the order of the genes between *unc-38* and *dpy-5* shown in Figure 4 relies on the assumption that the duplications are simple linear structures, these genes may not be correctly positioned. In general, gene positions that are based solely on duplication coverage may not be

accurate if the integrity of the duplication(s) is not known.

In contrast to the spontaneously derived duplications, the radiation-induced deficiencies that we analyzed, *mes-1(bn74)* and *hDf6*, appear to be devoid of gross rearrangements. The difference in the complexity of the structures of the deficiencies and duplications we analyzed may be due to the way they were generated, or it may represent a general difference between deficiencies and duplications. At least some of the molecular complexity observed in the duplications may be due to the fact that they were generated through a series of events (see above). Nonetheless, it should be noted that at least some radiation-induced deficiencies are complex: based on conventional RFLP analysis, the deficiency *hDf8* appears to be more complex than a simple deletion (McKIM *et al.* 1992).

Different approaches for molecular mapping duplications and deficiencies: The various techniques available for mapping chromosomal duplications and deficiencies have specific advantages and limitations that are directly correlated with the resolution the technique allows. *In situ* hybridization to meiotic chromosomes has low resolution but allows large regions of the genome to be analyzed. Each experiment is informative, since each hybridization can reveal whether a particular DNA sequence is located on the deficiency chromosome or is present on a free duplication (ALBERTSON 1993). However, *in situ* hybridization is unlikely to detect internal rearrangements within duplications or deficiencies unless they encompass very large regions. The macro RFLP technique described here has an intermediate level of resolution; endpoint polymorphisms can be detected from hundreds of kilobases away and internal rearrangements of moderate sizes can also be detected. Conventional RFLP analysis (KRAMER *et al.* 1988, 1990; VON MENDE *et al.* 1988) has very high resolution, allowing for the detection of small deletions or rearrangements. However, to obtain information, the probe used has to be at or very near the breakpoint of the deletion or rearrangement. With quantitative Southern hybridization analysis of DNA from animals heterozygous for a deficiency or duplication and PCR analysis of animals homozygous for a deficiency (TAX *et al.* 1994; HENDERSON *et al.* 1994), each experiment is informative in that it demonstrates whether a specific region is covered by the duplication or deficiency. However, these approaches do not reveal the overall structure of the duplication or deficiency. Deciding which technique to use depends on the resolution that is desired. In some cases, it may be advantageous to use more than one mapping procedure.

We gratefully acknowledge K. McKIM, R. ROSENBLUTH, A. HOWELL, and A. ROSE for the duplication strains and the *hDf6* strain; ALAN COULSON, JOHN SULSTON and other members of the genome project for cosmid clones; and the Caenorhabditis Genetics Center, which is funded by the National Instutes of Health (NIH) National Center for Research Resources (NCRR), for various *C. elegans* strains. This research was funded by National Science Foundation grant IBN9117559 and NIH grant GM-34059 to S.S. H.B. was supported by NIH predoctoral training grant GM-07227, L.B. by NIH postdoctoral grant GM-14599, and M.E.Z. by NIH grant GM-43930.

LITERATURE CITED

ALBERTSON, D. G., 1993 Mapping chromosome rearrangement breakpoints to the physical map of *Caenorhabditis elegans* by fluorescent *in situ* hybridization. Genetics **134:** 211–219.

ALBERTSON, D. G., and J. N. THOMSON, 1982 The kinetochores of *Caenorhabditis elegans*. Chromosoma **86:** 409–428.

AUSUBEL, F. M., R. BRENT, R. E. KINGSTON, D. D. MOORE, J. G. SEIDMAN *et al.*, (Editors), 1988 *Current Protocols in Molecular Biology*. Greene Publishing Associates and Wiley-Interscience, New York.

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

CAPOWSKI, E. E., P. MARTIN, C. GARVIN and S. STROME, 1991 Identification of grandchildless loci whose products are required for normal germ-line development in the nematode *Caenorhabditis elegans*. Genetics **129:** 1061–1072.

COHEN, D., I. CHUMAKOV and J. WEISSENBACH, 1993 A first-generation physical map of the human genome. Nature **366:** 698–701.

COULSON, A., J. SULSTON, S. BRENNER and J. KARN, 1986 Toward a physical 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.

COULSON, A., Y. KOZONO, B. LUTTERBACH, R. SHOWNKEEN, J. SULSTON *et al.*, 1991 YACs and the *C. elegans* genome. BioEssays **13:** 413–417.

COULSON, A., C. HUYNH, Y. KOZONO and R. SHOWNKEEN, 1995 The physical map of the *Caenorhabditis elegans* genome. Methods Cell Biol. **48:** 533–550.

FAN, J.-B., Y. CHILASHIGE, C. L. SMITH, O. NIWA, M. YANAGIDA *et al.*, 1988 Construction of a *NotI* restriction map of the fission yeast *Schizosaccharomyces pombe* genome. Nucleic Acids Res. **17:** 2801–2818.

FIRE, A., and R. H. WATERSTON, 1989 Proper expression of myosin genes in transgenic nematodes. EMBO J. **8:** 3419–3428.

GUO, S., and K. J. KEMPHUES, 1995 *par-1*, a gene required for establishing polarity in *C. elegans* embryos, encodes a putative ser-thr kinase that is asymmetrically distributed. Cell **81:** 611–620.

HARTL, D. L., J. W. AJIOKA, H. CAI, A. R. LOHE, E. R. LOZOVSKAYA *et al.*, 1992 Towards a *Drosophila* genome map. Trends Genet. **8:** 70–75.

HENDERSON, S. T., D. GAO, E. J. LAMBIE and J. KIMBLE, 1994 *lag-2* may encode a signaling ligand for the GLP-1 and LIN-12 receptors of *C. elegans*. Development **120:** 2913–2924.

HERMAN, R. K., 1984 Analysis of genetic mosaics of the nematode *Caenorhabditis elegans*. Genetics **108:** 165–180.

HERMAN, R. K, D. G. ALBERTSON and S. BRENNER, 1976 Chromosome rearrangements in *Caenorhabditis elegans*. Genetics **83:** 91–105.

HERMAN, R. K., J. E. MADL and K. KARI, 1979 Duplications in *Caenorhabditis elegans*. Genetics **92:** 419–435.

HODGKIN, J., R. H. A. PLASTERK and R. H. WATERSTON, 1995 The nematode *Caenorhabditis elegans* and its genome. Science **270:** 410–414.

HOHEISEL, J. D., E. MAIER, R. MOTT, L. MCCARTHY, A. V. GRIGORIEV *et al.*, 1993 High resolution cosmid and P1 maps spanning the 14 Mb genome of the fission yeast *S. pombe*. Cell **73:** 109–120.

HOWELL, A. M., and A. M. ROSE, 1990 Essential genes in the *hDf6* region of chromosome I in *Caenorhabditis elegans*. Genetics **126:** 583–592.

JOHNSON, K., and D. HIRSH, 1979 Patterns of proteins synthesized during development of *Caenorhabditis elegans*. Dev. Biol. **70:** 241–248.

KOHARA, Y., K. AKIYAMA and K. ISONO, 1987 The physical map of the whole *E. coli* chromosome: application of a new strategy for the rapid analysis and sorting of a large genomic library. Cell **50:** 495–508.

KRAMER, J. M., J. J. JOHNSON, R. S. EDGAR, C. BASCH and S. ROBERTS, 1988 The *sqt-1* gene of *C. elegans* encodes a collagen critical for organismal morphogenesis. Cell **55:** 555–565.

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

LAWRANCE, S. K., C. L. SMITH, R. SRIVASTAVA, C. R. CANTOR and S. M. WEISSMAN, 1987 Megabase-scale mapping of the HLA gene complex by pulsed field gel electrophoresis. Science **235:** 1387–1390.

LINK, A. J., and M. V. OLSON, 1991 Physical map of the *Saccharomyces cerevisiae* genome at 110-kilobase resolution. Genetics **127:** 681–698.

MCKIM, K. S., and A. M. ROSE, 1990 Chromosome I duplications in *Caenorhabditis elegans*. Genetics **124:** 115–132.

MCKIM, K. S., and A. M. ROSE. 1994 Spontaneous duplication loss and breakage in *Caenorhabditis elegans*. Genome **37:** 595–606.

MCKIM, K. S., T. STARR and A. M. ROSE, 1992 Genetic and molecular analysis of the *dpy-14* region in *Caenorhabditis elegans*. Mol. Gen. Genet. **233:** 241–251.

MENEELY, P. M., and R. K. HERMAN, 1979 Lethals, steriles, and deficiencies in a region of the *X* chromosome of *Caenorhabditis elegans*. Genetics **92:** 99–115.

MIZUKAMI, T., W. I. CHANG, I. GARKAVTSEV, N. KAPLAN, D. LOMBARDI *et al.*, 1993 A 13 kb resolution cosmid map of the 14 Mb fission yeast genome by nonrandom sequence-tagged site mapping. Cell **73:** 121–132.

MORTIMER, R. K., C. R. CONTOPOULOU and J. S. KING, 1992 Genetic and physical maps of *Saccharomyces cerevisiae*, edition 11. Yeast **8:** 817–902.

OLSON, M. V., J. E. DUTCHIK, M. Y. GRAHAM, G. M. BRODEUR, C. HELMS *et al.*, 1986 Random-clone strategy for genomic restriction mapping in yeast. Proc. Natl. Acad. Sci. USA **83:** 7826–7830.

PAULSEN, J. E., E. E. CAPOWSKI and S. STROME, 1995 Phenotypic and molecular analysis of *mes-3*, a maternal-effect gene required for proliferation and viability of the germ line in *C. elegans*. Genetics **141:** 1383–1398.

SAMBROOK, J., E. F. FRITSCH and T. MANIATIS, 1989 *Molecular Cloning: A Laboratory Manual*, Ed. 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

SIGURDSON, D. C., G. J. SPANIER and R. K. HERMAN, 1984 *Caenorhabditis elegans* deficiency mapping. Genetics **108:** 331–345.

SMITH, C. L., J. G. ECONOME, A. SCHUTT, S. KLCO and C. R. CANTOR, 1987 A physical map of the *Escherichia coli* K12 genome. Science **236:** 1448–1453.

SPIETH, J., M. MACMORRIS, S. BROVERMAN, S. GREENSPOON and T. BLUMENTHAL, 1988 Regulated expression of a vitellogenin fusion gene in transgenic nematodes. Dev. Biol. **130:** 285–293.

STROME, S., P. MARTIN, E. SCHIERENBERG and J. PAULSEN, 1995 Transformation of the germ line into muscle in *mes-1* mutant embryos of *C. elegans*. Development **121:** 2961–2972.

SULSTON, J., Z. DU, K. THOMAS, R. WILSON, L. HILLIER *et al.*, 1992 The *C. elegans* genome sequencing project: a beginning. Nature **356:** 37–41.

TAX, F. E., J. J. YEARGERS and J. H. THOMAS, 1994 Sequence of *C. elegans lag-2* reveals a cell-signalling domain shared with *Delta* and *Serrate* of *Drosophila*. Nature **368:** 150–154.

VILLENEUVE, A. M., and B. J. MEYER, 1990 The role of *sdc-1* in the sex determination and dosage compensation decisions in *Caenorhabditis elegans*. Genetics **124:** 91–114.

VON MENDE, N., D. M. BIRD, P. S. ALBERT and D. L. RIDDLE, 1988 *dpy-13*: a nematode collagen gene that affects body shape. Cell **55:** 567–576.

WILSON, R, R. AINSCOUGH, K. ANDERSON, C. BAYNES, M. BERKS *et al.*, 1994 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. Nature **368:** 32–38.

Communicating editor: R. K. HERMAN