Chromosomal Walking and Jumping to Isolate DNA from the *Ace* and *rosy* Loci and the Bithorax Complex in *Drosophila melanogaster*

Welcome Bender†, Pierre Spierer‡ and David S. Hogness§

Department of Biochemistry
Stanford University School of Medicine
Stanford, Calif. 94305, U.S.A.

A chromosomal walk is described that covers 315×10^3 base-pairs of DNA from the 87DE region of the third chromosome of *Drosophila melanogaster*. The walk includes the DNA for the *rosy* and *Ace* loci, which code for xanthine dehydrogenase and acetylcholinesterase, respectively. Several dispersed repetitive elements were encountered in the walk. In every case, their positions in the chromosome differed in different strains, and so they are all presumed to be transposable elements. Several rearrangement breakpoints have been localized within the walk, including the break for In(3R) Cbx^{+R1} (87E1, 2–89E1, 2). One breakpoint fusion fragment of this inversion was isolated to jump from 87E into the cluster of homeotic genes of the bithorax complex, at 89E1–4.

1. Introduction

The decades of research with *Drosophila melanogaster* have produced numerous mutations with unusual genetic properties or striking developmental effects. These strains are now particularly promising resources for molecular biologists who are trying to understand how developmental decisions are controlled by the actions of macromolecules. The gene products of loci with interesting mutants can be studied directly if the mutated region of the chromosome can be isolated on recombinant DNA molecules. Until recently, it has been necessary to identify the gene product in advance to acquire a probe for the DNA sequence. But in *D. melanogaster*, we can take advantage of the polytene chromosomes and the excellent cytogenetics to isolate genes by virtue of their chromosomal position, when nothing is known of their molecular function. We describe here the isolation of the DNA for three loci of particular interest: *rosy*, the gene for xanthine

---

† Present address: Department of Biological Chemistry, Harvard Medical School, Boston, Mass. 02115, U.S.A.

‡ Present address: Department of Molecular Biology, University of Geneva, 30 quai Ernest-Ansermet, 1211 Geneva 4, Switzerland.

§ Author to whom reprint requests should be addressed.

dehydrogenase, which has been the subject of intensive genetic fine-structure mapping (Chovnick *et al.*, 1977); *Ace*, the gene for acetylcholinesterase (Hall & Kankel, 1976), an enzyme that is hormone-inducible in tissue culture cells (Cherbas *et al.*, 1977); and the bithorax complex, a cluster of genes affecting segmental determination (Lewis, 1978). The gene products of the bithorax complex are unknown, while the protein products of *rosy* and *Ace* are present in low amounts, so that isolation and translation of their messenger RNAs would be quite difficult.

The strategy we have used is called chromosomal walking and jumping; it is shown diagrammatically in Figure 1. The chromosomal origin of any non-repeated segment of *D. melanogaster* DNA (Dm segment) can be determined by *in situ* hybridization of that DNA to polytene chromosomes. When the sites of hybridization are visualized by tritium autoradiography, the position is usually confined to one or a few bands, which is similar to the precision of the cytological localizations of rearrangement breakpoints or the localizations of well-mapped genes. If a DNA sequence is found within a few bands of a gene of interest, that sequence can be used as the starting point for a chromosomal walk to the gene. A "step" in the walking procedure involves screening a recombinant DNA library of random large Dm segments to collect those that overlap the starting point. The
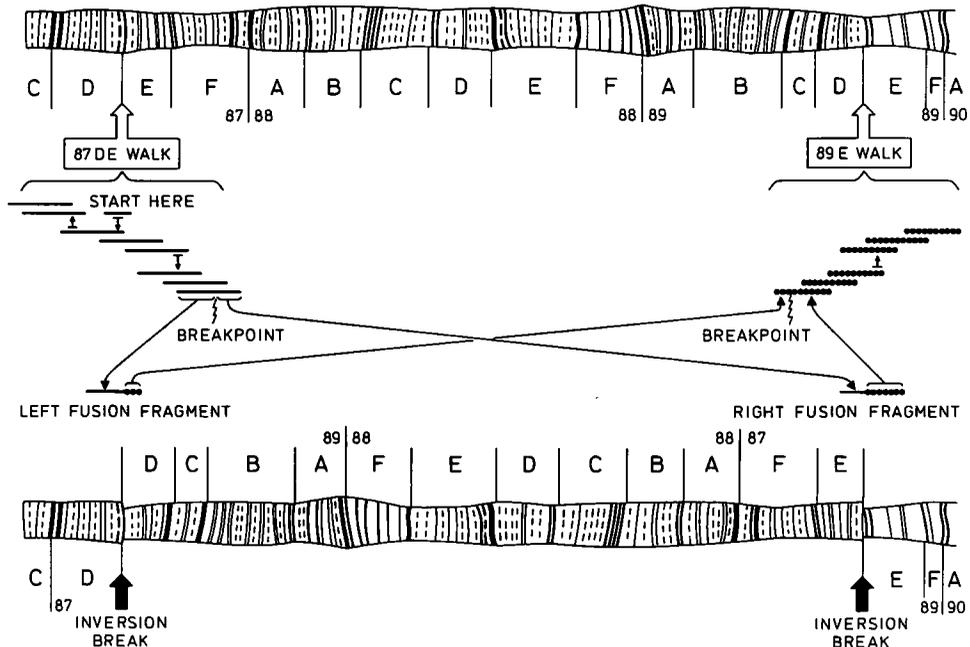

FIG. 1. The strategy for walking and jumping. The upper chromosome represents a portion of the right arm of the third chromosome with normal cytology (drawn from the map of Bridges, 1941), and the lower chromosome has an inversion of the region from 87E to 89E. A few steps of a chromosomal walk are shown diagrammatically below the 87E region (not to scale with the chromosome). When the walk reached the site of the inversion breakpoint, the DNA from that position could be used to identify the two fusion fragments isolated from the inversion chromosome. The foreign DNA in the fusion fragments (tandem circles) was homologous to normal chromosomal DNA at the right or distal inversion breakpoint, and thus it served as the origin of a chromosomal walk in 89E.

new segments are mapped and aligned, and the sequences farthest from the starting point in both directions are used as new starting points. A succession of such steps will collect the DNA from a long contiguous region of the chromosome. Once the walk covers more than one or two bands, the chromosomal orientation of the walk can be determined by *in situ* hybridization with probes from either end of the walk.

Jumping uses chromosomal rearrangements to shift the walk to new chromosomal positions. An inversion is shown diagrammatically in Figure 1, but translocations and duplications are equally useful. One can determine when the DNA sequence of a walk crosses the position of a cytological breakpoint by hybridizing individual segments from along the walk to the rearranged chromosome. If a breakpoint site is found, the DNA of the animal with the rearranged chromosome can be used to construct a library, and the rearrangement library screened with a probe from the wild-type chromosome at the site of the breakpoint. The resulting clones should include a breakpoint fusion fragment. The new DNA sequence in the fusion fragment will, of course, hybridize to a new position on the wild-type chromosome, and this sequence can be used as the starting point for a new chromosomal walk at that position.

2. Materials and Methods

(a) *Preparation of probes*

Phage DNA (10 to 30 μg) was digested to completion with a suitable restriction enzyme and the fragments were separated on a preparative agarose gel. The desired fragment was electroeluted into a dialysis bag (Maniatis *et al.*, 1982), and the DNA was recovered by successive precipitations with isopropanol and ethanol. The DNA was resuspended in 10 to 20 μl of nick-translation buffer (Rigby *et al.*, 1977) with 50 μCi of [^{32}P]dCTP and 5 to 10 units of polymerase I; nick-translations were left at 12°C for 3 h. Labeled DNA was separated from unincorporated nucleotides by chromatography on Biogel P60; typically, 10^7 cts/min were incorporated.

(b) *Screening phage libraries*

The Canton S and Oregon R libraries (see the text) were plated on 10 cm Petri dishes so that the plaques were nearly confluent (5 to 10,000 plaques/plate). For each step, a total of 20 plates was screened. Single nitrocellulose filter replicas were prepared from each plate by the method of Benton & Davis (1977). The filters were prehybridized in 50% formamide, 5 × SSPE (SSPE is 0·15 M-NaCl, 10 mM-sodium phosphate, pH 7·0, 2 mM-EDTA), 1 × Denhardts (1966) solution, 0·1 mg salmon sperm DNA/ml, and 0·1% sodium dodecyl sulfate (Maniatis *et al.*, 1982). The nick-translated DNA was melted and prepared in 10 ml of the same solution. Hybridizations were at 42°C for 12 to 24 h. The filters were washed (Benton & Davis, 1977) and autoradiographed.

The positive spots on the film could not be assigned to a single plaque on the first round of plates, and so a plug of agar about 5 mm in diameter was removed from positive sites on the plates. These plugs were eluted in 1 ml PSB (0·1 M-NaCl, 10 mM-Tris, pH 7·4, 10 mM-MgCl$_2$, 0·05% gelatin), and the phage suspension was replated at a density of about 100 plaques/plate. Filters were prepared from the second plates and probes as before. Single positive plaques from the second plates were stabbed with a toothpick, and the toothpick was then touched repeatedly to a freshly poured bacterial lawn to create a zone of confluent lysis 2 to 3 cm in diameter. The top agar of the lysis zone was scraped off the plate and

mixed in 2 ml PSB. This suspension usually contained about 10^9 plaque-forming units (p.f.u.)/ml.

(c) Phage plate stocks

Phage were plated on three fresh 15 cm LB plates at about 10^7 p.f.u./plate. Phage were harvested by the procedure of Davis *et al.* (1980). The phage suspension was further purified by banding on 2 sequential CsCl step-gradients according to Davis *et al.* (1980), except that the phage were banded between CsCl step-densities of 1·50 and 1·45 g/cm^3.

(d) Heteroduplex and restriction mapping

To prepare heteroduplexes for electron microscopy, about 0·1 μl of each phage preparation in CsCl (containing about 10 ng DNA) was added to 10 μl of 75% formamide, 0·25 M-NaCl, 0·1 M-Tris, 0·01 M-EDTA (pH 8·5). The mixture was heated to 90°C for 30 s, then cooled to room temperature and left for 15 to 30 min. The mixture was then diluted to 50 μl, to give a final composition of 55% formamide, 0·05 M-NaCl, 0·06 M-Tris, 0·006 M-EDTA, and 60 μg cytochrome *c*/ml (pH 8·5). This solution was spread onto a hypophase of 27% formamide, 10 mM-Tris, 1 mM-EDTA (pH 8·5), according to the procedure of Davis *et al.* (1971). Single-stranded circles of ϕX phage were included for a single-stranded length standard; double-stranded lengths were compared to duplex regions formed by the phage vector arms.

Each new phage was tested against the most distal phage from the previous step. When the orientation relative to the vector arms differed between the new phage and the test phage, the heteroduplex was repeated with a phage from the previous step in the opposite orientation. When possible, a phage in the Charon 4 vector was compared with one in the Sep 6 vector. The two vectors have a region of non-homology in their right arms next to the *D. melanogaster* sequences. Measurement of the 2 single strands of the resulting "substitution loop" can resolve the left–right ambiguity in the relative chromosomal positions of the 2 Dm segments.

Phages were lysed to give phage DNA solutions by dialysis against 50% formamide (Davis *et al.*, 1980). All the phages in any given step were digested with *Eco*RI, and the restriction fragment patterns together with the heteroduplex information were usually sufficient to align all the phages. The phage from the new step that extended furthest along the chromosome was carefully mapped with *Eco*RI, *Bam*HI, *Hind*III and *Sal*I, usually by running all possible single and double digests on the same gel. *Hind*III digests of phage lambda (Davis *et al.*, 1980) were used as length standards. Restriction fragments less than 0·5 kb† were usually not visible on the gels; hence, such small fragments would have been missed in most of our mapping.

(e) DNA from mutant flies

DNA was prepared from larvae or adults of various mutant stocks by the procedure of R. Lifton (personal communication). Up to 100 flies were added to 2 ml of 0·1 M-NaCl, 0·2 M-sucrose, 0·1 M-Tris·HCl, 0·05 M-EDTA (pH 9·1), with 0·5% sodium dodecyl sulfate and 1% diethyl pyrocarbonate. The flies were ground quickly in a glass homogenizer and the slurry was incubated at 65°C for 30 min. Then 0·3 ml of 8 M-potassium acetate was added and the mixture was kept at 0°C for 30 min. The slurry was centrifuged for 5 min at 10,000 *g* and the supernatant was recovered and mixed with an equal volume of ethanol. The ethanol mixture was left for 5 min at 25°C, then centrifuged for 5 min at 10,000 *g*. The precipitate was washed once in 80% ethanol, dried under vacuum, and resuspended in 0·4 ml 0·01 M-Tris, 0·001 M-EDTA (pH 7·5). The suspensions were cloudy white or pink,

† Abbreviation used: kb, 10^3 base-pairs.

presumably because of carbohydrates remaining in the preparation, but the mixture usually worked for restriction digestions without further purification. The typical yield from 100 adult flies was 30 μg of DNA. For one track on a whole genome southern†, 40 μl (3 μg DNA) were digested simultaneously with RNase (2 μg/ml) and restriction enzyme.

(f) In situ hybridizations

The procedures for *in situ* hybridization experiments are described in the accompanying paper (Spierer et al., 1983).

3. Results

(a) Walking in 87D–E

We were fortunate to have as a starting point for the walk a plasmid named cDm758, which contained a Dm segment from the 87E region of the chromosome. cDm758 was isolated as an accidental product of a screen for another genetic locus (Lis et al., 1978). By *in situ* hybridization, it failed to hybridize to the two deletion chromosomes, Df(3R)126c and Df(3R)126d (Fig. 2), and hence it was located
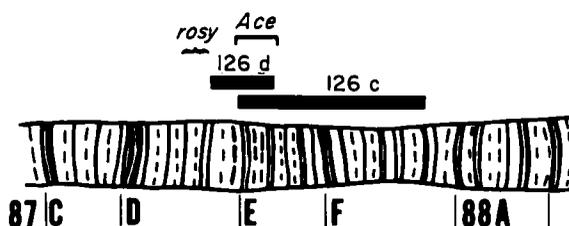

FIG. 2. Cytological localizations in 87D–E. The extents of the two deficiencies Df(3R)126c and Df(3R)126d are indicated by horizontal bars. The region where they overlap (87E1–6) includes the locus coding for acetylcholinesterase (*Ace*) and also included DNA homologous to cDm758, the starting point of the chromosomal walk. The locus coding for xanthine dehydrogenase (*rosy*) is just to the left of the 126d deficiency, in 87D10–12.

within the bands 87E1–6 (J. T. Lis, E. Meyerowitz & P. Spierer, unpublished results). cDm758 was used to screen 2 libraries of Dm segments in lambda phage vectors. The first library contained segments of DNA from the Oregon R strain inserted into the vector Sep 6 by tailing with A·T or G·C (Meyerowitz & Hogness, 1982). In the second library, Dm segments from the Canton S strain were inserted into the vector Charon 4 with *Eco*RI linkers (Maniatis et al., 1978). For both libraries, the Dm segments were generated by shearing high molecular weight DNA, and so the fragment endpoints should be random. Both libraries were used throughout the subsequent walk, although the Canton S library proved more useful because it had more independent recombinants, and because the *Eco*RI linkers allowed us to isolate the DNA at the very edge of a Dm segment without contamination by sequences of the vector arms.

The walk is shown diagrammatically in Figure 3. The central line in the Figure is marked off in base-pairs $\times 10^3$ (kb) from the starting point, which is the position of homology to cDm758. cDm758 was used to probe the Oregon R library, and the

† A whole genome southern is a Southern blot (Southern, 1975) of restricted genomic DNA.

FIG. 3. Dm segments of the walk with composite restriction map. The two boldface lines in the center of the Figure are the composite restriction maps for Canton S (upper line) and Oregon R (lower, discontinuous line). The lines are marked off in kb from the starting point of the walk. The restriction enzyme sites are designated: B (BamHI), H (HindIII), R (EcoRI), and S (SalI). Where the Oregon R map differs from the Canton S map, extra sites or missing sites (minus sign) are marked below the Oregon R line. The different Dm segments are indicated above and below the restriction maps; segments numbered in the 2100 or 2800 series are from the Canton S library, while the 1900 series is from Oregon R. The L or R at the right end of each clone line indicates whether that end joins with the left or right arm of the lambda phage vector. Insertions of extra DNA relative to the Canton map are indicated by triangles, where the horizontal side of the triangle indicates the insert length. DNA present in Canton S but absent in Oregon R is indicated by parentheses.

clone λbDm1901 was isolated and its Dm1901 segment mapped relative to Dm758. All of the fragments of Dm1901 were used to screen both the Oregon R and Canton S libraries; this screening yielded the phage λbDm1905–08 and λcDM2101–10. This first step covered about 30 kb. (Dm segments from the Oregon R library have numbers in the 1900 series, while Canton S segments are numbered in the 2100 and 2800 series; λb and λc refer to the Sep 6 and Charon 4 vectors, respectively. Henceforth, only the numbers will be used to refer to the phage and their Dm segments, except where they need to be distinguished.)

For the second step, the leftmost restriction fragment of the 1907 segment was used as probe, and we recovered 1911 to 15 and 2112 to 17, which extended the walk 14 kb to the left. Likewise, the rightmost fragment of 2106 was used to screen the libraries for a rightward step, but here we stepped into our first repeat. We found many more plaques hybridizing to the probe than we expected. A few of these phage were grown up (1916 to 1922) and compared to 2106 and to each other. As Figure 3 illustrates schematically, 1916, 1920 and 1921 all matched 2106 for about 2·5 kb and matched each other for about 5 kb, as determined by electron microscopy of heteroduplexes. But the DNA flanking the homology region did not match in any pair tested. The probe from the right end of 2106 was hybridized to HindIII-cut Oregon R DNA on a whole genome southern, and an array of about 30 bands was labeled. Adjacent regions of 2106 labeled only single bands on whole genome southerns, as expected for single-copy DNA. Thus it was clear that the right end of the 2106 insert contained sequences repeated within the genome, and we had isolated in 1916 to 1922 copies of the repeated DNA from other chromosome locations. We named the repeated element Kermit.

The portion of 2106 with repetitive DNA could be guessed by its homology to 1916 etc. We therefore retreated backwards and isolated for a probe the EcoRI fragment spanning +2·4 to +8·7 kb; with this we screened a collection of phages isolated by homology to the repetitive fragment, looking for the copy of the repeat flanked by DNA of 87E. One phage, 2111, was recovered that contained the complete 5 kb Kermit element (defined by homology with 1916 etc.) and also included 4·5 kb of single-copy DNA of the far side (Fig. 3). A restriction fragment from this region was then used as probe for the next step to the right, which proceeded without difficulty.

In the analysis of 2111 and the phage of the subsequent step, we discovered that the Kermit repeat of 2111 was not present at this chromosomal position in the Oregon R strain. The 1908 segment was missing the repetitive DNA but had 0·5 kb of single-copy sequence from beyond the repeated sequence of 2111. A whole genome southern with HindIII-cut Oregon R DNA was probed with sequences flanking the repeat to confirm that the repeat was missing. In retrospect, we could have avoided the Kermit repeat by switching to the Oregon R library. Other blocks of repeated DNA encountered in the walk also proved to be evolutionarily mobile; they were missing from one strain at the chromosomal location where they were found in the other. We continued to use both libraries throughout the walk, and took advantage of this mobility to identify repeats and to step over them.

The walk was continued in both directions until we reached landmarks that

were known to be beyond the *rosy* locus on the left and beyond the *Ace* locus on the right. On the left, we walked past the distal endpoint of Df(3R)*kar*^{1G27} at about -190 kb; this large deficiency ends in 87D just to the left of (proximal to) *S12*, the complementation group just proximal to *rosy* (Hilliker *et al.*, 1980). On the right, the walk spanned the distal endpoint of Df(3R)126d at about $+90$ kb; this deficiency deletes *Ace* and two complementation groups distal to *Ace* (Hilliker *et al.*, 1980). These deficiency endpoints were identified by *in situ* hybridizations and whole genome southerns with various probes along the walk, as described in the accompanying paper (Spierer *et al.*, 1983).

A step-by-step description of the walk is unnecessary, but several interesting features along the map should be noted. These will be listed in chromosomal order going from left (centromere proximal) to right (distal) on Figure 3.

-185.5. *Hind*III site present in Oregon R only.

-178. *Bam*HI site present in Oregon R only.

-163.5 to -160.5. Insertion of repeated DNA element in the strain ry^{+11}. This element, which we name "Jiminy", is also found in Oregon R DNA at position -92 in the walk (see below). While genome southerns to *Eco*RI-cut DNA from Oregon R or Canton S labeled over 100 bands.

-159. Position of a heat-shock cognate gene analogous to that defined by Ingolia & Craig (1982). Their clone designated MG21 contains an 8.4 kb *Bam*HI fragment that matches the *Bam*HI fragment from -163 to -154.7 kb. MG21 was isolated from a library of Schneider cell DNA fragments by homology to the coding region for the 70,000 molecular weight heat-shock protein. The *Eco*RI site at -159 cuts within the region of homology.

-131. Insertion of 8 kb mobile repeat present in only some chromosomes of Oregon R. The element is an exact copy of the "roo" element (Meyerowitz & Hogness, 1982) as judged by restriction mapping and electron microscopy of heteroduplexes with the prototype copy of roo in the phage λbDm 2030. The Oregon clones 1965, 1967, 1968 and pPW126 have no insert here.

-127. Insertion of mobile repeat present in only some chromosomes of Oregon R. The clone pPW126 was described by Wensink *et al.* (1979) as a "scrambled cluster" of repeat elements. A small single-copy fragment from this plasmid was shown by Wensink *et al.*, to hybridize to the chromosomal region 87D10–E1. We searched the clones of the walk for homology to this fragment of pPW126, and found hybridization to the region of -130 kb. Subsequent comparison by restriction mapping and heteroduplex analysis showed that pPW126 matched the walk exactly from -135 to -127 kb. (This corresponds to the right half of the insert of pPW126 as shown in Figure 1 of Wensink *et al.* The 0.9 kb *Eco*RI fragment noted in that Figure should be located to the right of the 2.4 kb *Eco*RI fragment; these correspond to the 0.8 and 2.7 kb *Eco*RI fragments in the -134 to -130 kb region of our map.) The remaining 8.5 kb of the pPW126 insert (the portion nearest the tetracycline-resistance region of pMB9) does not hybridize with the 2823 segment and is presumed to be repetitive from its homology to the other "scrambled clusters". The 2199 and 2821 segments both span the insertion site of the pPW126 mobile repeat, and both gave only single-copy bands when used to probe whole genome southerns.

−112. The 2801 segment matches the other segments from −115 to −112 kb, but then has 15 kb of non-homologous sequences. This either reflects heterogeneity within the Canton S DNA (a large mobile repeat?) or an artefact of cloning in which two unrelated genomic fragments were fused together. The latter possibility seems more likely, since five other Canton phage covering this region show no such anomaly, and since there is generally little heterogeneity in the Canton S clones (see Discussion).

−92 and −88. Mobile repetitive sequences present in only some Oregon chromosomes. In 1957, there is an insertion of the Jiminy element at −92 kb. At −88 kb, in the phages 1957 and 1956, there is an insertion of the rightmost 2 kb of the Jiminy element again, but in inverted orientation. Thus, single strands of the phage 1957 visualized in the electron microscope form a stem-and-loop structure. The loop between the two inverted repeats consists of the normal single-copy DNA found in Canton chromosomes between −92 and −88. Probes from 1958 spanning both insert sites show single-copy bands with whole genome southerns of Canton S DNA and doublet bands with Oregon R reflecting the heterogeneity within Oregon R chromosomes.

−66. Insertion of 2·3 kb mobile repeat in some chromosomes of Oregon R. This is a shortened copy of the 4·4 kb repetitive element named "Doc", which has been found twice in spontaneous mutants of the bithorax complex (W. Bender *et al.*, 1983). Doc element probes reveal approximately 20 bands in whole genome southerns of Oregon R DNA cut with *Bam*HI. With 2176 as probe, only the expected single-copy bands are seen with Canton S DNA while doublets are seen with Oregon R.

−62·5. Site for *Sal*I absent in some Oregon R chromosomes. 1944 and 1955 lacked this site, but whole genome southerns on *Sal*I-cut Oregon R DNA with 2176 showed that the site was present in about half the Oregon R chromosomes.

−25. Site for *Hind*III present in Oregon R but not in Canton S clones. The strain difference was confirmed by whole genome southerns with 2112 as probe.

−18 to −17. *Bam*HI to *Hind*III fragment larger by 0·5 kb in some Oregon R chromosomes. The insert has not been checked for genomic copy number.

−13 to −11. *Bam*HI to *Sal*I fragment larger by 0·5 kb in some Oregon R chromosomes. The insert has also not been checked for genomic copy number.

+9 to +14. Kermit mobile repeat present in Canton S only, as described above. Whole genome southerns showed about 30 copies present in Oregon R DNA. *In situ* hybridizations with Kermit to Oregon R chromosomes show about 30 sites of labeling including the chromocenter (A. Spierer, personal communication). Six copies were isolated from the Oregon library on 1916 to 1918 and 1920 to 1922. The Kermit copies on 1916, 1920 and 1921 matched the 2111 copy exactly, as judged by microscopy of heteroduplexes; 1917 and 1918 both showed an internal deletion of about 1·3 kb, beginning about 1·1 kb from the left end of the element. These deleted Kermit copies appear analogous to the deleted variants described for the copia element (Young & Schwartz, 1981). The copy in 1922 is flanked by inverted repeats each about 2·6 kb in length. The resulting stem-and-loop structures prevented accurate comparison of the repeats in 2111 and 1922. The copia element has also been found flanked by inverted repeats

(Carlson, 1979). The Kermit element has apparent terminal repeats, in that a probe from the end of 2110 overlapping the left boundary weakly labeled the *Eco*RI fragment of 2111 overlapping the right boundary. No small hybrid regions were seen in heteroduplex molecules designed to reveal stable inverted or direct repeats, and so the terminal repeat must be quite small or badly mismatched. The 2130 phage deleted the Kermit element during growth in *Escherichia coli*. The phage stock grown from a single plaque gave two bands in a cesium chloride density gradient; the phage in the lighter band lacked the 5 kb repeat, but was otherwise identical to the heavy phage. 2121 was the only Canton S phage from this region that completely lacked the repeat. We suspect 2121 also lost Kermit on growth in *E. coli*, but it could reflect heterogeneity in the Canton S DNA from which the library was made.

+ 37·5. Apparent fusion of unrelated fragments in 2129. The right-hand 3·5 kb of this phage shows no homology to the other Oregon R or Canton S phage covering this region. The discontinuity begins at an *Eco*RI site. A whole genome southern of *Eco*RI-cut Canton S DNA with 2127 as probe shows no 4 kb band as found in 2129.

+ 38. Heterogeneity in Oregon R. The two small *Hind*III fragments here measure 0·5 and 0·55 kb in the Canton S segments and in 1930. In 1933, they measure 0·5 and 0·62 kb.

+ 42·5. *Hind*III site present in Oregon R but not Canton S DNA.

+ 42·9 to + 47·5. Insertion of 4·6 kb of repetitive element in Canton S but not Oregon R. Whole genome southerns to Oregon R DNA suggest about 100 copies per genome. The position of the insertion site is not certain relative to the extra *Hind*III site at + 42·5 mentioned above.

+ 49·5. *Hind*III site present in Oregon R but not in Canton S DNA. The *Eco*RI fragment including this *Hind*III site is slightly larger in Oregon R than in Canton S, 1·50 *versus* 1·45 kb.

+ 86 to + 88·5 and + 90 to + 92·5. Tandem repeat. A subclone containing the region from + 85 to + 89 hybridized to a second subclone containing + 90 to + 97. In heteroduplexes between these subclones, there were two short duplex regions separated by a single-stranded bubble, indicating that the central 0·8 to 0·9 kb of the tandem repeat is badly mismatched. Whole genome southerns probed with 2186, which includes both copies, show only the bands expected from the 2186 restriction map; hence, the elements of this tandem repeat are not found elsewhere in the genome.

+ 89. *Hind*III site present in Oregon R but not in Canton S DNA.

(b) *Jumping into the bithorax complex*

The original incentives for carrying out the 87DE walk were twofold: to isolate DNA containing the *rosy* and *Ace* loci, and to develop the technique, as this was the first attempt at a long chromosomal walk. While the walk was in progress, we learned from Ed Lewis of inversions with breakpoints in 87D–E and in the bithorax complex, and this suggested to us the possibility of jumping from 87D–E into the complex. Three inversions were available; all were *Ubx* mutants

generated by X-ray irradiation of a *Cbx* stock. Breakpoints with a *Ubx* phenotype *cis*-inactivate the dominant *Cbx* phenotype (Lewis, 1964). *Cbx Ubx*[21987A] and *Cbx Ubx*[21988B] were inversions isolated by T. Ramey & E. B. Lewis, with the proximal inversion breakpoints located cytologically in distal 87E and distal 87D, respectively. *Cbx*[+R1] was isolated by T. Kaufman and had an inversion breakpoint in proximal 87E. All three had the distal inversion breakpoint in 89E, as expected from the *Ubx* phenotypes.

The clearest method for locating potential rearrangement breakpoints within the walk was to do *in situ* hybridizations of clones from the walk to the rearranged chromosomes. Unfortunately, all three inversions were lethal when homozygous, certainly because of the *Ubx* mutations, and potentially because of other lethal hits. When such large inversions are heterozygous over non-inverted chromosomes, the chromosomes pair in an inversion loop. The two ends of the inversion are held together at the base of the loop, and so it is difficult to determine by *in situ* hybridizations which end of the inversion is labeled. In order to make the chromosomes cytologically homozygous, we attempted to rescue the known *Ubx* lethality with a duplication for the bithorax complex. When Dp(3;2)P10 (Lewis, 1978) was added, the *Cbx Ubx*[21987A] and *Cbx*[+R1] inversions
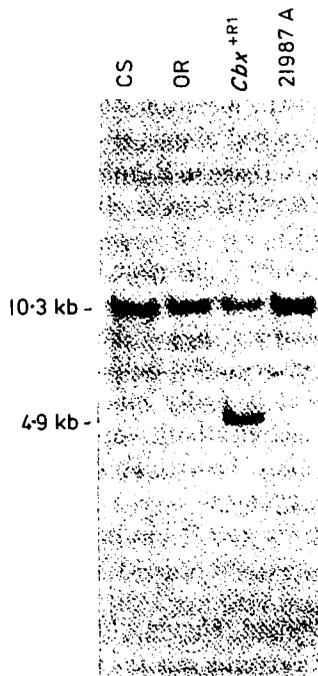

FIG. 4. Whole genome southern of inversion breakpoint. The 4 lanes had *Eco*RI-digested DNA from Canton S, Oregon R, the *Cbx*[+R1] inversion and the 21987A inversion, respectively. The blot was probed with the 10·3 kb *Eco*RI fragment from −6·4 to −16·7 in the walk. In the *Cbx*[+R1] lane, the normal 10·3 kb band is replaced by 2 bands at 4·9 and 10·4 kb. Both inversions were induced on the same background chromosome (*Cbx*¹). Both inversions were made homozygous in the presence of Dp(3;2)P32, which complements the lethal breakpoint in the bithorax complex.

could be made homozygous. *Cbx Ubx*[21988B] could not, and it was subsequently found to fail to complement lethal alleles at the *pic* locus (Hilliker *et al.*, 1980). The *pic* breakpoint is at about −150 kb, which was beyond the boundaries of the walk at the time the jump experiments were initiated.

Squashes were prepared from the two homozygous inversion stocks, and these were hybridized with two probes, which represented the ends of the walk at that time (2179 and 2175 at −80 and +85). Both probes hybridized to the proximal end of the *Cbx Ubx*[21987A] inversion, indicating that the inversion breakpoint in 87E was distal to the existing walk. But with the *Cbx*[+R1] inversion, 2179 gave grains at the proximal end and 2175 gave grains at the distal end, demonstrating that the inversion broke somewhere within the walk. We then repeated the *in situ* hybridizations to *Cbx*[+R1] with three probes distributed along the walk (1933 at +40 kb, 2121 at +10 kb, and 2113 at −10 kb) to locate the inversion break more precisely. The 1933 and 2121 probes both hybridized only to the distal endpoint, but 2113 gave grains at both ends of the inversion. Thus the inversion splits the walk within the 2113 segment (−18 to −1 kb).

We had been simultaneously looking for the inversion breakpoints by doing whole genome southerns to the rearrangement DNA using probes along the walk. This approach was confusing because strain heterogeneities produced anomalous bands in several places. But once the breakpoint was located by the *in situ* results, we repeated the southerns with 2113 and found the apparent breakpoint fusion fragments. Figure 4 shows a southern of *Eco*RI-digested DNA from two wild types and the two homozygous inversion stocks probed with the 10·3 kb *Eco*RI fragment from the middle of 2113. In the *Cbx*[+R1] track, the expected 10·3 kb band is replaced by two fainter bands, one slightly above 10·3 kb and one at about 4·9 kb. We chose to reclone the larger (10·4 kb) fusion fragment to jump across the inversion breakpoint.
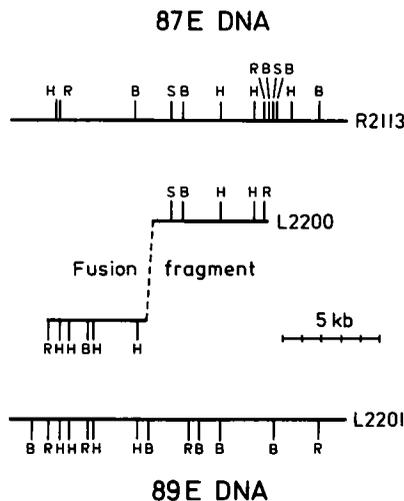

FIG. 5. Restriction map of the cloned fusion fragment. The 2200 segment, containing the distal breakpoint fusion fragment of the *Cbx*[+R1] inversion, is compared to clones from the Canton S library covering the breakpoint regions in 87E (2113) and 89E (2201). The maps are drawn as in Fig. 3.

FIG. 6. *In situ* hybridization of the cloned fusion fragment. DNA from the 2200 phage was nick-translated with $[^{125}I]dCTP$ and hybridized to salivary chromosomes from a gt^1/gt^{x11} larva. Grain clusters are indicated at both 87E and 89E.

The lambda phage vector Charon 13 was used to make a small library of recombinant molecules including the fusion fragment. Charon 13 has a capacity for *Eco*RI fragments greater than 2·4 and less than 16·7 kb (Blattner *et al.*, 1977), which was ideal for the desired 10·4 kb fragment. Cbx^{+R1} DNA was cut to completion with *Eco*RI and fragments of approximately 10 kb were separated by sucrose gradient centrifugation. The two arms of the Charon 13 vector were also purified by sucrose gradients. Arms and inserts were ligated with bacteriophage T4 ligase, transfected into $CaCl_2$-shocked cells (Davis *et al.*, 1980), and plated. About 3000 plaques were produced, and these were screened with the labeled 10·3 kb *Eco*RI fragment isolated from 2113. One plaque was positive, and this phage (2200) was compared with the 2113 DNA region by restriction mapping and by electron microscopy of heteroduplexes. As shown in Figure 5, 2200 matches 2113 for about 5·5 kb, but then has about 5 kb of DNA unrelated to the 87E map, presumably because it comes from the bithorax complex in 89E. *In situ* hybridization with 2200 (Fig. 6) showed grains at 87E and 89E, as expected. There were considerably fewer grains at the 89E constriction than over the 87E1–2 band in most nuclei; we suspect that the constriction reflects underreplication of the bithorax region in salivary polytene chromosomes (A. Spierer & P. Spierer, unpublished results).

The 5 kb of new DNA from 2200 was used as a probe to isolate DNA from the wild-type bithorax complex. In the first screening of the Canton S library, we recovered the phage 2201, which matched the map of 2200 as indicated in Figure 5. *In situ* hybridization with 2201 gave grains only at 89E, as expected. The 2201 phage served as the starting point for a walk within the bithorax complex. Many rearrangement breakpoints, insertions and deletions with various bithorax mutant phenotypes have been subsequently mapped in the region around 2201, confirming that the jump landed within the bithorax complex (Bender *et al.*, 1983).

4. Discussion

When this walk was started, it was not clear whether the strategy would be frustrated by sequences that could not be cloned in *E. coli* or by large blocks of repetitive DNA. At every step, as it happened, the new overlapping phages were recovered from the Canton S library with about the expected frequency of one in 10^4. Some long sequences of repeated DNA were encountered, but in every case, the repetitive block was mobile in the evolution of *D. melanogaster* strains; hence, repeats could be avoided by extending the walk in a different strain. Similar walks have recently isolated comparable DNA stretches from several other regions of the *D. melanogaster* chromosomes (89E, Bender *et al.*, 1983; 74E, W. Jones, K. Burtis & D. Hogness, personal communication; 84AB, M. Scott, A. Weiner, B. Polisky & T. Kaufman, personal communication).

We know that the walk includes the loci for *rosy* and *Ace* because it extends beyond deficiency endpoints that bracket these loci. The accompanying paper (Spierer *et al.*, 1983) locates more accurately the positions of *rosy* and *Ace*, as well as other lethal complementation groups covered by the walk.

The jumping strategy worked easily for the Cbx^{+R1} inversion, but it can be frustrated by repetitive DNA at the rearrangement breakpoints. For example, the distal end of the deficiency Df(3R)126d was reached early in the walk, and we attempted to jump to the proximal endpoint to get more quickly to the *rosy* locus (see Fig. 2). The deficiency endpoint was defined first by *in situ* hybridizations and then by whole genome southerns, and the expected fusion fragment was cloned and mapped. The cloned fragment was homologous to the walk for about 4 kb (+85 to +89 kb), and had another 7 kb of new DNA. Unfortunately, all the new DNA appeared to be repetitive by whole genome southerns, and so it was useless as a probe to complete the jump. Since every rearrangement is a unique chromosome, the repeats cannot be avoided by going to another strain. It is not clear how often X-ray breaks are associated with repetitive DNA, but this does not appear to be a frequent event as several other X-ray rearrangements that are without repeats at the breakpoints have been recloned more recently (Bender *et al.*, 1983; Atavanis-Tsakonas *et al.*, 1983; Scott *et al.*, unpublished data; Wolfner & Baker, unpublished data). Spontaneous deficiencies associated with unstable mutations such as white-crimson (Green, 1967) or white-Dzl (Bingham, 1980), or rearrangements produced in hybrid dysgenesis crosses (Bingham *et al.*, 1982) seem more likely to be associated with repetitive elements at the endpoints.

A long continuous stretch of DNA covering at least 12 unrelated genes (Spierer *et al.*, 1983) provides a sampling of the sequence organization that should be representative. Of the 315 kb of DNA in the walk, all but the rightmost 8 kb has been tested for sequence repetition by whole genome southerns, typically one phage at a time. With the exception of the evolutionarily mobile elements mentioned, all the sequences appeared to be single-copy, in that the only major bands on the southern were those expected from the restriction maps of the phage clones. The southerns occasionally had faint minor bands that could reflect partial digestion of the genomic DNA, heterogeneity in the restriction map within a stock, or weak homology of the probe to sequences at other sites. Such a whole

genome southern analysis would have missed very small repeated sequences (less than a few hundred bases), and recent results suggest that such small repeats are indeed present in several places (L. Hall & P. Spierer, unpublished results). We saw no evidence of clustered repeats (Wensink *et al.*, 1979) and little indication of preferred regions for insertion of mobile repeats (the region around −130 was the only "warm spot").

A comparison of the restriction maps between Canton S and Oregon R can give a rough measure of single-copy sequence divergence. Of 161 sites mapped in single-copy Canton S DNA that was also cloned from Oregon R, only one site was missing in Oregon R. Similarly, six out of 166 Oregon R sites were missing in Canton S. All the restriction enzymes used recognized six base-cleavage sites, and so the mismatch is 1/966 and 6/996 for the two comparisons, or an average of 0·4%.

We noted a surprising amount of heterogeneity within the Oregon R library. In six cases in which we found an insertion in Oregon R DNA relative to Canton S, the insertion was absent from one or more of the Oregon R clones covering the insertion site. The two Canton S insertions were invariant in all clones, with the exception of the 2121 phage, which probably deleted the Kermit element during growth in *E. coli*. The embryonic DNA for the Oregon R library was derived from the Stanford Oregon population, maintained with over 10^5 flies for several years. The Canton S library was derived from the Caltech strain of Canton S, maintained in milk bottles with the population restricted to about 100 flies at each transfer. There is a greater chance of contamination associated with maintaining large populations, but we suspect that the larger population size itself contributes to the greater diversity found in the Oregon R chromosomes.

REFERENCES

Artavanis-Tsakonas, S., Muskavitch, M. A. & Yedvobnick, B. (1983). *Proc. Nat. Acad. Sci.*, *U.S.A.* **80**, 1977–1981.

Bender, W., Akam, M., Karch, F., Beachy, P. A., Peifer, M., Spierer, P., Lewis, E. B. & Hogness, A. S. (1983). *Science*, in the press.

Benton, W. D. & Davis, R. W. (1977). *Science*, **196**, 180–182.

Bingham, P. M. (1980). *Genetics*, **95**, 341–353.

Bingham, P. M., Kidwell, M. G. & Rubin, G. M. (1982). *Cell*, **29**, 995–1004.

Blattner, F. R., Williams, B. G., Blechl, A. E., Denniston-Thompson, K., Faber, H. E., Furlong, L., Grunwald, D. J., Kiefer, D. O., Moore, D. M., Schumm, J. W., Sheldon, E. L. & Smithies, O. (1977). *Science*, **196**, 161–169.

Bridges, C. B. (1941). *J. Hered.* **32**, 299–300.

Carlson, M. (1979). Ph.D. thesis, Stanford University.

Cherbas, P., Cherbas, L. & Williams, C. M. (1977). *Science*, **197**, 275–277.

Chovnick, A., Gelbart, W. & McCarron, M. (1977). *Cell*, **11**, 1–10.

Davis, R. W., Simon, M. & Davidson, N. (1971). In *Methods in Enzymology* (Grossman, L. & Moldave, K., eds), vol. 21, pp. 413–428, Academic Press, New York.

Davis, R. W., Botstein, D. & Roth, J. R. (1980). *Advanced Bacterial Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Denhardt, D. T. (1966). *Biochem. Biophys. Res. Commun.* **23**, 641–646.

Green, M. M. (1967). *Genetics*, **56**, 467–482.

Hall, J. C. & Kankel, D. R., (1976). *Genetics*, **83**, 517–535.

Hilliker, A. J., Clark, S. H., Chovnick, A. & Gelbart, W. M. (1980). *Genetics*, **55**, 95–110.

Ingolia, T. D. & Craig, E. A. (1982). *Proc. Nat. Acad. Sci., U.S.A.* **79**, 525–529.

Lewis, E. B. (1964). In *Role of Chromosomes in Development* (Locke, M., ed.), pp. 231–252, Academic Press, New York.

Lewis, E. B. (1978). *Nature (London)*, **276**, 565–570.

Lis, J. T., Prestidge, L. & Hogness, D. S. (1978). *Cell*, **14**, 901–919.

Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K. & Efstratiadis, A. (1978). *Cell*, **15**, 687–701.

Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular Cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Meyerowitz, E. & Hogness, D. S. (1982). *Cell*, **28**, 165–176.

Rigby, P. W. J., Dieckman, M., Rodes, C. & Berg, P. (1977). *J. Mol. Biol.* **113**, 237–251.

Southern, E. M. (1975). *J. Mol. Biol.* **98**, 503–517.

Spierer, P., Spierer, A., Bender, W. & Hogness, D. S. (1983). *J. Mol. Biol.* **168**, 35–50.

Wensink, P. C., Tabata, S. & Pachl, C. (1979). *Cell*, **18**, 1231–1246.

Young, M. W. & Schwartz, H. E. (1981). *Cold Spring Harbor Symp. Quant. Biol.* **45**, 629–640.

Edited by P. Chambon

Note added in proof: We note that the Jiminy element found in our walk is identical with the ribosomal insertion-like element 101 F (Dawid, I. B., Long, E. O., Dinocera, P. P. & Pardue, M. L. (1981). *Cell*, **25**, 399–408; Dinocera, P. P., Digan, M. E. & Dawid, I. B. (1983). *J. Mol. Biol.* In the press) as seen by comparison of restriction maps and electron microscopy of heteroduplexes.

2