Cell, Vol. 17, 415–427, June 1979, Copyright © 1979 by MIT

Transposition of Elements of the *412*, *copia* and *297* Dispersed Repeated Gene Families in Drosophila

S. Steven Potter,* William J. Brorein, Jr., Pamela Dunsmuir and Gerald M. Rubin
Department of Tumor Biology
Sidney Farber Cancer Institute
and Department of Biological Chemistry
Harvard Medical School
Boston, Massachusetts 02115

Summary

The stability of elements of three different dispersed repeated gene families in the genome of Drosophila tissue culture cells has been examined. Different amounts of sequences homologous to elements of *412*, *copia* and *297* dispersed repeated gene families are found in the genomes of D. melanogaster embryonic and tissue culture cells. In general the amount of these sequences is increased in the cell lines. The additional sequences homologous to *412*, *copia* and *297* occur as intact elements and are dispersed to new sites in the cell culture genome. It appears that these elements can insert at many alternative sites. We also describe a DNA sequence arrangement found in the D. melanogaster embryo genome which appears to result from a transposition of an element of the *copia* dispersed repeated gene family into a new chromosomal site. The mechanism of insertion of this *copia* element is precise to within 90 bp and may involve a region of weak sequence homology between the site of insertion and the direct terminal repeats of the *copia* element.

Introduction

A novel class of repeated genes of Drosophila melanogaster has been described in which the elements are widely dispersed throughout the genome, rather than tandemly arranged (Rubin, Finnegan and Hogness, 1976). Two such dispersed repeated gene families, *412* and *copia*, have been studied in detail and share several characteristic features although they show no sequence homology (Finnegan et al., 1978). *412* and *copia* elements are each found at approximately thirty widely scattered locations in the chromosomes of D. melanogaster and both code for abundant poly(A)-containing cytoplasmic RNAs. While the sequences of the *412* and *copia* elements are very closely conserved at each chromosomal site, the sequences adjacent to the elements differ at each site. *412* and *copia* elements are terminally redundant: *412* elements carry direct repeats of 0.5 kb and *copia* elements have direct repeats of 0.3 kb. Such terminal direct repeats are a feature of certain bacterial transposons (Kleckner, 1977) and the integrated genomes of retroviruses (Hughes et al., 1978; Sabran et al., 1979). Moreover, dispersed repeated gene families are not limited to the genus Drosophila; repeated elements similar in structure of *412* and *copia* have recently been found in the genome of the yeast Saccharomyces cerevisiae (Cameron, Loh and Davis, 1979).

412 and *copia* are not the only examples of dispersed repeated genes in D. melanogaster. We have demonstrated the existence of a third family of dispersed repeated genes, *297*, which shares the features outlined above for *412* and *copia* (G. M. Rubin and B. Backner, unpublished results). Furthermore, results obtained in other laboratories suggest the existence of several additional families (Ilyin et al., 1978; Tchurikov et al. 1978; Carlson and Brutlag, 1979; W. Bender, P. Spierer, M. Goldberg, R. Lifton, R. Karp and D. S. Hogness, unpublished results), and together, these dispersed repetitive gene families account for nearly 1% of the D. melanogaster genome. Data presented here and by Strobel, Dunsmuir and Rubin (1979) suggest that elements of these dispersed repeated gene families may share another common feature; elements of the *412*, *copia* and *297* gene families are mobile within the genome and undergo significant rearrangement in both fly populations and tissue culture cell lines.

To study the molecular mechanism by which such transpositions take place, we sought an experimental system in which transpositions of elements from the repeated gene families occur. We reasoned that tissue culture cells would provide a permissive environment for such transpositions since these cells should tolerate more mutation than embryo cells, which must develop to form a fertile adult. Moreover, all nonlethal transposition events would remain in the population, not just those that occurred in the small group of cells set apart to form the germ line. We therefore measured the stability of elements of the *412*, *copia* and *297* dispersed repeated gene families in the genomes of Drosophila tissue culture cells and found that these elements may change in both number and location. This paper also describes an embryo DNA sequence arrangement which appears to have resulted from the insertion of an element of the *copia* dispersed repeated gene family into a new chromosomal site.

Results and Discussion

There Are Different Amounts of Sequences Homologous to the *412*, *copia* and *297* Elements in the Genomes of Embryonic and Tissue Culture Cells

We have compared the amount of sequences homologous to the *412*, *copia* and *297* elements in DNA isolated from three sources: D. melanogaster embryos (Oregon R), Schneider's cell line 2 and Echalier's Kc_0 cell line. We measured the reassociation kinetics of trace amounts of highly labeled DNA from each of

* Present address: Department of Biology, Wesleyan University, Middletown, Connecticut 06475.

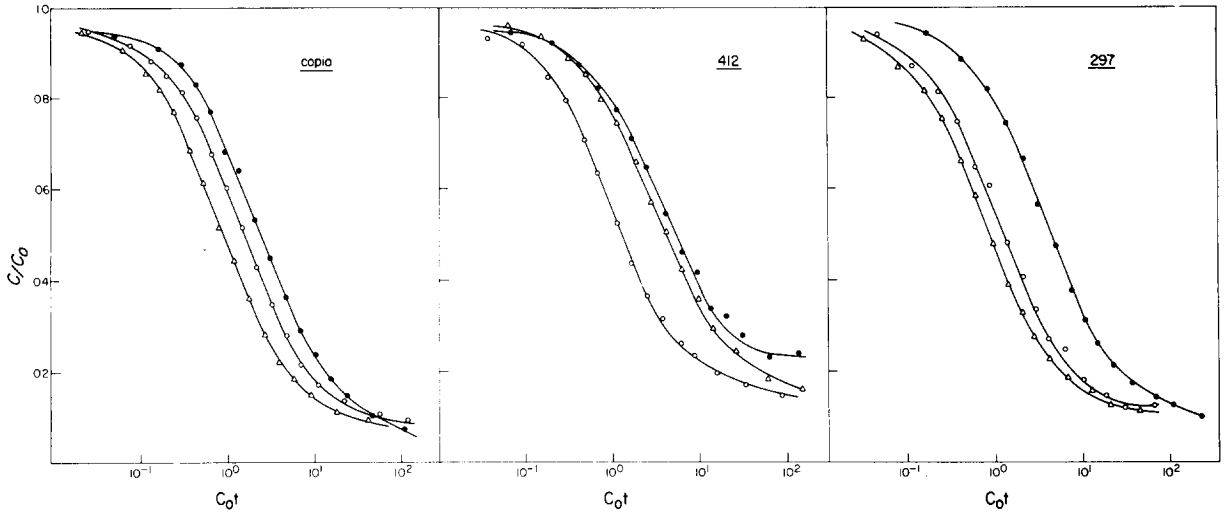

Figure 1. Reassociation Kinetics of *412*, *copia* and *297* Sequences in the Presence of D. melanogaster DNA Isolated from Embryos, the Kc_0 Cell Line or Schneider's Cell Line 2

A restriction fragment from elements of each of the dispersed repeated gene families was labeled to a high specific activity by nick translation (Rigby et al., 1977) and allowed to reassociate in the presence of a vast excess of nonradioactive DNA isolated from either embryos (●), Kc_0 cells (○) or Schneider's line 2 cells (△).

Table 1. Copies per Haploid Genome [Relative Copy Number]

	Oregon R Embryo	Kc_0 Cell Line	Schneider's Cell Line 2
copia	60 [1.0]	90 [1.51 ± 0.09; (n = 3)]	170 [2.69 ± 0.21; (n = 2)]
412	40 [1.0]	120 [3.52 ± 0.82; (n = 3)]	40 [1.16 ± 0.33; (n = 3)]
297	30 [1.0]	110 [3.4; (n = 1)]	170 [5.1; (n = 1)]

The numbers given for copies per haploid genome represent our best estimates and were calculated from the ratio of the $Cot_{1/2}$ of E. coli DNA and the $Cot_{1/2}$ of the *412*, *copia* or *297* sequences which were allowed to reassociate in the same tube. Genome sizes were taken as 4×10^3 kb for E. coli and 1.65×10^5 kb for D. melanogaster, of which 72% was assumed to be nonrepetitive (Schachat and Hogness, 1974). The relative copy numbers represent the ratios of $Cot_{1/2}$ determined for each driver DNA in the same experiment and are therefore independent of assumptions involving genome sizes, salt or fragment length corrections. In cases where multiple determinations were performed, the mean, standard deviation and number of determinations are indicated. [Schneider's cell line 2 was established from Oregon R embryos (Schneider, 1972), whereas the Kc_0 line was established from embryos of the F_2 generation from a cross between *sepia* and *ebony* flies derived from a different wild-type (Echalier and Ohanessian, 1969, 1970; Barigozzi, 1971). Comparisons between Schneider's line 2 cells and embryo cells may therefore be more significant than comparisons to Kc_0 cells since, in the case of Kc_0 cells, we may be observing strain differences, rather than differences which have arisen during passage of the cells in culture.]

these dispersed repeated gene families in the presence of a vast excess of DNA from each of the three DNA sources. Examples of the results of these experiments are summarized in Table 1. The amounts of DNA homologous to the elements of each gene family differ among the three DNA sources; generally, however, they are more abundant in tissue culture DNA. Sequences homologous to the *412*, *copia* and *297* elements appear to be independently modulated in the different genomes. For example, the *412* sequences show no detectable differences in amount between embryo and Schneider's line 2 cells, whereas the *copia* and *297* sequences show a 2.7 and 5.1 fold increase, respectively. In the case of *297*, this represents an increase from approximately 0.1% of the cell's DNA in embryos to 0.5% in Schneider's cell line 2.

It should be noted that our estimates of copy number for the *412*, *copia* and *297* elements in embryo DNA approximate the number of chromosomal sites at which these elements occur as determined by in situ hybridization to polytene chromosomes: 40 copies versus 26 sites for *412*, 60 copies versus 30 sites on the chromosome arms plus homologous sequences at the chromocenter for *copia*, and 30 copies versus 20 sites on the chromosome arms plus homologous sequences at the chromocenter for *297* (Strobel et al., 1979). [Sequences found in heterochromatin are predominantly located at the chromocenter of salivary gland polytene chromosomes (Peacock et al., 1978).] This correlation suggests that at each chromosomal site in the embryo where *412*, *copia* and *297* sequences occur, only a single copy of the element is present.

The Sequences Homologous to *412*, *copia* and *297* Occur as Intact Elements

Two lines of evidence lead to the conclusion that most, if not all, of the cell culture sequences which are homologous to the *412*, *copia* and *297* elements occur as intact elements, indistinguishable from those found in the embryo. The first set of experiments examines the pattern of restriction enzyme cleavage sites within those sequences homologous to the elements of each family in DNA from embryos, Schneider's cell line 2 and Kc_0 cells. In the embryo genome, nearly all the elements of a given dispersed repeated gene family are alike with respect to their restriction enzyme cleavage patterns (Rubin et al., 1976; Finnegan et al., 1978; G. M. Rubin and B. Backner, unpublished results). If such restriction patterns were observed for the homologous sequences in the genomes of the cell culture lines, they would provide evidence that the additional sequences occur as complete elements similar to those found in the embryo. An experiment demonstrating such conservation of restriction enzyme cleavage sites within sequences homologous to the *412* element is shown in Figure 2. Three characteristic restriction enzyme fragments were generated from within the *412* element on the cloned embryo DNA segment Dm412 (Rubin et al., 1976). The five restriction enzyme cleavage sites needed to generate these fragments are within the element, and each one is conserved in the *412* elements found in the embryo genome (Rubin et al., 1976; Finnegan et al., 1978). Consider fragment C, generated by cleavage with the restriction endonucleases Hind III and Eco RI: Hind III plus Eco RI digests of DNA isolated from embryos, Schneider's cell line 2 and the Kc_0 cell line were fractionated in parallel by electrophoresis in an agarose gel, and the resultant distributions of fragments were transferred to a nitrocellulose filter (Southern, 1975). Hybridization with ^{32}P-labeled fragment C DNA reveals that homologous genomic sequences are confined to fragments identical in size to fragment C in each DNA. Similar results were obtained with fragments B and E of Dm412 (Figure 2) and with fragments from within the *copia* and *297* elements (Figure 3). In each case the characteristic restriction enzyme cleavage sites within the sequences homologous to *412*, *copia* and *297* elements are closely conserved in the genomes of embryos, Schneider's cell line 2 and Kc_0 cells. Moreover, the relative amounts of the internal, restriction enzyme-generated fragments are consistent with the amount of sequences homologous to the element in each genome determined by DNA reassociation kinetics. These are the results we would expect if all, or nearly all, of the sequences homologous to the *412*, *copia* and *297* families occurred as intact *412*, *copia* and *297* elements.

Independent evidence in support of the integral nature of these elements comes from examination of fragments of Schneider's cell DNA which have homology to *412*, *copia* and *297*. Restriction enzyme maps of four Bam HI fragments selected by virtue of their homology to *copia* are shown in Figure 4. Each one contains the pattern of restriction sites characteristic of an intact *copia* element. Since there are about 3 times as many *copia* sequences in Schneider's cell DNA as in embryo DNA, we expect that some of these examples represent *copia* elements from sites where *copia* is not found in the embryo genome. This assumption has been confirmed by the experiments described below. Similar analysis of other Bam HI fragments of Schneider's cell DNA has been extended to include a total of seven fragments with homology of *copia*, four fragments with homology to *412* and nine fragments with homology to *297* (P. Dunsmuir and E. Young, unpublished observation). All fragments examined contain restriction enzyme cleavage patterns characteristic of intact elements.

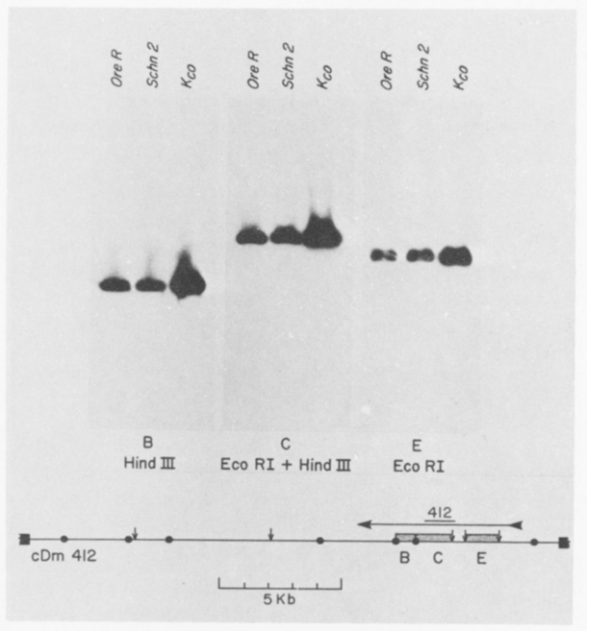

Figure 2. Comparison of Restriction Fragments in cDm412 with Restriction Fragments from Total DNA of D. melanogaster Embryos (Oregon R), Kc_0 Cells and Schneider's Line 2 Cells Which Contain *412* Sequences

For each set of results, either Hind III, Eco RI or Hind III plus Eco RI digests of each total genome DNA were separated on 1.4% agarose gels. Each panel shows the autoradiograph obtained when these restriction fragments are subsequently hybridized according to the method of Southern (1975) with ^{32}P-labeled restriction fragments derived from cDm412 as indicated. The degree of hybridization is as expected from the amount of *412* sequences estimated by reassociation kinetics. A physical map of cDm412 (Rubin et al., 1976) is shown below the autoradiograms. The thin horizontal line represents Drosophila DNA. The thick blocks at the end of the map represent Col E1 DNA. Restriction enzyme cleavage sites for Eco RI (↓) and Hind III (●) are indicated. The shaded regions delineate the restriction fragments used as hybridization probes. The position of the *412* element is indicated above the map.

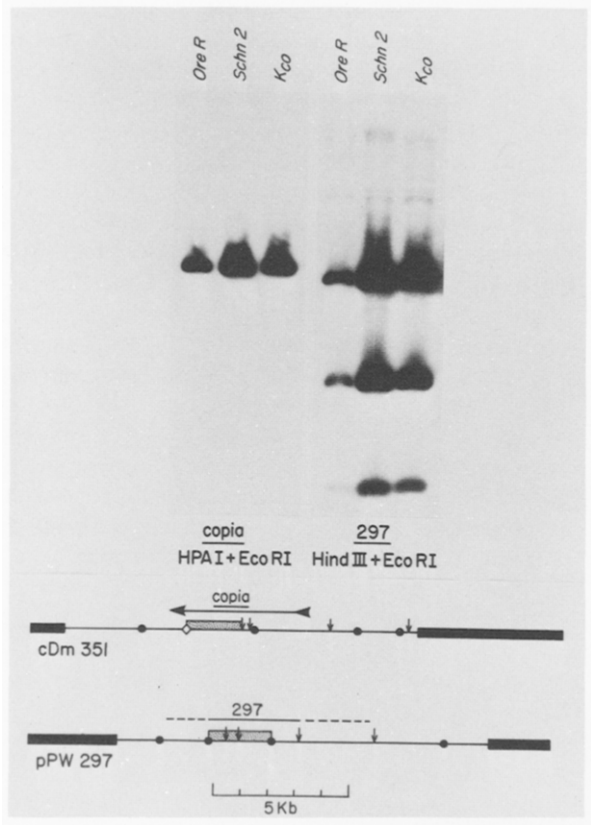

Figure 3. Comparison of Restriction Fragments in cDm351 and pPW297 with Restriction Fragments from Total DNA of D. melanogaster Embryos (Oregon R), Kc_0 cells and Schneider's Line 2 Cells Which Contain *copia* or *297* Sequences

For *copia* sequences, Hpa I plus Eco RI digests of each total genome DNA were separated on a 1% agarose gel and subsequently hybridized with the indicated fragment of cDm351 according to the method of Southern (1975). For *297* sequences, Hind III plus Eco RI digests were separated on a 1.4% agarose gel and hybridized with the indicated Hind III fragment of pPW297. Autoradiograms are shown. The degree of hybridization is as expected from the amount of *copia* and *297* sequences estimated by reassociation kinetics. In the maps of cDm351 (D. J. Finnegan, G. M. Rubin and D. S. Hogness, unpublished results) and pPW297 (G. M. Rubin and B. Backner, unpublished results), the thin horizontal lines represent Drosophila DNA and the heavy lines represent DNA of the plasmid vector (Col E1 for cDm351 and pMB9 for pPW297). Restriction enzyme cleavage sites for Eco RI (↓), Hind III (●) and Hpa I (◇) are indicated. The shaded regions delineate the restriction fragments used as hybridization probes. The positions of the *copia* element in cDm351 and the *297* element in pPW297 are indicated above the maps. The precise end points of the *297* element have not been determined, but fall within the limits shown by the dashed portion of the line.

The Additional *412*, *copia* and *297* Elements Are Dispersed to New Sites in the Tissue Culture Cell DNA

How are the additional *412*, *copia* and *297* elements arranged in the genome? To answer this question, we have examined the DNA sequences surrounding the elements in the cell culture genome. DNA from embryos, Schneider's cell line 2 and Kc_0 cells was digested with the restriction endonuclease Eco RI, and the sizes of the resultant fragments homologous to the

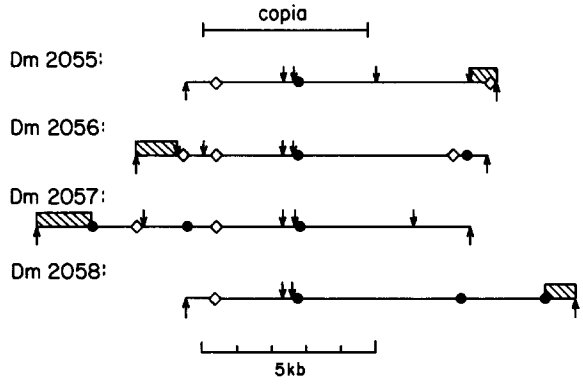

Figure 4. Restriction Enzyme Maps of Four Bam HI Fragments of Schneider's Cell Line 2 DNA Which Contain *copia* Elements

Schneider's cell line 2 DNA was digested with the restriction endonuclease Bam HI, and the fragments were ligated into the Bam HI site of the plasmid vector pBR322. Bacteria containing such hybrid plasmids were screened using a modified colony hybridization procedure (P. Gergen and P. C. Wensink, manuscript submitted) for those containing sequences homologous to *copia*. Restriction enzyme maps of four selected fragments are shown. The horizontal lines represent Drosophila DNA, and the cleavage sites for Bam HI (↑), Eco RI (↓), Hind III (●) and Hpa I (◇) are indicated. The shaded regions delineate portions of each Bam HI fragment that were subsequently used as hybridization probes (see text). The approximate end points of the *copia* elements were determined by comparison of restriction sites with those of cDm351 (see Figure 3) and are indicated by the line above the maps.

Hind III/Eco RI fragment C of cDm412 were determined (Figure 5). A single such Eco RI fragment will be generated from each site in the genome where a *412* element occurs. The size of this fragment is equal to the distance from the Eco RI site within *412*, which forms one end of fragment C, to the nearest Eco RI site occurring in sequences flanking the element. This latter Eco RI site will be characteristic of the location in the genome where that particular element is found. Comparison of the patterns generated in this way from embryo and Schneider's cell line 2 DNA reveals numerous differences suggesting differences in the distributions of *412* elements in these genomes. Similar results were obtained when analogous experiments were carried out with the *copia* (data not shown) and *297* (Figure 6) elements. In contrast, when the genomic distribution patterns of other D. melanogaster middle repetitive sequences which do not exhibit the characteristic sequence structure of *412*, *copia* and *297* elements are examined, they are found to be similar in embryo, Schneider's cell line 2 and Kc_0 cell DNAs (G. M. Rubin, unpublished results; P. Bingham, unpublished results). As discussed below, we believe that these differences between the embryo and cell culture DNAs in the pattern of restriction enzyme sites in those sequences flanking the *412*, *copia* and *297* elements are evidence of the transposition of these elements.

Four distinct models can be envisaged for the origin and distribution of the additional *412*, *copia* and *297*

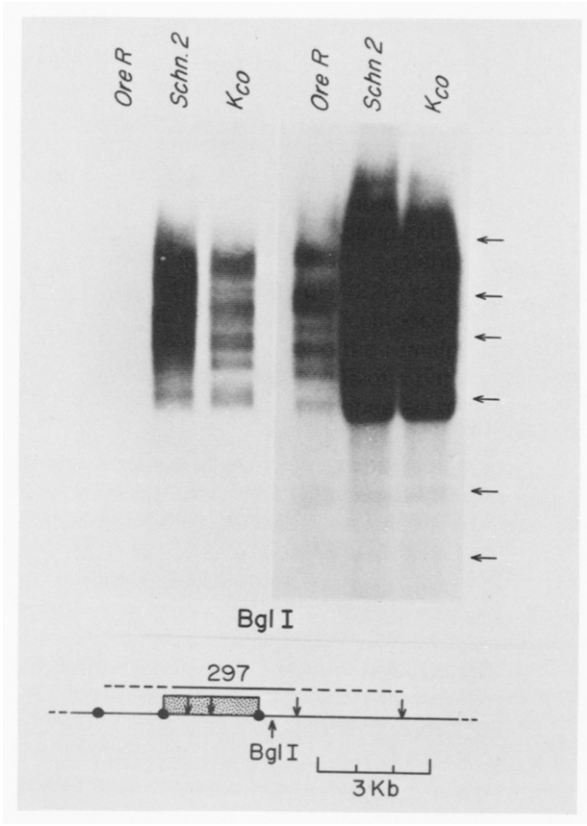

Figure 5. Distribution of Eco RI Restriction Sites in Sequences Adjacent to *412* Elements in the Genomes of Drosophila Embryos (Oregon R), Kc_0 Cells and Schneider's Line 2 Cells

Eco RI digests of each total genome DNA were separated on a 0.5% agarose gel. The autoradiograph obtained when these fragments were hybridized according to the procedure of Southern (1975) with the indicated ^{32}P-labeled Eco RI plus Hind III restriction enzyme fragment of cDm412 is shown. A portion of the physical map of cDm412 is also shown. The horizontal line represents Drosophila DNA, and restriction enzyme cleavage sites for Eco RI (↓) and Hind III (●) are indicated. The shaded region delineates the restriction fragment used as a hybridization probe. The position of the *412* element is indicated above the map. The arrows indicate the positions of Hind III fragments of bacteriophage λ DNA which are 23, 9.8, 6.6, 4.5 and 2.5 kb long (Murray and Murray, 1975). (The observed patterns of hybridization appear to be reproducible in that three independent DNA preparations made at 2 week intervals from the Kc_0 cell line gave identical results.)

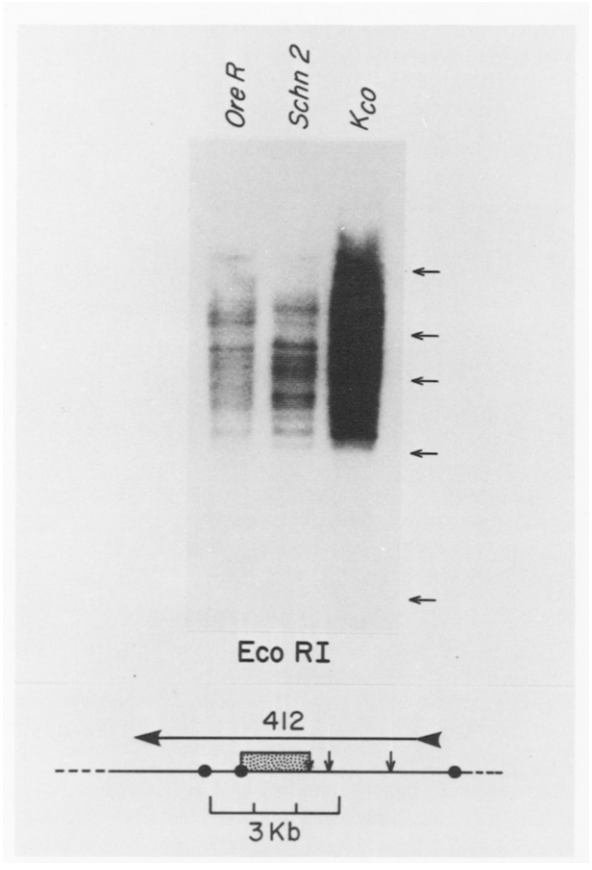

Figure 6. Distribution of Bgl I Restriction Sites in Sequences Adjacent to *297* Elements in the Genomes of Drosophila Embryos (Oregon R), Kc_0 Cells and Schneider's Line 2 Cells

Bgl I digests of each total genome were separated on a 0.5% agarose gel. Two autoradiographs obtained when these restriction fragments were hybridized according to the method of Southern (1975) with the indicated ^{32}P-labeled Hind III fragment of pPW297 are shown. These differ only in length of exposure and are presented to facilitate comparisons not possible with any single exposure. A portion of the physical map of pPW297 is presented. The horizontal line represents Drosophila DNA, and restriction enzyme cleavage sites for Eco RI (↓), Hind III (●) and Bgl I (↑) are indicated. These shaded region delineates the restriction enzyme fragment used as a hybridization probe. The position of the *297* element is shown above the map. The arrows indicate the positions of Hind III fragments of bacteriophage λ DNA which are 23, 9.8, 6.6, 4.5, 2.5 and 2.2 kb long (Murray and Murray, 1975).

elements which accumulate in the genomes of cells in culture. These models are diagrammed in Figure 7. Consider the >5 fold increase in the number of *297* elements observed in the Schneider's cell line 2 genome. In model 1, each *297* element would duplicate locally to produce, on the average, five tandem copies of *297*. In model 2, the environment of most elements would remain unchanged, but one (or a few) elements would duplicate to produce a long tandem array(s) of *297* elements. Model 3 also involves tandem duplication, but unlike the above models, a chromosomal region larger than the element, rather than the element itself, duplicates. This is the proposed mechanism by which the gene for dihydrofolate reductase duplicates to produce methotrexate resistance in cells exposed to that drug (Schimke et al., 1978). In model 4, *297* elements duplicate and transpose to new chromosomal sites. Only model 4 postulates that elements will occur at previously uninhabited chromosomal sites.

Two lines of evidence allow us to distinguish among these four models. The first compares the sizes of element-containing fragments obtained when DNA from different sources is digested with restriction endonucleases which do not cleave within the element. For example, the experimentally determined size dis-

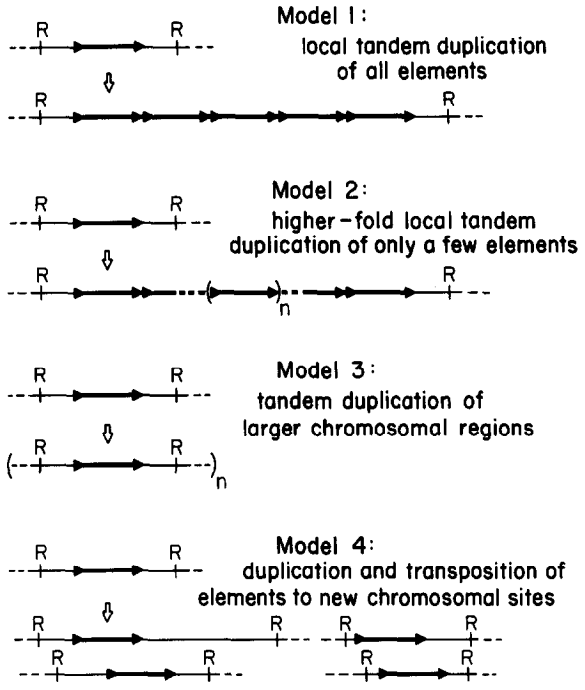

Figure 7. Models for the Origin and Distribution of the Additional *412*, *copia* and *297* Elements Which Accumulate in the Genomes of Cells in Culture

The thin horizontal lines represent Drosophila DNA, except for DNA of the elements of a dispersed repeated gene family which are shown as heavy horizontal lines. Cleavage sites for the restriction endonucleases Bam HI and Kpn I, which do not cleave within the elements, are indicated (R).

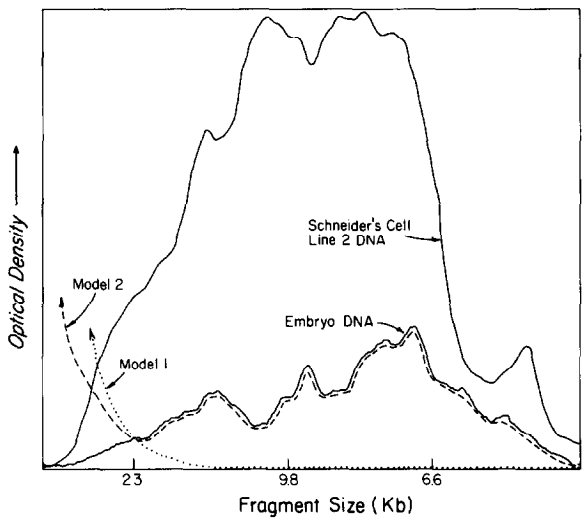

Figure 8. Size Distributions of *297*-Containing Restriction Fragments in Different Genomes

Bam HI plus Kpn I digests of embryo and Schneider's cell line 2 DNA were separated on a 0.5% agarose gel. Densitometer tracings of the autoradiographs obtained when these fragments were hybridized according to the procedure of Southern (1975) with the ^{32}P-labeled 2.3 kb Eco RI fragment of pPW297 (see Figure 3) are shown. The distributions predicted by models 1 and 2 are indicated. The large (>20 kb) fragments of DNA predicted by models 1 and 2 would be underrepresented due to inefficiency of transfer (Southern, 1975); however, the observed distribution of fragments differs from that predicted by these models for fragments of all sizes.

tributions of *297*-containing fragments in digests of Schneider's cell line 2 and embryo DNA are shown in Figure 8. The fragment size distributions are similar in both DNAs although the number of fragments is much greater in the cell line, consistent with there being 5 fold more *297* elements in Schneider's cell line 2 DNA than in embryo DNA. Similar results were obtained when the fragment size distributions of *copia* sequences were assayed in these two DNAs (data not shown). These observed differences are consistent with either model 3 or 4, but rule out models 1 and 2. Further evidence against model 1 is provided by restriction site maps of cloned segments of Schneider's cell line 2 DNA; each segment shown in Figure 4 contains only a single copy of the *copia* element. Similar results have been obtained with DNA segments containing *412* and *297* elements (P. Dunsmuir and E. Young, unpublished observations).

Model 4 predicts that some or all of the four segments of Schneider's cell line 2 DNA shown in Figure 4 are the result of transposition of *copia* elements to new sites in the genome during passage of the Schneider's cells. If this prediction is correct, then the corresponding segments of the embryo genome should lack the *copia* element. The following experiment was performed to test this prediction. Each of the segments of Schneider's cell DNA shown in Figure 4 is a Bam HI fragment whose size can be measured directly. If, as predicted by model 3, the corresponding embryo Bam HI fragment contains a *copia* element, it should be identical in size to that isolated from the Schneider's cell genome. Alternatively, if the *copia* element had inserted into this fragment during passage of the cell cultures, then the embryo fragment would not contain the *copia* element and would be smaller than the cloned Schneider's fragment by the size of a *copia* element. We measured the size of the Bam HI fragments in the embryo genome which correspond to each of the four segments shown in Figure 4. To do this, we hybridized only a unique sequence from each cloned fragment, indicated by the shaded area, to a Bam HI digest of embryo DNA which had been fractionated by electrophoresis on an agarose gel. In all four cases, the embryo Bam HI fragment was smaller in size than the corresponding cloned segment of Schneider's cell line 2 DNA. The average of these size differences was 4.5 ± 0.8 kb. This number is not very precise since it is the result of subtracting two large numbers, but it approximates the 5 kb increase expected for an insertion of a *copia* element. The simplest interpretation of these results is that *copia* elements have inserted into new sites in the Schneider's cell line 2 genome as predicted by model 4. This interpretation has been confirmed for one of the Schneider's cell line 2 Bam HI fragments,

Dm2056 (diagrammed in Figure 4). The corresponding embryo Bam HI fragment has been isolated and compared to Dm2056. Restriction site maps suggest that the two fragments differ only by the insertion of a *copia* element, and electron microscope analysis of heteroduplex molecules formed between the fragments shows that they are similar along their entire lengths except for the presence of a single insertion of approximately 5 kb (Figure 9). These data are consistent only with the model that predicts that elements are duplicated and transposed to new chromosomal locations during passage of Drosophila cell cultures. Elements also appear to be capable of transposition without concomitant duplication. The genome of Schneider's cell line 2 shows no increase in the number of *412* elements relative to the embryo, yet *412* elements are in different locations in these two genomes as judged by the pattern of restriction sites in adjacent DNA (Figure 5). Additional evidence suggesting that transposition can occur without a net increase in the number of elements present is presented by Strobel et al. (1979).

Elements Appear to Be Able to Insert at Many Alternative Sites in the Genome

Elements of the dispersed repeated gene families might be capable of integrating at a very large number of sites in the genome, or there may be a limited number of highly preferred sites where integration occurs. In Schneider's cell line 2 DNA there must be at least 170 genomic locations of *copia* since we have no evidence of tandem repetitions of elements. If there are additional possible integration sites for *copia* elements, these would be reflected by heterogeneity in the genomic distributions of elements in the cell population. To test whether the Schneider's cell line 2 fragments shown in Figure 4 are identical to all corresponding Bam HI fragments from the population of tissue culture cells, we hybridized a unique sequence from each cloned fragment to a Bam HI digest of DNA from the population of Schneider's cells. In the case of segment Dm2057, we found that the vast majority of cells in the population contain a Bam HI frgment of the same size as Dm2057; for the other three fragments, however, the majority of cells have a Bam HI fragment identical in size to the corresponding embryo fragment, that is, smaller than the cloned fragment by the length of a *copia* element. Similar analyses of other fragments from Schneider's cell DNA have been extended to include a total of seven fragments with homology to *copia*; four of these were similar to Dm2057, but the other three showed differences of an element length between the cloned fragment and the corresponding fragment in the population. These results demonstrate that there is considerable heterogeneity in the genomic locations of *copia* elements in Schneider's line 2 tissue culture cells, and suggest that there are many more than 170 sites where integration can occur.

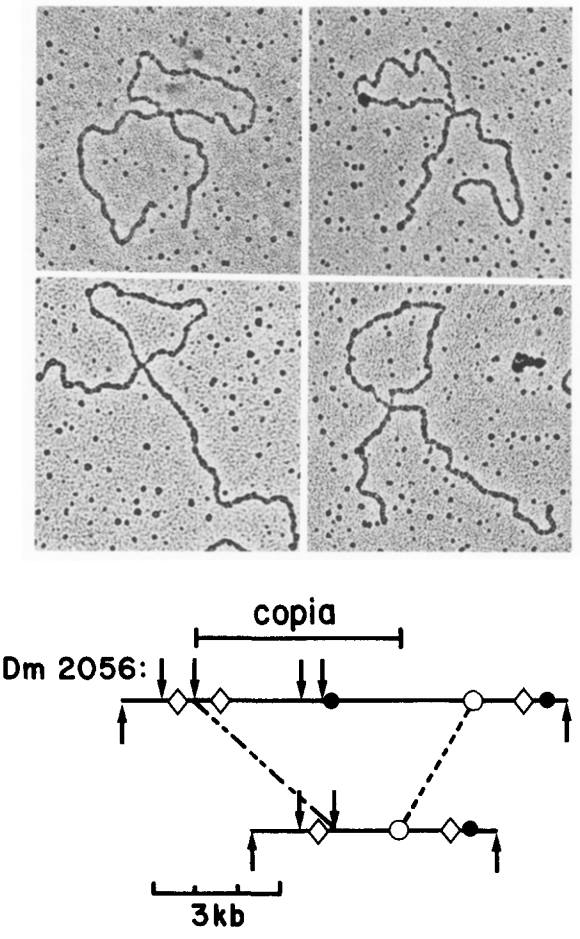

Figure 9. Comparison between a Schneider's Cell Line 2 Bam HI Fragment and the Corresponding Fragment from Oregon R Embryo DNA

Heteroduplex molecules were prepared after Bam HI cleavage of the two pBR322 hybrid plasmids containing the indicated Schneider's cell line 2 and embryo DNA Bam HI fragments. [The embryo DNA fragment was selected by using the shaded region of the Schneider's cell line 2 fragment Dm2056 (Figure 4) to screen bacteria containing hybrid plasmids between Bam HI-digested embryo DNA and pBR322.] The long and short double-stranded arms measure 3.50 ± 0.34 kb and 1.78 ± 0.16 kb, respectively; the single-stranded insertion loop measures 5.07 ± 0.33 kb. Nine molecules were measured. Restriction enzyme maps of these two Drosophila DNA fragments are shown, and the cleavage sites for Bam HI (↑), Eco RI (↓), Hind III (●), Hpa I (◇) and Xba I (○) are indicated. The approximate end points of the *copia* element were determined by comparing restriction sites with those of cDm351 (see Figure 3) and are indicated by the line above the maps.

Evidence for Transposition of a *copia* Element in the Embryo Genome

The organization of *412*, *copia* and *297* elements in the genomes of tissue culture cells indicates that these elements are transposable. The first evidence we obtained suggesting that transposition of *412*, *copia* and

297 is not limited to cells in culture came from the analysis of a segment of the Drosophila embryo genome which appears to have resulted from the insertion of a *copia* element into a tandem array of 3.0 kb units. We had previously shown that sequences homologous to these 3.0 kb units are found at the termini, or telomeres, of each chromosome arm (Rubin, 1978). In the course of those experiments, we isolated two plasmids containing cloned segments of D. melanogaster DNA, cDm356 and pPW220; regions of each of these DNA segments are composed of tandem repeats of the 3.0 kb unit, and the remaining Drosophila DNA in each plasmid consists of a portion of an element of the *copia* dispersed repeated gene family.

The D. melanogaster DNA segments cloned in pPW220 and cDm356 were most probably derived from a single chromosomal site containing a *copia* element flanked on either side by a tandem array of 3.0 kb repeating units. [There are approximately 10 copies of the 3.0 kb repeat per haploid genome (S. S. Potter and G. M. Rubin, unpublished results), six of which can be found on cDm356 and pPW220.] Figure 10 shows a model for the structure of such a chromosomal site and a proposal for its generation. As is shown below, the 3.0 kb unit of the tandem repeat is interrupted at the same site in both cDm356 and pPW220, as expected for a simple insertion such as the one described in Figure 10. We have also been able to isolate a Bam HI restriction fragment of embryo DNA, predicted by the model shown in Figure 10, which contains a complete *copia* element flanked on each side by 3.0 kb repeat sequences. To isolate this fragment, we screened approximately 20,000 independent hybrid plasmids made between Bam HI fragments of embryo DNA and the plasmid vector pBR322 for those plasmids which contained both 3.0 kb repeat and *copia* sequences. One plasmid, cDm2087, contained a Bam HI fragment of the predicted structure (Figure 10).

The Insertion of *copia* Is Precise to within 90 bp

If the sequences found on cDm356 and pPW220 resulted from a simple insertion of a *copia* element into a tandem array of identical 3.0 kb units as shown in Figure 10, then the position at which the 3.0 kb unit sequences are interrupted by the *copia* element on each plasmid should correspond precisely. Restriction enzyme maps of the regions of cDm356 and pPW220 where junctions between the *copia* element and 3.0 kb units occur and of the corresponding region of an uninterrupted 3.0 kb unit are shown in Figure 11. The data used to map the Hae III cleavage sites are shown in Figure 12. Two Hae III cleavage sites, 90 base pairs (bp) apart, are present in each uninterrupted 3.0 kb unit at distances of 35 and 125 bp to the left of the Bgl I site. The Hae III site 35 bp

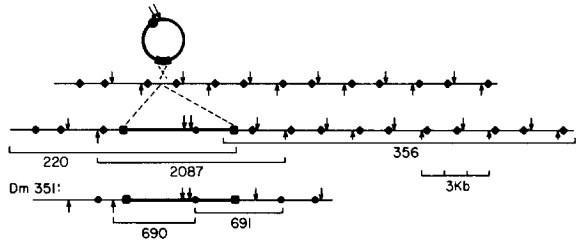

Figure 10. Model for the Proposed Transposition of a *copia* Element into a Tandem Series of 3.0 kb Repeat Units

Most simply, integration would result in the insertion of a *copia* element into one of the 3.0 kb repeats, thereby dividing that repeat into two noncontiguous segments without otherwise altering its original nucleotide sequence. The regions of this D. melanogaster chromosomal site which were cloned in cDm356, pPW220 and cDm2087 are indicated. Also included in this figure is a restriction map of the D. melanogaster DNA segment contained in cDm351, a plasmid which contains an entire *copia* element, but from a different chromosomal location (Finnegan et al., 1978). Indicated on this map are the 3.5 kb Bam HI × Hind III and the 4.0 kb Hind III fragments which were recloned in pMB9 to create the plasmids pmDm690 and pmDm691, respectively. Restriction endonuclease cleavage sites for Eco RI (↓), Bam HI (↑), Kpn I (◆) and Hind III (●) are indicated.

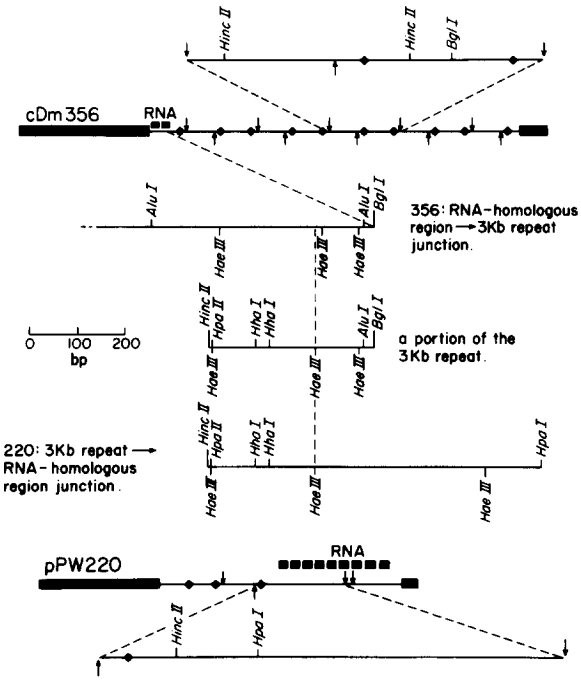

Figure 11. Restriction Site Maps of Junctions between *copia* and the 3.0 kb Repeat Unit on cDm356 and pPW220

A map of cDm356 is presented at the top of the figure, with one 3.0 kb repeat expanded to show finer detail. At the bottom of the figure is a map of pPW220, with a 3.9 kb segment expanded to show the junction between the left end of *copia* and the 3.0 kb repeat. Restriction enzyme cleavage sites for Kpn I (◆), Eco RI (↓) and Bam HI (↑) are indicated. In the center, a much finer scale map of a 350 bp Bgl I × Hinc II fragment of the 3.0 kb repeat is presented, and this region is compared with the regions of cDm356 and pPW220 which contain the junctions between the 3.0 kb repeat and the right and left ends of *copia*, respectively. There are two Hae III sites, 90 bp apart, in the 3.0 kb repeat fragment; at each junction, the transition from 3.0 kb repeat to *copia* occurs at some point between these two Hae III sites.

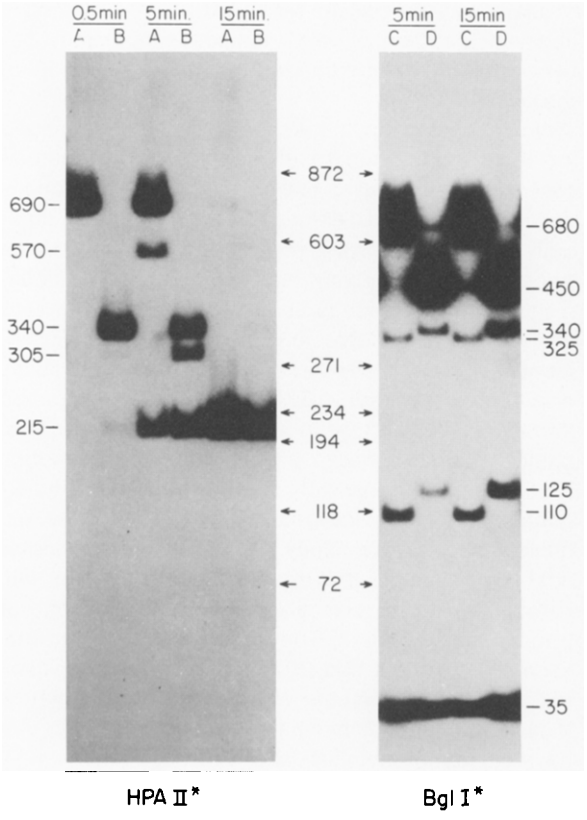

Figure 12. Partial Digestion with Hae III Localizes the Transition from the 3.0 kb Repeat to *copia*

The autoradiograph in the left panel compares a 690 bp Hpa II/Hpa I fragment of pPW220 (lanes A) which contains the junction of the left end of *copia* with a 3.0 kb repeat unit, and a 340 bp Hpa II/Bgl I fragment from the homologous region of an uninterrupted 3.0 kb repeat unit of cDm356 (lanes B). Each fragment was ^{32}P-labeled at the Hpa II end and subjected to partial digestion with Hae III for the indicated times (Smith and Birnstiel, 1976). One Hae III site, 215 bp from the Hpa II end, is present in both fragments, but the second Hae III site of the 3.0 kb fragment (lanes B), 305 bp from the Hpa II end, is absent from the pPW220 fragment (lanes A). The autoradiograph of the right panel compares a 680 bp Bgl I/Hinf I fragment of cDm356 (lanes C), which contains the junction of the right end of *copia* within the 3.0 kb repeat, and a 450 bp Bgl I/Hinf I homologous fragment of the 3.0 kb repeat from cDm356 (lanes D). Each fragment was ^{32}P-labeled at the Bgl I end and subjected to partial digestion with Hae III for the indicated times. These two fragments share a common Hae III site, 35 bp from the Bgl I site, but the second Hae III site of the 3.0 kb repeat fragment, 125 bp from the Bgl I site (lanes D), is not present in the junction fragment (lanes C). All fragment lengths were determined by comparison with markers, which consisted of ^{32}P end-labeled fragments of ϕX174 digested with Hae III, electrophoresed in adjacent lanes of these 10% acrylamide gels.

from the Bgl I site is also present in cDm356 near the junction with the *copia* element, but the Hae III site 125 bp from the Bgl I site is absent (Figure 12, right panel). This Hae III site is found, however, on the other side of the *copia* element near the junction between the element and the 3.0 kb unit sequences on pPW220 (Figure 12, left panel). The integration of *copia* into this tandem array of 3.0 kb repeating units therefore occurred in the 90 bp region between these two Hae III sites and altered or destroyed no more than 90 bp, and possibly much less, of the sequence originally present at the integration site.

Weak Sequence Homology Exists between the Direct Terminal Repeats of *copia* and the Region of Its Insertion

If *copia* sequences are capable of transposition, then it is of interest to determine whether potential integration sites possess a distinctive structure which enables *copia* to recognize them. cDm356 and pPW220 afford a system for examining this question because they contain a tandem series of 3.0 kb units apparently identical to the one disrupted by the integration of *copia*. As such, each 3.0 kb repeat harbors an unoccupied integration site with its characteristic features, if any.

The results of the experiment presented in Figure 13 demonstrate the existence of sequence homology between *copia* and the 3.0 kb repeat units. *Copia* sequences from cDm351 were hybridized to several different restriction enzyme digests of cDm356 by the method of Southern (1975). The intense labeling of the largest fragment in each digest represents hybridization to the approximately 500 bp of *copia* present in cDm356. A very much weaker hybridization to certain fragments derived from the 3.0 kb repeat units of cDm356 is detected. These results indicate that there are sequences homologous to the *copia* element present in the 865 bp region between the Hinc II and Kpn I cleavage sites on each of the 3.0 kb units, and that these *copia*-homologous sequences are present on both sides of the Bgl I cleavage site. This 865 bp region includes the 90 bp region between Hae III series within which, in one of the 3.0 kb units, *copia* has inserted. The extent and/or fidelity of this observed homology appears to be low; the amount of hybridization is at least 20 fold less than that observed between well paired duplexes 300 bp long. Furthermore, our attempts to map this region of homology by electron microscope heteroduplex techniques have been unsuccessful. We have therefore attempted to map the position of those sequences on the *copia* element which are homologous to the 3.0 kb unit by repeating the experiment shown in Figure 13 using particular regions of the *copia* element as probes, rather than the entire element. These experiments suggest that the homologous sequences are within the 300 bp terminal repeats of the *copia* element. Fragments of cDm351 containing either of the *copia* terminal repeats, such as Dm690 and Dm691 (Figure 10), hybridized to the 3.0 kb repeat. Dm690 was cut by Hpa I to yield a 2.5 kb *copia*-internal fragment and a 0.9 kb fragment which contains one of the 300 bp terminal repeats; only the 0.9 kb fragment containing the terminal repeat showed homology to the 3.0 kb unit.

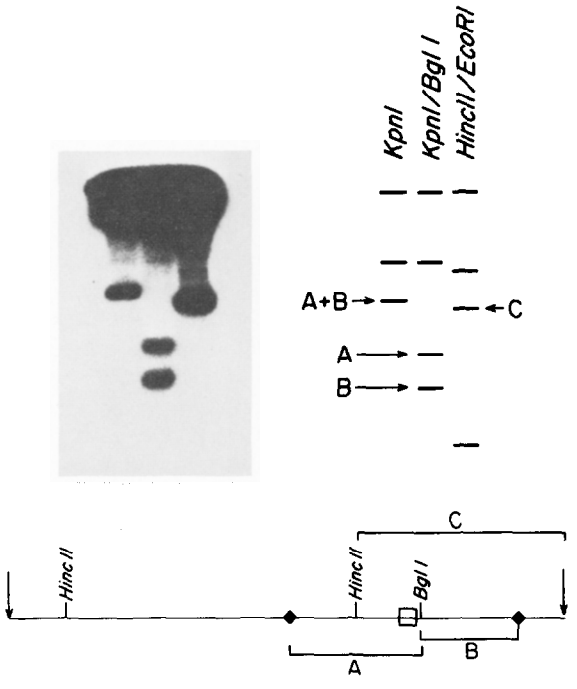

Figure 13. Hybridization of *copia* Sequences to Restriction Fragments of cDm356

Kpn I, Kpn I plus Bgl I and Hinc II plus Eco RI digests of cDm356 were separated on 1.4% agarose gels. The right panel shows a diagram of the separated fragments. The left panel shows the autoradiograph obtained when these restriction fragments are subsequently hybridized according to the method of Southern (1975) with the ^{32}P-labeled, 4.2 kb Hha I fragment of cDm351 (Finnegan et al., 1978). This fragment contains nearly the entire *copia* element. A map of one 3.0 kb repeat unit of cDm356 is shown. Restriction enzyme cleavage sites for Eco RI (↓), Kpn I (◆), Bgl I and Hinc II are indicated. The location of the fragments A, B and C, which show homology to *copia*, are indicated on the map. The open box represents the 90 bp region within which, in one of the 3.0 kb units, *copia*, has inserted.

Concluding Remarks

We have shown that there are different amounts of sequences homologous to the elements of the *412*, *copia* and *297* dispersed repeated gene families in Drosophila embryonic and tissue culture cells, and that the amount of these sequences is generally increased in the cell lines. Analysis of restriction enzyme cleavage sites within genomic sequences homologous to these elements indicates that most, if not all, of these sequences occur as intact elements in the cell culture genome, indistinguishable from those found in the embryo. Analysis of restriction enzyme cleavage sites in DNA flanking these elements indicates that the local genomic environments in which the elements occur differ not only among embryos and cell cultures, but also among individual cells in the same cell culture population. These differences were shown to be the result of transpositions of *412*, *copia* and *297* elements. We have also described a DNA sequence arrangement in the embryo genome which apparently resulted from the transposition of a *copia* element to a new chromosomal site. The insertion of this *copia* element was precise to within 90 bp and may have involved a region of weak sequence homology between the site of insertion and the terminal repeats of the *copia* element.

Numerous molecular mechanisms can be envisaged for the transposition of elements of the dispersed repeated gene families during passage of Drosophila cell cultures. These elements have structures strikingly similar to certain procaryotic drug-resistance transposons, and many models for the the mechanisms of transposition of such elements in procaryotes are directly applicable to *412*, *copia* and *297* elements. These models have been discussed in detail elsewhere (Kleckner, 1977). They can be grouped into two classes: those in which the element, such as bacteriophage λ, has a highly preferred integration site(s), and those in which the element, such as insertion sequences and drug-resistance transposons, exhibits little site specificity. Both classes of models share the feature that the integration occurs with respect to a particular site on the bacteriophage or transposable element. The latter feature is also true for the *412*, *copia* and *297* elements; the restriction maps of these elements at each of the sites where they occur are not circular permutations but rather have the same beginning and end.

What degree of site preference in the target DNA is exhibited by elements of the dispersed repeated gene families? The data we have presented suggest that in at least one case, weak sequence homology exists between the terminal repeats of *copia* and its integration site. If such homology is a general feature, a high degree of specificity would be expected; however, the observation that the *412*, *copia* and *297* elements can insert at many places in the genome suggests that little site specificity is involved. It may be that homologous sites are preferred, but not required, for integration, or there may be very many homologous sites in the genome. We are now characterizing additional genomic integration sites.

It is possible to estimate a minimum rate for transposition of elements of the dispersed repeated gene families. We have shown that most, if not all, of the additional copies of *copia* in Schneider's cell line 2 are inserted into new sites in the genome. The number of additional *copia* elements, 110, is a minimum estimate of the average number of transposition events which have occurred in the lineage of each cell during the ten years that Schneider's line 2 cells have been grown in culture. Assuming a uniform rate, this would correspond to 10^{-3}–10^{-4} transpositions per element per generation.

Transpositions of elements of the dispersed repeated gene families is not limited to cells in culture. We have described a sequence arrangement in the embryo genome which almost certainly originated by the transposition of a *copia* element. We have also shown (Strobel et al., 1979) that different strains of D.

melanogaster have strikingly different distributions of the *412*, *copia* and *297* elements in their genomes. Moreover, the chromosomal locations of these elements are highly polymorphic even when individuals from the same laboratory stock are compared.

Experimental Procedures

Nucleic Acid Preparations

D. melanogaster (Oregon R) embryonic DNA and plasmid DNAs were isolated as described previously (Wensink et al., 1974) with minor modifications (D. J. Finnegan, G. M. Rubin, D. J. Bower and D. S. Hogness, manuscript in preparation). Restriction endonuclease digests and agarose gel electrophoresis were carried out as previously described (Finnegan et al., 1978). Individual restriction fragments were purified by electrophoresis of digests of the appropriate plasmid in horizontal agarose gels followed by elution of the desired fragment by a modification of the method of Thuring, Sanders and Borst (1975). The purified restriction fragments used as probes in the DNA-DNA reassociation experiments for the *412*, *copia* and *297* dispersed repeated gene families were obtained in this manner as follows. For *412*, we used the 4.9 kb Hind III fragment of cDm412 (Finnegan et al., 1978) which contains the right end of the *412* element (see map in Figure 2), including one terminal repeat sequence and approximately 0.5 kb of nonrepetitive DNA from the Dm412 chromosomal site. For *copia*, the 4.2 kb Hha I fragment which contains nearly the entire *copia* elements was used (Finnegan et al., 1978). The only Hha I sites within *copia* are within or near each terminal repeat sequence. For *297*, the 2.3 kb Eco RI fragment of pPW297 was used (see map in Figure 3). These DNAs were ^{32}P-labeled by nick translation (Rigby et al., 1977).

DNA was isolated from Drosophila cell cultures as follows. 10 ml of tissue culture cells (~10^8 cells) were harvested by centrifugation for 5 min at 3000 × g, washed with 10 ml of 0.15 M NaCl, 0.05 M Tris (pH 7.5) and resuspended in 10 ml of 0.03 M Tris (pH 8.0), 0.1 M EDTA, 100 μg/ml proteinase K (E. M. Biochemicals) at 0°C. 2.5 ml of 5% Sarkosyl were then added and the resultant lysate was incubated for 5–10 hr at 37°C. CsCl was added to the lysate to a final density of 1.701 and the DNA was purified by isopycnic centrifugation. Fractions containing DNA were pooled and dialyzed against 10 mM Tris (pH 7.5), 1 mM EDTA.

Enzymes

Restriction endonucleases were obtained from either Bethesda Research Laboratories or New England Biolabs. E. coli DNA polymerase I was a gift from M. Goldberg. E. coli DNA ligase was prepared as described by Panasenko (1977). T4 polynucleotide kinase was prepared as described by Richardson (1965) from T4amN82-infected E. coli B cells purchased from New England Biolabs.

Restriction Maps

Restriction sites were located by determining the lengths of fragments generated by single, double and partial digests. Partial digestion of end-labeled fragments was carried out by a modification of the procedure of Smith and Birnstiel (1976).

Hybridization Procedures

DNA reassociation rates were measured as follows. 10–100 μg of nonradioactive DNA isolated from either Oregon R embryos, Schneider's cell line 2 or the Kc_0 cell line were combined with 20,000 dpm of nick-translated, ^{32}P-labeled DNA (spec. act. >10^7 dpm/μg) of the described segments of *412*, *copia* or *297* plus 50 μg of ^{3}H-labeled E. coli DNA and 500 μg of salmon sperm DNA. DNA concentrations were determined by the diphenylamine reaction as described by Burton (1956). The mixture of DNAs was adjusted to 3 ml of 0.15 M sodium acetate, 10 mM Tris–HCl (pH 7.5), 1 mM EDTA, and sonicated with four 30 sec pulses using the microtip of a Branson Sonifier model 185 at a setting of 4. Following sonication, the DNA was ethanol-precipitated, redissolved in 160 μl of 2.5 mM EDTA, 5 mM Tris–HCl (pH 7.5), transferred to a small conical tube, covered with mineral oil and denatured by immersing in boiling water for 10 min. The tube was transferred to a constant temperature block set at 65°C and allowed to equilibrate for 2 min, and 40 μl of 1 M sodium phosphate buffer were then added (time zero). At each time point, 10 μl of solution were removed, added to 0.5 ml of 0.12 M sodium phosphate buffer, quickly frozen in liquid nitrogen and stored at −20°C. The fraction of single-stranded DNA at each time point was determined by hydroxyapatite (HAP, Biorad HTP) chromatography. 1 ml HAP columns were poured into 3 ml disposable plastic syringes held at 60°C by a surrounding water bath. After loading the samples, single-stranded DNA was eluted with 4.5 ml of 0.12 M sodium phosphate buffer, and double-stranded DNA was then eluted with 5 ml of 0.3 M sodium phosphate buffer. Samples were mixed with 13 ml Aquasol (New England Nuclear), and radioactivity was monitored in a Beckman LS8000 scintillation counter using automatic quench control. The E. coli DNA was included as an internal standard to control for fluctuations in either the extent of sonication, the final salt concentration or the temperature of incubation. Values obtained for the $Cot_{1/2}$ of E. coli in different tubes in the same experiment varied by <5%, and by <20% when results obtained in different experiments were compared. In controls in which no Drosophila driver DNA was included, the ^{32}P-labeled probe remained >90% single-stranded at the last time point. Because the reassociations were carried out in 0.2 M sodium phosphate, the resulting Cot values were multiplied by 2.3 to give the equivalent Cot values shown in Figure 1 (Britten, Graham and Neufeld, 1974).

When nick-translated DNAs were hybridized to DNA transferred from agarose gels to nitrocellulose filters (Schleicher & Schuell, BA85) as described by Southern (1975), the strips were pretreated by soaking in 0.02% ficoll, 0.02% polyvinyl pyrrolidone, 0.02% bovine serum albumin (Denhardt, 1966) dissolved in 4 × SSCP (Rubin, 1978) for 6–8 hr at 65°C. Hybridization was carried out for 12–24 hr in the above solution containing 0.5% SDS, 200 μg/ml sonicated salmon sperm DNA and 10^5–10^6 cpm/ml of nick-translated DNA (spec. act. 1–10 × 10^7 cpm/μg). Filters were washed in 1 × SSC, 0.02% ficoll, 0.02% polyvinyl pyrrolidone, 0.02% bovine serum albumin for 1–2 hr at room temperature, and then in 3 mM Tris base for an additional 1–2 hr. In the hybridizations shown in Figure 13, the filter was washed in 2 × SSC at 65°C rather than in 3 mM Tris base.

Construction and Identification of Hybrid Plasmids

Hybrid plasmids were constructed between the plasmid pBR322 (Bolivar et al., 1977) and segments of Schneider's cell line 2 DNA by the following modification of the procedure described by Glover et al. (1975). The Bam HI-digested DNAs were mixed to give final concentrations of 150 μg/ml Schneider's cell line 2 DNA and 40 μg/ml pBR322 in ligation buffer [0.1 mM NAD, 1 mM EDTA, 10 mM Tris (pH 7.5), 10 mM $(NH_4)_2SO_4$, 10 mM $MgSO_4$, 100 μg/ml BSA]. The mixture was heated at 37°C for 5 min to melt self-annealed ends and cooled to 14°C, and E. coli ligase was added to a final concentration of 30 μg/ml. After 30 min, the DNA was diluted 50 fold with ligation buffer, more DNA ligase was added (final concentration 6 μg/ml) and the incubation was continued. After 16 hr, the salt was increased by adding 20 μl/ml of 1 M Tris (pH 7.5), and the ligation reaction was continued for another 24 hr. EDTA (pH 7.5) was then added to a final concentration of 50 mM, and the DNA was concentrated to 1 ml by dialysis against PEG 20,000 (Fisher Scientific) at 4°C. The concentrated ligation mixture was sedimented for 17 hr at 21,500 rpm in an SW27 rotor on a 38 ml, 10–25% sucrose gradient in 10 mM Tris (pH 7.5), 1 mM EDTA, 0.2 M NaCl at 20°C. Fractions were dialyzed against 10 mM Tris (pH 7.5) and used for transformations of the E. coli K12 strain HB101 (Boyer and Roulland-Dussoix, 1969) as described (D. J. Finnegan, G. M. Rubin, D. J. Bower and D. S. Hogness, manuscript in preparation). Fractions from the lower third of the gradient contained greater than two thirds hybrid plasmids and were used to generate colonies which were screened using the P. Gergen and P. S. Wensink (manuscript submitted) modification of the colony hybridization procedure of Grunstein and Hogness (1975).

Hybridization probes consisting of portions of the cloned Bam HI

fragments shown in Figure 4 which did not contain any *copia* sequences were isolated as follows: pBR322 contains only one Eco RI and one Hind III cleavage site, each located near the Bam HI site into which the Bam HI fragments of Schneider's line 2 DNA have been inserted as described above. Each hybrid plasmid was digested with either Hind III or Eco RI. The enzyme was then heat-inactivated, and the reaction mixture was diluted to a DNA concentration of 0.5 μg/ml in ligation buffer. E. coli DNA ligase was added to 30 μg/ml and the mixture was incubated overnight at 14°C. At this low DNA concentration, the primary reaction was self-circularization of the molecules, which were then used to transformation of E. coli HB101. The net result of this procedure is the subcloning of one end of the original inserted Bam HI fragment, which generates a DNA segment adjacent to, but not containing any, *copia* sequences.

Heteroduplex Mapping

The hybrid plasmids were digested with the restriction endonuclease Bam HI, phenol-extracted, ethanol-precipitated and then mixed together (5 μg/ml). Denaturation and renaturation in formamide were exactly as described by Davis, Simon and Davidson (1971). Heteroduplexes were spread from a hypophase containing 50% formamide. φX single-stranded DNA (Sanger et al., 1977) and T7 double-stranded DNA (McDonnell, Simon and Studier, 1977) were used as internal length standards to calibrate length measurements of heteroduplex molecules.

Acknowledgments

This work was supported by a grant from the NIH. S.S.P. was an NIH postdoctoral fellow, and W.J.B. is an NIH predoctoral trainee. P.D. was supported by a grant from the Damon Runyon-Walter Winchell Cancer Fund. We would like to thank Pieter Wensink for allowing us access to his library of cloned Drosophila DNA segments from which pPW297 was isolated; we are also grateful to Brian Backner for his excellent technical assistance.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received February 1, 1979; revised March 26, 1979

References

Barigozzi, C. (1971). *Drosophila* cell culture and its application for the study of genetics and virology. I. *Drosophila* cells *in vitro*: behavior and utilization for genetic purpose. Curr. Topics Microbiol. Immunol. *55*, 209–222.

Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., Crosa, J. H. and Falkow, S. (1977). Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene *2*, 95–113.

Boyer, H. W. and Roulland-Dussoix, D. (1969). A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. *41*, 459–472.

Britten, R. J., Graham, D. E. and Neufeld, B. R. (1974). Analysis of repeating DNA sequences by reassociation. In Methods in Enzymology, *29E*, L. Grossman and K. Moldave, eds. (New York: Academic Press), pp. 363–419.

Burton, K. (1956). A study of the conditions and mechanisms of the diphenylamine reaction for the colorimetric estimation of deoxyribonucleic acid. Biochem. J. *62*, 315–323.

Cameron, J. R., Loh, E. Y. and Davis, R. W. (1979). Evidence for transposition of dispersed repetitive DNA families in yeast. Cell *16*, 739–751.

Carlson, M. and Brutlag, D. (1979). A gene adjacent to satellite DNA in *Drosophila melanogaster*. Proc. Nat. Acad. Sci. USA *75*, 5898–5902.

Davis, R. W., Simon, M. and Davidson, N. (1971). Electron microscope heteroduplex methods for mapping regions of base sequence homology in nucleic acids. In Methods in Enzymology, *21*, L. Grossman and K. Moldave, eds. (New York: Academic Press), pp. 413–428.

Denhardt, D. T. (1966). A membrane-filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Commun. *23*, 641–646.

Echalier, G. and Ohanessian, A. (1969). Isolement, en cultures *in vitro*, de lignées cellulaires diploides de *Drosophila melanogaster*. CR Acad. Sci. (Paris) Series D *268*, 1771–1773.

Echalier, G. and Ohanessian, A. (1970). *In vitro* culture of *Drosophila melanogaster* embryonic cells. In Vitro *6*, 162–172.

Finnegan, D. J., Rubin, G. M., Young, M. W. and Hogness, D. S. (1978). Repeated gene families in *Drosophila melanogaster*. Cold Spring Harbor Symp. Quant. Biol. *42*, 1053–1063.

Glover, D. M., White, R. L., Finnegan, D. J. and Hogness, D. S. (1975). Characterization of six cloned DNAs from Drosophila melanogaster, including one that contains the genes for rRNA. Cell *5*, 149–157.

Grunstein, M. and Hogness, D. S. (1975). Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Nat. Acad. Sci. USA *72*, 3961–3965.

Hughes, S. H., Shank, P. R., Spector, D. H., Kung, H.-J., Bishop, J. M., Varmus, H. E., Vogt, P. K. and Breitman, M. L. (1978). Proviruses of avian sarcoma virus are terminally redundant, co-extensive with unintegrated linear DNA and integrated at many sites. Cell *15*, 1397–1410.

Ilyin, Y. V., Tchurikov, N. A., Ananiev, E. V., Ryskov, A. P., Yenikolopov, G. N., Limborska, S. A., Maleeva, N. E., Gvozdev, V. A. and Georgiev, G. P. (1978). Studies on the DNA fragments of mammals and Drosophila containing structural genes and adjacent sequences. Cold Spring Harbor Symp. Quant. Biol. *42*, 959–969.

Kleckner, N. (1977). Translocatable elements in procaryotes. Cell *11*, 11–23.

McDonnell, M. W., Simon, M. N. and Studier, F. W. (1977). Fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. J. Mol. Biol. *110*, 119–146.

Murray, K. and Murray, N. E. (1975). Phage lambda receptor chromosomes for DNA fragments made with restriction endonuclease III of *Haemophilus influenzae* and restriction endonuclease I of *E. coli*. J. Mol. Biol. *98*, 551–564.

Panasenko, S. M. (1977). Ph.D. thesis, Stanford University, Stanford, California.

Peacock, W. J., Lohe, A. R., Gerlach, W. L., Dunsmuir, P., Dennis, E. S. and Appels, R. (1978). Fine structure and evolution of DNA in heterochromatin. Cold Spring Harbor Symp. Quant. Biol. *42*, 1121–1135.

Richardson, C. C. (1965). Phosphorylation of nucleic acid by an enzyme from T4 Bacteriophage-infected *Escherichia coli*. Proc. Nat. Acad. Sci. USA *54*, 158–163.

Rigby, P. W. J., Dieckmann, M., Rhodes, C. and Berg, P. (1977). Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. *113*, 237–251.

Rubin, G. M. (1978). Isolation of a telomeric DNA sequence from *Drosophila melanogaster*. Cold Spring Harbor Symp. Quant. Biol. *42*, 1041–1046.

Rubin, G. M., Finnegan, D. J. and Hogness, D. S. (1976). The chromosomal arrangement of coding sequences in a family of repeated genes. Prog. Nucl. Acid Res. Mol. Biol. *19*, 221–226.

Sabran, J. L., Hsu, T. W., Yeater, C., Kaji, A., Mason, W. S. and Taylor, J. M. (1979). Analysis of integrated avian RNA tumor virus DNA in transformed chicken, duck and quail fibroblasts. J. Virol. *29*, 170–184.

Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. D.,

Fiddes, J. C., Hutchison, C. A., Slocombe, P. M. and Smith, M. (1977). Nucleotide sequence of bacteriophage ϕX174 DNA. Nature *265*, 687–695.

Schachat, F. H. and Hogness, D. S. (1974). Repetitive sequences in isolated Thomas circles from *Drosophila melanogaster*. Cold Spring Harbor Symp. Quant. Biol. *38*, 371–381.

Schimke, R. T., Kaufman, R. J., Alt, F. W. and Kellems, R. F. (1978). Gene amplification and drug resistance in cultured murine cells. Science *202*, 1051–1055.

Schneider, I. (1972). Cell lines derived from late embryonic stages of *Drosophila melanogaster*. J. Embryol. Exp. Morphol. *27*, 353–365.

Smith, H. O. and Birnstiel, M. (1976). A simple method for DNA restriction site mapping. Nucl. Acids Res. *3*, 2387–2397.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. *98*, 503–517.

Strobel, E., Dunsmuir, P. and Rubin, G. M. (1979). Polymorphisms in the chromosomal locations of elements of the *412*, *copia* and *297* dispersed repeated gene families in Drosophila. Cell *17*, 429–439.

Tchurikov, N. A., Ilyin, Yu.V., Ananiev, E. V. and Georgiev, G. P. (1978). The properties of gene Dm225, a representative of dispersed repetitive genes in *Drosophila melanogaster*. Nucl. Acids Res. *5*, 2169–2187.

Thuring, R. W. J., Sanders, J. P. M. and Borst, P. (1975). A freeze-squeeze method for recovering long DNA from agarose gels. Anal. Biochem. *66*, 213–220.

Wensink, P. C., Finnegan, D. J., Donelson, J. E. and Hogness, D. S. (1974). A system for mapping DNA sequences in the chromosomes of Drosophila melanogaster. Cell *3*, 315–325.