Volume 9 Number 24 1981 Nucleic Acids Research

A complex repeated DNA sequence within the Drosophila transposable element copia*

David L.Fouts and Jerry E.Manning
Department of Molecular Biology and Biochemistry, University of California at Irvine, Irvine, CA 92717, and

Gary M.Fox and Carl W.Schmid
Department of Chemistry, University of California at Davis, Davis, CA 95616, USA

Received 1 September 1981

ABSTRACT

A 320 nucleotide repeated DNA sequence within the copia coding element of Drosophila melanogaster has been identified and characterized. This sequence has been localized by DNA-DNA hybridization and electron microscopic analysis of heteroduplexes to the approximate middle of the 5 kb copia coding region. The primary sequence of this repeated DNA has been determined. The sequence is composed of three related subunits, 35-37 nucleotides in length (A, B and C). This 105 nucleotide higher order repeat has apparently been duplicated twice to yield a complex repeated sequence, ABCA'B'C'A"B"C", which exhibits divergence among the individual subunits. This sequence is AT rich, as are the direct terminal repeats which flank the copia coding region, but does not contain any apparent homology with the terminal repeats. This repeated sequence contains three presumptive polyadenylation signals and two 25 nucleotide, imperfectly matched, inverted repeat sequences adjacent to two of the polyadenylation sequences.

INTRODUCTION

Copia is a member of a class of repetitious, transposable DNA sequences in Drosophila characterized by widely scattered chromosomal locations (30 to 50) and direct repeated sequences which flank a conserved coding region (1). The coding region is approximately 5 kilobases (kb) in length and is bounded by direct repeats, 276 nucleotides long, whose sequence has been determined (2). The transposable nature of copia is reflected in the observation that the gene is found in different chromosomal locations and with different frequencies in selected strains of Drosophila, and in varying frequencies between tissue culture and embryo DNA's (3,4). Copia also has several features in common with transposable elements and insertion sequences in procaryotes (see: 5,6). For example, copia has short, imperfectly matched, terminal inverted repeats in each direct flanking repeat (2) and insertion of copia generates a five base pair duplication at the site of integration (7).

Copia codes for at least two abundant poly(A)-containing RNAs, approximately 2 kb and 5 kb in length (8,9,10), which constitute approximately 3% of

the total tissue culture cell polyadenylated RNA (1). The two RNAs are initiated at sites near the 5' end of the copia coding sequence (but not in the flanking direct repeat sequence) with the 2 kb RNA terminating in the approximate middle of the coding region (9). It is, however, not known whether the 2 kb RNA arises from a processing event of the 5 kb RNA, or whether there are different RNA polymerase initiation or termination sites in the coding sequence. Some heterogeneity in the electrophoretic migration of the 2 kb size RNA indicates that there may be multiple RNAs that differ slightly in size (9).

A number of protein products, ranging in size from 51,000 daltons to 18,000 daltons, have been made using RNA homologous to copia in in vitro translation systems (9,10). Studies employing hybrid arrested translation of the copia RNAs with individual restriction fragments from selected regions of the coding sequence indicated that the RNA's which code for these proteins all appear to come from the most 5' 2 kb of the copia coding sequence (9). These studies also show that the largest protein coded for by copia has a different initiation site from the smaller proteins and, at least in this case, the smaller proteins are not premature termination products of a single, larger molecule. These results suggest that copia codes for an mRNA like other abundant RNA's in Drosophila. Consistent with this are the observations that copia RNA's are modulated during development (9) and that the half-lives of these RNAs in the cytoplasm resemble other abundant mRNA species in Drosophila (10).

In this study, we report the existence of a complex repeated DNA sequence near the middle of the copia coding region which differs in sequence from the direct repeats that flank the copia gene. The repeat consists of three related, but distinct subunits (A, B and C), each approximately 35 nucleotides in length, and has the general arrangement ABCA'B'C'A"B"C". This repeat sequence contains a number of polyadenylation signals (AAUAAA) and inverted repeat sequences, and may represent the region in which termination or processing of the 2 kb mRNA(s) occurs.

MATERIALS AND METHODS

Selection of clones

The clones used in this study were selected from a library of Drosophila melanogaster genomic DNA cloned in the lambda phage Charon 4 (11), a gift of Dr. T. Maniatis (Harvard University, Boston, Mass.). The library was screened by the technique of Benton and Davis (12) with polyadenylated RNA

from L_2F81 tissue culture cells (a gift from R. Kanatka, California Institute of Technology L.A., CA). Selected clones were further analyzed by restriction analysis and in situ hybridization to Drosophila polytene chromosomes.

Filter hybridizations

All hybridizations were carried out in sealed plastic bags in 6 X SSC, 1 X Denhardt's (13) solution at 68°C, with constant agitation for 24-48 hrs. Filters were prehybridized for 2 hrs at 68°C in 6 X SSC (1X = 0.15 M sodium chloride, 0.015 M sodium citrate) 1 X Denhardt's, followed by the addition of fresh hybridization buffer and a mixture of sonicated calf thymus DNA (100 µg/ml), yeast tRNA (100 µg/ml), polyadenylic acid (20 µg/ml), and radiolabeled probe. The nucleic acids were boiled for 5 minutes prior to addition.

The filters were washed in 0.1 X SSC, 0.1% SDS (dodecyl sodium sulfate) for 4 x 30 minutes at 52°C with constant agitation, followed by two 30 minute washes in 0.1 X SSC at 52°C. The filters were air dried and exposed at room temperature to Kodak XRP-5 X-ray film using a single Dupont lighting-plus intensifying screen.

Restriction enzymes

All restriction enzymes were purchased from Bethesda Research Labs, Inc. (BRL) or New England Biolabs, and were used as recommended.

Radiolabeling of DNA

All DNA's were radiolabelled by nick translation using a nick translation kit (BRL) with α^{32}P-dCTP (2000-3000 Ci/mM) from Amersham.

Electrophoresis and Southern transfers

Restricted DNAs were electrophoresed overnight at 40 mAmps in 1.5% horizontal agarose (Seakem, ME) gels (5 mm thick), submerged 3 mm in Tris-borate (14) electrophoresis buffer. DNA was visualized using a UV transilluminator following staining of the gel for 1 hr in electrophoresis buffer containing 0.5 µg/ml ethidium bromide. DNA was transferred to nitrocellulose using the procedure of Southern (15), with a transfer solution of 6 X SSC.

Isolation of DNA fragments from Agarose gels

EcoRI restricted DNA fragments to be subcloned were first removed from agarose gels by transfer onto Whatman DE81 paper (R. Reeder, personal communication). The band to be removed was visualized after ethidium bromide staining and a slit was cut in the gel directly in front of the band. A strip of DE81 paper, moistened in electrophoresis buffer, was placed in the slit. The DNA was then electrophoresed at 60 mA until it had been transferred to the paper (determined by monitoring with UV light). The paper was removed and put in

the barrel of a 1 cc plastic syringe. The paper was washed four times by overlaying 100 µl each time of a solution of 0.1 M NaCl, 0.1 mM EDTA, 0.01 M Tris, pH 8.0, and centrifuging the syringe in a 15ml Corex tube at 2000 RPM for 30 seconds in a swinging bucket rotor. The DNA was eluted from the paper with four washes of 1.0 M NaCl in the same manner. After ethanol precipitation the DNA was resuspended and used without further treatment.

Subcloning of DNA fragments

The isolated EcoRI Drosophila DNA fragments were ligated into the plasmid pACYC 184 (16). The plasmid (a gift of K. Oishi, Univ. of Calif., Irvine CA), was restricted with EcoRI, and 1 µg of it was mixed with an equal amount of a Drosophila DNA fragment in 25 µl of ligation buffer (66 mM Tris, pH 7.5, 6.6 mM $MgCl_2$, 0.4 mM ATP and 10 mM DTT). One-half unit of ligase (BRL) was added and the mixture incubated at 4°C for 8 hours. The reaction mix was then diluted 10-fold with ligation buffer, another one half unit of ligase was added, and the ligation continued overnight at 4°C. The ligated DNA was used to transform the bacteria HMS174 (W3110, $R_{K12}^{-}M_{K12}^{+}$ RecA^{-}) (a gift of K. Oishi) exactly as described by Dagert and Erlich (17). Transformants were selected and potential recombinants were screened by replica plating onto nitrocellulose filters, with subsequent treatment and hybridization as described by Hanahan and Messelson (18). Selected recombinants were grown in liquid culture and the plasmid DNA's were isolated and purified (19).

3'-end-labeling of DNA for Sequencing

Recombinant plasmids containing the desired Drosophila fragments were digested with EcoRI or XbaI in preparation for labeling. All fragments were 3'-end-labeled using the large Klenow fragment of E. coli DNA polymerase I (22). DNA (approximately 50 pmole of 3' ends) was dissolved in 40 µl H_2O and 5 µl of buffer stock (0.5 M Tris, 0.05 M $MgCl_2$, 0.1 M mercaptoethanol, 0.5 mg/ml BSA) and then added to dried α^{32}P-dNTPs (25 µCi of each dNTP, 400 Ci/mmole, Amersham). The Klenow fragment of DNA polymerase I was then added (New England Biolabs, 5 µl of 0.7 units/µl), and the reaction mixture was incubated at 20°C for 30 min. Four microliters of a solution, 2.5 mM in each dNTP, was then added and the mixture was incubated an additional 15 min. The labeling reaction was terminated by the addition of 250 µl of 0.3 M sodium acetate, followed by ethanol precipitation. The DNA fragments labeled at both ends were then either strand separated on 5% polyacrylamide gels (23) or cleaved with another restriction enzyme and the resulting fragments separated on 10% acrylamide or 1% agarose gels. Fragments were isolated from agarose or acrylamide gels by excision of the band, followed by electrophoresis into a dialysis bag (24).

Bands were visualized by autoradiography and/or staining of the gel with ethidium bromide.

DNA sequencing

DNA sequencing was by the method of Maxam and Gilbert (23). In addition, reactions using $KMnO_4$ to cleave at T only, and hydroxylamine to cleave at C only, were performed (25). Gel electrophoresis was done on 30 cm long 0.3 mm thick 20% or 12% polyacrylamide gels for bases near the labeled end. In order to read further into the sequence, 8% gels, 1 M long, and 0.3 mm thick were used (23,26). Autoradiography was with Kodak XAR-5 or Dupont Cronex X-ray film at -80°C. In some cases, Dupont lighting-plus intensifying screens were used.

RESULTS

Subcloning of the 0.5kb Eco RI fragment from Copia

A lambda phage (clone DM311) containing the complete coding region of the *copia* gene was isolated from a library of *Drosophila* chromosomal DNA segments (11). A partial restriction endonuclease digestion map of lambda DM311 is shown in Figure 1. The endonuclease restriction sites within the *copia* coding sequence are identical to those reported by Dunsmuir *et al*. (7). To further confirm that DM311 contains *copia* this clone was hybridized *in situ* to *D. melanogaster* polytene chromosomes and hybridization to approximately 35 locations on the polytene chromosomes was observed. This observation is consistent with that previously seen by Young and Hogness (27) and Finnegan *et al*. (1).

For the purpose of experiments to be described elsewhere (Fouts, Levy and Manning, in preparation) the 0.5 kb Eco RI fragment within the coding region of *copia* was isolated and subcloned into the bacterial plasmid pACYC184 (Fig. 1). To insure the identity of the 0.5 kb fragment present in the recombinant, the plasmid was labeled by nick translation and hybridized to an EcoRI restriction digest of lambda DM311 by the method of Southern (15). The 0.5 kb DNA fragment hybridized to itself, and, susprisingly, to the 7.0 kb Eco RI fragment which flanks the 0.5 kb Eco RI fragment.

To pursue the possibility of homology between the 0.5 kb EcoRI fragment and its flanking sequences, clones of both the 3.2 kb and 7.0 kb EcoRI fragments were constructed (Fig. 1). Hybridization studies of Southern transfers of these clones probed with the 0.5 kb EcoRI fragment (data not shown) demonstrated that the 0.5 kb, and its adjacent 7.0 kb, EcoRI fragment shared homologous sequences. The region of this homology was further localized by restriction mapping to the 2.1 kb EcoRI/HpaI subfragment immediately adjacent to the 0.5 kb

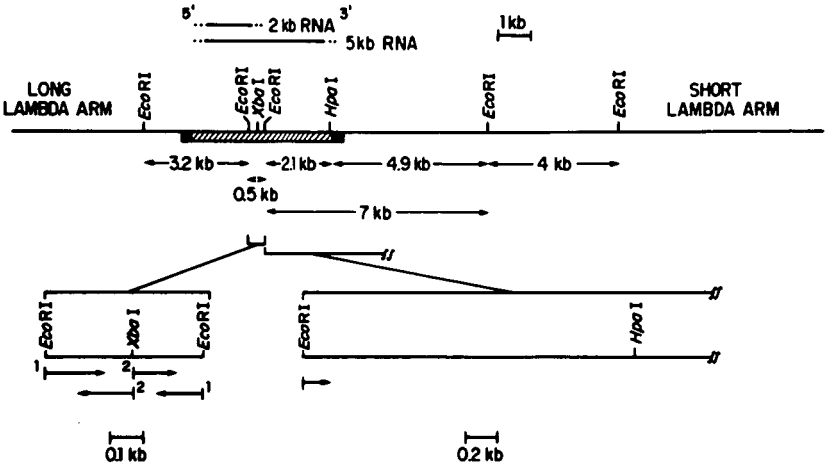

FIGURE 1. Restriction Map and Sequencing Strategy of Lambda Clone DM311.
The top diagram is a partial restriction map of clone DM311. The relative positions of the endonuclease restriction fragments mentioned in the text are shown. The position and polarity of the RNA molecules shown above the map are from Flavell et al. (9). The copia coding region is shown as the hatched bar with the terminal direct repeats shaded (2). The lower diagram shows the sequencing strategy employed. The purified 0.5 kb Eco RI fragment was either 3' end-labeled and then restricted with Xba I before sequencing (arrows labeled 1), or the plasmid containing the 0.5 kb Eco RI fragment was restricted with Xba I, 3' end-labeled, further restricted with Eco RI and the two Xba I-Eco RI fragments isolated before sequencing (arrows labeled 2). The purified 7.0 kb Eco RI fragment was 3' end-labeled, restricted with Hpa I and the 2.1 kb Eco RI-Hpa I fragment isolated and sequenced (arrow).

EcoRI fragment (Fig. I). There was no apparent homology between the 0.5 kb and its adjacent 3.2 kb EcoRI fragment. These results also imply that the direct repeats present on the ends of copia are not the source of homology to the 0.5 kb EcoRI fragment. This point is confirmed by our sequence results. (vide infra).

To localize the region of homology more precisely, heteroduplexes constructe between the 0.5 kb and 7.0 kb fragments were examined in the electron microscope using the T4-gene 32 protein technique. Within the resolution of this technique the 0.5 kb repeat was homologous to a 100 to 150 nucleotide region on the end of the 7.0 kb EcoRI fragment. In summary, these physical methods served to identify the presence of a repeated sequence within the copia coding region, apparently unrelated to the previously known terminal direct repeats. This internal repeat spans the EcoRI site between the 0.5 kb and 7.0 kb EcoRI fragments.

DNA sequencing

We have determined the base sequence of a portion of the copia coding

region near its two internal Eco RI sites to confirm the presence of the internal repeat. The sequencing strategy is shown in the lower portion of Figure 1. The end of the coding region closest to the Hpa I site has been chosen as the 3' end. This corresponds to the polarity of the major poly(A)-containing transcripts of copia (9). The overall spatial arrangement of the repeated sequences is shown in the upper portion of Figure 2. The base sequence starting at the 5' Eco RI site of the 0.5 kb fragment and continuing 170 bases past the 3' Eco RI site of the 0.5 kb fragment is shown in the middle portion of Figure 2.

As predicted by blot hybridization and electron microscopy data, the sequence contains an internal tandem repeat spanning the 3' EcoRI site of the 500 nt fragment. This repeat consists of three related, but distinguishable monomer subunits, which we will call A, B, and C, tandemly repeated in the arrangement $^{5'}$ABCA'B'C'A"B"C"$^{3'}$ (Fig. 2). Each monomer subunit is 35-37 nucleotides in length (except for C", which is missing 10 bases at its 3' end), so that each ABC-type trimer is about 107 nucleotides long. The first and second trimer units (ABC, positions 287-394, and A'B'C', positions 395-502) are nearly exact duplicates, with only one base change. The third trimer unit (A"B"C", positions 503-598) is about 25% divergent from the first two in sequence and is 10 nucleotides shorter due to a deletion at the 3' end of C". The monomer subunits A,B, and C are related, but not identical repeat units, with an average sequence divergence of about 33%. The divergence between any pair of monomer subunit types (A vs B, B vs C, A vs C) is approximately the same. In contrast, comparisons of A vs A', B vs B', and C vs C' show nearly perfect homology. Monomer subunits within the third trimer (A", B" and C") are considerably more divergent from their A, B, and C counterparts (about 25%), but still are of three distinct types. Detailed comparison of A with A", B with B", and C with C" shows that mutations are concentrated in the 3' side of the monomers. If we consider the 21 nucleotides of the 5' end of each monomer unit, the comparison A vs A" shows no mutations, B vs B" has 3 base changes, and C vs C" has 3 base changes. The 14-16 nucleotides in the 3' end are much more divergent: A vs A" has two changes and a 3-base deletion, B vs. B" has 8 base changes, and C vs. C" has 4 changes and a 9 base deletion. In summary, the 21 bases on the 5' ends are, on the average, 10% divergent [(6 changes) - (3 x 21 total bases) X 100] and the 14-16 bases on the 3' ends are 56% divergent [(25 changes) - (16 + 15 + 14 bases) X 100].

In addition to the repeat pattern described above, what appears to be another monomer unit is found starting at position 599. This monomer is

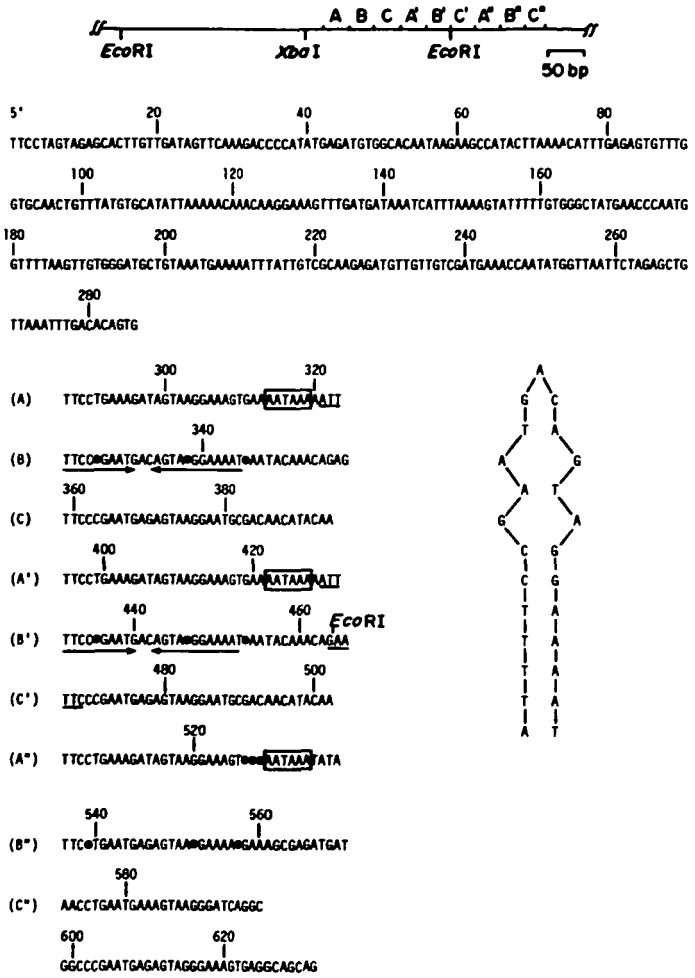

FIGURE 2. DNA Sequence of the Internal 0.5 kb Eco RI Fragment and the 3' Flanking Portion of the 7.0 kb Eco RI Fragment. The drawing above the sequence shows the spatial orientation of the repeats in the copia coding region (see Fig. 1). The sequence is oriented so that nucleotide 1 is the middle of the Eco RI restriction site located at the 5' end of the internal Eco RI fragment. The 3' internal Eco RI fragment is underlined and labeled. The three presumptive polyadenylation signals are indicated by the boxes. The inverted repeats are underlined with facing arrows. The black dots in the sequences B, B', A" and B" represent deletions added to yield the maximum register for the repeat sequences. The possible complementarity of the inverted repeat sequence is also shown.

closely related to the A,B, and C subunits through 21 bases, after which no recognizable homology with any subunit exists. We cannot, however, eliminate the possibility of additional repeat units further to the 3' end of the sequenced region.

The trimer repeat unit is somewhat AT-rich (68% AT vs 57% AT in total D. melanogaster DNA) (28). The A-type monomer units are especially AT-rich (76%) and have an A-T cluster at their 5' end (16 of 18 bases A or T), such as that found in the terminal direct repeats of copia (2). The three A-type monomers each contain a presumptive polyadenylation signal (bases 314-319, 422-427, and 527-532). Beginning one base 3' of each of the first two polyadenylation signals is a short, imperfectly matched, inverted repeat sequence. This sequence includes the end of the A and A' monomers and extends into the B and B' monomers, respectively (bases 321-345 and 429-453). These repeats are indicated in the sequence by the facing arrows and are shown base-paired in Figure 2.

DISCUSSION

Our principal result is that >320 nucleotides, located near the middle of the transcribed region of the 5.0 kb copia element, are arranged as tandem repeat units of length 35-37 nucleotides. A comparison of the nucleotide sequence of this internal repeat unit and the direct terminal repeats flanking the coding region indicates no apparent sequence homology between these two classes of repeats. Both repeats are, however, AT rich, 68 and 72%, respectively, and contain short, inverted repeat regions (2).

Although tandem repeat sequences have not previously been observed in either copia or other repeated sequence families in Drosophila which are similar to copia, it is interesting that short tandem repeats are present in at least one member (FB3) of the foldback (FB) family of transposable elements in Drosophila (29). In FB3, the inverted repeats consist of multiple copies of repeated sequences with monomer lengths of 10, 20, or 31 nucleotides. The 31 nucleotide repeat represents an expansion of the shorter repeat units and is found in tandem throughout the inner halves of the inverted repeats. There are, however, no obvious similarities between the tandem repeats in copia and those of the FB family. Although the 35 nt monomer unit of copia is similar in size to the 31 nt repeat in FB3, it is not generated by expansion of a simpler repeat, since no smaller repeated sequences are apparent in the 5' or 3' flanking sequences.

One feasible, but not exclusive mode of construction of the complex repeat

(i.e., ABCA'B'C'A"B"C") is that the three closely related sequences, A, B, and C, were generated by duplication events of the A sequence to yield AAA, followed by divergence of the sequence to give ABC. This 105 nt sequence could then have been duplicated to generate ABCABC. Following further sequence divergence, to yield ABCA"B"C", a second duplication of ABC could have occurred, yielding the final sequence, ABCA'B'C'A"B"C". We favor this order of duplication since the set ABC and A'B'C' differ by only one base (position 466), while the third set, A"B"C", shows marked divergence from the other two.

While the terminal direct repeats of copia appear to be involved in the mechanism of transposition of the copia element (4), no obvious function can be assigned to the internal tandem repeats. However, since the 2.0 kb RNA transcript of the copia element is known to terminate in or near the 0.5 kb Eco RI fragment (8,9), and potential polyadenylation recognition sites exist within the A, A', and A" subunits, it is possible that this repeat may serve as the termination site for the 2.0 kb transcript. Several observations favor this possibility. Multiple polyadenylation or termination sites have been observed in both eucaryotic viral genes (30,31,32), as well as the mouse α-amylase gene (33). Also, polyadenylation of eucaryotic mRNA frequently occurs 20-25 nucleotides downstream of the AATAAA sequence at a site where one or more A residues are present in the sequence (34,35). Within the A subunit, a 77% A-rich sequence stretch is found 21 to 29 nucleotides downstream from the putative poly A recognition site. It is also interesting to note that the most conserved subunit, A, contains the presumptive polyadenylation sequences.

It is informative to examine the general organization of middle repetitive DNA in Drosophila with regard to the structural arrangement of the multiple repetitive sequence elements (i.e., direct terminal and internal tandem repeats) in the copia sequence. A previous electron microscopic study of middle repetitive (MR) heteroduplex structures indicated that most of the middle repetitive sequences in Drosophila are long, having an average length of 5.6 kb (36). Clearly, copia contains short, multiply-repeated sequences, both internally as well as on both termini. However, assuming both of these types of repeat units maintain a consistent structural order among the dispersed members of the copia family, the heteroduplex structures generated by reassociation of the dispersed copia genes would be precisely those previously reported by Manning et al. (36). Young (37) found that at least 18 of the approximately 70 repetitive families that constitute the MR DNA are present in mobile repetitive elements. As discussed by Young, these results are consistent with the possi-

bility that most of the MR DNA in Drosophila is present in an arrangement similar to that found in copia.

In contrast to these results, Wensink et al. (38) has examined the sequence homology between four independently isolated repetitive sequences in Drosophila and concluded that the long repetitive sequences are clusters of a scrambled arrangement of short, moderately repetitive elements. Their conclusion is based on the observation of complex patterns of cross homology between restriction fragments of repeated sequences (38). The presence of internal repeats in copia, in addition to the previously known terminal repeats provides an alternate interpretation to some of their observations. The data presented here show that internal restriction fragments from within the copia sequence cross hybridize. In addition, restriction fragments which share the terminal direct repeats also cross hybridize (1). It is not unreasonable to assume that other repetitive families may contain similar tandemly repeated sequences. If this is the case, then restriction enzyme fragments of even discrete single family sequences would cross hybridize in a complex homology pattern.

ACKNOWLEDGEMENTS

This investigation was supported by P.H.S. Grant No. GM 22207 (JEM), by P.H.S. Grant No. CA 00522 awarded by the National Cancer Institute, DHHS (JEM), and by P.H.S. Grant No. GM 21346 (CWS).

*This paper is dedicated to Norman Davidson on his 65th birthday.

REFERENCES

1. Finnegan, D.J., Rubin, G.M., Young, M.W. and Hogness, D.S. 1977. Cold Spring Harbor Symp. Quant. Biol. 42: 1053-1063.
2. Levis, R., Dunsmuir, P. and Rubin, G.M. 1980. Cell 21: 581-588.
3. Strobel, E., Dunsmuir, P. and Rubin, G.M. 1979. Cell 17: 429-439.
4. Potter, S.S., Brorein, Jr., W.J., Dunsmuir, P. and Rubin, G.M. Cell 17: 415-427.
5. Kleckner, N. 1977. Cell 11: 11-23.
6. Calos, M.P. and Miller, J.H. 1980. Cell 20: 579-595.
7. Dunsmuir, P., Brorein, Jr., W.J., Simon, M.A. and Rubin, G.M. 1980. Cell 21: 575-579.
8. Carlson, M. and Brutlag, D. 1978. Cell 15: 733-742.
9. Flavell, A. J., Ruby, S. W., Toole, J.J., Roberts, B.E. and Rubin, G.M. 1980. Proc. Natl. Acad. Sci. USA 77: 7107-7111.
10. Falkenthal, S. and Lengyel, J.A. 1980. Biochemistry 19: 5842-5850.
11. Maniatis, T., Hardison, R.C., Lacy, E., Quon, D., Sim, G.K. and Efstratiadis, A. 1978. Cell 15: 687-701.
12. Benton, W.D. and Davis, R.W. 1977. Sci. 196: 180-182.
13. Denhardt, D.T. 1966. Biochem. Biophys. Res. Commun. 23: 641-646.

14. Greene, P. J., Betlach, M.C., Goodman, H.M. and Boyer, H.W. 1974. In DNA Replication. Methods in Molecular Biology. R. B. Wickner, editor; Marcel Dekker, Inc., New York. 7: 87-105.
15. Southern, E.M. 1975. J. Mol. Biol. 98: 503-517.
16. Chang, A.C.Y. and Cohen, S.N. 1978. J. of Bacteriology. 134: 1141-1156.
17. Dagert, M. and Ehrlich, S.D. 1979. Gene 6: 23-28.
18. Hanahan, D. and Messelson, M. 1980. Gene 10: 63-67.
19. Clewell, D.B. and Helinski, D.R. 1970. Biochemistry 9: 4428-4440.
20. Davis, R.W., Simon, M. and Davidson, N. 1971. In Methods in Enzymology. L. Grossman and K. Moldave, eds. Academic Press, New York, 21D: 413-428.
21. Holmes, D.A., Cohn, R.H., Kedes, L.H. and Davidson, N. 1977. Biochemistry 16: 1504-1512.
22. Klenow, H. and Henningsen, I. 1970. Proc. Natl. Acad. Sci. USA 65: 168-175.
23. Maxam, A.M. and Gilbert, W. 1980. In Methods in Enzymology, L. Grossman and K. Moldave, eds. Academic Press, New York 65: 499-560.
24. McDonell, M.W., Simon, M.N. and Studier, F.W. 1977. J. Mol. Biol. 110: 119-146.
25. Rubin, C.M. and Schmid, C.W. 1980. Nucl. Acid, Res. 8(20): 4613-4619.
26. Sanger, F. and Coulson, A.R. 1978. FEBS Letters 87: 107-110.
27. Young, M.W. and Hogness, D.S. 1977. In Eukaryotic genetics system. ICN-UCLA Symposia on Molecular and Cellular Biology. G. Wilcox et al., eds. Academic Press, New York. 8: 315-318.
28. Hasting, J.R.B. and Kirby, K.S. 1966. Biochem. J. 11: 532-538.
29. Truett, M.A., Jones, R.S. and Potter, S.S. 1981. Cell 24: 753-763.
30. Dijkema, R., Maat, J., Dekker, B.M.M., van Ormondt, H. and Boyer, H.W. 1981. Gene 13: 375-385.
31. Engler, J.A. 1981. Gene 13: 387-394.
32. Akusjarui, G., Zabielski, J., Perricaudet, M. and Pettersson, U. 1981. Nucl. Acids Res. 9(1): 1-17.
33. Tosi, M., Young, R.A., Hagenbucke, O. and Shiber, U. 1981. Nucl. Acids Res. 9: 2313-2323.
34. Alestrom, P., Akusjarvi, G., Perricaudet, M., Mathews, M.B., Klessig, D.F. and Pettersson, U. 1980. Cell 19: 671-681.
35. Benoist, C., O'Hare, K., Breathnach, R. and Chambon, P. 1980. Nucl. Acids Res. 8: 127-142.
36. Manning, J.E., Schmid, C.W. and Davidson, N. 1975. Cell 4: 141-155.
37. Young, M.W. 1979. Proc. Natl. Acad. Sci. USA 76: 6274-6278.
38. Wensink, P.C., Tabata, S. and Pachl, C. 1979. Cell 18: 1231-1246.