Cell, Vol. 10, 167-176, February 1977, Copyright © 1977 by MIT

A Novel Arrangement of the 18S and 28S Sequences in a Repeating Unit of Drosophila melanogaster rDNA

David M. Glover* and David S. Hogness†
Department of Biochemistry
Stanford University School of Medicine
Stanford, California 94305

Summary

The sequences corresponding to the 18S and 28S rRNAs have been mapped within a cloned 17 kilobase (kb) fragment formed by Eco R1 cleavage of Drosophila melanogaster rDNA. This fragment, Dm103, represents the longer of two major types of repeating units that are present in the rDNA of this fly, and was cloned as a hybrid plasmid, pDm103, consisting of Dm103 inserted at the Eco R1 site of the pSC101 vector (Glover et al., 1975). Mapping of the 18S and 28S rDNA in Dm103 was accomplished by quantitative determination of the amount of these rDNAs in each member of an ordered set of restriction fragments obtained by Hind III and Eco R1 cleavage of pDm103. The amounts of 18S and 28S rDNAs were determined by hybridization of the rRNAs to fragments that were purified by cloning, and an unambiguous order of the fragments within pDm103 was established by heteroduplex mapping and from the stoichiometry of the fragment lengths. The resulting map revealed that the 4 kb of 28S rDNA within the long repeating unit represented by Dm103 is divided into two blocks that are separated by 5.4 kb of DNA of unknown function. It is this unusual arrangement of the 28S rDNA that distinguishes the long repeating units (17 kb) from the short units (11.5 kb), whose 4 kb of 28S rDNA is confined to a single block, as is shown in the accompanying paper (White and Hogness, 1977). The remainder of the DNA in this long unit appears to be typically arranged, with the 2 kb of 18S rDNA confined to a single block that is separated by about 1 kb from the closest block of 28S rDNA.

Introduction

The genes that are transcribed to yield the 18S and 28S rRNAs of Drosophila melanogaster are clustered in two nucleolus organizers (NO)—one in the X chromosome and one in the Y (Rittossa et al., 1966). The number of genes per NO, though variable, is normally about 250 in the X chromosome and somewhat less in Y (Tartof, 1973; Spear, 1974). Electron microscopic observations of Drosophila rDNA in the act of transcription indicate that these genes are located in gene spacer units that are tandemly repeated in the NO (Hamkalo, Miller, and Bakken, 1973; Meyer and Henig, 1974; Glätzer, 1975; W. Y. Chooi and C. D. Laird, personal communication). Transcribed regions alternate with nontranscribed regions, or spacers, in this tandem array, and a unit consists of a spacer plus an adjacent transcribed region. The primary transcript obtained from this region is about 8 kilobases (kb) long and is the precursor of the 18S (2 kb) and 28S (4 kb) rRNAs (Loening, 1968; Greenberg, 1969; Perry et al., 1970; Meyer and Hennig, 1974; Jordan, 1975). [kb is a unit of length equal to 1000 bases or base pairs in single- or double-stranded nucleic acids, respectively].

In this and the accompanying paper (White and Hogness, 1977), we demonstrate that the repeating units in D. melanogaster rDNA divide into two major classes that differ in length and in the arrangement of the 28S rDNA sequences [Figure 1; see Glover et al. (1975) for the first mention of these two classes]. The members of one class are about 11.5 kb in length and exhibit the conventional arrangement of 18S and 28S rDNA sequences first worked out for Xenopus rDNA (Brown and Sugimoto, 1973). The second class differs from the first in that the 28S rDNA is divided into two blocks separated by about 5 kb of DNA of unknown function which increases the length of these units to approximately 17 kb. This unusual arrangement of the 28S rDNA has not been observed previously and forms the principal subject of these papers.

We have previously reported on the isolation of a cloned segment of D. melanogaster DNA (Dm segment) that belongs to this second class (Glover et al., 1975). This segment (Dm103) is a 17 kb Eco R1 fragment of embryonic rDNA that was cloned as a hybrid plasmid (pDm103) formed by insertion of Dm103 into the tetracycline-resistant plasmid, pSC101, at its Eco R1 cleavage site. In this paper, we map the 18S and 28S rDNA in Dm103 by a quantitative determination of the amount of these rDNAs in each member of a set of restriction endonuclease fragments that have been ordered within pDm103 DNA molecules.

Results

Both Eco R1 Sites and Four of the Six Hind III Sites in pDm103 Are Mapped by Comparing the Lengths of Fragments Produced by These Restriction Endonucleases

Figure 2 shows the two Eco R1 cleavage sites in pDm103 that link the 17 kb Dm103 segment to the 9 kb pSC101 plasmid DNA (Glover et al., 1975). Circular pDm103 also contains six Hind III sites which, when cut with this enzyme, yield the six fragments given in Table 1. pSC101 DNA is known to contain a

* Present address: Department of Biochemistry, Imperial College of Science and Technology, London SW7, England.
† To whom requests for reprints should be addressed.

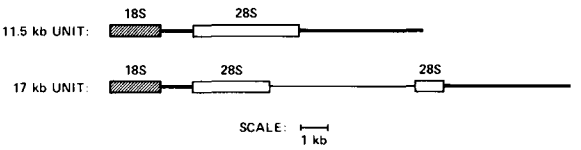

Figure 1. Two Classes of Repeating Units in D. melanogaster rDNA

The map for each unit was constructed from data presented in this and the accompanying paper (White and Hogness, 1977). The blocks labeled 18S and 28S refer to sequences in the rDNA that hybridize to 18S and 28S rRNAs isolated from D. melanogaster cell cultures.

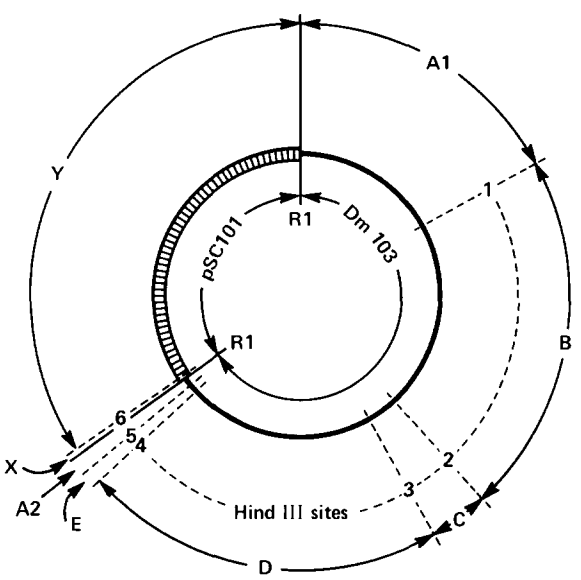

Figure 2. Eco R1 and Hind III Cleavage Sites in pDm103

The two Eco R1 sites are designated by R1, and the six Hind III sites by Arabic numerals. The eight lettered segments correspond to the fragments created by cleavage with both enzymes; A1, A2, B, C, D, and E derive from Dm103, and X and Y from pSC101 (see Table 1 for fragment lengths).

single Hind III site that is located only 0.03 kb from the single Eco R1 site in this plasmid (Carroll and Brown, 1976). In pDm103, this Hind III site should therefore be located 0.03 kb from one of the two Eco R1 sites and is designated site 6 in Figure 2. The other five Hind III sites must be located within the Dm103 segment.

Table 1 shows that when pDm103 is cut by both Hind III and Eco R1, four of the six Hind III fragments are untouched (that is, B, C, D, and E), and two are cleaved by Eco R1 to yield four new fragments (Y and A1; X and A2). The 0.30 kb Hind III fragment yields the expected 0.03 kb Hind III-Eco R1 fragment of pSC101 DNA (X) and a 0.27 kb fragment (A2) which must derive from one end of Dm103 and be adjacent to X. This locates one of the Hind III sites in Dm103 (site 5) at a distance of 0.30 kb counterclockwise from site 6 and 0.27 kb from the Eco R1 site that lies between Hind III sites 5 and 6 (Figure 2).

Table 1. Fragments of pDm103 Produced by Cleavage with Hind III and Eco R1

Lengths of Fragments Produced with:[a]		Nomenclature for
Hind III (kb)	Hind III + Eco R1 (kb)	Fragments Produced with Hind III + Eco R1
13.5	9.	Y
	4.5	A1
5.8	5.8	D
5.6	5.6	B
0.9	0.9	C
0.30	0.27	A2
	0.03 (approximately)	X
0.25	0.25	E

[a] The lengths of fragments ≥0.9 kb were determined by electrophoresis in 0.7% agarose gels using the following DNA molecules as length standards: Eco R1-cleaved pSC101 (Wensink et al., 1974); Eco R1 fragments of pDm103 and pDm102 (Glover et al., 1975); Eco R1 and Hpa II limit digestion fragments and Hpa I partial digestion fragments of SV40 (Nathans and Smith, 1975). The lengths of fragments <0.9 kb were determined in 6% acrylamide gels with reference to the mobilities of the Hind III fragments of PM2 DNA (Streeck, Philippsen, and Zachau, 1974). Fragment X was detected in 8% acrylamide gels.

The second Eco R1 site in pDm103 lies within the 13.5 kb Hind III fragment, since this fragment is cut by Eco R1 to yield the expected long 9 kb Hind III-Eco R1 fragment of pSC101 DNA (Y) and a 4.5 kb fragment (A1), which must derive from the other end of the Dm103 segment. This places another Hind III site in Dm103 (site 1) 4.5 kb from the second Eco R1 site (Figure 2).

A third Hind III site in Dm103 (site 4) can be located from the fact that partial digestion of pDm103 with Hind III yields a fragment with a length of 0.55 kb, as detected by electrophoresis in 6% polyacrylamide gels. A fragment of this length could be obtained by partial digestion only if the E fragment (0.25 kb) and the 0.30 kb Hind III fragment (X + A2) are adjacent to each other in pDm103. Hind III site 4 must therefore be 0.25 kb counterclockwise from site 5 (Figure 2).

The Remaining Hind III Sites Are Mapped by Electron Microscopy of Heteroduplexes Formed between Dm103 and Cloned Hind III Fragments

The Hind III fragments B, C, and D were purified by inserting each into the Col E1-kan plasmid (pML2 of Hershfield et al., 1974) at its single Hind III cleavage site, and cloning the resultant plasmids, which we choose to call ckDm103B, ckDm103C, and ckDm103D, respectively. The isolation and identification of these plasmids is described in Experimental Procedures.

The Col E1-kan vector is itself a hybrid plasmid formed by the insertion of an Eco R1 fragment

determining kanamycin resistance into the colicinogenic plasmid, Col E1, at its Eco R1 cleavage site (Hershfield et al., 1974). The lengths of Col E1-kan and Col E1 DNAs that we have obtained by electron microscopy are 13.71 kb (SE = ±0.13 kb; N = 34) and 6.54 kb (SE = ±0.06 kb; N = 46), respectively, and by difference yield a length of 7.2 kb for the Eco R1 kan-fragment, in reasonable agreement with the value of 7.3 kb obtained previously by gel electrophoresis (Cohen et al., 1973). This fragment contains the single Hind III cleavage site at which the B, C, and D fragments were inserted to form ckDm103B, ckDm103C, and ckDm103D, respectively.

These insertions are readily revealed when Eco R1 fragments of each plasmid are denatured, allowed to anneal briefly, and then examined in the electron microscope. Figure 3A shows that the single strands of the Eco R1 kan-fragment of Col E1-kan snap back upon themselves to form a hairpin that consists of a 1.14 ± 0.02 kb duplex segment with a 0.99 ± 0.04 kb single-stranded loop at one end, and two single-stranded tails of 1.53 ± 0.06 kb and 2.44 ± 0.03 kb at the other (mean values ± SE for N = 21 are given in each case; the sum of the lengths of the single-stranded regions plus twice the length of the duplex segment is 7.24 kb, consistent with the length of the kan-fragment determined above). The inverted repeats responsible for this hairpin have previously been identified with the region containing the kanamycin-resistant genes in R6-5 (Sharp, Cohen, and Davidson, 1973), the drug resistance factor that was the source of this Eco R1 kan-fragment (Cohen et al., 1973; Hershfield et al., 1974). Figures 3B and 3C show that when the Eco R1 kan-fragments of ckDm103B, ckDm103C, and ckDm103D are examined in this way, each yields a similar hairpin structure in which the length of the single-stranded loop is the only variable; and this equals, within experimental error (±10%), the length of the Col E1-kan loop plus the length of the inserted Hind III fragment (Table 1). Evidently the Hind III insertion site in Col E1-kan is located between the inverted repeats of its kan-fragment.

This conclusion has been confirmed and the position of four Hind III sites in Dm103 (sites 1, 2, 3, and 4) determined by examining the heteroduplexes formed between these hairpins and single strands of Dm103. We expect that the sequence corresponding to a Hind III fragment of Dm103 that is contained in the hairpin loop will interact with the complementary sequence in a Dm103 strand to form a duplex region within the loop. The length of this duplex region should equal the length of the inserted Hind III fragment, and the lengths of the single-stranded tails of Dm103 DNA at either end of this region should determine the location of the fragment within the Dm103 segment.

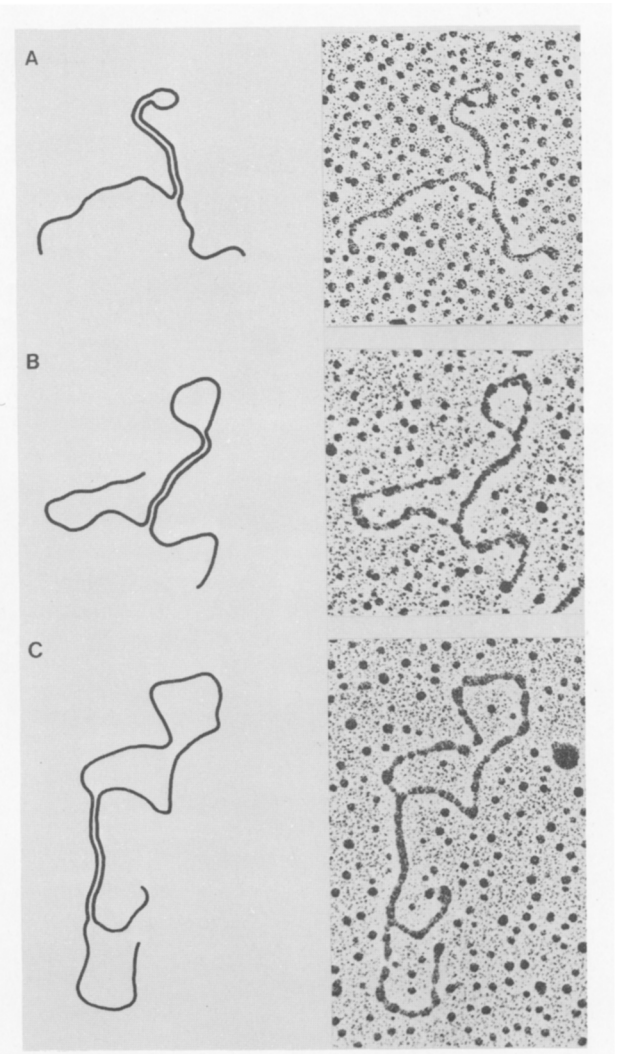

Figure 3. Hairpins Formed by Single Strands of the kan Fragments Obtained from the Co1 E1-kan Plasmid and the ckDm Hybrids

The circular DNAs were cleaved with Eco R1, denatured, and allowed to reanneal under the same conditions used for heteroduplex formation (Experimental Procedures). (A) Co1 E1-kan; (B) ckDm103C; (C) ckDm103B. The lengths of the single- and double-stranded regions are given in the text. ckDm103D forms a hairpin (not shown) that is almost identical to that formed by ckDm103B.

Figure 4 shows that the observed heteroduplexes conform to these expectations. The mean lengths (± SE) of the duplex regions within the hairpin loops formed from ckDm103B, ckDm103C, and ckDm103D are 5.59 ± 0.12 kb (N = 18), 0.95 ± 0.03 kb (N = 20), and 5.71 ± 0.10 kb (N = 17), respectively—in good agreement with the electrophoretically determined lengths of 5.6, 0.9, and 5.8 kb for fragments B, C, and D, respectively (Table 1). The lengths of the single-stranded Dm103 DNA tails that emanate from the duplex within the ckDm103B loop are 4.3 kb and 7.2 kb (Figure 4A), and are to be compared with the electrophoretically determined

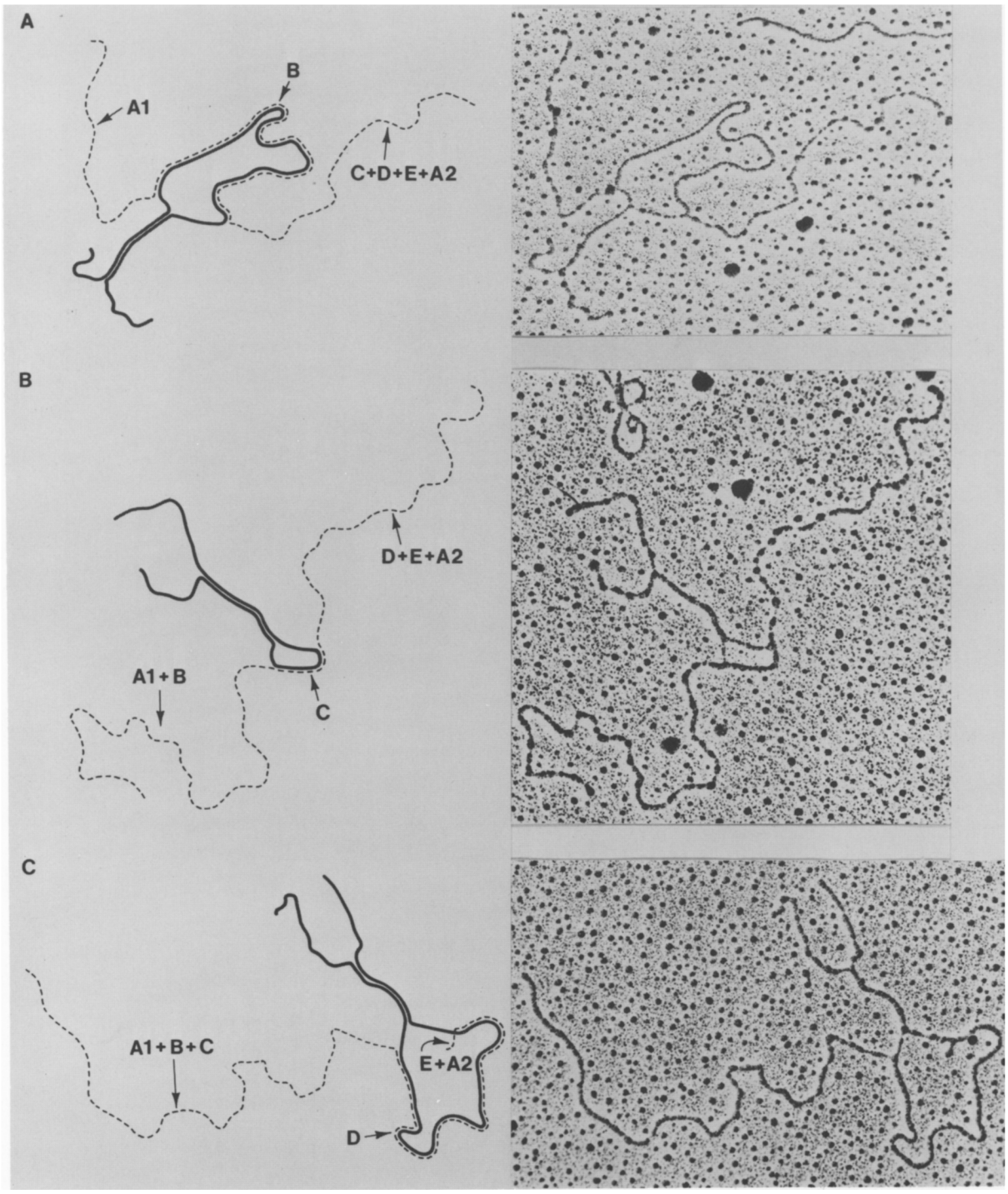

Figure 4. Heteroduplexes Formed between Dm103 and the Hairpins of the Eco R1 kan-Fragments from ckDm103B, ckDm103C, or ckDm103D

pDm103 and the indicated ckDm hybrid were cleaved with Eco R1, denatured, and annealed to form heteroduplexes (Experimental Procedures). The dashed line represents a strand from the Dm103 segment, and the continuous line indicates the hairpin formed by a strand from the Eco R1 kan-fragment of the ckDm hybrid. The letters associated with the dashed line refer to the fragments of Dm103 given in Table 1 and Figure 2. The mean lengths of the duplex region within the hairpin loops were measured relative to a pSC101 internal standard. The lengths of the two Dm103 single-stranded tails were determined from the ratio of their measured contour lengths and the sum of their lengths, as determined by subtraction of the above duplex length from the known total length of Dm103 (17.1 kb; Glover et al., 1975). These ratios are given below, and the tail lengths calculated from them appear in the text. The plasmid used to form the hairpins in the above heteroduplexes and the mean ratio (± SE) for the Dm103 tail lengths are: (A) ckDm103B—1.69 ± 0.07 (N = 18); (B) ckDm103C—1.46 ± 0.03 (N = 20); (C) ckDm103D—24.8 ± 2.9 (N = 17).

lengths of the A1 fragment (4.5 kb) and the sum of the C, D, E, and A2 fragments (7.2 kb), respectively (Table 1). These measurements place the B fragment adjacent to A1, and locate the Hind III site 2 at a position that is 5.6 kb (that is, the length of the B fragment) clockwise from site 1 (Figure 2). The D fragment is similarly placed adjacent to the E fragment because the lengths of the single-stranded Dm103 DNA tails in the ckDm103D heteroduplex are 0.44 kb and 11.0 kb (Figure 4C), which compare favorably with the expected values of 0.52 kb (fragments A2 + E) and 11.0 kb (A1 + B + C), respectively (Table 1). The Hind III site 3 must then be located 5.8 kb (that is, the length of the D fragment) counterclockwise from site 4, and fragment C must correspond to the 0.9 kb of DNA that lies between sites 2 and 3. This last conclusion is confirmed by the observation that the single-stranded Dm103 DNA tails in the ckDm103C heteroduplexes exhibit lengths of 6.6 and 9.6 kb (Figure 4B), in reasonable agreement with the expected values of 6.3 kb (A2 + E + D) and 10.1 kb (A1 + B) calculated from Table 1. These heteroduplex measurements therefore complete the map given in Figure 2 by locating sites 2 and 3, and confirm the previously assigned positions of sites 1 and 4.

The 18S and 28S rDNA Sequences Are Mapped by Hybridization of the rRNAs to the Ordered Fragments of Dm103

A map of the 18S and 28S rDNA in Dm103 has been constructed by determining the amount of these rDNAs in fragments B, C, and D, and in the combined A1 + A2 fragments. This has been accomplished by hybridizing saturating amounts of ^{3}H-labeled 18S and 28S rRNAs to nitrocellulose filters containing denatured ^{32}P-labeled ckDm103B, ckDm103C, ckDm103D, and pDm103A DNAs. These plasmid DNAs were used because quantitative hybridization under conditions of RNA saturation is critically dependent upon the purity of the DNA. The first three of these plasmids have been defined in the preceding section as sources of fragments B, C, and D that are uncontaminated with other sequences from Dm103. pDm103A consists of the A2-X-Y-A1 segment in pDm103 (Figure 2) that has been circularized by joining the Hind III termini of A1 and A2, and is described below.

pDm103A was obtained by treating a Hind III digest of pDm103 with E. coli DNA ligase, and then transforming tetracycline-sensitive E. coli K12 to resistance with the ligated DNA (Experimental Procedures). Six independent transformants were tested and each yielded the same plasmid, as judged by the common pattern of fragments observed when Eco R1 and Eco R1 + Hind III digests of each were subject to electrophoresis in 0.7% agarose gels. pDm103A was arbitrarily selected from these six. The lengths of the Eco R1, Hind III, and Eco R1-Hind III fragments obtained from this circular plasmid were determined electrophoretically, as described in the footnote to Table 1. One of the two Eco R1 fragments is 4.8 kb long and equals the sum of the lengths of A1 and A2; the length of the other is identical to pSC101—that is, to the X-Y Eco R1 fragment of pDm103 (Figure 2). Hind III similarly divides pDm103A into two fragments. The lengths of these fragments are 0.30 kb and 13.5 kb, and equal the lengths of the A2-X and Y-A1 Hind III fragments of pDm103 (Table 1). Finally, cleavage of pDm103A with both Eco R1 and Hind III yields four fragments whose lengths are equal to X, Y, A1, and A2—the four Eco R1-Hind III fragments obtained from pDm103 (Table 1). pDm103A must therefore have the structure,

where Eco R1 sites link X to A2 and Y to A1, and Hind III sites link X to Y and A1 to A2. The length of pDm103A determined by electron microscopy (13.63 ± 0.11 kb; N = 32) agrees, within experimental error, with that obtained from the sum of the electrophoretically determined lengths of its two Hind III fragments (13.8 kb) or of its two Eco R1 fragments (14.0 kb). Heteroduplexes formed between randomly nicked strands of pDm103A and pSC101, and pDm103A and pDm103 were examined; they exhibited the expected figure-eight configuration of a duplex ring attached to a single-stranded ring [not shown; see Figure 2 of Wensink et al. (1974) for an example of such a structure]. In each case, the length of the duplex ring equaled the length of the shorter of the two components making up the heteroduplex, whereas the single-stranded ring equaled the difference between their lengths. These results confirm the above structure since they indicate that pDm103A contains one equivalent of the pSC101 sequences in a contiguous segment (X-Y), and that all the sequences in pDm103A are included in a contiguous segment (A2-X-Y-A1) of pDm103.

The results obtained when either 28S rRNA or a mixture of 18S and 28S rRNAs was hybridized to filters containing pDm103A, ckDm103B, ckDm103C, or ckDm103D DNAs are given in Figure 5. A mixture of 18S and 28S rRNAs was used rather than 18S rRNA alone because Drosophila 28S rRNA consists, for the most part, of two hydrogen-bonded components, each about the size of the 18S rRNA, which readily dissociate and often contaminate preparations of 18S rRNA (Greenberg, 1969; Shine and Dalgarno, 1973; Jordan, 1975).

The amount of 18S rRNA bound to each DNA has therefore been calculated from the difference between the saturation values obtained for the mixture and for the 28S rRNA alone.

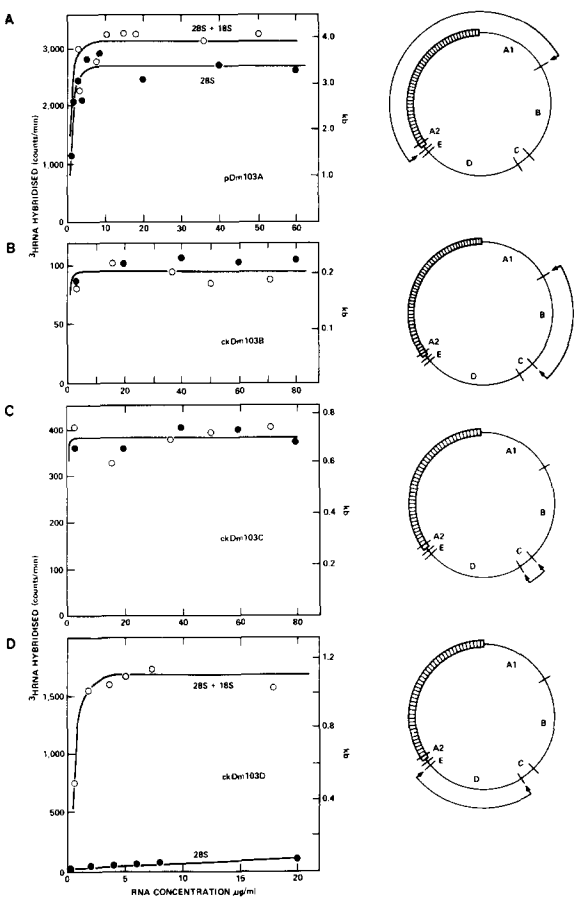

Figure 5. Saturation Hybridization of the Plasmid DNAs Containing Dm103 Fragments with 18S and 28S rRNA

^{3}H-labeled D. melanogaster rRNA was isolated from cell cultures (Glover et al., 1975) and had specific radioactivities of 2.20 × 10^5, 1.74 × 10^5, 1.74 × 10^5, and 2.04 × 10^5 cpm/μg for the experiments given in A, B, C, and D, respectively. The ^{32}P-labeled pDm103A, ckDm103B, ckDm103C, and ckDm103D DNAs had specific radioactivities of 4.80 × 10^3, 2.46 × 10^4, 2.04 × 10^4, and 1.76 × 10^4 cpm/μg, respectively. Hybridization was carried out on filters as described by Glover et al. (1975). The left-hand ordinate represents the mean ^{3}H cpm of triplicate (in A) or duplicate (in B, C, and D) determinations of the hybridized RNA at each RNA concentration, after correction for the cpm obtained with blank filters (or with filters containing DNA from pDm102, a non-rDNA plasmid; Glover et al., 1975), and for the ^{32}P cpm that are registered in the tritium channel, and after normalization to 0.1 μg DNA per filter (in A, B, and C) or to 0.33 μg DNA per filter (in D). The corrections typically involved a subtraction 40 cpm from the observed values. The right-hand ordinate represents the amount of RNA, in kb, that is hybridized per plasmid DNA molecule. Filled and open circles represent values obtained with 28S rRNA and with the mixture of 18S and 28S rRNAs, respectively.

A map of pDm103 showing its Hind III sites (radial lines) and the pSC101 segment (crosshatched region) is reproduced alongside each set of saturation curves. The arrows indicate the region of pDm103 contained in the plasmid used in each experiment.

The data given in Figure 5 indicate that the A1–A2 and D fragments contain 0.54 kb and 1.1 kb of 18S rDNA sequences, respectively, whereas no significant amounts of 18S rDNA are found in fragments B and C. While these results do not in themselves localize the 18S rDNA within the A1–A2 or D fragments, such a localization is possible if one adds the reasonable restriction that the 18S rDNA is confined to a single block in each repeating unit of the rDNA. Given this restriction, the 18S rDNA is distributed as shown in the map of Dm103 given in Figure 6. The 1.1 kb of 18S rDNA in fragment D is placed at the end adjacent to fragment E, and the 0.54 kb in A1–A2 occupies all the 0.27 kb in the A2 fragment and the 0.27 kb at the end of A1 adjacent to the Eco R1 site. This arrangement requires that the 0.25 kb in the untested E fragment consists of 18S rDNA and yields a total of 1.1 + 0.54 + 0.25 = 1.9 kb of 18S rDNA in Dm103. This value is in reasonable agreement with the size of Drosophila 18S rRNA determined by gel electrophoresis [2.0 kb (Loening, 1968); 1.9 kb (Perry et al., 1970); and 2.0 kb (Meyer and Hennig, 1974)].

Figure 5 indicates that the fragments A1–A2, B, and C contain, respectively, 3.3, 0.2, and 0.7 kb of 28S rDNA, or a total of 4.2 kb—in good agreement with the value of 4.1 kb found for the length of Drosophila 28S rRNA by gel electrophoresis (Loening, 1968; Perry et al., 1970; Meyer and Hennig, 1974). This distribution is surprising because there is no arrangement of the 28S rDNA sequences within each of these fragments that will allow them to be confined to a single block in Dm103 or in a repeating unit. This is because only 0.2 of the 5.6 kb in fragment B is 28S rDNA, and this fragment is located between fragments A1 and C, both of which must contain 28S rRNA.

In Figure 6, all the 3.3 kb of 28S rDNA in A1–A2 has been placed within A1 to conform with our previous assignment of the 18S rDNA. (A1 must contain at least 3 kb of 28S rDNA regardless of the

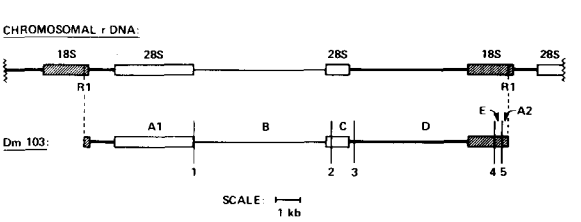

Figure 6. Map of Dm103 Showing the Arrangement of the 18S and 28S rDNAs and Indicating Its Origin from the Chromosomal rDNA

Dm103 is shown here as a complete repeating unit excised from the chromosomal rDNA by cleavage with Eco R1 at the sites indicated by R1. These sites correspond to the two Eco R1 sites in pDm103 (Figure 2). The five Hind III sites in Dm103 are represented by the numbered vertical lines. The positions of these sites on the map are determined by the lengths of the Eco R1-Hind III fragments (A1 and A2) and of the Hind III fragments (B, C, D, and E) given in Table 1.

distribution of 18S rDNA, since A2 is only 0.27 kb long.) The 0.2 kb of 28S rDNA in fragment B could be located at either of its ends to form a minimum of two blocks of 28S rDNA that are separated by at least the remaining 5.4 kb of DNA in fragment B. The arrangement that gives the best agreement with the R loop maps described in the accompanying paper (White and Hogness, 1972) is to place this 0.2 kb adjacent to the 0.7 kb of 28S rDNA in fragment C so as to form the 0.9 kb block shown in Figure 6. We have, for the same reason, placed the remaining 3.3 kb of 28S rDNA in a single block at the end of A1 next to fragment B. In this position, it is separated from the smaller 0.9 kb block of 28S rDNA by 5.4 kb, and from the 18S rDNA by 0.9 kb.

Discussion

Topography of the 18S rDNA

The map of Dm103 shown in Figure 6 was constructed according to the principle of minimizing the number of blocks of 18S and 28S rDNA per repeat unit. A single block of 18S rDNA per unit is sufficient to include all the 18S rDNA found in the different Dm103 fragments. If such a single block is assumed, then it follows that the Eco R1 termini of Dm103 derive from cleavage sites located within the 18S rDNA. This is so because the length of the 18S rDNA in A1+A2 (that is, in pDm103A; Figure 5A) is greater than the length of A2, and a single block can be obtained only if this 18S rDNA occupies all of A2 and the Eco R1 terminal region of A1. This conclusion that the 18S rDNA is confined to the two Eco R1 terminal regions in Dm103 is confirmed in the accompanying paper (White and Hogness, 1977).

The Eco R1 termini in Dm103 could result from cleavage of a single site within the 18S rDNA block in each of two adjacent repeat units, in which case Dm103 would represent a complete unit. Alternatively, Dm103 might result from cleavage at two different Eco R1 sites both located within the block, and in this case Dm103 would contain less than one equivalent of 18S rDNA. The first alternative is favored for the following reasons. The amount of 18S rDNA in Dm103 determined from saturation hybridization experiments is 2.1 (± 0.2) kb, and this equals, within experimental error, one equivalent of 18S rDNA as defined by the length of Drosophila 18S rRNA determined electrophoretically (Loening, 1968; Perry et al., 1970; Meyer and Hennig, 1974). The value given for the 18S rDNA in Dm103 is the mean (and range) of the 1.9 kb determined here from the sum of the amounts in the Dm103 fragments, and of the 2.3 kb obtained previously by hybridization to the entire Dm103 segment (Glover et al., 1975). These two values are not significantly different since the fractional standard error for each is about ±9%. (This is somewhat greater than the ±6% error associated with the 28S rDNA determinations since the measurement of 18S rDNA, but not of 28S rDNA, depends upon the difference between two saturation plateau values.)

The simplest interpretation of these data is that there is but one Eco R1 site within the 18S rDNA block. However, the precision of the hybridization measurements is not sufficient to exclude two or more sites that are closely spaced within the block. This is improbable since randomly sheared segments of D. melanogaster rDNA that have been cloned by the poly(dA)/poly(dT) technique of Wensink et al. (1974) and in which the 18S rDNA block is centrally located have been shown to contain only one Eco R1 site at a ± 0.03 kb level of resolution (D. J. Finnegan, G. M. Rubin, and D. S. Hogness, manuscript in preparation; D. M. Glover, unpublished experiments). Hence we conclude that Dm103 represents a complete unit of rDNA.

Topography of the 28S rDNA

The amount of 28S rDNA within Dm103 determined from the sum of the amounts in its fragments is 4.2 kb and agrees with the value of 4.1 kb obtained previously by hybridization of 28S rRNA to the entire Dm103 segment (Glover et al., 1975). As expected, this corresponds to one equivalent of 28S rDNA, since the length of 28S rRNA obtained by several groups is 4.1 kb (Loening, 1968; Perry et al., 1970; Meyer and Hennig, 1974). What is unexpected is that this 28S rDNA cannot be confined to a single block within the repeating unit represented by Dm103. The minimum number of blocks required to contain the 28S rDNA is two. A reduction of this minimum to one would necessitate a change in the order of the fragments within Dm103, and we are satisfied that the order shown in Figures 2 and 6 is unambiguous.

Given that the number of blocks is two, there are only two ambiguities in the map shown in Figure 6. One of these is the location of the 0.2 kb of 28S rDNA in fragment B, which could be included either with the 0.7 kb block in C or with the 3.3 kb block in A1. We have chosen the first of these alternatives simply because it conforms to the R loop maps given in the accompanying paper (White and Hogness, 1977). The other is the location of the 3.3 kb block in A1, which, for the same reason, we have so placed as to maximize the separation between it and the 18S rDNA in this fragment.

Regardless of how these ambiguities are resolved, the two blocks cannot correspond to the two large fragments of RNA that are obtained by heating Drosophila 28S rRNA (Greenberg, 1969; Shine and Dalgarno, 1973; Jordan, 1975). This is because each of these RNA fragments is about 2 kb in length, significantly shorter than the long block

of 28S rDNA in A1 and longer than short blocks in B + C. Hence it is not possible to imagine that the two RNA fragments are independently transcribed from the two 28S rDNA blocks. In addition, Jordan, Jourdan, and Jacq (1976) have shown that when the 28S rRNA first appears, it does not contain the central hidden break responsible for these fragments. This observation indicates that the transcription of all sequences within the 28S rRNA results from a single initiation event and, unless novel ad hoc mechanisms are introduced, eliminates the repeating unit shown in Figure 6 as the source of this 28S rRNA. In this regard, it should be emphasized that the 28S rRNAs referred to here have been isolated from D. melanogaster cell cultures. We do not wish to exclude the possibility that another 28S rRNA with an altered sequence is transcribed from the Dm103 type unit at some stage of D. melanogaster development. Indeed this possibility is treated in the accompanying paper (White and Hogness, 1977).

Experimental Procedures

Bacteria and Plasmids

The pSC101, pML2 (Col E1-kan), and pDm103 plasmids were obtained from the E. coli K12 strains C600 (pSC101) (Cohen et al., 1973), C600 (pML2) (Hershfield et al., 1974), and HB101 (pDm103) (Glover et al., 1975). HB101 is hsm^-, hrs^-, $recA^-$, gal^-, pro^-, str^R (Boyer and Roulland-Dussoix, 1969) and was used as the host for all new hybrid plasmids constructed here. Plasmids were isolated from these strains as described by Wensink et al. (1974).

Enzymes

The E. coli enzymes Eco R1 (Greene et al., 1974), RNA polymerase (Berg, Barrett, and Chamberlin, 1971), and DNA ligase (Modrich, Anraku, and Lehman, 1973) were prepared according to the indicated references and were gifts from P. Wensink, W. Wickner, and P. Modrich, respectively. Hind III (Smith and Wilcox, 1970) and Hae III (Roberts et al., 1975) were provided by S. Goff and D. Charney, respectively.

Electron Microscopy

DNA was spread for electron microscopy by the 40% formamide or aqueous methods (Davis, Simon, and Davidson, 1971) according to whether the DNA examined contained single-stranded regions or consisted only of duplexes, respectively. Contour lengths were measured (Schachat and Hogness, 1973) relative to reference DNAs on the same grid; M13 DNA (6.6 kb; Marvin and Hohn, 1969) and pSC101 DNA (9.2 kb; Wensink et al., 1974) were used for single- and double-stranded references, respectively. Heteroduplexes were formed and examined as described by Davis et al. (1971). Standard errors (SE; standard deviation in the mean) were calculated according to Davis et al. (1971).

Construction and Cloning of Hybrid Plasmids

Col E1-kan Hybrids Containing Hind III Fragments of pDm103

A mixture of 2.5 μg of pDm103 and 0.15 μg of Col E1-kan DNAs in 75 μl of Hind III reaction buffer [120 mM NaCl, 13.4 mM $MgCl_2$, 0.1 mM dithiothreitol, 12 mM Tris-HCl (pH 7.4)] was treated with sufficient Hind III to complete the digestion of these DNAs in 90 min at 37°C. After inactivating the Hind III by raising the temperature to 65°C for 5 min, DPN, $(NH_4)_2SO_4$, and BSA were added to obtain concentrations of 0.1 mM, 10 mM, and 100 μg/ml, respectively, in a final volume of 100 μl. E. coli DNA ligase (1 μg) was added, and ligation was allowed to proceed for 90 min at 15°C. The reaction mixture was then diluted to 500 μl with ligase reaction buffer [0.1 mM DPN, 100 μg BSA/ml, and 10 mM each of $(NH_4)_2SO_4$, $MgSO_4$, and Tris-HCl (pH 7.4)] containing 1 μg of DNA ligase, and the reaction continued at 15°C for an additional 15 hr.

E. coli K12 strain HB101 was transformed to colicin E1 immunity by this mixture of ligated DNAs according to the method described by Glover (1976). Those transformants that contain, in addition to the Col E1-kan DNA, one or more of the pDm103 Hind III fragments were identified by colony hybridization, using ^{32}P-cRNA transcribed from pDm103 DNA with E. coli RNA polymerase as the probe (Grunstein and Hogness, 1975). 79 clones, or 36% of the screened transformants, contained pDm103 sequences by this test, and 19% of these were tetracycline-resistant (Tc^R). Plasmid DNAs were isolated from 15 of the 79 clones, including two that were Tc^R. Determination of their contour lengths in the electron microscope and examination of their Hind III fragments by gel electrophoresis revealed that 10 of the 15 consisted of a *single* pDm103 Hind III fragment inserted into Col E1-kan. Hybrids that contain either the B (5.6 kb) or D (5.8 kb) fragments are difficult to distinguish on the basis of the electrophoretic mobilities of these fragments, and were further delineated by the distinctive pattern that their multiple Hae III restriction fragments exhibit upon gel electrophoresis [Hae III digestions were carried out at 37°C in 0.5 mM dithiothreitol, 10 mM $MgCl_2$, 10 mM Tris-HCl (pH 7.4)]. ckDm103B, ckDm103C, and ckDm103D were selected as representatives of hybrids containing the B, C, and D fragments, respectively. Each yields two fragments after Hind III digestion—one equal in length to Col E1-kan and the other to the respective fragment from Dm103. The means and standard errors of their contour lengths are: ckDm103B—19.13 ± 0.25 kb (N = 37); ckDm103C—15.39 ± 0.22 kb (N = 38); ckDm103D—19.57 ± 0.21 kb (N = 31).

A fourth hybrid obtained from one of the Tc^R clones was similarly examined and demonstrated to consist of the Y-A1 Hind III fragment of pDm103 inserted into Col E1-kan. It is designated ckDm103A1 and exhibits a contour length of 28.11 ± 0.25 kb (N = 29). Although not used in the experiments reported here, it is noted because of its relevance to the explanation given below for the isolation of pDm103A.

Isolation of pDm103A

The cloning of pDm103A (that is, the circularized A2-X-Y-A1 segment of pDm103; see text) was the unexpected result of a procedure designed to clone the circularized Y-A1 Hind III fragment of pDm103. Thus a solution of pDm103 DNA (325 μg/ml) in Hind III reaction buffer (see above) was treated with sufficient Hind III to complete its digestion in 90 min at 37°C, as judged by gel electrophoresis of the resulting fragments (Table 1). After raising the temperature to 65°C for 5 min to inactivate the Hind III, 2 μl of this reaction mixture (0.65 μg of DNA) were diluted into 500 μl of ligase reaction buffer (see above) that contained 1 μg of DNA ligase, and the mixture was incubated at 15°C for 15 hr. HB101 was then transformed to Tc^R by this mixture of ligated DNAs according to the procedure described by Wensink et al. (1974). Hybrid plasmids were isolated from each of six independent Tc^R transformants, and all exhibited the structure of pDm103A (see text), which was arbitrarily chosen from this set.

An Explanation for the Isolation of pDm103A

The low concentration of pDm103 Hind III fragments (7×10^{-11} M) used during the above ligation reaction favors intramolecular circularization over intermolecular joining. Hence the circular Y-A1 molecules are expected to be formed at a higher frequency than are the pDm103A molecules, whose formation requires either the joining of two fragments (A2-X and Y-A1) or the circularization of the small number of A2-X-Y-A1 molecules that escape final Hind III cleavage. Since the observed ratio of Tc^R transformants containing Y-A1 plasmid to those containing the pDm103A plasmids is zero to six, we infer that the circular Y-A1 molecules cannot carry out this transformation. Yet we know that the Y-A1 segment contains the tet gene(s) (that is, the gene(s) coding for

the polypeptide(s) responsible for the Tc^R phenotype), as these are expressed in the ckDm103A1 plasmid noted above.

These results can be explained by a model proposed by Alain Rambach which postulates that the promoter for tet transcription is inactivated by deletion of the X segment or by insertion at the Hind III site joining X and Y (A. Rambach and D. S. Hogness, manuscript in preparation). The expression of tet in hybrids that exhibit either characteristic will then depend upon whether the added DNA can provide a substitute promoter for this transcription. Thus it is assumed that tet expression fails in circular Y-A1 because A1 cannot provide such a promoter, but succeeds in ckDm103A1 because the Col E1-kan segment can. Similarly, the tet expression that Rambach has observed for a hybrid consisting of the Y segment and only one half the kan segment present in Col E1-kan is assumed to result from a promoter in this half.

The model can also be used to explain why we did not obtain Tc^R transformants with hybrids consisting the B, C, D, or E fragments of Dm103 inserted at the Hind III site of pSC101, even though Carroll and Brown (1976) used Tc^R transformation to clone hybrids consisting of Xenopus laevis 5S DNA inserted at this site. We assume that the 5 S DNA, but not the Dm 103 fragments, provides a sequence that is recognized as a promoter by E. coli RNA polymerase. Indeed, this assumption accounts for their observation that the orientation of the insertion was the same in all the independently cloned hybrids that they examined. Such a sequence could be contained within the 5S DNA (as is the case for the eucaryotic SV40 viral DNA; Dhar et al., 1974), or it could be formed when the pSC101 and inserted DNAs are joined.

Finally we note that Cabello, Timmis, and Cohen (1976) have mapped the replication origin of pSC101 at the center of Y. Hence one can dismiss an alternative explanation that X deletion and Hind III insertion interfere with Tc^R transformation by altering the origin sequence that is required for plasmid replication.

Containment Conditions

All plasmids containing Dm segments were propagated in EK1 host-vector systems under P2 physical containment, as defined and recommended for this class of experiments by Guidelines for Research Involving Recombinant DNA Molecules issued in June, 1976 by the National Institutes of Health, Bethesda, Maryland.

Acknowledgments

This work was supported by research grants from the NSF and the NIH. D.M.G. was a Damon Runyon fellow.

Received September 13, 1976, revised October 22, 1976

References

Berg, D., Barrett, K., and Chamberlin, M. (1971). Purification of two forms of Escherichia coli RNA polymerase and of sigma component. In Methods in Enzymology, *21D*, L. Grossman, and K. Moldave, eds. (New York: Academic Press), pp. 506–519.

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.

Brown, D. D., and Sugimoto, K. (1973). The structure and evolution of ribosomal and 5S DNAs in Xenopus laevis and Xenopus mulleri. Cold Spring Harbor Symp. Quant. Biol. *38*, 501–505.

Cabello, F., Timmis, K., and Cohen, S. N. (1976). Replication control in a composite plasmid constructed by in vitro linkage of two distinct replicons. Nature *259*, 285–290.

Carroll, D., and Brown, D. D. (1976). Adjacent repeating units of Xenopus laevis 5S DNA can be heterogeneous in length. Cell *7*, 477–486.

Cohen, S. N., Chang, A. C. Y., Boyer, H. W., and Helling, R. B. (1973). Construction of biologically functional bacterial plasmids in vitro. Proc. Nat. Acad. Sci. USA *70*, 3240–3244.

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, *21D*, L. Grossman and K. Moldave, eds. (New York: Academic Press), pp. 413–428.

Dhar, R., Weissman, S. M., Zain, B. S., and Pan, J. (1974). Nucleotide sequence preceding an RNA polymerase initiation site on SV40 DNA: II. The sequence of the early strand transcript. Nucl. Acids Res. *1*, 595–614.

Glätzer, K. H. (1975). Visualization of gene transcription in spermatocytes of Drosophila hydei. Chromosoma *53*, 371–379.

Glover, D. M. (1976). The construction and cloning of hybrid DNA molecules. In New Techniques in Biophysics and Cell Biology, *3*, R. Pain and B. Smith, eds. (London: John Wiley), in press.

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.

Greenberg, J. R. (1969). Synthesis and properties of ribosomal RNA in Drosophila. J. Mol. Biol. *46*, 85–98.

Greene, P. J., Betlach, M. C., Goodman, H. M., and Boyer, H. W. (1974). The EcoR1 restriction endonuclease. In Methods in Molecular Biology, *7*, R. Wickner ed. (New York: Academic Press), pp. 87–111.

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.

Hamkalo, B. A., Miller, O. L., and Bakken, A. H. (1973). Ultrastructure of active eukaryotic genomes. Cold Spring Harbor Symp. Quant. Biol. *38*, 915–919.

Hershfield, V., Boyer, H. W., Yanofsky, C., Lovett, M. A., and Helinski, D. R. (1974). Plasmid ColE1 as a molecular vehicle for cloning and amplification of DNA. Proc. Nat. Acad. Sci. USA *71*, 3455–3459.

Jordan, B. R. (1975). Demonstration of intact 26S ribosomal RNA molecules in Drosophila cells. J. Mol. Biol. *98*, 277–280.

Jordan, B. R., Jourdan, R., and Jacq, B. (1976). Late steps in the maturation of Drosophila 26S ribosomal RNA: generation of 5.8S and 2S RNAs by cleavages occurring in the cytoplasm. J. Mol. Biol. *101*, 85–105.

Loening, U. E. (1968). Molecular weights of ribosomal RNA in relation to evolution. J. Mol. Biol. *38*, 355–365.

Marvin, D. A., and Hohn, B. (1969). Filamentous bacterial viruses. Bacteriol. Rev. *33*, 172–193.

Meyer, G. F., and Hennig, W. (1974). The nucleolus in primary spermatocytes of Drosophila hydei. Chromosoma *46*, 121–144.

Modrich, P., Anraku, Y., and Lehman, I. R. (1973). Deoxyribonucleic acid ligase. J. Biol. Chem. *248*, 7495–7501.

Nathans, D., and Smith, M. O. (1975). Restriction endonucleases in the analysis and restructuring of DNA molecules. Ann. Rev. Biochem. *44*, 273–293.

Perry, R. P., Cheng, T.-Y., Freed, J. J., Greenberg, J. R., Kelley, D. E., and Tartof, K. D. (1970). Evolution of the transcription unit of ribosomal RNA. Proc. Nat. Acad. Sci. USA *65*, 609–616.

Rittossa, F. M., Atwood, K. C., Lindsley, D. L., and Spiegelman, S. (1966). On the chromosomal distribution of DNA complementary to ribosomal and soluble RNA. Nat. Cancer Inst. Monograph *23*, 449–472.

Roberts, R. J., Breitmeyer, J. B., Tabachnik, N. F., and Myers, P. A. (1975). A second specific endonuclease from Haemophilus aegyptius. J. Mol. Biol. *91*, 121–123.

Schachat, F. H., and Hogness, D. S. (1973). Repetitive sequences in isolated Thomas circles from Drosophila melanogaster. Cold Spring Harbor Symp. Quant. Biol. *38*, 371–381.

Sharp, P. A., Cohen, S. N., and Davidson, N. (1973). Electron

microscope heteroduplex studies of sequence relations among plasmids of Escherichia coli. J. Mol. Biol. *75*, 235–255.

Shine, J., and Dalgarno, L. (1973). Occurrence of heat-dissociable ribosomal RNA in insects: the presence of three polynucleotide chains in 26S RNA from cultured Aedes aegypti cells. J. Mol. Biol. *75*, 57–72.

Smith, M. O., and Wilcox, K. W. (1970). A restriction enzyme from Hemophilus influenzae: I. Purification and general properties. J. Mol. Biol. *51*, 379–391.

Spear, B. B. (1974). The genes for ribosomal RNA in diploid and polytene chromosomes of Drosophila melanogaster. Chromosoma *48*, 159–179.

Streeck, R. E., Philippsen, P., and Zachau, H. G. (1974). Cleavage of small bacteriophage and plasmid DNAs by restriction endonucleases. Eur. J. Biochem. *45*, 489–499.

Tartof, K. D. (1973). Regulation of ribosomal RNA gene multiplicity in Drosophila melanogaster. Genetics *73*, 57–71.

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.

White, R. L., and Hogness, D. S. (1977). R loop mapping of the 18S and 28S sequences in the long and short repeating units of Drosophila melanogaster rDNA. Cell *10*, 177–192.