Detection and Resolution of Closely Related Satellite DNA Sequences by Molecular Cloning

Kirk Fry† and Douglas Brutlag

Department of Biochemistry
Stanford University Medical Center
Stanford, Calif. 94305, U.S.A.

(Received 28 December 1978, and in revised form 21 August 1979)

Highly repeated satellite DNAs often consist of mixtures of DNAs with closely related repeating sequences. By cloning individual molecules we have resolved the 1.705 g/cm³ satellite DNA of Drosophila melanogaster into two distinct components: poly[d(A-G-A-G-T-T-C-C)] and poly[d(A-G-A-G-A-G-T-T-C-C)]. The presence of two distinct sequences within this physically homogeneous satellite DNA had not previously been detected by standard physical, chemical, or sequencing techniques. Both cloning and direct sequence analysis suggest that the five-base-pair and seven-base-pair repeating units reside on separate molecules and are not interspersed with each other.

1. Introduction

The organization and function of Drosophila heterochromatin has been studied extensively at the genetic and cytogenetic level. Analyses of mitotic chromosomes and comparison with polytene interphase chromosomes have shown that the pericentric heterochromatin constitutes 25% of the genome and is distributed among all three autosomal pairs as well as the sex chromosomes (Rudkin, 1969; Peacock et al., 1977). Genetic analysis of heterochromatic deletions indicates that these regions are essential for germ-line functions and not for somatic processes. Male Drosophila carrying deletions of X heterochromatin, for example, are viable but have reduced meiotic pairing of the X and Y chromosomes, abnormal chromosome segregation, defective spermatogenesis and reduced fertility (Gershenson, 1933; Sandler & Braver, 1954; Cooper, 1964; Peacock et al., 1975).

At the molecular level, the bulk of the heterochromatic DNA can be isolated as four satellite DNAs in CsCl gradients (Peacock et al., 1973; Endow et al., 1975; Brutlag et al., 1977a). These satellites consist of short nucleotide sequences, 5 to 359 base-pairs in length which are tandemly repeated in arrays over 275 x 10³ bases long (Goldring et al., 1975). Each chromosome has a unique arrangement of satellite DNA around its centromere (Peacock et al., 1977). For example, three large blocks of the 1.705 g/cm³ DNA (1.705 g/cm³ satellite) have been resolved on the Y chromosome, two blocks on the second chromosome, and much smaller amounts are distributed among the other chromosomes.

† Present address: Department of Radiology, Stanford Medical Center, Stanford, Calif. 94305, U.S.A.

0022-2836/79/350581-13 $02.00/0 © 1979 Academic Press Inc. (London) Ltd.
Despite the physical homogeneity of most of the simple sequence satellites, many of them are mixtures of DNA species (Skinner et al., 1970; Hatch & Mazrimas, 1974). The 1-672 g/cm³ satellite of Drosophila melanogaster for example, contains both poly
\[
\text{d}(\text{A-A-T-A-T-A})
\]
and poly
\[
\text{d}(\text{T-T-A-T-A})
\]
which are so similar as to preclude separation by the classical physical techniques (Brutlag & Peacock, 1979). We therefore developed methods for cloning individual molecules of satellite DNA in order to separate these sequences and to resolve “cryptic” satellites (Brutlag et al., 1977b). We have now applied this technique to the 1-705 g/cm³ satellite DNA of D. melanogaster, which is homogeneous by a variety of criteria including sedimentation in both neutral and alkaline CsCl gradients, by denaturation of renatured 1-705 g/cm³ DNA within 1 deg. C of the native satellite, and by isolation of pyrimidine tracts of average length 750 bases (Endow et al., 1975; Birnboim & Sederoff, 1975; Sederoff et al., 1975; Brutlag et al., 1977a). This satellite contains the repeating sequence poly
\[
\text{d}(\text{A-A-G-A-G})
\]
as determined by analysis of cRNA from both satellite DNA and from pyrimidine tracts. This sequence has been verified by digestion with the restriction endonuclease MboII, which recognizes the sequence \(\text{G-A-A-G-A}\)
\(\text{C-T-T-C-T}\) and cleaves the satellite into fragments 5, 10 and 15 base-pairs long (Endow, 1977).

Some sequence heterogeneity was suggested by the presence of small amounts of free GMP and 5’ A-U-A-G 3’ in T₄ RNase digests of 1-705 g/cm³ complementary RNA (Endow et al., 1975) and by the presence of the sequence 5’ A-A-G-A-A-G-A-G 3’ detected among partial T₄ digests (Brutlag & Peacock, 1975). Such sequence heterogeneity could result from either variations within a poly
\[
\text{d}(\text{A-A-G-A-G})
\]
region or by mixtures of closely related repeated sequences each with a density of 1-705 g/cm³.

In this work, we show by sequence analysis of clones containing the 1-705 g/cm³ DNA in hybrid plasmids, that this satellite actually contains two distinct repeated sequences. The major species is poly
\[
\text{d}(\text{A-A-G-A-G})
\] while 15% of the satellite is composed of the sequence poly
\[
\text{d}(\text{A-A-G-A-G})
\]. We conclude that the major source of sequence variation in this satellite, as in the 1-672 g/cm³ satellite, results from a mixture of distinct DNA species. The similarity between the two repeat units found in each of these satellites extends the observation originally made in D. virilis, that simple sequence satellites within one species contain related sequences (Gall & Atherton, 1974). Resolution of these similar sequences by cloning will now allow localization of each independently by hybridization in situ under stringent criteria.

2. Materials and Methods

Satellite DNA was purified by the procedures previously described (Brutlag et al., 1977a). The 1-705 g/cm³ DNA from the actinomyein D/CS gradient was repurified by another CsCl gradient. Bacterial hybrid plasmids containing 1-705 g/cm³ satellite sequences were constructed and the DNA purified using published procedures (Brutlag et al., 1977b).
The MboII restriction endonuclease was purified from *Moraxella bovis* kindly supplied to us by R. Roberts who suggested the following outline for purification. Cell lysis was carried out using a French pressure cell followed by high-speed centrifugation, and nucleic acids were precipitated by streptomycin sulfate. The restriction endonucleases were then concentrated with ammonium sulfate followed by phosphocellulose, DEAE-cellulose and DNA-agarose column chromatography (Gelinas et al., 1977). If polyamine P was used to precipitate the nucleic acids (Bickle et al., 1977) the final enzyme preparation contained more 3' exonuclease as assayed by the method of England (1972). Commercial preparations of MboII were used when they became available (New England Biolabs), but these two had small amounts of 3' exonuclease (see Results).

*MboII* digestions were carried out at 37°C in 60 mM NaCl, 10 mM-Tris·HCl (pH 8-9), 1-0 mM-dithiothreitol, 7 mM MgCl₂ for periods up to 4 h. The extent of digestion was monitored by gel electrophoresis of parallel samples containing simian virus 40 DNA in addition to the satellite DNA.

*MboII* endonuclease digestion fragments in the digestion buffer were dephosphorylated after inactivating the *MboII* enzyme for 2 min at 100°C, by addition of bacterial alkaline phosphatase (4 units/ml) with incubation at 37°C for 30 min. The reaction mixture was then diluted with the appropriate salts and [γ-³²P]ATP (1000 to 1600 Ci/mM) to give final concentrations of 70 mM-Tris·HCl (pH 8-9), 20 mM-NaCl, 6.7 mM-K₂HPO₄/KH₂PO₄ (pH 8-9), 12.3 mM-MgCl₂, 16.0 mM-2-mercaptoethanol, 0.33 mM-dithiothreitol, and 1.6 µM-ATP. Polyribonucleotide kinase (Richardson, 1965) was added to 33 units/ml and the reaction incubated at 37°C for 30 min. This method had the advantage that no immediate purifications had to be used since bacterial alkaline phosphatase is strongly inhibited by the phosphate.

The *MboII* digestion fragments were resolved by electrophoresis of the complete reaction mixture in a 20% polyacrylamide gel at 15 V/cm (Maxam & Gilbert, 1977). In some cases, in order to lower the background of ATP and ³²P in the gel, the oligonucleotides were purified from the kinase reaction mixture by gel filtration on a 0.7 cm x 10 cm Biogel P4 column (+400 mesh, Bio-Rad Lab.) equilibrated in 10 mM-triethylammonium bicarbonate (pH 9-5). All fractions containing radioactivity eluting before the ATP were combined and concentrated in a vacuum desiccator before electrophoresis. Marker compounds (p(dA-dG)₂, p(dA-dG)₃, p(dA-dG)₄ and p(dA-dG)₅) purchased from Collaborative Research were labeled as above and used as size markers on the gels.

Autoradiography was carried out by wrapping the gel in clear plastic wrap, and freezing the gel at −70°C during exposure to prevent diffusion of the small oligonucleotides. After developing the film, the appropriate radioactive oligonucleotides were located by aligning the X-ray film over the gel using radioactive ink marks and cutting out the gel slice containing the oligonucleotide.

Oligonucleotides were eluted by placing the gel slice in 4 ml of 0.01 M-Tris·HCl (pH 7-5), 0.001 M-EDTA overnight or longer. Elution from the slice was monitored by counting the eluted radioactivity by Cerenkov radiation. Removal of salts and urea, and concentration of the oligonucleotides were accomplished by binding the oligonucleotides eluted from the gel slices to a 5 mm x 5 mm DEAE-cellulose column. This was followed by a distilled water wash and elution of the oligonucleotides with 0.4 ml of 2.2 M-triethylammonium bicarbonate (pH 9-5). The triethylammonium bicarbonate was evaporated under vacuum and the oligonucleotides redissolved in 0.05 ml of water. Homochromatography in one dimension was then used to further purify the oligonucleotides and separate complementary sequences. After elution of the pure oligonucleotides from the homochromatography plates by standard methods (Barrell, 1971), partial digestion using venom phosphodiesterase was carried out. Samples eluted from the homochromatography plates were split into 6 to 9 fractions and dried on plastic sheets. Three samples were dissolved in 2 µl each containing, respectively, 0.5 mg/ml, 0.1 mg/ml or 0.05 mg/ml venom phosphodiesterase (Worthington) in 0.05 M-Tris·HCl (pH 8-9), 0.01 M-MgCl₂. Incubation was carried out in glass capillary tubes at 37°C for times ranging from 15 min for oligonucleotides 5 to 7 nucleotides long to 2 h for those 12 to 20 nucleotides long. The 3 samples were then combined and partial products were analyzed by standard 2-dimensional methods using electrophoresis at pH 3-5 and homochromatography (Brownlee & Sanger, 1969;
3. Results

S. Endow (1977) has shown that the restriction nuclease MboII is a powerful tool for analyzing the 1.705 g/cm³ satellite of D. melanogaster because its recognition sequence (5' G-A-A-G-A 3') is the same as the basic repeating unit of the satellite. Since MboII cleaves DNA eight nucleotides to the right of its recognition site as shown in Figure 1, one would expect a series of oligonucleotides to result from MboII digestion of the 1.705 g/cm³ satellite which would be multiples of five base-pairs. Further, one would expect that limit digests of poly d(G-A-A-G-A
C-T-C-T-C-T) would contain fragments 5, 10 and 15 base-pairs in length all beginning with 5' G-A-G . . . and terminating with . . . G-A-A 3' (Brutlag et al., 1978). Longer oligonucleotides should only exist as intermediates in the digestion. Endow (1977) determined the repeating unit of the 1.705 g/cm³ satellite by sequence analysis of these major MboII digestion products.

We used MboII to analyze various hybrid plasmids containing segments of the 1.705 g/cm³ satellite (Brutlag et al., 1977b) and to compare individual regions within the total satellite DNA. Figure 2 shows MboII digestion products from satellite DNA and 12 independently isolated hybrid plasmids. Only oligonucleotides 5 to 100 base-pairs in length are resolved on the 20% polyacrylamide gel, so that the larger MboII fragments from the vector DNA remain near the top of the gel. The length of the major fragments from the 1.705 g/cm³ satellite DNA are integral multiples of five base-pairs. In order to prove the mechanism shown in Figure 1, we determined the nucleotide sequence of both the purine and the pyrimidine strands of each oligo-

![Diagram](image)

**Fig. 1.** Generation of the 5, 10 and 15-base-pair DNA fragments by digestion of poly d(A-A-G-A-G
T-T-C-T-C) with MboII restriction nuclease.

The MboII recognition site is shown in the box and the cleavage site, which is 10 base-pairs from the center of the recognition site (1 turn of the DNA helix), is shown by the broken arrows. Cleavage of poly d(A-A-G-A-G
T-T-C-T-C) by this mechanism releases oligonucleotides which are integral multiples of 5 base-pairs. Fragments of 15 base-pairs or shorter are limit products since they no longer contain both a recognition site and a cleavage site. Hyphens omitted for clarity.
nucleotide. Each oligonucleotide was eluted from the gel, resolved into the two complementary strands by homochromatography, and the sequence determined by partial exonuclease digestion from the 3' end. Figures 3, 4 and 5 show twodimensional chromatograms of the partial exonuclease digests of each strand of the 5, 10 and 15 base-pair oligonucleotides, respectively. The sequences of the complementary strands of each fragment can base-pair except for the 3'-terminal T of the pyrimidine strand and the 3'-terminal A in the purine strand. These results confirm

**Fig. 2.** DNA fragments produced by cleavage of 1.705 g/cm³ satellite and plasmid carrying cloned segments of the satellite by MboII endonuclease.

This Figure is an autoradiogram of a 20% polyacrylamide gel showing DNA fragments produced by MboII digestion of the satellite and 12 recombinant plasmids containing 1.705 g/cm³ sequences as described in Materials and Methods. Polynucleotide kinase was used to label the 5' ends of the fragments with ³²P. The satellite in the pDm705 plasmids was cloned in the R1 site of pSC101 as described (Brutlag et al., 1977a). The satellite in the plasmid crDm705-5 was cloned in the single R1 site present in the vector pCR1 by the same method (Covey et al., 1976). The numbers beside the gel indicate the lengths of the oligonucleotides (in base-pairs) as determined by external standards and by direct sequence analysis of the fragments.
Fig. 3. Sequence analysis of the pentanucleotide fragment from MboII digestion of 1.705 g/cm³ satellite DNA. (a) Chromatogram of 2-dimensional fractionation of partial digests generated by snake venom phosphodiesterase on the pentanucleotide fragments from the purine strand. The mobility changes in the 1st and 2nd dimensions are characteristic of the nucleotides lost from the 3’ end as shown by the letter next to the arrow (Galbert et al., 1974). The sequence read from the top and proceeding from the 5’ nucleotide to 3’ nucleotide would be 5’ pG-A-G-A-A 3’. (b) Same analysis as in (a) except the pyrimidine strand was analyzed. The sequence determined is 5’ pT-C-T-C-T 3’.

the mechanism shown in Figure 1, and demonstrate that MboII makes a staggered cut eight nucleotides to the 3’ side of the recognition site 5’ G-A-G-A 3’ but only seven nucleotides away on the complementary strand. Because of the protruding 3’ A or T residue, these fragments can only be arranged next to each other with the purines on one strand and the pyrimidine on the other. Any arrangements of these three major fragments results exclusively in the repeating sequence poly-
\[
\begin{bmatrix}
T-T-C-T-C
\end{bmatrix}
\] This confirms the earlier conclusions of Endow et al. (1975), Sederoff et al. (1975) and Endow (1977).

The origin of the major oligonucleotide 20 base-pairs in length is obscure. If such a fragment resulted from digestion of poly-
\[
\begin{bmatrix}
T-T-C-T-C
\end{bmatrix}
\] it should have been cleaved to 5 and 15-base-pair fragments as shown in Figure 1. The 20-base-pair fragments could have resulted from a region of satellite containing an altered sequence and hence missing one MboII recognition site or from partial digestion. Control digestions containing simian virus 40 DNA as an internal marker indicated that digestion was complete. However, the overlapping nature of the recognition and cleavage sites in the satellite DNA might make its digestion more difficult. Sequence
Fig. 4. Sequence analysis of the decanucleotide fragments from MboII digestion of 1.705 g/cm³ satellite DNA (a) Sequence analysis of purine strand as described in Fig. 3. Sequence determined is 5’ pG-A-G-A-G-A-G-A-A 3’. (b) Sequence analysis of pyrimidine strand. The sequence is 5’ pT.C.T.C.T.C-T.C-T.C-T 3’.

Variations in the MboII recognition site at the 5’ end of a DNA molecule would invariably result in a 20-base-pair fragment. However, sequence variations internal or near the 3’ end of a satellite molecule could result in either a 5, 10, 15 or 20-base-pair product according to its location and the mechanism of Figure 1. While we have not examined sequence variation within the 20-base-pair fragment directly, minor sequence variations were detected in both the 5 and 10-base-pair fragments (Figs 3 and 4). These minor sequences could have arisen either by sequence variation of the five-base-pair repeat or by exonuclease digestion of longer MboII fragments resulting in oligonucleotides of unusual sequence. We have noted low levels of exonuclease in all preparations of MboII restriction nuclease.

While the major MboII digestion products of the 1.705 g/cm³ satellite are readily explained as originating from poly [d(A-A-G-A-G)] or from single base alterations of this sequence, the origin of the 7 and 14-base-pair fragments is not obvious. Endow (1977) noted the 14-base-pair fragment and suggested that it may have resulted from exonucleolytic digestion of the 15-base-pair fragment. Even limited exonuclease action during the MboII digestion can result in shortened fragments (see Fig. 2, pDm705.71). All of the hybrid plasmids we examined except one contained satellite regions which gave rise to oligonucleotides that were multiples of five base-pairs after MboII digestion (Fig. 2, pDm705.28 through pDm705.72). The exceptional plasmid, crDm705.5, contained the fragments 7, 14 and 21 base-pairs in length and

Fig. 6. Sequence analysis of the heptanucleotide fragment obtained by MboII digestion of crDm705-5 plasmid DNA. (a) The sequencing method described in Fig. 3 was used to determine the sequence 5' pA-G-A-G-A-G-A-G 3' for the purine strand. (b) The sequence of pyrimidine strand is 5' pT-C-T-C-T-C-T-C 3'.
the five-base-pair series was missing entirely. These data indicate that fragments which are multiples of seven base-pairs are derived from a novel sequence which is segregated in regions separate from the poly[d(A-A-G-A-G)/T-T-C-T-C]] sequence within the Drosophila genome.

In order to determine the repeating unit of this DNA we analyzed the sequence of the 7 and 14-base-pair fragments from both the satellite DNA and from the cloned plasmid crDm705.5. The sequences of the cloned fragments are shown in Figures 6 and 7. The seven-base-pair oligonucleotide consists of 5' A-G-A-A-G-A-G 3' and 5'T-C-T-C-T-C-T-C3'. These sequences are complementary except for the terminal dGMP or dCMP. When the 7 and 14-base-pair fragments are aligned adjacent to each other, the repeating sequence poly[d(A-A-G-A-G-A-G)/T-T-C-T-C-T-C]] is obtained (Fig. 8). The analysis of the 7 and 14-base-pair oligonucleotides isolated from satellite DNA gave the same major sequences as shown in Figures 6 and 7; however, minor sequences not present within the cloned segments were detected (Brutlag et al., 1978). As described above, such minor components could represent sequence variation present within the total satellite DNA and absent in the cloned region, or they could be due to other distinct DNA species within the 1.705 g/cm³ satellite fraction.

The relative amounts of poly[d(A-A-G-A-G)/T-T-C-T-C]] and poly[d(A-A-G-A-G-A-G)/T-T-C-T-C-T-C]] were determined by quantitative alkaline hydrolysis using the alkali-stable DNA at pH 3.5 as a standard (see Figure 8).

**Fig. 7.** Sequence analysis of the 14-nucleotide fragment obtained by MboII digestion of crDm705-5 plasmid DNA. (a) The sequence of the purine strand is 5' pA-G-A-A-G-A-G-A-A-G-A-G 3'. (b) The sequence of the pyrimidine strand is 5' pT-C-T-C-T-C-T-C-T-C-T-C 3'.

---

**Fig. 8.** Electrophoresis of poly[d(A-A-G-A-G)/T-T-C-T-C]] at pH 3.5. The left lane shows the high-molecular-weight satellite DNA, while the right lane shows the cloned segments of the satellite DNA. The gel is stained with ethidium bromide.
Fig. 8. Fragments produced by MboII digestion of the 7-base-pair repeated sequence present in the 1.705 g/cm^3 satellite DNA. When MboII cleaves the 7-nucleotide repeat by the mechanism described in Fig. 1, all the products are integral multiples of 7 base-pairs. Fragments 21 base-pairs or longer contain both a recognition site and a cleavage site and are subject to further digestion to the limit products 7 and 14 nucleotides in length. Hyphens omitted for clarity.

present in the satellite DNA were estimated by determining the amounts of radioactivity present in the five-base-pair series and in the seven-base-pair series of fragments. After correcting for the relative length of each fragment and assuming equal labeling efficiency of both series, we estimate that 10 to 15% of the satellite consists of poly[d(A-A-G-A-G-A-G)T.T.C.T.C.T.C]. The error in this estimate arises from a background of [32P]ATP in the gels (for which corrections were made) and from the small amount of exonuclease which can reduce 15-base-pair fragments to 14-base-pair fragments.

4. Discussion

Highly repeated satellite DNAs often contain several distinct DNA species. The presence of multiple species can be detected by the resolution of more than two strands in alkaline CsCl gradients (Gall et al., 1971), by further resolution of the satellite in cesium gradients containing base-specific binding agents (Skinner & Beattie, 1973), or by direct sequence analysis (see preceding paper, Brutlag & Peacock, 1979). The 1.705 g/cm^3 satellite DNA from D. melanogaster seemed to be homogeneous by all of these as well as by many other physical criteria. By cloning individual molecules, however, we have shown that this satellite actually contains two DNA species, poly[d(A-A-G-A-G)T.T.C.T.C] and poly[d(A-A-G-A-G)T.T.C.T.C]. It is clear that such closely related DNA species are readily resolved only by molecular cloning.

Most of the sequence variation present within the 1.705 g/cm^3 satellite can be accounted for by the mixture of poly[d(A-A-G-A-G)T.T.C.T.C] and poly[d(A-A-G-A-G)T.T.C.T.C] or by single-base changes of these sequences. Single base substitutions in either the five or seven-base-pair repeat would give rise to 20 and 21-base-pair MboII digestion products. These fragments could also be a result of partial digestion with MboII and a complete sequence analysis of these fragments is required to determine whether they contain variant sequences. The two other sequences previously observed within
the 1.705 g/cm³ satellite (5′ A-U-A-G 3′; Endow et al., 1975; and 5′ A-A-G-A-A-G-A-G-A-G 3′, Brutlag & Peacock, 1975) have not been detected in either our restriction nuclease analysis nor in our sequence analysis of cloned segments. These sequences could have arisen either from sequence variation within the two DNA species described here or from other DNAs present within the 1.705 g/cm³ satellite. We have estimated that the seven-base-pair repeating sequence constitutes about 10 to 15% of the 1.705 g/cm³ satellite DNA. While only one hybrid plasmid out of 16 examined contained the seven-base-pair sequence, we initially selected our plasmids by a colony hybridization procedure which would favor the recovery of the majority type (Grunstein & Hogness, 1975; Brutlag et al., 1977b).

The chromosomal regions in the Drosophila genome containing the 1.705 g/cm³ satellite have been estimated to be 275 to 1000 × 10⁹ base-pairs in length (Goldring et al., 1975; Brutlag et al., 1977a). The sizes of the regions containing the individual sequences are difficult to determine. The longest regions present in our hybrid plasmids are only 1000 base-pairs. For example, the plasmid crDm705-5 contains about 140 copies of the seven-base-pair repeat. Since regions of simple-sequence satellite DNA longer than this are not stable in bacterial hosts it will be difficult to determine if regions longer than 10⁹ base-pairs contain only a single sequence by cloning. Thus it is still possible that chromosomal regions 10⁶ base-pairs in length could contain both sequences.

The similarity of the two 1.705 g/cm³ sequences to each other and to the two sequences present in the 1.672 g/cm³ satellite of D. melanogaster(\[\text{poly} \left\{ \begin{array}{c} A-A-T-A-T \\ T-T-A-T-A-T \end{array} \right\} \]) and \[\text{poly} \left\{ \begin{array}{c} A-A-T-A-T-A \\ T-T-A-T-A-T \end{array} \right\} \] extend the interesting observation made by Gall & Atherton (1974). They showed that the three major simple-sequence satellite DNAs from another species, D. virilis, were each composed of heptamer repeating units. The three repeating units differed from each other by one or two base changes. The 1.672 and the 1.705 g/cm³ satellites of D. melanogaster both contain five and seven-base-pair repeat sequences which are related to each other by two or three T-A→G-C transversions. The relationship of these sequences to the 1.686 g/cm³ satellite sequence (\[\text{poly} \left\{ \begin{array}{c} A-A-T-A-A-C-A-T-G \\ T-T-A-T-G-T-A-T-C \end{array} \right\} \]) is not as simple because of the differences in the lengths of the repeat unit. All of these repeats can be represented in the form proposed by Endow et al. (1975) of (A-A-N)ᵯ (A-M)ᵯ, where N and M represent T, C or G, and i and j are small integers. This pattern distinguishes the D. melanogaster satellites from the major satellites of D. virilis, all of which contain three consecutive adenine bases (5′ A-C-A-A-A-C-T 3′; 5′ A-T-A-A-A-C-T 3′; 5′ A-C-A-A-A-T-T 3′). Recently, however, Mullins & Blumenfeld (1979) have determined the sequence of a fourth minor satellite from D. virilis (5′ A-A-T-A-A-T-G 3′) which seems more closely related to the sequences of D. melanogaster than to those of D. virilis. They propose that this new sequence may be an evolutionary intermediate between D. virilis and D. melanogaster satellites.

The strong similarity between the major satellite DNAs in any one species might be required for heterochromatin function. It has already been shown that each chromosome of D. melanogaster has a unique distribution of these simple-sequence satellites around its centromere (Peacock & Steffensen, 1975; Peacock et al., 1977).
This suggests a role for these DNAs in chromosome-specific recognition processes. It is interesting that the 1·705 and the 1·672 g/cm³ satellites, which have similar DNA sequences, also have similar chromosome distributions, both being located primarily on the Y chromosome and on chromosome 2. With cloned segments of each of the satellite sequences it is now possible to determine by hybridization in situ whether one particular sequence has a more limited chromosomal distribution and whether the five and seven-base-pair repeats of each satellite also have a parallel chromosome distribution. The five and seven-base-pair sequences should not cross-hybridize since, when aligned, they show 17 A-C or G-T mismatches every 35 base-pairs.

We thank Dr Sharyn Endow for communicating her results prior to publication. This work was supported by a Basil O'Connor Starter Grant from the National Foundation March of Dimes. One of us (K. F.) was a postdoctoral fellow of the National Institute of General Medical Sciences.

REFERENCES


