Volume 10 Number 2 1982 Nucleic Acids Research

Simple repeated sequences in human satellite DNA

M.Frommer, J.Prosser, D.Tkachuk, A.H.Reisner* and P.C.Vincent

Kanematsu Memorial Institute, Sydney Hospital, Macquarie Street, Sydney, N.S.W. 2000, and *CSIRO, Molecular and Cellular Biology Unit, P.O. Box 184, North Ryde, N.S.W. 2113, Australia

Received 4 November 1981; Accepted 4 December 1981

ABSTRACT

In an extensive analysis, using a range of restriction endonucleases, HinfI and TaqI were found to differentiate satellites I, II and III & IV. Satellite I is resistant to digestion by TaqI, but is cleaved by HinfI to yield three major fragments of approximate size 770, 850 and 950bp, associated in a single length of DNA. The 770bp fragment contains recognition sites for a number of other enzymes, whereas the 850 and 950bp fragments are "silent" by restriction enzyme analysis. Satellite II is digested by HinfI into a large number of very small (10-80bp) fragments, many of which also contain TaqI sites. A proportion of the HinfI sites in satellite II have the sequence 5'GA($^{G}_{C}$)TC. The HinfI digestion products of satellites III and IV form a complete ladder, stretching from 15bp or less to more than 250bp, with adjacent multimers separated by an increment of 5bp. The ladder fragments do not contain TaqI sites and all HinfI sites have the sequence 5'GA($^{A}_{T}$)TC. Three fragments from the HinfI ladder of satellite III have been sequenced, and all consist of a tandemly repeated 5bp sequence, 5'TTCCA, with a non-repeated, G+C rich sequence, 9bp in length, at the 3' end.

INTRODUCTION

Human highly repeated DNA can, in part, be isolated as four cryptic satellites, each of which is separated from main band DNA by isopycnic centrifugation in an appropriate density gradient (1-4). Satellites I, II, III and IV are A+T rich compared to main band DNA (4) and constitute about 5% of the human genome (5).

Satellites II, III and IV have been characterised in terms of the products of digestion with restriction endonucleases EcoRI and HaeIII, whereas satellite I is resistant to digestion by these two enzymes. Several fractions of satellites II, III and IV have been identified by digestion with EcoRI and HaeIII (6). These include: (i) a "ladder" of fragments which are exact multiples in length of a 170 base pair (bp) monomer, (ii) various fragments which do not bear any obvious size relationship to the ladder fragments, (iii) a 3400bp male specific fragment (7, 8), and (iv) sequences

0305-1048/82/1002-0547 $2.00/0

that are resistant to digestion by either EcoRI or HaeIII or both enzymes. Satellites III and IV have been found to be identical by a range of criteria, including restriction enzyme analyses. Satellite II is similar to satellites III and IV in the size classes of products of digestion with EcoRI and HaeIII, but has been distinguished from satellite III by the observation that different sequence variants of the 170bp ladder material are incorporated into satellite II.

The repeated sequences which make up the 170bp ladder in satellites II and III, particularly the 340bp and 680bp components, are present in large quantities in unfractionated human DNA (9). A 340bp fragment and part of a 680bp fragment, from an EcoRI digest of total DNA, have been directly isolated and sequenced, to provide consensus sequences for a portion of the 170bp ladder material (10). This work has yielded no evidence for any internal repeat within the 170bp monomer.

There is evidence that shorter repeated sequences are also present in the human DNA satellites. The sequence determined for part of a single cloned 1770bp fragment, isolated from an EcoRI digest of satellite III, shows an obvious, though irregular, 4-9bp internal repeat and numerous HinfI and TaqI sites (11). Evidence for relatively simple repeated sequences in satellites I and II has been obtained from DNA fingerprinting studies (12).

We report here that preparations of satellites II, III and IV may vary considerably in their content of the EcoRI or HaeIII 170bp ladder components, but that all preparations of satellites III and IV contain, as a major component, a conserved repeated sequence, 5bp in length. This sequence appears as a ladder, with a 5bp periodicity, in HinfI digests of satellites III and IV. Satellite II, like satellite III, consists mainly of sequences which are identified by the presence of frequent HinfI sites, but the size distribution of fragments obtained by digestion with HinfI is entirely different. In satellite II, unlike satellite III, sequences which contain HinfI sites also contain numerous TaqI sites. Satellite I is characterised by a relative lack of HinfI sites. This satellite contains a sequence which consists of DNA resistant to restriction enzyme digestion, associated with a sequence containing clustered recognition sites for a number of enzymes.

METHODS

DNA extraction and preparation of satellite DNAs

DNA was extracted by the method of Marmur (13), from male placentae and

from preparations of leukaemic cells obtained from male patients (with acute myeloblastic or chronic granulocytic leukaemia) who were undergoing leucapheresis for therapeutic purposes. Isolation of human satellites I, II, III and IV, using Ag^{+} or $Hg^{++}/Cs_2\ SO_4$ density gradients, was carried out as previously described (5).

Restriction endonuclease digestion

EcoRII endonuclease was a gift of Dr. Keith Brown. All other restriction endonucleases were purchased from New England Biolabs. Digestions were carried out under the following conditions:-
AvaI, AvaII, BglII, DdeI, HgaI, HindIII, HinfI, MboI, Sau3AI, Sau96I: 6mM Tris-HCl, pH 7.5, 50mM NaCl, 6mM $MgCl_2$, 6mM β-mercaptoethanol, 0.1mg/ml bovine serum albumin (BSA).
BamHI, SalI, XbaI, XhoI: 6mM Tris-HCl, pH 7.9, 150mM NaCl, 6mM $MgCl_2$, 0.1mg/ml BSA.
Fnu4HI, HaeIII, TaqI: 6mM Tris-HCl, pH 7.4, 6mM NaCl, 6mM $MgCl_2$, 6mM β-mercaptoethanol, 0.1mg/ml BSA.
EcoRII, MspI: 10mM Tris-HCl, pH 7.5, 6mM KCl, 10mM $MgCl_2$, 6mM β-mercaptoethanol, 0.1mg/ml BSA.
EcoRI: 100mM Tris-HCl, pH 7.5, 50mM NaCl, 5mM $MgCl_2$, 0.1mg/ml BSA.
Digestions were carried out for 2 hours at 37°, except for TaqI digestions, which were carried out at 65°. Reaction mixtures, of total volume 10µl, contained 0.05-3µg DNA and sufficient enzyme to ensure complete digestion within two hours. Reactions were terminated by heating to 65° for 5 min, followed by rapid cooling to 0°.

Labelling of restriction fragments and gel electrophoresis

With the exception of HaeIII, all the restriction enzymes used in this study cleaved double stranded DNA, to produce single stranded 5' ends which were then filled in by the action of reverse transcriptase, using one or two appropriate ^{32}P-nucleoside triphosphates (14). When digests contained a large number of small fragments, the incubation with labelled nucleoside triphosphate was followed by incubation with an excess (0.1-0.2mM) of all four unlabelled nucleoside triphosphates. The reaction was terminated by heating to 65°, followed by rapid cooling to 0°.

Restriction fragments were separated by electrophoresis in vertical gels of 4% or 12% polyacrylamide (0.5mm thick) or 1.5% agarose (0.8mm or 3mm thick). ^{32}P-labelled fragments were visualised by autoradiography. HaeIII digestion products were stained with ethidium bromide and visualised under ultraviolet light.

Various restriction endonuclease digests of the plasmid PBR322 were used as standards (15). When calculating sizes of small, end-labelled fragments, it was necessary to take into account differences between restriction enzymes in the length of the 5'→ 3' staggered cut and thus in the number of nucleoside triphosphate units added by reverse transcriptase.

DNA sequencing

For sequencing, HinfI fragments of satellite III, containing [α-^{32}P]ATP at the 3' end of each strand, were eluted from 12% polyacrylamide gels. The fragments were denatured, the strands were separated by electrophoresis in 5% polyacrylamide gels at low ionic strength, and the single stranded DNA was sequenced according to Maxam and Gilbert (16). In each case, the more slowly moving strand was eluted from the denaturing gel and sequenced.

Southern transfers and filter hybridisation

DNA fragments were transferred from 1.5% agarose gels to nitrocellulose filters by the method of Southern (17). Hybridisation was carried out using a DNA probe labelled with [α-^{32}P]ATP by nick translation (18).

RESULTS

Satellites I, II and III are extensively cleaved by HinfI

Satellites I, II and III can be effectively differentiated by comparison of the products of digestion with HinfI (Figure 1). Satellite I is broken down into a number of high molecular weight fragments (Figure 1a). The gel shows three prominent bands of approximate size 770bp, 850bp and 950bp, a small amount of undigested material at the origin and a large band at about 3500bp, which, on further analysis in 1.5% agarose gels, proved to be a smear of high molecular weight material. Satellites II and III are almost entirely digested by HinfI (Figure 1, b and c). There is very little undigested material, either at the origin or at the 3500bp position. The low molecular weight fragments in HinfI digests of satellites II and III have been further resolved by electrophoresis in 12% polyacrylamide gels (Figure 2). Satellite II is cleaved by HinfI into a large number of very small fragments (Figure 2b); the smallest fragments are less than 10bp in length, and there are very few fragments of size greater than 80bp. The HinfI digestion products of satellite III form a complete ladder stretching from 15bp or less to more than 250bp (Figure 2j). Adjacent components of the ladder are separated from each other by an increment of 5bp.

Although the HinfI digests shown in Figure 2 were obtained under

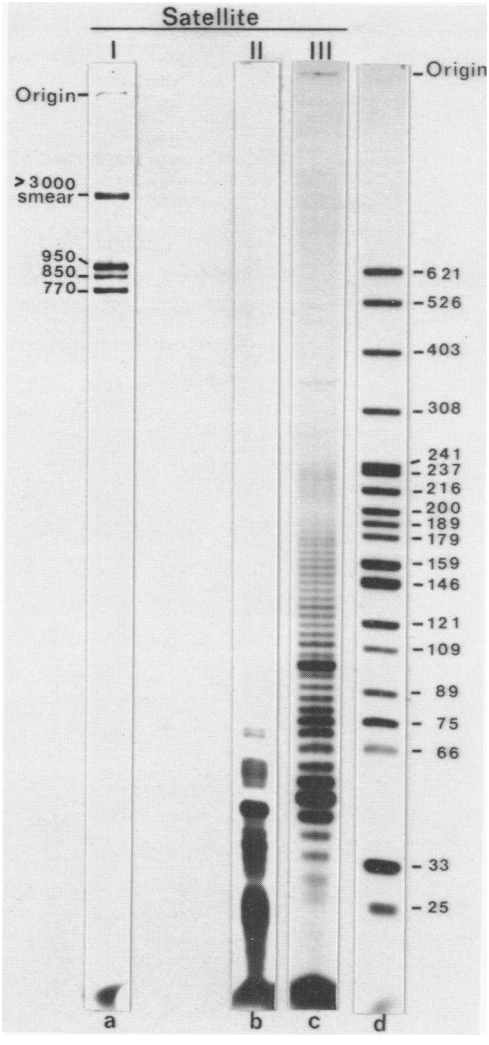

Figure 1. 4% polyacrylamide gels of [α-^{32}P]ATP-labelled fragments produced by digestion of satellites I, II and III with HinfI.
a: HinfI digested satellite I. Fragment sizes, in base pairs, are marked on the left side.
b: HinfI digested satellite II.
c: HinfI digested satellite III.
d: [α-^{32}P]CTP-labelled fragments from digestion of PBR322 with MspI. Fragment sizes, in base pairs, are marked on the right side; these have been adjusted to make them comparable in size with end-labelled HinfI fragments, and give an indication of sizes of fragments in lanes b and c.

conditions which normally result in a complete reaction, the satellite III HinfI ladder has the appearance of a partial digest. However, when samples of both satellites II and III were incubated with increased concentrations of HinfI or for longer times, there was no change in the amount of any of the satellite III ladder components or the satellite II fragments. In addition, individual HinfI ladder components, eluted from a gel, were not further digested by reincubation with an excess of HinfI. We conclude, therefore, that the digestions shown in Figures 1 and 2 were complete, and that the

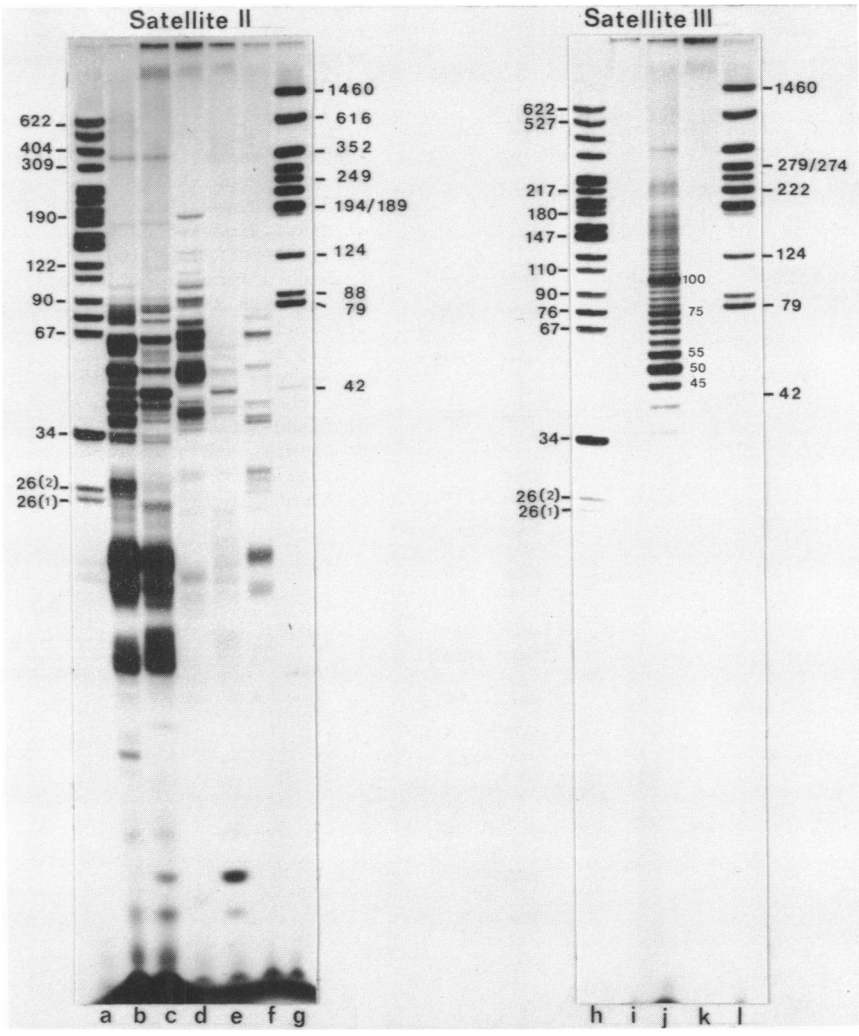

Figure 2. 12% polyacrylamide gels of HinfI and TaqI digestion products of satellites II and III. a and h: [α-^{32}P]CTP-labelled fragments from digestion of PBR322 with MspI. Fragment sizes, in base pairs, are marked. b: HinfI digest of satellite II, fragments labelled with [α-^{32}P]ATP. c: HinfI/TaqI double digest of satellite II, fragments labelled with [α-^{32}P]ATP and [α-^{32}P]CTP. d: TaqI digest of satellite II, fragments labelled with [α-^{32}P]CTP. e: HinfI/TaqI double digest of satellite II, fragments labelled with [α-^{32}P]CTP. f: HinfI digest of satellite II, fragments labelled with [α-^{32}P]CTP. g and l: [α-^{32}P]CTP-labelled fragments from digestion of PBR322 with Sau96I. i: HinfI digest of satellite III, fragments labelled with [α-^{32}P]CTP. j: HinfI digest of satellite III, fragments labelled with [α-^{32}P]ATP. k: TaqI digest of satellite III, fragments labelled with [α-^{32}P]CTP.

fragment sizes and concentrations accurately reflect the distribution of HinfI sites in satellites II and III.

HinfI, with recognition site 5'G$^{\downarrow}$ANTC, will cleave four possible sequences, and it was of interest to determine whether satellites II and III could be further differentiated by analysis of the sequence of HinfI sites. HinfI digests were incubated either with [α-^{32}P]ATP, which will label all fragments, or with [α-^{32}P]CTP, which will label only fragments produced by cleaving the sequences 5'GA($\frac{G}{C}$)TC. Satellite III contains virtually no HinfI sites with a central ($\frac{G}{C}$) pair (Figure 2i). Short term autoradiographs (15-60 minute exposure), as used in all other HinfI analyses, failed to reveal any bands in the gels of [α-^{32}P]CTP-labelled HinfI digests of satellite III. Exposure times of 15 hours or more were required to reveal the HinfI 5bp ladder in these gels, indicating that very few of the HinfI sites in all size classes of the ladder have the base composition 5'GA($\frac{G}{C}$)TC. By contrast, a considerable proportion of the HinfI sites in satellite II contain a central ($\frac{G}{C}$) pair (Figure 2f). Comparisons of band intensities, in autoradiographs with increasing exposure times, showed that gels of satellite II HinfI digests contain a few bands which are essentially unlabelled by [α-^{32}P]CTP, some bands which incorporate a large amount of the CTP label and a majority of bands which are labelled to an intermediate extent by [α-^{32}P]CTP.

TaqI digestion products of satellites II and III

DNA sequences which are digested into small fragments by HinfI might be expected also to contain recognition sites for TaqI(5'T$^{\downarrow}$CGA), since part of two adjacent HinfI sites will form a TaqI site. Products of TaqI digestion of satellites II and III are shown in Figure 2 (d and k). Satellite II is extensively cleaved by TaqI, to yield a large number of small fragments, the majority of which range from 35bp to 190bp. The double digests of satellite II with TaqI and HinfI (Figure 2, c and e) contain very few fragments of the same size as those in the TaqI single digest, indicating that most of the sequences which contain TaqI recognition sites also contain HinfI recognition sites. A number of new bands appear in the gels of HinfI/TaqI double digests, including a prominent band of approximately 5bp. Satellite III, on the other hand, is almost entirely resistant to digestion by TaqI (Figure 2k). In autoradiographs exposed for 15 hours, a faint 5bp ladder could be seen in the TaqI digest of satellite III, with fragments of the same size as those in the HinfI ladder. The amount of material cleaved by TaqI could not be increased by increasing the enzyme concentration or the digestion time. Thus, it appears that the resistance

of satellite III to digestion by TaqI indicates a real lack of TaqI sites in this DNA.

Comparison of satellite preparations by digestion with EcoRI, HaeIII and HinfI

Figure 3A shows an autoradiograph of a 1.5% agarose gel in which [α-^{32}P]ATP-labelled fragments from EcoRI digests of a number of satellite II, III and IV preparations have been separated. Although the 170bp ladder can be identified in all preparations, there is considerable variation between preparations in the concentrations of the 170bp ladder components.

Satellites III and IV: Seven preparations of satellites III and IV have been assayed. Three preparations of satellite III, digested with EcoRI, yield only faint bands at 340bp and 680bp (Figure 3A, c and i, one preparation is not illustrated). One preparation of satellite IV (Figure 3A, f) yields a complete and prominent 170bp ladder, with the monomer just visible, and the 2X, 3X, 4X, 5X, 6X, 7X and 8X multimers present in the gel, along with 2200bp and 3400bp fragments. One preparation of satellite IV and two preparations of satellite III fall into an intermediate class, where a number of the ladder components can be detected in low concentrations (Figure 3A, g and h; Figure 4).

A number of other experimental procedures were used in attempts to detect more of the 170bp ladder material in all satellite III and IV preparations: (i) Digestion of satellite preparations with EcoRI and HaeIII, electrophoresis in 4% polyacrylamide and 1.5% agarose gels, and staining with ethidium bromide. (ii) Southern transfer of fragments, from 1.5% agarose gels of EcoRI and HaeIII digests of satellites III and IV, to nitrocellulose strips, followed by hybridisation with an [α-^{32}P]ATP-labelled, nick translated satellite IV probe, known to contain large amounts of 170bp ladder components. (iii) Direct digestion of [α-^{32}P]ATP-labelled, nick translated DNA with EcoRI and HaeIII. None of these procedures yielded any evidence of greater concentrations of 170bp ladder components than those found in the original, end-labelled, EcoRI digests shown in Figure 3A. We conclude that different amounts of 170bp ladder material are contained in different preparations of the same satellite, sometimes from the same individual, isolated to the same standards of ultracentrifugal purity.

Figure 3B shows an autoradiograph of a 4% polyacrylamide gel, containing [α-^{32}P]ATP-labelled fragments from HinfI digests of the same five

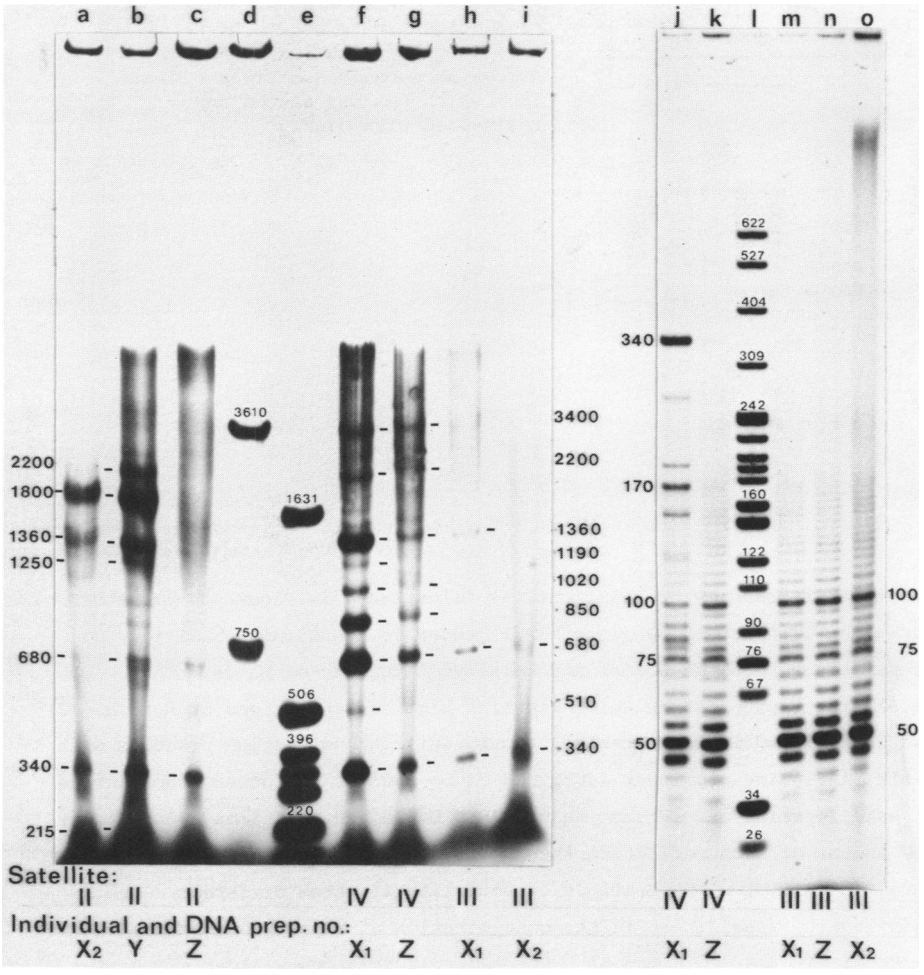

Figure 3. Variability between satellite preparations.
A: 1.5% agarose gel of EcoRI digestion products of satellites II, III and IV. B: 4% polyacrylamide gel of HinfI digestion products of satellites III and IV. The satellite type is noted below each lane. The individual, from whose DNA the satellite was prepared, is identified by a letter, X, Y or Z. Where two DNA preparations were made from one individual, the preparation number is identified by a subscript. Fragment sizes, in base pairs, are marked at either side of the autoradiographs. Lanes d, e and l are standard digests of PBR322 with EcoRI/PstI, HinfI and MspI, respectively. All digests were end-labelled with [α-^{32}P]ATP, except the MspI standard digest (lane l), which was labelled with [α-^{32}P]CTP.

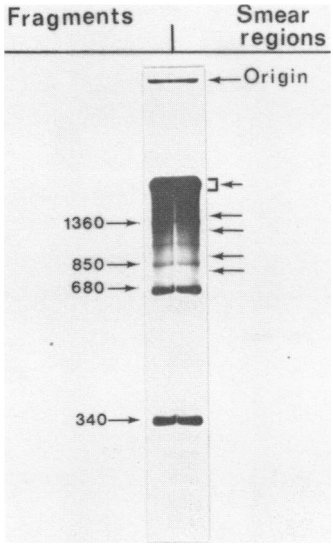

Figure 4. 4% polyacrylamide gel of EcoRI digestion products of satellite III, labelled with [α-^{32}P]ATP. Arrows show bands and smear regions from which DNA was eluted and redigested with HinfI.

satellite III and IV preparations, as illustrated in Figure 3A after digestion with EcoRI. The 5bp HinfI ladder is identical in all five preparations, both in number and relative intensity of bands. The satellite IV preparation which contains high concentrations of both HaeIII and EcoRI 170bp ladder material, shows additional digestion products with HinfI, including prominent 170bp and 340bp bands. The other preparations of satellite III and IV show faint 340bp bands and sometimes 170bp bands, the amount of each reflecting the amount of material digested by EcoRI.

Satellite II: Two preparations of satellite II, from different individuals, also differ in the concentration of repeated sequences cleaved by EcoRI (Figure 3A, a and b). The same preparations yielded virtually identical fragment patterns when digested with HinfI, with perhaps some small differences in the concentration of three or four larger HinfI bands (70-100bp).

Satellite I: Three preparations of satellite I from different individuals were all resistant to digestion by EcoRI, and showed identical patterns for the major products of HinfI digestion.

Fate of the HinfI 5bp ladder in an EcoRI digest of satellite III

EcoRI digestion products of satellite III, separated in a 4% polyacrylamide gel, are illustrated in Figure 4. In addition to the bands detected in 1.5% agarose gels, a concentrated smear of material at around the 3500bp position and a more diffuse smear of material behind all the

larger EcoRI bands can be seen. To determine the fate of the HinfI 5bp ladder in an EcoRI digest of satellite III, DNA from each band and smear region, as marked in Figure 4, was eluted and redigested with HinfI. The origin and 3500bp smear were almost entirely digested by HinfI, to yield a pattern of fragments indistinguishable from that seen in HinfI digests of total satellite III. The HinfI ladder was also detected in redigested smear regions from either side of the 1360bp and 850bp EcoRI bands. Table 1 lists the series of bands obtained by digestion of eluted EcoRI fragments with HinfI. The redigested 1360bp and 850bp EcoRI fragments contained a trace of the HinfI 5bp ladder, similar in concentration to that seen in adjacent smear regions. The relative intensities of smear regions and bands, in the EcoRI digest shown in Figure 4, give an indication of the relative proportions of 5bp HinfI ladder material and 170bp EcoRI ladder material in this satellite III preparation.

Repeated sequence components of satellite I

Preparations of satellites I, II, III and IV were analysed with a range of restriction enzymes, as listed in the Methods section. Of the enzymes tested, only HinfI cleaved all preparations of satellites III and IV to a major extent, and only HinfI and TaqI cleaved both preparations of satellite II to a major extent. Satellite I was cleaved extensively by AvaII, DdeI, EcoRII, Fnu4HI, HinfI and MboI. Fragment sizes for ^{32}P-nucleoside triphosphate labelled digests of satellite I with these

Table 1: HinfI digestion products of eluted fragments from EcoRI digest of satellite III.

Eluted fragment size (bp)	340	680	850	1360
Major products of digestion with HinfI (bp)	171,169	340 171,169 195,100,50	340 170 125,50 95,80	340 195,100,50
Minor products of digestion with HinfI (bp)	340 295,50 95,80	510 125 80	-	170
HinfI 5bp ladder	-	-	faint	faint

enzymes are displayed in Figure 5. Double digests were carried out using all possible pairs of the five enzymes AvaII, EcoRII, Fnu4HI, HinfI and MboI. In addition, all bands, including undigested origins and 3500bp smear regions, were cut out of 4% polyacrylamide gels of satellite I digested with Fnu4HI and HinfI. The DNA eluted from these bands was redigested with a number of enzymes. These mapping experiments have indicated that the five enzymes AvaII, EcoRII, Fnu4HI, HinfI and MboI all cleave the same component

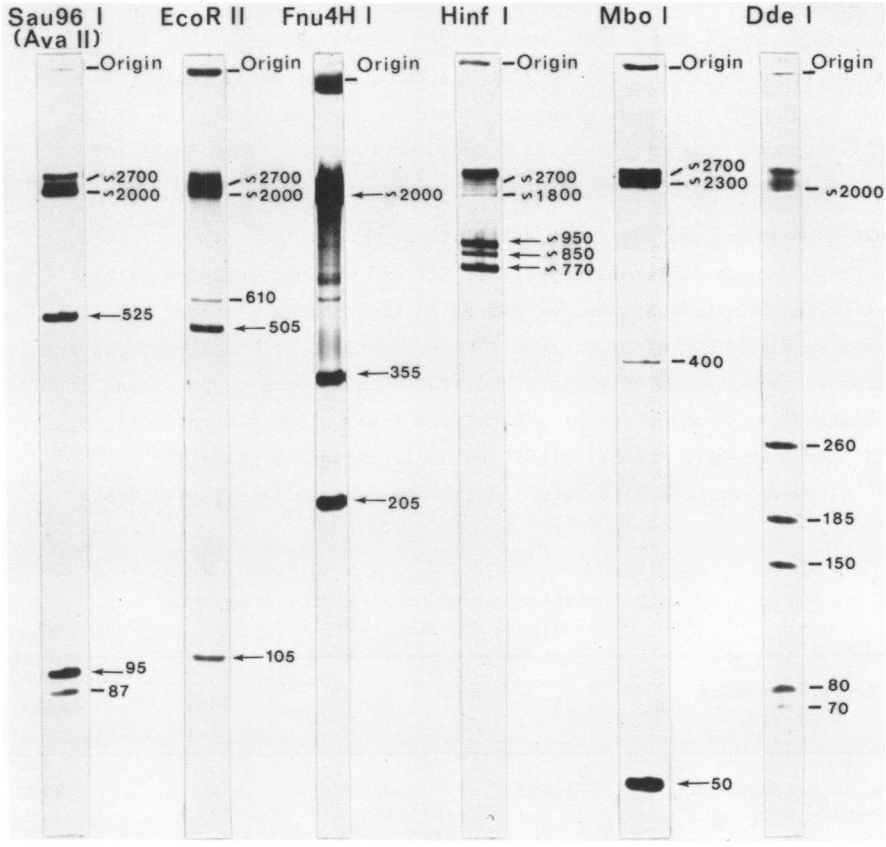

Figure 5. 4% polyacrylamide gels of restriction endonuclease digestion products of satellite I. Fragments were end-labelled with appropriate ^{32}P-nucleoside triphosphates. Fragment sizes are given in base pairs. Sizes of large fragments (1000-3000bp) were determined by electrophoresis of satellite I digests in 1.5% agarose. All gels also contain an approximately 3500bp band, which indicates an accumulation of undigested material. Fragments which have been located in the mapped segment of satellite I are marked with arrows. The location of the remainder of the fragments is not known.

of the satellite. This component appears to consist of a region which is "noisy" by restriction enzyme analysis, associated with a "silent" region. The "noisy" region is bounded by two HinfI recognition sites, 770bp apart, and contains recognition sites for the four other enzymes studied. A preliminary map of the HinfI fragments of satellite I is provided in Figure 6; more detailed sequence analysis of satellite I will be provided in a later publication.

Base sequence of HinfI restriction fragments of satellite III

Base sequences were determined for three fragments, of approximate length 45bp, 50bp and 100bp, from the HinfI 5bp ladder of satellite III. The base sequences are presented from the 3' end, for the first 42 bases of the 45bp fragment, the first 47 bases of the 50bp fragment and the first 91 bases of the 100bp fragment (Figure 7). The exact sequences at the 5' end of the fragments were not determined and therefore the exact size of each fragment is not known. All three sequences consist of a tandemly repeated unit, 5'TTCCA, with an apparently identical, G+C rich sequence, 9bp in length, at the 3' end of each fragment. The 100bp fragment is clearly a dimer, consisting of two nearly identical subunits, one of which is 49bp in length. The 9bp G+C rich sequence is present twice. The

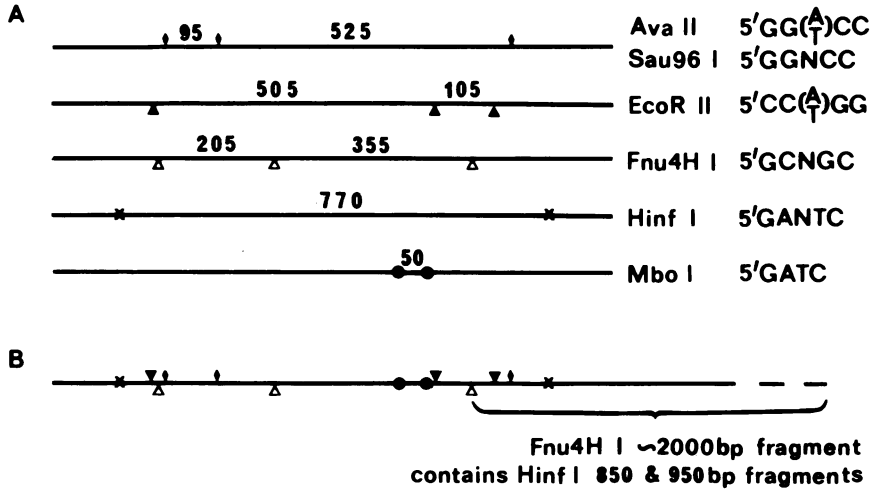

Figure 6. Restriction site map of satellite I.
A: Recognition sites and fragment sizes for individual enzymes.
B: Restriction site map of satellite I for AvaII/Sau96I, EcoRII, Fnu4HI, HinfI and MboI.

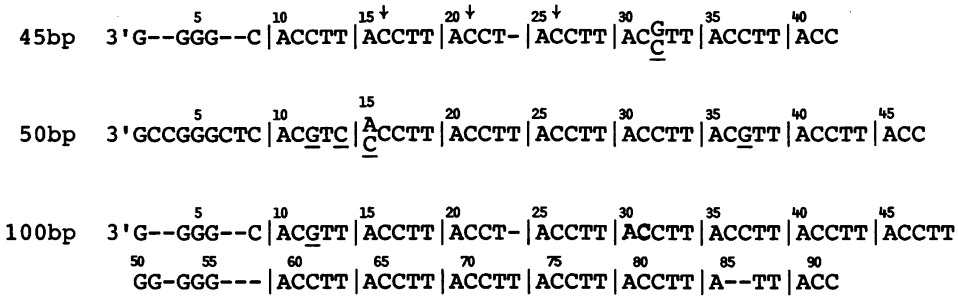

Figure 7. Base sequences of the 45bp, 50bp and 100bp HinfI fragments of satellite III. The 5bp monomers, 5'TTCCA, are marked and positions where the consensus sequence has clearly diverged from the monomer sequence are underlined. Positions where the consensus sequence could not be easily read are left blank. Positions in the 45bp sequence where a single base change would generate a HinfI and TaqI site are marked with arrows. The 100bp fragment is presented as two tandem repeats of 49bp each.

sequence at the junction between the subunits is 5'GGTTCCA, so the absence of an A between positions 49 and 50 explains the lack of a HinfI site at the centre of the dimer.

DISCUSSION

Criteria of purity of satellite preparations

The data presented here confirm that the human satellites I, II and III are distinct DNA fractions. All satellite preparations will contain a certain proportion of contaminating repeated sequences; for instance, satellite I contains a very small amount of the satellite III HinfI ladder, and EcoRII digests of all satellites may contain very small amounts of a series of fragments which appear as prominent repeated sequence components in EcoRII digests of total DNA. In addition, variable amounts of the repeated sequences which form the EcoRI and HaeIII 170bp ladders are isolated in satellites II, III and IV. The 170bp ladder forms a major part of the repeated sequences in total human DNA, and appears to consist of a collection of related sequences, variants of which have different chromosomal locations (19) and may be isolated in different satellite preparations (6). We have observed that sequence variants, which differ in number and position of HinfI sites (Table 1), may be isolated in satellite III. The 170bp ladder material cannot be termed a contaminant of satellite

preparations, since the relative amount of the ladder fragments in EcoRI digests does not correlate with the relative purity of satellite preparations, as determined by buoyant density analysis, with subsequent curve analysis of the resultant DNA peak. Preparation of satellites enriched for these sequences has proved to be a useful method for studying them (6,19). However, we locate the 170bp ladder material as a relatively minor component in most satellite II, III and IV preparations, and its presence cannot be used as a criterion of purity for any satellite preparation.

Restriction enzyme analysis of satellite I

No data from restriction enzyme analysis of human DNA satellite I have been published to date. The restriction site data for satellite I, presented here, allow a comparison of satellite I with satellites II and III.

Satellite I has been classed as one of a "family of A+T rich" (FAT) satellites, with simple sequence characteristics, in human DNA (10,12). The lack of restriction enzyme sites in large portions of the satellite supports the DNA fingerprinting data (12), that satellite I consists of a simple sequence or a set of simple sequences. Yet this A+T rich satellite contains a region with frequent sites for a number of restriction enzymes with related, G+C rich, recognition sequences - AvaII: 5'GG($\frac{A}{T}$)CC, EcoRII: 5'CC($\frac{A}{T}$)GG, Fnu4HI: 5'GCNGC. The sequence relationships between the "noisy" region and the adjacent "silent" regions of satellite I remain to be established.

Organisation of simple sequences in satellites II and III

Satellite III has been characterised in terms of the short repeated sequence, 5'TTCCA, which is a major component of all satellite III and IV preparations. The sequence can be isolated from HinfI digests of the satellite, where it appears as a ladder of fragments, with adjacent fragments separated by an increment of 5bp. The presence of HinfI sites within the 5bp repeated sequence is not surprising; only one base substitution is required for each HinfI site, 5'TTCCATTCCA→5'TTCGATTCCA. However, the lack of TaqI sites in the satellite III HinfI fragments is unexpected, since such single base substitutions must yield one TaqI site (5'TCGA) for every HinfI site. The lack of TaqI sites presumably indicates that any sequence change leading to the generation of a HinfI site in the 5bp repeated sequence has occurred as part of a larger change in the region immediately 5'

to the newly generated HinfI site. In the three fragments sequenced, this change has involved the insertion of an 8 or 9bp, G+C rich sequence into the 5bp repeated sequence, immediately 5' to the HinfI site. Further studies will establish the distribution of the 9bp sequence in all fragment size classes of the HinfI ladder, and thus yield information about the origin and relationships of the ladder components.

Satellites II and III are similar in that both fractions of DNA contain frequent HinfI sites, but are differentiated by the observations that: (i) HinfI digests of satellite II contain a far greater proportion of very small fragments, (ii) most of the HinfI fragments of satellite II show no obvious size relationship to each other, and (iii) a significant fraction of HinfI sites in satellite II contain a central ($\frac{G}{C}$) pair. The high concentration of TaqI recognition sites in the HinfI fragments of satellite II provides further evidence that the sequence organisation for major parts of satellites II and III is quite different.

A fragment of approximate length 5bp is prominent in the HinfI/TaqI double digest of satellite II, indicating that, in this highly repeated DNA, HinfI and TaqI sites are often closely associated. Satellite II, by restriction enzyme analysis, appears in total to be very similar to the cloned 1770bp EcoRI fragment, isolated from satellite III and partially sequenced by Cooke and Hindley (11). This sequence also contains frequent, and often overlapping, HinfI and TaqI sites, which are scattered through the sequenced regions of the fragment, and are found in lengths of DNA obviously related to the 5bp repeated sequences, 5'TTCCA or 5'TTCGA. Wu and Manuelidis (10) have reported that the sequence of the cloned 1770bp satellite III EcoRI fragment is compatible with DNA fingerprints of satellite II (12), and Cooke and Hindley (11) have reported that the cloned 1770bp EcoRI fragment will hybridise to satellite II as well as satellite III. Satellite II, prepared in our laboratory, contains a prominent 1800bp EcoRI fragment. If the sequences of HinfI and TaqI fragments of satellite II do prove to be similar to those found in the cloned 1770bp EcoRI fragment, an interesting problem emerges, about the origins and evolutionary relationships of the simple sequence components of satellites II and III, two highly repeated DNA fractions which are closely related in sequence, but very different in organisation.

ACKNOWLEDGEMENTS

We wish to thank Carolyn Bucholtz for her co-operation and useful advice on technical aspects of this work. The work was supported by a project grant, to M.F., J.P. and P.C.V., from the National Health and Medical Research Council of Australia.

REFERENCES

1. Corneo, G., Ginelli, E. and Polli, E. (1968) J. Mol. Biol. 33, 331-335.
2. Corneo, G., Ginelli, E. and Polli, E. (1970) J. Mol. Biol. 48, 319-327.
3. Corneo, G., Ginelli, E. and Polli, E. (1971) Biochim. Biophys. Acta 247, 528-534.
4. Corneo, G., Ginelli, E. and Polli, E. (1972) Biochim. Biophys. Acta 269, 201-204.
5. Prosser. J.. Reisner. A.H.. Bradley. M.L.. Ho, K. and Vincent, P.C. Biochim. Biophys. Acta (in press).
6. Mitchell, A.R., Beauchamp, R.S. and Bostock, C.J. (1979) J. Mol. Biol. 135, 127-149.
7. Cooke, H.J. (1976) Nature 251, 288-292.
8. Bostock, C.J., Gosden, J.R. and Mitchell, A.R. (1978) Nature 272, 324-328.
9. Manuelidis, L. (1976) Nucl. Acids Res. 3, 3063-3076.
10. Wu, J.C. and Manuelidis, L. (1980) J. Mol. Biol. 142, 363-386.
11. Cooke, H.J. and Hindley, J. (1979) Nucl. Acids Res. 6, 3177-3197.
12. Manuelidis, L. (1978) Chromosoma (Berl.) 66, 1-21.
13. Marmur, J. (1961) J. Mol. Biol. 3, 208-218.
14. Both, G.W. and Sleigh, M.J. (1980) Nucl. Acids Res. 8, 2561-2575.
15. Sutcliffe, J.G. (1978) Nucl. Acids Res. 5, 2721-2728.
16. Maxam, A.M. and Gilbert, W. (1980) in Methods in Enzymology, Vol. 65, pp 499-560, Academic Press, New York and London.
17. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
18. Maniatis, T., Jeffrey, A. and Kleid, D.G. (1975) Proc. Nat. Acad. Sci. USA. 72, 1184-1188.
19. Beauchamp, R.S., Mitchell, A.R., Buckland, R.A. and Bostock, C.J. (1979) Chromosoma (Berl.) 71, 153-166.