Volume 11 Number 10 1983 Nucleic Acids Research

Unusual domains of human alphoid satellite DNA with contiguous non-satellite sequences: sequence analysis of a junction region

S.Steven Potter and Richard S.Jones

Biology Department, Wesleyan University, Middletown, CT 06457, USA

Received 18 February 1983; Revised and Accepted 15 April 1983

ABSTRACT

The sequence organization of cloned segments of Human DNA carrying unusual domains of alphoid satellite was studied by restriction mapping, electron microscopy and base sequence analysis. In some cases restriction mapping revealed the absence of the typical 340 bp EcoR 1 dimer, although blot hybridizations showed the extensive presence of alphoid satellite. A variant monomeric construction was demonstrated by DNA sequencing. Furthermore, inverted repeats within these domains were detected by electron microscopy. In one case these were shown to be the result of interruptions in the satellite sequence by members of a family of repetitive, conserved elements.

INTRODUCTION

The satellite sequences of eukaryotic genomes are indeed enigmatic (for reviews see 1-3). Their tandem repeats of simple sequence can comprise the bulk of the genome (kangaroo rat), a small fraction of the genome (human) or esssentially none of the genome (yeast). These generally centromeric and heterochromatic sequences are not transcribed (although recently some very interesting exceptions have been found, [4,5]) and presumably noncoding (even though there are genes interspersed in heterochromatin at low density [3]). Many roles for satellite DNA have been proposed, but none have been established.

Nevertheless, the satellite sequences do exist and often in large quantities, even though their tandem repeat construction would expedite a rapid elimination by simple unequal crossover, if such removal were advantageous. It seems reasonable to conclude, therefore, that their function is real, and merely awaits experimental confirmation.

Contrasting with our relative ignorance of the purpose of

© IRL Press Limited, Oxford, England.

satellite DNA is the considerable store of information that we now have concerning its structure. For example, the alphoid satellite sequences, originally discovered in the African green monkey (6,7), have now been found in a wide variety of primates (8-12). In some cases these sequences constitute as much as 20% of the genome, and in other cases (e.g., Homo sapiens) only about 2%. Combinations of restriction enzyme analyses and DNA sequence analyses have shown that in some organisms (e.g., African green monkey) the satellite exists as tandem arrays of 172 bp monomers, with individual units diverging about 3% from the consensus sequence (13,14). But in other cases (Human, baboon and bonnet monkey) the unit of construction is a dimer of about 340 bp that is itself made of two subunits that are about 27% divergant (15-18). Dimer to dimer comparison within a species reveals only approximately 1% variation (16).

The regions of satellite DNA contiguous with non-satellite sequences are of particular interest for several reasons. First, these might represent terminal junction sequences between non-satellite DNA and the long tandem arrays of satellite. Second, the structure of such terminal satellite regions, according to certain models of satellite evolution (19), would be aberrant. And third, satellite sequences might serve as useful traps for the capture of human transposable elements. That is, non-satellite sequences interrupting tandem arrays of satellite might represent mobile elements within the human genome.

In this report we describe the purification and analysis of unusual domains of human alphoid satellite DNA. Non-dimeric alphoid satellite is demonstrated to exist in the human genome, and data relevant to the understanding of the evolution of these sequences is presented. Moreover, a dispersed repeated element with a conserved construction that is capable of interrupting the alphoid satellite is described.

MATERIALS AND METHODS

DNA Purification

Plasmid DNAs were purified by the rapid boil method of Holmes and Quigley (20), with the modifications previously

described (21). This DNA was excellent for restriction mapping, sequencing and electron microscopy, but it contained an inhibitor of the Bal31 exonuclease which could be removed by CsCl ethidium bromide isopycnic centrifugation.

Hela cell DNA was prepared by resuspending pelleted cells in 0.1 M Tris pH 7.4, 0.05 M EDTA and lysing with a final concentration of 1% sodium lauryl sulfate. The lysate was extracted three times with neutralized phenol and then once with chloroform, followed by exhaustive dialysis against 10 mM Tris pH 7.4, 1 mM EDTA. DNA concentration was determined by the method of Burton (22).

Library Construction and Screening

The library construction and screening were performed basically as previously described (23), with the following modifications. The pBR322 DNA, after linearization with BamH 1, was treated with calf intestinal alkaline phosphatase (24) to prevent recircularization of plasmid only molecules. The Hela DNA was partially digested with Sau3A 1 in a series of experiments yielding average DNA segment sizes ranging from 3 Kb to 30 Kb. The partial digests were pooled and electrophoresed through 0.4% low gel temperature agarose, poured and run at 4 degrees C. DNA segments larger than 8 Kb were cut from the gel and purified by melting the agarose at 65 degrees C, followed by two phenol extractions, one chloroform extraction and ethanol precipitation. After ligation of plasmid and Hela cell DNA , and transformation of *E. coli* HB101 (23), the culture was treated with ampicillin and the library was stored as frozen aliquots of these cells.

Filter Hybridization

Pre-soak, hybridization, and post-hybridization wash were all conducted in 4XSSCP (0.6 M NaCl, 0.06 M Na Citrate, 0.08 M Na phosphate, pH 7.0) 1XDenhardt's solution(0.02% bovine serum albumin, 0.02% polyvinylpyrrolidone, 0.02% ficoll), with shaking, at 65 degrees C. The pre-soak was for 6 hr, the hybridization for 14-18 hr, and the post-hybridization wash was for 16 hr, with at least four changes of wash solution.

Electron Microscopy

Electron microscopy was conducted exactly as previously

described (23).

Introduction of Restriction Sites

In order to facilitate the DNA sequencing, restriction sites were artificially introduced into the right inverted repeat region of pa 7, using the Bal31 exonuclease, as previously described (25). In one experiment pa 7 was first opened with BamH 1, then treated for various periods with Bal31, monitoring the digestion by gel electrophoresis. Polynucleotide linkers carrying Xho 1 recognition sites were blunt-end ligated on, with T4 ligase, and the molecules were cleaved with Xho 1, to generate "sticky ends", after which the molecules were re-circularized with T4 ligase. Following transformation of E. coli HB101, and selection on ampicillin plates, random colonies were used for DNA mini-preps (20) and restriction analyses, to determine the exact placements of the Xho 1 sites.

In a separate experiment the initial cleavage of pa 7 was to the left of the region of interest, and Eco R1 sites were introduced at varying positions, following Bal31 digestion.

DNA Sequencing

DNA sequencing was performed according to the partial chemical cleavage method of Maxam and Gilbert (26), using the G, A>C, T+C and C only reactions. DNA was labelled at the 5' ends by polynucleotide kinase and at the 3' ends by use of the Klenow fragment of DNA polymerase 1.

Materials

Isotopes were purchased from New England Nuclear and enzymes were from Bethesda Research Labs and New England Biolabs. Acrylamide was from Biorad and agarose and low gel temperature agarose were from Sigma.

RESULTS

Isolation and Restriction Analysis of Unusual Alphoid DNA Segments.

In our initial studies we screened a lambda library of human DNA segments (kindly provided by T. Maniatis) using as a probe an unusual (non-dimeric) cloned segment of human alphoid satellite DNA isolated from a plasmid library of Hela cell circular DNA (Jones and Potter, in preparation). Although this

preliminary work did provide some data of interest, suggesting regions of alphoid satellite sequence inverted relative to each other, we found that the lambda clones carrying satellite sequences were very unstable. Restriction digests routinely yielded submolar segments (faint bands in gels), suggesting heterogeneity in the DNA, likely caused by recombination between satellite subunits. Similar problems have been reported by others (27,28). We therefore discontinued work with the lambda cloning system, where satellite DNA segments studied might carry artifacts, and proceeded to construct a library of Hela cell DNA segments in the plasmid vector pBR322 (see Materials and Methods). We, and others, have found that inserted segments carrying tandem repeats are very stable in the plasmid pBR322 in *E. coli* HB101 (rec A mutant)(25,29,30).

Plasmid clones showing homology to alphoid satellite were selected at random and physically mapped by restriction endonuclease analysis. The molecules fell into two distinct categories. The first type consisted entirely of tandem repeats of the typical alphoid 340 bp Eco R1 dimer (see pa 1 in Fig. 1). This sort of construction has already been extensively studied (16) and we did not deal with it further. Molecules typical of the other category are also shown in Fig. 1 (see pa 3, pa 5, pa 7). In these cases there are no obvious repeat patterns in the restriction maps, and in particular there is no tandem repetition of 340 bp Eco R1 segments.

Distribution of Alphoid Sequences.

We were interested in learning the distribution of alphoid satellite sequences within the pa 3, pa 5, and pa 7 segments. For each clone we performed various single and multiple restriction digests which were then electrophoresed in agarose gels and blotted to nitrocellulose (31). These Southerns were then hybridized to the same nick-translated (32) alphoid satellite sequences used for the initial library screen, and the segments with homology were detected by autoradiography. The results for pa 3 and pa 7 were very simple. Every restriction segment with insert DNA hybridized (data not shown), indicating that the alphoid sequences are not limited to some small regions on these clones. The results with pa 5 did, however, reveal one

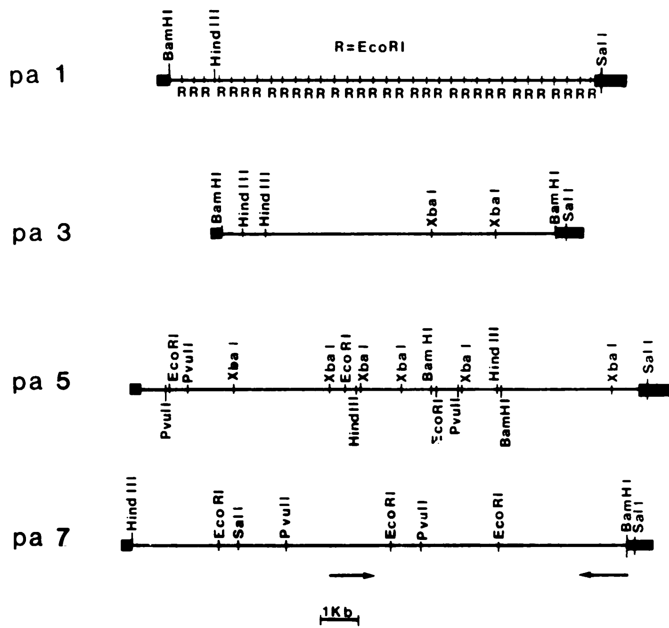

Figure 1. Restriction maps of cloned segments carrying alphoid satellite DNA. A plasmid (pBR322) library of human Hela cell DNA segments was constructed and screened with alphoid satellite probe. The thin lines represent the human insert segments and the thick lines represent the plasmid DNA. The arrows underneath pa 7 designate the positions of the inverted repeats. Insert sizes for pa 1, pa 3, pa 5 and pa 7 are 11.0, 8.8, 13.0 and 12.8 Kb respctively.

region of this insert (the left most 900 bp) that consisted of sequence that was not alphoid satellite (Fig. 2). In general though, even pa 5 carried alphoid sequence throughout most of its length. This data, plus the sequence data to be presented later, indicates that these cloned segments carry long arrays of unusual alphoid satellite DNA, generally lacking Eco R1 recognition sites and occasionally interrupted (or terminated) by non-satellite sequences.

Distribution of Inverted Repeats.

Our initial experiments with lambda clones had indicated that long inverted repeats might exist within alphoid satellite. We pursued this further by examining the plasmid inserts for the presence of inverted repeats. The circular molecules were

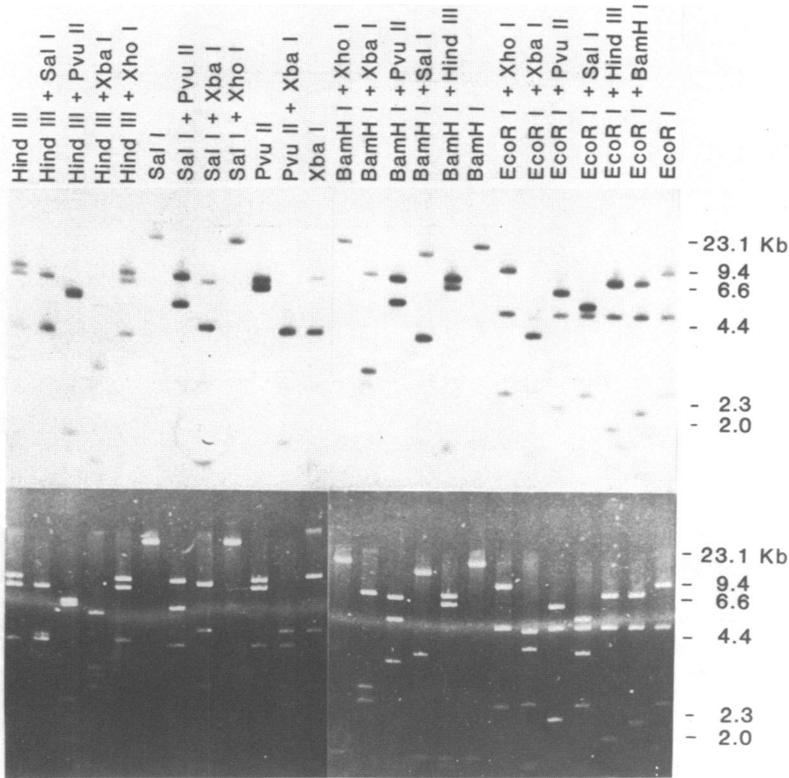

Figure 2. Autoradiograph and ethidium stain showing restriction segments of pa 5 homologous to alphoid satellite. The pa 5 DNA was digested with various restriction endonucleases, electrophoresed on 0.6% agarose gels, Southern blotted to nitrocellulose and hybridized to a nick-translated restriction segment of alphoid satellite. Only the pBR322 sequences and the leftmost 900 bp of the insert showed no homology to the satellite. For example, the third band down in the ethidium stain of the Pvu II digest, carrying pBR322 sequences and the left edge of the insert, clearly does not hybridize in the autoradiograph.

linearized by cleavage with a restriction enzyme, then denatured in alkali and, after neutralization, allowed to renature under dilute DNA conditions favoring intramolecular reannealing. For pa 5 we consistently observed partially collapsed structures, with a significant fraction of double-stranded DNA, but these were variable in appearance suggesting multiple possible arrangements or registers for homoduplexing (data not shown). In

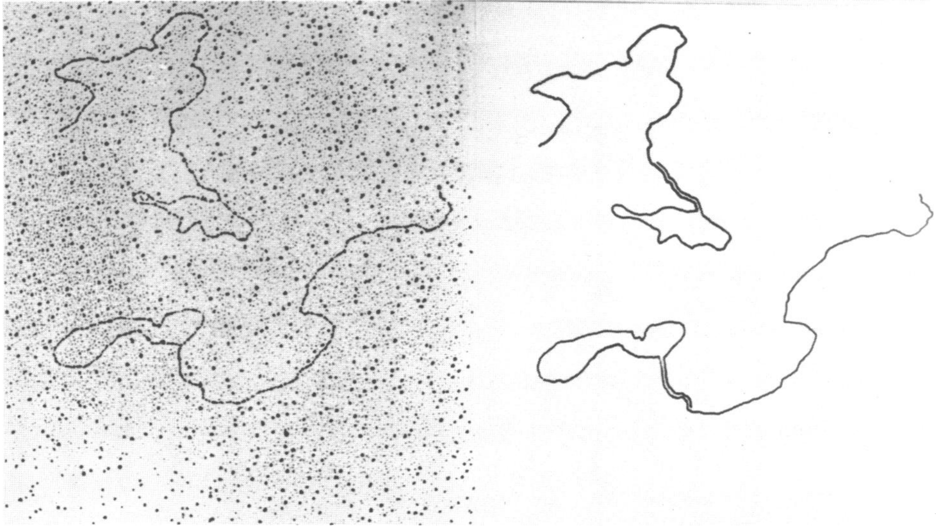

Figure 3. Electron micrograph of pa 7 DNA. The DNA was linearized by cleavage with Bam H1, then denatured in alkali and reanealed under dilute DNA conditions. The resulting stem-loop structure indicates the presence of inverted repeats.

the case of pa 7, however, the results were much more singular. A simple stem-loop structure, resembling a lariat (Fig. 3) was repeatedly seen. The positions of the inverted repeats of pa 7 were initially determined by measurements of electron micrographs of the molecules, and later fixed more exactly by DNA sequence analysis.

DNA Sequence Analysis.

We were particularly interested in learning more of the nature of these inverted repeats. Our Southern blot experiments indicated that all restriction segments of this pa 7 clone carried alphoid satellite. But if these inverted repeats consisted of alphoid satellite then why weren't other homoduplex structures also generated, since the alphoid satellite was also present at other locations in the insert, and many possible duplex registers should exist? Perhaps these inverted repeats consisted of a pair of very aberrant satellite regions that hybridize well to each other but not to the other alphoid sequences of this clone. Or, alternatively, perhaps the inverted

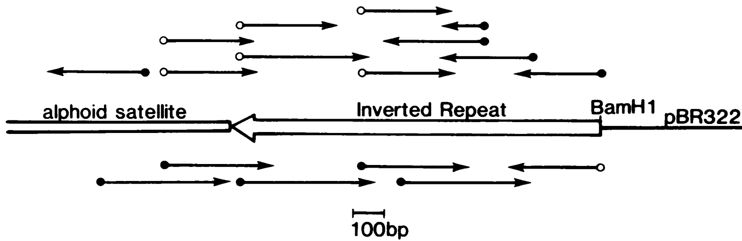

Figure 4. Sequencing strategy for the right inverted repeat of pa 7. Each arrow represents a separate end label-sequencing experiment. Open circle tails indicate 5' end labelling by polynucleotide kinase and closed circle tails indicate 3' end labelling using the Klenow fragment. Except for the Bam H1 site all restriction sites labelled were introduced by the method of Bal 31 *in vitro* mutagenesis described in Materials and Methods.

repeats are made of DNA that is not even related to satellite sequence, and the hybridization of these restriction segments to alphoid satellite in the Southern blot experiments is the result of flanking satellite that resides on the same restriction segments.

To resolve this matter we determined the base sequence of the right inverted repeat and its flanking region (Fig. 1). The sequencing strategy is shown in Fig. 4. To facilitate the sequencing in general, and in particular to make it possible to sequence the simple repeats of satellite DNA, a method of in vitro mutagenesis was used (see Materials and Methods). By using a combination of restriction digestion, Bal31 exonuclease digestion, and polynucleotide linker ligation, it was possible to introduce known restriction sites into useful positions for sequencing. As shown in Fig. 4, most of the region of interest was sequenced at least three independent times, including both DNA strands.

The resulting sequence is shown in Fig. 5. The first 511 bp of this sequence represents flanking region to the left of the inverted repeat. Analysis of this sequence reveals that it consists of three alphoid satellite subunits of 171, 171 and 169 bp. In Fig. 6 these three unusual alphoid subunits are displayed for comparison with the standard subunits of the Human alphoid dimer and the African green monkey monomer. Recall that

```
                                                            50
GAGCAGTTTG GAAACAGTGT TTTTCTAGTA TCTGCAAATG TATATTTGGA
                                                           100
GTGCTTTGAG GCTCATGGTG GAAAAGGAAA TATCTTCACA TAAAAACTAG
                                                           150
ACAAGAGCAC TCTGAGAAAC TTCTTTGTGA TATGTGCATT GAACTCACAT
                                                           200
AGCTAAACCT TTCTTTTCAT TGAGCAGTTT TGATACCCTC TTTTTACAGA
                                                           250
ATCTTCAAGT GGATATTTGG AGCACTTTGT GGCCTCTACT GGAAACGGAA
                                                           300
ATATCTTCAC ATAAAAACTG GACAGAAGCA TTCTGAGAAC CTTCTTTTTG
                                                           350
ATGGATGCAT TCATCTCACT GAGTTGAACG TTTCTTTTGA CTCAGCAATT
                                                           400
TTGAAAAACT CTTTTTGTAG AATCTGCAAC TGGATATTTT AGCACTTTGA
                                                           450
GGCCTATGTT CAAAAGGAAA TATCTTCACA TAAAAACTAC ACAGAACCAT
                                                           500
TCTGAGAACC TTCATTGTGA TATGTGCATT CCTATCACAG AATTGAACTT
                                                           550
TTCTTTTGCT TTTCTTGTTT TTCTTTTTAT TATACTTTAA GTTCTAGGGT
                                                           600
ACATGTGCAC ATTGTGCGGT TAGTTACATA TGTATACATG TGCAATGCTG
                                                           650
CACATGTAAC TCGTCATCTA GCATTAGTTA TATCTCCCAA TGCCATCCCT
                                                           700
CCCCCCTCCC CCCACCCCAC AACAGTCCCC AGAGTGTGAT ATTCCCCTTC
                                                           750
CTGTGTCCAT GTGATCTCAT TGTTCAATTC CCACCTATGA GTGAGAACAT
                                                           800
GTGGTGTTTG GTTTTTTGTC CTTGTGATAG TTTGCTGAGA ATGATGGTTT
                                                           850
CCAGCTTCAT CCATGTCCCT GCAAAGGACA TGAACTCATC ATTTTTTATG
                                                           900
GCTGCATAGT ATTCCATGGT GTATATGTGC CACATTTTCT TAATCCAGTC
                                                           950
TATCATTGTT GGACATTTGG GTTGGTTCCA AGTCTTTGCT TTTGTGAATA
                                                          1000
ATGCCACAAT AAACATACGT GTGCATGTGT CTTTATAGCA GTATGATTTA
                                                          1050
TAGTCCTTTG GGTATATACT CAGTAATGGG ATGGCTGGGT CAAATGGTAT
                                                          1100
TTCTAGTTCT AGATCCCTGA GGAATCGCCA CACTGACTTC CACAATGGTT
                                                          1150
GAACTAGTTT ACAGTCCCAC CAACAGTGTA AAAGTGTTCC TATTTCTCCA
                                                          1200
CATCCTCTCC AGCACCTGTT GTTTCCTGAC TTTTTAATGA TTGCCATTCT
                                                          1250
AACTGGTGTG AGATGGTATC TCATTGTGGT TTTGATTTGC ATTTCTCTGA
                                                          1300
TGGCCAGTGA TGATGAGCAT TTTTTCATTG TTTTTTGGCT GCATAAATGT
                                                          1350
CTTCTTTTGA GAAGTGTCTG TTCATGTCCT TCGCCCACTT TTTGATGGGG
                                                          1400
TTGTTTGTTT TTTTCTTGTA AATTTGTTGG AGTTCATTGT AGATTCTGGA
                                                          1450
TATTAGCACT ATGTCAGATG AGTAGGTTGC AAAAATTTTC TCCCATTTTG
                                                          1500
TAGGTTGCCT GTTCACTCTG ATGGTAGTTT CTTTTGCTGT GCAGAAGCTC
                                                          1550
TTTAGTTTAA TTAGATCCCA TTTGTCAATT TTGGCTTTTG TTACCATTGC
                                                          1600
TTTTGGTGTT TTAGACATGA AGTCCTTGCC CATGCCTGTG TCCTGAATGG
                                                          1650
TAATGCCTAG GTTTTCTTCT AGGGTTTTTA TGGTTTTAGG TCTAACGTTA
                                              1689
AAGTCTTTTA ATCCATCTTG AATTGATTTT TATACAAGG
```

Figure 5. Base sequence of the right inverted repeat and flanking alphoid satellite. The first 511 bases (approximately since it is impossible to determine the exact boundary at this time) are monomeric alphoid sequence to the left of the inverted repeat. The remainder represents the inverted repeat itself.

most of the Human alphoid satellite has been shown to consist of 340 bp dimers that are made of two divergent monomer subunits. The dimer itself though, is highly conserved, with only about 1% sequence difference from copy to copy (16). It is interesting to note that the three unusual alphoid subunits to the left of the inverted repeat are about 21% divergent from each other in all possible pairwise comparisons, and therefore exhibit no dimeric

Figure 6. Comparison of alphoid satellite sequences. The first three unusual alphoid sequences are the monomeric units presented in Fig. 6, flanking the inverted repeat. The sequence of the human consensus Eco RI dimer was determined by Wu and Manuelidas (16), and the sequence of the consensus African green monkey monomer was determined by Rosenberg et al. (13).

construction (where every other unit is identical). Furthermore these unusual monomers are still more divergent, about 26%, from both of the subunits of the standard human dimer, and perhaps not surprisingly, even more divergent, about 35%, from the standard African green monkey monomer. That is, these unusual units are divergent from each other, are not dimeric in construction, and are quite distinct from both of the typical subunits of the Human alphoid dimer. A summary of the levels of divergence found in these various comparisons is presented in Table I. A similar situation has recently been found on a cloned segment of alphoid satellite DNA isolated from a library of Hela cell circular DNA (Jones and Potter, in preparation).

Analysis of the inverted repeat itself shows it is made of sequences that are unrelated to alphoid satellite. That is, two nearby copies of this sequence, in inverted orientation, interrupt the tandem repeats of this unusual domain of alphoid satellite. We were interested in further investigating the nature of these interrupting elements. Are they repeated

Table I

		Monomers		Dimers		
		2	3	I	II	AGM
	1	22%	22%	27%	25%	33%
Monomers	2	-	20%	27%	25%	35%
	3	-	-	25%	28%	35%

Alphoid Sequence Comparisons

elsewhere in the genome? Are they dispersed? Do other copies of this element exist in a conserved construction?

Genomic Organization of the Interrupting Sequence.

To begin to answer these questions we performed whole genome Southern blot experiments. Total genomic DNA purified from Hela cells was digested with various restriction endonucleases and electrophoresed on agarose gels. The DNA was then Southern blotted and hybridized to a radioactive (nick-translated) probe with sequences from an internal portion of the right inverted repeat element (1035 bp to 1689 bp of Fig. 5). The results of such an experiment are shown in Fig. 7. When some restriction enzymes are used a smear of hybridization is observed. This is the result expected for a dispersed repeated element not cleaved (or broken only once) by the restriction enzyme used. The variable flanking sequences will result in DNA segments of different sizes with homology to the probe (giving a smear of hybridization). For other enzymes the result is quite different, however. In several cases we find a discrete band or bands. This is the result expected for a dispersed repeated element of conserved construction. When the restriction enzyme used cleaves at least twice within the element, and when the elements of the family have the same internal arrangement of restriction sites (a conserved construction), then DNA segments of discrete sizes with homology to the probe will be generated. In summary these results suggest that this inverted repeat is made of two members of one family of dispersed, repeated, and conserved elements.

The member of this family that was sequenced in this

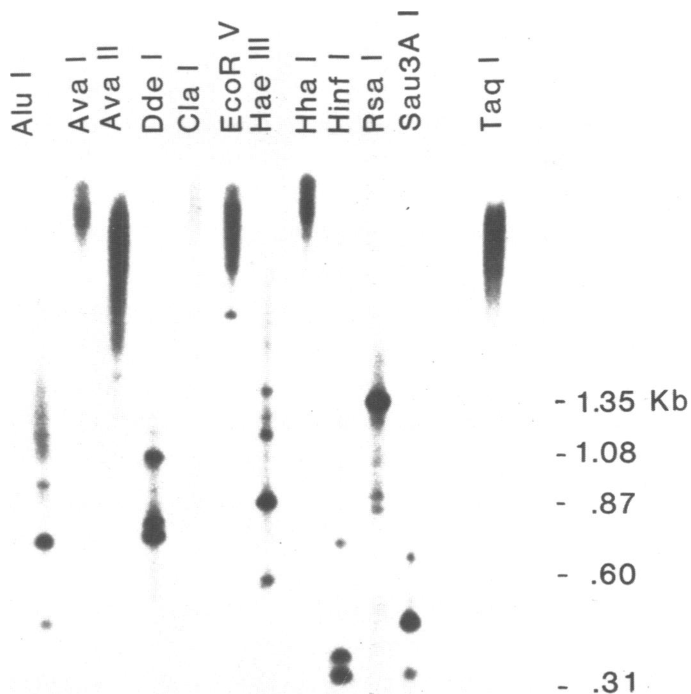

Figure 7. Genomic organization of the inverted repeat sequences. Whole genomic Hela cell DNA was digested with various restriction enzymes, electrophoresed on an agarose gel, blotted to nitrocellulose and hybridized to nick translated inverted repeat sequences from pa 7 (see text for details).

study was apparently truncated by the cloning process, since at one end we find pBR322 sequence before we again encounter alphoid satellite. Nevertheless, from the data at hand it is possible to infer the approximate size of the intact elements. First, the sequence data indicates a minimal size of 1178bp. Second, since the SalI-PvuII segment adjacent to the left inverted repeat carries alphoid satellite we can assume that this element ends somewhere prior to the left end of this segment, setting a maximal size of about 3 Kb.

Computer analysis of the inverted repeat coding potential

reveals two significant open reading frames. Both are read from the complement of the strand shown in Fig. 5. One extends from base 1679 to base 1277, giving an open reading frame 402 nucleotides long, and the other extends from 1383 to 726, resulting in an open frame of 657 nucleotides. This establishes that these inverted repeat sequences are at least potentially capable of encoding proteins. Finally, results of library screens, using the conserved element sequences as a probe, suggest an approximate copy number of 30,000, indicating that the element constitutes about 2% of the Hela cell genome. A search of the Gen Bank genetic sequence data bank failed to locate any recorded elements of the size and sequence reported here. That is, this appears to represent a previously unsequenced family of elements.

DISCUSSION

We have isolated, by plasmid cloning, a series of Human DNA segments carrying unusual domains of alphoid satellite sequence. Restriction mapping demonstrated the absence of the typical alphoid 340 bp Eco R1 dimer, but blot hybridizations revealed the widespread presence of alphoid sequence. Electron microscopy suggested the presence of inverted repeats capable of forming variable homoduplexes in one clone (pa 5), while another clone gave a single stem-loop structure (pa 7). The right inverted repeat of pa 7 and its flanking sequence were further investigated by base sequence analysis.

The sequence confirmed and extended the previous data suggesting the presence of unusual alphoid sequence. A tandem array of monomers quite different from either of the two components of the consensus dimer was found. These monomers are more similar to each other than to the typical dimer subunits, but they are nevertheless still quite variant. Preliminary sequence data from other regions of pa 7 and from another alphoid satellite clone (Jones and Potter, unpublished data) supports the conclusion that the construction is monomeric and not trimeric, or tetrameric, etc., with a higher order repeat. The prevalence of this aberrant alphoid sequence within the human genome is, however, somewhat difficult to determine because of cross

hybridization with the normal dimer. It is interesting to note though that our cloning procedure (using Sau3A 1 partials) should not significantly select for or against monomers or dimers, which appear to have approximately normal densities of Sau3A 1 sites, and the majority of our "randomly selected" alphoid clones were non-dimeric in restriction site arrangement.

Several excellent previous studies have also demonstrated the presence of unusual domains of various satellites (33-38), generally by using restriction analysis, so our finding of peculiar monomeric regions of alphoid satellite is not totally unexpected. At this point it is perhaps useful to briefly consider the evolution of satellite DNA. If unequal crossing over does play an important role in this process, as suggested by Smith (19), and supported by a number of investigators (1,2), then the central regions of long arrays of satellite would be expected to undergo rectification more often, and hence to be more homogeneous than the terminal regions of satellite. This suggests that the unusual domains studied in this report are perhaps from the end regions of the tandem repeats of the alphoid satellite.

Alternatively, these clones may represent a long array(s) constructed entirely of unusual monomers of alphoid satellite. In at least one case the predominant population of a single species of satellite of an entire individual chromosome has been shown to deviate from the "standard" organization (27). Therefore, it is certainly plausible that the unusual domains studied in this report represent extensive regions of monomeric alphoid satellite.

It is interesting to note that novel regions of alphoid satellite in the African green monkey have been previously studied with lambda cloning and restriction analysis by McCutchan et al.(28). In this earlier work junctions between satellite and non-satellite were also found, and at least in some cases these interruptions were caused by a dispersed, repeated, conserved family of elements that had been independently discovered and studied by a number of groups (39-41).

Another example where a repeated, dispersed, conserved element inserted into satellite can be found in Drosophila (42).

In this case the sequence was a well-studied transposable element named copia.

Because satellite sequences are subject to continual rectification by gene conversion and unequal crossing over, it seems likely that foreign sequences inserted into them would tend to be eliminated. Therefore the presence of a member of a repeated family of sequences within satellite provides at least tentative evidence suggesting that the members of the family are mobile and the insertion event occurred relatively recently in evolutionary time. Alternatively, the insertion of these elements may have disrupted the rectification process, allowing the surrounding satellite repeats to diverge.

The sequence data presented in this paper demonstrate that the pa 7 inverted repeats are constructed of sequence that is unrelated to alphoid satellite. Whole genome Southern blots indicate that these inverted repeats are elements of a family of dispersed, repeated, conserved elements.

ACKNOWLEDGEMENTS

Computer analysis of sequence was conducted using the Stanford Sumex-Aim facility, MOLGEN project. We wish to thank Dr. T. Maniatis for providing a library of Human DNA. This work was supported by NIH grants GM 30816 and CA 00852.

REFERENCES

1. Singer, M.F. (1982) Int. Rev. Cytol. 76, 67-112.
2. Brutlag, D.L. (1980) Ann. Rev. Genet. 14, 121-144.
3. Hilliker, A.J., Appels, R. and Schalet, A. (1980) Cell 21, 607-619.
4. Diaz, M.O, Barsacchi-Pilone, G., Mahon, K.A. and Gall, J.G. (1981) Cell 24, 649-659.
5. Varley, J.M., Macgregor, H.C. and Erba, H.P. (1980) Nature 283, 686-688.
6. Maio, J.J. (1971) J. Mol. Biol. 56, 579-595.
7. Kurnit, D.M. and Maio, J.J. (1974) Chromosoma 45, 387-400.
8. Donehower, L. and Gillespie, D. (1979) J. Mol. Biol. 134, 805-834.
9. Gillespie, D. (1977) Science 196, 889-891.
10. Singer, D.S. and Donehower, L. (1979) J. Mol. Biol. 134, 835-842.
11. Musich, P.R., Brown, F.L. and Maio, J.J. (1980) Chromosoma 80, 331-348.
12. Maio, J.J., Brown, F.L. and Musich, P.R. (1981) Chromosoma 83, 103-125.
13. Rosenberg, H., Singer, M.F. and Rosenberg, M. (1978) Science 200, 394-402.

14. Thayer, R.E., Singer, M.F. and McCutchan, T. (1981) Nucl. Acids Res. 9, 169-181.
15. Manuelidas, L. and Wu, J.C. (178) Nature (London) 276, 92-94.
16. Wu, J.C. and Manuelidis, L. (1980) J. Mol. Biol. 142, 363-386.
17. Donehower, L., Furlong, L. , Gillespie, D. and Kurnit, D. (1980) Proc. Nat. Acad. Sci. (U.S.A.) 77, 2129-2133.
18. Rubin, C.M., Deininger, P.L., Houck, D.M. and Schmid, C.W. (1980) J. Mol. Biol. 136, 151-167.
19. Smith, G.P. (1976) Science 191, 528-535.
20. Holmes, D.S. and Quigle, M. (1981) Analytical Biochem. 114, 193-197.
21. Potter, S.S. (1982) Mol. Gen. Genet. 188, 107-110.
22. Burton, K. (1956) Biochem. J. 62, 315-323.
23. Potter, S.S., Brorein, W.J., Dunsmuir, P. and Rubin, G.M. (1979) Cell 17, 415-427.
24. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning (Cold Spring Harbor Laboratory).
25. Potter, S.S. (1982) Nature 297, 201-204.
26. Maxam, A.M. and Gilbert, W. (1980) Meth. Enzym. 65, 499-560.
27. Lee, T.N.H. and Singer, M.F. (1982) J. Mol. Biol. 161, 323-342.
28. McCutchan, T., Hsu, H., Thayer, R.E. and Singer, M.F. (1982) J. Mol. Biol. 157, 195-211.
29. Levis, R. and Rubin, G.M. (1982) Cell 30, 543-550.
30. Levis, R., Collins, M. and Rubin, G.M. (1982) Cell 30, 551-565.
31. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
32. Rigby, P.W.J., Dieckman, M., Rhodes, C. and Berg, P. (1977) J. Mol. Biol. 113, 237-251.
33. Roizes, G.P., Pages, M. and Lecou, C. (1980) Nucleic Acids Res. 8, 3779-3792.
34. Brown, S.D.M and Dover, G.A. (1980) Nucleic Acids Res. 8, 781-792.
35. Beauchamp, R.S., Mitchell, A.R., Buckland, R.A. and Bostock, C.J. (1979) Chromosoma 71, 153-166.
36. Cooke, H.J. and Hindley, J. (1979) Nucleic Acids Res. 6, 3177-3197.
37. Kunkel, L.M., Smith, K.D. and Boyer, S.H. (1979) Biochemistry 18, 3343-3353.
38. Darling, S.M., Crampton, J.M. and Williamson, R. (1982) J. Mol. Biol. 154, 51-63.
39. Adams, J.W., Kaufma, R.E., Kretschmer, P.J., Harrison, M. and Niennins, A.W. (1980) Nucleic Acis Res. 8, 6113-6128.
40. Schmeckpeper, B.J., Willard, H.F. and Smith, K.D. (1981) Nucleic Acids Res. 9, 1853-1872.
41. Maio, J.J., Brown F.L., McKenna, W.G. and Musich, P.R. (1981) Chromosoma 83, 127-144.
42. Carlson, M. and Brutlag, D.L. (1978) Cell 15, 733-742.