MOLECULAR AND CELLULAR BIOLOGY, June 1983, p. 967–973
0270-7306/83/060967-07$02.00/0

Vol. 3, No. 6

Interruption of an α-Satellite Array by a Short Member of the *Kpn*I Family of Interspersed, Highly Repeated Monkey DNA Sequences

RONALD E. THAYER* AND MAXINE F. SINGER

Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205

Received 23 December 1982/Accepted 2 March 1983

We describe here the interruption of a cloned African green monkey α-satellite array by an 829-base-pair-long nonsatellite DNA segment. Hybridization experiments indicate that the sequences within the interruption are homologous to segments frequently found in the 6-kilobase-pair-long members of the *Kpn*I family of long, interspersed repeats. These data confirm and extend earlier results suggesting that sequences common to the *Kpn*I family can occur independently of one another and in segments of variable lengths. The 829-base-pair-long segment, which is termed *Kpn*I-RET, contains a terminal stretch of adenosine residues preceded by two typical but overlapping polyadenylation sites. *Kpn*I-RET is flanked by direct repeats of a 14-base-pair-long segment of α-satellite that occurs only once in the satellite consensus sequence. These structural features suggest that *Kpn*I-RET was inserted into the satellite array as a movable element.

Recent investigations have demonstrated the presence in primate genomes of a family of interspersed repeated sequences whose members are up to about 6 kilobase pairs (kbp) in length (1, 10, 13–17, 22, 24–26). The family is generally referred to as the *Kpn*I family, although in some reports it is called the 6.4-kbp repeat unit (1, 13). Estimates of the reiteration frequency of the repeat unit in humans and African green monkeys vary from about 4×10^3 to 5×10^4. However, recent data (10) suggest that in the monkey genome, distinct portions of the long repeat unit may occur independently of one another and have differing repeat frequencies. These results are consistent with earlier observations on cloned human genomic segments which indicated that different family members have lengths ranging from 3.6 to 6.4 kbp (1). Altogether, these data suggest that the *Kpn*I family may include a complex assortment of members that share some but not all segments.

We describe here a very short member of the complex monkey *Kpn*I family. The sequence was isolated from a genomic segment selected from a monkey library (19) because it contained a different type of repeated sequence, namely, α-satellite. The *Kpn*I family member, which we term *Kpn*I-RET, interrupts the α-satellite in the genomic segment, and we were, therefore, able to define its two ends. The structure of the surrounding α-satellite sequences, as well as that of *Kpn*I-RET itself, suggests that *Kpn*I-RET was inserted after the amplification of α-satellite.

MATERIALS AND METHODS

Enzymes. Restriction endonucleases were purchased from either Boehringer-Mannheim or New England Biolabs; digestions were carried out in enzyme excess under the conditions recommended by the manufacturer. Polynucleotide kinase and calf intestinal phosphatase used for 5′ end labeling with [γ-^{32}P]ATP were obtained from P-L Biochemicals, Inc., and Boehringer-Mannheim, respectively. *Escherichia coli* DNA polymerase I large fragment, used for labeling 3′ ends with α-^{32}P-labeled nucleotide triphosphate was obtained from New England Nuclear Corp. T4 DNA ligase was purchased from New England Biolabs.

Subcloning fragments from λCaα1. All subcloning of fragments of λCaα1 was carried out by standard procedures with *E. coli* RR1 (PRC 400 from the Plasmid Reference Center, Stanford University, Stanford, Calif.), pBR322 (3), and T4 DNA ligase.

Sequencing. Fragments used for sequence determination were labeled either with polynucleotide kinase and [γ-^{32}P]ATP (ICN Pharmaceuticals, Inc.) at their 5′ termini or with *E. coli* DNA polymerase I large fragment and the appropriate α-^{32}P-nucleotide triphosphate (New England Nuclear Corp.) at their 3′ termini. All sequencing was done by the method of Maxam and Gilbert (18).

Gel electrophoresis and hybridization. All gel electrophoresis (except for sequencing gels) was on 1% agarose in 0.04 M Tris–0.02 M sodium acetate–0.002 M EDTA (pH 7.8). Southern transfers (28) and DNA hybridizations were performed as previously de-

scribed (19) except that the specific radioactivity of the probes used was at least 10^8 cpm/μg.

RESULTS

Isolation of a cloned monkey segment containing the interrupted α-satellite array. The α-satellite of the African green monkey (*Cercopithecus aethiops*) includes 3×10^6 to 5×10^6 copies of a 172-bp-long repeat unit arranged in long tandem arrays at the centromeres of all chromosomes (reviewed in references 19 and 25). Recently, we described a set of recombinant phage that were selected from a monkey genomic library by virtue of their hybridization to α-satellite (19). The library had been designed to enrich for α-satellite segments joined to other DNA sequences, and most of the 17 randomly selected α-satellite-containing phage had such junctions (8, 9, 17a, 19). The α-satellite close to the junctions was atypical, as predicted if amplification of satellite arrays were by unequal crossing-over (27). One phage in the collection (λCaα1), however, was unusual (19). It contained a typical satellite array as judged from its *Hin*dIII fragments (i.e., a ladder of fragments 172, 344, and 516 bp long in decreasing amounts), but it also produced a *Hin*dIII fragment 1 kbp long that hybridized to α-satellite (Fig. 1A). Given the conformity of all the other cloned segments to the predictions of the unequal crossing-over model of amplification, the association of an atypical α-containing segment with an otherwise regular satellite array was interesting, and we undertook further study.

The 1-kbp-long *Hin*dIII fragment was subcloned into pBR322 (p1, Fig. 1B). Hybridization of restriction endonuclease digests of p1 with a cloned α-satellite probe indicated that the bulk of the monkey DNA segment was not homologous to the satellite. The detectable satellite sequences comprised less than 200 bp close to that end of the 1-kbp-long segment that contained a *Bam*HI site (Fig. 1B; data not shown). By using the *Bam*HI site as a marker, the position of the 1-kbp *Hin*dIII fragment within λCaα1 itself (Fig. 1A) was determined by restriction endonuclease digestion and hybridization (data not shown).

Sequence of the interruption and the flanking α-satellite sequences. We determined the sequence of p1 (1,015 bp) as well as the sequence of the α-satellite to the right of the 1-kbp *Hin*dIII fragment within λCaα1 (Fig. 1A). The latter was obtained by sequencing from the *Acc*I site in the 1-kbp region across the boundary with satellite. For this purpose, the entire monkey insert of λCaα1 was first recloned in pBR322 (pCaα1; Fig. 1D). The insert in pCaα1 is somewhat shorter than that in λCaα1; cloned regions of α-satellite are frequently unstable in *E. coli* (19) because of recombination. The rearrangements did not involve the junction sequences since there were no sequence changes observed in the 100 bp of overlapping data obtained with p1 and pCaα1. A total of 1,265 contiguous bp was obtained from the *Hin*dIII site at the left of the interruption to 249 bp beyond the *Hin*dIII site at the right of the interruption (Fig. 1A). The sequence is shown in Fig. 2 in the same orientation as in the diagrams in Fig. 1.

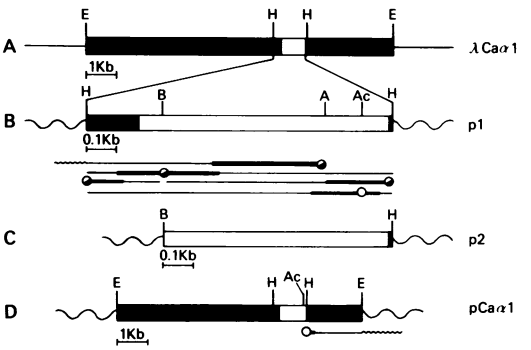

FIG. 1. Map of λCaα1 and subclones showing restriction endonuclease sites used for subcloning and sequencing. α-Satellite sequences are the filled bars; the nonsatellite sequence (*Kpn*I-RET) is shown in open bars. Straight lines show the λ arms, and wavy lines indicate the pBR322 sequences. Symbols: E, *Eco*RI; H, *Hin*dIII; B, *Bam*HI; A, *Ava*I; Ac, *Acc*I. (A) The phage λCaα1; there are numerous *Hin*dIII sites in the α-satellite that are not shown (19). (B) A subclone (p1) of the 1,015-bp *Hin*dIII fragment from λCaα1. The restriction map is not complete; only sites used in this work are shown. The sequencing strategy is shown below. The thin lines indicate the fragments used for sequencing, and the thick lines indicate the portions for which sequence information was obtained. The fragment ends labeled with ^{32}P are shown as circles. An open circle indicates 5′ end labeling; half-filled circles indicate that both strands were sequenced from that point, one being 5′ end labeled and the other 3′ end labeled. (C) A subclone (p2) of the indicated *Bam*HI-*Hin*dIII fragment of p1; all of the α-satellite except the 14 bp at the right end of p1 is removed. The first 76 bp of *Kpn*I-RET upstream of the *Bam*HI site is also missing. (D) pCaα1 contains the entire monkey insert in λCaα1 cloned into pBR322.

Starting at the leftmost *Hin*dIII site of the 1-kbp piece, there are 173 bp of α-satellite (lower-case letters, Fig. 2) followed by the 829 bp we designate *Kpn*I-RET (upper-case letters, Fig. 2) and then another 14 bp of α-satellite up to the *Hin*dIII site of the 1-kbp piece. Beyond this *Hin*dIII site are the additional 249 bp of α-satellite in the sequenced region. There is some ambiguity in the precise location of the borders between *Kpn*I-RET and α-satellite because of

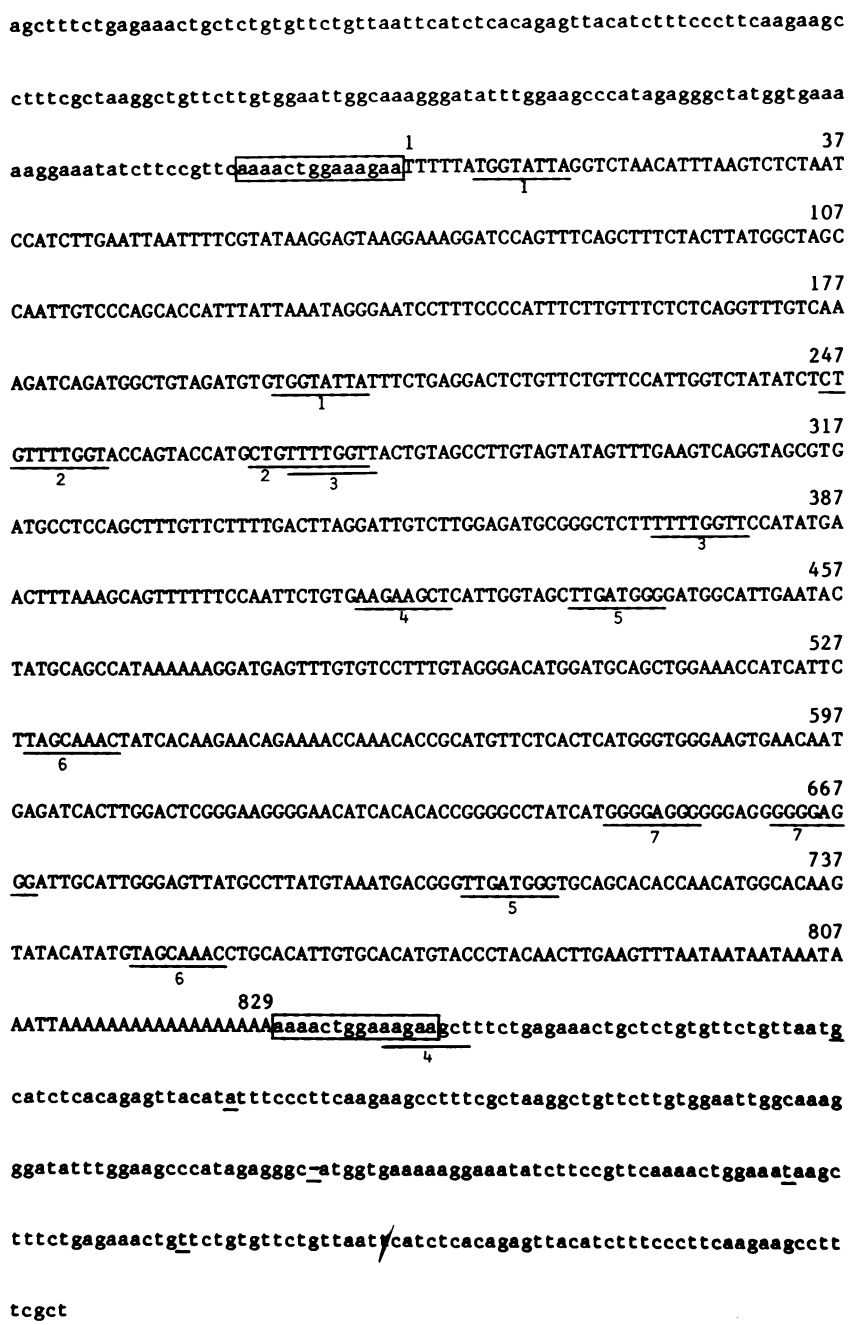

FIG. 2. Sequence of *Kpn*I-RET (upper case) and flanking α-satellite (lower case). Only the *Kpn*I-RET sequence is numbered. The first five bases of the sequence (agctt) are bases 2 through 6 of the left-hand *Hin*dIII site in λCaα1 (Fig. 1). The direct repeats of α-satellite sequence are boxed. Other perfect direct repeats are underlined and numbered. Variations from the α-satellite average sequence (21) are also underlined.

the series of A residues; we have selected one possible location. Within λCaα1, the 1,265-bp segment is embedded in a total of about 9 kbp of a typical α-satellite array (19).

There are 14 bp of α-satellite sequence at the left junction with *Kpn*I-RET that are repeated at the right junction (boxes, Fig. 2), although they occur only once in a typical α-satellite consensus sequence (21). Thus, *Kpn*I-RET is flanked by direct repeats of its apparent insertion target

site. The *Kpn*I-RET sequence ends with a long, A-rich stretch (residues 794 to 829) that includes two perfect overlapping polyadenylation sites (AATAAA, residues 800 to 809).

We used the computer program of Queen and Korn (20) to analyze the *Kpn*I-RET sequence. Six perfect repeats are underlined in Fig. 2, as is one set of repeats (no. 4) that involves the internal *Kpn*I-RET sequence and the downstream junction with α-satellite. The significance, if any, of the repeats is unknown, but the analysis suggests that *Kpn*I-RET did not evolve from multiple, short-tandem repeats. There are several dyad symmetries within the sequence. The region between residue 760 and 783 is imperfectly inversely symmetrical to residues 510 to 492 and 738 to 726; in each case, 13 bp or more can be formed. Additional imperfect dyad symmetries occur in the following pairs: 1–9/473–465 and 70–80/148–137. Also, residues 397 to 405 form a dyad symmetry to the last two residues of *Kpn*I-RET and the first seven residues of the flanking α-satellite. Most strikingly, a stem and loop structure involving as many as 50 bp between residues 260–330 and 600–511 (in the stem) can be written; different stem structures are possible, depending on the way the strand is aligned. Understanding of the significance of these secondary structures must await further information about the *Kpn*I family.

The longest open reading frames are on the opposite strand from that shown in Fig. 2 and go (i) from the ATG at residues 744 to 742 through the TAG triplet at residues 459 to 457 and (ii) from the ATG at residues 427 to 425 through the junction with α-satellite to the TGA triplet that starts 14 bp upstream from the satellite–*Kpn*I-RET junction.

From residue 494 through the polyadenylated stretch at the end of *Kpn*I-RET, the sequence is about 85% homologous to the 325 bp in a randomly cloned and otherwise uncharacterized human repeated sequence called pPD16 (6). *Kpn*I-RET contains no long significant homologies to the sequence of human 7S RNA (30), monkey *Alu* sequences (8, 9), α-satellite (21), or the 1.9-kbp-long *Hin*dIII fragment (15) that is part of the human and monkey *Kpn*I family of long interspersed repeats (10, 15, 17, 24, 26).

Genomic frequency of sequences within *Kpn*I-RET. Sequences within *Kpn*I-RET are highly repeated in the monkey genome. Of the phages in a newly made genomic library (kindly made available by Michael Lerman and constructed as previously described [19]), 16% hybridized with the ^{32}P-labeled subclone p2 (Fig. 1C), which contains no α-satellite sequence. Six filters were used in these experiments and were prepared (2) from plates containing between 500 and 1,000 plaques. When p2 was used as a probe against digests of monkey genomic DNA with *Hin*dIII, *Hin*dIII plus *Xho*I, *Eco*RI, and *Eco*RI plus *Sal*I (after separation on 1% agarose gels and transfer to nitrocellulose), heavy smears of hybridization were seen from the top of the gel down to the level of 500-bp fragments. Some discrete bands were discernible in the *Eco*RI (about 1.4 and 4.8 kbp) and *Hin*dIII (0.9, 1.0, 1.3, and 1.5 kbp) digests. These data, as well as the library screening, are consistent with there being interspersed in the genome at least 10^4 copies of sequences included in *Kpn*I-RET.

Relation of *Kpn*I-RET to the *Kpn*I family of long, interspersed, repeated sequences. Experiments similar to those just described were also carried out, using p2 as a probe against *Kpn*I digests of monkey DNA (Fig. 3A). Again, a heavy smear of hybridization was seen in the region of large fragments. However, distinct bands of hybridization were apparent at 1.2 and 2.8 kbp. The 1.2-kbp band coincides with a band seen upon staining the original gel with ethidium bromide (Fig. 3A). Since both 1.2- and 2.8-kbp-long fragments were previously identified as components of the complex *Kpn*I family of human (17, 24, 25) and monkey DNA (10), we investigated directly the relationship of *Kpn*I-RET to this family.

Two phages termed λCaα6 and λCaα7 were among those isolated from the same monkey genomic library from which λCaα1 was obtained. Each of the two phages contains α-satellite sequences linked to distinct members of the *Kpn*I family (10). In each case, the *Kpn*I family members are about 6 kbp long. The *Kpn*I family member in λCaα7 includes a 1.2- and a 1.5-kbp-long *Kpn*I segment; in λCaα6, sequences homologous to the 1.2- and 1.5-kbp segments of λCaα7 are included in a single 2.8-kbp *Kpn*I segment (Fig. 3C). When the p2 subclone of λCaα1 (Fig. 1C) was used as a probe against *Kpn*I digests of λCaα6 and λCaα7, the 2.8- and 1.2-kbp bands, respectively, hybridized (Fig. 3B). Therefore, *Kpn*I-RET includes sequences that hybridize to the 1.2-kbp element of the *Kpn*I family but not the 1.5-kbp element. This finding is consistent with the fact that we found no significant homology between the *Kpn*I-RET sequence and the reported sequence of a cloned human 1.9-kbp *Hin*dIII fragment that is part of the *Kpn*I family and includes the 1.5-kbp *Kpn*I fragment (15). It should be noted that in most monkey *Kpn*I family members, the 1.9-kbp human *Hin*dIII fragment is replaced by a 2.5-kbp *Hin*dIII fragment (10); this corresponds to the 2.5-kbp *Hin*dIII fragment at the left of λCaα6 and λCaα7 in Fig. 3C. In addition, p2 hybridized to approximately 10-kbp-long *Kpn*I fragments from both λCaα6 and λCaα7 (Fig. 3B). These fragments include the right-hand

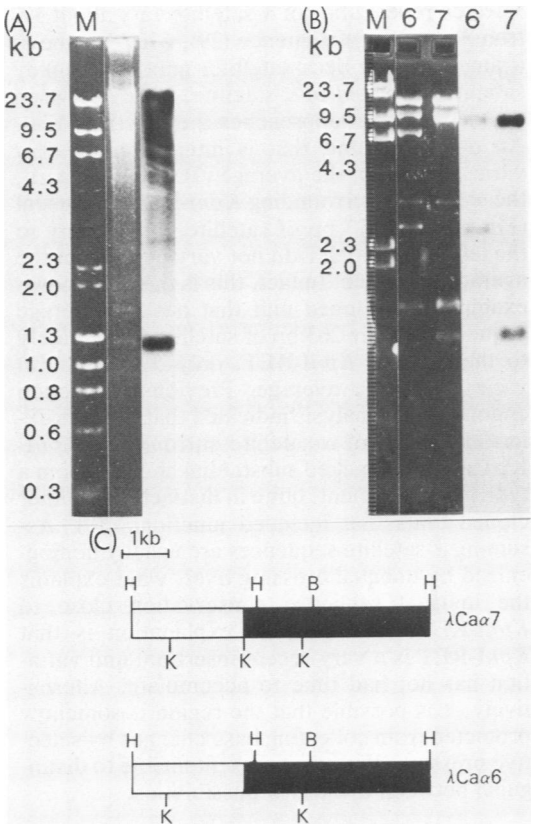

FIG. 3. Hybridization of the p2 subclone of λCaα1 (Fig. 1C) to *Kpn*I digests of total monkey DNA, λCaα6, and λCaα7. Digestion and *Kpn*I, electrophoresis, transfer to nitrocellulose, and hybridization conditions are described in the text. (A) African green monkey liver DNA (1 μg); the dark field shows the ethidium bromide-stained gel, and the light field shows the autoradiogram. (B) λCaα6 and λCaα7 (1 μg) (8, 10, 19). (C) Schematic diagrams of the *Kpn*I family segments in λCaα6 and λCaα7 (8, 10, 19). Only a few restriction endonuclease sites are shown. Regions of cross-hybridization with *Kpn*I-RET are shown by filled boxes. Not indicated is an *Alu* family sequence that is placed somewhere to the right of the *Bam*HI site in both λCaα6 and λCaα7.

portion of the *Kpn*I family members (Fig. 3C) and a portion of the right arm of the λ Charon 4A vector; in λCaα6 it also includes a short stretch of α-satellite sequence (8, 10).

DISCUSSION

The *Kpn*I family was initially detected in two ways. First, digestion of human (and other primate DNAs) with *Kpn*I revealed a set of DNA bands (typically 1.2, 1.5, 2.8, 3.4, and 4.6 kbp in length), some of which are abundant enough to be seen by ethidium bromide staining (14–17, 22, 24–26). Hybridization of cloned probes representing the 1.2- and 1.5-kbp bands to restriction digests of total human DNA indicated that the different probes each contain sequences homologous to several of the genomic fragments and, further, that sequences within the two cloned fragments are often contiguous to one another in the genome (24). Similarly, the sequences contained in a 1.9-kbp *Hin*dIII fragment that is also abundant in human DNA were found to be included within particular *Kpn*I bands (17). The situation in the African green monkey genome is similar, except that the 1.9-kbp *Hin*dIII fragment is of low abundance and there is instead a 2.5-kbp *Hin*dIII fragment; the 2.5-kbp *Hin*dIII fragment also occurs in human DNA, but relatively infrequently (10).

At about the same time that the *Kpn*I and *Hin*dIII bands were discovered, Adams and coworkers described a 6.4-kbp-long stretch of repeated DNA downstream from the human β-globin gene (1, 13). These workers showed that sequences within the 6.4-kbp segment are interspersed elsewhere in the genome, and they studied the structure of several additional segments isolated from a human genomic library; the size of the homologous regions in the cloned segments varied from 3.6 to 6.4 kbp.

More recent work has shown directly that cloned *Kpn*I fragments from human and monkey DNA hybridize with the 6.4-kbp repeat unit from near the human β-globin gene (10, 24). Furthermore, segments of about 6 kbp in length have been described in monkey DNA (10). Subcloned regions covering almost the entire 6 kbp of the monkey DNA are all repeated on the order of 10^4 times in the genome and include several of the typical *Kpn*I fragments as well as approximately 2 kbp of additional repetitive sequence. Therefore, it appears that many *Kpn*I family members are 6 kbp or more in length and include within them the shorter sequences originally detected as discrete *Kpn*I or *Hin*dIII segments, as well as additional sequences. The work of Adams et al. (1) suggested further that some human *Kpn*I family members are shorter than 6 kbp in length. Also, hybridization of the subcloned portions of the monkey family members to restriction endonuclease digests and to a genomic library of monkey DNA indicated that some members of the monkey *Kpn*I family might contain only portions of the 6 kbp (10).

In this report we describe a 10-kbp-long cloned monkey genomic segment that contains an 829-bp stretch of DNA that is homologous to regions within the longer monkey *Kpn*I family members. We conclude that *Kpn*I-RET is a distinctive member of the large and complex *Kpn*I family. The sequences within *Kpn*I-RET are homologous to those in the right-hand por-

tion of the longer *Kpn*I family members shown in Fig. 3C. However, since *Kpn*I-RET is only 829 bp long and the segments to which it hybridizes encompass over 3.5 kbp in λCaα6 and λCaα7, it seems likely that the homology extends only over a limited portion of the 3.5 kbp. We are currently trying to localize the homologous sequences more precisely. In any case, the structure of *Kpn*I-RET clearly demonstrates that elements within long *Kpn*I family members occur independently of one another in the monkey genome, as predicted from more indirect experiments (10). We do not know how many times, if at all, the *Kpn*I-RET segment itself is repeated in the genome.

We were able to define the precise size of *Kpn*I-RET because it interrupts a 9-kbp-long stretch of α-satellite. Moreover, *Kpn*I-RET is flanked by direct repeats of a 14-bp-long sequence that occurs only once in the α-satellite consensus sequence (21). As reviewed previously (9, 25), the monkey α-satellite represents a relatively recent amplification, since the α-satellites of baboons and humans are distinctly different. Any sequence that interrupts the satellite array is likely to have been inserted after amplification, and there are features in the structure of *Kpn*I-RET which are suggestive of a movable element. These features include the flanking direct repeats of the target site and the long terminal stretch of A · T base pairs. These structures are also typical of *Alu* repeats (23, 25), a known *Drosophila* movable element (5), and some pseudogenes (11, 31). The similarity between *Kpn*I-RET and pseudogenes includes the polyadenylation signals just upstream from the stretch of A · T base pairs. It has been suggested that these elements may represent a class of movable elements (5, 26) distinct from the classical transposable elements (4) and that an RNA intermediate may be involved in the transposition mechanism (7, 12, 31). That model implies that *Kpn*I-RET may have come from a transcript. To our knowledge, however, no transcripts homologous to *Kpn*I family sequences have been reported; Adams and co-workers (1) describe negative results in this regard, and we are currently investigating the possibility further. In any case, the structural features and position of *Kpn*I-RET suggest that at least portions of the *Kpn*I family may be movable.

Finally, the α-satellite sequences surrounding *Kpn*I-RET are interesting. Satellites are likely to be subject to amplification and deletion by unequal crossing over (27). Monkey α-satellite provides a number of examples bearing out one prediction of that model, namely, that satellite segments close to junctions with other sequences will vary from the average sequence by more than will a random repeat unit (27). Randomly selected repeat units of α-satellite vary about 3% from the average sequence (29), whereas one at a junction with deca-satellite, another monkey satellite unrelated to α-satellite, varies between 20 and 40% as it approaches the junction (17a). An α-satellite unit that is interrupted by *Alu* varies 10% from the average (9). In this regard, the α-satellite surrounding *Kpn*I-RET is unusual (Fig. 2). The 173 bp of satellite immediately to the left of *Kpn*I-RET do not vary at all from the average sequence. In fact, this is the first known example of a cloned unit that has the average sequence. Of the 263 bp of satellite immediately to the right of *Kpn*I-RET, only 5 bp (or 2%) varies from the average. Previous restriction endonuclease analysis indicated that the approximately 9 kbp of α-satellite surrounding *Kpn*I-RET in λCaα1 lacked substantial changes from a typical arrangement, quite in distinction to other cloned units that included junctions (19). Assuming α-satellite sequences are usually homogenized by unequal crossing over, what explains the unusual sequence conservation close to *Kpn*I-RET? One possible explanation is that *Kpn*I-RET is a very recent insertion, and variation has not had time to accumulate. Alternatively, it is possible that the region is somehow protected from collecting base changes by selective pressure. We are at present unable to distinguish between these two possibilities.

ACKNOWLEDGMENTS

We are grateful to Michael Lerman for the library and to him and Giovanna Grimaldi for continuing fruitful discussions. Jake Maizel kindly helped us with the analysis of dyad symmetries. We thank May Liu for her careful preparation of the manuscript.

LITERATURE CITED

1. **Adams, J. W., R. E. Kaufman, P. J. Kretschmer, M. Harrison, and A. W. Nienhuis.** 1980. A family of long reiterated DNA sequences, one copy of which is next to the human beta globin gene. Nucleic Acids Res. **8**:6113–6128.
2. **Benton, W. D., and R. W. Davis.** 1977. Screening λgt recombinant clones by hybridization to single plaques in situ. Science **196**:180–182.
3. **Bolivar, F., R. L. Rodriquez, P. J. Green, M. C. Betlach, H. L. Heyneker, H. W. Boyer, J. H. Crosa, and S. Falkow.** 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2**:95–113.
4. **Calos, M. P., and J. H. Miller.** 1980. Transposable elements. Cell **20**:579–595.
5. **Dawid, I. B., E. O. Long, P. P. DiNocera, and M. L. Pardue.** 1981. Ribosomal insertion-like elements in *Drosophila melanogaster* are interspersed with mobile sequences. Cell **25**:399–408.
6. **Deininger, P. L., D. J. Jolly, C. M. Rubin, T. Friedmann, and C. W. Schmid.** 1981. Base sequence studies of 300 nucleotide renatured repeated human DNA clones. J. Mol. Biol. **151**:17–33.
7. **Denison, R. A., and A. M. Weiner.** 1982. Human U1 RNA pseudogenes may be generated by both DNA- and RNA-mediated mechanism. Mol. Cell. Biol. **2**:815–828.
8. **Grimaldi, G., C. Queen, and M. F. Singer.** 1981. Interspersed repeated sequences in the African green monkey

genome that are homologous to the human Alu family. Nucleic Acids Res. **9**:5553–5568.

9. **Grimaldi, G., and M. F. Singer.** 1982. A monkey Alu-sequence is flanked by 13 base pair direct repeats of an interrupted α-satellite DNA sequence. Proc. Natl. Acad. Sci. U.S.A. **79**:1497–1500.
10. **Grimaldi, G., and M. F. Singer.** 1983. Members of the KpnI family of long interspersed repeated sequences join and interrupt α-satellite in the monkey genome. Nucleic Acids Res. **11**:321–338.
11. **Hollis, G. F., P. A. Hieter, O. W. McBride, D. Swan, and P. Leder.** 1982. Processed genes: a dispersed human immunoglobulin gene bearing evidence of RNA-type processing. Nature (London) **296**:321–325.
12. **Jagadeswaren, P., B. G. Forget, and S. M. Weissman.** 1981. Short interspersed repetitive DNA elements in eukaryotes: transposable DNA elements generated by reverse transcription of RNA PolIII transcription? Cell **26**:141–142.
13. **Kaufman, R. E., P. J. Kretschmer, J. W. Adams, H. C. Coon, W. F. Anderson, and A. W. Nienhuis.** 1980. Cloning and characterization of DNA sequences surrounding the human γ-, δ-, and β-globin genes. Proc. Natl. Acad. Sci. U.S.A. **77**:4229–4233.
14. **Maio, J. J., F. L. Brown, W. G. McKenna, and P. R. Musich.** 1981. Toward a molecular paleontology of primate genomes. II. The KpnI families of alphoid DNAs. Chromosoma **83**:127–144.
15. **Manuelidis, L.** 1982. Nucleotide sequence definition of a major human repeated DNA, the HindIII 1.9 kb family. Nucleic Acids Res. **10**:3211–3219.
16. **Manuelidis, L.** 1982. Repeated DNA sequences and nuclear structure, p. 263–285. *In* G. A. Dover and R. B. Flavell (ed.), Genome evolution. Academic Press, Inc., New York.
17. **Manuelidis, L., and P. Biro.** 1982. Genomic representation of the HindIII 1.9 kb repeated DNA. Nucleic Acids Res. **10**:3221–3239.

17a. **Maresca, A., and M. F. Singer.** 1983. Deca-satellite: a highly polymorphic satellite that joins α-satellite in the African green monkey genome. J. Mol. Biol. **164**:493–511.

18. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
19. **McCutchan, T., H. Hsu, R. E. Thayer, and M. F. Singer.** 1982. Organization of African green monkey DNA at junctions between α-satellite and other DNA sequences. J. Mol. Biol. **157**:195–211.
20. **Queen, C., and L. Korn.** 1980. Computer analysis of nucleic acids and proteins. Methods Enzymol. **65**:595–609.
21. **Rosenberg, H., M. F. Singer, and M. Rosenberg.** 1978. Highly reiterated sequences of SIMIANSIMIAN SIMIANSIMIANSIMIAN. Science **200**:394–402.
22. **Schmeckpeper, B. J., H. F. Willard, and K. D. Smith.** 1981. Isolation and characterization of cloned human DNA fragments carrying reiterated sequences common to both autosomes and the X chromosome. Nucleic Acids Res. **9**:1853–1872.
23. **Schmid, C. W., and W. R. Jelinek.** 1982. The Alu family of dispersed repetitive sequences. Science **216**:1065–1070.
24. **Shafit-Zagardo, B., J. J. Maio, and F. L. Brown.** 1982. KpnI families of long interspersed repetitive DNAs in human and other primate genomes. Nucleic Acids Res. **10**:3175–3193.
25. **Singer, M. F.** 1982. Highly repeated sequences in mammalian genomes. Int. Rev. Cytol. **76**:67–112.
26. **Singer, M. F.** 1982. SINES and LINES: highly repeated short and long interspersed sequences in mammalian genomes. Cell **28**:433–434.
27. **Smith, G. P.** 1976. Evaluation of repeated DNA sequences by unequal crossover. Science **191**:528–535.
28. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
29. **Thayer, R. E., M. F. Singer, and T. F. McCutchan.** 1981. Sequence relationships between single repeat units of highly reiterated African green monkey DNA. Nucleic Acids Res. **9**:169–181.
30. **Ullu, E., S. Murphy, and M. Melli.** 1982. Human 7SL RNA consists of a 140 nucleotide middle-repetitive sequence in an Alu sequence. Cell **29**:195–202.
31. **Van Arsdell, S. W., R. A. Denison, L. B. Bernstein, A. M. Weiner, T. Mancer, and R. F. Gesteland.** 1981. Direct repeats flank three small nuclear RNA pseudogenes in the human genome. Cell **26**:11–17.