Recent amplification of an alpha satellite DNA in humans

Kendall M.Gray, James W.White, Carl Costanzi[1], David Gillespie[1], Wanda T.Schroeder, Bruno Calabretta and Grady F.Saunders

Dept. Biochemistry and Molecular Biology, Univ. Texas System Cancer Center, M.D. Anderson Hospital and Tumor Institute, Houston, TX 77030, and [1]Barry Ashbee Leukemia Research Labs., Orlowitz Cancer Inst., Hahnemann Medical Univ., Philadelphia, PA 19102, USA

Received 20 August 1984; Revised and Accepted 26 December 1984

ABSTRACT

A repeat sequence 682 base pairs (bp) long produced by cleavage of human DNA with Xba I restriction enzyme is composed of four tandemly arranged subunits with lengths of 171, 170, 171, and 170 bp each. The sequence organization of the 682 bp Xba I repeat bears a striking resemblance to other complex satellite DNAs of primates, including the Eco RI human alpha satellite family which also occurs as a 170 bp repeat. The Eco RI tetramer and the 682 bp Xba I repeat show a sequence divergence of 21%. The 682 bp Xba I repeat sequence is restricted to humans and is only distantly related to the previously reported 340 bp Xba human repeated DNA sequence.

These finding are consistant with the concept of occasional amplifications of members or groups of members of alpha satellite DNA during human evolution. Amplifications apparently occurred after humans, apes and gibbons diverged from Old World monkeys (Eco RI satellite), after humans and apes diverged from gibbons (340 bp Xba I satellite) and after humans diverged from the great apes (682 bp Xba I satellite).

INTRODUCTION

A sizeable portion of the eukaryotic genome is comprised of families of repetitive sequences (1). These repeat families may have two prevalent patterns of arrangement in the eukaryotic genome. They can be interspersed with single copy or other repetitive sequences or alternatively they may be clustered in a tandem array (1) as in satellite DNA. Interspersed and satellite DNA can each be further subdivided on the basis of average sequence length or complexity (2). Short, interspersed repeated sequences (SINES) average 300 bp in length (3) while long, interspersed repeated sequences (LINES) average 6000 bp in length (4). Simple sequence satellite DNA consists of tandemly repeating units less than 10 bp long (5) while the monomer of complex satellite sequences in primates is around 170 bp long (6,7). No known sequence relatedness exists among the four types of primate repeated DNA, except that LINES may consist, in part, of satellite-related sequences (8).

Families of repeated sequences are often named for certain restriction endonuclease recognition sites present within them. The Alu family of primate

SINES is present in approximately 500,000 copies per haploid genome (3) and is the predominant repeat family in primate DNA. Alu repeats may be found in the vicinity of known genes or even within a transcription unit (9,10). The Kpn family of primate LINES (11,12) is present in approximately 50,000 copies and members are also located near structural genes (13). Other less abundant species of SINES and LINES exist in primate DNA (14). Several related complex satellite DNA species collectively termed "alpha" satellite DNA, have been reported in various primates (2,7,15, 16). They are all based on a repeating unit of about 170 bp and are localized to the centromeres and telomeres of chromosomes. Rosenberg et al. (15) reported the nucleotide sequence of the Hind III family of guenon alpha satellite and this was quickly followed by the Wu and Manuelides report (16) of the sequence of the Eco RI family of human alpha satellite which is shared by apes and gibbons. Later, other primate alpha satellites were reported, including the Bam HI family restricted to baboons (7), a different Bam HI family restricted to colobus monkeys (7) and an Xba I 340 bp alpha satellite found in humans and apes but not in gibbons (8).

In characterization of a human Eco RI α-satellite, Wu and Manuelides (16) reported results of an experiment in which they hybridized the Eco RI 680 bp fragment with Xba I - digested human DNA that had been electrophoretically fractionated and immobilized on nitrocellulose paper. They found hybridization with a family of repeated DNA fragments 340n base pairs long, where n = a small integer. From these results and the observation that four positions in the Eco RI 340 bp sequence could generate an Xba I recognition site by a single base change, the authors concluded that the Eco RI and the Xba I satellite families were closely related.

The results of some later studies suggested a more distant relatedness. Using an Eco RI 340 bp satellite fragment as probe of Xba I digested human DNA, hybridization occurred with the 1020 bp fragment of the Xba I satellite family but there was only weak hybridization with the Xba I 680 bp satellite tetramer (12). Hybridization with a gel-purified 850 bp repeat fragment generated by digestion of human DNA with Xba I showed little homology with the Eco RI α-satellite and a very close homology with the 340 bp DNA Xba I satellite (8). Similarly gel-purified Xba I and Eco RI 340 bp DNA fragments failed to cross hybridize (D.G., unpublished).

We report here the cloning and DNA sequencing of a 682 bp fragment generated by digestion of human DNA with Xba I. Its nucleotide sequence is related but not identical to the previously published Eco RI 680 bp fragment (16). By molecular hybridization it was shown to be related to other Xba I-cleaved repeated sequences in human DNA, but its absence in other primates was unlike any other alpha satellite sequence. The sequence is enriched in the centromeres of a set of human chromo-

somes as shown by in situ hybridization. These results suggest that the Xba I 682 bp fragment represents a very recently amplified subset of the α-satellite family.

MATERIALS AND METHODS

Extraction of total genomic DNA

Human DNA was extracted from frozen biopsy samples. Approximately 5 g of tissue was homogenized three times at high speed in a 50 ml volume of 60 mM Tris (pH 7.3), 25 mM KCl, 0.1% Triton X-100, 10 mM EDTA, and 1 mM spermidine for 15 seconds each time. The homogenate was filtered through gauze and centrifuged at 1500 rpm for 10 min. The pellet was resuspended in 5-10 ml saline (0.9% NaCl) + 0.2% SDS and treated with 10 mg Proteinase K overnight at 37°C. The DNA was then extracted with phenol and Sevag (chloroform:isoamyl alcohol, 24:1) and precipitated with ethanol.

Isolation and purification of Xba I repeat fragments

A bulk (~200 µg) Xba I digest of genomic DNA was run on a 1.5% agarose horizontal slab gel. After staining the gel with ethidium bromide, bands representing the desired fragments were visible under UV light. The gel was sliced above and below each DNA band, and a strip of Whatman DE 81 (DEAE cellulose) paper, folded to double thickness, was inserted into each slot. The gel was run an additional 90 min at 100-200 V. The paper strips were removed from the gel, and the DNA was eluted three times in 300 µl of 1 M NaCl at room temperature in the dark for 30 min. each time. Ethidium bromide was extracted with 2 volumes of isoamyl alcohol, and the DNA was precipitated with ethanol. The initial DNA isolate was subsequently run on a 4% polyacrylamide gel and stained with ethidium bromide. The desired band of DNA was located by UV light and cut from the gel with a scalpel. The excised piece of polyacrylamide was pushed through an 18 gauge needle and the DNA shaken overnight at 37°C with an equal volume (~2 ml) of 500 mM NaCl, 100 mM Tris (pH 8.0), and 5 mM EDTA. The sample was then centrifuged, the supernatant fraction collected, and the pellet rinsed with an extra 0.5 volume of the same elution buffer. The pooled supernatants were filtered through a 0.22 µm Millipore filter, and the DNA precipitated with 2 volumes of ethanol.

Clone construction, screening of recombinant bacteria and plasmid purification

End-filling of the isolated Xba I fragments was performed using a dNTP nucleotide mix and Klenow ("large") fragment DNA polymerase I in a solution of 100 mM Tris (pH 7.4), 75 mM $MgCl_2$, 10 mM DTT, and 0.5 mg/ml BSA. The reaction proceeded for 30 min at room temperature. Ligation of the blunt-ended fragments into a Hind III cut pBR322 plasmid vector (also blunt-ended by the end-filling procedure) was performed according to Maniatis et al. (17). The Xba I fragments were first

phosphorylated with T4 polynucleotide kinase in a reaction mix of 66 mM Tris (pH 7.6), 1 mM ATP, 1 mM spermidine, 10 mM $MgCl_2$, 15 mM DTT, and 0.2 mg/ml BSA for 1 hr at 37°C. The phosphorylated fragments were then ligated to the plasmid vector with T4 DNA ligase in the same reaction mixture for 6 hr at room temperature. Recombinant plasmids were transformed into Escherichia coli RR1 (18) and transformants were first screened for tetracyclene sensitivity and then grown on Millipore filters on agar containing ampicillin and chloramphenicol (250 μg/ml). The filters were then hybridized with ^{32}P-labeled Xba I repeat fragments isolated from genomic DNA to screen for positive recombinants. Positive transformants were grown in large scale by following the conventional procedure (17).

Restriction digests of genomic DNA

Restriction endonucleases Xba I, Eco RI, Hind III, Kpn I, Hinf I, Pst I, and Bam HI were purchased from Bethesda Research Laboratory. Additional samples of Xba I were purchased from New England Biolabs and Collaborative Research Labs. DNA digestions with various restriction enzymes were carried out with 4 U enzyme/ μg DNA at 37°C for 90 min. A buffer of 20 mM Tris, pH 7.4, 5 mM $MgCl_2$, 10 mM DTT, 10 mg/ml BSA, and 50 mM NaCl was used in each case.

Southern blot transfer

DNA fragments were size fractionated by electrophoresis on 1.2% agarose horizontal slab gels and stained with ethidium bromide at 0.5 μg/ml. Prior to transfer, the DNA in the gel was partially depurinated for 5 min. in 0.25 M HCl, denatured for 45 min in 0.2 N NaOH, 0.6 M NaCl, and neutralized with 1 M Tris (pH 7.4), 0.6 M NaCl. Transfer was performed overnight according to Southern (18) using a reservoir of 20xSSC (3 M NaCl, 0.3 M Na_3citrate, pH 6.9), wicks consisting of a double layer of Whatman 3 MM paper, and a filter of nitrocellulose paper soaked in 1xSSC.

^{32}P labeling of plasmid clones

Recombinant plasmid DNA was nick translated in the presence of (α-^{32}P)-dCTP in a 25 μl reaction volume containing 50 mM Tris (pH 7.8), 7.5 mM $MgCl_2$, 10 mM DTT, and 50 μg/ml BSA. Pancreatic deoxyribonuclease (1 mg/ml) was freshly diluted 10^7-fold, and 1 μl was used in each reaction, along with 2 μl of DNA polymerase I (New England Biolab). Reactions were incubated 90 min at 15°C, terminated by addition of 10 μl of 100 mM EDTA, 10% SDS, and the labeled DNA separated from unincorporated free nucleotides by fractionation on an AG-10D Sephadex column.

Hybridization

DNA-containing membranes in Figures 1 and 2 were prehybridized for 16-24 hours at 68°C in a solution of 5x Denhardt's (0.1% BSA, 0.1% ficoll, 0.1% PVP), 6xSSC. For each filter, 1-2x10^7 cpm of nick translated probe was denatured by

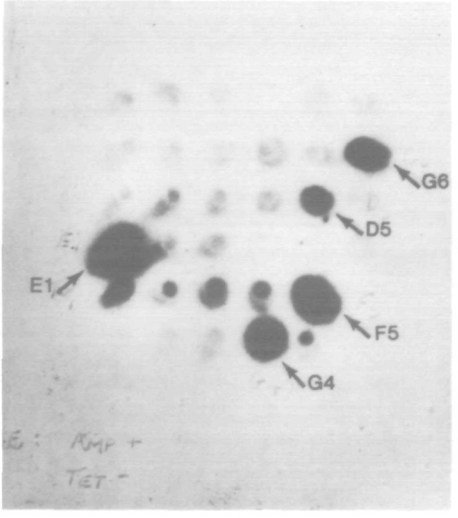

Figure 1. Identification of recombinant clones homologous to a human Xba I repeat. A nitrocellulose filter containing 36 recombinants obtained as described in the text was hybridized with (^{32}P) Xba I fragment 680 bp in length, purified as described in the methods section.

340 bp to 1300 bp. These fragments appeared as bright bands following gel electrophoresis and staining with ethidium bromide. Such multiple copy bands are generally indicative of repetitive sequences.

After isolation from agarose gels, the repeat fragments were cloned into plasmid vector pBR322 which had been cut with Hind III. Insertion into the unique Hind III site on pBR322 interrupts the region encoding tetracycline resistance. The ampicillin resistance gene of the plasmid was left intact so that bacterial colonies successfully transformed with these recombinant clones were ampicillin resistant and tetracyclene sensitive. The Hind III cleavage of pBR322 was well suited for this study because the ligation of blunt-ended Xba I and Hind III cuts results in the regeneration of functional Xba I restriction sites, permitting release of a complete insert from the final clone.

Transformation of E. coli RRI cells with the ligation mix produced 384 transformant colonies, 34 of which showed the proper complement of antibiotic resistance. These 34 colonies were grown on nitrocellulose filters and hybridized with the Xba I repeat DNA isolated from the agarose gel digests of total DNA to screen for desired recombinants. Previous hybridizations had demonstrated that the 682 bp fragment was of sufficient purity for use as probe for screening clones. Six of the 34 recombinant clones were positive for the Xba I repeat (Fig. 1) and two of

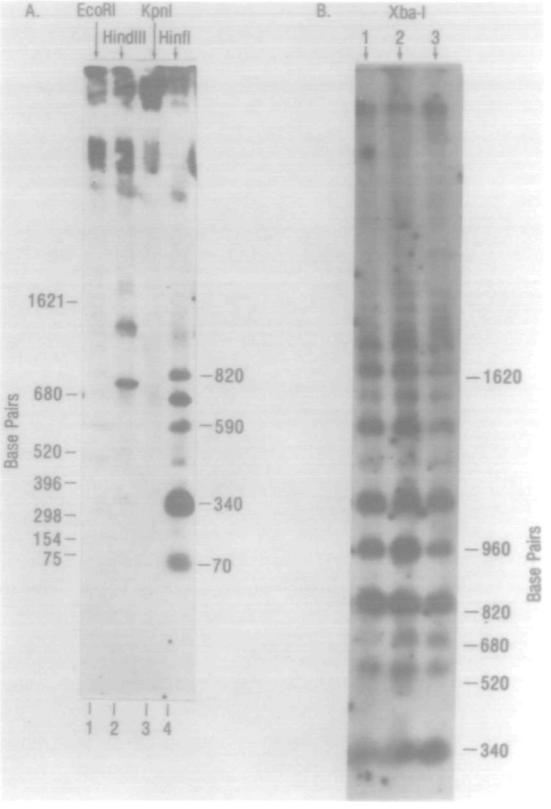

Figure 2. Hybridization of digested human DNA with (^{32}P) pE1. 5 µg human DNA, prepared as described, was digested with the restriction enzymes indicated, electrophoresed through a 1.2% agarose gel, transferred to a nitrocellulose filter according to Southern (18) and hybridized with the (^{32}P) Xba I repeat clone pE1. The lengths of the hybridizing fragments were determined by using Hinf I digested pBR322 as size markers.

these (E1 and F5) were grown for nucleotide sequence analysis.

Hybridization of clones E1 and F5 to Xba I digests of human genomic DNA (Fig. 2) showed preferential binding to DNA fragments with length multiples of ~340 bp, with strongest hybridization to fragments of 680 bp, 1020 bp and 1360 bp. There was little or no hybridization with the 340 bp fragment generated by Xba I digestion. There was unusually strong hybridization to DNA fragments of 850 bp and 1870 bp (see also Fig. 5), both of which coincidentally correspond to positions of two interspersed repeats generated by cleavage with Xba I (8).

The cloned Xba I repeat fragments were also used as probes in hybridization with genomic DNA cleaved with restriction endonucleases Eco RI, Hind III, Kpn I,

boiling for 5 min and added to a 5 ml volume of 5x Denhardt's, 6xSSC, 0.5% SDS, and 1 mM EDTA. This hybridization mix was then added to the filter in a sealable plastic bag and placed at 68°C overnight. The hybridized filters were washed four times at room temperature in 2xSSC, 0.1% SDS for 5 min each time, then washed three times in 0.1xSSC, 0.1% SDS for 30 min each time at 42°C before exposing to X-ray film at -80°C.

DNA-containing membranes in Figures 4 and 5 were prehybridized for 16-24 hrs at 42°C in a solution of 1x Denhardt's. For each filter ^{32}P labeled probe was denatured by boiling for 10 min and added to formamide hybridization cocktail (50% formamide, 6xSSC, .05 M $NaPO_4$, pH 7, 1% SDS). This hybridization mix was added to the membrane in a sealable bag and placed at 42°C overnight.

Nucleotide Sequence Analysis

Repeat fragments were released from selected recombinant clones by digestion with Xba I. Digested plasmid DNAs were separated by electrophoresis through a 4% polyacrylamide gel, and the 680 bp bands were isolated as described above. The DNA was eluted overnight at 37°C into 0.5 M NaCl, 0.05 M Tris-HCl (pH 7.3), 0.01 M EDTA, and then precipitated with ethanol. The purified inserts were ligated into the Xba I site of bacteriophage M13 mp10 (19) and sequenced by the dideoxy-chain termination method (20).

In situ Hybridization

Recombinant plasmid pE1 was ^{3}H-labeled by the method of Lai et al. (21) in the presence of ^{3}H-dATP (51.1 Ci/mmol) and ^{3}H-dTTP (87.5 Ci/mmol) (New England Nuclear Corp.) to a specific activity of 1.7×10^{7} cpm/μg DNA. Free nucleotides were removed from labeled plasmid DNA by column chromatography.

Human mitotic figures were prepared from whole blood cultures synchronized with 0.1 μM methotrexate as described (22). Metaphase preparations were stored at room temperature for one week and then used within 14 days for hybridization experiments.

In situ hybridization experiments were performed by the method of Harper and Saunders (23). Slides were hybridized with ^{3}H-labeled pE1 in the presence of 10% dextran sulfate at concentrations of 2.5 ng/ml and 6.2 ng/ml. Hybridizations were carried out at 37°C for 10-12 hours and rinsed in 50% formamide, 2xSSC at 39°C. Slides were then coated with Kodak NTB2 nuclear track emulsion and exposed for 3 days to 3 weeks at -80°C. Exposed slides were developed in Kodak Dektol at 15°C and G-banded with 0.25% Wright stain (22).

RESULTS

The digestion of human genomic DNA with restriction endonuclease Xba I released amplified copy numbers of certain fragments of DNA ranging in size from

```
          Xba
                  10          20          30          40          50
Xba-1   TCTAGAGAGA AGCATTCTCA GGAACTTCTT TGTGATGTTT GCATTCACGT
Xba-2   **A***C*** ********** *A********C ********** *******AC*
Eco RI: -*****C*** **A******* *T******C* ****T***G* *T*****AC*
Monkey: A**G**A*** ***T****G* *A****G**C ****T*C*G* TA*****TC*

                  60          70          80          90          100
Xba-1   CACAGAACTG AACATTCCCT TTCATAGAGC ATGTTTGAAA CACTCTTTCT
Xba-2   **T***GT** ****C*T**C ***G**C*** *G******** ********T*
Eco RI: ******GT** ***GA***T* *A**C***** *GAC****** ********T*
Monkey: ******GT*A C*TC**T**C ****AGA*** C*T*CGCT** GG**G**CT*

                  110         120         129         139         149
Xba-1   GTAGTATCTA CAAACGGACA TTTCAAA-CG CTTTCAGGCC TATGGTGAGA
Xba-2   ***AC**T*G G**GT***** ***GC*-G** ****G***** ****T***A*
Eco RI: **G*A**T*G ***GT***G* *****GC-** ****G***T* A*****AGA*
Monkey: **G*A**TGG ****G***T* ***GG**G*C *A*AG***G* ********A*

                  159         169
Xba-1   AAGGAAATAT CTTCAAATAA AA
Xba-2   ********** ****TCC*** **
Eco RI: T********* ****CT***G **
Monkey: ********** ****CGT*C* **


                  181         191         201         211         221
Xba-1   ACTAGACAGA AGCATTCTCA GAAACTTATT TGCGATGTGT GTCCTCAACT
Xba-2   **C******* ********** *******CC* **T******* **A***G*G*
Eco RI: ********** *TG******* ******CC** **T******* *CGT******
Monkey: ***G**A*** ***T****G* ******GC*C **T*T*C*** TAAT***T**

                  231         241         251         261         271
Xba-1   AACAGAGTTG AACCTTTCTT TTGATACAAC ATTTTGGAAA CACTCTTTTT
Xba-2   ********** ******C*** ****C*G*G* *G***T***G **********
Eco RI: C********T ********** **C***G*G* *G**A***** ******G***
Monkey: C********A C*T*****CC **C*AGA*G* C***C*CT** GG**G**C**

                  281         291         301         310         319
Xba-1   GTAGAATCTG CAAGTGGATA TTTGAATAGC TTT-GAAGGT TTCGTTGGA-
Xba-2   ********** ********** ***TG***C* ***-**G*A* *********C
Eco RI: ***A*G**** ********** **CAG*CCTC ***-**G*CC *********A
Monkey: **G****TG* ***AG***** ****G*AGC* CA*A**G**C *AT*G**A*A

                  329         339
Xba-1   ACGGGAATAT CTTCATATAA AA
Xba-2   *****-**** *G******** **
Eco RI: *****-**T* ********T* TG
Monkey: *A**A***** ****CGT*C* **
```

Figure 3. Comparison of nucleotide sequences of Xba I repeat clone pEl, Eco RI repeat family, and monkey component α. For each position, an asterisk (*) represents the same residue found in the first half (Xba-I) of the 682 bp Xba I repeat clone, and a dash (-) signifies a space inserted to allow the best possible match with the Xba-I sequence.

Hinf I, Pst I, and Bam HI (as a control), which are known to release other human repeat sequence families. The results indicated a strong homology between the Xba I repeat clones and certain sizes of fragments in the Hinf I and Hind III digests, but little or no homology with repetitive fragments produced by the other restriction enzymes.

Sequence analysis revealed more about the possible relatedness of the Xba I repeats to known repeat sequence families. The sequence of the 682 bp fragment in clone pEl (Fig. 3) shows it to be composed of four tandemly arranged subunits (labeled A through D, respectively) with lengths of 171, 170, 171, and 170 bp. Each subunit begins with an Xba I restriction site (TCTAGA) which has, except in the case

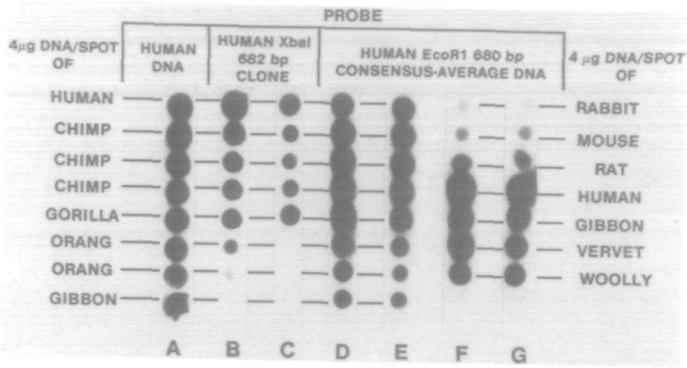

Figure 4. Phylogeny of Xba I 682 bp repeat. 4 μg of human, chimpanzee, gorilla, orang-utan, and gibbons DNA (A-E) and 4 μg of rabbit, mouse, rat, human, gibbon, vervet, and woolly monkey DNA (F,G) were spotted onto nylon membranes in NaCl and baked at 80°C for 2 hr in vacuo. A, hybridized at 42°C with nick translated human genomic DNA probe; B, same filter washed clean and hybridized at 42°C with nick translated Xba I 682 bp repeat clone probe; C, filter from B washed for 2 hr at 65°C in 0.1 x SSPE (15 mM NaCl, 1.15 mM Na_2HPO_4, 125 μM EDTA); D, F, hybridized at 42°C with nick translated gel purified human Eco RI 680 bp DNA fragment probe; E, G, filter from D wash for 2 hr at 65°C in 0.1 x SSPE.

of the first subunit, one or two base changes, rendering the site non-functional. This opening sequence is followed in each case by five base pairs, and then an Eco RI site containing a single base mutation. The average deviation in sequence homology between adjacent subunits is greater than 32%, but the variation between alternating subunits is much lower: 21% in the case of subunits A and C, and 19% for subunits B and D. The sequence of a second 682 bp Xba I repeat clone (pF5) differed from that of the first (pE1) by 15 nucleotides; 2 in the first subunit, 1 in the second, 1 in the third, and 11 in the fourth. The sequence organization of this Xba I repeat bears a striking resemblance to the Eco RI human repeat family, which has a 340 bp repeat and is composed of two tandem subunits of 169 and 171 bp. This sequence also occurs as a 680 bp tetramer in which adjacent subunits show 27% base sequence variance with each other. Comparison of the Eco RI tetramer to the Xba I repeat shows a sequence divergence between the two of 21.2%. A comparison of the published sequence of the monkey component α (Fig. 3) and that of the human Xba I repeat revealed a sequence divergence ranging from 30.4% to 36.5% variance between component α and each of the four Xba I 680 bp repeat subunits.

To study the phylogenetic distribution of sequences complementary to the Xba I 682 bp repeat clones, 4μg of genomic DNA from human, chimpanzee, gorilla, orang-utan, and gibbon were spotted onto nitrocellulose filters. Hybridization of this immobilized DNA with radioactive human genomic DNA probe resulted in an

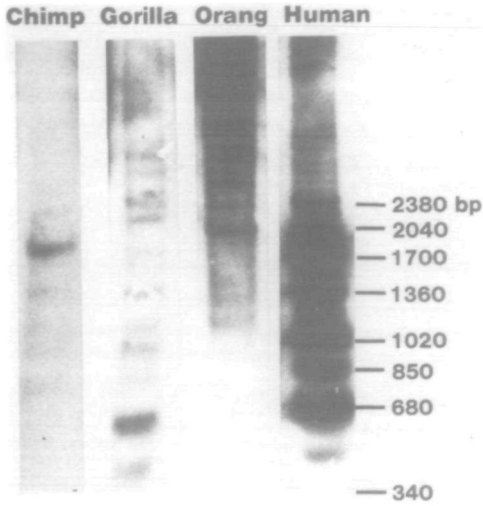

Figure 5. Xba I 682 bp satellite sequence organization. Approximately 0.5 μg of Xba I digested chimpanzee, gorilla, orang-utan, and human genomic DNA's were electrophoretically fractionated, transferred to nylon filter and hybridized with nick translated Xba I 682 bp repeat clone. High molecular weight ape DNAs were graciously donated by Dr. Oliver Ryder.

equal hybridization signal to each DNA spot (Fig. 4A). The filter was then washed clean of probe and rehybridized under the same stringency as the first probe, using the Xba I 682 bp repeat clone as probe (Fig. 4B). Hybridization was strongest with human DNA, evenly strong with the chimpanzee and gorilla DNA, weak with orangutan DNA and absent to the gibbon spot. No hybridization was obtained with DNA from Old or New World monkeys (not shown). The filter was then washed at a high temperature and a low salt concentration for two hours which caused the hybridization signal with orang-utan to disappear. This procedure also revealed a stronger signal with gorilla DNA than with chimpanzee DNA (Fig. 4C).

For comparison a ^{32}P-labeled gel purified human Eco RI 680 bp DNA probe was hybridized to the same DNA preparations under the same conditions as with the Xba fragment probe. A strong hybridization signal was observed over each DNA spot (Fig. 4D). The hybridization signals decreased uniformly with all DNAs after a low salt, high temperature wash under the exact conditions of 4C (Fig. 4E). To determine the limit of homology of the Eco RI fragment to mammalian genomic DNA, the gel purified probe was hybridized to membranes spotted with total DNA from rabbit, mouse, rat, human, gibbon, vervet, and woolly monkey under the same conditions as the previous experiment. A strong hybridization signal was observed with DNA from the primates, a weak signal with DNA from rat and mouse, and a

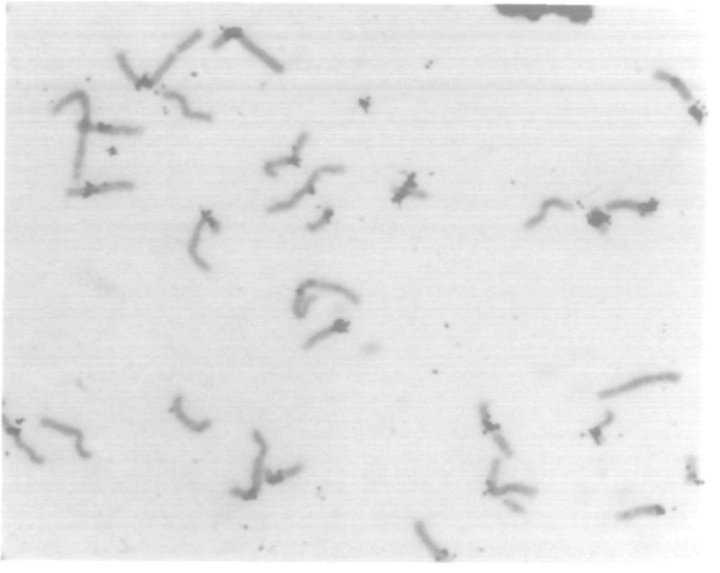

Figure 6. Human metaphase cell hybridized with the ^{3}H-Xba I probe pEl, at a concentration of 0.1 µg/ml for 10 hr and exposed to autoradiographic emulsion for 21 days. Overexposure to emulsion was used to illustrate clustering of silver grains on the centromeres of chromosomes . Cells used for analysis were hybridized with a lower probe concentration and exposed for a shorter period of time.

very weak signal with rabbit DNA (Fig. 4F). A low salt high temperature wash of this filter again revealed a uniform melting of all hybrids (Fig. 4G). In a comparable experiment with the human Xba I 682 bp clone El, a hybridization signal was obtained only with the human DNA (data not shown).

To further compare the Xba I repeat sequence organization in humans, chimpanzees, gorillas and orang-utans, Xba I digested genomic DNAs from all four animals were electrophoretically fractionated and transferred to a nitrocellulose filter. This filter was then hybridized with Xba I 682 bp repeat clone pEl. The autoradiogram (Fig. 5) revealed a different hybridization pattern in each case. The chimpanzee DNA lane showed a barely visible ladder with the strongest hybridization signal over a band migrating at about 1700 bp. The gorilla DNA lane showed a similar pattern with a very weak signal at 1700 bp and a strong signal at 680 bp. The orang-utan DNA lane showed a different pattern with the strongest signal at about 2040 bp.

Chromosomal localization by in situ hybridization demonstrated that sequences corresponding to the cloned Xba I 682 bp fragment are located in the centromeres of all human chromosomes except the Y (Fig. 6). Slides hybridized at 25 ng/ml and exposed for 3 days exhibited an average of 10 labeled sites per cell. Clusters of

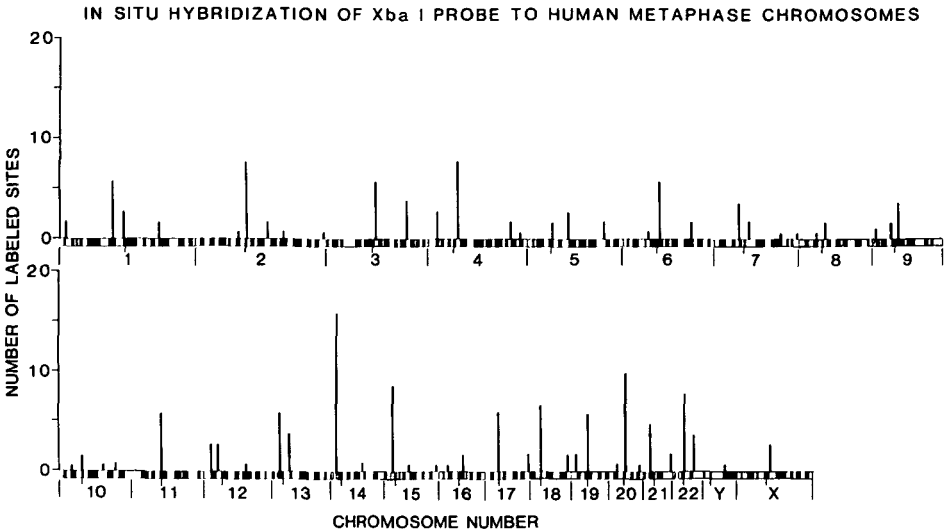

Figure 7. Localization of Xba I 682 bp repeat by in situ hybridization. Metaphase chromosomes prepared from peripheral blood cultures of a human male were hybridized with pE1. Analysis of 210 labeled sites indicates that the Xba I 682 bp repeat is localized on the centromeres of all human chromosomes except the Y, with increased labeling on the D and G group chromosomes. Small lines below each chromosome denote the position of the chromosome centromeres.

silver grains were typically observed at each labeled site. Out of 210 total labeled sites analyzed, 64% were located at the centromeres of chromosomes. In addition, over half (43%) of centromeric label was observed on D and G group chromosomes, suggesting that the centromeres of these chromosomes contain more copies of the 682 bp Xba I fragment than other groups. The remaining labeled sites were widely scattered over chromosomal regions as shown in Fig. 7. The abscissa, composed of the human haploid karyotype including the X and Y chromosomes, was constructed by placing ideograms of the chromosomes head to tail followed by division into 160 equal segments. The number of labeled sites was plotted along the ordinate. These data suggest that the 682 bp Xba I repeat is a centromeric satellite in the human genome. No significant labeling was observed at the telomeres of chromosomes as has been observed with other α-satellite sequences.

DISCUSSION

Electrophoretically fractionated, Xba I digested human DNA stained with ethidium bromide revealed visible, regularly spaced bands upon exposure to UV light. The banding pattern was indicative of a tandemly organized repeat family and resembled the human Eco RI satellite family by the size of the fragments

generated. Conflicting data has been presented in the past as to the relationship that these two repeat families hold to each other. Studies on a cloned member of the Xba I repeat family have clarified some of the apparent discrepancies.

The cloned 682 bp DNA fragment from Xba I digested human DNA was characterized as a member of the human α-satellite family by the following criteria. First, the cloned fragment hybridized to the ladder bands of Xba I digested human DNA which were electrophoretically separated. Second, sequence analysis revealed a structural organization remarkably similar to that of the Eco RI α-satellite family (16). The 682 bp Xba I fragment is composed of four subunits of 171, 170, 171, 170 bp, respectively, with adjacent subunit sequence divergence of greater than 32% and alternating subunit sequence divergence of about 20% suggesting a I-II-I-II subunit composition. Third, the nucleotide sequence of the Xba I 682 bp cloned DNA matched the consensus Eco RI α-satellite sequence in about 80% of the positions. Finally, the 682 bp Xba I repeat sequence was localized to chromosomal centromeres. These properties are those expected of a true α-satellite sequence.

The Eco RI and Xba I 682 bp α-satellites differ in their phylogeny and chromosomal locations. At high stringency radioactive Eco RI α-satellite hybridized with α-satellite DNA in gibbons, apes and humans but not significantly with α-satellite from old and new world monkeys (21). Related sequences were detectable in all monkeys by molecular hybridization at moderate stringency. Labeled Xba I 682 bp satellite hybridized only with human α-satellite DNA (Fig. 5). Related sequences were found in great apes but not in gibbons or monkeys. The related sequences in apes were in low copy number in orangutans, higher copy number in chimpanzee, higher still in gorillas and highest in humans (Fig. 4). By in situ hybridization, Eco RI α-satellite sequences were localized primarily to chromosomes 19,7,1,10,3,6 and 5 (25) while Xba I 682 bp α-satellite sequences hybridized mainly with the centromeres of chromosomes 14,20,22,15,2,4 and 18.

Conceptually, we define a satellite DNA "family" as a group of sequences all derived from the same amplification event or template. In practice this is not a measurable quantity. There is little doubt that the appearance of alpha satellite sequences has occurred suddenly and repeatedly during primate evolution (6-8). When such amplifications occurred is measurable by scoring the phylogenetic distribution of given sequences. However, the results depend upon the scoring methodology. Phylogenetic distributions can be scored by measuring the production of DNA fragments of certain lengths (about 170 bp) by restriction endonucleases, i.e. restriction fragment length measurements. Satellite DNA fragments produce by a given restriction endonuclease represent a collection of sequences derived from

one or more amplifications using templates with the given restriction sites. The most ancestral amplification is scored. Thus, the original amplification leading to the Eco RI alpha satellite family or families occurred after apes and gibbons diverged from monkeys while the original amplification leading to the Xba I alpha satellite families occurred after great apes and gibbons diverged from each other (8).

Phylogenetic distributions scored by measuring hybridization of a "consensus-average" probe comprised of uncloned repeat fragments to total DNA of other animals using moderate hybridization stringency measure in addition, related sequences which resulted from other amplifications. This method has invariably yielded broader phylogenetic distributions with a phylogenetic gradient indicative of earlier amplifications. Thus, the Eco RI consensus-average probe hybridized significantly with DNA from primates and even detected related sequences in rodents.

Similar hybridization measurements using cloned probes and the same moderate hybridization stringencies often produce a narrower phylogenetic distribution than the restriction fragment length measurements, as was the case with the Xba I 682 bp cloned probe reported here. Presumably, the sequences detected with this probe were a subset of the Xba I alpha satellite families which had been amplified since great apes and gibbons diverged or since orang-utans diverged from man and the remaining great apes.

Finally, hybridization of a cloned probe to restriction fragments of DNA yield the narrowest phylogenetic distribution of all. Under moderate hybridization stringency the Xba I 682 bp alpha satellite DNA probe detected a ladder of 170n bp fragments only in human DNA and not in DNA of great apes. We speculate that this result signifies that an amplification of one or more Xba I alpha satellite DNA families occurred in the human lineage since humans diverged from great apes. It would follow that an amplification of alpha satellite DNA related to the Xba I 682 bp probe also occurred at a similar time in the chimp-gorilla lineage.

In conclusion, these experiments clearly distinguish the human Xba I and Eco RI satellite DNA components as resulting from separate amplifications and suggest further that the Xba I satellite DNA component is heterogenous, having resulted from several amplifications which occurred after the establishment of the ape lineage.

ACKNOWLEDGEMENT

This research was supported by grants from the Leukemia Society of America, Inc., American Cancer Society (CD-223), the Robert A. Welch Foundation (G-267), and National Institute of Health (CA 20124, GM 14368 and CA 29545). J.W.W. is supported by a NIH NRSA (1 F2 AM 07091-01). W.T.S. is a Rosalie B. Hite Predoctoral Fellow.

REFERENCES

1. Britten, R.J. and Kohne, D.E. (1968) Science **161**:529-540.
2. Singer, M. (1982) Int. Rev. Cytol. **76**:67-112.
3. Rubin, C.M., Houck, C.., Deininger, P.L., Friedman, T. and Schmid, C.W. (1980) Nature (London) **284**:372-374.
4. Adams, J.W., Kaufman, R.E., Kretschmer, P.J., Harrison, M. and Nienhuis, A.W. (1980) Nucl. Acids Res. **8**:6113-6127.
5. Southern, E.M. (1970) Nature (London) **227**:794-796.
6. Musich, P.R., Brown, F.L. and Maio, J.J. (1980) Chromosoma **80**:331-348.
7. Donehower, L. and Gillespie, D. (1980) J. Mol. Biol. **134**:805-834.
8. Gillespie, D., Adams, J.W., Costanzi, C. and Caranfa, M.J. (1982) Gene **20**: 409-414.
9. Tashima, M., Calabretta, B., Torelli, G., Scofield, M., Maizel, A. and Saunders G.F. (1981) Proc. Natl. Acad. Sci. USA **78**:1508-1512.
10. Houck, C.M., Rinehart, F.P. and Schmid, C.W. (1979) J. Mol. Biol. **132**: 289-306.
11. Maio, J.J., Brown, F.L., McKenna, W.G. and Musich, P.R. (1981) Chromosoma **83**:127-144.
12. Gillespie, D., Donehower, L. and Strayer, D. (1982) in Genome Evolution, eds. Dover, G.A. and Flavell, R.B., eds. (Academic Press, London) pp. 113-133.
13. Strayer, D., Heitz, N., Roeder, R. and Gillespie, D. (1983) Proc. Natl. Acad. Sci. **80**:4770-4774.
14. Deininger, P.L., Jolly, D.J., Rubin, C.M., Friedman, T. and Schmid, C.W. (1981) J. Mol. Biol. **151**:17-33.
15. Rosenberg, H., Singer, M. and Rosenberg, M. (1978) Science **200**:394-399.
16. Wu, J.C. and Manuelidis, L. (1980) J. Mol. Biol. **142**:363-386.
17. Maniatis, T., Fritsch, F.E. and Sambrook, J. (1982) Molecular Cloning (Cold Spring Harbor eds.).
18. Southern, E.M. (1975) J. Mol. Biol. **98**:503-517.
19. Messing, J. (1983) Methods in Enzymology **101**:20-78.
20. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. **74**: 5463-5467.
21. Lai, E.C., Woo, S.L.C., Dugaiczyk, A. and O'Malley, B.W. (1979) Cell **16**: 201-211.
22. Yunis, J.J. and Chandler, M.E. (1977) in Progress in Clinical Pathology Stefanini, M., Hossaine, A.A., and Isenberg, H.D., eds., (Grune and Stratton, New York), vol. **7**, pp. 267-288.
23. Harper, M.E. and Saunders, G.F. (1981) Chromosoma **83**:431-439.
24. Maio, J.J., Brown, F.L. and Musich, P.R. (1981) Chromosoma **83**:103-125.
25. Manuelidis, L. (1978) Chromosoma **66**:23-32.