Different subfamilies of alphoid repetitive DNA are present on the human and chimpanzee homologous chromosomes 21 and 22

A.Lund Jørgensen, Carol Jones¹, C.J.Bostock² and A.Leth Bak

Institute of Medical Microbiology, The Bartholin Building, University of Aarhus, DK-8000 Aarhus C, Denmark, ¹Eleanor Roosevelt Institute for Cancer Research, Inc., 1899 Gaylord Street, Denver, CO 80206, USA, and ²Institute for Animal Disease Research, Pirbright Laboratory, Pirbright, Woking, Surrey GU24 0NF, UK

Communicated by A.P.Bird

The alphoid repeat DNA on chimpanzee chromosome 22 was compared with alphoid repeat DNA on its human homologue, chromosome 21. Hybridization of different alphoid probes under various conditions of stringency show that the alphoid repeats of chimpanzee chromosome 22 are not closely related to those of human chromosome 21. Sequence analysis of cloned dimer and tetramer *EcoRI* fragments from chimpanzee chromosome 22 confirm the low overall level of homology, but reveal the presence of several nucleotide changes which are exclusive to the chromosome 21 subfamily of human alphoid DNA. Southern blot analysis of alphoid repeat DNA on the chimpanzee X chromosome suggests this subfamily has been strongly conserved during and since the separation of chimpanzee and man although the two subfamilies can be distinguished on the basis of *Taq I* restriction fragments.

Key words: Evolution/human/primate/repeated sequences/satellite DNA

Introduction

The genomes of primates contain a characteristic family of repeated DNA sequences, called the alphoid repeats, which are found to be localized mainly in pericentromeric regions of chromosomes and, in the case of human, to comprise a few per cent of the total genome. The monomeric unit of alphoid DNA is 170 bp, although many families of alphoid repeats are based on arrays of dimeric 340-bp units in which there are two distinguishable monomeric sequences. Related monomers and dimers are arranged in tandem in arrays with lengths of between 2 and 5.5 kb. Particular arrays seem to predominate on different chromosomes and constitute higher-order units of alphoid DNA characteristic for individual chromosomes. Several hundred-fold amplification of a particular higher-order unit on a chromosome explains the existence of different amplification units and varying amounts of alphoid DNA on the different individual chromosomes of a primate chromosome set. Several alphoid repeat subfamilies have now been identified and assigned to individual chromosomes (Reis *et al.*, 1985; Waye and Willard 1985, 1986a,b; Wolfe *et al.*, 1985; Devilee *et al.*, 1986; Jørgensen *et al.*, 1986, 1987).

Since the structure of an alphoid sequence tends to be characteristic for a particular chromosome, comparison of the alphoid sequences on homologous chromosomes can give some insight into the rates of formation of alphoid subfamilies and their evolutionary relationships. We describe here an alphoid repeat from chimpanzee chromosome 22 and compare it with the alphoid

repeat of human chromosome 21, its human homologue (Seuánez, 1979). The subfamilies of these two chromosomes are found to be distinctly different which indicates that the recombination events which initially created the specific arrays of alphoid repeats on each chromosome, and the subsequent processes which amplified the subfamily, took place during or after separation of the two species.

Results

Following digestion with several different restriction enzymes, total human genomic DNA and DNAs from several hybrid cell lines were blotted and hybridized either at low stringency with a cloned alphoid repeat — $\alpha R1-6$ (Jørgensen *et al.*, 1986, 1987) — representing an average alphoid repeat sequence, or at high stringency with an alphoid repeat — $\alpha R1(680)13,5$ (Jørgensen *et al.*, 1987) — representing the alphoid subfamily on human chromosome 21 (Figure 1). Each hybrid cell carried a single human (X,7,13,21 or 22) or chimpanzee (22) chromosome. The hybridization patterns at both low and high stringency shown in Figure 1 indicate that human chromosome 21 carries only one detectable subfamily. In contrast to the results with human chromosome 21, the low stringency hybridization to *EcoRI* digests of DNA from the chimpanzee chromosome 22 (Figure 1a) reveals a typical alphoid 340-bp ladder of strongly hybridizing fragments. These bands are not detectable in the chimpanzee chromosome 22 lane following hybridization at high stringency with the human chromosome 21 probe. This indicates that the alphoid sequences present on chimpanzee chromosome 22 are not closely related to those on the homologous human chromosome 21. Hybridization under the same conditions of high stringency to digests of total genomic chimpanzee DNA did not reveal any band patterns characteristic of the human chromosome 21 subfamily (data not shown).

Sequence analysis of cloned *EcoRI* fragments of chimpanzee chromosome 22 alphoid DNA confirms the low level of homology between the alphoid subfamilies of human chromosome 21 and chimpanzee chromosome 22. Figure 2A displays the sequence of six individual 340-bp dimers and six individual 680-bp tetramers isolated and cloned from the corresponding *EcoRI* restriction bands of chimpanzee chromosome 22 DNA. The sequences have been aligned to allow maximum homology to the 340-bp sequence of the human average alphoid repeat (Wu and Manuelidis, 1980; Jørgensen *et al.*, 1986, 1987). The sequences of all the tetramers are nearly identical, each consisting of two distinct, but clearly related, dimers identified as I and II. The sequence of the tetrameric dimer I is also found in five out of the six sequences of 340-bp dimer fragments and dimer II is represented in the sequence of the sixth 340-bp dimer fragment. Thus the *EcoRI* dimers and tetramers are clearly part of the same subfamily. Presumably the tetramers were derived from a part of the subfamily in which the first nucleotide in the first monomer section of dimer II was deleted, thereby abolishing the *EcoRI* site.

Together the sequences of dimers I and II contain 60 subfamily-

Fig. 1. Hybridization at low (left) and high (right) stringency of alphoid repeat DNA in Southern blots of different DNAs digested with various restriction endonucleases: (a) *EcoRI*; (b) *XbaI*; (c) *TaqI*; (d) *BamHI*. The lanes contain as indicated from left to right, 0.2 μ g of total human genomic DNA, 2 μ g of DNA from the hamster/chimpanzee hybrid cell 588-3a-9 containing the chimpanzee chromosome 22 as the only chimpanzee chromosome, 2 μ g of DNA from the mouse/human hybrid cells PGMe25/8, ThyB1-33-6, PGMe25/9, C121, ThyB1-33-12 containing human chromosome 22, 21, 13, 7 and X, respectively, as the only human chromosome. The numbers on the left refer to molecular sizes in bp. Hybridization at low stringency (30% mismatch) was done with 32 P-labelled α -RI-6 probe prepared from a recombinant M13 clone, which has a sequence nearly identical to that of the average alphoid repeat sequence. Hybridization at high stringency was done with 32 P-labelled α -RI(680)13-5 probe prepared from a recombinant M13 clone which has a sequence representative of the alphoid repeat subfamily on human chromosome 21.

Fig. 2. Nucleotide sequences of clones of alphoid repeats from the chimpanzee chromosome 22 (A). Six dimeric alphoid repeat units [CaRI(340)-23, -33, -72, -110, -130, -168] were isolated from *EcoRI*-digested DNA of chimpanzee chromosome 22 DNA by excision from gels and cloning in phage M13 of the 340-bp *EcoRI* band (Figure 1a). Six tetrameric alphoid repeat units [CaRI(680)-37, -63, -86, -100, -104, -130] were isolated by excision from gels and cloning in phage M13 of the 680-bp *EcoRI* band (Figure 1a). Each tetramer consists of two dimers (I and II). The complete sequence of the dimers and tetramers was obtained by sequencing from both ends of the units using the dideoxy chain-termination method. Asterisks indicate identical nucleotides, and a dash (—) indicates a deletion of a nucleotide, relative to an average human alphoid repeat sequence, av.αRI seq. (Wu and Manuelidis, 1980; see text). An insertion of an A has occurred between position 92 and 93 in the first monomer of dimer II in two tetramers. A deletion of A at position 1 in dimer II of the tetramers abolishes the characteristic *EcoRI* site. (B) Consensus sequence of each of the two dimers in the human chromosome 21 alphoid subfamily, taken from Jørgensen *et al.* (1987). The reference sequence (av.αRI seq.) is the same as in (A).

specific changes, i.e. identical nucleotide changes at the same positions in two or more dimer sequences. Only 22 of these changes are represented among the 108 changes which define the human chromosome 21-specific subfamily. Of these 22, six are found as subfamily-specific changes *exclusively* in the human chromosome 21 subfamily. These are A in 1'-monomer position 45, G in 1'-monomer position 66, T in 1'-monomer position 101, A in 1'-monomer position 132, G in 1'-monomer position 162 and A in 2'-monomer position 14. In addition two,

C in 1'-monomer position 31 and A in 1'-monomer position 124, are found exclusively in the more distantly related human chromosome 22 subfamily. None of the chimpanzee chromosome 22 subfamily-specific changes are found to be present as specific changes exclusive to any of the unassigned human alphoid subfamilies identified previously (Jørgensen *et al.*, 1986); the human chromosome 7 subfamily (Jørgensen *et al.*, 1986) or human X or 17 subfamilies (Waye and Willard, 1986a,b). From the accumulating sequence data from the increasing number of iden-

Fig. 3. Hybridization at increasing stringency (a–c) of alphoid repeat DNA to Southern blots of *Bam*HI digest of human and chimpanzee X chromosomal and total genomic DNAs. **Lane 1:** 2.5 μ g of DNA from the hamster/chimpanzee hybrid cells 800-3A containing the chimpanzee X chromosome as the only chimpanzee chromosome. **Lane 2:** 2.5 μ g of DNA from the mouse/human hybrid cells ThyB1-33-12 containing the human X chromosome as the only human chromosome. **Lane 3:** 2.5 μ g of total chimpanzee DNA. **Lane 4:** 2.5 μ g of total human DNA. Numbers to the right refer to molecular size in kb. Hybridization was to a 32 P-labelled probe prepared from a recombinant M13 clone that contains a 2-kb *Bam*HI fragment representative of the alphoid repeat sequences found in the human X-chromosome (lane 2). Hybridization and washing conditions were as follows: (a) 45.5°C in 6 $\times$ SSC; 44°C in 1 $\times$ SSC allowing for 30% mismatch. (b) 75.5°C in 6 $\times$ SSC; 74°C in 1 $\times$ SSC corresponding to T_m –12°C. (c) 75.5°C in 6 $\times$ SSC; 82°C in 1 $\times$ SSC corresponding to T_m –4°C.

tified subfamilies, it has become clear that the so-called average sequence of the human alphoid DNA does not represent the authentic average sequence of all human alphoid DNA. However, comparison of subfamily-specific sequences with a conventional reference does allow easy visualization of sequence homology between subfamilies. The great difference (24%) between the chimpanzee chromosome 22 and the human chromosome 21 is revealed by comparing Figure 2A with the consensus sequence of each of the two different dimers of the human chromosome 21 subfamily (Figure 2B). Hybridization of the human chromosome 21 probe [α -RI(680)13,5] to Southern blots of restriction enzyme-digested total chimpanzee DNA showed that this distinct subfamily is not present elsewhere in the chimpanzee genome (data not shown). Nor did hybridization of one of the cloned chimpanzee 22 alphoid repeats reveal any additional alphoid repeats on chromosome 22.

The low level of overall sequence homology between the alphoid subfamilies on the homologous human 21 and chimpanzee 22 chromosomes, while retaining evidence of common origin, suggests rapid reorganization of alphoid repeat sequences on these

Fig. 4. Hybridization of alphoid repeat DNA to Southern blots of *Taq*I digests of human and chimpanzee X chromosomal and total genomic DNAs. Numbers on the left refer to molecular sizes in bp. The probe and DNAs were as described in the legend to Figure 3, and hybridization conditions were the same as those for Figure 3b.

chromosomes during primate evolution. Similar evidence has been obtained from studies on the alphoid repeats of human and chimpanzee Y chromosomes (Wolfe *et al.*, 1985). Our preliminary studies on the alphoid repeats of human and chimpanzee X chromosomes suggest that these sequences have been much more tightly conserved during primate and hominid evolution. Figure 3 shows the patterns obtained by hybridizing Southern blots of *Bam*HI-digested human and chimpanzee X chromosome DNA with a probe derived from the 2-kb *Bam*HI fragment of human X chromosome DNA (see the fragment at 2040 bp in the low stringency hybridization in Figure 1d). The same patterns are observed in both human and chimpanzee X chromosome DNA, even at high stringency. This indicates that the organization of the X-chromosomal alphoid repeat DNA must be very similar in the two species, and, since hybridization to the homologous X-chromosome alphoid fragments is only favoured at stringencies of T_m –4°C (Figure 3c), the sequence divergence can only be a few per cent. Nevertheless, as Figure 4 shows, the *Taq*I restriction patterns of the two X chromosome DNAs differ. The chimpanzee X chromosome DNA has a single prominent band at 1530 bp, whereas human X chromosome DNA has two prominent bands at 850 and 1020 bp.

Discussion

Our results show that alphoid repeat DNA is subjected to genetic mechanisms which result in the presence of distinct subfamilies on individual chromosomes and that during evolution these subfamilies can either remain largely unchanged, as in the case of the X chromosome subfamily, or can diverge dramatically to the 24% difference seen between the human and chimpanzee homologous chromosomes 21 and 22. It seems unlikely that such

a remarkable sequence difference can be attributed to homogenization by fixation of accumulated point mutations throughout the subfamily, given the low neutral mutation rate of 0.13% per Myr (Britten, 1986; Koop *et al.*, 1986), and an estimated branch-point for the separation of human and chimpanzee of between 6 and 8 Myr ago (Koop *et al.*, 1986). A great diversity, however, may arise as a consequence of the processes (molecular drive; Dover, 1982) leading to the concerted evolution of a given repeat subfamily. One important component in these evolutionary processes is unequal crossing-over. By this mechanism, variant nucleotides can be spread along tandem repeats at a rate much faster than, and independent of, the mutation rate. This has been indicated in studies with other organisms (Strachan *et al.*, 1985). These studies also showed that the degree of diversity of a given repeat family may vary between closely related species, and the turn-over rate of homogenization may vary for different repeat families with distinct chromosomal locations within the same species. Evidence for reorganization also follows from the findings that the nucleotide changes which constitute the pattern defining a given subfamily are often present as identical changes but in different combinations in other subfamilies (Jørgensen *et al.*, 1986, 1987).

The chimpanzee data reported here and the previous studies of the human subfamilies indicate that amplification also plays an important role in the evolution of a repeat subfamily. This follows from the finding of a precise alternating pattern of different repeat units (monomers, dimers) in larger amplification units (*Bam*HI fragments >2 kb, Figure 1d). The presence of partially homogenized changes (changes found in only two or a few repeats) indicates that the creation and fixation of new variant repeats by unequal crossing-over and homogenization are occurring simultaneously.

Periodic reorganization of the genome accompanied by amplification of different families of repetitive DNA has been suggested to occur prior to the emergence of new species (Gillespie *et al.*, 1982). The presence of widely different subfamilies of alphoid repetitive DNA on the human and chimpanzee homologous chromosomes 21 and 22 would appear to support this hypothesis. On the other hand, the close similarity between the 2-kb *Bam*HI alphoid repeats from the X chromosomes from the two species would suggest an ancient origin of the subfamily, before the two species separated. The different distributions of *Taq*I recognition sequences on the two chromosomes would indicate some degree of reorganization after the two species diverged.

Materials and methods

Cell lines

Human chromosome-specific DNAs were extracted from the mouse/human hybrid cells C121, PgMe25/9, Thy B1-33-6, PgMe25/8 and Thy B1-33-12 which contain, respectively, human chromosomes 7, 13, 21, 22 and X as the only detectable human chromosomes or sub-chromosomal fragments (see Lund *et al.*, 1984; Jørgensen *et al.*, 1987, and the references cited therein for a full description of these cells). Chimpanzee chromosome-specific DNAs were extracted from the Chinese hamster/chimpanzee hybrid cells 588-3a-9 and 800-3a, which contain, respectively, chimpanzee chromosomes 22 and X as the only detectable chimpanzee chromosome or sub-chromosomal fragment (Jones *et al.*, 1985). The Giemsa 11 differential staining technique showed a single blue-staining primate chromosome in each of these hybrids and cytogenetic analysis confirmed that these had the correct R- and G-banding patterns for the 22 and X chromosomes. The hybrid 588-3a-9 expresses the chimpanzee isozyme SOD 1, complements the ADE⁻G auxotrophic mutant, expresses the chimpanzee enzyme PFK-L and is positive for several human chromosome 21-specific RFLPs. The identification of the chimpanzee X chromosome in hybrid 800-3a is presently based entirely on cytogenetic criteria, since it is negative for two X-linked isozymes presumably because it was the inactive X chromosome that was retained. Fur-

ther characterization using human X-linked DNA probes is in progress. Total human DNA was isolated from normal human embryonic fibroblasts.

DNA extraction and gel electrophoresis

DNAs were extracted from cell lines as described previously (Jørgensen *et al.*, 1986). They were digested to completion with various restriction enzymes under conditions recommended by the supplier. Restriction enzymes *Xba*I, *Taq*I, *Eco*RI and *Bam*HI, T4 ligase and Klenow DNA polymerase were obtained from Boehringer Mannheim. For the Southern blot analysis shown in Figure 1, 2 µg of hybrid cell DNAs or 0.2 µg of total genomic DNAs were loaded into each slot of a 1.5% agarose gel. For the Southern blots shown in Figure 3, 2.5 µg of both hybrid cell and total genomic DNAs were loaded into the slots of a 1.0% agarose gel.

Isolation of restriction fragments

*Eco*RI fragments (340 and 680 bp) containing the alphoid repeat sequences from chimpanzee chromosome 22 and 2.0-kb *Bam*HI fragments containing alphoid repeat sequences from the human X chromosome were electrophoresed from the agarose gel onto a piece of NA45 membrane (Schleicher and Schuell). DNA fragments were eluted in 1 M NaCl, cloned into phage M13 mp10 vector (Amersham) and those clones carrying the alphoid sequences identified by hybridization as described previously (Jørgensen *et al.*, 1986).

DNA sequencing

Using the dideoxy chain-termination procedure of Sanger *et al.* (1977) the sequence was read on both strands or read unambiguously in two or more independent preparations of single-stranded DNA of each clone to be sequenced.

³²P-labelled DNA probes

The probe used at low stringency in Figure 1 was αRI-6 (see Jørgensen *et al.*, 1986) which represents the average sequence of the alphoid repeat family (Wu and Maniatis, 1980). The probe used at high stringency in Figure 1 was αRI (680) 13-5 which represents the human chromosome 21-specific alphoid repeat subfamily (Jørgensen *et al.*, 1987). The probe used in Figure 3 represents alphoid repeat sequences isolated from the 2-kb *Bam*HI fragment from the human X chromosome (Figure 3, lane 2). The alphoid repeat sequences of the probes were labelled with [α -³²P]dATP (Amersham, sp. act. 3000 Ci/mmol) by chain extension on single-stranded recombinant M13 DNA using sequencing primer (Jørgensen *et al.*, 1986). For the probes used in Figure 1, the double-stranded labelled inserts were excised by *Eco*RI digestion and isolated by gel electrophoresis, as described above. Specific activities of 3–5 × 10⁸ c.p.m./µg DNA were obtained.

Southern blot

Southern blotting to nitrocellulose filters (Schleicher and Schuell) was done as described previously (Jørgensen *et al.*, 1986). Probes were added to all hybridization mixes at a concentration of 1 × 10⁶ c.p.m./ml hybridization buffer. Low stringency hybridizations (corresponding to 30% mismatch) were done in 6 × SSC at 45.5°C (1 × SSC is 0.15 M NaCl, 0.015 M sodium citrate, pH 7) with post-hybridization washes in 1 × SSC at 44°C. High stringency hybridizations were done in 6 × SSC at 75.5°C, with post-hybridization washes in 1 × SSC at 74°C (Figure 1) or in 1 × SSC at 82°C (Figure 3c). The filters were air-dried, placed in plastic bags against X-ray films (Fuji Photo Film Co., Japan) with an intensifying screen and exposed at -70°C for a few hours to a few days.

Acknowledgements

We thank Mrs Hanne Puibaraud for technical assistance. This work was supported by in part by Grant nos. 12-5734 from the Danish Medical Research Council and HD17449 from the National Institutes of Health, USA.

References

- Britten, R.J. (1986) *Science*, **231**, 1392–1398.
- Devilee, P., Slagboom, P., Cornelisse, C.J. and Pearson, P.L. (1986) *Nucleic Acids Res.*, **14**, 2059–2073.
- Dover, G. (1982) *Nature*, **299**, 111–117.
- Gillespie, D., Donhower, L. and Strayer, D. (1982) In Dover, G.A. and Flavell, R.B. (eds), *Genome Evolution*. Academic Press, London, pp. 113–133.
- Jones, C., Morse, H.G. and Palmer, D.K. (1985) *Cytogenet. Cell Genet.*, **40**, 662–663.
- Jørgensen, A.L., Bostock, C.J. and Bak, A.L. (1986) *J. Mol. Biol.*, **187**, 185–196.
- Jørgensen, A.L., Bostock, C.J. and Bak, A.L. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 1075–1079.
- Koop, B.F., Goodman, M., Xu, P., Chan, K. and Slightom, J.L. (1986) *Nature*, **319**, 234–237.
- Lund, E., Bostock, C., Robertson, M., Christie, S., Mitchen, J.L. and Dahlberg, J.E. (1983) *Mol. Cell Biol.*, **3**, 2211–2222.

- Reis,R.J.S., Srivastava,A., Beranek,D.T. and Goldstein,S. (1985) *J. Mol. Biol.*, **186**, 31–41.
- Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
- Seuáñez,H.N. (1979) *The Phylogeny of Human Chromosomes*. Springer, Berlin, p. 39.
- Strachan,T., Webb,D. and Dover,G.A. (1985) *EMBO J.*, **4**, 1701–1708.
- Waye,J.S. and Willard,H.F. (1985) *Nucleic Acids Res.*, **13**, 2731–2743.
- Waye,J.S. and Willard,H.F. (1986a) *Mol. Cell Biol.*, **6**, 3156–3165.
- Waye,J.S. and Willard,H.F. (1986b) *Nucleic Acids Res.*, **14**, 6915–6927.
- Wolfe,J., Darling,S.M., Erikson,R.P., Craig,I.W., Buckle,V.J., Rigby,P.W.J., Willard,H.F. and Goodfellow,P.N. (1985) *J. Mol. Biol.*, **182**, 477–485.
- Wu,J.C. and Manuelidis,L. (1980) *J. Mol. Biol.*, **142**, 363–386.

Received on February 20, 1987; revised on March 26, 1987