

Structure of repeated sequences in the centromeric region of the human Y chromosome

CHRIS TYLER-SMITH

CRC Chromosome Molecular Biology Group, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

Summary

Alphoid satellite DNA is a family of sequences with an approximately 170 bp periodicity which is found near the centromere of all human chromosomes. The structure of the human Y-chromosome alphoid DNA has been studied in two somatic cell hybrids, 3E7 and 853 (Tyler-Smith & Brown, 1987). The 170 bp alphoid subunits are tandemly repeated and are organized into units approximately 5.7 kb long. A few variant units on the 3E7 Y chromosome contain two extra 170 bp subunits and are approximately 6.0 kb long; the variant units are present in two clusters at least 90 kb apart on the chromosome. On each Y chromosome

there is a single major block of alphoid DNA: on the 3E7 Y chromosome it is approximately 440 kb long and on the 853 Y chromosome it is approximately 540 kb long. A long-range restriction map of the 853 block has been constructed covering approximately 1.1 mb of DNA. The distribution of restriction sites suggests that the sequences on one side of the alphoid block may be typical euchromatic DNA, while the sequences on the other side may be another satellite sequence.

Key words: centromere, human, alphoid satellite DNA.

(A) Introduction

The centromere is the part of the chromosome that interacts with the spindle apparatus and ensures correct segregation of the chromosome at cell division. During the last 7 years, centromeric DNA from several chromosomes of *Saccharomyces cerevisiae* has been analysed in detail (Clarke & Carbon, 1980; reviewed by Murray & Szostak, 1985). More recently, centromeric DNA from *Schizosaccharomyces pombe* has been characterized (Nakaseko *et al.* 1986). However, the DNA sequences that function as centromeres in mammalian chromosomes have not yet been identified. In both yeast species, an important step in the identification of the centromeric sequences was the construction of a physical map of a region of the chromosome which included the centromere. We would like to construct such maps for human chromosomes. The development of the technique of pulsed-field gel electrophoresis for analysing DNA molecules up to approximately 2000 kb in size (Schwartz & Cantor, 1984; Carle & Olson, 1984) has made it possible to begin this task. This paper will

discuss the DNA sequences which have been found near the centromere of the human Y chromosome and will present a restriction map covering approximately 1.1 mb of DNA in this region.

(B) Tandemly repeated sequences found near human centromeres

The genomes of mammals are more than 100 times larger than those of yeasts and the currently available genetic and physical maps are less detailed. Consequently, known genes or anonymous unique sequences have not provided a good starting point for mapping mammalian centromeric region DNA. However, it has been known since the time of the first *in situ* hybridization experiments to mammalian chromosomes (Pardue & Gall, 1970; Jones & Corneo, 1971) that large blocks of tandemly repeated 'satellite' or 'simple sequence' DNA are commonly found near mammalian centromeres. These sequences can therefore provide probes with which we can begin to map these regions.

Table 1. *Tandemly repeated DNA sequence families found in the centromeric regions of human chromosomes**

Sequence family	Repeat length of smallest subunit (bp)	Location in classical satellites
alphoid	170	variable minor component of II, III & IV
satellite 2 }	5	II
satellite 3 }		III & IV
satellite 1	17 & 25	I

*Based on the data of Prosser *et al.* (1986)

In human chromosomes, three families of tandemly repeated DNA sequences have been found in the centromeric regions. They are listed in Table 1 and each will be discussed briefly.

Alphoid DNA is characterized by a subunit size of approximately 170 bp. It has been found by *in situ* hybridization at or close to the centromere of all the chromosomes (Manuelidis, 1978; Mitchell, Gosden & Miller, 1985). Manuelidis (1978) reported hybridization to the centromeres of all the autosomes and to the X chromosome when she used a purified but uncloned 340 bp *EcoRI* fragment as a probe. The extent of hybridization differed considerably on different chromosomes: 1, 3, 7, 10 and 19 were heavily labelled, while the Y chromosome was virtually unlabelled. Using a cloned alphoid probe of unknown chromosomal origin called p82H, Mitchell *et al.* (1985) found a more even distribution of the sequence family: all chromosomes, including the Y, were labelled.

Satellites 2 and 3 are the major simple sequence components of the classical satellites II and III/IV respectively (Prosser, Frommer, Paul & Vincent, 1986). Both consist of tandemly repeated variants of the 5 bp sequence TTCCA, although they differ in the extent to which they have diverged from this prototype repeat. Their similarities are more striking than their differences and here they will be considered as the single family satellite 2, 3. Using gradient-purified satellite DNA, Gosden *et al.* (1975) found two major sites of hybridization: in the pericentric region of chromosome 9 and on the long arm of the Y chromosome, and minor sites near the centromeres of chromosomes 1, 5, 7, 10, 12, 13, 14, 15, 16, 17, 20, 21 and 22. Using a cloned sequence from chromosome 15, Higgins *et al.* (1985) found hybridization to the pericentric regions of chromosomes 1, 9, 13, 14, 15, 16, 21, and 22, and to the Y-chromosome long arm.

Satellite 1 is the major simple sequence component of the very AT-rich classical satellite I and consists in part of variants of 17 bp and 25 bp units (Prosser *et al.* 1986). Its chromosomal distribution was determined

using gradient-purified DNA (Gosden *et al.* 1975) and was similar to that of satellite 2, 3, although satellite 1 did not hybridize to chromosome 16. The hybridization of satellite 1 to the Y-chromosome long arm may be due to a second component present in this uncloned probe: the 2.4 kb sequence (Frommer, Prosser & Vincent, 1984), which is known to be located in this region of the genome (Cooke, Schmidke & Gosden, 1982). Thus the *in situ* hybridization experiments show that all the human chromosome centromeric regions share some features: they all contain alphoid DNA close to the primary constriction. However, they differ in the amounts of satellites 1 and 2, 3 that are present nearby: the satellites may form a large cytologically visible C band such as that on chromosome 9 or may be undetectable by *in situ* hybridization. Since the Y chromosome falls into the latter category, investigations of the Y-chromosome centromeric region DNA have begun with the study of the alphoid sequences. The Y chromosome has been particularly suitable for this work because it contains a relatively small amount of alphoid DNA.

(C) Y-chromosome alphoid DNA: short-range structure

The use of somatic cell hybrids containing single human chromosomes on a rodent background has made it possible to study the chromosome-specific organization of repeated sequences in detail (Beauchamp, Mitchell, Buckland & Bostock, 1979). The structure of Y-chromosome alphoid DNA has been studied in this way (Willard, 1985; Wolfe *et al.* 1985; Tyler-Smith & Brown, 1987) and the results are summarized as a diagram in Fig. 1. The basic structure is the 170 bp subunit, which is tandemly repeated 34 times to form an approximately 5.7 kb unit, which in turn is tandemly repeated to form an array several hundred kb long. This section will discuss the structures of the 170 bp subunit, the 5.7 kb unit and their variants; section D will discuss the structure of the whole block.

(1) Cloning the alphoid DNA

The somatic cell hybrids used in the study of the Y-chromosome alphoid DNA were 3E7 (Marcus *et al.* 1976) and 853 (originally called 7631: Burk, Ma & Smith, 1985), independently derived cell lines each of which contains a different Y chromosome as its major but not sole human component. Wolfe *et al.* (1985) constructed a cosmid library from 3E7 DNA and used an X-chromosome-derived alphoid probe to identify two cosmids containing the major Y-chromosome alphoid sequence. Similarly, Tyler-Smith & Brown (1987) probed a 3E7 cosmid library with p82H

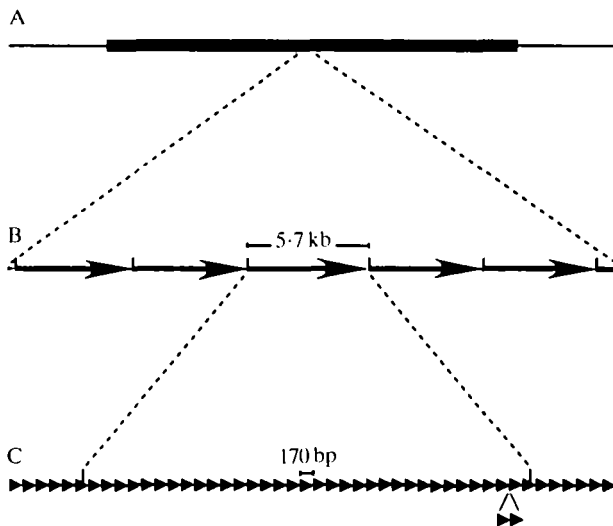


Fig. 1. Structure of the major block of alphoid satellite DNA on the human Y chromosome. (A) There is a single major block of alphoid DNA, about 440 kb in size on the 3E7 Y chromosome and about 540 kb in size on the 853 Y chromosome. (B) The block consists of tandemly repeating units which are approximately 5.7 kb in size. A few variant units on the 3E7 Y chromosome are approximately 6.0 kb long. (C) The 5.7 kb unit is made up of tandemly repeating 170 bp subunits. The variant 6.0 kb units contain two additional 170 bp subunits. (Data from Tyler-Smith & Brown, 1987.)

(Mitchell *et al.* 1985) and identified nine cosmids containing the major alphoid sequence.

(2) The 170 bp subunit

Restriction enzyme mapping of the cloned DNA with *HaeIII* showed a repeating structure with a periodicity of approximately 170 bp and sequencing of about 2 kb of DNA from each of two different units revealed the details of the structure (Tyler-Smith & Brown, 1987). The sequence data reported by Wolfe *et al.* (1985) contained errors and will not be discussed further. Individual 170 bp subunits were between 76 and 86 % homologous to one another. A Y-chromosome consensus sequence was derived: individual subunits were between 87 and 92 % homologous to the consensus. The consensus sequence was compared with alphoid sequences from other chromosomes. For example, it was about 82 % homologous to a chromosome-7 sequence (Jorgensen, Bostock & Bak, 1986) and was about 89 % homologous to the X-chromosome consensus sequence (Waye & Willard, 1985).

These results show that the Y-chromosome alphoid subunits are heterogeneous, but are generally more similar to one another than they are to alphoid sequences from other chromosomes.

(3) The 5.7 kb unit

Restriction enzyme mapping of both cosmid clones and of genomic DNA showed that the 170 bp subunits were organized into a higher order structure with a periodicity of approximately 5.7 kb. Higher order structure has been described on other chromosomes: for example, the X-chromosome unit is approximately 2 kb in size and consists of 12 subunits (Waye & Willard, 1985). Three pieces of evidence demonstrated that there was very little sequence divergence between different Y-chromosome 5.7 kb units. (i) Partial sequencing of the same 220 bp region from 14 independently cloned units showed sequence differences at only two positions. (ii) More extensive sequencing of >2 kb of DNA from two different units showed only one difference outside an extra region of 340 bp present in one of the units, discussed below. (iii) Long-range restriction site mapping of the 853 alphoid block (section D & Fig. 3) showed that a site that was absent from one unit was absent from all units: no enzyme cut just once or a few times within the block. These pieces of evidence lead to the conclusion that different units are >99 % homologous to one another.

(4) Variant units

Nevertheless, two variant regions of the unit were found in the 3E7 cell line. About one tenth of the units were approximately 6.0 kb long. Sequencing of the extra region showed that it consisted of two additional 170 bp subunits of typical Y-chromosome alphoid DNA. Restriction enzyme mapping with *AvaII* revealed that most of the 6.0 kb units contained an *AvaII* site, while all of the 5.7 kb units lacked the site. The site was not situated in the two extra subunits, but was a point mutation nearby. Mapping of both cosmid clones and genomic DNA showed that the variant units were organized into two clusters at least 90 kb apart on the chromosome. Neither of the two variants was detected on the 853 Y chromosome or on the additional Y chromosomes examined by Wolfe *et al.* (1985).

(D) Y-chromosome alphoid DNA: long-range structure

Initial studies of the long-range structure of the alphoid DNA in 3E7 and 853 cells were reported by Tyler-Smith & Brown (1987). Five restriction enzymes were identified which did not cut within either block of alphoid DNA; a sixth enzyme, *AvaII*, cut within the 3E7 alphoid DNA at two clusters of sites but did not cut within the 853 alphoid DNA. These enzymes were used to digest high molecular weight DNA and the products of digestion were analysed by pulsed-field gel electrophoresis. The five non-cut

enzymes each produced a single major alphoid DNA fragment. This suggested that most of the alphoid DNA was organized as a single block or possibly as multiple blocks with identical restriction site patterns. However, the quantitative estimate of the amount of alphoid DNA per Y chromosome reported by Wolfe *et al.* (1985), together with the finding that alphoid DNA is not interspersed with other sequences, suggested that there was almost certainly only a single block per Y chromosome. An upper limit to the size of this block was provided by the size of the smallest single restriction fragment. This was estimated at approximately 475 kb for the 3E7 Y chromosome and 575 kb for the 853 Y chromosome. Thus there was a size difference of about 100 kb between the two blocks.

Sites outside the block were mapped by carrying out double digests. The resulting maps suggested that the flanking sequences were similar on the two chromosomes. A striking feature of these maps was the lack of some restriction sites which would be expected to occur frequently in human DNA. For example, on one side of the alphoid DNA there was between 85 and 100 kb of DNA lacking both a *Bam*HI and a *Bcl*I site. Their absence suggested that the flanking DNA was not typical human DNA, but might be another simple sequence. This is discussed further in section E. However, the true nature of the flanking sequences was not determined since no cosmid clones extended into them.

The observation of a large difference in size between the alphoid DNA blocks on the two Y chromosomes examined suggests that the tandemly repeated sequences may be very variable on a large scale in the human genome. The *Ava*II site variants provide evidence for variability on a smaller scale, similar to that found on chromosomes 17 and X by Willard *et al.* (1986).

(1) Detailed restriction site mapping of the 853 alphoid block

The long-range structure of the 853 alphoid DNA has recently been mapped in more detail. A total of 36 restriction enzyme specificities that do not occur within the alphoid DNA were identified. Genomic digests with 30 of these enzymes were analysed by pulsed-field gel electrophoresis. The analysis was facilitated by the use of an apparatus designed by E. M. Southern and known as a 'waltzer' (Anand, 1986; Southern, Anand, Brown & Fletcher, 1987). This apparatus has the advantage over the original designs of Schwartz & Cantor (1984) and Carle & Olson (1984) that the DNA tracks are straight. Consequently, the mobilities of DNA fragments in different tracks can be easily and accurately compared.

DNA samples digested singly with each of 26 restriction enzymes, including 24 different specificities, are shown in Fig. 2A,B. In most tracks, a single major alphoid band is seen. The sizes estimated for these bands range from approximately 540 kb for the *Ava*II digest to approximately 1200 kb for the *Nae*I digest. Sizes measured using this apparatus were commonly 5 to 10 % smaller than those of the same fragment measured using a Carle & Olson apparatus (Tyler-Smith & Brown, 1987). Some restriction enzymes, such as *Cla*I, produced multiple bands. The pattern did not change when the enzyme:DNA ratio was increased (results not shown); double digests showed that there were no sites within the alphoid block. The multiple bands may therefore reflect partial blockage at some sites e.g. by DNA methylation. Other enzymes, such as *Not*I, did not produce any bands within the limit of size fractionation of the gel. These enzymes therefore lack an accessible pair of sites within the region analysed.

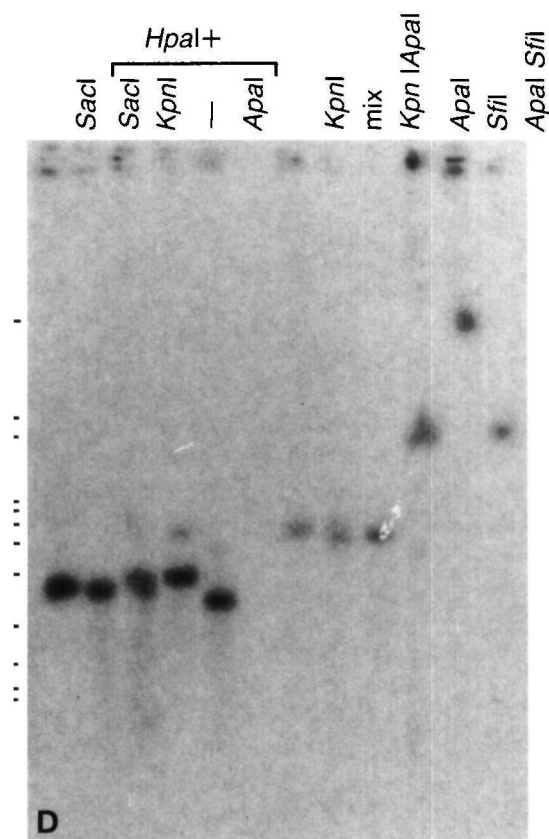
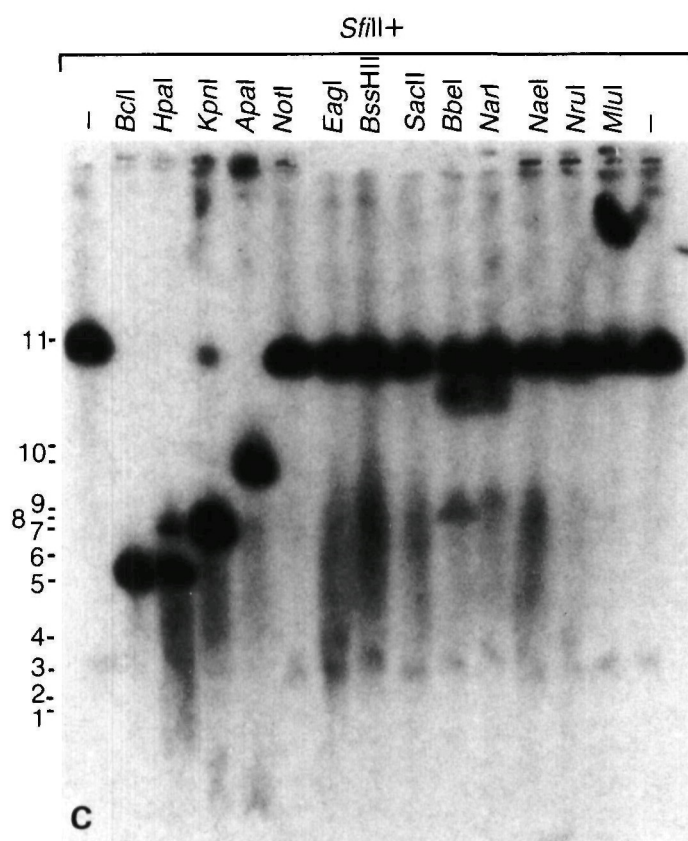
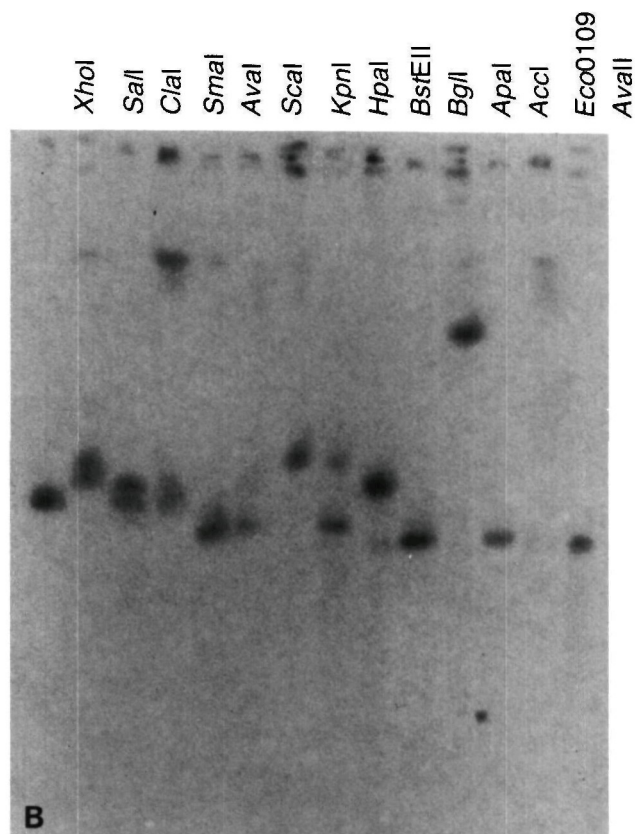
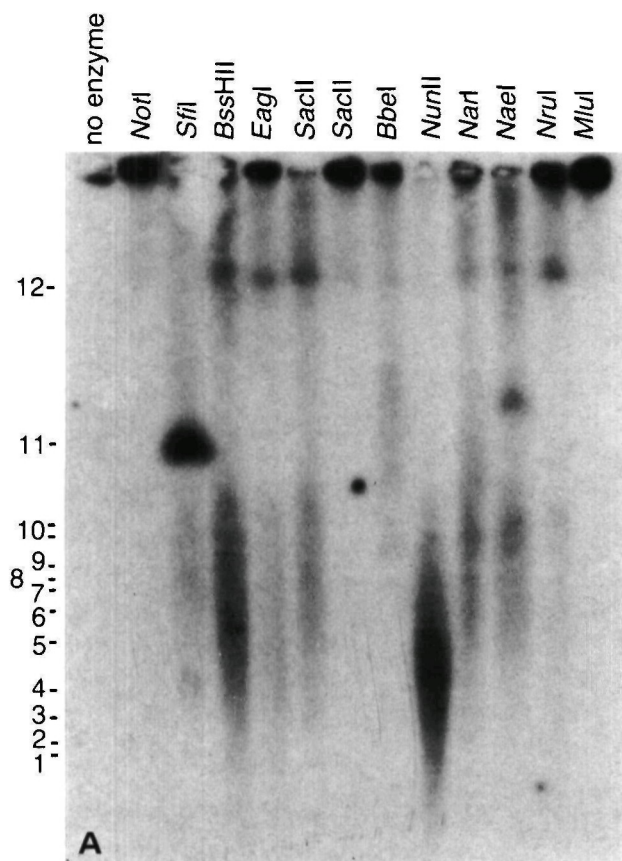
Double digests were carried out with pairs of restriction enzymes and the sizes of the resulting alphoid DNA fragments were measured. Examples of the double digests are shown in Fig. 2C,D. From these and other such digests, a map of 1.1 mb of DNA was constructed, on which sites for 20 enzymes could be placed (Fig. 3). An additional seven enzymes were shown to lack even one accessible site within this region (Fig. 2C). The map shown in Fig. 3 provides some indirect evidence about the nature of the sequences flanking the alphoid DNA. Its implications will be discussed in the next section.

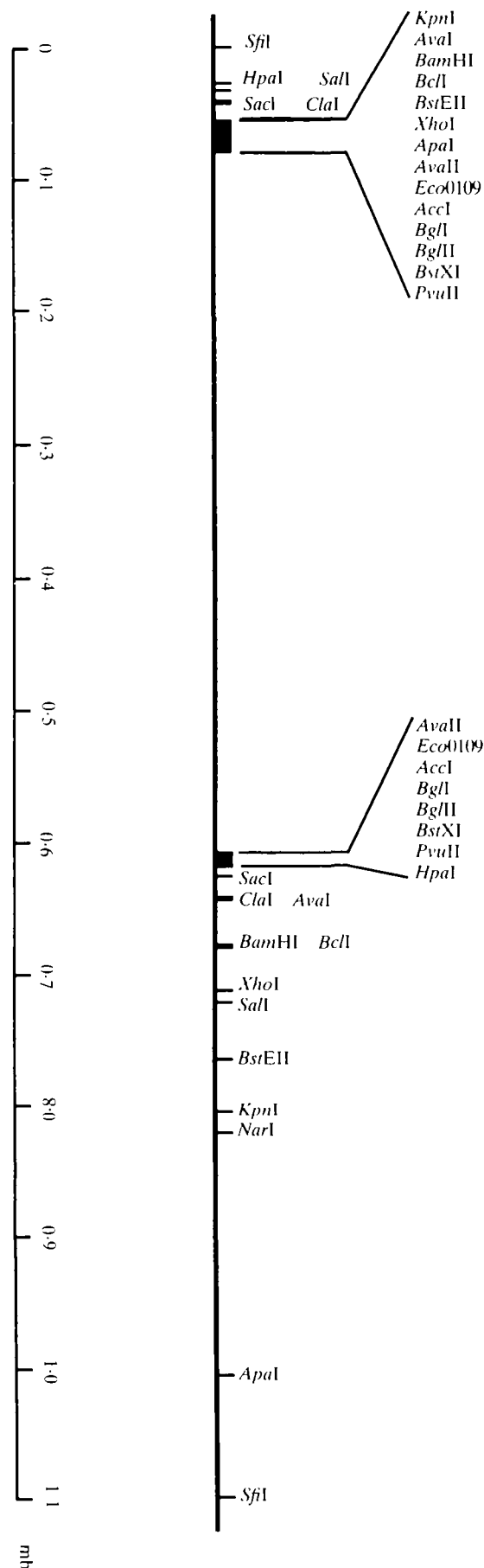
(E) What other DNA sequences are present in the Y-chromosome centromeric region?

(1) The left-hand flanking region

The restriction site map shown in Fig. 3 can be divided into three regions: the alphoid block and its

Fig. 2. Analysis of the long-range structure of the 853 Y-chromosome alphoid DNA. A and B show single digests; C and D show double digests of 853 DNA. The enzymes used are indicated at the top of each track, and the positions of *Saccharomyces cerevisiae* chromosome size markers are shown on the left-hand side according to the nomenclature of Carle & Olson (1985). DNA preparation in agarose plugs, restriction enzyme digestion, and pulsed-field gradient gel electrophoresis on a 'waltzer' were as described by Anand (1986). Effective pulse times were (A) 85 s, (B) 60 s, (C) 75 s, (D) 70 s. All gels were run at 150 V for 33 h at 20°C. DNA was transferred to nitrocellulose and hybridized to a ³²P-labelled plasmid containing one 6.0 kb unit of Y-chromosome alphoid DNA. Filters were washed in 3×SSC at 65°C and autoradiography was carried out overnight without intensification.





Typical Alphoid Simple sequence

Fig. 4. DNA sequences present in the 853 Y chromosome centromeric region. A 540 kb block of alphoid DNA (black line) is flanked on the left-hand side by sequences which have a restriction site distribution typical of most human DNA containing unique and interspersed repeated sequences (open line), and on the right-hand side by a region of at least 400 kb which is thought to contain simple sequence DNA (stippled line).

left-hand and right-hand flanking sequences (Fig. 4). In the left-hand flanking region, a site for 13 of the 15 enzymes that are expected to cut frequently in human DNA occurs within approximately 20 kb of the alphoid-nonalphoid boundary, forming a striking cluster. The two exceptions are *SacI* and *HpaI*; these sites are approximately 35 and 55 kb, respectively, from the boundary. Sites for some enzymes that are expected to cut less frequently in human DNA are dispersed over about 80 kb (*ClaI*, *SalI* and *SfiI*). This region of DNA, therefore, has a restriction site distribution similar to that expected in a stretch of typical human DNA consisting of a mixture of unique sequences and interspersed repeated sequences. It is possible that the left-hand boundary is the boundary between euchromatic and heterochromatic DNA.

(2) The right-hand flanking region

In contrast, the right-hand flanking region does not have the restriction site distribution expected for typical human DNA. For example, the closest *ApaI* site to the boundary is approximately 400 kb away, while the closest *KpnI* site is approximately 200 kb away. These enzymes are expected to cut on average every 6.4 and 8.5 kb, respectively (Drmanac, Petrovic, Glisin & Crkvenjakov, 1986). This region is therefore likely to consist of a simple sequence. The dispersion of sites suggests that the sequence is less homogeneous than the alphoid DNA. The small cluster of sites at the boundary might indicate a small stretch of more typical DNA or, alternatively, that the proposed simple sequence is rich in these sites throughout.

What could be the identity of this putative block of simple sequence DNA? The most obvious candidates are the satellites 2, 3 and 1, since they are known to be located near alphoid DNA on other chromosomes.

Fig. 3. Restriction site map of the 853 Y-chromosome alphoid DNA. The positions of the cleavage sites closest to the alphoid DNA are shown for each enzyme. Additional distal sites may be present. Sites in the two large clusters were not positioned relative to one another within the cluster. No sites were detected within this region for *BssHII*, *EagI*, *MluI*, *NaeI*, *NotI*, *NruI* or *SacII*.

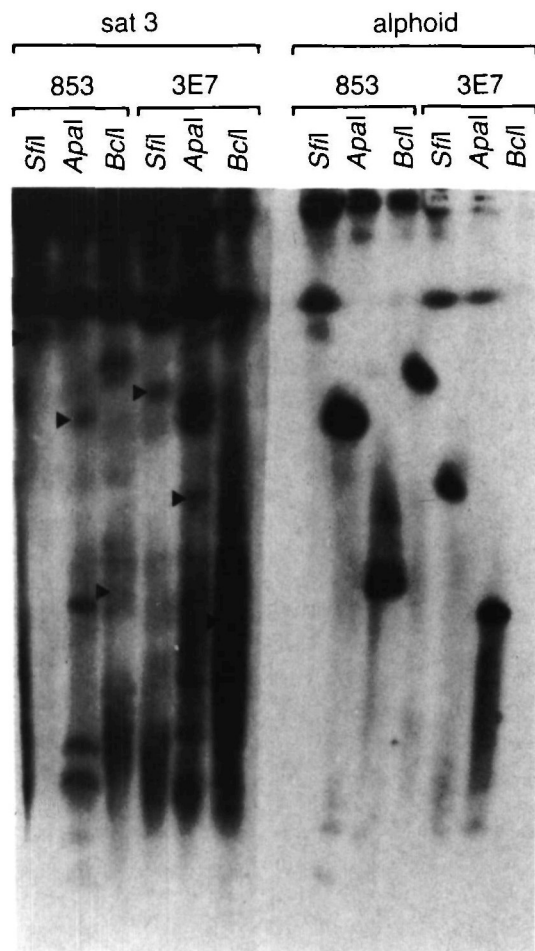


Fig. 5. Alphoid DNA and satellite 3 DNA hybridize to large DNA fragments of the same size. DNA from 853 or 3E7 cells was digested with the enzyme shown and electrophoresed as described in the legend to Fig. 2, using an effective pulse time of 70 s. After transfer to nitrocellulose, the right-hand half of the filter was hybridized to a Y-chromosome alphoid probe, while the left-hand half was hybridized to the cloned satellite 3 probe λ HS5 (Cooke & Hindley, 1979). The arrows mark the position in each track that would be expected to hybridize to alphoid DNA. In each case, a satellite band is present at this position.

They were not detected by *in situ* hybridization near the centromere of the Y chromosome, but this failure is easily explained: the very strong signal from the long arm of the Y chromosome would mask a weak signal from the centromere. Fig. 5 shows an experiment that investigates the possibility that the flanking sequence is related to satellite 2, 3. 3E7 DNA or 853 DNA was digested with *BclI*, *ApaI* or *SfiI*, enzymes that produce DNA fragments in which the alphoid DNA is linked to its right-hand flanking sequence. One half of the gel was probed with an alphoid sequence and the other half with λ HS5 (Cooke & Hindley, 1979), a satellite 2, 3 sequence. In each track, there is some satellite 2, 3 hybridization at

the position of the alphoid fragment. This result is consistent with the hypothesis that satellite 2, 3 DNA may be physically linked to alphoid DNA. However, it does not demonstrate linkage conclusively because two different DNA fragments, one alphoid and one satellite, might comigrate by chance in each case. Additional experiments are required to resolve this issue.

(F) Conclusions

A complete map of a human centromeric region would be expected to extend from the euchromatic DNA on one arm of the chromosome, through the heterochromatic DNA, to the euchromatic DNA on the other arm. The technique of pulsed-field gel electrophoresis has made it possible to construct a restriction site map of the Y-chromosome centromeric region, which covers about 2.0% of the chromosome, and to identify three distinct areas on the map, including a block of alphoid DNA and a block of a second simple sequence. Thus, a substantial portion of the Y-chromosome centromeric region has been mapped. This map will form the basis for a more detailed structural analysis of the region and a systematic search for functional mammalian centromeric sequences.

I thank Lesley Taylor for growing cells, Martin Johnson for constructing the pulsed-field gel electrophoresis apparatus, Rakesh Anand for protocols for the preparation and analysis of high molecular weight DNA, Rakesh Anand and Ed Southern for reading the manuscript and Frank Caddick for photography. This work was funded by the Cancer Research Campaign.

References

- ANAND, R. (1986). Pulsed field gel electrophoresis: a technique for fractionating large DNA molecules. *Trends in Genetics* **2**, 278–283.
- BEAUCHAMP, R. S., MITCHELL, A. R., BUCKLAND, R. A. & BOSTOCK, C. J. (1979). Specific arrangements of human satellite III DNA sequences on human chromosomes. *Chromosoma* **71**, 153–166.
- BURK, R. D., MA, P. & SMITH, K. D. (1985). Characterization and evolution of a single-copy sequence from the human Y chromosome. *Molec. cell. Biol.* **5**, 576–581.
- CARLE, G. F. & OLSON, M. V. (1984). Separation of chromosomal DNA molecules from yeast by orthogonal-field-alternation gel electrophoresis. *Nucleic Acids Res.* **12**, 5647–5664.
- CARLE, G. F. & OLSON, M. V. (1985). An electrophoretic karyotype for yeast. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3756–3760.

- CLARKE, L. & CARBON, J. (1980). Isolation of a yeast centromere and construction of functional small circular chromosomes. *Nature, Lond.* **287**, 504–509.
- COOKE, H. J. & HINDLEY, J. (1979). Cloning of human satellite III DNA: different components are on different chromosomes. *Nucleic Acids Res.* **6**, 3177–3197.
- COOKE, H. J., SCHMIDKE, J. & GOSDEN, J. R. (1982). Characterisation of a human Y chromosome repeated sequence and related sequences in higher primates. *Chromosoma* **87**, 491–502.
- DRMANAC, R., PETROVIC, N., GLISIN, V. & CRKVENJAKOV, R. (1986). A calculation of fragment lengths obtainable from human DNA with 78 restriction enzymes: an aid for cloning and mapping. *Nucleic Acids Res.* **14**, 4691–4692.
- FROMMER, M., PROSSER, J. & VINCENT, P. C. (1984). Human satellite I sequences include a male specific 2.47 kb tandemly repeated unit containing one Alu family member per repeat. *Nucleic Acids Res.* **12**, 2887–2900.
- GOSDEN, J. R., MITCHELL, A. R., BUCKLAND, R. A., CLAYTON, R. P. & EVANS, H. J. (1975). The location of four human satellite DNAs on human chromosomes. *Exp Cell Res.* **92**, 148–158.
- HIGGINS, M. J., WANG, H., SHTROMAS, I., HALIOTIS, T., RÖDER, J. C., HOLDEN, J. J. A. & WHITE, B. N. (1985). Organization of a repetitive human 1.8 kb KpnI sequence localized in the heterochromatin of chromosome 15. *Chromosoma* **93**, 77–86.
- JONES, K. W. & CORNEO, G. (1971). Location of satellite and homogeneous DNA sequences on human chromosomes. *Nature New Biology* **233**, 268–271.
- JØRGENSEN, A. L., BOSTOCK, C. J. & BAK, A. L. (1986). Chromosome-specific subfamilies within human alphoid repetitive DNA. *J. molec. Biol.* **135**, 127–149.
- MANUELIDIS, L. (1978). Chromosomal localization of complex and simple repeated human DNAs. *Chromosoma* **66**, 23–32.
- MARCUS, M., TANTRAVAHU, R., DEV, V. G., MILLER, D. A. & MILLER, O. J. (1976). Human-mouse cell hybrid with human multiple Y chromosomes. *Nature, Lond.* **262**, 63–65.
- MITCHELL, A. R., GOSDEN, J. R. & MILLER, D. A. (1985). A cloned sequence, p82H, of the alphoid repeated DNA family found at the centromeres of all human chromosomes. *Chromosoma* **92**, 369–377.
- MURRAY, A. W. & SZOSTAK, J. W. (1985). Chromosome segregation in mitosis and meiosis. *A. Rev. Cell Biol.* **1**, 341–367.
- NAKASEKO, Y., ADACHI, Y., FUNAHASHI, S., NIWA, O. & YANAGIDA, M. (1986). Chromosome walking shows a highly homologous repetitive sequence present in all the centromere regions of fission yeast. *EMBO J.* **5**, 1011–1021.
- PARDUE, M. L. & GALL, J. G. (1970). Chromosomal localization of mouse satellite DNA. *Science* **168**, 1356–1358.
- PROSSER, J., FROMMER, M., PAUL, C. & VINCENT, V. C. (1986). Sequence relationships of three human satellite DNAs. *J. molec. Biol.* **187**, 145–155.
- SCHWARTZ, D. C. & CANTOR, C. R. (1984). Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell* **37**, 67–75.
- SOUTHERN, E. M., ANAND, R., BROWN, W. R. A. & FLETCHER, D. S. (1987). A model for the separation of large DNA molecules by crossed field gel electrophoresis. *Nucleic Acids Res.* **15**, 5925–5943.
- TYLER-SMITH, C. & BROWN, W. R. A. (1987). Structure of the major block of alphoid satellite DNA on the human Y chromosome. *J. molec. Biol.* **195**, 457–470.
- WAYE, J. S. & WILLARD, H. F. (1985). Chromosome-specific alpha satellite DNA: nucleotide sequence analysis of the 2.0 kilobasepair repeat from the human X chromosome. *Nucleic Acids Res.* **13**, 2731–2743.
- WILLARD, H. F. (1985). Chromosome-specific organisation of human alpha satellite DNA. *Am. J. hum. Genet.* **37**, 524–532.
- WILLARD, H. F., WAYE, J. S., SKOLNICK, M. H., SCHWARTZ, C. E., POWERS, V. E. & ENGLAND, S. B. (1986). Detection of restriction fragment length polymorphisms at the centromeres of human chromosomes by using chromosome-specific satellite DNA probes: implications for development of centromere-based genetic linkage maps. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5611–5615.
- WOLFE, J., DARLING, S. M., ERICKSON, R. P., CRAIG, R. W., BUCKLE, V. J., RIGBY, P. W. J., WILLARD, H. F. & GOODFELLOW, P. N. (1985). Isolation and characterization of an alphoid centromeric repeat family from the human Y chromosome. *J. molec. Biol.* **182**, 477–485.