© 1992 Oxford University Press
Nucleic Acids Research, Vol. 20, No. 18 4781–4787

Characterisation of a boundary between satellite III and alphoid sequences on human chromosome 10

M.S.Jackson, S.E.Mole and B.A.J.Ponder
CRC Human Cancer Genetics Research Group, Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK

Received June 30, 1992; Revised and Accepted August 19, 1992 EMBL accession no. X67031

ABSTRACT

Alphoid and satellite III sequences are arranged as large tandem arrays in the centromeric regions of human chromosomes. Several recent studies using *in situ* hybridisation to investigate the relative positions of these sequences have shown that they occupy adjacent but non-overlapping domains in metaphase chromosomes. We have analysed the DNA sequence at the junction between alphoid and satellite III sequences in a cosmid previously mapped to chromosome 10. The alphoid sequence consists of tandemly arranged dimers which are distinct from the known chromosome 10-specific alphoid family. Polymerase chain reaction experiments confirm the integrity of the sequence data. These results, together with pulsed field gel electrophoresis data place the boundary between alphoid and satellite III sequences in the mapping interval 10 centromere-10q11.2. The sequence data shows that these repetitive sequences are separated by a partial L1 interspersed repeat sequence less than 500bp in length. The arrangement of the junction suggests that a recombination event has brought these sequences into close proximity.

INTRODUCTION

Several families of repetitive sequences are associated with the centromeric regions of human chromosomes (1). Alpha satellite, or alphoid, DNA appears to be present at the centromeres of all human chromosomes (2), where it is arranged in large tandem arrays of repeats based on a monomeric repeat length of approximately 170bp. Many chromosome specific subfamilies with distinct higher order repeat structures have been identified by restriction enzyme periodicities or from primary sequence (reviewed in ref. 3). However, not all alphoid repeats show this chromosomal specificity; among the acrocentric chromosomes, several alphoid subfamilies are present on more than one chromosome (4). The advent of Pulsed Field Gel Electrophoresis (PFGE) has facilitated the construction of centromeric maps for several chromosomes (5, 6) and it has recently been shown that chromosome 7 contains two linked alphoid arrays which can be spatially resolved by PFGE (6). This, together with the extensive long range polymorphism associated with these sequences, suggests that their organisation may be more complex than was originally thought.

Repetitive sequences based on the 5bp sequence 5'GGAAT3' (the satellite III sequence family) appear to be present in the centromeric regions of some, but not all, human chromosomes. Early *in situ* hybridisation studies demonstrated that chromosomes Y, 9, 15 and the acrocentric chromosomes contain large quantities of satellite III DNA, while others, including chromosome 10, exhibited low, but detectable, levels of hybridisation to satellite III probes (1). Although not as well characterised as alphoid sequences, some chromosome-specific higher order repeat structures have been identified for satellite III sequences. Beauchamp *et al.* (7) found that the distribution of restriction sites within satellite III sequences differed among chromosomes. Since then satellite III subfamilies with characteristic repetitive structures have been found on chromosome 15 (a 1.8kb repeat, ref. 8), chromosome 14 (a polymorphic 5.0kb or 4.8kb repeat, ref. 9) and one subfamily present on both chromosomes 14 and 22 (10). Thus satellite III sequences appear to exhibit the same organisational features as alphoid sequences.

Recently, several studies have used *in situ* hybridisation to demonstrate that satellite III and alphoid DNA are closely linked on human chromosomes, but that they do not appear to interdigitate with each other (11, 12). In addition, preliminary PFGE experiments have suggested that satellite III and alphoid sequences occur on the same large restriction fragments (5).

One way to demonstrate conclusively linkage between alphoid and satellite III sequences would be to sequence directly from one repeat type to the other. Although genomic clones have been identified which contain sequences from both the *Sau*3A and *Hae*III families of repeats (13), no clones have been described which contain both alphoid and satellite III sequences. Analyses of 'novel' DNA sequences present in alphoid clones has, to date, only identified the presence of long interspersed repetitive elements (L1 elements) which are presumed to have inserted into existing alphoid arrays (14, 15).

In addition to providing linkage and mapping data, sequence information from the terminal repeats of alphoid and satellite III arrays may improve our understanding of the processes which maintain the sequence identity of tandemly repeated arrays. Computer simulations (16) and sequence data both from short tandem repeats in the rDNA spacer region of wheat and maize (17,18) and from minisatellites (19) have shown that the repeats at the termini of tandem arrays are often degenerate relative to the consensus repeat. This is believed to be due to the mechanics of processes such as unequal crossing over and gene conversion,

which are thought to maintain the sequence identity of tandemly repeated arrays (16, 20). However, the generality of these observations has not been tested in very long tandemly arranged sequences such as alphoid DNA. We describe here a detailed analysis of one genomic cosmid clone cMEN375, previously mapped to chromosome 10 (21), which contains a boundary between alphoid and satellite III sequences.

MATERIALS AND METHODS

Derivation of cosmid clone cMEN375

The construction of the cosmid library from which cMEN375 was isolated is described elsewhere (22). Briefly, DNA from a subline of the radiation hybrid R244-3A (which contains only human chromosomes 10 and Y in a hamster background) was partially digested with *Sau*3A, size fractionated, and cloned into the cosmid vector pWEX15. Identification of clones containing human DNA and their subsequent mapping are described elsewhere (21, 22).

Electrophoresis, Southern transfer and hybridisations

Electrophoresis and Southern blotting was carried out using standard methods (23).

Pulsed Field Gel Electrophoresis (PFGE) was performed using the CHEF DRII system (Bio-Rad) in 0.5×TBE buffer (45mM Tris, 45mM Boric acid, 1mM EDTA, pH 7.6). Gels were run at 200V with a pulse time of 60sec. for 15hrs. followed by 90sec. for 9hrs. Chromosomes from the *S.cerevisiae* strain YNN295 were used as molecular weight size standards. Gels were depurinated for 30 minutes in 0.25M HCl prior to denaturation and transfer.

DNA probes were labelled using the random oligonucleotide-priming method (24). Filters were washed in 2×SSC, 42°C for low stringency and 0.1×SSC, 0.1%SDS at 65°C for high stringency.

Cosmid fragments were subcloned into pBluescriptIISK$^-$ (Stratagene) and transformed into DH5α bacteria. The subclone mC375H was obtained by digesting cMEN375 to completion with HindIII, religating, transforming and screening for clones with the correct restriction pattern in suitable digests.

Generation of nested deletions and DNA sequencing

Nested deletion series for DNA sequencing were prepared in both directions from two clones p375M2.4 and mC375H using the Erase-a-Base system (Promega) according to the manufacturer's instructions.

p375M2.4 was double digested with 1. *Kpn*1 and *Sal*1 plus 2. *Sst*1 and *Bam*H1 prior to Exonuclease treatment to produce nested deletions in both orientations.

The protocols to produce nested deletions from mC375H were as follows:

(1)mC375H was digested with *Not*1, protected from ExoIII digestion with α-phosphorothioate dNTP's and then digested with *Eco*RV to produce an ExoIII sensitive end.

(2)mC375H was digested with *Bss*HII, protected from ExoIII digestion with α-phosphorothioate dNTP's and digested with *Hin*dIII to produce an ExoIII sensitive end.

ExoIII digestion, S1 digestion, ligation, transformation and the screening of transformants was performed according to manufacturer's recommendations. Double-stranded sequencing templates were prepared according to Kraft *et al.* (25) and dideoxy sequencing reactions were performed with [^{35}S]dATP as the labelled deoxynucleotide using Sequenase (U.S.B.) according to the manufacturer's instructions. Compressions were resolved by sequencing other templates from the same, or adjacent timepoint in the deletion series, by Taq sequencing (Pharmacia), or by the addition of formamide to the sequencing reaction (26).

The M13 Universal and Reverse primers were used with templates derived from p375M2.4. The following oligonucleotides were used as primers with templates derived from mC375H:

WEXT3 5'CGTCTTCAAGAATTCGC3'
BR4341 5'AGGCCCTTTCGTCTTCA3'
WEXR4502 5'TCGCCATGGGTCACGACGAGATC3'

In addition, the following oligonucleotides were used for Taq sequencing:

M13-40/24 5'CGCCAGGGTTTTCCCAGTCACGAC3' 24
BR4341/27 5'GCGTATCACGAGGCCCTTTCGTCTTCA3' 27

Sequence data was assembled using the Assembly Align software package (I.B.I) and was analysed using the MacVector software package (I.B.I.), and the GCG package of programmes (27).

PCR conditions and primers

The following oligonucleotides were used in PCRs.

375A 5'CCATGACACGAGAACTACATGATC3'
375B 5'CTCTGCGAGTTGAACATACACATC3'
375C 5'TGAATGGAATTGAATGGAATCAT3'
375D 5'CGGTTACTGAGGTTGTGAATTCCA3'
375E 5'TTGGGAAGGAGTATGTGTGTTGAGGA3'

PCRs were performed using 100ng of genomic DNA or 5ng of cloned DNA in a volume of 50μl in the presence of 1.25U of Taq polymerase (N.B.S.). The reaction conditions were 50mM KCl; 10mM Tris pH8.4; 1.5mM $MgCl_2$; 200μM each of dATP, dCTP, dGTP and dTTP; and 100μg/ml gelatin. Primer concentration varied, being 0.5μM for each of the L1 primers 375A, 375D and 375E, 1.0μM for the alphoid primer 375B and between1.0 μM and 2.0μM for the Satellite III primer 375C. A 4 minute denaturation at 95°C was followed by 30 cycles of 1 minute denaturation at 94°C, 1 minute annealing at 58°C and 1 minute extension at 72°C, with a final 10 minute extension at 72°C.

Somatic cell hybrids

R342A4 (28). A somatic cell hybrid containing chromosome 10 as its only human material in a hamster background.

64034p6-1C (29). A somatic cell hybrid containing the long arm of chromosome 10 and chromosome 12 as its only human material in a hamster background.

TG3 and TK2 (30). Two subclones derived from a fusion between the mouse RAG line and a human fibroblast cell line containing a balanced X;10 translocation. Hybrid TG3 carries 10pter-10q11.2, and hybrid TK2 carries 10q11.2-qter, as their only chromosome 10 material respectively.

RESULTS

Characterisation of a satellite III/alphoid boundary

Our initial analysis involved the hybridisation of DNA from the chromosome 10-only hybrid R342A4 with cloned human repetitive sequences to identify homologous sequences on

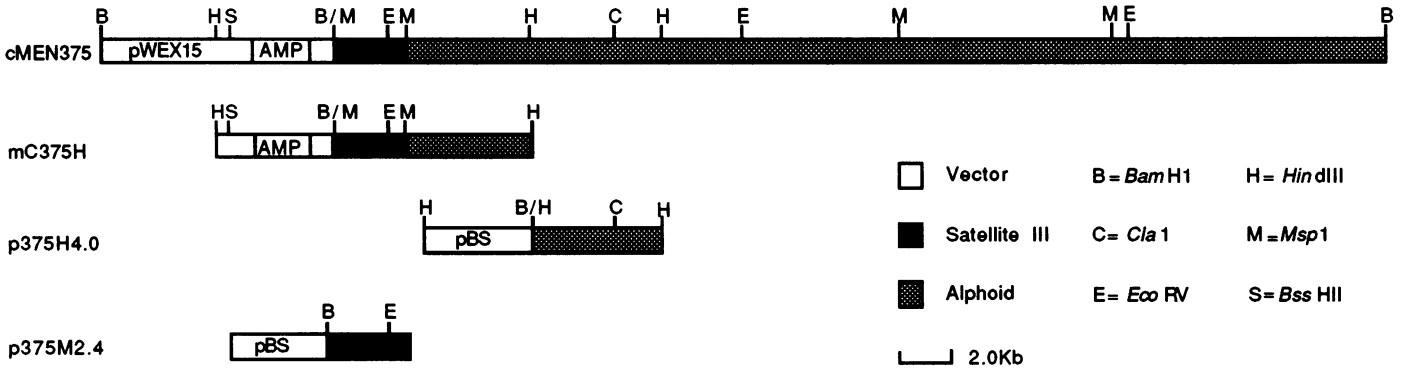

Figure 1. Restriction map of cMEN375 and subclones. The location of vector, alphoid and satellite III sequences within each clone is shown. Additional, unmapped *Hin*dIII sites exist in cMEN375. *Msp*1 sites within the vectors are not shown.

chromosome 10. A satellite III clone, pHS3, derived from chromosome 1 (31) gave a strong positive result. Two clones, p22hom48.4 and pKFC88, which are members of the 42bp family and the Sau3A family respectively (32, 33) gave negative results. We then screened 43 chromosome 10 derived cosmids, 19 of which mapped to the pericentromeric region (21), with the satellite III clone, pHS3, and a chromosome 10 specific alphoid clone, pα10RP8 (34). Fifteen of these cosmids, 13 of which mapped to the pericentromeric region, hybridised to the chromosome-10 specific alphoid clone pα10RP8, and one of these, cMEN375, also hybridised with pHS3.

The arrangement of the alphoid and satellite III sequences in cMEN375 was determined by hybridising pα10RP8 and pHS3 to restriction digests of cMEN375 and by using gel purified *Eco*R1 fragments from the cosmid as probes against digested total human DNA (data not shown). A partial restriction map of this cosmid and the location of sequences homologous to both probes is given in Figure 1. The cosmid consists of approximately 2.5kb of satellite III homologous sequence which appears to abut directly onto 34−36kb of alphoid related sequences. To analyse the junction between these sequences in more detail the subclones shown in Figure 1 were derived from cMEN375. Repeated attempts to clone the junction fragment into plasmid vectors were unsuccessful and so the mini-cosmid mC375H (Fig. 1) was derived directly from cMEN375 for further analysis.

Sequence of the satellite III/alphoid boundary

We sequenced 3.9kb of DNA spanning the satellite III/alphoid boundary by creating a series of Exonuclease III deletions from the subclones mC375H and p375M2.4 (Materials and Methods). The sequence consists of 603bp of satellite III sequence and 2818bp of alphoid DNA separated by 485bp of sequence which is unrelated to either repeat. A partial restriction map and diagon plot of the sequence compared against itself (Figure 2) illustrates the repetitive nature of the satellite III and alphoid sequences and their relative positions.

The partial restriction map (Figure 2a) shows that the satellite III sequence (nucleotides 1−603) contains multiple sites for the restriction enzymes *Taq*I and *Hin*fI. The diagon plot demonstrates the highly repetitive nature of this sequence (Figure 2b) with multiple sequence matches being identified within the first 600bp of sequence. A search of the EMBL data base using nucleotides 1−599 of the cMEN375 sequence revealed 93% sequence identity with a satellite III sequence derived from chromosome 1 (35).

The alphoid DNA (nucleotides 1089−3906) consists of tandemly arranged monomers of approximately 170bp which are cleaved by the restriction enzymes *Dde*I, *Eco*RI and *Mae*I with the same, or related periodicity (Figure 2a). This periodicity is clear from the diagon analysis (Figure 2b) where the distance between the parallel lines created by multiple sequence matches corresponds to the length of the repeat unit.

The alphoid sequence was used to search the EMBL library for homologous sequences. Of the sequences with known chromosomal origin, the highest similarity scores (80%−85%) were obtained with alphoid clones which map to chromosomes containing the Type 1 Suprafamily of alphoid sequences (chromosomes 1, 3, 5, 6, 7, 10, 12, 16 and 19) defined by Alexandrov *et al.* (36). The individual clones were derived from chromosomes 7 (37), 16 (38), 1, 5 or 9 (39), 3 (40) and 12 (41). Comparison of the alphoid sequence of cMEN375 with the sequence of the chromosome 10-specific clones pα10RP8 and pα10RR6 (42) gave sequence identities of 84%−85%, no higher than the scores obtained with sequences from other chromosomes. In addition, the cosmid cMEN375 does not contain the 1.0kb and 1.35kb *Rsa*1 fragments characteristic of the alphoid array from which pα10RP8 and pα10RR6 were derived (data not shown). This confirms that the alphoid sequences presented here are distinct from those already characterised for chromosome 10 (34).

To investigate further the sequence organisation of the alphoid monomers in cMEN375 we aligned and compared the 16 complete alphoid monomers which were sequenced. The boundaries between individual monomers within a tandem array are normally arbitrarily defined either by an enzyme site which recognises a higher order repeat structure, or an enzyme site used to obtain clones from which sequence information was obtained (discussed in ref. 43). Since we have obtained a junction fragment, the start of the monomeric register has been chosen as the first nucleotide after the end of the 485bp sequence separating the satellite III from the alphoid sequences (see below).

The results of the sequence comparisons are presented in Figure 3 and highlight the dimeric organisation of this sequence. For instance, monomer 1 shares between 83% and 89% sequence similarity with monomers 3, 5, 7, and 9, but only 67% to 72% sequence similarity with monomers 2, 4, 6, 8 and 10. No higher order repeat structures are apparent from this analysis, or from a comparison of the dimers present in this sequence (data not shown).

The data in Figure 3 also suggests that a single unequal crossover event has occurred within monomers 11 to 13 as the

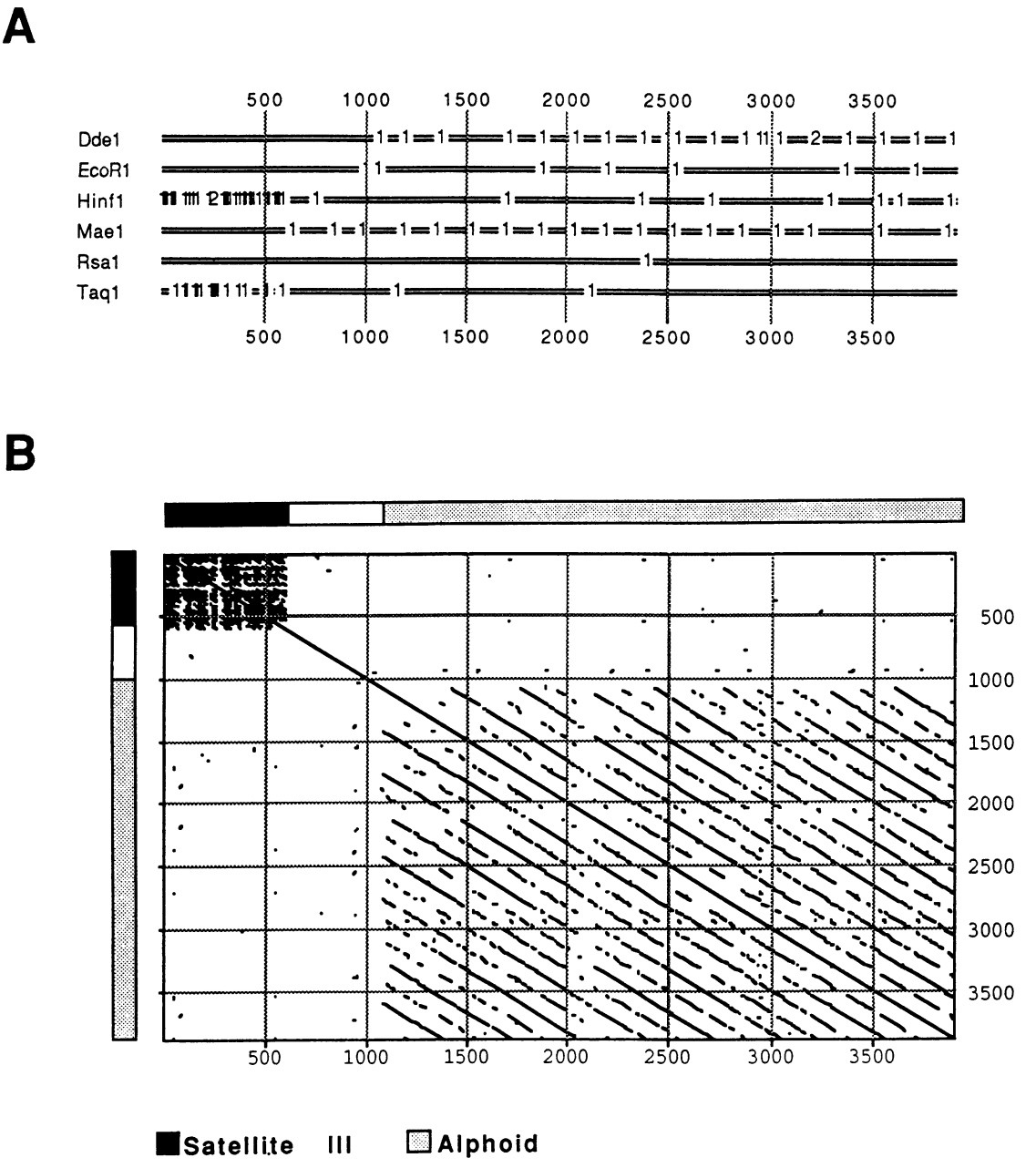

Figure 2. A. Restriction map of sequence obtained from the cMEN375 subclones p375M2.4 and mC375H. **B.** Dot matrix of sequence derived from cMEN375 subclones (self comparison). The location of satellite III and alphoid DNA within the sequence are shown on both axes. The window size was 25bp, with a minimum match of 70% required to register a dot.

regular dimeric organisation seen in monomers 1 to 10 changes register by one monomeric unit at this point. This is apparent as monomers 11 and 14 share 82%−92% sequence similarity with monomers 1, 3, 5, 7 and 9, whereas monomers 12 and 13 share only 66%−75% similarity to these monomers. A similar pattern is observed when monomers 10 and 13, and 11 and 12 are compared to monomers 2, 4, 6, and 8. (The relevant similarity scores are underlined in Figure 3). Although a change of register could be caused by a deletion or a duplication, these events would have a high probability of resulting in altered monomer lengths which is not observed. A single crossover event within monomers 11 to 13 is more consistent with the sequence arrangement.

It is also clear from Figure 3 that there is no evidence of sequence divergence close to the end of the alphoid array as monomers adjacent to the satellite III sequences give sequence similarity scores comparable to the other monomers in the array.

The 485bp sequence separating the satellite III and alphoid repeats was used to search the EMBL data base and was found to share 85−90% sequence identity with long interspersed repetitive (L1) elements. The L1 sequence precisely defines the boundary between the satellite III and alphoid sequences, and is both truncated and rearranged relative to full length L1 sequences (44, 45). Complex rearrangements in L1 sequences have been previosly reported (reviewed in ref. 46). In addition,

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16		Nucleotides
1.0	0.67	0.85	0.70	0.88	0.70	0.83	0.73	0.87	0.72	0.85	**0.66**	**0.72**	0.84	0.72	0.87	**1**	1089-1257
	1.0	0.67	0.84	0.72	0.87	0.68	0.88	0.71	0.88	**0.74**	**0.74**	0.79	0.70	0.89	0.73	**2**	1258-1429
		1.0	0.69	0.90	0.72	0.87	0.72	0.89	0.72	0.84	**0.67**	**0.73**	0.89	0.71	0.89	**3**	1430-1598
			1.0	0.74	0.90	0.70	0.89	0.71	0.89	**0.73**	**0.73**	0.79	0.72	0.89	0.74	**4**	1599-1769
				1.0	0.77	0.90	0.77	0.93	0.78	0.90	**0.71**	**0.77**	0.92	0.76	0.94	**5**	1770-1938
					1.0	0.70	0.91	0.73	0.91	**0.71**	**0.77**	0.80	0.75	0.91	0.78	**6**	1939-2096
						1.0	0.72	0.87	0.74	0.82	**0.68**	**0.72**	0.85	0.72	0.89	**7**	2097-2263
							1.0	0.75	0.89	**0.78**	**0.75**	0.82	0.76	0.92	0.79	**8**	2264-2434
								1.0	0.75	0.87	**0.67**	**0.75**	0.90	0.74	0.93	**9**	2435-2601
									1.0	**0.78**	**0.77**	0.81	0.76	0.92	0.79	**10**	2602-2772
										1.0	0.72	0.75	0.83	**0.78**	0.87	**11**	2773-2930
											1.0	0.78	**0.69**	**0.78**	**0.71**	**12**	2931-3101
												1.0	**0.76**	0.82	**0.77**	**13**	3102-3271
													1.0	0.75	0.95	**14**	3272-3437
														1.0	0.78	**15**	3438-3608
															1.0	**16**	3609-3777

Figure 3. Genetic distances between the 16 complete alphoid monomers sequenced from cMEN375. Comparisons which imply an unequal crossing over event are underlined (see text). The monomers were aligned using the PILEUP programme with gaps being introduced to maximise alignment. Numbers indicate% sequence similarity between monomers after alignment.

the L1 element sequenced here has no direct repeats which are commonly found flanking L1 elements (46).

Confirmation of sequence arrangement by PCR

We confirmed the integrity of the sequence information by performing PCR experiments using the primers schematically represented in Figure 4a. Results from two of these experiments are shown in Figures 4b and 4c. The repetitive nature of the sequences under study can make the pattern of PCR products complex. For instance, PCR products are obtained when primers C (satellite III primer) and B (alphoid primer) are used separately (Fig. 4b lanes 9 and 10 respectively) and many of these products are also produced when both primers are used together (Fig. 4b lane 11). However, a product of the expected size (1294bp) is only obtained when the primers are used in combination (lanes 11 and 12). The PCR products obtained when somatic cell hybrids are used instead of total human DNA are much less complex (lanes 2 to 6) and a band of the expected size is clearly visible in the chromosome 10-only hybrid R342A4 and the hybrids TG3 and 64034p6 (lanes 3, 4 and 6). The result with the L1 primers D and E is clearer (Fig. 5c), with a major PCR product of the expected size (476bp) being obtained in the cosmid and human controls (lanes 10 and 11) and in the hybrids R342A4, TG3 and 64034p6 (lanes 2, 3 and 5).

The results of both these experiments map the alphoid/satellite III boundary to the interval 10cen-10q11.2. This result is consistent with the hybridisation mapping for the alphoid clone p375H4.0 (not shown) and the total cosmid mapping of Mole *et al.* (1991). However, preliminary hybridisation data indicated that satellite III sequences from p375M2.4 mapped more distal in 10q11.2-10q23 (data not shown).

To investigate further the disposition of satellite III sequences in chromosome 10 PFGE experiments were performed.

Figure 5a shows various digests of R342A4 DNA probed with satellite III sequences from pMEN375. A single band of hybridisation approximately 900kb in size is obtained with any combination of the enzymes *Bgl*II, *Bst*eII and *Hin*dIII demonstrating that there is a single major block of satellite III sequence of approximately 900kb on chromosome 10 which is defined by these enzymes. The enzyme *Bam*HI cuts several times within this block to produce fragments of 200−350kb.

Digests of DNA from R342A4, TG3 and TK2 hybridised with the same probe are presented in Figure 5b. The R342A4 digests (lanes 1, 4 and 7) give the same 900kb fragment seen in Figure 5a. The hybrid TK2 also contains a single large fragment homologous to satellite III sequences, of approximately 750kb in addition to many smaller weaker hybridising bands (lanes 3, 6 and 9). The hybrid TG3 also contains several much smaller fragments that hybridise with this probe. The interpretation of these results is complicated by the fact that chromosomes other than chromosome 10 are present in the hybrids TG3 and TK2; however, it is clear that satellite III sequences are present in TG3. Therefore, it is likely that although the majority of satellite III sequences are present distal to 10q11.2, the satellite III sequences of cMEN375 are localized to 10cen-q11.2.

DISCUSSION

We have described here a junction between alphoid and satellite III sequences which maps to chromosomal region 10cen-q11.2. This is the first sequence information from a junction between two major repeat families which has been mapped to the centromeric region of a human chromosome. The organisation of the sequences at the alphoid/satellite III boundary provides several clues as to the formation of the present sequence arrangement. First, the sequence identity of large tandem arrays such as alphoid DNA is thought to be maintained by unequal crossing over (16) and several studies have provided evidence of unequal crossing over in the formation of existing alphoid arrays (43, 47). One prediction of an unequal crossing over model of maintenance of sequence identity is that sequences at the end of a large array would be diverged relative to the consensus sequence of the array as they would be involved in fewer unequal crossing over events. However, the dot matrix analysis (Fig. 2b) and the sequence analysis between alphoid monomers (Fig. 3) clearly show that there is no evidence of sequence divergence towards the end of the alphoid array. Second, a 485bp L1 element is present at the junction between the two sequence types.

Both of these observations, that the alphoid monomers are not degenerate at the end of the alphoid array and that the L1 sequence precisely separates the alphoid and satellite III sequences, suggest that the present sequence arrangement is not due to a simple L1

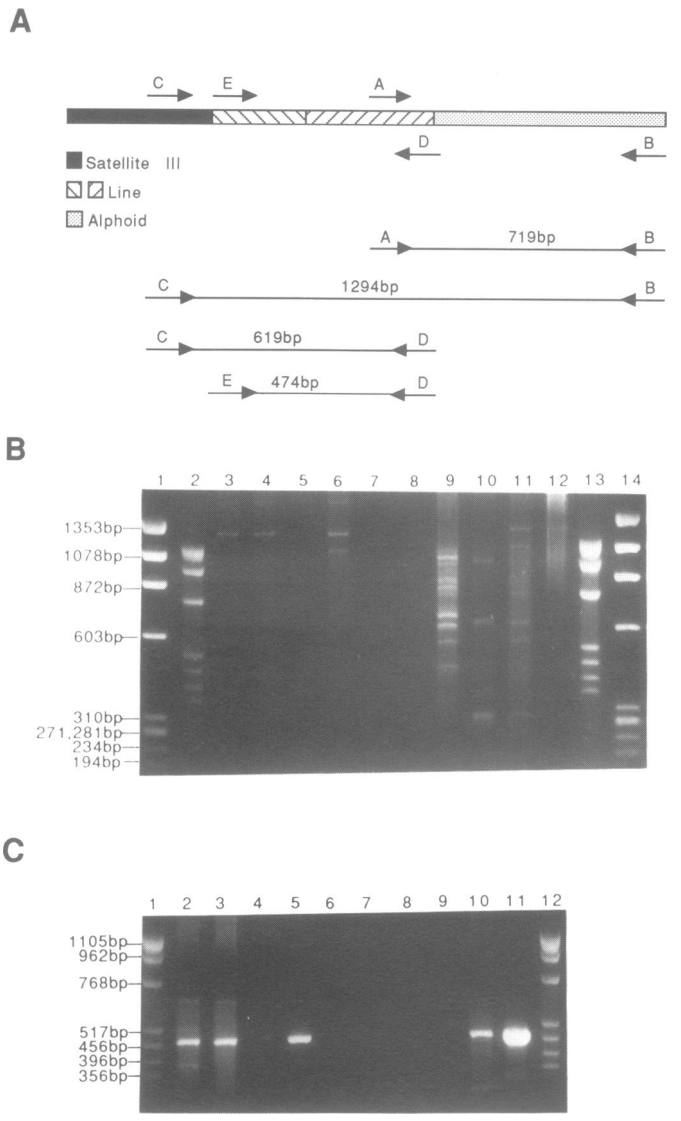

Figure 4 A. Schematic representation of PCR primers and products used to confirm sequence arrangement for cMEN375. The position of satellite III, L1 and alphoid sequences are shown. The different hatching patterns within the L1 sequence represent a rearrangement of these sequences relative to published sequences. For primers see Materials and Methods. **B.** PCR from satellite III to alphoid sequences. Lanes 1 and 14; ϕX174 *Hae*III marker. Lanes 2 and 13; Marker—equimolar mixture of pBSKS$^-$ *Xho*I digest and pBSKS$^-$ *Hin*fI digest. Lane 3; 64034p6, Primers B+C. 4; R342A4, Primers B+C. 5; TK2, Primers B+C. 6; TG3, primers B+C. 7; Hamster, primers B+C. 8. Mouse, primers B+C. 9; Human, Primer C. 10; Human, primer B. 11; Human, primers B+C. 12; cMEN375, primers B+C. **C.** PCR from L1 primers in cMEN375. Primers D+E were used throughout. Lanes 1 and 12 pBSKS$^-$ molecular weight marker. 2; 64034p6. 3. R342A4. 4; TK2. 5;TG3. 6; Hamster. 7; Mouse. 8. No DNA. 9; no DNA. 10; Human. 11; cMEN375. For details of hybrids used see Materials and Methods.

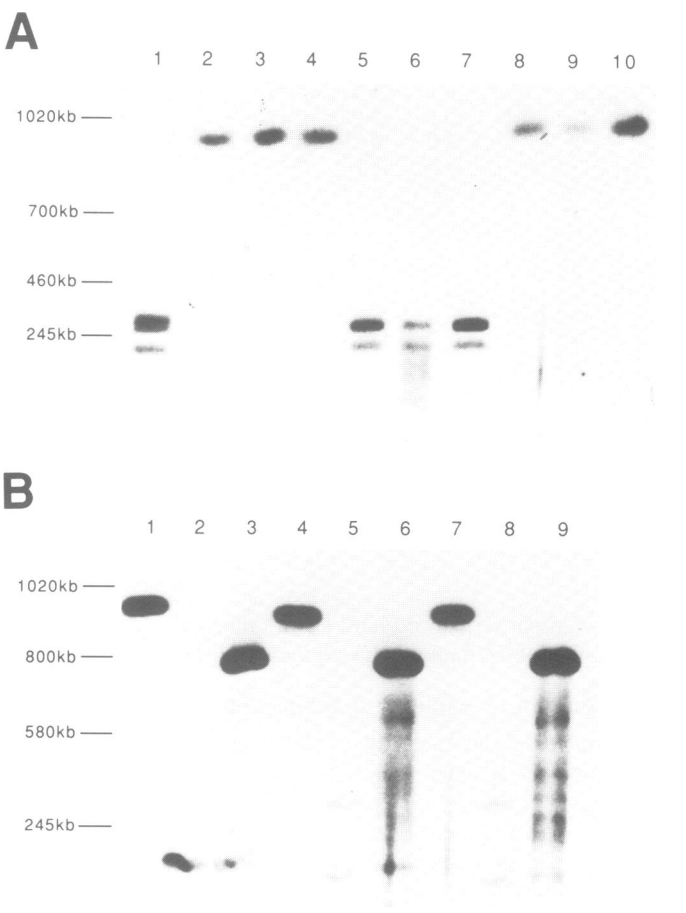

Figure 5 A. PFGE analysis of satellite III sequences in R342A4. The 2.1kb *Bam*HI/*Eco*RV fragment (Figure 1) of p375M2.4 was used to probe DNAs from the chromosome 10-only hybrid R342A4. Digests were as follows: Lane 1; *Bam*HI, 2; *Bgl*II,. 3; *Bst*eII, 4; *Hin*dIII, 5; *Bam*HI/*Bgl*II, 6; *Bam*H1/*Bst*eII, 7; *Bam*HI/*Hin*dIII, 8; *Bgl*II/*Bst*eII, 9; *Bgl*II/*Hin*dIII. 10; *Bst*eII/*Hin*dIII. The filter was washed at low stringency prior to exposure. **B.** PFGE analysis of satellite III sequences present in TG3 and TK2. Lanes 1; R342A4, 2; TG3, 3; TK2, digested with *Bgl*II. Lanes 4; R342A4, 5; TG3, 6; TK2, digested with *Hin*dIII. Lanes 7; R342A4, 8; TG3, 9; TK2, digested with *Bgl*II/*Hin*dIII. The probe was the 2.1kb *Bam*HI/*Eco*RV fragment of p375M2.4. The filter was washed at low stringency prior to exposure.

insertion at an alphoid/satellite III boundary. It seems more likely that it is a result of a recombination event between two L1 elements, one embedded in satellite III sequences and one embedded in alphoid sequences. An event such as this is particularly likely at the end of a long tandem array in which unequal crossing over is occurring. We propose that misalignment of the alphoid monomers (or satellite III sequences) during an unequal crossing over event is likely to result in non-homologous sequences at the ends of the repetitive arrays being brought into close proximity. Interspersed repetitive elements within these sequences could provide short regions of homology in otherwise unrelated sequences. These could then act as a substrate for a recombination event which would result in the juxtaposition of satellite III sequences alongside alphoid sequences with only a rearranged interspersed repeat element separating the two sequence types.

Examples of genetic rearrangement involving interspersed repetitive elements have been extensively documented at the LDL receptor locus (48) and the STS locus (49). These euchromatic rearrangements are known to be deleterious, but the extensive polymorphism found in alphoid arrays (eg: 6, 50–52) and satellite DNA (9) suggests that the same is unlikely to be true for rearrangements involving elements in repetitive sequences. Interspersed repetitive elements have been found in alphoid sequences before (14, 15, 53, 54). For instance Thayer and Singer (54) and Potter (15) describe L1 elements which interrupt alphoid

DNA, and which are clearly due to simple insertion events. Thus it seems plausible that rearrangements involving interspersed repetitive elements may be observed at a much higher frequency in alphoid and satellite sequences than in euchromatic sequences.

The alphoid sequences described here are arranged in a dimer formation characteristic of this DNA family. There is no evidence of a higher order repeat structure in the 16 monomers sequenced, although the existence of a higher order structure too large to be recognised from this sequence cannot be ruled out. Despite this, the organisation is distinct from that seen in the chromosome 10 clones pα10RP8 and pα10RR6 (34) which recognises a chromosome 10-specific alphoid family characterised by higher order repeat units of 6 and 8 monomers. The alphoid sequences from cMEN375 share as much sequence identity with alphoid DNA from other chromosomes as they do with these chromosome 10 alphoid clones demonstrating that they do not belong to the same chromosome 10 specific alphoid family.

Several authors have reported the existence of more than one alphoid family with distinct higher order repeat lengths on a single chromosome (6, 55). In addition, many alphoid clones have been obtained which cross-hybridise to arrays on more than one chromosome (3). The discovery of a second family of alphoid DNA on chromosome 10 is, therefore, not unexpected. However, it has yet to be determined if the cMEN375 sequences are contiguous with the alphoid array defined by pα10RP8 or with the 900kb satellite III block defined by p375M2.4.

We have shown here that the identification and sequencing of a boundary between two repetitive sequence types can produce a unique PCR tag which can complement and verify the mapping information produced by hybridisation techniques. It is possible that the identification of such unique tags from within complex sequences may provide a tool for mapping areas of the genome otherwise refractory to detailed analysis.

ACKNOWLEDGEMENTS

We thank A.Colomer and J.K.Moore for their technical assistance, Y.Nakamura for making available the cMEN series cosmids, C.Tyler-Smith for access to repetitive clones and D.Love and A.Tunnacliffe for critical reading of the manuscript. This work was funded by Medical Research Council H.G.M.P. Grant number G8916653 to B.A.J.Ponder. B.A.J.P. is a Gibb Fellow of the Cancer Research Campaign.

REFERENCES

1. Gosden, R.J., Mitchell, A.R., Buckland, R.A., Clayton, R.P. & Evans, H.J. (1975) *Expl. Cell. Res.*, **92**, 148–158.
2. Mitchell, A.R., Gosden, J.R. & Miller, D.A. (1985) *Chromosoma*, **92**, 369–377.
3. Choo, K.H., Vissel, B., Nagy, A., Earle, E. & Kalitsis, P. (1991) *Nucleic Acids Res.*, **19**, 1179–1182.
4. Choo, K.H., Earle,E. & Vissel, B. (1989) *Genomics*, **5**, 322–344.
5. Tyler-Smith, C. (1987) *Development*, **101**, 93–100.
6. Weverick, R. & Willard, H.F. (1991) *Nucleic Acids Res.*, **19**, 2295–2301.
7. Beauchamp, R.S., Mitchell, A.R., Buckland, R.A. & Bostock, C.J. (1979) *Chromosoma*, **71**, 153–166.
8. Higgins, M.J., Wang, H., Shtromas, I., Haliotis, T., Roder, J.C., Holden, J.J.A. & White, B.N. (1985) *Chromosoma*, **93**, 77–86.
9. Choo, K.H., Earle, E., Vissel, B. & Kalitsis, P. (1992) *Am. J. Hum. Genet.*, **50**, 706–716.
10. Choo, K.H., Earle, E. & McQuillan, C. (1990) *Nucleic Acids Res.*, **18**, 5641–5648.
11. Rocchi, M., Archidiacono, N., Ward, D.C. & Baldini, A. (1991) *Genomics*, **9**, 517–523.
12. Mitchell, A.R., Jeppesen, P., Hanratty, D. & Gosden, J. (1992) *Chromosoma*, **101**, 333–341.
13. Agresti, A., Rainaldi, G., Lobbiani, A., Magnani, I., DiLernia, R., Menervi, R., Siccardi, A.G. & Ginelli, E. (1989) *J. Mol. Biol.*, **205**, 625–631.
14. Grimaldi, G. & Singer, M.F. (1983) *Nucleic Acids Res.*, **11**, 321–338.
15. Potter, S.S. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1012–1016.
16. Smith, G.P. (1976) *Science*, **191**, 528–535.
17. Barker, R.F., Harberd, N.P., Jarvis, M.G. & Flavell, R.B. (1988) *J. Mol. Biol.*, **201**, 1–17.
18. McMullen, M.D., Hunter, B., Phillips, R.L. & Rubenstein, I. (1986) *Nucleic Acids Res.*, **14**, 4953–4968.
19. Wong, Z., Wilson, V., Jeffreys, A.J. & Thein, S.L. (1986) *Nucleic Acids Res.*, **14**, 4605–4616.
20. Dover, G.A. (1982) *Nature (London)*, **299**, 111–117.
21. Mole, S.E., Papi, L., Telenius, H., Moore, J.K., Tunnacliffe, A., Nakamura, Y. & Ponder, B.A.J. (1991) *Cytogenet. Cell Genet.*, [A27104].
22. Tokino, T. Takiguchi, S., Tanigami, A., Bragg, T., Jones, C. & Nakamura, Y. (1991) *Genomics*, **12**, 401–402.
23. Sambrook, J., Fitch, E.F. & Maniatis, E.F. (1989) Molecular cloning: a laboratory manual. 2nd Ed. Cold Spring Harbor Laboratory Press, New York.
24. Feinberg, A. & Vogelstein, B. (1984) *B. Anal. Biochem.*, **137**, 266–267.
25. Kraft, R., Tardiff, J., Krauter, K.S. & Leinwand, L.A. (1988) *Biotechniques*, **6**, 544–6, 549.
26. Zhang, W., Hu, G. & Deisseroth, A. (1991) *Nucleic Acids Res.*, **19**, 6649.
27. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.*, **12**, 387–395.
28. Trask, B.J., Ger van den Engh, M.C., Massa, H.F., Gray, J.W. & Van Dilla, M. (1991) *Som. Cell Mol. Genet.*, **17**, 117–136.
29. Fisher, J.H., Kao, F.T., Jones, C., Tyler-White, R., Benson, B.J. & Mason, R.J. (1987) *Am. J. Hum. Genet.*, **40**, 503–511.
30. Mathew, C.G.P., Wakeling, W., Jones, E., Easton, D., Fisher, R., Strong, C., Smith, B., Chin, K., Little, P., Nakamura, Y., Shows, T.B., Jones, C., Goodfellow, P.J., Povey, S. & Ponder, B.A.J. (1990) *Ann. Hum. Genet.*, **54**, 121–129.
31. Cooke, H.J. & Hindley, J. (1979) *Nucleic Acids Res.*, **6**, 3177–3197.
32. Metzdorf, R., Gottert, E. & Blin, N. (1988) *Chromosoma*, **97**, 154–158.
33. Menervi, R., Agresti, A., DellaValle, G., Talarico, D., Siccardi, A.G. & Ginelli, E. (1985) *J. Mol. Biol.*, **186**, 483–489.
34. Devilee, P., Kievits, T., Waye, J.S., Pearson, P.L. & Willard, H.F. (1988) *Genomics*, **3**, 1–7.
35. Poltoratsky, V.P. & Podgornaya, O.J. (unpublished). X60726.
36. Alexandrov, A., Mitkevitch, S.P. & Yurov, Y.B. (1986) *Chromosoma*, **96**, 443–453.
37. Waye, J.S., England, S.B. & Willard, H.F. (1987) *Mol. Cell. Biol.*, **7**, 349–356.
38. Greig, G.M., England, S.B., Bedford, H.M. & Willard, H.F. (1989) *Am. J. Hum. Genet.*, **45**, 862–872.
39. Baldini, A., Smith, D.I., Rocchi, M., Miller, O.J. & Miller, D.A. (1989) *Genomics*, **5**, 822–828.
40. Baldini, A., Smith, D.I., Rocchi, M., Miller, O.J. & Miller, D.A. (unpublished). M29461.
41. Baldini, A., Rocchi, M., Archidiacono, N., Miller, O.J. & Miller, D.A. (unpublished). M28221.
42. Looijenga, L.H.J., Oosterhuis, J.W., Smit, V.T.H.B.M., Wessels, J.W., Mollevanger, P. & Devilee, P. (1992) *Genomics*, **13**, 1125–1132.
43. Waye, J.S. & Willard, H.F. (1986) *Mol. Cell. Biol.*, **7**, 3156–3165.
44. Skowronski, J., Fanning, T.G. & Singer, M.F. (1988) *Mol. Cell. Biol.*, **8**, 1385–1397.
45. Hattori, M., Hidaka, S. & Sakaki, Y. (1985) *Nucleic Acids Res.*, **13**, 7813–7827.
46. Hutchison, C.A., Hardies, S.C., Loeb, D.D., Shehee, W.R. & Edgell, M.M. (1989) In Berg, D.E. & Howe, M.M. (ed.), Mobile DNA, American Society for Microbiology., Washington, D. C., pp 593–617.
47. Waye, J.S. & Willard, H.F. (1986) *Nucleic Acids Res.*, **14**, 6915–6927.
48. Hobbs, H.H., Russell, D.W., Brown, M.S. & Goldstein, J.L. (1990) *Annu. Rev. Genet.*, **24**, 133–170.
49. Yen, P.H., Li, X-M., Tsai, S-P., Johnson, C., Mohandas, T. & Shapiro, L.J. (1990) *Cell*, **61**, 603–610.
50. Haaf, T. & Willard, H.F. (1992) *Genomics*, **13**, 122–128.
51. Wu, J. & Kidd, K.K. (1990) *Hum. Genet.*, **84**, 279–282.
52. Waye, J.S. & Willard, H. F. (1989) *Chromosoma*, **97**, 475–480.
53. Lerman, M.I., Thayer, R.E. & Singer, M.F. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3966–3970.
54. Thayer, R.E. & Singer, M.F. (1983) *Mol. Cell. Biol.*, **3**, 967–973.
55. Choo, K.H., Earle, E., Vissel, B. & Filby, R.G. (1990) *Genomics*, **7**, 143–151.