

A VNTR immediately adjacent to the human pseudoautosomal telomere

C.F.Inglehearn and H.J.Cooke¹

Department of Human Genetics, Newcastle University and ¹MRC Human Genetics Unit, Western General Hospital, Edinburgh, UK

Received November 2, 1989; Accepted November 22, 1989

EMBL accession no. X17009

ABSTRACT

The probe 29C1 detects a hypervariable locus 18kb from the telomere of the human X and Y chromosomes, in the pseudoautosomal region. Here we report that hypervariability of fragments containing this sequence in the human population arises by loss or gain of a 31 base pair GC rich repeat. Labelled 29C1 does not detect a DNA fingerprint at low stringency, though the consensus repeat sequence does show some similarity to previously reported minisatellites. Sequence within the repeat block has G and C rich strands, a feature associated with sequences at the telomeres of many higher organisms. The repeat block shows sequence characteristics normally associated with a low methylation island, though the locus is methylated and does not appear to be transcribed.

INTRODUCTION

The phenomenon of sequence hypervariability has been found at a number of loci in the human genome. By chance, Wyman and White [1] isolated an anonymous DNA fragment found to define a locus with at least 8 alleles in the human population. Other hypervariable loci have been described in humans near the light chain immunoglobulin locus [2] the insulin gene [3], the alpha globin gene cluster [4,5,6,7], the Harvey-ras oncogene [8], the type II collagen gene [9] and the apolipoprotein B gene [10]. Loci differ in degree of variability with reports of as few as 6 alleles [11] or as many as 80 [12]. In each case, hypervariability is the result of variation in copy number of simple repeats known as minisatellites. Two such repeats were found to detect families of related sequences in loci scattered throughout the human genome [13,14], the genomes of other vertebrate species [15,16,17,18,19]; and even in plants [20]. Several other repeats have since also been reported to detect dispersed families of minisatellites in animals and plants [7,21,22,23,24,25]. The hypervariability of these loci, and in some cases their similarity to the lambda 'Chi' recombination signal led Jeffreys *et al* (1985a) [13] to propose that these sequences are recombination hotspots. Consistent with this hypothesis, Steinmetz *et al* [26] found that a minisatellite locus in the mouse major histo-compatibility complex was a site of frequent recombination. Further, a Jeffreys 'polycore' probe has recently been shown to bind selectively to sites of chiasmata on 'in situ' meiotic chromosome spreads [27]. However, in one case generation of a new allele at a minisatellite

locus has been shown not be associated with exchange of flanking markers [28].

Several such hypervariable loci have been mapped to the pairing region of the human sex chromosomes, at their short arm tips [29,30,22]. These have been used to demonstrate a gradient of sex linkage, proving the existence of a 'pseudoautosomal' region on the X and Y where markers recombine relative to sex [31,32]. Map distances in the region are additive, suggesting that a single crossover is required for XY pairing. An obligatory crossover in a stretch of DNA less than 5 megabases in length [32,33] would constitute a recombination hotspot, consistent with the observed high level of hypervariability in the region.

One hypervariable probe, clone 29C1, defines a locus, DXYS 14, within 20 kb of the pseudoautosomal telomere [29]. In addition to contributing to the pseudoautosomal map, this probe has been used to study the terminal heterogeneity of the adjacent telomere in different tissues and at various stages of development [34]. Here we present sequence data from locus DXYS14 in order to identify the basis of polymorphism in the region.

MATERIALS AND METHODS

Cloning of Fragments Containing a DXYS14 Repeat Block

The isolation of cosmid CY29 and subclone 29C1 has been described elsewhere [29]. Clone Thy B17 was selected from a lambda L47 DNA library screened with probe 29C1. This library had been prepared from Hind III digested DNA of the cell line Thy B (described in 35), which had been size fractionated on a sucrose gradient by a method described elsewhere [36]. Clone A35 was isolated from a library prepared in the same way using DNA from the human lymphoblastoid cell line PES. These lambda clones were found to be unstable both in the rec A⁻ cell line Q359, and the rec BC⁻ sbc B⁻ cell line DL282. The same inserts were stable when subcloned into plasmids.

Blotting and Hybridisation

Genomic DNA's were digested by restriction enzymes, size fractionated on an agarose gel, then capillary blotted by the method of Southern, [37] onto Hybond-N nylon filters (Amersham). Filters were prehybridised in 7% SDS and 0.5M sodium phosphate buffer at 68° for 30 minutes. 29C1 insert labelled with (³²P) by the method of Feinberg and Vogelstein [38] was then added, and filters were hybridised overnight. High stringency washes were carried out at 65°C in 0.1×SSC and

0.1% SDS. Lower stringency washes were at 55°C in 2×SSC and 0.1% SDS. It should be noted that salmon sperm DNA used as a competitor in hybridisations will cause high background hybridisation, since 29C1, like many minisatellite probes, has homology to sequences in fish DNA [19].

Sequence Analysis

The complete sequence of the 1.3kb Taq I-Pst I subclone from clone 29C1 was obtained by chemical sequencing of the ends of this fragment, the 800 bp Hae III fragment within it, and two deleted fragments obtained by the method of Henicoff [39]. The ends to be sequenced were ³²P labelled either by filling in a 3' overhang with kleinow enzyme or kinase end labelling a 5' overhang. Sequence was obtained by the method of Maxam and Gilbert [40]. The repeat block had previously proved unreadable with the enzymic sequencing method of Sanger et al [41]. Data generated was analysed using programmes from the University of Wisconsin Genetics Computer Group sequence analysis package [42].

RESULTS

Isolation of DXYS14 Hypervariable Sequences

Cosmid CY29 has insert DNA derived from the human Y chromosome in hybrid cell line 3E7, as described elsewhere [29]. A subclone of CY29 known as 29C1 was found to detect a locus in the pairing region of the X and Y chromosomes, within 20kb of the telomere. When used as a probe at high stringency on a range of human DNAs it detects between two and four bands of variable lengths with different restriction enzymes. To further characterize locus DXYS14, 29C1 homologous DNA fragments were isolated from two cell lines. A 17kb Hind III fragment referred to as ThyB17 was isolated from a lambda L47 library of DNA from the Human-Mouse hybrid cell line Thy-B, [35], which contains a human X chromosome. Also a 5.7kb Hind III fragment given the name A35 was identified in an L47 library of DNA from human male lymphoblastoid cell line PES. This fragment mapped to the Y chromosome. Cosmid CY29 and these two lambda clones, each isolated from the DXYS14 locus of a different human sex chromosome, have been used to produce restriction maps of the subtelomeric regions on each of these chromosomes (Manuscript in preparation).

The DXYS14 minisatellite

Chemical sequencing of 29C1 revealed a tandem head to tail array of a GC rich repeat (Figure 1). Eleven almost identical copies were observed, with declining homology to the repeat sequence on the telomeric side. On the centromeric side of the block imperfect copies of the repeat continue to the end of the restriction fragment. Figure 2 displays as a dot plot the structure of the repeat block. Partial sequences of subclones from this locus on the Thy-B X chromosome and the PES Y chromosome were also obtained (data not shown). The same repeat unit was found in these clones. The sequence obtained showed that each repeat contains a site for the restriction enzyme Ava II. It was therefore possible from fragment size estimates and from complete digestion with Ava II to show that a 4kb Taq I fragment from the Thy-B X chromosome containing 29C1 homologous sequence consists largely of a head to tail array of around 100 repeats. Using a similar rationale, it was found that a 1.7kb Pst1 fragment from the PES Y chromosome contained an array of around 25 repeats. Data on these repeat blocks is shown in Table 2. Therefore

```

CGAGGTGATCCCGAGGGAGGAGCGGGGGTCTGAGGGTGTCCCGAGAGGACCGAGGGGTGGAGCGGGGG 70
GAGGGTCTCGAGATGGCCCGGAGGGAGTCCCGATPAGGAGGAGCGGCACTCTGGGGTGGTCCCGAGGGGAA 140
GAAGCGTCTGCTGTGGTCTGGAATGGGACGCGGGTCTGGGGTGGTCCCGAGGGGAGGCGGGGG 210
TCTGGGGTGTCCCGAGGGGAGGAGCGGGGGTCTGGGGTGGTCCCGAGGGGAGGCGGGGGTCTGGGGT 280
GSTCCCGAGGGGAGGAGCTGGGGTCTCTGGGTGGTCCCGAGGGGAGGCGGGGGTCTGGGGTGGTCC 350
TAAGAGGAGGAGCAGGGGTCTGGGGTGGTCCCGAGGGGAGGAGCGGGGGTCTGGGGTGGTCCCGAGGG 420
AGGAGCGGGGGTCTGGGGTGGTCCCGAGGGGAGGAGCGGGGGTCTGGGGTGGTCCCGAGGGGAGGAGCG 490
GGGGTCTGGGGTGGTCCCGAGGGGAGGAGCGGGGGTCTGGGGTGGTCCCGAGGGGAGGCGGGGGT 560
CTGGGATGGTCCCTAAGAGGAGGAGCGGGGGTCTGCAATGGTCTTCAGGGGTGGAGCAITGGGGTCTCC 630
TGTGGTCCGAGGGTGGAGCAGGGGGTCTGGGGTGGTACTCTTCGGGGGAGCAGCGCAITCTCTCTTT 700
TGGTCCGGTTCCTCAITCTGCTGATCTGGGGTCTCTGTGATCTCTGACCGGGGCGAGATGGAGGGTCAAG 770
GTAGGGAGGAGGAGTGGCAGCTTGGTCCCGAGGGGAGGAGCGGGGGTCTGGGGTGGTCCCGAGGGG 840
TGACCCATCCATAGGAGAAITGGACACCTCAGACTCTCTCAATCTTCGGCCAGTGGCAGGTCCCGAGTCTG 910
CCTTCCCTGGCTCTCTGAGGGTCTGAGGGTCTGAGGGTCTCTCTTTTCAITCTGGCAATTTTAAAAAATTC 980
TTCTADGATCTCAGTGTCAAGGCTGGCTGTGTGGCAGGCAAGGATCCGGTCTTTGCTGAGGTCTCCAC 1050
TCTAGTGGCTGAAGGAAAGGGTAATATAGCTGGAAGGTAATCTGGGGTGGTCCCGAGGGGAGGATCTACAT 1120
TTCATCTTAGAAGGGGATTTGACAGGAGACAGAACTCCAGACTCTCTGAAATTTCAAGAACTACTTTC 1190
CAGCCTGGACAAATTCGGGAGGCGCTCAITCTTACAAAATAAAAATTAAGAAATTCGCCAGTGGCATGG 1260
CACACTCTGTAGTCCCACTACTCTGGAGGCTGAGGGGAGGAGATCGCTTGGAGTGGGGTCCGAGGC 1330
    
```

Fig. 1 Sequence of the 1.3kb Taq I-Pst I fragment from 29C1. Sequence reads proximal to distal. Repeat units are noted by arrows.

Table I Table showing the repeat blocks cloned from locus DXYS14, the clones which contain them and the cell lines and chromosomes from which they derive.

Cell Line of Origin	Human Chromosome of Origin	Lambda Clone (length)	Subclone Analysed	No. Repeats in block
3E7	Y	Cy 29(21kb)	29C1 (1.8kb)	11
Thy B	X	Thy B 17(17kb)	Thy B4(4kb)	100
PES	Y	A 35(5.6kb)	—	25

variation in size of fragments at this locus on different chromosomes clearly arises from variation in copy number of the repeat identified.

Repeat units deriving from locus DXYS14 on the three chromosomes analysed were compared using the Wisconsin sequencing package programme 'Pretty'. The consensus repeat was found to be as follows:-

GGGAGGAGCGGGGGTCTGGGGTGGTCCCGAG

A separate consensus was obtained for repeats deriving from each clone. Consensuses for repeats from the 3E7 Y chromosome (clone 29C1) and from the Thy-B X chromosome are identical to that shown. The PES Y chromosome consensus repeat differs only by deletion of a C at position 28 in the sequence shown above, and addition of a G at position 20.

Similarity to Other Minisatellites

The DXYS14 repeat sequence bears varying degrees of similarity to several other GC rich minisatellites. Table 2 shows some of these sequences in best alignment to the DXYS14 repeat consensus. All of these repeats have been used to detect some sort of DNA 'fingerprint'. It may therefore be that locus DXYS14 is one of the hypervariable loci detected at low stringency by some of these probes, especially the M13 and 3' HVR consensus repeats. However, by comparing signal on the same southern

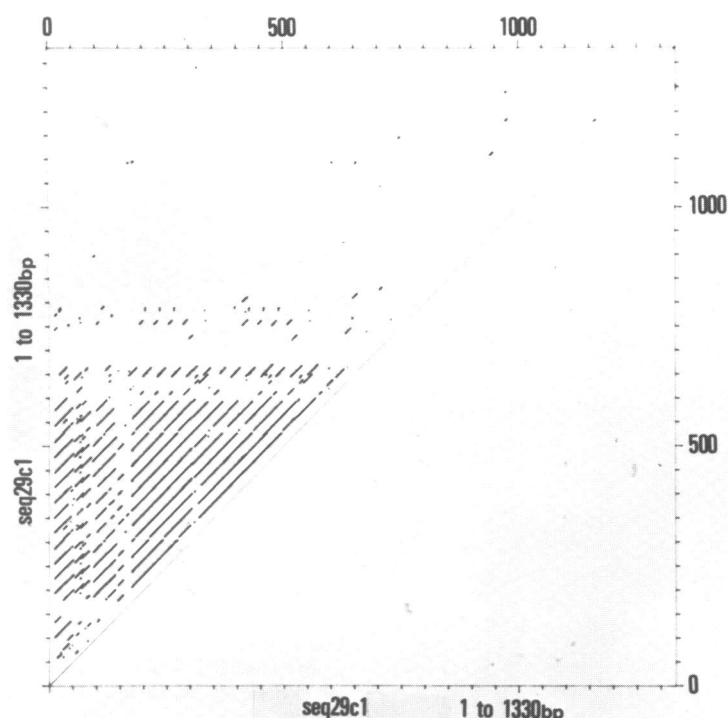


Fig. 2 Sequence shown in fig. 1, displayed as a dot plot. The repeat structure is visualised as a rectangular array of dots, each of which represents a match of 14 out of 21 basepairs in a comparison between the sequence and itself. The in register perfect homology appears as a continuous diagonal. Repeats are seen as out of register homologies parallel with the diagonal. The plot is symmetrical about the diagonal, so only one half of the square is drawn.

Table II Five sequences known to detect multiple loci scattered throughout the human genome, shown in their best possible alignments to the DXYS14 repeat consensus.

REPEAT	REFERENCE	REPEAT SEQUENCE
'Polycore' minisatellite	Jeffreys <i>et al</i> 1985a	G G A G G T G G G C A G G X G
M13 protein III gene repeat	Vassart <i>et al</i> 1987	G A G G G T G G X G G X T C T
Mouse repeat (3 possible alignments)	Georges <i>et al</i> 1987	A C N G G N A C N G G N A C N G G N
3' HVR consensus (2 possible alignments)	Jarman <i>et al</i> 1986	N G G G G N A C A G G N G G G G N A C A G
(CAC) ₅ oligo	Schafer <i>et al</i> 1988	G T G G T G G T G G T G G T G
DXYS14 consensus		A G G G G A G G A G C G G G G G T C T G G G G T G G T C C C G

blot probed alternately with polycore probes 33.15 and 33.6 with that of 29C1, these probes were found not to detect bands deriving from locus DXYS14 (data not shown). Low stringency blotting with 29C1 detects only a faint smear and some constant bands of less than 1kb, as well as bands known to derive from locus DXYS14. The location of sequences detected thus is unknown, and the signal is weak relative to intensity of the pseudoautosomal pattern.

Identification of a Novel Variant

Based on a survey of 29C1 homology in 43 unrelated individuals, and a further 51 individuals from four large pedigrees obtained from the CEPH DNA bank, the heterozygosity at locus DXYS 14 is estimated to be 97% (data not shown). Sequence analysis of the locus on three different chromosomes has clearly demonstrated that heterogeneity arises by variation in copy number of a minisatellite repeat. These observations imply a high

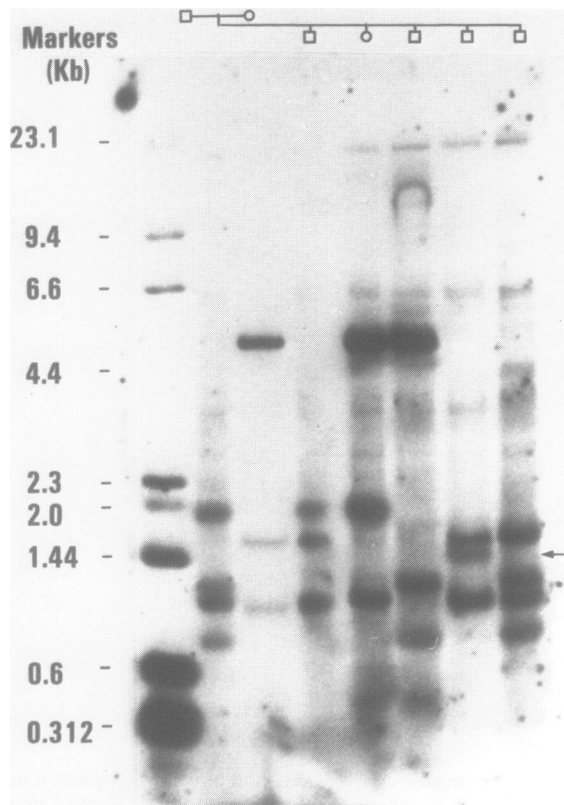


Fig. 3 C.E.P.H. family 104 DNAs digested with HaeIII, blotted, and probed with 29C1. Washes were performed at high stringency. Track 6 reveals the novel band, while other offspring represent the four possible combinations of unaltered parental alleles at this locus.

rate of spontaneous mutation to new length repeat blocks. In order to test this hypothesis, inheritance of 29C1 homology was analysed in the four large CEPH families studied. The products of 75 meiotic events were identified in these pedigrees. In one family a novel band was detected in an offspring which was not found in the patterns detected in either parent (figure 3). Analysis of 29C1 homology with a range of enzymes has confirmed that the new band is the result of a reduction in size of around 500bp (approximately 16 repeat units), in the length of a band derived from the father. Paternity was confirmed in this case using fingerprint probe 33.15. These figures suggest that the mutation rate at this locus is of the order of 1.3% per gamete. However, since the locus is duplicated in some individuals (manuscript in preparation), this figure overestimates the instability at each block of repeats.

In another CEPH family, not included in the data described above, six out of seven offspring had a band which appeared in neither parent. This may be the result of germline mosaicism, and is being further investigated.

Island-like features of the DXYS14 sequence

The cosmid CY29 was originally selected because it contained two sites for restriction enzyme Sst II. Sites for this enzyme are relatively rare, and are generally clustered at 'islands' of low methylation and high CpG, in mammalian genomic DNA [43]. These islands have in turn been shown to be associated with 5' ends of genes [44]. Sites for the enzymes Hpa II, Hha I and Sst II, each of which contain CpG in their recognition sequence, have subsequently been shown to be clustered in clone 29C1 and

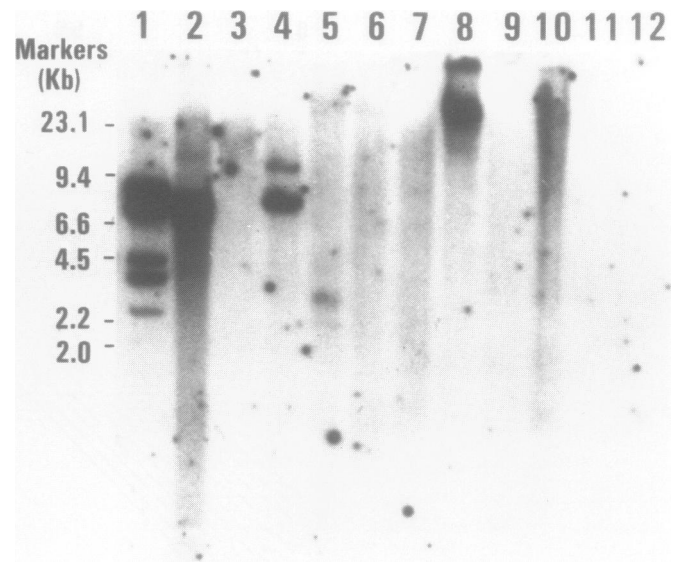


Fig. 4 Genomic DNAs of a range of organisms, EcoRI digested, blotted, and probed with 29C1. Washes were performed at low stringency. Sources are, from left to right, human, chimpanzee, orangutan, gorilla, pig, rat, mouse, chicken, xenopus toad, trout, *Schizosaccharomyces pombe* (fission yeast), and *Saccharomyces cerevisiae* (bakers yeast).

sequence immediately proximal to it [29]. The presence of these enzyme sites and a G + C content of 68% in and around the repeat block, are consistent with the presence of a CpG island at this locus. In addition, the dinucleotide CpG which normally occurs at only one fifth of its expected frequency in vertebrate DNA, occurs in the region of the repeat block at around half the level expected in a sequence of this composition, and at the same level as the equivalent CpG dinucleotide. However, the sequence is methylated in human genomic DNA prepared from blood, and neither clone 29C1, nor 29C4 from the proximal side of the repeat block, detected a transcript on a northern blot of Hela cell RNA (data not shown). It would appear therefore, that this locus does not contain a low methylation island, and that sequence characteristics alone cannot be taken as proof of the presence of an island.

Homology in Other Species

29C1 was used as a probe against EcoRI digested DNA of a variety of species (figure 4). Strong hybridisation signal can be seen in Gorilla and Chimp, which are both closely related to man. Orangutan, an old world monkey and so a more distant relative shows no homology. Two faint bands can be seen in Pig, while Mouse, Rat, Xenopus, *Saccharomyces cerevisiae*, and *Schizosaccharomyces pombe* gave no appreciable signal. Chicken DNA shows homology to 29C1 in undigested DNA at the top of the track, though the ethidium stained photo shows a complete digest (data not shown). Also in trout a smear can be seen running the length of the track, possibly implying homology to an interspersed repetitive sequence in that genome.

DISCUSSION

This report describes the hypervariable locus DXYS14, which is located immediately adjacent to the human pseudoautosomal telomere. Data presented proves that variation at this locus arises

by loss or gain of copies of a 31 basepair GC rich minisatellite repeat. The use of a probe from this locus in defining the end of the genetic map for the human pseudoautosomal region has been discussed elsewhere [29]. This probe is particularly useful since, like other VNTR loci, it is almost always informative due to high heterozygosity (97%)

The hypervariable loci first described were generally at internal sites on chromosomes, in or around genes, since these regions were the most closely studied. This observation together with fingerprinting studies [13,15] led researchers to believe that these loci were autosomal and dispersed. However, new evidence increasingly favours the hypothesis that hypervariable repeat blocks tend to be clustered in telomeric regions [45]. The telomeric location of locus DXYS14 constitutes further evidence for this hypothesis. In addition minisatellite loci are thought to be hotspots for recombination [13,26]. The pseudoautosomal region is known to have a high recombination rate in male meioses, since a single obligate crossover is required in this small region for XY pairing [31]. The presence of the DXYS14 minisatellite, together with other hypervariable loci [22] might therefore serve to facilitate this increased rate of crossover. However, the mere presence of minisatellites alone clearly cannot be the cause of increased recombination, since recombination in the pseudoautosomal region is not disproportionately high in female meioses [31]

It has been suggested that two distinct classes of minisatellites exist [22]. Those which are AT rich show only small length variation, while GC rich minisatellite loci vary greatly in length. Sequence from within the DXYS14 repeat block in clone 29C1 contains 72% GC, compared with 40% in bulk genomic DNA. Also the blocks studied in this paper contain between 11 and 100 repeat units while family studies have suggested that blocks of as many as 200 repeats at this locus exist in the human population (data not shown). Clearly the DXYS14 minisatellite is of the GC rich class and fulfills the prediction of varying greatly in length.

It is also interesting to note that within the repeat block studied there are clear G rich and C rich strands, similar in composition to those known to occur at the immediate telomeres of many species [46,47]. The G rich strand is the one which will terminate with a 3' end at the adjacent telomere, as is the case with known eukaryotic telomeres [48]. The sequence shown has no homology to known telomeric sequences. Neither is there any reason to suppose that blocks of GC rich repeats 10–20kb from the telomere are directly involved in telomere function. However it is interesting to speculate whether these repeats might function as 'reserve telomeres' for a chromosome which has lost its telomeric function. One way to further investigate this hypothesis would be to find whether the DXYS14 repeat unit could prime a 'telomerase' type addition reaction in tetrahymena extracts [49].

Though it is similar in composition to other GC rich repeats, the DXYS14 repeat does not detect sequences from other loci at high stringency. The weak smear and bands seen at very low stringency, together with the similarity demonstrated to other minisatellite repeats, especially the M13 protein III gene repeat and the 3' HVR minisatellite core, suggests that locus DXYS14 might be detected weakly by these probes, and may therefore be one of the loci which make up their fingerprint. Thus fingerprint bands detected by minisatellite probes may not in fact be exclusively autosomal, as was first observed [13,15] since this and other minisatellite loci [22] are now known to be in the XY pairing region.

Sequence homology of minisatellite loci in a range of animal and plant species is a common phenomenon, as stated above.

The observed cross-hybridization of the 29C1 probe with sequences in Gorilla, Chimpanzee, Pig, Chicken and Trout is therefore not in itself surprising, although perhaps lack of signal in Orangutan and rodents is. The observed similarity in intensity of signal in Human, Chimp and Gorilla suggests that this may represent the same locus. It would therefore be interesting to determine the location of hybridising sequences in Chimp and Gorilla, and to test whether these sequences are hypervariable. A study of the positions of a range of such loci in different species may enhance our understanding of the structure and evolution of the human genome and those of other organisms.

ACKNOWLEDGEMENTS

We thank Niolette McGill for technical assistance, Paul Sealey for help with sequence analysis, Susan Lindsey for constructive criticism and Adrian Bird and Nick Hastie for useful discussions. Thanks also to the Centre d'Etude du Polymorphisme Humaine (CEPH) for family DNA samples.

REFERENCES

- Wyman A.R., and White R., (1980) *Proc. Natl. Acad. Sci. USA*, 77 : 6754–6758
- Hieter P.A., Hollis G.F., Korsmeyer S.J., Waldmann T.A., and Leder P., (1981) *Nature*, 294 : 536–540
- Bell G.I., Selby M.J., and Rulter W.J., (1982) *Nature*, 295 : 31–35
- Higgs D.R., Goodburn S.E.Y., Wainscoat J.S., Clegg J.B., and Weatherall D.J., (1981) *Nucleic Acids Res.*, 9 : 4213–4224
- Proudfoot N.J., Gil A., and Maniatis T., (1982) *Cell*, 31 : 553–563
- Goodburn S.E.Y., Higgs D.R., Clegg J.B., and Weatherall D.J., (1983) *Proc. Nat. Acad. Sci. USA*, 80: 5022–5026
- Jarman, A., Nicholls R.D., Weatherall D.J., Clegg J.B., and Higgs D.R., (1986) *EMBO Journal*, 5 : 1857–1863
- Capon D.J., Chen E.Y., Levison A.D., Seeburg P.H., and Geoddel D.V., (1983) *Nature*, 302 : 33–37
- Stoker N.G., Cheah K.S.E., Griffin J.R., and Solomon E., (1985) *Nucleic Acids Res.*, 13 : 4613–4622
- Knott T.J., Wallis S.C., Pease R.J., Powell L.M., and Scott J., (1986) *Nucleic Acids Res.*, 14 : 9215–9216
- Sykes B.C., Ogilvie D.J., and Wardworth B.P., (1985) *Hum. Genet.*, 70: 35–37.
- Balazs I., Baird M.L., Wexler K., and Wyman A.R., (1986) *Am. J. Hum. Genet.*, 39 229.
- Jeffreys A.J., Wilson V., and Thein S.L., (1985)a. *Nature*, 314: 67–73.
- Jeffreys A.J., Wilson V., and Thein S.L., (1985)b. *Nature*, 316: 76–79
- Jeffreys A.J., Wilson V., Kelly R., Taylor B.A., and Bullfield G., (1987) *Nucleic Acids Res.*, 15: 2823–2826.
- Jeffreys A.J., and Morton D.B., (1987) *Animal Genetics*, 18: 1–15.
- Wetton J.H., Carter R.E., Parker D.T., and Walters D., (1987) *Nature*, 327: 147–152.
- Burke T., and Bruford M.W., (1987) *Nature*, 327 : 149–152
- Georges M., Lequarre A.S., Castelli M., Hanset R., and Vassart G., (1988) *Cytogenet. Cell Genet.*, 47: 127–131.
- Dallas J.F., (1988) *Proc. Natl. Acad. Sci. USA*, 85 : 6831–6835
- Georges M., Cochaux P., Lequarre A.S., Young M.W., and Vassart G., (1987) *Nucleic Acids Res.*, 15 7193.
- Simmler M.C., Johnsson C., Petit C., Rouyer F., Vergnaud G., and Weissenbach J., (1987) *EMBO Journal*, 6 : 963–969
- Vassart G., Georges M., Monsieur R., Brocas H., Lequarre A.S., and Christophe D., (1987) *Science*, 235: 683–684
- Schafer R., Zischler H., and Epplen J.T., (1988) *Nucleic Acids Res.*, 16: 5196
- Rogstad S.H., Patton J.C., and Schaal B., (1988) *Proc. Natl. Acad. Sci. USA*, 85 : 9176–9178
- Steinmetz M., Stephen D., and Lindahl K.F., (1986) *Cell*, 44: 894–904
- Chandley A.C., and Mitchell A.R., (1988) *Cytogenet. Cell Genet.*, 48 : 152–155
- Wolff R.K., Nakamura Y., and White R., (1988) *Genomics*, 3 : 347–351
- Cooke H.J., Brown W.R.A., and Rappold G.A., (1985) *Nature*, 317 : 687–692
- Goodfellow P.J., Darling S.M., Thomas S.N., and Goodfellow P.N., (1986) *Science*, 234 : 740–743

31. Rouyer F., Simmler M.C., Johnsson C., Vergnaud G., Cooke H.J., and Weissenbach J., (1986)a. *Nature*, 319 : 291–295.
32. Rouyer F., Simmler M.C., Vergnaud G., Johnsson G., Levilliers J., Petit C., and Weissenbach J., (1986)b. *Cold Spring Harbour Symp. Quant. Biol.*, 51 : 221–228
33. Pritchard C.A., Goodfellow P.J., and Goodfellow P.N., (1987) *Nature*, 328: 273–275
34. Cooke H.J., and Smith B.A., (1986) *Cold Spring Harbour Symp. Quant. Biol.*, 51 : 213–219
35. Lund E., Bostock C., Robertson M., Christie S., Mitchen J.L., and Dahlberg J.E., (1983) *Molec. Cell. Biol.*, 3 2211–2220
36. Maniatis T., Hardison R.C., Lacy E., Lauer J., O'Connell C., Quon D., Sim D.K., and Efstatidis A., *Cell*, 15 : 687
37. Southern E.M., (1975) *J. Mol. Biol.*, 98 : 503–517
38. Feinberg A.P., and Vogelstein B., (1983) *Analytical Biochem.*, 132 : 6–13
39. Henicoff S., (1984) *Gene*, 28 : 351–359
40. Maxam A.M., and Gilbert W., (1980) *Methods in Enzymology*, 65: 500–560
41. Sanger F., Nicklen S., and Coulson A.R., (1977) *Proc. Natl. Acad. Sci. USA*, 74 : 5463–5467
42. Deveraux J., Haerberli P., and Smithies O., (1984) *Nucleic Acids Res.*, 12: 387–395
43. Bird A.P., (1986) *Nature*, 321 : 209–213
44. Lindsay S., and Bird A.P., (1987) *Nature*, 327 : 336–338
45. Royle N.J., Clarkson R.E., Wong Z., and Jeffreys A.J., (1988) *Genomics*, 3 : 352–360
46. Allshire R.C., Gosden J.R., Cross S.H., Cranston G., Rout D., Sugawara N., Szostak J. W., Fantes P.A., and Hastie N.D., (1988) *Nature*, 332: 656–659.
47. Moyzis R.K., Buckingham J.M., Scott Cram L., Dani M., Deaven L.L., Jones D.J., Meyne J., Ratliff R.L., and Wu J.R., (1988) *Proc. Natl. Acad. Sci. USA*, 85 : 6622–6626
48. Blackburn E.H., (1984) *Cell*, 37 : 7–8
49. Greider C.W., and Blackburn E.H., (1985) *Cell*, 43 : 405–413.