The EMBO Journal vol.6 no.4 pp.963–969, 1987

Two highly polymorphic minisatellites from the pseudoautosomal region of the human sex chromosomes

M.C.Simmler, C.Johnsson, C.Petit, F.Rouyer, G.Vergnaud and J.Weissenbach

Unité de Recombinaison et Expression Génétique, INSERM U163, CNRS UA271, Institut Pasteur, 28 rue du Dr Roux, 75724 Paris Cédex, France

Communicated by P.Tiollais

Two pseudoautosomal loci DXYS15 and DXYS17 from the pairing region of the human sex chromosomes display a high variability with at least eight alleles each. The structural elements responsible for the polymorphisms have been isolated and sequenced. In both cases the variations result from DNA rearrangements occuring in tandemly repeated sequences (minisatellites) of 21–29 nucleotides for DXYS15 and 28–33 nucleotides for DXYS17. At reduced stringency, the DXYS15 minisatellite detects other hypervariable sequences located in other parts of the genome and hence represents a new family of minisatellites. In contrast to most other known hypervariable families, the DXYS15 hypervariable sequence displays a very high AT content.

Key words: minisatellite/hypervariable regions/human pseudoautosomal region

Introduction

Several polymorphic loci have recently been mapped to the tip of the short arms of the human X and Y chromosomes (Simmler *et al.*, 1985; Cooke *et al.*, 1985; Rouyer *et al.*, 1986a, b; Goodfellow *et al.*, 1986). These loci exhibit either partial or no sex linkage and therefore define a pseudoautosomal region in the terminal short arms of the human sex chromosomes (Burgoyne, 1982; Rouyer *et al.*, 1986a,b; Goodfellow *et al.*, 1986). The restriction fragment length polymorphisms (RFLPs) of many pseudoautosomal loci, i.e. DXYS14, DXYS15, DXYS17 and DXYS20 (Goodfellow *et al.*, 1985) are characterized by numerous allelic variations analogous to other multiallelic polymorphisms (Wyman and White, 1980; Higgs *et al.*, 1981; Bell *et al.*, 1982; Proudfoot *et al.*, 1982; Capon *et al.*, 1983; Jarman *et al.*, 1986). In all documented instances, such allelic variations result from differences in copy number of arrays of short, tandemly repeated sequences or minisatellites (Bell *et al.*, 1982; Proudfoot *et al.*, 1982; Ullrich *et al.*, 1982; Capon *et al.*, 1983; Goodbourn *et al.*, 1983; Jeffreys *et al.*, 1985a; Stoker *et al.*, 1985; Jarman *et al.*, 1986; Knott *et al.*, 1986). A number of these hypervariable minisatellites belong to larger families scattered in the human genome. Two such families have been identified to date and are each characterized by a common core sequence shared by the different members (Jeffreys *et al.*, 1985a; Jarman *et al.*, 1986). It has been suggested that the extreme variability of minisatellites from the myoglobin family is related to a high frequency of recombination in these sequences (Jeffreys *et al.*, 1985a). Recently a similar role was attributed to a minisatellite from the mouse major histocompatibility complex (Kobori *et al.*, 1986; Steinmetz *et al.*, 1986; Uematsu *et al.*, 1986). We have shown previously that the human pseudoautosomal region is a recombination hotspot in human male meiosis (Rouyer *et al.*, 1986a,b). It was thus of interest to determine if the hypervariability of pseudoautosomal loci also resulted from the presence of minisatellites. If so, these latter could be related to other hypervariable regions (HVRs) reported earlier, especially those with a putative role in recombination. Therefore the HVRs of loci DXYS15 and DXYS17 have been isolated, sequenced and used as probes in the search for other related minisatellite sequences.

Results

Multiallelic RFLPs of DXYS15 and DXYS17 detected by probes 113D and 601 respectively, have been described previously (Simmler *et al.*, 1985; Rouyer *et al.*, 1986a). In both cases a large number of alleles can be seen (Figure 1) among different individuals taken at random. This suggests that the variability could result from insertions and/or deletions occurring in a defined hypervariable region (HVR) as in many other reported instances. We therefore tried to locate accurately these HVRs, combining restriction mapping data and use of adjacent and/or overlapping probes.

Isolation of pseudoautosomal HVRs

Probe 601 was originally obtained as a subclone of a sequence isolated in clone 60 of a cDNA library synthesized from mRNA of the 3E7 cell line, a Y only somatic cell hybrid (Marcus *et al.*, 1976). To remove several Alu repeats present in the insert of clone 60, sonicated fragments of this latter were subcloned and screened for the absence of human repetitive DNA (see Materials and methods). Probe 601, from a subclone free of Alu sequences, was further used to probe Northern blots of human poly(A)$^+$ RNA from various tissues. However, no transcript was detected even in the 3E7 cell line, the original template mRNA source, and clone 60 may therefore result from artefactual cloning of a genomic DNA fragment in the cDNA library. Probe 601 detects multiple alleles in different restriction digests of human genomic DNA, indicating that this variability can be ascribed to DNA rearrangements. Five independent lambda clones (λ1–λ5) from a human genomic library were identified with probe 601. The size of the *Taq*I subfragment from these lambda clones hybridizing to probe 601 was identical to that of one allele in the DNA source. A *Bam*HI fragment from clone λ1 including this DXYS17 *Taq*I allele was subcloned in a plasmid and its restriction map was compared to that of clone 60 insert (Figure 2a). With the exception of the *Rsa*I–*Eco*RV fragment the two maps were identical, indicating that these sites encompassed the DXYS17 HVR (Figure 2a). This *Rsa*I–*Eco*RV fragment, probe 602 hereafter, was characterized further.

Probe 113D, a *Sau*3A fragment originally derived from a cosmid insert of human genomic DNA (Simmler *et al.*, 1985), maps between a *Bgl*II and a *Taq*I site of a *Taq*I fragment (Figure 2b) and detects a RFLP in *Taq*I, *Pst*I and *Xba*I digests of human male and female DNAs suggesting again that this RFLP results from insertion/deletion events. However, this probe did not detect

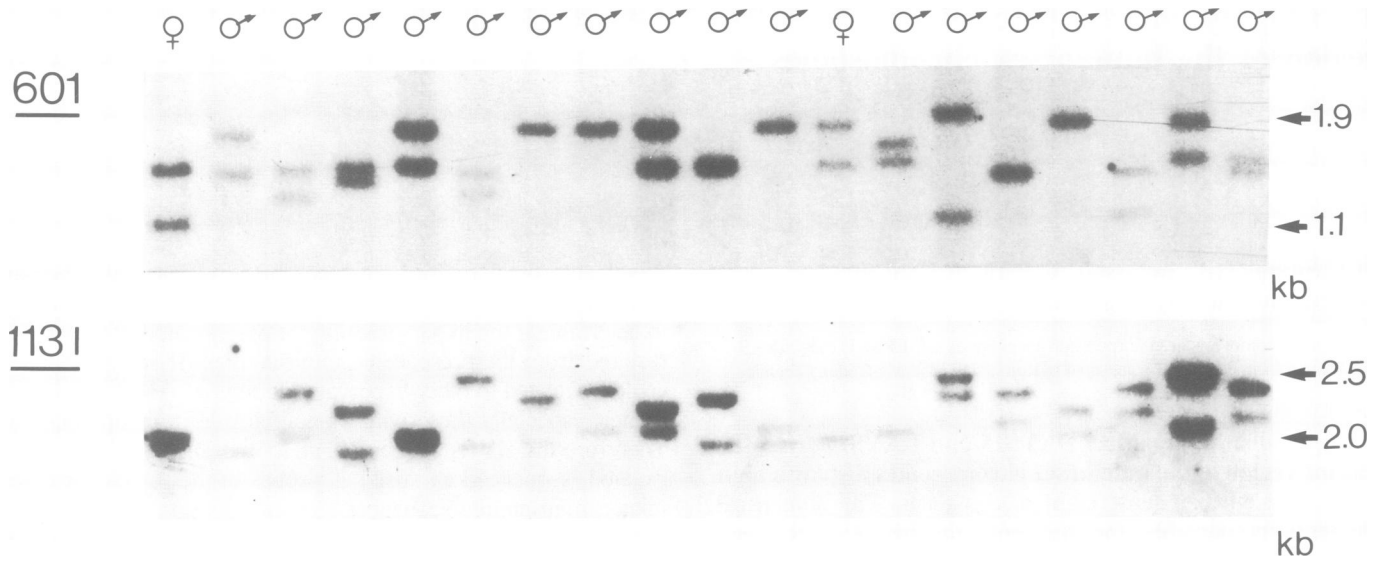

Fig. 1. Hypervariable pseudoautosomal fragments in DNAs from unrelated individuals. Analysis of human genomic DNA samples digested with *Taq*I and successively probed with (**a**) 601 (locus DXYS17) and (**b**) 113I (locus DXYS15).

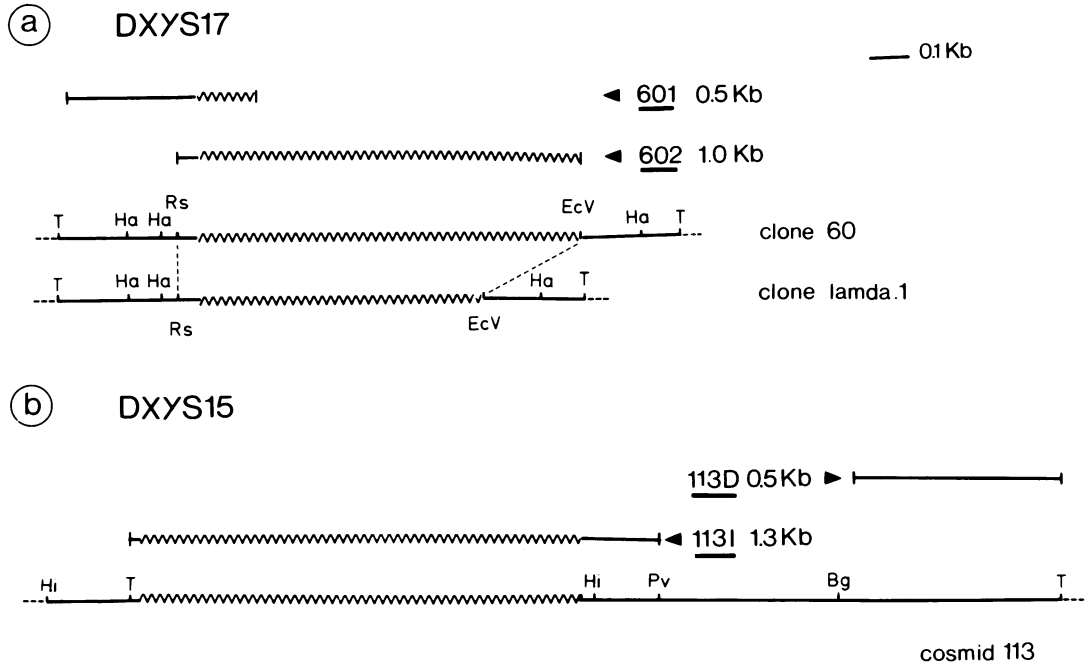

Fig. 2. Restriction mapping of pseudoautosomal HVRs. (**a**) DXYS17 HVR: location of probes 601 and 602 with respect to the hypervariable segments on maps of clones 60 and λ1. (**b**) DXYS15 HVR: location of probes 113D and 113I on a map representing a part of cosmid 113. Straight lines represent invariant DNA sequences and wavy lines represent parts displaying size variations. Enzyme sites shown are: T, *Taq*I; Ha, *Hae*III; Rs, *Rsa*I; EcV, *Eco*RV; Hi, *Hin*fI; Pv, *Pvu*II; Bg, *Bgl*II.

RFLPs in *Hin*fI and *Pvu*II DNA digests. Since both *Pvu*II and *Hin*fI restriction sites are located within the polymorphic *Taq*I fragment (Figure 2b), the HVR DXYS15 does not overlap with probe 113D. Conversely, the RFLP was detected in *Pvu*II and *Hin*fI digests when probed with the adjacent *Pvu*II−*Taq*I fragment, 113I (shown for *Hin*fI in Figure 3), confirming that the DXYS15 HVR was located between these *Pvu*II and *Taq*I sites. This variable 113I fragment was therefore subcloned into pBR for further manipulation.

Probes 601 and 113I have been used in Southern blot hybridization of various Caucasian male and female DNAs restricted with *Taq*I. Figure 1 shows the hybridization patterns obtained as a series of at least eight alleles at both loci. The sizes of the bands range from 1.1 to 1.9 kb when probed with 601 and from 2.0 to 2.5 kb when hybridized with 113I (Figure 1). In many instances a slight difference of hybridization intensity could be noticed for both polymorphisms, the larger band hybridizing more intensely. Moreover this difference in hybridization intensity was sharper between two alleles of extreme size and undetectable between very close alleles. This indicates that either the probe itself contains DNA sequences present in the higher allele and absent in the lower one (qualitative difference) or that the higher allele contains more copies of a repeated element present in both alleles and in the probe (quantitative difference). It follows that in any

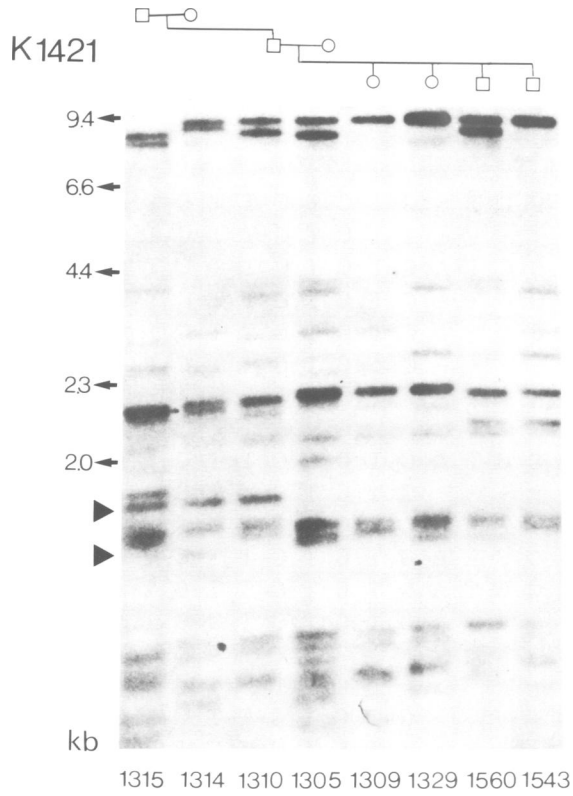

Fig. 3. Family analysis of HVRs related to DXYS15 HVR. Southern blot hybridization analysis of DNA samples from CEPH kindred K1421, digested with *Hin*fI, probed with 113I at low stringency (see Materials and methods). Arrowed fragments represent the cognate fragments from locus DXYS15, the only locus detected under standard hybridization conditions.

instance the probe recognizes and thus contains (at least partially) the variation causative element. This confirms the above mapping according to which both probes 601 and 113I overlap the HVRs to some extent. Another confirmation of this mapping comes from the fact that probe 113D does not detect any quantitative difference in hybridization intensity and thus does not recognize the source of the variation. Conversely, the differences in hybridization intensity appear even sharper with probe 602 as compared to the 601 pattern (not shown), suggesting that this latter contains less of the variable sequences or more of the non-polymorphic DNA, allowing a more accurate mapping of the DXYS17 HVR. In order to determine the origin of the high variability detected by and mapping within probes 602 and 113I, sequencing of the HVRs and adjacent regions has been performed using the M13 shotgun sequencing procedure.

DXYS17 HVR

Sequence analysis shows that ~1.0 kb of clone 60 and 0.75 kb of the lambda phage clone consist of a tandemly repeated unit of 28 or 33 bp (Figure 4). Elements of this repeated array are found both in 601 and 602 probes as shown in Figure 2a and represent ~95% of probe 602, hence accounting for the HVR of locus DXYS17. There is no obvious rule in alternating short (28) and long (33) repeats. The 33 monomer is built up of two adjacent subunits of 11 bp sharing many similarities followed by a third 11-bp sequence showing slight analogies with the middle unit (Figure 4). A somewhat similar organization, with two adjacent related domains is found in the IZ-HVR of the alpha globin complex (Goodbourn *et al.*, 1983). One of the terminal monomers misses the first two nucleotides of the consensus 28- and 33-bp monomers. Two consecutive 28-bp units adjacent to this terminal 25-bp unit differ also from the consensus 28-bp sequence by three nucleotides T, A and G instead of C, G and T in positions 19, 22 and 23. This stresses the analogies between the middle and third 11-bp subunits and also suggests that the 28 and 33 repeated monomers may have originated from one another through slippage or unequal exchange events. Whereas many positions of the monomer remain invariant, several others are more prone to variations. Variations are also more numerous in the edges of the array where homogenization had possibly less chances to occur. A similar observation has been made recently for an extremely polymorphic member of the myoglobin minisatellite family (Wong *et al.*, 1986) and for a tandemly repeated sequence of the human involucrin gene (Eckert and Green, 1986).

DXYS15 HVR

Sequence analysis of insert 113I, harbouring the HVR of locus DXYS15 shows again a tandemly repeated monomer rich in AT dinucleotides which represent 85% of this minisatellite. The consensus monomer can be subdivided into three or four elementary subunits among which the alternation of two or three GATATATA octanucleotides (Figure 5) accounts for the 21/29-bp alternation of the monomers. The sequence divergence is remarkably low in the major part of the array and has only been observed in two positions. However, in the vicinity of one edge there is an important drift versus the consensus sequence (Figure 5). It is again possible that homogenization processes occur less frequently near edges. The tandemly repeated array is flanked at each side by an inverted sequence of 7 bp. Some of these features are also found in the human alpha 1 type II collagen (COL2A1) minisatellite, another highly AT-rich minisatellite (Stoker *et al.*, 1985). Furthermore both sequences share extensive homologies as shown in Figure 5. To test whether these homologies could result in cross-hybridization, a Southern blot with several restriction digests of plasmid pSCG7 containing the collagen HVR was probed with 113I. However, no cross-hybridization could be detected in conditions of very low stringency (not shown).

A family of AT-rich HVRs

To test if the DXYS17 and DXYS15 minisatellites belonged to larger families like the myoglobin family (Jeffreys *et al.*, 1985a) or the insulin/alpha globin group (Jarman *et al.*, 1986) hybridizations and washings at low stringency were performed on *Taq*I digests of a sampling of unrelated human DNAs. No additional hybridization signal was observed with probe 602 when hybridized at very low stringency, which suggests that the DXYS17 HVR has rather unique features in the human genome.

On the other hand probe 113I detects 10–20 additional bands under conditions of reduced stringency (Figure 6). Some of these bands are apparently polymorphic and a minimum of six independent RFLPs have been detected by screening unrelated individuals and some families. These family studies also show that these RFLPs are inherited in a Mendelian fashion. Between three and five alleles have been observed so far for the different RFLPs. As for DXYS15, the size variations between two extreme alleles from a single series do not exceed a few hundred basepairs. The polymorphisms also appear in different digestions like *Eco*RI, *Hin*dIII but are best observed in smaller fragments like those generated by *Hin*fI (Figure 3) or other frequent cutter restriction enzymes. In some instances it was possible to define haplotypes by the co-segregation of a given allele from two different digests in a single family. Such haplotypes consist of the co-segregation of either the larger or the smaller alleles but not

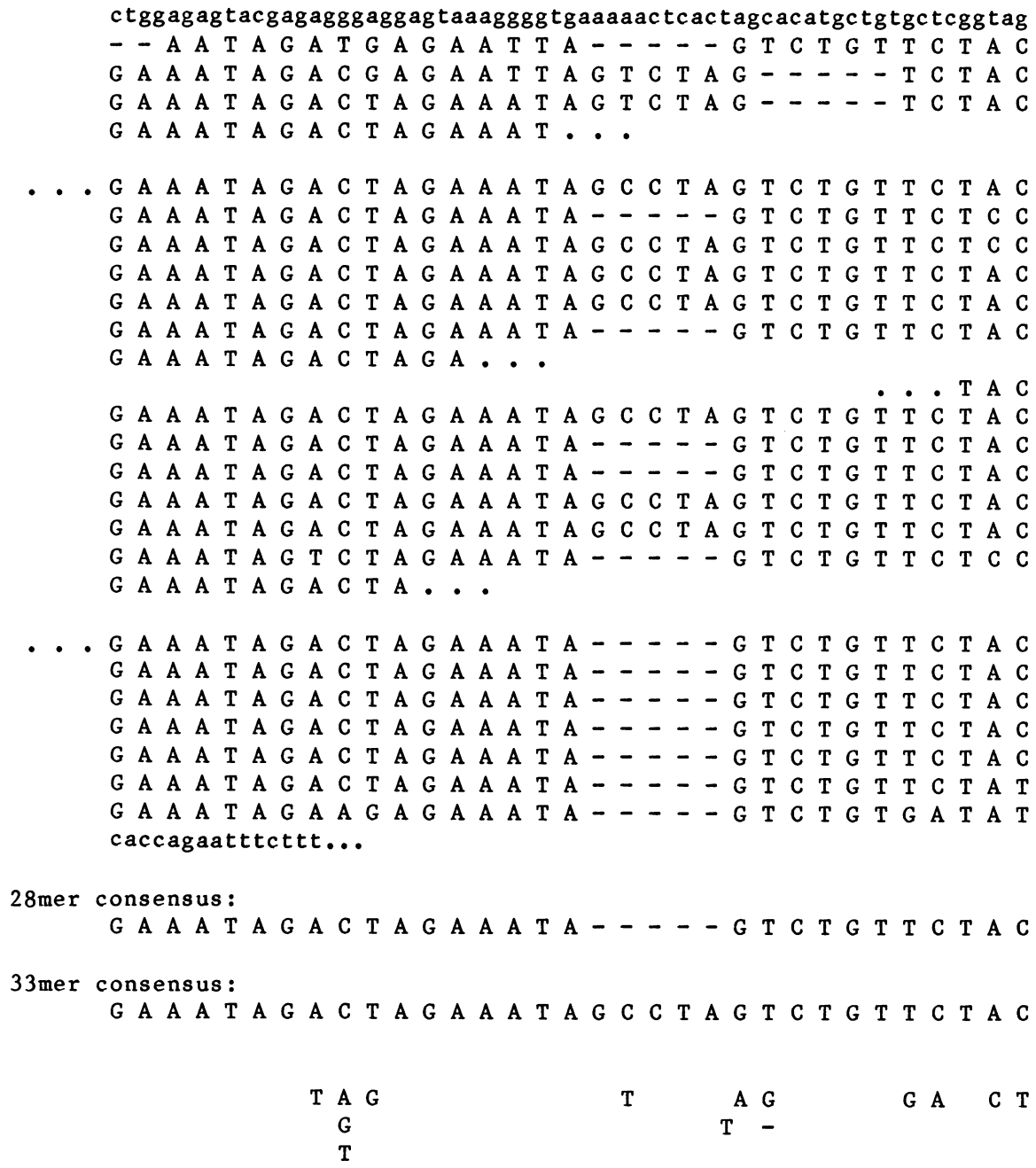

Fig. 4. DNA sequence of hypervariable region of locus DXYS17. The sequence of 5′ and 3′ flanking regions is shown together with the beginning, the end and the repeat sequence of two large random regions of the minisatellite. The consensus sequence and differences from this consensus are given.

of a mix of one small with one large allele, again indicating that these polymorphisms result from insertion/deletion events.

Hybridization to human cell lines with multiple X or Y chromosomes did not show any dosage effect for these additional polymorphic bands. Similarly, informative segregation analysis of different RFLPs also showed absence of sex linkage and absence of linkage to any of the pseudoautosomal loci. In addition there was an apparent independent segregation of the different RFLPs. These observations indicate that the DXYS15 HVR belongs to a family of sequences scattered in several autosomes.

Discussion

The two minisatellites described in this paper are precisely located within HVRs and represent a high percentage of these latter. There is very little doubt that the observed polymorphisms are generated by variations in the number of repeats of these two minisatellites. This is in agreement with many earlier findings associating genetic instability with the presence of simple repeated sequences (Brutlag *et al.*, 1977; Wyman *et al.*, 1985; Jarman *et al.*, 1986; Wong *et al.*, 1986). Copy number variation of tandemly repeated simple sequences remains to date the essential mechanism proposed accounting for hypervariability in mammalian genomes. When cloned in wild-type $RecA^+$ strains the present HVRs often underwent rearrangement showing their instability.

The sequence complexity of probe 113I is rather low and detection of single copy cross-hybridizing sequences giving hybridization signals almost as intense as the cognate repeated minisatellite

```
ttatcatc G A T A T A T A T T A C A G A T A T A T A                    21
         G A T A T A T A T T A C A G A T A T A T A G A T A T A T A    29
         G A T A T A T A T T A C A G A T A T A T A G A C A T A T A    29
         G A T A T A T A T T A C A G A T A T A T A G A T A T A T A    29
         G A T A T A T A T T A C A G A T A T A T A                    21
         G A T A T A T A T T A C A G A T A T A T A G A T A T A T A    29
         G A T A T A T A T T A C A G A T A T A T A                    21
         G A T A T A T A T T A C A G A T A T A T A G A T A T A T A    29
         G A T A T A T A T T A T A G A T A T A T A                    21
         G A T A T A T A T T A C A G A T A T A T A                    21
         G A T A T A T A T T A T T G A T A T A T A . . . .

. . . . G A T A T A T A T T A T - G A T A T A T A
                   T T A T - G A T A T A T -
         G A T A T A T A T T A T - G A T A T A T -                    19
         G A T A T A T A T T A T - G A T A T A T - G A T A T A T A
         - - - - T A T A T T A T - G A T A T A T - G A T A T A T A tagataa

consensus:
         G A T A T A T A T T A C A G A T A T A T A(G A T A T A T A)

              +                 G           +                 G
DXYS15 HVR    TATATATCTATATATCTATAATATATATCTATATATCTATATATCTATAATATATATC
                        °°°°  °°°°°°°°°° °°°°°°°°   °°°°
COL2A1 HVR              CAATATAGATAATATATACCTATATATTATTATA
```

Fig. 5. DNA sequence of hypervariable region of locus DXYS15. The inverted flanking repeats are shown together with the beginning and the end of the minisatellite. Two repeated units are shown on the complementary strand aligned with a 34-bp minisatellite monomer from the 3′ region of human type II alpha 1 collagen gene (Stoker *et al.*, 1985).

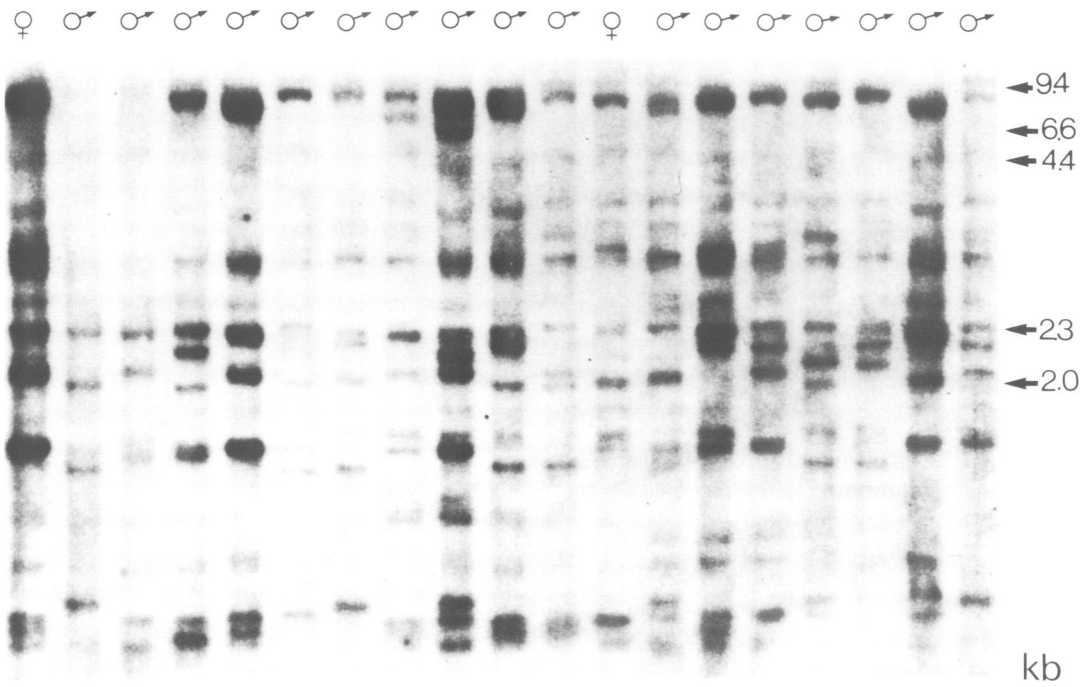

Fig. 6. DNA fingerprints of unrelated human DNA samples digested with *TaqI* probed with 113I at low stringency.

appears hardly possible. This is good evidence that most if not all cross-hybridizing bands actually also represent tandemly repeated sequences. Furthermore the additional polymorphisms are multiallelic, of insertion/deletion type and many of these polymorphic fragments are small with either little or no place left for an unrelated HVR. Taken together these arguments strongly suggest that the cross-hybridizing RFLPs also arose through copy number variations of minisatellites related to the DXYS15 minisatellite.

The RFLPs of the DXYS15 HVR family vary over a rather small range of at most 1 kb in a single instance, but more often around a few hundred basepairs or even less. This is a striking difference with the extreme hypervariability observed in a number of members of the myoglobin family. The elevated GC ratios of the common core (Jeffreys *et al.*, 1985a) and of other minisatellites (Bell *et al.*, 1982; Proudfoot *et al.*, 1982; Capon *et al.*, 1983; Goodbourn *et al.*, 1983; Jarman *et al.*, 1986) contrasts also with the very high AT ratio of the DXYS15

minisatellite. Type II collagen (Stoker *et al.*, 1985) and apolipoprotein B (Knott *et al.*, 1986) HVRs, the other AT-rich HVRs known to date, are also characterized by a small size variation of the different alleles. This suggests that an increased variability could result from a different behaviour of GC-rich versus AT-rich minisatellites. Either this difference in behaviour may indeed be related to a role in recombination for the more variable GC-rich minisatellites, or, more trivially, the higher stability of GC-rich heteroduplexes may facilitate the generation of new variants through any of the proposed mechanisms (unequal crossover, replication slippage). It should also be noted that the apparently high amount of sequences related to the myoglobin minisatellite family represents a large potential reservoir for the generation of new HVRs.

The occurrence of sequences similar to the DXYS15 HVR in different parts of the genome cannot be explained with the present amount of data. We do not know if those sequences arose independently or if they result from the spread of a prototype sequence. No obvious structural elements suggesting a mobility within the genome have been evidenced yet. Further investigation is also required to define conserved structural elements (common core) among the various members of the DXYS15 minisatellite. Although significantly homologous to the type II collagen minisatellite, some features of the DXYS15 HVR are not present in this former. Whereas the type II collagen minisatellite may be derived from an inversion of a simple sequence, organization of DXYS15 is rather reminiscent of duplications of a shorter sequence leading to a modular type of repeat. A modular organization with several domains has also been found in DXYS17-HVR and IZ-HVR in myoglobin-related HVRs (Goodbourn *et al.*, 1983). Recombination in the human pseudoautosomal region occurs 10–20 times more frequently in male than in female meiosis. In line with the function the myoglobin-like minisatellite may play in recombination, it is tempting to relate the high recombinational activity of the pseudoautosomal region to the important variability of many known pseudoautosomal DNA loci. The minisatellite sequences presented above are, however, totally unrelated to the myoglobin core sequence and do not share any feature in common with the bacterial Chi sequence (Smith *et al.*, 1981) nor with other reported minisatellites. The hypothesis of characteristics specific to the pseudoautosomal region can also be rejected in view of the numerous autosomal RFLPs related to 113I. It has recently been shown that some limited autosomal regions recombine more frequently in male meiosis (White *et al.*, 1985). If minisatellite structures are indeed more recombinogenic than others it would be of interest to examine if the mapping of the DXYS15-like HVRs corresponds to regions recombining more frequently at male meiosis.

An additional interest of the DXYS15 HVR resides in its potential use in human genetics. Six autosomal allelic series have been observed with this probe and represent DNA loci suitable for mapping. Although these loci have an apparently limited number of alleles, they appear easy to characterize and should provide a complementary tool to human DNA fingerprinting proposed by Jeffreys *et al.* (1985b).

Materials and methods

DNA blots and hybridization conditions

Human genomic DNAs were digested, fractionated, blotted and hybridized as previously described (Simmler *et al.*, 1985; Rouyer *et al.*, 1986a). Hybridizations at low stringency were in 35% formamide, 5 × SSC at 42°C followed by washings in 5 × SSC and 0.1% SDS at 50°C.

Cloning of hypervariable fragments

Clone 60, containing the DXYS17 hypervariable region, was isolated in a cDNA library in lambda NM1149. cDNA was synthesized from a polyA$^+$ mRNA template extracted from 3E7 hybrid cells. Double-stranded cDNA was cloned according to Gubler and Hoffman (1983). Insert DNA from the recombinant phage was subcloned into the *Eco*RI pBR322, sonicated and cloned after dG–dC tailing and recombinants were screened for the absence of human repetitive DNA as described previously (Bishop *et al.*, 1983). Insert of subclone 601 free of Alu sequences is a 0.5-kb *Pst*I fragment, probe 602 is a 1.0-kb fragment obtained after digestions with *Rsa*I and *Eco*RV endonucleases of clone 60.

A human lymphoblastoid cell line DNA, partially digested with *Sau*3A was cloned in lambda 47.1. The resulting library (1.5×10^6 clones) was screened with probe 601 and yielded five positive plaques: λ1–λ5. A *Bam*HI fragment of 2.3 kb, hybridizing to probe 601 and present in these lambda clones was further subcloned in a 'Bluescribe BS$^+$' vector (Stratagene).

Probes 113D and 113I originate both from cosmid 113 described previously (Simmler *et al.*, 1985). The 1.3-kb 113I *Taq*I–*Pvu*II fragment was subcloned in the *Cla*I site of pBR327. Probes were labelled either by nick-translation or by random priming labelling (Feinberg and Vogelstein, 1983) using the 'multiprime DNA labelling' system (Amersham).

DNA sequence analysis

DNA sequences were determined using the dideoxynucleotide chain termination method (Sanger *et al.*, 1977). Random subclones from clone 60 were obtained in the M13mp8 vector propagated in *Escherichia coli* JM101. Ten subclones hybridizing positively to probe 601 were used to determine the DXYS17 HVR. 113I fragment was inserted in M13tg130 vector and directly sequenced in both orientations, containing the 5′ and 3′ flanking regions and part of the tandem repeats of locus DXYS15.

Acknowledgements

We thank the Centre d'Etude du Polymorphisme Humain (CEPH) for the family DNA samples, Dr Ellen Solomon for the gift of plasmid pSCG7; Pierre Tiollais for support and encouragement and Colin Bishop for reading this manuscript. Gilles Vergnaud is a member of Délégation Générale de l'Armement.

References

Bell,G.I., Selby,M.J. and Rutter,W.J. (1982) *Nature*, **295**, 31–35.
Bishop,C.E., Guellaen,G., Geldwerth,D., Voss,R., Fellous,M. and Weissenbach,J. (1983) *Nature*, **303**, 831–832.
Brutlag,D., Fry,K., Nelson,T. and Hung,P. (1977) *Cell*, **10**, 509–519.
Burgoyne,P.S. (1982) *Hum. Genet.*, **61**, 85–90.
Capon,D.J., Chen,E.Y., Levinson,A.D., Seeburg,P.H. and Goeddel,D.V. (1983) *Nature*, **302**, 33–37.
Cooke,H.J., Brown,W.A.R. and Rappold,G. (1985) *Nature*, **317**, 688–692.
Eckert,R.L. and Green,H. (1986) *Cell*, **46**, 583–589.
Feinberg,A.P. and Vogelstein,B. (1983) *Anal. Biochem.*, **132**, 6–13.
Goodbourn,S.E.Y., Higgs,D.R., Clegg,J.B. and Weatherall,D.J. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 5022–5026.
Goodfellow,P.N., Davies,K.E. and Ropers,H.H. (1985) *Cytogenet. Cell Genet.*, **40**, 296–307.
Goodfellow,P.J., Darling,S.M., Thomas,S.N. and Goodfellow,P.N. (1986) *Science*, **234**, 740–743.
Gubler,U. and Hoffman,B.J. (1983) *Gene*, **25**, 263–269.
Higgs,D.R., Goodbourn,S.E.Y., Wainscoat,J.S., Clegg,J.B. and Weatherall,D.J. (1981) *Nucleic Acids Res.*, **9**, 4213–4224.
Jarman,A.P., Nicholls,R.D., Weatherall,D.J., Clegg,J.B. and Higgs,D.R. (1986) *EMBO J.*, **5**, 1857–1863.
Jeffreys,A.J., Wilson,V. and Thein,S.L. (1985a) *Nature*, **314**, 67–73.
Jeffreys,A.J., Wilson,V. and Thein,S.L. (1985b) *Nature*, **316**, 76–79.
Knott,T.J., Wallis,S.C., Pease,R.J., Powell,L.M. and Scott,J. (1986) *Nucleic Acids Res.*, **14**, 9215–9216.
Kobori,J.A., Strauss,E., Minard,K. and Hood,L. (1986) *Science*, **234**, 173–179.
Marcus,M., Tantravaki,R., Dev,V.G., Miller,D.A. and Miller,O.J. (1976) *Nature*, **262**, 63–65.
Proudfoot,N.J., Gil,A. and Maniatis,T. (1982) *Cell*, **31**, 553–563.
Rouyer,F., Simmler,M.C., Johnsson,C., Vergnaud,G., Cooke,H. and Weissenbach,J. (1986a) *Nature*, **319**, 291–295.
Rouyer,F., Simmler,M.C., Vergnaud,G., Johnsson,C., Levilliers,J., Petit,C. and Weissenbach,J. (1986b) *Cold Spring Harbor Symp. Quant. Biol.*, **51**, in press.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Simmler,M.C., Rouyer,F., Vergnaud,G., Nyström-Lahti,M., Ngo,K.Y., de la Chapelle,A. and Weissenbach,J. (1985) *Nature*, **317**, 692–697.
Smith,G.R., Kunes,S.M., Schultz,D.W., Taylor,A. and Triman,K.L. (1981) *Cell*, **24**, 429–436.

Steinmetz,M., Stephan,D. and Fisher-Lindahl,K. (1986) *Cell*, **44**, 895–904.
Stoker,N.G., Cheah,K.S.E., Griffin,J.R., Pope,F.M. and Solomon,E. (1985) *Nucleic Acids Res.*, **13**, 4613–4622.
Uematsu,Y., Kiefer,H., Schulze,R., Fisher-Lindahl,K. and Steinmetz,M. (1986) *EMBO J.*, **5**, 2123–2129.
Ullrich,A., Dull,T.J., Gray,A., Philips, J.A. and Peter,S. (1982) *Nucleic Acids Res.*, **7**, 2225–2240.
White,R., Leppert,M., Bishop,D.T., Barker,D., Berkowitz,J., Brown,C., Callahan,P., Holm,T. and Jerominski,L. (1985) *Nature*, **313**, 101–105.
Wong,Z., Wilson,V., Jeffreys,A.J. and Thein,S.L. (1986) *Nucleic Acids Res.*, **14**, 4605–4616.
Wyman,A.R. and White,R. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 6754–6758.
Wyman,A.R., Wolfe,L.B. and Botstein,D. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 2880–2884.

Received on January 21, 1987