# SHORT COMMUNICATION

# Human Sex-Chromosome-Specific Repeats within a Region of Pseudoautosomal/Yq Homology

ELIZABETH M. C. FISHER,* TIINA ALITALO,† SHIUH-WEN LUOH,*
ALBERT DE LA CHAPELLE,† AND DAVID C. PAGE*

*Whitehead Institute, 9 Cambridge Center, and Department of Biology, Massachusetts Institute of Technology, Cambridge,
Massachusetts 02142; and †Department of Medical Genetics, University of Helsinki, 00290 Helsinki, Finland

Several categories of human X–Y homologous DNA sequences have been recognized. We report that a locus (*DXYS77*) approximately 14 kb distal to the pseudoautosomal boundary (*PAB*) is 93% identical in nucleotide sequence to a locus (*DYS148*) on the long arm of the Y chromosome (Yq). Within this segment of pseudoautosomal/Yq homology we identified a member of a family of repeats that are concentrated in Xp22.3 and in the euchromatic portion of the Y chromosome. The repeat sequence structure—a dimer bounded by short terminal repeats—is reminiscent of retroposons derived from RNA polymerase III transcripts. © 1990
Academic Press, Inc.

In the pseudoautosomal region of the human sex chromosomes, virtual identity of X and Y sequences is maintained by X–Y recombination during male meiosis. During a chromosomal walk of the distal short arm of the human Y chromosome, we isolated phage λOX316. The human insert of λOX316 is 14.5 kb in length and is pseudoautosomal, spanning the region from approximately 10 to 25 kb distal to *PAB* (the proximal boundary of the pseudoautosomal region as defined by Ellis *et al.*, 1989). A 918-bp *Eco*RI fragment, pDP316, was subcloned from λOX316 (Fig. 1) and used as a hybridization probe to identify related sequences within a genomic library prepared from a 49,XYYYY male. λOX320 was among the cross-hybridizing clones identified. However, the restriction maps of phages λOX316 and λOX320 could not be aligned, suggesting that they might derive from two different loci. This suggestion was borne out by restriction mapping of additional phage clones from each of the two loci.

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. M33523 and M33524.

Portions of pseudoautosomal plasmid pDP316 were subcloned to yield pDP316a and pDP316b (Fig. 1). When hybridized to *Taq*I-digested genomic DNAs, pDP316a detected X–Y common restriction fragments [in males, females, and (not shown) human–rodent cell hybrids retaining only a human X or only a human Y chromosome],
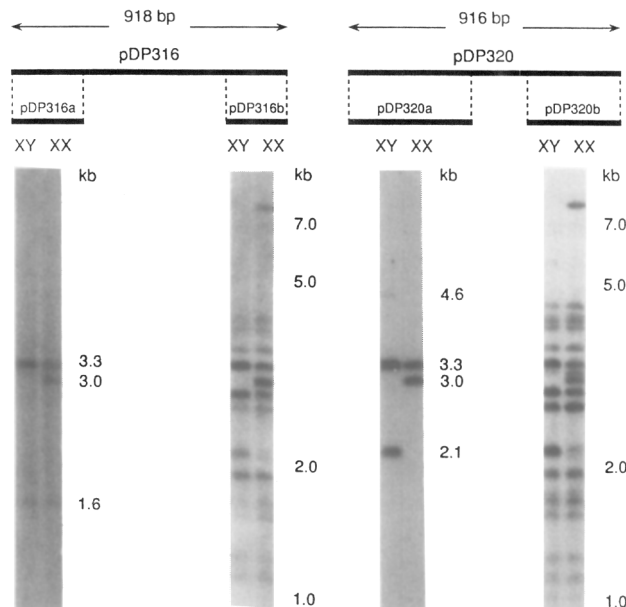


**FIG. 1.** Hybridization of pDP316 and pDP320 subclones to *Taq*I-digested human genomic DNAs. Plasmid pDP316 contains a 918-bp *Eco*RI fragment which maps approximately 14 kb distal to *PAB*, within the pseudoautosomal region. Plasmids pDP316a and pDP316b contain, respectively, 213-bp *Eco*RI/*Bam*HI and 181-bp *Acc*I/*Eco*RI fragments subcloned from pDP316. Plasmid pDP320 contains a 916-bp *Eco*RI fragment which derives from Yq. Plasmids pDP320a and pDP320b contain, respectively, 395-bp *Eco*RI/*Dra*II and 298-bp *Acc*I/*Eco*RI fragments subcloned from pDP320. Conditions were as previously described [(8); hybridization at 47°C in a solution containing 50% formamide, 5× SSC; wash at 60°C in 0.1× SSC, 0.1% SDS].
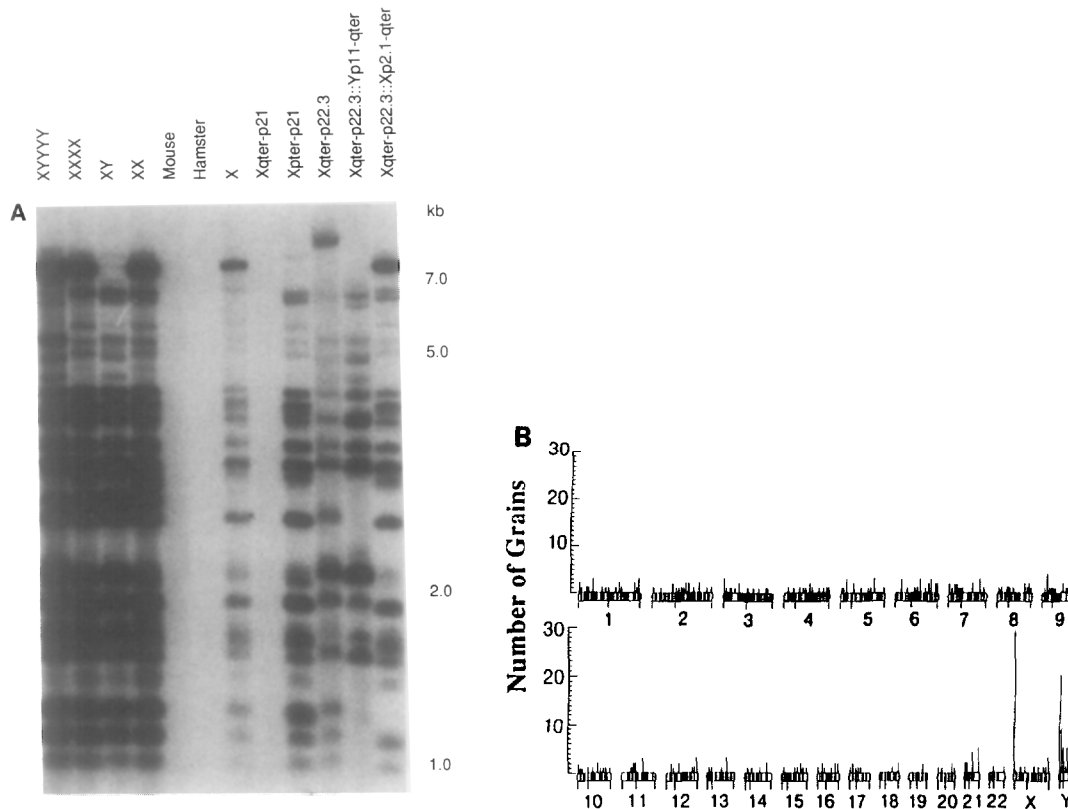
**FIG. 2.** Physical mapping of pDP320-related sequences on the human sex chromosomes. (**A**) pDP320-related sequences map to Xp21.2–pter by deletion analysis. Plasmid pDP320 was hybridized to *Taq*I-digested DNAs from (left to right) XYYYY male OXEN, XXXX female GM1416, a normal human male, a normal human female, mouse, hamster, and a series of human–rodent hybrids retaining the indicated portions of the human sex chromosomes [see Fig. 5 in Ref. (8) for details]. (**B**) *In situ* hybridization. Metaphase spreads of a normal male were incubated with $^3$H-labeled pDP320 (sp act, 4.4 × $10^6$ cpm/μg; concentration 20–40 ng/ml) as previously described (1). The slides were exposed for 21–26 days. The histogram summarizes the sites of hybridization in the 80 metaphase cells scored; grain clusters were scored as single grains.

as is characteristic of pseudoautosomal probes. While the hybridizing 1.6-kb *Taq*I fragment appears to be invariant, the 3.0- and 3.3-kb fragments are allelic. The characteristics of this *Taq*I RFLP are consistent with a location in the most proximal portion of the pseudoautosomal region: 3.3-kb homozygotes, 3.0-kb homozygotes, and heterozygotes have been observed among both males and females, but in 20 male meioses examined, no exception to sex-linked inheritance was seen.

Pseudoautosomal plasmid pDP316 cross-hybridized strongly to a 916-bp *Eco*RI fragment from phage λOX320 (from the second locus). This 916-bp *Eco*RI fragment and portions thereof were subcloned to yield plasmids pDP320, pDP320a, and pDP320b (Fig. 1). When hybridized to *Taq*I-digested genomic DNAs, pDP320a detected the allelic (3.0- or 3.3-kb) fragments recognized by pDP316a. In addition, pDP320a hybridized to 2.1-kb and (barely detectable) 4.6-kb fragments which are male-specific rather than pseudoautosomal. (These male-specific fragments were not detected by pDP316a, which is 182 bp shorter than pDP320a, but were detected by other portions of pDP316, e.g., pDP316b; Fig. 1.) The 2.1- and 4.6-kb male-specific fragments detected by pDP320a were mapped to in-

tervals 5–6, on the long arm of the Y chromosome, by hybridization to DNAs from individuals with previously characterized deletions (Vergnaud *et al.*, 1986). We conclude that a pseudoautosomal locus defined by pDP316 and its derivatives is homologous to a Yq locus defined by pDP320 and its derivatives. These pseudoautosomal and Yq loci have been designated, respectively, *DXYS77* and *DYS148*.

In contrast to pDP316a and pDP320a, which yield rather simple patterns of hybridization, both pDP316b and pDP320b detected 15 to 20 bands of varying intensity when hybridized to *Taq*I-digested human genomic DNAs (Figs. 1 and 2A). Restriction fragment length polymorphisms were observed but, with the exception of the 2.1-kb male-specific fragment previously detected by pDP320a, no fragments appeared to be specific to males or females. Similar patterns of hybridization were observed with human genomic DNAs digested with *Eco*RI, *Hin*dIII, or *Pst*I. We conclude that pseudoautosomal clone pDP316b and Yq clone pDP320b contain members of a repetitive sequence family.

We then assessed the chromosomal distribution of these repeats. First, pDP320 was hybridized to DNAs from a panel of human–rodent cell hybrids retaining
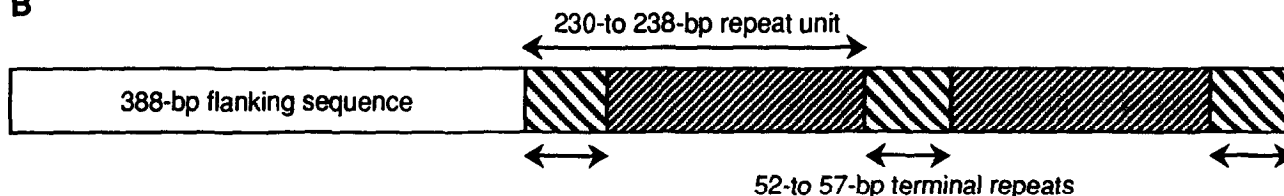
**A**

```
pDP316   GAATTCAGGCCTCACTGTCTGTCTGTAACCCAACAGACGGTGTCTGCAGAGATCGAAGTATTTTGTCGTCGAAGAGGAAGGAATGATCATTCATCACAAAAAGCAAGACA 110
pDP320   ...................A..........A........A.G........G.G...............A..A...........................T.............

         TCTTTGGTGCAAGGAAAACTCGAGGAAAATACCGCAGACCATGCAATGAGGCACTGGTTGACGGTGTGTTATAAACCCGTCTTCCCAGAGTGGCATGCACACGGATCCCTCAGGACATGG 230
         .....................A..A........................................C..T..A.....G.........................T...........C......

         GTGACACACAGACTATGCTTCAGCAGGTCTGTCTGGGCCCAAGCACACAGTGTTTCTCATCAGCTCCCAGGGGATGTCAAGGCTGCAGATCCATGGATCTCACTTTGCAGGACAGAGACTT 350
         ..................................................T.................................................................

         GGTAATGGCTTCCCAGAGTTGTTACAATGCAATCCCAAAGACTGGGCAGCTTAAACAACAACCTTGATTCTCCCACAGTCCTGGAAGCTGGAAGTCTGAGATCAAGGTGTGGGCAGGGC- 470
         ....................A.A.......   .........CC...............--.......T......A.   ...............C.A...C.....T

         GGTTCCTCCTGAGTCCTCTCTCCTGGGCTTGTAGATGCCGTCTTCTCCCTGAGTCCCCACGTGGTCATCCCTCTGTGTGCGTCTGTGTCCTCATCTCCTCTTCTTATGAGGTGTCTTAGT 590
         ..........G..........T..G..........T.............T....T...AG...TG.............T.................................

         CCATTTCAGGCTGCTGTCACAGCATACCATAGACTGGGTGGCTTATAAGCAACAGACATTGATTCTCCCACAGCCCTGGAGGCTGGACGTCTTGAGATCAGGATATGGGCAAGGCTGTTT 710
         .......................G..G.......................................T......A.......  ...-.........G........G.....C..

         CCTCCTGAGGCCTCTGTCCTGGGCTTGTAGACACCATCTTCTCCCTGTGTCCCCACGTGGTCATCCCTCTATGTGCATGTCTGTGTCCTCATCTGCTCTTCTTATGAGATGTCTTAGTCC 830
         ...................TG.TG........A..........T...............G..G.TG.......TT.................A.................

         ATTGCAGGCTGCTATCACAGAATACCATAGGCTGGGTGGCTTACAAACCACAGACTTTTATTCTCCCACAGTCCTGGAGGCTGGAATTC
         .........................T..........G...........T....................
```

**B**



230-to 238-bp repeat unit

388-bp flanking sequence

52-to 57-bp terminal repeats

FIG. 3. (A) Comparison of the nucleotide sequences (10) of the human genomic inserts of pDP316 and pDP320. In pDP320, dots indicate identity to pDP316. Dashes indicate gaps. Three direct repeats of 52 to 57 bp are boxed. RNA polymerase III promoter A-box motifs are underlined. (B) Diagram of the tandem repeats and single-copy flanking sequence as found in both pDP316 and pDP320.

various sets of human chromosomes. Only hybrids retaining human X or Y chromosomes gave appreciable signals. Second, deletion mapping using hybrids retaining partial human X chromosomes demonstrated that substantial numbers of the repeats are located in Xp21.2-pter; there is no evidence of such repeats elsewhere on the X chromosome (Fig. 2A). Third, in situ hybridization of pDP320 revealed a concentration of repeats in the euchromatic portion of the Y chromosome, principally Yp, and refined the X chromosome location to Xp22.3, the most distal band on the short arm (Fig. 2B). Of a total of 308 autoradiographic grains located on chromosomes, 30 (10%) were found on Xp22.3, while 40 (13%) were on the euchromatic portion of the Y (30 of these grains on Yp). In light of these findings, the similarity of male–female hybridization patterns (e.g., Southern blots of Figs. 1 and 2A) suggests that most of the repeats are pseudoautosomal. However, we cannot exclude the possibility that some repeats map to the strictly X-linked portion of Xp22.3. As already described, at least one repeat maps to Yq.

Cross-hybridizing repetitive sequences occur in other mammals. When pDP320 was hybridized to Southern transfers of genomic DNAs from primates, rabbits, dogs, goats, and horses, multiple fragments were readily detected in both males and females (not shown). However, no hybridization to mouse DNA was observed.

Nucleotide sequence analysis of pseudoautosomal clone pDP316 and Yq clone pDP320 clarified the nature of the repeats. The sequences of pDP316 and pDP320 differ by only three insertion/deletions, each of one or two nucleotides, and are otherwise 93% identical (Fig. 3A). No long open reading frames were found. In addition to a 388-bp flanking sequence, each clone contains a tandem duplication of 230 to 238 bp, with 52 to 57 bp terminally repeated (Fig. 3B). (The inserts of pDP316 and pDP320 do not cross-hybridize to restriction fragments adjoining them in the genome, confirming that the repeats are contained entirely within pDP316 and pDP320 and are dimers.) These dimeric repeats may be pseudogenes derived from RNA polymerase III-transcribed sequences; they have a structure grossly similar to Alu repeats. Such pseudogenes frequently contain an internal RNA polymerase III promoter (Deininger, 1989). Indeed, the 230- to 238-bp tandem repeats contain sequences similar to the A-box motif of RNA polymerase III promoters (Fig. 3A).

Sequence analysis also shed light on the relationship of the repeats to the pseudoautosomal/Yq homology. Within pDP316, the two 230- to 238-bp monomers are 90% identical; within pDP320, the two monomers are 86% identical. The pDP316 dimer is 92% identical to that of pDP320. Thus, the interclone similarity is as great as if not greater than the intraclone similarity. This suggests the following sequence of events during evolution: (1) tandem duplication of a 230- to 238-bp monomer to form a dimer, followed by (2) interarm duplication of a larger unit containing the dimer.

Not only the dimeric repeats but also the flanking sequences of pseudoautosomal clone pDP316 and Yq clone pDP320 are highly similar: the two clones are 92 and 95% identical in, respectively, their repeat and flanking regions. This suggests that the interarm duplication involved a block of DNA larger than the inserts of the two clones. Indeed, other pseudoautosomal DNA sequences in the immediate vicinity of pDP316 (as well as other Yq sequences in the immediate vicinity of pDP320) show pseudoautosomal/Yq patterns of hybridization on Southern blots of human genomic DNAs, suggesting that the duplicated segment is at least several kb in length (E. M. C. Fisher, unpublished results). The modest divergence of pDP316 and pDP320 indicates that the duplication occurred during primate evolution. Hybridization of pDP320a to Southern transfers of male and female chimpanzee and gorilla DNAs (not shown) suggests that the duplication occurred prior to the divergence of humans and apes: in apes, as in humans, pDP320a detects one fragment common to males and females and one fragment specific to males.

Five categories of human X–Y homologous sequences have previously been identified: (1) pseudoautosomal (e.g., Cooke et al., 1985; Simmler et al., 1985); (2) Xq/Yp (e.g., Page et al., 1984); (3) Xp/Yq (e.g., Koenig et al., 1985); (4) Xq/Yq (Cooke et al., 1984); and (5) Xp/Yp non-pseudoautosomal (Page et al., 1987). We have identified a sixth class of human X–Y homology—between the pseudoautosomal locus DXYS77 and the Yq locus DYS148.

Within this duplicated segment, we identified a short interspersed repetitive sequence whose relatives are, as judged by a variety of hybridization techniques, concentrated on the X and Y chromosomes. However, computer searches revealed sequences distantly related to these repeats in a number of autosomal locations, including PAD (A4 amyloid protein), BCR, and the mu–delta intron of IGH (E. M. C. Fisher, unpublished results). Thus, it may be that a narrow offshoot of a widely dispersed family of short interspersed repeats has proliferated on the sex chromosomes.

Rouyer and colleagues (1990) have recently characterized a family of interspersed repetitive sequences (STIR) specific to human subtelomeric regions, especially the pseudoautosomal regions of the X and Y chromosomes. Although the chromosomal distribution of these STIR repeats and of the repeats described in this paper overlap considerably, nucleotide sequence comparison suggests that the two families of repeats are not closely related.

## REFERENCES

1. ANDERSSON, M., PAGE, D. C., PETTAY, D., SUBRT, I., TURLEAU, C., DE GROUCHY, J., AND DE LA CHAPELLE, A. (1988). Y;autosome translocations and mosaicism in the etiology of 45,X maleness: Assignment of fertility factor to distal Yq11. Hum. Genet. **79**: 2–7.

2. COOKE, H. J., BROWN, W. R. A., AND RAPPOLD, G. A. (1985). Hypervariable telomeric sequences from the human sex chromosomes are pseudoautosomal. Nature (London) **317**: 687–692.

3. COOKE, H. J., BROWN, W. R. A., AND RAPPOLD, G. A. (1984). Closely related sequences on human X and Y chromosomes outside the pairing region. Nature (London) **311**: 259–261.

4. DEININGER, P. L. (1989). SINEs: Short interspersed repeated DNA elements in higher eucaryotes. In "Mobile DNA" (D. E. Berg and M. M. Howe, Eds.), pp. 619–636, American Society for Microbiology, Washington, DC.

5. ELLIS, N. A., GOODFELLOW, P. J., PYM, B., SMITH, M., PALMER, M., FRISCHAUF, A.-M., AND GOODFELLOW, P. N. (1989). The pseudoautosomal boundary in man is defined by an Alu repeat sequence inserted on the Y chromosome. Nature (London) **337**: 81–84.

6. KOENIG, M., MOISAN, J. P., HEILEG, R., AND MANDEL, J. L. (1985). Homologies between X and Y chromosomes detected by DNA probes: Localization and evolution. Nucleic Acids Res. **13**: 5485–5501.

7. PAGE, D. C., HARPER, M. E., LOVE, J., AND BOTSTEIN, D. (1984). Occurrence of a transposition from the X-chromosome long arm to the Y chromosome short arm during human evolution. Nature (London) **311**: 119–123.

8. PAGE, D. C., MOSHER, R., SIMPSON, E. M., FISHER, E. M. C., MARDON, G., POLLACK, J., McGILLIVRAY, B., DE LA CHAPELLE, A., AND BROWN, L. G. (1987). The sex-determining region of the human Y chromosome encodes a finger protein. Cell **51**: 1091–1104.

9. ROUYER, F., DE LA CHAPELLE, A., ANDERSON, M., AND WEISSENBACH, J. (1990). An interspersed repeated sequence specific for human subtelomeric regions. EMBO J. **9**: 505–514.

10. SANGER, F., NICKLEN, S., AND COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**: 5463–5467.

11. SIMMLER, M.-C., ROUYER, F., VERGNAUD, G., NYSTROM-LAHTI, M., NGO, K. Y., DE LA CHAPELLE, A., AND WEISSENBACH, J. (1985). Pseudoautosomal DNA sequences in the pairing region of the human sex chromosomes. Nature (London) **317**: 692–697.

12. VERGNAUD, G., PAGE, D. C., SIMMLER, M.-C., BROWN, L., ROUYER, F., NOEL, B., BOTSTEIN, D., DE LA CHAPELLE, A., AND WEISSENBACH, J. (1986). A deletion map of the human Y chromosome based on DNA hybridization. Amer. J. Hum. Genet. **38**: 109–124.