

# Single-copy sequence hybridizes to polymorphic and homologous loci on human X and Y chromosomes

(restriction endonuclease/genetic marker/sex linkage)

DAVID PAGE\*, BERENGERE DE MARTINVILLE†, DAVID BARKER‡, ARLENE WYMAN\*, RAY WHITE‡, UTA FRANCKE†, AND DAVID BOTSTEIN\*

\*Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139; †Department of Human Genetics, Yale University School of Medicine, New Haven, Connecticut 06510; and ‡Department of Cellular, Viral, and Molecular Biology and Howard Hughes Medical Institute, University of Utah Medical School, Salt Lake City, Utah 84132

Contributed by David Botstein, May 28, 1982

**ABSTRACT** Use of a 4.5-kilobase-pair (kb) segment of single-copy DNA from a human genomic library as a hybridization probe of genomic human DNAs revealed allelic *Taq* I restriction fragments 10.6, 11.8, and 14.6 kb long. Among 12 unrelated individuals, all 6 males exhibited the 14.6-kb fragment in addition to one of the other fragments. Three of the females displayed 10.6- and 11.8-kb fragments, and the other three displayed only one fragment length; none had the 14.6-kb fragment. Hybridization of this probe to *Taq* I-digested DNAs from human-rodent hybrid cell lines (which have partial complements of human chromosomes) demonstrated segregation of the 14.6-kb fragment with the human Y chromosome and segregation of the 10.6- and 11.8-kb fragments with the human X chromosome. Furthermore, hybridization of this probe to *Taq* I-digested DNAs from 48 members of a single kindred revealed Y-linked inheritance of the 14.6-kb fragment and X-linked inheritance of the 10.6- and 11.8-kb fragments. These experiments demonstrate homology between single-copy sequences on the human X and Y chromosomes.

Heterogeneity in the lengths of defined DNA restriction fragments occurs within normal human populations, and it provides the basis for a new class of markers for human genetic studies. Botstein *et al.* (1) have argued that a collection of several hundred such restriction fragment length polymorphisms (RFLPs) could facilitate the construction of a detailed linkage map of the human genome and, in particular, could be used to map inherited traits for which no related gene product or coding sequence has been identified.

Here we report the finding of an unusual RFLP arising from a site of single-copy sequence homology between the human X and Y chromosomes and characterized by the existence of X-specific and Y-specific alleles.

## MATERIALS AND METHODS

**Origin of Hybridization Probes.** Phage  $\lambda$ -rHs4813 was obtained from a human genomic library provided by Edward Fritsch and Tom Maniatis (2). It was among approximately 3% of library phage showing weak plaque hybridization (3) to <sup>32</sup>P-labeled total human DNA. (Approximately 90% of library phage hybridized strongly and 7% failed to hybridize; ref. 4 and unpublished results.) A 4.5-kilobase pair (kb) *Eco*RI restriction fragment and a 2.2-kb *Bgl* II/*Eco*RI fragment from the human insert of  $\lambda$ -rHs4813 (Fig. 1) were purified from 1% agarose gels by the NaI/glass slurry method of Vogelstein and Gillespie (5). In this paper, these *Eco*RI and *Bgl* II/*Eco*RI fragments are referred to as probe A and probe B, respectively. For conven-

ience, the 4.5-kb *Eco*RI fragment (probe A) was subcloned in plasmid vector pBR322 by using methods described previously (4). The recombinant plasmid containing the 4.5-kb human insert was designated pDP31 and was used instead of the purified 4.5-kb *Eco*RI fragment in some experiments involving probe A.

**Preparation of Human DNA.** Human DNA was prepared from peripheral leukocytes as described (4).

**Human-Rodent Hybrid Cell Lines.** Six series of human-Chinese hamster somatic cell hybrids were used. The origin, propagation, and chromosome and enzyme analyses of hybrid series X, XII, XIII, XV, and XVII have been described (6). Human-mouse hybrid line IaA9498, the gift of D. R. Cox, was the product of the fusion of A9 cells with normal human male fibroblasts (7). Somatic cell hybrid DNA was prepared as described (6).

**Digestion, Electrophoresis, Transfer, and Hybridization of DNA.** Restriction digestion, agarose gel electrophoresis, and gel transfer of DNA followed by hybridization with nick-translated probe DNAs and autoradiography were carried out as reported (4), except that some gel transfers were made to Schleicher & Schuell BA85 nitrocellulose filters (8) or to New England Nuclear Gene-Screen paper.

**Inheritance Studies.** Leukocyte samples from Utah kindred 1085 were obtained in collaboration with Mark Skolnick.

## RESULTS

**A *Taq* I Polymorphism.** Recombinant phage  $\lambda$ -rHs4813, isolated from the Maniatis human genomic library (2), contains a human DNA insert 13.5 kb in length. Two insert-derived restriction fragments (probes A and B in Fig. 1), showing no detectable hybridization to labeled total human DNA, were in turn used as probes of total human DNAs digested with various restriction enzymes. The lengths and numbers of hybridizing genomic fragments were as predicted by the restriction map of phage  $\lambda$ -rHs4813, implying that these probes consist essentially of single-copy sequences.

An unusual RFLP was detected by hybridizing probe A to human DNAs digested with *Taq* I and gel-transferred by the method of Southern (8). Among 12 unrelated individuals (6 females and 6 males), three *Taq* I fragments (10.6, 11.8, and 14.6 kb long) could be distinguished (Fig. 2). All six males exhibited the 14.6-kb fragment in addition to either the 10.6- or 11.8-kb fragment (Table 1). Three females had both the 10.6- and 11.8-kb fragments, three had only the 10.6-kb fragment, and none

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: kb, kilobase pair(s); RFLP, restriction fragment length polymorphism.

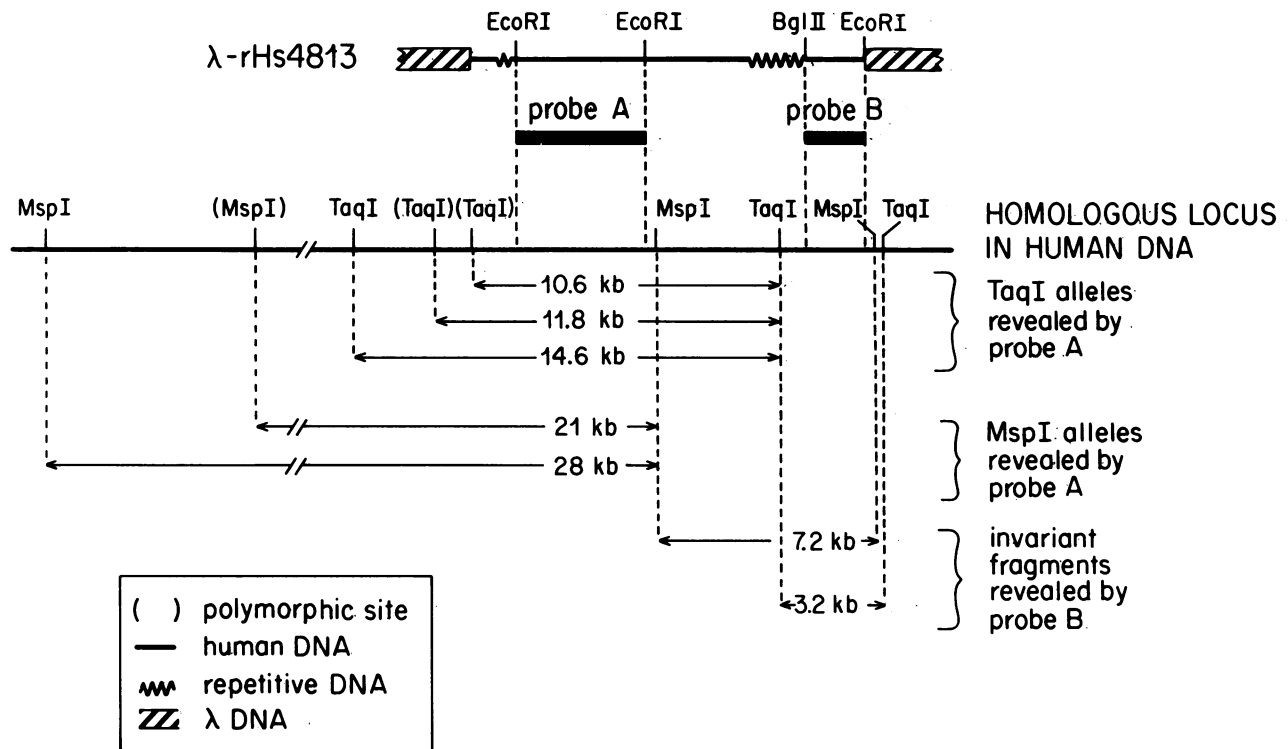


FIG. 1. Human insert of recombinant phage  $\lambda$ -rHs4813 aligned with the homologous locus in human DNA. The lengths of the allelic genomic *Taq* I and *Msp* I fragments revealed by probe A are shown, as are the invariant genomic *Msp* I and *Taq* I fragments detected by probe B. The extended *Taq* I/*Msp* I restriction map of the homologous locus in human DNA is inferred from these and other patterns of hybridization (with the restriction map of  $\lambda$ -rHs4813 providing points of reference).

had the 14.6-kb fragment. These results suggest that the 14.6-kb fragment derives from the Y chromosome.

**Segregation of the *Taq* I Alleles in Human-Rodent Somatic Cell Hybrids.** In order to ascertain the chromosomal location(s) of the human sequences homologous to probe A, DNAs prepared from a total of 25 human-rodent somatic cell hybrids were digested with either *Eco*RI or *Taq* I and then hybridized to probe A. Each hybrid line contained a different partial complement of human chromosomes. The hybrids derived from seven human donors.

Probe A was first hybridized to a panel of *Eco*RI-digested

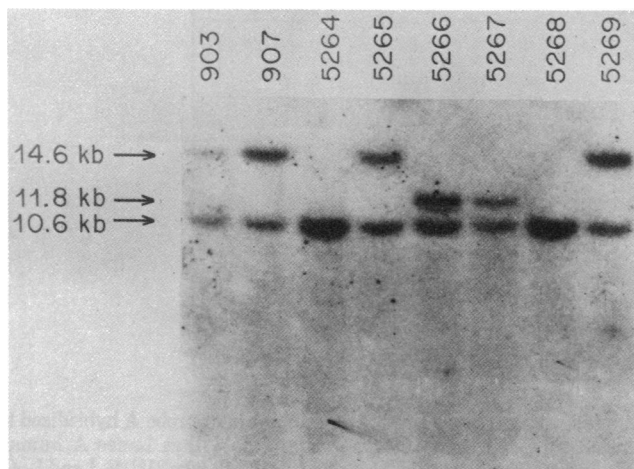


FIG. 2. Autoradiogram of hybridization of probe A to *Taq* I-digested DNAs from unrelated individuals. Numbers above lanes identify the subjects tested; numbers to the left indicate lengths of DNA fragments.

DNAs from human-Chinese hamster (V79/380-6) hybrids which has been described elsewhere (6). Each hybrid was scored for the presence or absence of a homologous 4.5-kb *Eco*RI fragment. Although Chinese hamster (380-6) DNA itself exhibited no homology to probe A under the stringent conditions used, DNA from 12 of 17 human-hamster hybrids tested contained a homologous 4.5-kb *Eco*RI fragment. When compared with the human chromosome complements of the various hybrids, the results strongly suggest the presence, on the X chromosome, of sequences homologous to probe A. However, two hybrids were positive in the absence of the human X, suggesting the presence of homologous sequences on a second chromosome. Both of these hybrids retained the human Y chromosome. The presence of homologous sequences on any autosome was excluded by the results of these and subsequent

Table 1. Lengths of *Taq* I and *Msp* I fragments homologous to probe A

Individual	Sex	Length, kb	
		<i>Taq</i> I fragments	<i>Msp</i> I fragments
902	M	10.6; 14.6	ND
903	M	10.6; 14.6	21
5260	F	10.6; 11.8	21
5261	M	11.8; 14.6	21
5262	F	10.6	21
5263	M	10.6; 14.6	21; 28
5264	F	10.6	21
5265	M	10.6; 14.6	21
5266	F	10.6; 11.8	21
5267	F	10.6; 11.8	21
5268	F	10.6	ND
5269	M	10.6; 14.6	ND

ND, not done.

experiments—i.e., negative hybridization results were invariably obtained with hybrids lacking the human X and Y chromosomes, regardless of the human autosomal complements of those hybrids.

To confirm the presence of homologous sequences on the X chromosome, probe A was hybridized to *Taq* I-digested DNAs from human-rodent hybrids that had retained the human X alone or in addition to some autosomes. Lanes E, F, and J in Fig. 3 show that the 10.6- and 11.8-kb fragments segregate with the human X chromosome (in the presence of zero to three autosomes). Furthermore, the hybrid in lane G, which had lost the human X chromosome under counterselection in 8-azaguanine, did not exhibit the 10.6-kb fragment present in the hybrid of lane F, from which it was derived. Similarly, the hybrid in lane H grown in the presence of 8-azaguanine, exhibited the 14.6-kb fragment but not the 11.8-kb fragment present in the hybrid of lane I; both hybrids retained the human Y chromosome.

Finally, to examine further the apparent segregation of *Taq* I fragments with the X and Y chromosomes, probe A was hybridized to *Taq* I-digested DNAs from four additional human-Chinese hamster hybrids, all derived from the same human male donor. These hybrids contained (in addition to several autosomes) either the X or the Y chromosome, or both, or neither. The hybridization patterns are shown in lanes L through O of Fig. 3. The 10.6-kb fragment segregated with the X chromosome and the 14.6-kb, with the Y chromosome. The relative intensities of the 14.6- and 10.6-kb bands in lane N, when compared with those of lane K, are consistent with the frequencies of the X and Y chromosomes in that hybrid line (1.0 Y chromosomes per cell; 0.6  $X^T$  chromosome per cell).

Regional mapping of the homologous sequences on the X excludes the distal part of the short arm. Hybrids of series XIII and XII were derived from a mother and her son: both had a  $t(X;14)(p22;q21)$  translocation. The pattern of hybridization of *Taq* I-digested DNAs (i.e., the 10.6-kb fragment) was identical in hybrids containing either the morphologically normal X chro-

mosome (Fig. 3, lane E) or the translocation product  $X^T$  without the reciprocal part of the translocation,  $14^T$  (lane N; also hybrids XIII-7A and XIII-1B, data not shown). The presence or absence of each translocation product was confirmed by electrophoretic assays for the expression of human glucose-6-phosphate dehydrogenase (on  $X^T$ ) and nucleoside phosphorylase (on  $14^T$ ) in hybrid strains. These results assign the sequences homologous to probe A to region Xp22.2→qter.

**Inheritance of the *Taq* I Fragments.** In order to examine the heritability of the *Taq* I fragment lengths, DNAs from 48 members of Utah kindred 1085 (see Fig. 4) were digested with *Taq* I and tested with probe A. Among these related individuals (24 females and 24 males), *Taq* I fragments 10.6, 11.8, and 14.6 kb long were again observed. Each female had the 10.6- or the 11.8-kb fragment or both, and none exhibited the 14.6-kb fragment. Each of the 24 males showed two fragments, including the 14.6-kb fragment in every case. The fragment lengths exhibited by each individual are consistent with their inheritance as codominant Mendelian alleles. Furthermore, without exception, the inheritance of the 14.6-kb allele is consistent with Y-linkage and that of the 10.6- and 11.8-kb alleles, with X-linkage.

The apparent sex linkage of the *Taq* I alleles was statistically assessed by calculating a "lod score," which expresses the relative probabilities of observing this set of data given the following two alternative hypotheses: (i) the *Taq* I alleles are strictly sex linked, or (ii) the *Taq* I alleles segregate randomly with respect to sex (i.e., according to an autosomal mode). The informative family units within kindred 1085 (Fig. 4) are those in which the *Taq* I alleles have been identified in the father and two or more of his children. For example, both sons of male 5386 received their father's 14.6-kb allele, whereas all three daughters inherited their father's 11.8-kb allele. The probability of this outcome is  $(1/2)^4$  by autosomal inheritance, but, given complete sex linkage, the probability is 1; the ratio of probabilities for the two hypotheses (sex linked to autosomal) is thus 16:1. Extending this analysis to include all informative family

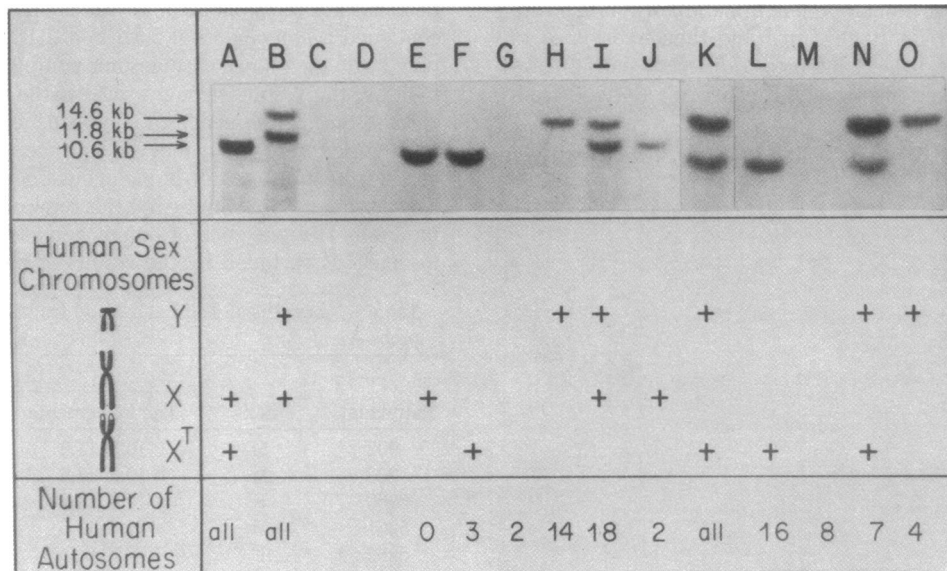


FIG. 3. Segregation of *Taq* I fragments with human sex chromosomes in human-rodent hybrid lines. Autoradiogram of probe A hybridized to *Taq* I-digested DNAs from human, rodent, and hybrid cell lines is shown with the human chromosomal complements of those lines. Lanes: A, human fibroblast MN-4 [46,X,t(X;14)(p22;q21)]; B, human fibroblast TH-5 (46,XY); K, human fibroblast MN-3 [46,XY,t(X;14)(p22;q21)]; E-I and L-O, hybrids derived from fusions of Chinese hamster cell line V79/380-6 (lane C) with MN-4 (E, hybrid XIII-2A; F, hybrid XIII-3A HAT; G, hybrid XIII-3A aza) or TH-5 (H, hybrid XV-16B-F4 aza; I, hybrid XV-18B-7a) or MN-3 (L, hybrid XII-2D HAT; M, hybrid XII-2D aza; N, hybrid XII-4A HAT; O, hybrid XII-4A aza); D, mouse cell line 3T3; J, mouse-human hybrid 1 $\alpha$ A9498. The  $X^T$  chromosome lacks region Xp22.2→pter. In hybrid XII-4A HAT (lane N), the absence of chromosome  $14^T$  (containing region Xp22.2→pter) was confirmed by the absence of human purine nucleoside phosphorylase activity in that hybrid (7).

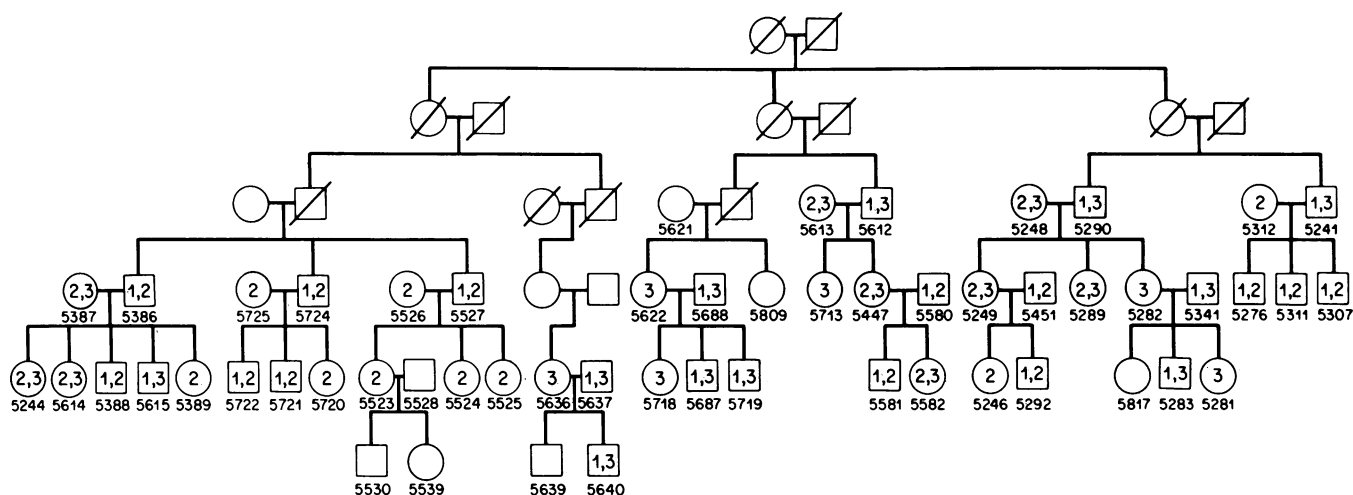


FIG. 4. The inheritance, in Utah kindred 1085, of *Taq* I fragments detected by probe A. The four-digit numbers identify the individuals. ○, Female; □, male; empty symbols, not tested; ∅, ∅, deceased. *Taq* I fragments are identified as follows: 1, 14.6 kb; 2, 11.8 kb; 3, 10.6 kb.

units in kindred 1085, this ratio increases to 260,000:1, corresponding to a lod score (the logarithm of the ratio) of 5.42 in favor of complete sex linkage.

**Basis of the *Taq* I Polymorphism.** Such a RFLP could result from either the creation or elimination of restriction sites by base-pair substitution or the displacement of sites by rearrangements such as insertions or deletions. RFLPs due to base-pair substitutions within *Taq* I sites (or small rearrangements) should not be detected in human DNAs digested with other enzymes; RFLP due to insertion/deletion might. Table 1 compares the lengths of *Taq* I and *Msp* I fragments homologous to probe A in the DNAs of nine unrelated individuals (see also Fig. 1). Except for male 5263, all samples displayed a single 21-kb *Msp* I fragment. Hybridization of probe B to these same *Taq* I- and *Msp* I-digested human DNAs revealed a single *Taq* I fragment (3.2 kb) and a single *Msp* I fragment (7.2 kb) in all individuals (see Fig. 1). From these and other hybridization data (not shown) one can infer certain restriction sites in human genomic sequences neighboring the human insert of recombinant phage  $\lambda$ -rHs4813 (Fig. 1). These data suggest that the *Taq* I RFLP is the result of base-pair substitutions within *Taq* I sites (or perhaps small rearrangements) occurring within a stretch of X-Y sequence homology at least 28 kb long.

## DISCUSSION

We have isolated from a human genomic library (2) a single-copy 4.5-kb *Eco*RI fragment (probe A) that defines a locus exhibiting length polymorphism of the *Taq* I fragment. The 14.6-kb allele is apparently male-specific; among 30 unrelated persons and five human fibroblast lines, the 14.6-kb allele is displayed by all 19 males and by none of the 16 females (not all data shown). The 10.6- and 11.8-kb alleles, however, are found in both males and females. Furthermore, in human-rodent somatic cell hybrids bearing partial complements of human chromosomes, the 14.6-kb allele segregates with the human Y chromosome and both the 10.6- and 11.8-kb alleles segregate with human X chromosomes. Finally, when their inheritance is traced through 48 members of one kindred, the 14.6-kb allele is found to be Y-linked and the 10.6- and 11.8-kb alleles X-linked; the odds in favor of complete sex linkage (as opposed to autosomal inheritance) or the *Taq* I alleles are  $2.6 \times 10^5$  to 1. We conclude that a Y-specific 14.6-kb *Taq* I fragment and allelic X-specific 10.6- and 11.8-kb *Taq* I fragments are derived from homologous sites on the X and Y chromosomes. Moreover, hybridizations of

probes A and B to human DNAs digested with *Msp* I or *Taq* I reveal certain restriction fragments that are of invariant length among most males and females, thereby extending the detected region of sequence homology between the X and Y chromosomes to 28 kb at this locus.

Further studies (results not shown) demonstrated that several restriction enzymes in addition to *Taq* I detect sequence polymorphisms at this locus. Like the *Taq* I polymorphism, a polymorphism revealed by *Mbo* I is characterized by a male-specific allele. One of two alleles detected by *Msp* I and one of two alleles detected by *Bgl* II have been observed only in a subset of males [e.g., the 28-kb *Msp* I fragment seen in male 5263 (Table 1)]. The male-limited *Bgl* II allele exhibits Y-linked inheritance.

We do not know of any reports of homology between single-copy sequences on the mammalian X and Y chromosomes. Ohno (9) has argued that the mammalian X and Y chromosomes evolved from a pair of homologues. One could speculate that we have detected an evolutionary remnant of once-extensive homology. Pairing of terminal portions of the X and Y chromosomes at meiosis in human spermatocytes has been interpreted as evidence of specific "homology" there (10-14). It is not known, however, whether DNA sequence homology itself is responsible for the accurate pairing of "homologous" chromosomes (15). Although the precise positions, on the X and Y chromosomes, of the sequences homologous to probe A have not yet been determined, we have excluded region Xp22.2  $\rightarrow$ pter as a possible site.

Darlington *et al.* (16) were the first to suggest that, should there exist such a region of homology between the human X and Y chromosomes, it might be a site of crossing-over, as in certain fish (17, 18), mosquito (19), and brine shrimp (20). Our Utah kindred 1085 provides no example of such recombination between corresponding loci on the X and Y chromosomes. Moreover, as mentioned, various restriction enzymes detect apparently male-specific and, in other cases, male-limited restriction fragment length variants at this locus. These findings seem inconsistent with recombination between the X and Y chromosomes at this locus within recent evolutionary time.

In considering the evolutionary and functional significance of homologous single-copy sequences on the X and Y chromosomes, it is relevant to note that, at high stringency, the single-copy sequences in the human insert of  $\lambda$ -rHs4813 bear no detectable homology to hamster or mouse DNA (not all data shown). We do not know if the human insert of  $\lambda$ -rHs4813 con-

tains or neighbors (in the human genome) any structural gene.

The X-specific *Taq* I alleles of lengths 10.6 and 11.8 kb may be of use as genetic markers for human X-linked traits; the Y-limited RFLPs may prove of value in population studies.

We thank Mark Leffert and Jon Berkowitz for preparing DNAs from human samples obtained in collaboration with Mark Skolnick at the University of Utah. This work was supported by National Institutes of Health Grants GM-27313 and GM-26105 and the Howard Hughes Medical Institute. D. Barker and A.W. are Postdoctoral Fellows of the National Institutes of Health. D.P. received research stipends from the Harvard Medical School and the Whitaker College at the Massachusetts Institute of Technology.

1. Botstein, D., White, R. L., Skolnick, M. & Davis, R. W. (1980) *Am. J. Hum. Genet.* **32**, 314-331.
2. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) *Cell* **15**, 1157-1174.
3. Benton, W. & Davis, R. (1977) *Science* **196**, 180-182.
4. Wyman, A. R. & White, R. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 6754-6758.
5. Vogelstein, B. & Gillespie, D. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 615-619.
6. deMartinville, B., Wyman, A. R., White, R. & Francke, U. (1982) *Am. J. Hum. Genet.* **34**, 216-226.
7. Cox, D. R., Francke, U. & Epstein, C. J. (1981) *Am. J. Hum. Genet.* **33**, 495-512.
8. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503-517.
9. Ohno, S. (1967) *Sex Chromosomes and Sex-linked Genes* (Springer-Verlag, New York), pp. 1-23.
10. Koller, P. C. (1937) *Proc. R. Soc. Edinburgh* **57**, 194-214.
11. Ford, C. E. & Hamerton, J. L. (1956) *Nature (London)* **178**, 1020-1023.
12. Solari, A. J. & Tres, L. L. (1970) *J. Cell Biol.* **45**, 43-53.
13. Pearson, P. L. & Bobrow, M. (1970) *Nature (London)* **226**, 959-961.
14. Moses, M. J., Counce, S. J. & Paulson, D. F. (1974) *Science* **187**, 363-365.
15. Lewin, B. (1980) *Gene Expression 2: Eukaryotic Chromosomes* (Wiley, New York), pp. 104-109.
16. Darlington, C. D., Haldane, J. B. S. & Koller, P. C. (1934) *Nature (London)* **133**, 417.
17. Aida, T. (1921) *Genetics* **6**, 554-573.
18. Winge, O. (1922) *J. Genet.* **12**, 145-162.
19. Gilchrist, B. M. & Haldane, J. B. S. (1947) *Hereditas* **33**, 175-190.
20. Bowen, S. T. (1965) *Genetics* **52**, 695-710.