Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 2824–2828, May 1985
Genetics

The telomeric region of the human X chromosome long arm: Presence of a highly polymorphic DNA marker and analysis of recombination frequency

(restriction fragment length polymorphism/genetic linkage/gene cluster)

ISABELLE OBERLÉ*, DENNIS DRAYNA†, GIOVANNA CAMERINO*‡, RAY WHITE†, AND JEAN-LOUIS MANDEL*

*Laboratoire de Génétique Moléculaire des Eucaryotes du Centre National de la Recherche Scientifique, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'Institut National de la Santé et de la Recherche Médicale, Faculté de Médecine, 11, rue Humann 67085 Strasbourg Cédex, France; and †Howard Hughes Medical Institute and Department of Cellular, Viral and Molecular Biology, University of Utah Medical Center, Salt Lake City, UT 84132

Communicated by L. L. Cavalli-Sforza, December 3, 1984

ABSTRACT **A DNA fragment (named *St14*) derived from the human X chromosome reveals a small family of related sequences that have been mapped to the Xq26–Xq28 region by using a panel of rodent–human somatic cell hybrids. The probe detects in human DNA digested by *Taq* I a polymorphic system defined by a series of at least eight allelic fragments with a calculated heterozygosity in females of 80%. With *Msp* I, we found three additional restriction fragment length polymorphisms, each of them being defined by two alleles. These polymorphisms are also common in Caucasian populations. The genetic locus defined by probe *St14* has been localized more precisely to the distal end of the X chromosome (in band q28) by linkage analysis to other polymorphic DNA markers. The results obtained suggest that the frequency of recombination is distributed very unevenly in the q27–qter region of the X chromosome, with a cluster of seven tightly linked loci in q28 showing about 30% recombination with the gene for coagulation factor IX located in the neighboring q27 band. Probe *St14* reveals one of the most polymorphic loci known to date in the human genome, and 17 different genotypes have already been observed. It constitutes the best marker on the X chromosome and should be of great use for the genetic study of three important diseases: hemophilia A, mental retardation with a fragile X chromosome, and adrenoleukodystrophy.**

Restriction-fragment-length polymorphisms (RFLPs) are codominant genetic markers that provide powerful tools for the analysis of a variety of genetic problems (1–3). RFLPs associated with genes directly involved in genetic diseases can be used for prenatal or carrier diagnosis of hemoglobinopathies (4, 5), phenylketonuria (6), hemophilia B (7, 8), and Lesch–Nyhan syndrome (9). It has been proposed more generally that a complete linkage map of the human genome could be established by using RFLPs detected by random DNA sequences, which could incorporate loci corresponding to genetic diseases even if the basic defect of some of these diseases is unknown (1). This prediction is beginning to be fulfilled, since linkage maps have been constructed for the human X chromosome (10) and for part of chromosome 11 (11, 12). RFLPs markers have been found for sex-linked recessive diseases [Duchenne muscular dystrophy, mental retardation with a fragile X chromosome, retinitis pigmentosa (13–16)] and for Huntington chorea, an autosomal dominant disease whose chromosomal localization was previously unknown (17). More than 160 DNA probes have been isolated that detect polymorphisms in the human genome (18). However, most of these detect polymorphisms with two alleles, which have an informativeness limited in the best cases to 0.5 for X-linked markers and 0.375 for autosomal ones (19). By testing a large number of different restriction enzymes, it is often possible to find several such RFLPs within a given genomic region as shown in the case of the β-globin gene cluster (5), but this complicates the use of the marker. However, a few probes reveal complex polymorphisms with a large number of alleles (20–23), often generated through sequence rearrangements rather than point mutations. These loci are extremely informative and useful for genetic analysis. We describe here an X chromosome-specific probe, isolated from the library of Davies *et al.* (24), that detects several polymorphisms in human DNA digested with *Taq* I or *Msp* I and defines 64 potential haplotypes using only these two enzymes.

MATERIALS AND METHODS

The cell lines used in gene-dosage experiments or for subchromosomal localizations have been described (25). DNA isolation, electrophoresis, blotting, and hybridization (in the presence of 40% formamide/5% dextran sulfate at 42°C) were performed as described (25). DNAs were blotted either on diazobenzyloxymethyl (DBM) paper or Zetapor membranes (10). After hybridization, filters were usually washed at 65°C in the presence of 0.5× NaCl/Cit/0.1% $NaDodSO_4$ (1× NaCl/Cit = 0.15 M NaCl/0.015 M Na citrate, pH 7).

RESULTS

A Probe That Detects a Set of Related Sequences in Xq26–Xq28. A 9.3-kilobase (kb) *Eco*RI fragment [laboratory acronym *St14*; designated *DXS52* in the catalogue of cloned human DNA segments (18)] devoid of repetitive sequences was isolated from a DNA library specific for the human X chromosome (24) by using a systematic screening procedure (unpublished data). X linkage was assessed by a gene-dosage experiment with genomic DNAs obtained from normal males (46 XY), normal females (46 XX), and human cell lines with 48 XXXX, 49 XXXXY, and 49 XYYYY karyotypes (25). Probe *St14* detected six *Eco*RI fragments in human DNA; the fragments differed in size and intensity of the hybridization signal, including one with the size of the cloned probe (9.3 kb) (Fig. 1). All of these fragments (total size ≈ 65 kb) are X chromosome-specific as shown by their dosage, which paralleled the concentration of X-chromosome sequences in

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: RFLP, restriction-fragment-length polymorphism; kb, kilobase(s); *HEMA*, *G6PD*, *ALD*, *CBD*, and *CBP*, respective genes for hemophilia A, glucose-6-phosphate dehydrogenase, adrenoleukodystrophy, deutan color blindness, and protan color blindness.

‡Present address: Department of Genetics and Microbiology, University of Pavia, Pavia, Italy.

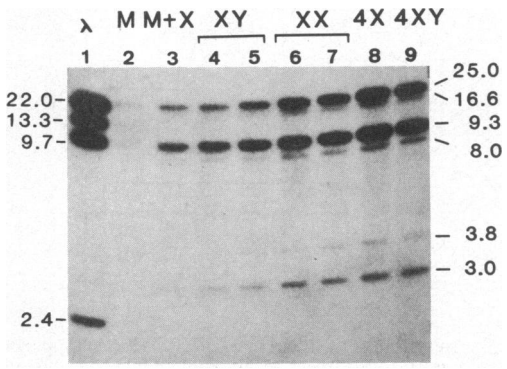

FIG. 1. Gene-dosage analysis of fragments hybridizing to probe *St14* in human DNA digested with *Eco*RI. The DNAs from the various rodent or human cell lines were digested with *Eco*RI and analyzed as described (25). After hybridization with the probe *St14*, the blot was washed in 1× NaCl/Cit containing 0.1% $NaDodSO_4$ at 60°C (low stringency). Origin of DNAs in lanes: 1, phage λ DNA digested by *Bgl* II; 2, mouse A9 cells (M); 3, hybrid line A9-HBRC2 containing an almost complete human X chromosome in addition to the mouse genome (M+X); 4 and 5, unrelated males (XY); 6 and 7, unrelated females (XX); 8, a 48 XXXX female; 9, a 49 XXXXY male. Sizes of the human DNA fragments and of the λ DNA markers are shown in kb.

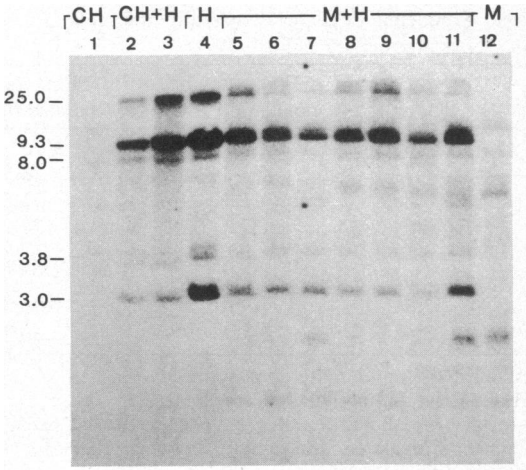

FIG. 2. Regional localization on the X chromosome of the fragments detected by the *St14* probe. The DNAs from the various rodent–human hybrid lines and from control rodent or human cells were digested and analyzed as in Fig. 1. Origin of DNAs in lanes: 1, hamster CH cells (CH); 2 and 3, hamster–human cell lines (CH + H); 2, CH 35A1 (Xq23–Xqter); 3, CH 34X (Xp22.3–Xqter); 4, human leukocyte DNA (H); 5–11, human–mouse cell lines containing different portions of the human X chromosome (M+H): GM194 RAG7 (Xpter–Xq28, including the *G6PD* gene; lane 5), GM97 RAG8 13 (Xq26–Xqter; lane 6), GM89 A9 (Xq22–Xqter; lane 7), Anly RAG1 (Xq12–Xqter; lane 8), GO RAG4 (Xp11–Xqter; lane 9), PI RAG7-2 (Xp22–Xqter; lane 10), A9-HRBC2 (Xp22.3(?)–Xqter; lane 11), and mouse A9 cells (M; lane 10). Sizes of the human DNA fragments detected by the probe *St14* are indicated in kb.

the various DNAs (the hybridization of the same blot with autosomal probes showed that equal amounts of DNA were present in each lane; results not shown).

To determine the regional localization of the fragments detected by *St14* on the X chromosome, we used a panel of rodent–human cell lines derived from human parental cells carrying X–autosome translocations with various breakpoints on the X chromosome (ref. 25; unpublished data). The DNAs from the various lines were digested with *Eco*RI and blotted onto diazobenzyloxymethyl paper. Hybridization with the *St14* probe showed that at least five of the six X-chromosome fragments detected cosegregate in the hybrid lines (Fig. 2). The signal obtained with the 16.5-kb fragment was very faint in the normal human DNA control (two X chromosomes) and, thus, was below the detection level in the hybrids. The results indicate that the sequences homologous to the *St14* probe are present in the region distal to the Xq26 breakpoint of the cell line GM97 translocation and proximal to the breakpoint of the cell line GM194 translocation in q28. It should be noted, however, that all markers tested [including the gene for glucose-6-phosphate dehydrogenase (*G6PD*) (26) and many DNA probes (unpublished data)] are present in the GM194 hybrid, so that this line might lack only a tiny portion (if any) of the q28–qter region, also undetectable cytogenetically.

***St14* Probe Reveals Complex Polymorphisms with *Taq* I and *Msp* I.** *Taq* I. The *St14* probe was tested for its ability to detect RFLPs, by using 11 restriction enzymes and a panel of 5–10 independent female DNA samples. With DNA digested by *Taq* I (Fig. 3), a complex pattern of bands was observed [12 fragments in male DNAs, which all appear to be X-linked based on gene-dosage experiments (data not shown)]. Eleven of the fragments (ranging from 5.5 to 0.7 kb) were always present with the same migration in all DNAs tested, whereas a variable pattern was observed in the 3.4- to 6.6-kb region. A series of at least eight allelic fragments defined a polymorphic system showing X-linked inheritance. Only one such fragment was present in male DNAs, whereas one or two could be detected in female DNAs. We classified the alleles according to the size of the variable fragment, from 6.6 kb for allele *1* to 3.4 kb for allele *8*. Allele *2* was often difficult to discriminate from the constant 5.5- to 5.4-kb doublet. This could be achieved by using as probe a 1.3-kb fragment isolated from the original 9.3-kb genomic sequence (results not shown). In fact, allele *2* was heterogeneous, and the example shown in lane 12 corresponds to a fragment slightly smaller than that seen in other families. It was also sometimes difficult to differentiate between alleles *5* and *6* for samples that were on different blots because of the small size difference of the fragments. The frequency of alleles was determined in two populations. In Strasbourg, we analyzed a set of families that segregate for various X-linked diseases (hemophilia A and B, fragile X chromosome mental retardation, etc.), while a panel of normal families (10, 12) was analyzed in Salt Lake City.

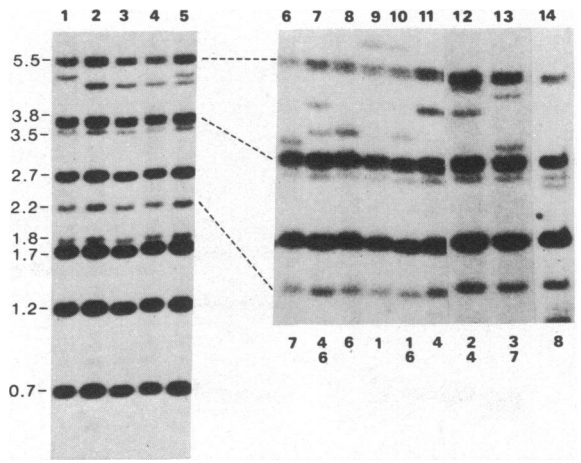

FIG. 3. Polymorphic pattern of *Taq* I fragments hybridizing to the *St14* probe. Human DNAs were digested with *Taq* I and analyzed by blot hybridization with the complete 9.3-kb insert. Lanes: 1–5, complete *Taq* I pattern observed in a family, showing segregation of alleles *3* and *4* (sizes of the constant fragments are shown in kb); 6–14, characterization of the various allelic fragments. The corresponding genotype is indicated below each lane.

Table 1. Frequency of alleles detected by the probe *St14* in human DNAs digested by *Taq* I

Allele	Size, kb	Frequency in population 1,* %	Frequency in population 2,† %
1	6.6	0.7	0
2	5.3	5.5	3.9
3	4.8	11.7	11.8
4	4.5	35.9	37.2
5	4.1	21.4‡	19.6‡
6	4.0		
7	3.9	9.0	13.7
8	3.4	15.9	13.7

*Strasbourg; 145 unrelated chromosomes.
†Salt Lake City; 54 unrelated chromosomes.
‡The results for alleles *5* and *6* have been pooled because it is sometimes difficult to differentiate between them for samples present on different blots.

As shown in Table 1, allele *4* was the most frequent (approximately one-third of the chromosomes analyzed), while other alleles had frequencies of <0.16. Allele *1* was observed only once. No significant difference was found between the two populations, which can both be defined as Caucasian (although a few non-European families were included in population 1). From the frequencies of alleles, it can be calculated that the probability of being heterozygous for this single polymorphism is 0.78.

Msp I. A variable pattern of bands also was observed in human DNA digested with the enzyme *Msp* I. These variable bands correspond to three separate polymorphisms, each characterized by two alleles (Fig. 4). The first system corresponds to allelic fragments of 4.4 and 3.6 kb (named *1* and *2*, respectively). It should be noted that the variant site appears to be sensitive to underdigestion by *Msp* I (see Fig. 4*B*, fourth and fifth lanes), which in some cases made it difficult to determine unambiguously the genotype for this polymorphism, especially in females: incomplete digestion could generate a pattern suggesting heterozygosity (*1/2*) in a female with a *2/2* genotype. The second system corresponds to allelic fragments of 2.0 and 1.6 kb (named *3* and *4*). The genotype at this second locus could be determined unambiguously by using as probe a 1.3-kb fragment rather than the complete *St14* insert (Fig. 4*C*). Finally, a third independent

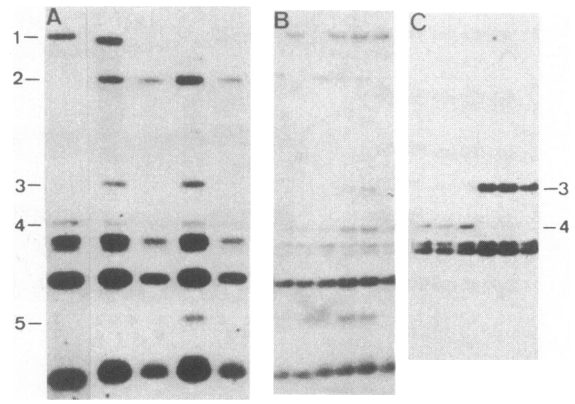

FIG. 4. Polymorphic pattern of *Msp* I fragments hybridizing with the probe *St14*. DNA samples were digested by *Msp* I and analyzed by blot hybridization. The 9.3-kb insert was used as probe in *A* and *B*. A 1.3-kb *Taq* I fragment was used in *C* to probe the same samples as in *B*. Male and female DNA samples are analyzed in *A*, while only male samples are present in *B* and *C*. Two samples in *B* present both bands 1 and 2 because of underdigestion at the relevant site (see text). Bands corresponding to allelic fragments are numbered. Their size is given in Table 2.

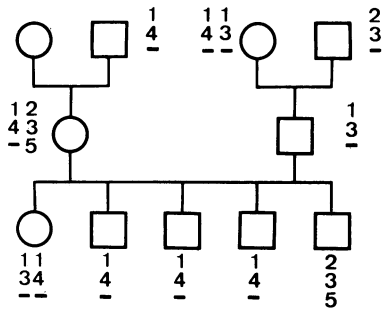

FIG. 5. Inheritance of the *Msp* I RFLP. DNA samples from pedigree no. 1340 were analyzed by blot hybridization with the 9.3-kb probe (alleles *1*, *2*, and *5*) or a 1.3-kb subfragment (alleles *3* and *4*).

polymorphism corresponds to the presence or absence of a 1.0-kb band (named *5*). It is likely that the fragment corresponding to the second allele of this polymorphism was either too small or hidden behind a constant fragment and, thus, escaped detection. As a result, it is not possible to distinguish between the genotypes *5/5* and *5/–* in females in the absence of additional family data.

The segregation of the five variable fragments was analyzed in several families, demonstrating that the three polymorphic systems are distinct (Fig. 5). The absence of recombination between them and the presence of specific combinations of alleles between the three RFLPs (see below) suggest that they belong to a single genetic system and, thus, can define haplotypes.

The frequency of the alleles detected with *Msp* I is presented in Table 2. Each polymorphic system had a frequency in the useful range for genetic analysis. Furthermore, the analysis in 57 unrelated chromosomes of the haplotypes for the *1/2* and *3/4* allelic systems showed that they should be informative in about 70% of the cases. On the other hand, fragment *5* was found to be associated in 10 cases out of 11 to the *Msp* I haplotype *2/3*, an indication of linkage disequilibrium. Six of the eight potential haplotypes defined by the three *Msp* I RFLPs were detected so far (four of them segregate in the family presented in Fig. 5). This suggests that the combination of these polymorphisms should be very useful for genetic analysis.

Localization of the *St14* Probe by Linkage Analysis to RFLPs. A previous study of linkage between various RFLPs on the X chromosome has defined the respective positions of three such markers in the q27–q28 region (10). Probe *DX13* revealed a *Bgl* II polymorphism showing 28% recombination with the gene for coagulation factor IX [hemophilia B locus (*HEMB*)] and 30% recombination with the random probe *52A* (which detects a *Taq* I RFLP). The data suggested the

Table 2. Frequency of alleles and haplotypes detected by the probe *St14* in human DNAs digested by *Msp* I

Allele or haplotype	Size, kb	Frequency, %	Unrelated chromosomes analyzed, no.
1	4.4	45.6	68
2	3.6	54.4	68
3	2.0	69.9	73
4	1.6	30.1	73
5	1.0	36.8	38*
1/3	—	35.1	57
1/4	—	10.5	57
2/3	—	35.1	57
2/4	—	19.3	57

*Only males were analyzed because genotype cannot be determined unambiguously in females.

order centromere-*52A-HEMB-DX13*. Segregation of the *Taq* I (or *Msp* I) alleles at the *St14* locus was analyzed in families that were informative for either of the three markers. No recombination was observed between the *St14* and *DX13* loci (in 41 informative meioses), while a 30% recombination fraction was found between *St14* and the gene for coagulation factor IX (*HEMB*) or *52A* (Table 3). Thus, *St14* can be localized at the distal end of the long arm of the X chromosome, close to the *DX13* probe.

DISCUSSION

The probe *St14* reveals one of the most polymorphic loci presently known in the human genome. This is due in part to the fact that this 9.3-kb DNA fragment hybridizes to a series of bands amounting to 65 kb, located in the q26–q28 region of the X chromosome. It is likely that a sequence contained within the cloned fragment is repeated at least three to four times with a varying degree of homology (or of length of the repeat unit) to account for the differing intensities of bands detected by the probe. The *St14* probe hybridizes to sequences in the mouse and hamster genome under conditions of reduced stringency (washing at 60°C in 1× NaCl/Cit; see Fig. 2). Since nonexpressed sequences are unlikely to be conserved between man and rodents (divergence time = 80 × 10^6 yr), this could suggest that the *St14* probe detects a small cluster of related genes.

The complex *Taq* I RFLP revealed with the *St14* probe is defined by a series of at least eight allelic fragments, ranging from 6.6 to 3.4 kb, which could be generated by sequence rearrangement. In the similar case of the polymorphisms at the insulin and c-Ha-*ras* loci, it has been shown by sequencing that an allelic series is created by variation in the number of tandemly repeated units (21, 22). The measurement of sizes of the *Taq* I alleles is too imprecise to support or refute such a mechanism. It should be noted that if sequence rearrangements were the cause of such a polymorphism, it should be possible to detect it by using other restriction enzymes. This has not been found with the few restriction enzymes (*Eco*RI, *Pst* I, and *Msp* I) that have been tested so far under conditions that allow the detection of cross-hybridizing fragments. This negative result might be due to the location of relevant restriction sites with respect to both the hybridizing region and the polymorphic region. Alternatively, this polymorphism might be created by independent point mutations in a region that contains a high density of *Taq* I sites, as in the case of the complex *Sst* I RFLP found in the switch region of the immunoglobulin heavy chain genes (23). The three independent *Msp* I RFLPs are more likely to be due to separate point mutations. One can assume that the four polymorphic systems define a single genetic locus for the following reasons. We recently cloned the two smaller *Eco*RI fragments homologous to the *St14* probe, and this allowed us to demonstrate that the region of the *Taq* I RFLP is present within 5 kb of the site corresponding to *Msp* I alleles *3* and *4*. The presence of specific associations between alleles belonging to different RFLPs is another indication that they are located close to each other: *Taq* I alleles *4* and *7* have been found to be associated to *Msp* I allele *3*, while *Taq* I allele *8* was associated to *Msp* I allele *4*, and *Msp* I allele *5* is preferentially found in a *Msp* I *2/3/5* haplotype (results not shown). Structural analysis of the *St14* locus and mapping of the polymorphic sites (or regions) should give a definitive answer to this point.

Table 3. Linkage between *St14* and markers in the Xq27–Xqter region

Pairwise cross	No. of recombinants/ total no. of chromosomes	Recombination fraction ± SEM
St14–DX13	0/41	0 (≤0.056)
*St14–HEMB**	11/37	0.297 ± 0.075
St14–52A	12/41	0.292 ± 0.071
*DX13–HEMB**	6/21	0.285 ± 0.098

All of the recombination data were obtained from three-generation phase-known families. SEM for the recombination fraction was estimated by the formula $\text{SEM} = \left[\frac{P(1-P)}{n}\right]^{1/2}$ where P is the recombination fraction and *n* is the number of meioses analyzed. For the *St14–DX13* cross, the upper limit (at 90% confidence) for the recombination fraction was calculated as proposed by Morton (27).
*Locus for coagulation factor IX gene.

Table 4. Association of the most common *Taq* I alleles with various *Msp* I haplotypes

Taq I alleles	Frequency of *Msp* I haplotypes							
	1/3/5	*1/3/−*	*1/4/5*	*1/4/−*	*2/3/5*	*2/3/−*	*2/4/5*	*2/4/−*
4	0	9	0	0	4	3	0	0
7	0	2	0	0	3	1	0	0
8	0	0	0	7	0	0	0	4

Complete haplotype analysis was performed on 44 independent X chromosomes; 15 different combinations were found. Frequency of those combinations that carry *Taq* I alleles *4*, *7*, or *8* is presented.

Whatever the mechanisms generating the various RFLPs, the *St14* probe constitutes one of the best genetic markers in the human genome and certainly the best on the X chromosome. About 80% of all females are heterozygous, considering only the eight *Taq* I alleles, and a similar figure is found for the combination of the three *Msp* I RFLPs. The *Taq* I and *Msp* I RFLPs would be expected to generate 64 haplotypes if there were no linkage disequilibrium. Although our preliminary studies indicate that some preferential associations are present in the population, we have found that the most common *Taq* allele *4* can be associated to three different *Msp* allele combinations (1/3/−, 2/3/5, 2/3/−), and *Taq* allele *8* is found to be associated to *Msp* haplotypes 1/4/− and 2/4/− (Table 4). Although a detailed analysis will be required to determine the ultimate informativeness of the probe, our present results suggest that heterozygosity is at least 90%. Thus, *St14* is an almost ideal marker for the genetic analysis of the loci for important diseases present in the subtelomeric region of the X chromosome: Hemophilia A (*HEMA*), adrenoleukodystrophy (*ALD*), and mental retardation with a fragile X chromosome (fragile site *FS*).

The linkage analysis presented here shows that the *St14* locus is very close (in terms of recombination distance) from the previously described locus *DX13*. Five other loci in this region were already known to be very closely linked: *HEMA, G6PD, CBP* and *CBD* (protan and deutan color blindness, respectively), and *ALD* (see Table 5). Recent data demonstrate that *DX13* and *St14* are part of this cluster, since no recombinants were found between *DX13* and *HEMA* (28) or between *St14* and *HEMA* (38; see Table 5). On the other hand, the gene for factor IX (*HEMB* locus), which is located in Xq27 (25, 33, 34), shows about 30% re-

Table 5. Linkage data for loci in the Xq28 region

Cross	lod score*	θ_{max}	References
DX13—St14	12.3	0	This paper
DX13—HEMA	5.4	0	28
St14–HEMA	9.6	0	38
G6PD–HEMA	17.5	0	29, 30
G6PD–ALD	5.4	0	31
CBD–G6PD	37.8	0.02	32
CBD–CBP	15.9	0.06	32
CBD–HEMA	17.1	0.03	32

*Logarithm of odds in favor of linkage.

combination with *DX13* or *St14* and did not show in previous analysis measurable linkage with the color blindness (*CBD/CBP*)–*HEMA* cluster (a total of 16 recombinants were found in 37 meioses analyzed, cited in ref. 35). With the exception of *FS*, the fragile X chromosome locus (ref. 36; unpublished data), no locus has been found at an intermediate distance of the *G6PD* cluster. We suggest that the presence of such a cluster of seven tightly linked loci and the large genetic distance between loci in q28 and the *HEMB–52A* loci indicate a highly nonrandom distribution of recombination events in this region. Crossing-overs would be highly favored in a region between *HEMB* (in Xq27) and the loci in Xq28: the 30% recombination figure corresponds to a frequency of at least 0.6 crossover per meiosis. It is interesting to note that such a recombination hot spot would be close to a region that can show chromosome breakage in certain individuals—i.e., the fragile site *FS* at the Xq27–Xq28 interface, as also proposed independently by Szabo *et al.* (35). Our segregation data demonstrate that *FS* is located between the *HEMB* and *St14* loci, at approximately equal distances (≈12% recombination) from these two markers (unpublished data). In contrast, few recombination events would occur in the distal q28 region, perhaps as a consequence of interference. If such a heterogeneity of recombination frequency occurs at other places in the human genome, as is also suggested by chiasmata mapping (32, 37), it might be expected that segregation analysis would demonstrate close linkages [such as found in the case of Huntington's chorea (17)] more often than expected.

Note Added in Proof. No recombinants were found in 57 meioses between the *St14* locus and the gene for coagulation factor VIII, which confirms the very close linkage between *St14* and *HEMA* (39).

We thank Dr. K. Davies and R. Williamson for the generous gift of the genomic library, C. Kloepfer for precious help, and C. Aron and B. Boulay for preparing the manuscript. D. Drayna is a fellow of the Muscular Dystrophy Association. R.W. is an investigator of the Howard Hughes Medical Institute. G.C. was supported by a European Molecular Biology Organization long-term fellowship. This work was supported by grants from the Institut National de la Santé et de la Recherche Médicale (PRC 134029), from the Ministère de l'Industrie et de la Recherche (C0671), from the Fondation pour la Recherche Médicale Française, and from Agir (to J.-L.M.) and by grants from the Muscular Dystrophy Association and from the National Institutes of Health (Grant GM 29789).

1. Botstein, D., White, R. L., Skolnick, M. & Davis, R. W. (1980) *Am. J. Hum. Genet.* **32,** 314–331.
2. Kan, Y. W. & Dozy, A. M. (1980) *Science* **209,** 388–391.
3. Cavenee, W. K., Dryja, T. P., Phillips, R. A., Benedict, W. F., Godbout, R., Gallie, B. L., Murphree, A. L., Strong, L. C. & White, R. L. (1983) *Nature (London)* **305,** 779–784.
4. Little, P. F. R., Annison, G., Darling, S., Williamson, R., Camba, L. & Modell, B. (1980) *Nature (London)* **285,** 144–147.
5. Antonarakis, S. E., Boehm, C. D., Giardina, P. V. J. & Kazazian, H. H. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 137–141.
6. Woo, S. L. C., Lidsky, A. S., Güttler, F., Chandra, T. & Robson, K. J. H. (1983) *Nature (London)* **306,** 151–155.
7. Grunebaum, L., Cazenave, J. P., Camerino, G., Kloepfer, C., Mandel, J. L., Tolstoshev, P., Jaye, M., De la Salle, H. & Lecocq, J. P. (1984) *J. Clin. Invest.* **73,** 1491–1495.
8. Gianelli, F., Choo, K. H., Winship, P. R., Rizza, C. R., Anson, D. S., Rees, D. J. G., Ferrari, N. & Brownlee, G. G. (1984) *Lancet* **i,** 239–241.
9. Nussbaum, R. L., Crowder, W. E., Nyhan, W. L. & Caskey, C. T. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 4035–4039.
10. Drayna, D., Davies, K., Hartley, D., Mandel, J. L., Camerino, G., Williamson, R. & White, R. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 2836–2839.
11. Antonarakis, S., Philipps, J., Mallonee, R., Kazazian, H., Fearon, E., Waher, P., Kronenberg, H., Ullrich, A. & Meyers, D. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 6615–6619.
12. White, R., Leppert, M., Bishop, D. T., Barker, D., Berkowitz, J., Brown, C., Callahan, P., Holm, T. & Jerominski, L. (1985) *Nature (London)* **313,** 101–105.
13. Murray, J. M., Davies, K. E., Harper, P. S., Meredith, L., Mueller, C. R. & Williamson, R. (1982) *Nature (London)* **300,** 69–71.
14. Davies, K. E., Pearson, P. L., Harper, P. S., Murray, J. M., O'Brien, T., Sarfarazi, M. & Williamson, R. (1983) *Nucleic Acids Res.* **11,** 2303–2312.
15. Camerino, G., Mattei, M. G., Mattei, J. F., Jaye, M. & Mandel, J. L. (1983) *Nature (London)* **306,** 701–704.
16. Bhattacharya, S. S., Wright, A. F., Clayton, J. F., Price, W. H., Philipps, C. I., McKeown, C. M. E., Jaye, M., Bird, A. C., Pearson, P. L., Southern, E. M. & Evans, H. J. (1984) *Nature (London)* **309,** 253–255.
17. Gusella, J. F., Wexler, N. S., Conneally, P. M., Naylor, S. L., Anderson, M. A., Tanzi, R. E., Watkins, P. C., Ottina, K., Wallace, M. R., Sakaguchi, A. Y., Young, A. B., Shoulson, I., Bonilla, E. & Martin, J. B. (1983) *Nature (London)* **306,** 234–238.
18. Skolnick, M. H., Willard, H. F. & Menlove, L. A. (1984) *Cytogenet. Cell Genet.* **37,** 210–273.
19. Skolnick, M. H. & White, R. (1982) *Cytogenet. Cell Genet.* **32,** 58–67.
20. Wyman, A. R. & White, R. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 6754–6758.
21. Bell, G. I., Selby, M. J. & Rutter, W. J. (1982) *Nature (London)* **295,** 31–35.
22. Barker, D. & White, R. (1984) *Cytogenet. Cell Genet.* **37,** 412–413.
23. Migone, N., Feder, J., Cann, H., Van West, B., Hwang, J., Takahashi, N., Honjo, T., Piazza, A. & Cavalli-Sforza, L. L. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 467–471.
24. Davies, K. E., Young, B. D., Elles, R. G., Hill, M. E. & Williamson, R. (1981) *Nature (London)* **293,** 374–376.
25. Camerino, G., Grzeschik, K. H., Jaye, M., De La Salle, H., Tolstoshev, P., Lecocq, J. P., Heilig, R. & Mandel, J. L. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 498–502.
26. Hellkuhl, B., De la Chapelle, A. & Grzeschik, K. H. (1982) *Hum. Genet.* **60,** 126–129.
27. Morton, N. E. (1955) *Am. J. Hum. Genet.* **9,** 55–75.
28. Harper, K., Winter, R. M., Pembrey, M. E., Hartley, D., Davies, K. E. & Tuddenham, E. G. D. (1984) *Lancet* **ii,** 6–8.
29. Tsevrenis, H., Mandalaki, T., Volkers, W. S. & Khan, P. M. (1979) *Cytogenet. Cell Genet.* **25,** 213 (abstr.).
30. Filippi, G., Mannucci, P. M., Coppola, R., Farris, A., Rinaldi, A. & Siniscalco, M. (1984) *Am. J. Hum. Genet.* **36,** 44–71.
31. Migeon, B. R., Moser, H. W., Moser, A. B., Axelman, J., Sillence, D. & Norum, R. A. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 5066–5070.
32. Keats, B. (1983) *Hum. Genet.* **64,** 28–32.
33. Chance, P. F., Dyer, K. A., Kurachi, K., Yoshitake, S., Ropers, H. H., Wieacker, P. & Gartler, S. M. (1983) *Hum. Genet.* **65,** 207–208.
34. Boyd, Y., Buckle, V. J., Munro, E. A., Choo, K. H., Migeon, B. R. & Craig, I. W. (1984) *Ann. Hum. Genet.* **48,** 145–152.
35. Szabo, P., Purrello, M., Rocchi, M., Archidiacono, N., Alhadeff, B., Filippi, G., Toniolo, D., Martini, G., Luzzatto, L. & Siniscalco, M. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 7855–7859.
36. Filippi, G., Rinaldi, A., Archidiacono, N., Ricchi, M., Balazs, I. & Siniscalco, M. (1983) *Am. J. Med. Genet.* **15,** 113–119.
37. Hulten, M. (1974) *Hereditas* **76,** 55–78.
38. Oberlé, I., Camerino, G., Heilig, R., Grunebaum, L., Cazenave, J. P., Crapanzano, C., Mannucci, P. M. & Mandel, J. L. (1985) *N. Engl. J. Med.*, in press.
39. Gitschier, J., Drayna, D., Tuddenham, E. G. D., White, R. L. & Lawn, R. M. (1985) *Nature (London)*, in press.