Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 1796–1800, March 1990
Genetics

Additional polymorphisms at marker loci *D9S5* and *D9S15* generate extended haplotypes in linkage disequilibrium with Friedreich ataxia

RICARDO FUJITA, ANDRÉ HANAUER, GIORGIO SIRUGO, ROLAND HEILIG, 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 Institut National de la Santé et de la Recherche Médicale, Faculté de Médecine, 11 Rue Humann, 67085 Strasbourg Cédex, France

Communicated by Jean Dausset, December 13, 1989

ABSTRACT **The gene for Friedreich ataxia (FA), a severe recessive neurodegenerative disease, has previously been shown to be tightly linked to the polymorphic markers *D9S15* and *D9S5* on human chromosome 9. In addition, the observation of linkage disequilibrium suggested that *D9S15* is within 1 centimorgan (cM) of the disease locus, *FRDA*. Although *D9S5* did not show recombination with *FRDA*, its localization was less precise (0–5 cM) due to its lower informativeness. We have now identified additional polymorphisms at both marker loci. Two cosmids spanning 50 kilobases around *D9S5* were isolated, and a probe derived from one of them detects an informative three-allele polymorphism. We have found a highly polymorphic microsatellite sequence at *D9S15* which is rapidly typed by the DNA polymerase chain reaction. The polymorphism information contents at the *D9S5* and *D9S15* loci have been increased from 0.14 to 0.60 and from 0.33 to 0.74, respectively. With the additional polymorphisms the lod ($\log_{10}$ odds ratio) score for the *D9S15–FRDA* linkage is now 48.10 at recombination fraction $\theta = 0.005$ and for *D9S5–FRDA*, the lod score is 27.87 at $\theta = 0.00$. We have identified a recombinant between *D9S15* and *FRDA*. However, due to the family structure, it will be of limited usefulness for more precise localization of *FRDA*. The linkage disequilibrium previously observed between *D9S15* and *FRDA* is strengthened by analysis of the haplotypes using the microsatellite polymorphism, while weaker but significant disequilibrium is found between *D9S5* and *FRDA*. Extended haplotypes that encompass *D9S5* and *D9S15* show a strikingly different distribution between chromosomes that carry the FA mutation and normal chromosomes. This suggests that both marker loci are less than 1 cM from the *FRDA* gene and that a small number of mutations account for the majority of FA cases in the French population studied. *D9S5* and *D9S15* are thus excellent start points to isolate the disease gene.**

Friedreich ataxia (FA) is a progressive neurodegenerative disorder, involving both central and peripheral nervous systems, with autosomal recessive inheritance. The disease is characterized by gait ataxia, muscle weakness, areflexia of the lower limbs, dysarthria, and sensory disturbances. It is usually manifest, in its classical form, around puberty (1, 2). At present, there is no clue to the basic biochemical defect in FA (3). Therefore, efforts to understand the disorder are directed towards isolating the disease gene on the basis of its genomic map position. Chamberlain *et al.* (4) mapped the FA gene (*FRDA* in the Human Gene Mapping Workshop nomenclature) to chromosome 9 by virtue of its tight linkage to the polymorphic locus *D9S15* (probe pMCT112). We confirmed this linkage in another set of families and identified a second closely linked marker locus, *D9S5* (5). So far, *D9S15* and *D9S5* have displayed no recombination with *FRDA* or between themselves in all informative meioses examined. The combination of our data with those published by Chamberlain *et al.* (6) showed that *D9S15* is within one centimorgan (cM) of the disease locus (7). The linkage data, combined with physical mapping of *D9S5* by *in situ* hybridization, support the assignment of *FRDA* to the proximal long arm of chromosome 9 (7). The very close proximity of *D9S15* to *FRDA* was further indicated by the finding of a significant allelic association between *FRDA* and the *Msp* I restriction fragment length polymorphism (RFLP) at *D9S15*. In contrast, no linkage disequilibrium was noted between *FRDA* and previously described RFLPs at *D9S5* (7). Several families in our initial linkage analysis were partially or completely uninformative, especially for *D9S5*. In an attempt to refine the genetic map around the *FRDA* locus and to further explore the linkage disequilibrium, a search has been conducted for new informative polymorphisms at the marker loci. We have isolated two overlapping cosmid clones which extend the *D9S5* locus over 50 kilobases (kb). Additional RFLPs have been revealed by single-copy probes derived from these cosmids. One of these (probe 26P) detects a three-allele *Bst*XI RFLP that increases considerably the heterozygosity of *D9S5*. We have also characterized within probe MCT112 (*D9S15*) a highly polymorphic "microsatellite" sequence (8, 9). This allowed us to confirm that *D9S5* and *D9S15* are very tightly linked to *FRDA*. By using these polymorphisms at both loci, extended haplotypes can be generated, and some of them show highly significant linkage disequilibrium with *FRDA*.

MATERIALS AND METHODS

Linkage Studies. Forty-one FA families (34 from France), each with at least two affected children, were studied. Thirty-eight of these families have been previously analyzed for the *Msp* I RFLPs at *D9S5* and *D9S15* (7). All of these families conform to the diagnostic criteria for typical FA (1, 2). Southern blot analysis was performed as described (7). Linkage analysis was carried out by using the LINKAGE program (10) and the lod-1 support interval (lod = $\log_{10}$ odds ratio) was taken as approximating 95% probability limits. Parental haplotypes were established by comparison with those of the affected or normal children. χ^2 statistics were used to assess the significance of the distributions observed. Polymorphism information content (PIC) was calculated by using allele or haplotype frequencies determined in normal chromosomes from parents of FA patients.

Analysis of the Microsatellite Polymorphism at *D9S15*. A sequence of about 160 base pairs (bp) was amplified by using the primers MX57 (5′-CCCTCAGGTCTATTGAAGAAG-

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: FA, Friedreich ataxia; cM, centimorgan; RFLP, restriction fragment length polymorphism; PIC, polymorphism information content.

*To whom reprint requests should be addressed.

TATAAATG-3′) and MZ68 (5′-GATGGTGGTAATCATT-TCACAG-3′).

One of the primers was end-labeled for 1 hr at 37°C in a 100-μl reaction mixture containing 300 pmol of primer, 40 pmol of [γ-^{32}P]ATP at 5000 Ci/mmol (1 Ci = 37 GBq), 10 mM $MgCl_2$, 5 mM dithiothreitol, 50 mM Tris·HCl (pH 7.6), and 15 units of T4 polynucleotide kinase. Polymerase chain reactions were carried out in 25 μl with 1 μl of end-labeled primer reaction mixture, 3 pmol of the same unlabeled primer and 6 pmol of the second unlabeled primer, each dNTP at 200 μM, 50 mM KCl, 10 mM Tris·HCl (pH 8.2), 1.5 mM $MgCl_2$, 100 ng of genomic DNA, and 1 unit of *Thermus aquaticus* (*Taq*) DNA polymerase (Perkin–Elmer/Cetus). The samples were subjected to 20–25 cycles consisting of 0.75 min at 92°C (denaturation), 0.75 min at 65°C (annealing), and 1 min at 72°C (extension). One-tenth of each reaction mixture was analyzed on an 8% polyacrylamide/8 M urea sequencing gel.

RESULTS

Cloning of a 50-kb Region Around *D9S5* and Characterization of RFLPs. To expand the *D9S5* locus, probe DR47 was used to screen a nonamplified genomic library (11) from a (49XXXXY) human lymphoblastoid cell line. Two overlapping clones, designated c25 and c26, were identified with insert size of 40 and 29 kb, respectively, together extending over 50 kb around probe DR47 (Fig. 1). The end fragments of both cosmids were subcloned by digestion with restriction enzymes that have a single site in the polylinker region of the cosmid vector, followed by ligation at low concentration, as described by Heilig *et al.* (11). We generated 10 end-fragment "miniclones" (some of them contain overlapping inserts), which were analyzed for the presence of repetitive sequences. The miniclones showing no or faint signal when hybridized to labeled total genomic DNA were further analyzed for their ability to identify RFLPs.

A rapid screening procedure for RFLPs was used, as described by Murray *et al.* (12). DNAs from 10 unrelated individuals were pooled and compared to DNA isolated from the hybrid cell line CF11.4, which contains an X/9 translocation chromosome Xqter→Xq13::9q34→9pter with no other human chromosome (13). Potential polymorphisms were checked for mendelian inheritance. Our search involved 42 restriction enzymes and five probes, and it allowed us to identify five polymorphisms at the *D9S5* locus with the subclones 26P, 25X, and 25B (Fig. 1). The subclone 26P is the most polymorphic, revealing RFLPs with at least three different restriction endonucleases. In *Bst*XI digests, 26P identifies a three-allele system with allelic fragments of 5.1 kb, 4.1 kb + 1.0 kb, and 3.5 kb + 1.6 kb present in 20%, 30.7%, and 49.3% of normal chromosomes, respectively (Table 1). The pattern of fragments (Fig. 2) is consistent with site variation at each of two adjacent *Bst*XI sites in genomic DNA. The PIC is 0.55, indicating that this polymorphism will be useful in a large proportion of matings. We determined the frequencies of the haplotypes generated by combination of the 26P/*Bst*XI RFLP and the *DR47/Msp* I RFLP (Table 1). Of the six possible haplotypes, only four were observed in both normal and FA chromosomes, and a single haplotype (*B3M1*) accounts for about half of normal chromosomes. The minor allele (*M2*) of *DR47/Msp* I was always associated with allele *B2* of 26P/*Bst*XI, but haplotypes *B2M2* and *B2M1* were both observed. Although linkage disequilibrium is present between these RFLPs on normal chromosomes, analysis of both RFLPs gives additional information, and the combined PIC is 0.60, compared with 0.55 when the *Bst*XI RFLP is used alone. Subclone 26P reveals at least two other two-allele polymorphisms, with *Dra* I and *Xba* I, that have not been further characterized. Subclone 25X detects a two-allele *Taq* I RFLP (Fig. 2) with a minor allele frequency of 0.037 and a PIC value of 0.075. This infrequent RFLP only modestly increases the informativeness at the *D9S5* locus. A rare *Nco* I RFLP was also found for subclone 25B.

Table 1. Haplotype frequencies at *D9S5*

RFLPs		Haplotype frequency, %	
26P *Bst*XI (*B*)	DR47 *Msp* I (*M*)	Normal chromosomes (n = 75)	FA chromosomes (n = 78)
1	*1*	20.0	29.5
2	*1*	24.0 } 30.7	34.6 } 43.6
2	*2*	6.7 }	9.0 }
3	*1*	49.3*	26.9*

*$P < 0.05$ that frequencies are the same.

A Highly Polymorphic Microsatellite at *D9S15*. To identify additional polymorphisms at D9S15 by using the polymerase chain reaction (PCR) coupled with direct sequencing, the nucleotide sequences of both ends of MCT112 were determined. Unexpectedly, the sequence close to one end reveals the presence of a tandem repeat of the CA dinucleotide. Weber and May (9) and Litt and Luty (8) have reported that such sequences (termed microsatellites) often display polymorphic variation in the number of repeat units. Two primers that flank the $(CA)_n$ repeat were synthesized. To reveal the amplification products on polyacrylamide/urea sequencing gels, one of the primers was end-labeled prior to the polymerase chain reaction as described (9). Analysis of samples of genomic DNAs revealed a highly polymorphic marker with at least six alleles that differ in size by 2 bp (Fig. 3). The allele frequencies were determined in normal parental chromosomes from the FA families (Table 2). This revealed three frequent alleles: allele *MS2* (37.9%), allele *MS5* (33.8%), and allele *MS4* (16.9%). Among the infrequent alleles, allele *MS6*

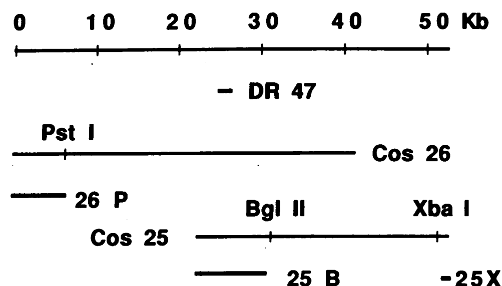

FIG. 1. Probes and cosmids at the *D9S5* locus. The relative position of cosmids 25 and 26 and single-copy probes are indicated, as well as restriction sites used for generating end clones.

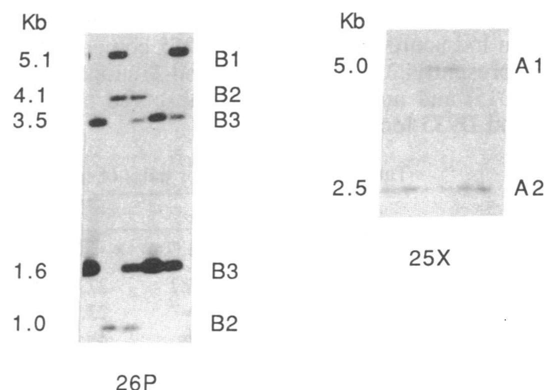

FIG. 2. Polymorphic patterns detected by probes 26P and 25X. (*Left*) The fragments corresponding to alleles *B1*, *B2*, and *B3* of the *Bst*XI RFLP detected by probes 26P. (*Right*) *A1* and *A2* correspond to the alleles detected by probe 25X in *Taq* I digests of human DNA.

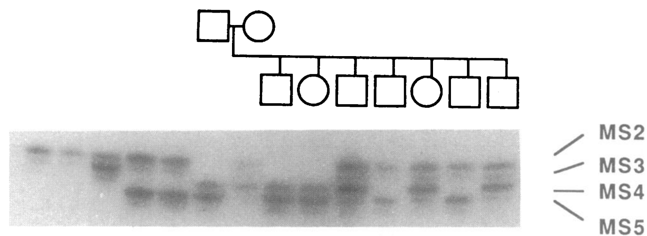

FIG. 3. Polymorphic pattern detected at the MCT112 microsatellite. The polymorphism was analyzed by polymerase chain reaction using a single end-labeled primer. Four of the six alleles are represented in the samples analyzed, which include a family (pedigree diagrammed) from the Centre d'Etude du Polymorphisme Humain (Paris) for demonstration of mendelian segregation.

was observed on only 3 FA chromosomes (out of 78). This polymorphism is very informative, with a PIC value of 0.65.

We examined for the presence of nonrandom association between the *Msp* I RFLP and the microsatellite polymorphism at *D9S15*. Haplotype frequencies were estimated in 149 parental chromosomes from the FA families (Table 2). Evidence for linkage disequilibrium between the polymorphisms is observed on both normal and FA chromosomes ($P < 0.001$ for normal chromosomes). For instance, the minor *Msp* I allele *2* is associated preferentially with allele *MS2* of the microsatellite (in a total of 57 chromosomes). *MS5* is always associated to allele *1* of the *Msp* I RFLP. Moreover, of the 12 possible haplotypes, only 9 were observed (one of which, *MS6 M1*, was observed only on FA chromosomes), and 4 haplotypes account for 87% and 92% of normal and FA chromosomes, respectively. The combined PIC of the two polymorphic sites at *D9S15* is 0.74, which makes them very useful for genetic analysis.

Linkage Analysis. Our previous data provided strong support that *D9S15* and *D9S5* are within 1 and 5 cM, respectively, of the *FRDA* gene (7). The identification of additional polymorphisms at both marker loci that extend dramatically the number of informative meioses for both markers allowed us to more accurately assess the linkage between *FRDA* and both *D9S15* and *D9S5*. Segregation of the new polymorphisms was analyzed in 41 FA families, which include the 38 FA families studied previously.

In 39 FA families one parent at least was informative for the 26P/*Bst*XI RFLP and 19 families were completely informative. The peak lod score for linkage between the 26P/*Bst*XI RFLP and *FRDA* was 23.18 at a recombination fraction $\theta = 0$ (not shown). A substantial increase in the lod score was obtained when the haplotypes composed of the three RFLPs at *D9S5* (26P/*Bst*XI, DR47/*Msp* I, and 25X/*Taq* I) were studied. This is justified as the polymorphic sites are within 50 kb and no intralocus recombination was observed. The maximum lod score for the *FRDA*–*D9S5* compound haplotype is presently 27.87 at $\theta = 0$ (lod-1 support interval 0.000–0.015) and no crossover was observed between the *FRDA* and *D9S5* loci (Table 3).

Table 2. Haplotype frequencies at *D9S15*

Alleles		Haplotype frequency, %			
MCT112 microsatellite (*MS*)	MCT112 *Msp* I (*M*)	Normal chromosomes ($n = 71$)		FA chromosomes ($n = 78$)	
1	*1*	0.0	4.2	0.0	2.5
1	*2*	4.2		2.5	
2	*1*	14.0	37.9	7.7	59.0
2	*2*	23.9*		51.3*	
3	*1*	5.6	7.0	1.3	1.3
3	*2*	1.4		0.0	
4	*1*	15.5	16.9	20.5	20.5
4	*2*	1.4		0.0	
5	*1*	33.8	33.8*	12.8	12.8*
5	*2*	0.0		0.0	
6	*1*	0.0	0.0	3.8	3.8
6	*2*	0.0		0.0	

*$P < 0.01$ that the frequencies of the haplotype are the same in normal and FA chromosomes.

We examined further the linkage relationship between *FRDA* and *D9S15* with the microsatellite polymorphism. Thirty-nine families were found to be informative for the analysis with the microsatellite polymorphism at *D9S15* in that at least one parent was heterozygous. Fourteen families were completely informative for this marker. One obligate crossover was detected between the microsatellite polymorphism and *FRDA* in an Italian family that was not included in our previous linkage studies. When analyzed for the *Msp* I/MCT112 RFLP, the recombination was confirmed (Fig. 4). A maximum lod score of 18.84 was obtained at $\theta = 0.01$. When the *Msp* I/MCT112 RFLP and the microsatellite polymorphism were combined, the lod score reached 23.51 at $\theta = 0.005$.

In the family exhibiting the recombination event, the two affected children have unequivocal FA that meets the criteria of Harding (ref. 2) (G. Nappi and A. P. Verri, Pavia, Italy, personal communication), and a possible sample mixup has been ruled out. The father is deceased and his genotype can be only partially deduced. The mother is homozygous for both *D9S15* markers and since her two affected children have inherited different alleles from their father, one of the two is recombinant between *FRDA* and *D9S15*. The recombination event must have occurred on a paternal chromosome 9. Because the genotype of father cannot be determined it is not possible to infer whether or not a recombination event has occurred between *D9S5* and *FRDA*. The two sibs might have inherited the same genotypes at *D9S5* either because no recombination has taken place or due to homozygosity of the father at this locus. According to the haplotype frequencies at *D9S5* given in Table 1, there is a 50% chance of the father's being homozygous for the *B3M1* haplotype, the most frequent one. By combining our data with those previously

Table 3. lod scores for pairs of loci at various values of recombination fraction (θ)

Pair of loci	lod score						
	$\theta = 0.000$	$\theta = 0.005$	$\theta = 0.01$	$\theta = 0.02$	$\theta = 0.03$	$\theta = 0.04$	$\theta = 0.05$
*D9S5**/*FRDA*	27.87	27.60	27.14	26.42	25.69	24.96	24.22
D9S15†/*FRDA*	$-\infty$	23.51	23.49	23.16	22.70	22.19	21.65
D9S15/*FRDA*‡	$-\infty$	48.10§	47.71	46.70§	45.60§	44.40§	43.20
D9S5¶/*D9S15*†	26.17	25.90	25.63	25.09	24.54	23.98	23.41

*Combined analysis of the 26P/*Bst*XI, DR47/*Msp* I, and 25X/*Taq* I RFLPs.
†Combined analysis of the MCT112/*Msp* I RFLP and the microsatellite polymorphism.
‡Combined data from present study and Chamberlain *et al.* (6).
§Using values interpolated from ref. 6.
¶26P/*Bst*XI RFLP.

reported by Chamberlain *et al.* (6), a peak lod score of 48.10 was obtained at a recombination fraction $\theta \leq 0.005$ (Table 3).

Finally, we examined the linkage relationship between the *D9S15* and *D9S5* markers in the FA families. For this purpose the MCT112/*Msp* I RFLP and the MCT112 microsatellite polymorphism were combined for *D9S15* and tested for linkage against the 26P/*Bst*XI RFLP for *D9S5*. A more complete analysis using all RFLPs at *D9S5* was not possible because of computer memory limitation. Complete linkage was found between the two marker loci, and a lod score of 26.17 was obtained at $\theta = 0.00$ (confidence interval 0.0–0.02).

Linkage Disequilibrium Between Markers and *FRDA*. As shown in Table 2, 59% (46 of 78) of FA chromosomes are associated with allele *2* (*MS2*) of the *D9S15* microsatellite, which is found in only 38% (27 of 71) of normal chromosomes. This reveals a significant allelic association ($P < 0.01$). On the other hand, we have previously shown a specific association between the minor allele *2* (*M2*) of the MCT112/*Msp* I RFLP and *FRDA* ($P < 0.01$) (7). In fact, of the FA chromosomes, 51% (40 of 78) are of haplotype *MS2 M2*, while this haplotype accounts for only 24% (17 of 71) of the normal chromosomes. Thus, it is the *MS2 M2* haplotype that is specifically associated with the *FRDA* allele ($P < 0.01$). Neither *MS* (not 2) *M2* nor *MS2 M1* shows specific association with *FRDA*. Conversely, haplotype *MS5 M1* is underrepresented on FA chromosomes compared to normal chromosomes ($P < 0.01$). None of the other haplotypes showed significant difference between normal and FA chromosomes.

Linkage disequilibrium was also investigated at *D9S5*. Forty-four percent (34 of 78) of parental chromosomes carrying the *FRDA* allele also carry allele *2* (*B2*) of the 26P/*Bst*XI RFLP, while only 31% (23 of 75) of normal chromosomes carry this allele. On the other hand, allele *3* represents 49% (37 of 75) of the normal chromosomes, but only 27% (21 of 78) of FA chromosomes (Table 1). Thus, even though there appears to be nonrandom allelic association ($P < 0.05$), the linkage disequilibrium between *D9S5* and *FRDA* alleles is less obvious than between *D9S15* and *FRDA* alleles.

We investigated the distribution of extended *D9S5–D9S15* haplotypes among normal and FA chromosomes. Such haplotypes could be inferred in 63 normal and 72 FA chromosomes. We observed important differences in haplotype distribution between FA and normal chromosomes (Table 4).

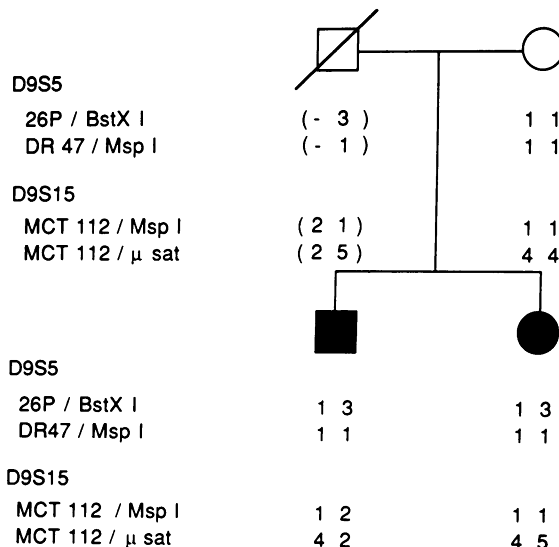

FIG. 4. Analysis of a family showing recombination between *D9S15* and *FRDA*. Segregation of alleles is shown for two RFLPs at *D9S5* (26P/*Bst*XI and DR47/*Msp* I) and for the two polymorphisms at *D9S15*. μ sat, microsatellite.

Table 4. Frequencies of extended haplotypes

Alleles			Haplotype frequency, %	
D9S5 26P *Bst*XI (*B*)	*D9S15* MCT112 *Msp* I (*M*)	*D9S15* MCT112 microsatellite (*MS*)	Normal chromosomes (n = 63)	FA chromosomes (n = 72)
1	*1*	*4*	4.7	9.7
1	*1*	Others	4.7	5.5
1	*2*	*2*	9.5	12.5
1	*2*	Others	0.0	0.0
2	*1*	*4*	4.7	8.4
2	*1*	*5*	9.5	8.4
2	*1*	Others	4.7	6.9
2	*2*	*2*	4.7*	23.6*
2	*2*	Others	4.7	0.0
3	*1*	*4*	8.0 }	4.2 }
3	*1*	*5*	22.2* } 39.7*	4.2* } 8.4*
3	*1*	Others	9.5 }	0.0 }
3	*2*	*2*	9.5	16.7
3	*2*	Others	3.2	0.0

*$P < 0.005$ that frequencies are the same on normal and FA chromosomes.

A total of 18 haplotypes was observed on normal chromosomes, while only 13 were seen on FA chromosomes. Six haplotypes account for 80% (57 of 72) of FA chromosomes, whereas they account for only 43% (27 of 63) of normal chromosomes. The most significant disease-associated haplotype is *B2 M2 MS2*, which is 5 times more frequent in FA chromosomes (23.6%) than in normal ones (4.7%) ($P < 0.005$). Haplotype *B3 M2 MS2* also shows an increased representation in FA chromosomes (16.7% versus 9.5%) but the difference does not reach significance. Conversely, the *B3 M1 MS5* haplotype is 1/5th as frequent on FA chromosomes as on normal chromosomes ($P < 0.005$). Thus the underrepresentation of allele *B3* at *D9S5* in FA chromosomes (Table 1) might be in fact the resultant of a slightly increased *B3 M2 MS2* haplotype and of an important decrease of the *B3 M1 MS5* haplotype.

DISCUSSION

The search for additional polymorphisms at both *D9S15* and *D9S5* fully succeeded in that highly informative markers have been identified in each case. The *D9S15* polymorphism belongs to a class of genetic markers, termed microsatellites (8, 9), that show variation in the number of tandem repeats of the dinucleotide CA. We have observed six alleles in the Caucasian population studied. The high informativeness of this marker is reflected by a PIC value of 0.65, which is in the range observed for similar sequences (8, 9). This polymorphism alone provides linkage information in almost all FA families, and the fact that it is very rapidly typed is an added advantage. Although there is significant linkage disequilibrium between the *Msp* I RFLP previously described for MCT112 (14) and the microsatellite polymorphism, the combination of the two polymorphisms provides additional information for genetic analysis and the PIC calculated from the haplotype frequencies reaches a value of 0.74. It has been reported that the spontaneous mutation rate to new length alleles due to unequal recombination or replication slippage at extremely variable human minisatellites can be sufficiently high to be directly measurable in human pedigrees (15). The presence of a strong linkage disequilibrium between the *Msp* I RFLP and the microsatellite polymorphism within MCT112 shows that this is not the case at this locus. The microsatellite polymorphism is thus a reliable genetic marker.

Genomic walking around the original DR47 probe at *D9S5* allowed us to characterize two RFLPs. The most useful is the *Bst*XI RFLP, since its three alleles generate a PIC of 0.55. This RFLP is due to the presence of two variable *Bst*XI sites, and one could expect four possible allelic combinations. However, the fourth genotype (3.5- + 1.0- + 0.6-kb fragments) has never been observed. When analyzed in combination with the *Msp* I or *Taq* I RFLPs detected by DR47, four haplotypes are generated and the PIC increases to 0.60. Probe 25X detects an infrequent *Taq* I RFLP. In the four instances where the rare allele was observed, it was associated with the *B1 M1* haplotype, and this polymorphism is thus useful when one parent is homozygous for this haplotype. The high informativeness due to the polymorphisms described here is well reflected in the results of our linkage analysis, since with the same family panel, maximum lod scores rose from 6.30 to 27.90 for the *D9S5–FRDA* linkage and from 14.30 to 23.51 for *D9S15–FRDA*.

One recombinant was observed with both *D9S15* polymorphisms. However, because of the family structure, this event (to our knowledge the first observed between *D9S15* and *FRDA*) will not be very useful for more precise localization of the disease gene. The recombination occurred in the germ cells of the father, who is deceased, and no other sibs are available. Thus for *D9S5*, where the two affected sibs have an identical genotype, it is not possible to distinguish between homozygosity of the father at this locus or absence of recombination between *D9S5* and *FRDA*.

The new linkage data narrows considerably (0 to 1.5 cM) the confidence interval (max lod-1) for the *D9S5–FRDA* genetic map distance. For *D9S15–FRDA* the confidence limit is inferior to 2 cM (using the combined data, Table 3). Finally, as could be expected, *D9S5* appears very highly linked to *D9S15*, with a lod score of 26.17 at $\theta = 0.00$ (support interval 0 to 2 cM). The precision on the genetic map distance between the marker loci is limited by the number of families available. However, the finding of significant linkage disequilibrium between both marker loci and *FRDA* is an indication that *D9S5* and *D9S15* are in fact less than 1 cM from *FRDA*.

We have previously reported significant ($P < 0.01$) preferential association between allele *2* of the *Msp* I RFLP at *D9S15* and *FRDA* (7). This is now reinforced by analysis of the haplotypes using the microsatellite polymorphism. Haplotype *MS2 M2* is much more frequent ($P < 0.01$) in FA chromosomes (51.3%) than on the normal parental chromosomes (23.9%). This is very likely the original haplotype on which the most frequent mutation in the French population occurred, since neither haplotype *MS2 M1* nor *MS* (not 2) *M2* shows preferential association to *FRDA*. Two other haplotypes, *MS4 M1* and *MS6 M1*, show a suggestion of positive association with *FRDA* without reaching significance, and this needs to be studied in a larger population. On the other hand, haplotype *MS5 M1* has a very significantly lower representation on FA chromosomes (12.8% versus 33.8%, $P < 0.01$).

While no linkage disequilibrium had been found previously with the *D9S5 Msp* I RFLP (7), this is no longer true when the *Bst*XI RFLP is analyzed. Allele *3* appears increased on normal chromosomes (49.3% versus 26.9%) ($P < 0.05$). In fact, this is shown even more clearly when extended haplotypes (*D9S5–D9S15*) were analyzed (Table 4). The lower frequency of allele *B3* on FA chromosomes is in fact the resultant of a striking underrepresentation of haplotype *B3 M1* (8.4% versus 39.7%) ($P < 0.005$) and more particularly of *B3 M1 MS5*, and an overrepresentation of *B3 M2 MS2*. Haplotype *B2 M2 MS2* is also much increased on the FA chromosomes (23.6% versus 4.7%) ($P < 0.005$). From these results we suggest that a mutation originally appeared on haplotype *B2 M2 MS2* and spread by recombination with *D9S5* to haplotypes *B3 M2 MS2* and *B1 M2 MS2*, which also show some preferential but less marked association with *FRDA*. On the other hand, mutation appears to be rare in the population whose *B3 M1* haplotypes were studied. That the number of mutations responsible for FA in the French population is relatively small is also indicated by the smaller number of extended haplotypes found on the FA chromosomes compared to the normal ones. The six more frequent haplotypes that account for 80% of FA chromosomes are present on only 44% of normal chromosomes. Our results show that the three loci *D9S5*, *D9S15*, and *FRDA* must be extremely close, and we have obtained evidence, based on pulsed-field gel electrophoresis, that *D9S5* and *D9S15* are within a *Not* I fragment of 450 kb (unpublished data).

The fact that linkage disequilibrium is very significant with both marker loci suggests that *FRDA* might be located between *D9S5* and *D9S15*. More extended study of recombination in this region, and analysis of extended haplotypes in other populations, especially in inbred populations that show high incidence of FA, such as in the French Canadian, Cajun, or Cypriot populations previously described (6, 16–18), should allow more precise orientation of the search for the FA gene, as such studies did for the cystic fibrosis gene (19).

We thank Drs. A. Seck and Y. Agid (Paris), Drs. C. Tommasi-Davenas and P. Trouillas (Lyon), and the Association Française pour l'Ataxie de Friedreich for invaluable contribution to the family analysis, through collection of blood samples and verification of diagnosis. We thank M. Cometto and Dr. J. C. Kaplan for establishment of lymphoblastoid cell lines. We are grateful to Drs. G. Nappi and A. P. Verri (Pavia) for providing blood samples and clinical information on the recombinant family. We thank Dr. Y. Nakamura and Dr. A. Driesel for probes MCT112 and DR47, Dr. T. Mohandas for cell line CF11.4, and A. Staub for oligonucleotide primers. This work was supported by grants from the Association Française contre la Myopathie, the Association Française de Ataxie de Friedreich, the Ministère de la Recherche (87C.0397), and the Groupe des Populaires d'Assurances.

1. Geoffroy, G., Barbeau, A., Breton, G., Lemieux, B., Aube, M., Lepert, C. & Bouchard, J. B. (1976) *Can. J. Neurol. Sci.* **3**, 279–287.
2. Harding, A. E. (1981) *Brain* **104**, 589–620.
3. Barbeau, A. (1984) *Can. J. Neurol. Sci.* **11**, 646–660.
4. Chamberlain, S., Shaw, J., Rowland, A., Wallis, J., South, S., Nakamura, Y., von Gabain, A., Farrall, M. & Williamson, R. (1988) *Nature (London)* **334**, 248–250.
5. Fujita, R., Agid, Y., Trouillas, P., Seck, A., Tommasi-Davenas, C., Driesel, A. J., Olek, K., Grzeschik, K.-H., Nakamura, Y., Mandel, J. L. & Hanauer, A. (1989) *Genomics* **4**, 110–111.
6. Chamberlain, S., Shaw, J., Wallis, J., Rowland, A., Chow, L., Farrall, M., Keats, B., Richter, A., Roy, M., Melancon, S., Deufel, T., Berciano, J. & Williamson, R. (1989) *Am. J. Hum. Genet.* **44**, 518–521.
7. Hanauer, A., Chery, M., Fujita, R., Driesel, A. J., Gilgenkrantz, S. & Mandel, J. L. (1990) *Am. J. Hum. Genet.* **46**, 133–137.
8. Litt, M. & Luty, J. A. (1989) *Am. J. Hum. Genet.* **44**, 397–401.
9. Weber, J. L. & May, P. E. (1989) *Am. J. Hum. Genet.* **44**, 388–396.
10. Lathrop, G. M., Lalouel, J. M., Julier, C. & Ott, J. (1985) *Am. J. Hum. Genet.* **37**, 482–498.
11. Heilig, R., Lemaire, C. & Mandel, J. L. (1987) *Nucleic Acids Res.* **15**, 9129–9142.
12. Murray, J. C., Shiang, R., Carlock, L. R., Smith, M. & Buetow, K. H. (1987) *Hum. Genet.* **76**, 274–277.
13. Hsu, L. C., Yoshida, A. & Mohandas, T. (1986) *Am. J. Hum. Genet.* **38**, 641–648.
14. Carlson, M., Nakamura, Y., Krapcho, K., Fujimoto, E., O'Connell, P., Leppert, M., Lathrop, G. M., Lalouel, J. M. & White, R. (1987) *Nucleic Acids Res.* **15**, 10614.
15. Jeffreys, A. J., Royle, N. J., Wilson, V. & Wong, Z. (1988) *Nature (London)* **332**, 278–281.
16. Dean, G., Chamberlain, S. & Middleton, L. (1988) *Lancet* **ii**, 587.
17. Barbeau, A., Sadibelouiz, M., Roy, M., Lemieux, B., Bouchard, J. P. & Geoffroy, G. (1984) *Can. J. Neurol. Sci.* **11**, 526–533.
18. Richter, A., Melancon, S., Farrall, M. & Chamberlain, S. (1989) *Cytogenet. Cell Genet.* **51**, 1066.
19. Kerem, B. S., Rommens, J. M., Buchanan, J. A., Markiewicz, D., Cox, T. K., Chakravarti, A., Buchwald, M. & Tsui, L. C. (1989) *Science* **245**, 1073–1080.