Hum Genet (1992) 89:531–538

Human Genetics

A 530 kb YAC contig tightly linked to the Friedreich ataxia locus contains five CpG clusters and a new highly polymorphic microsatellite

Ricardo Fujita[1,*], Giorgio Sirugo[1], Franck Duclos[1], Hadi Abderrahim[2], Denis Le Paslier[2], Daniel Cohen[2], Bernard H. Brownstein[3], David Schlessinger[2], Jean-Louis Mandel[1], and Michel Koenig[1]

[1]LGME-CNRS, U.184-INSERM, Institut de Chimie Biologique, Faculté de Médecine, Université Louis Pasteur, 11, Rue Humann, F-67085 Strasbourg Cedex, France
[2]Centre d'Etudes du Polymorphisme Humain, 27, Rue Juliette Dodu, F-75010 Paris, France
[3]Department of Genetics and Center for Genetics in Medicine, Box 8031, Washington University School of Medicine, St. Louis, MO 63110, USA

Received October 28, 1991

Summary. Friedreich ataxia (FA) is a severe autosomal recessive neurodegenerative disease. The defective gene has been previously assigned to chromosome 9q13–q21 by demonstration of tight linkage to the two independent loci D9S15 and D9S5. Linkage data indicate that FRDA is at less than 1 c*M* from both markers. Previous physical mapping has shown that probes defining D9S15 (MCT112) and D9S5 (26P) are less than 260 kb apart and are surrounded by at least six CpG clusters within 450 kb, which might indicate the presence of "candidate" genes for FA. We isolated and characterized a 530 kb YAC (yeast artificial chromosome) contig that contains five of the CpG clusters. The YACs were used to search for new polymorphic markers needed to map FRDA precisely with respect to the cloned segment. In particular, we found a $(CA)_n$ microsatellite polymorphism, GS4, that detects 13 alleles with a PIC value of 0.83 and allows the definition of haplotypes extending over 310 kb when used in combination with polymorphic markers at D9S5 and D9S15.

Introduction

Friedreich ataxia (FA) is an autosomal recessive disease that is characterized by a progressive degenerative disorder of the central and peripheral nervous system. The incidence in Caucasian populations is approximately 1 in 50000. The basic biochemical defect is unknown. Therefore, efforts to understand the disorder are directed toward isolating the disease gene on the basis of its genomic map position. A tight linkage has been found between the FA locus (FRDA) and two independent 9q13–q21 loci, D9S15 (probe MCT112) and D9S5 (probes DR47 and 26P) (Chamberlain et al. 1988, 1989; Fujita et al. 1989, 1990; Hanauer et al. 1990; Pandolfo et al. 1990). Addition of lod scores obtained from different laboratories (Hanauer et al. 1990; Chamberlain et al. 1989; Pandolfo et al. 1990; and unpublished data) gives a maximum lod value of over 61.53 between D9S15 and FRDA and over 35.69 between D9S5 and FRDA, both at a recombination fraction of 0. This indicates that FRDA is no more than 0.7 c*M* from D9S15 and 1.1 c*M* from D9S5, using the lod-1 confidence interval. Pulsed field mapping revealed that the two loci D9S15 and D9S5 are less than 450 kb apart (Fujita et al. 1990, 1991; Pandolfo et al. 1990, Wilkes et al. 1991). Taken together the data indicate that the most likely location for FRDA is in a 1.5 Mb interval centered around D9S5 and D9S15. However additional polymorphic markers spread within this interval are needed to bracket and narrow down the disease locus by genetic analysis.

The utilization of yeast artificial chromosomes (YACs) allows the isolation of several hundred kilobases in a single screening step and each clone is small enough to be a convenient starting point for subcloning. We describe here the isolation and mapping of eight YAC clones containing loci D9S5 and/or D9S15. Alu-PCR (polymerase chain reaction) fragments were amplified from YAC DNA and subcloning of the YACs into cosmids was initiated. Alu-PCR and cosmid subclones were used to identify RFLPs (restriction fragment length polymorphisms) and multiallelic microsatellite polymorphisms.

Materials and methods

YAC library screening

We screened two different human YAC libraries (Brownstein et al. 1989; Albertsen et al. 1990) by hybridization with probes DR47

* *Present address:* Department of Ophthalmology, University of Michigan, W. K. Kellogg Eye Center, 1000 Wall Street, Ann Arbor, MI 48105, USA
Correspondence to: M. Koenig

Table 1. Sequence of primers used for inter-Alu and Alu-vector polymerase chain reaction (PCR), and for GS1 and GS4 microsatellite PCR amplification

Primer	Amplified DNA	Primer sequence
NA71	Alu primer	5′-CCTCGGCCTCCCAAAGTGCTGGGATTACAG-3′
NL10	Right arm vector primer	5′-CTCCCGGGGGCGAGTCGAACGCCCG-3′
NL12	Left arm vector primer	5′-CAATTAAATACTCTCGGTAGCCAAG-3′
GS1D	$(TAAA)_6$ (151 bp)	5′-GCCTCTGCACTCCAGCCTGGGCAACAGAG-3′
GS1R		5′-GATACCTGATAGCTATATAGCTATAAGC-3′
GS4D	$(CA)_{17}$ (131 bp)	5′-GGGAAGAGCAAATTCCCTGAACCCCG-3′
GS4R		5′-CCTGGGCGACAGAGTGAGACTCG-3′

(D9S5) (Orzechowski et al. 1987) and MCT112 (D9S15) (Carlson et al. 1987), and by PCR amplification with MCT112 primers (Fujita et al. 1990). Both libraries were constructed with the pYAC4 vector and yeast strain AB1380 (Burke et al. 1987). The following YAC clones were isolated from the library of Brownstein et al. (1989): B120A12 (YAC1), P11H8 (YAC2), B206G11 (YAC8), A207F7 (YAC9), B229A12 (YAC10), and from the CEPH library (Albertsen et al. 1990): 185G8 (YAC3), 145B5 (YAC4), 253F8 (YAC5), 267A7 (YAC6), and 48D9 (YAC7).

Block preparation

Yeast cells were grown by conventional methods. Agarose blocks were prepared according to the method of Bellis et al. (1987), with the following modifications: blocks were treated with Proteinase K and 1% Sarkosyl, without a preincubation step, and were subsequently treated twice with phenylmethylsulfonyl fluoride (PMSF) at 40 µg/ml in 0.5 *M* EDTA at 42°C for 30 m to inactivate proteinase K. Blocks were rinsed twice in 0.5 *M* EDTA at 42°C for 30 m and stored at 4°C. Standard low melting point (LMP) agarose, without particular treatment, was used and was appropriate for restriction enzyme digestion. Typically, one 100 µl block was incubated with 20 to 40 units (u) of restriction enzyme in 250 µl final volume for 4 to 12 h.

Pulse field gel electrophoresis (PFGE) and blotting

DNA was fractionated by PFGE with an LKB pulsaphor with CHEF electrode array or Rotaphor type 3 apparatus, or Waltzer (Southern et al. 1987). DNA was usually run at 150 V for 24 h with pulses of 15 s. DNA was briefly depurinated in the gel and transferred on to two Hybond N^+ membranes layered as a "sandwich" on each side of the gel.

Inter-Alu and Alu-vector amplification

Yeast DNA was extracted from agarose blocks with a Geneclean kit (Bio 101, San Diego). DNA (100 ng) was amplified in a 100 µl reaction mixture containing 200 µ*M* of each dNTP, 50 m*M* KCl, 10 m*M* Tris-HCl, pH 8.2, 1.5 m*M* $MgCl_2$, 0.45 µ*M* Alu primer, 1.8 µ*M* vector primer and 2 units of *Taq* DNA polymerase (Cetus). Alu primer and pYAC4 vector primers are indicated in Table 1. Inter-Alu amplifications were performed in a Perkin-Elmer thermocycler with 35 cycles consisting of 1 m at 92°C for denaturation, 1 m at 45°C for annealing and 4 m at 72°C for elongation. Alu-vector amplifications were performed with 10 cycles of 1 m at 92°C, 1 m at 60°C and 4 m at 72°C followed by 25 cycles where annealing was performed for 30 s at 45°C. Annealing at 60°C during the first ten cycles favors priming by vector primers compared with Alu primers. Control reactions without DNA and with AB1380 yeast DNA (containing no artificial chromosomes) were performed. PCR products were analyzed by running one-tenth of the reaction on a 1% agarose gel. The rest of the amplification was subsequently fractionated on a preparative LMP agarose gel and the fragments of interest were excised. DNA samples (10 ng) were labeled by the random priming method (Feinberg and Vogelstein (1984). Most Alu-PCR probes were hybridized in the presence of human placental DNA at 100 µg/ml for competition of repetitive sequences. Probes were hybridized in 50% formamide at 42°C and washed in 1 × SSC at 60°C.

Genomic DNA and cell lines used for the chromosome 9 localization panel

The CF11.4 hybrid cell line contains an X;9 translocation (Xqter–q13::9q34–pter) as the only human component in a hamster genomic background (Hsu et al. 1986). Anly is a human-mouse hybrid cell line containing an X;9 translocation (Xqter–q13::9p24–qter) (Zonana et al. 1988). The CHPJ hybrid cell line contains an X;13 translocation (Xqter–q21::13q34–pter) (a gift from S. Gilgenkrantz, Centre Régional de Transfusion Sanguine et d'Hématologie de Nancy, France). Hamster DNA was extracted from the Chinese hamster ovary (CHO) cell line, mouse DNA was extracted from the LMTK-cell line and control human DNA was obtained from blood of a normal male and from the GM1204 cell line established from a 49 XXXXY individual (Human Genetic Mutant Cell Repository, Camden, N.J.). Digested DNAs were blotted on Hybond N^+ membranes (Amersham). Fragments present only in CF11.4, Anly and total human genomic DNA, but not in DNA of the other cell lines, were definitively assigned to chromosome 9.

Subcloning into cosmids

YAC blocks were partially digested with 36 u of *Mbo*I in a 300 µl reaction containing limiting concentrations of $MgCl_2$ (Albertsen et

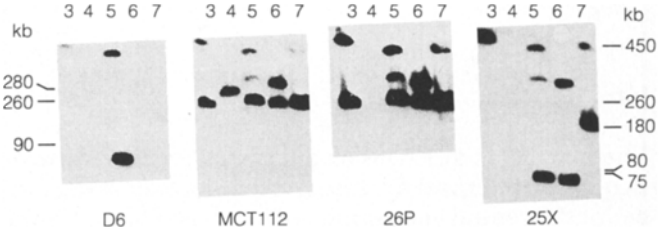

Fig. 1. Hybridization of *Not*I-digested YACs (yeast artificial chromosomes) to probes D6, MCT112, 26P and 25X. YAC number is indicated above each lane; probes used are shown below each panel and were used successively on the same blot. Size of completely digested fragments is given in kb on the left and right of the figure. Larger fragments correspond to partial digestion. YAC4 contains an abnormal-sized fragment of 280 kb hybridizing with MCT112 that is negative with 26P, indicating a rearrangement on the D9S5 side. In YAC3, probe 25X detects a 450 kb fragment instead of the predicted 190 kb fragment (Fujita et al. 1991), also indicating a rearrangement beyond D9S5

al. 1989). Conditions for partial digestion were tested on an LKB Pulsaphor in a 1% agarose, 0.5 × TAE gel with switching time of 10 s at 150 V during 17 h. These conditions offered good resolution in the 20–100 kb range. A preparative 1% Sea plaque agarose (FMC BioProducts) gel was run in the same conditions. Lambda and plasmid size markers run in parallel were ethidium bromide stained and used to identify the region of the gel that contains fragments ranging from 35 to 60 kb (Anand et al. 1989). Agarose pieces were equilibrated in 100 m*M* NaCl, melted at 67°C, cooled down at 37°C and incubated with 40 u/ml of agarase (Calbiochem) for 1 to 4 h or more to render the agarose unable to gel at low temperature. DNA was ethanol precipitated together with 1 mg/ml of glycogen as carrier. DNA (100 ng) was ligated with 200 ng of *Bam*HI-cut, dephosphorylated cosmid vector Lawrist4 (Nizetic et al. 1991) or SpCos2 (Heilig et al. 1987). Packaging was done with the Gigapack Gold kit (Stratagene). *Escherichia coli* strains ED6878 (Lawrist) and DH1 (SpCos2) were infected and spread at a density of 500 to 3000 colonies per plate onto 130 mm Hybond filters (Amersham) lying on LB agar plates with the appropriate antibiotic. Filters and replicas were treated as described by Heilig et al. (1987).

Detection of microsatellite polymorphisms

GS1 Alu-vector and EDY2 inter-Alu PCR fragments were subcloned blunt-ended in the *Sma*I site of the Bluescript plasmid vec-

Table 2. Summary of restriction fragment sizes of YACs spanning D9S5 and D9S15. *Not*I, *Mlu*I, *Bss*HII, *Eag*I and *Sac*II fragments detected with D9S5, D9S15 and YAC vector probes are listed. Sizes are indicated in kb. nt, Not tested. YACs 1 and 2 were not tested by *Mlu*I digestion. A dash (–) indicates absence of hybridization

	YAC							
Probe	1	2	3	4	5	6	7	8
*Not*I								
Right arm	170	195[a]	< 50[b]	280[b]	90	75	255	70[b]
Left arm	170	195[a]	255	40	80	255	180	90[b]
25X/DR47	170	195[a]	450[b]	–	80	75	180	nt
26P	–	195[a]	255	–	260	255	255	–
MCT112	nt	nt	255	280[b]	260	255	255	290[b]
D6	nt	nt	–	–	90	–	–	140[b]
*Mlu*I								
Right arm			190[b]	180[b]	190	235	95	400[b]
Left arm			95	140	240	95	70	120[b]
25X/DR47/26P			270	–	240	235	270	nt
MCT112			95	140	190	95	95	400[b]
D6			–	–	190	–	–	400[b]
*Bss*HII								
Right arm	130	100	< 40[b]	< 40[b]	90	35	95	70[b]
Left arm	40	55	95	40	40	95	140	< 50[b]
25X	130	55	160[b]	–	40	35	140	nt
DR47	40	40	40	–	40	40	40	nt
26P	–	100	140	–	140	140	140	–
MCT112	nt	nt	95	100	100	95	95	260[b]
*Eag*I								
Right arm	70	70[a]	< 40[b]	< 40[b]	55	75	95	70[b]
Left arm	100	95	95	≤ 10	80	95	50[a]	90[b]
25X/DR47	100	95	100	–	80	95	100	nt
26P	–	< 40[a]	160	–	160	160	160	–
MCT112	nt	nt	95	100	100	95	95	280[b]
D6	nt	nt	–	–	55	–	–	100[b]
*Sac*II								
Right arm	50[a]	100	40[b]	40[b]	90	< 20	25	< 40[b]
Left arm	40	25	25	40	< 20	25	80	< 40[b]
25X	30	30	30	–	30	30	30	nt
DR47	40	40	40	–	40	40	40	nt
26P	–	100	160	–	160	160	160	–
MCT112	nt	nt	25	30	30	25	25	30
D6	nt	nt	–	–	90	–	–	150[b]

[a] A fragment that reveals a polymorphism. YAC1 contains a *Sac*II site that is absent in overlapping YAC7. YAC2 contains an *Eag*I site and lacks a *Not*I site that are respectively absent and present in YACs 3, 5, 6 and 7. YAC7 contains an *Eag*I site that is absent in YAC1

[b] Presence of non-chromosome 9 sequences

tor (Stratagene) and sequenced. An *Eco*RI fragment containing the GS4 $(CA)_n$ repeat was cut with *Sau*3A and subcloned in the *Bam*HI site of M13 vector mp19. Positive clones were identified with a synthetic poly(dC-dA) poly(dT-dG) (Pharmacia) probe and sequenced. Primers used to amplify the GS1 and GS4 markers are described in Table 1. Amplifications were performed in the same reaction mix as for the Alu-PCR reactions described above, except that the non-Alu primers were used at a concentration of 0.2 μ*M* in 50 μl final volume. In addition, one of the two primers was end-labeled with γ^{32}P-ATP (Hanauer et al. 1990). GS1 marker was amplified through 30 cycles of 1 m at 94°C, 1 m at 50°C and 1 m at 72°C. GS4 was subjected to 30 cycles consisting of 1 m at 94°C, 1 m at 65°C and 1 m at 72°C.

Results

PFGE map of YACs containing D9S5 and D9S15 loci

The human YAC library of Brownstein et al. (1989) was screened with probes DR47 and MCT112 and the CEPH human YAC library (Albertsen et al. 1990) was screened with probe MCT112 only. Two DR47 positive clones (YACs 1 and 2) and eight MCT112 positive clones (YACs 3 to 10) were obtained. YAC size ranged from 80 to 800 kb. YACs 3, 5, 6, 7 and 9 were positive with both probes. YACs 8, 9 and 10 were isolated later and only YAC8 was further analyzed. YACs were mapped by digestion with the rare cutter enzymes *Bss*HII, *Eag*I, *Mlu*I, *Not*I and *Sac*II and hybridization with the YAC vector left and right arms, with the D9S15 probe MCT112 and with D9S5 probes 26P, DR47 and 25X (Fujita et al. 1990). Hybridization of *Not*I-digested YACs is shown in Fig. 1. Sizes of *Bss*HII, *Eag*I, *Mlu*I, *Not*I and *Sac*II fragments are given in Table 2; results are summarized on the map of Fig. 2. Four of the clones (YACs 2, 5, 6 and 7) showed complete colinearity between overlapping clones and with the map derived from cell lines and cosmids previously analyzed (Fujita et al. 1991). The other four YACs are chimeric as indicated by the presence of foreign sequences in addition to the segment from 9q13–q21. The yeast containing YAC1 has a smaller 100-kb co-transforming YAC that lacks the vector right arm and does not contain DR47 sequences (not shown). YAC3, YAC4 and YAC8 are the products of ligation or recombination between two genomic segments, one of which contains the D9S15 locus. In YAC3, the *Not*I fragment detected by probes 25X and DR47 is 450 kb in length instead of the 190 kb predicted from the genomic map (Fujita et al. 1991). YAC3 was analyzed on request by H. Dauwerse and G.J.B. Van Ommen (Leiden University, The Netherlands) by in situ hybridization of biotinylated YAC DNA on metaphase chromosomes. YAC3 indeed showed two chromosomal localizations: one on 9q13–q21 and the other on 13q22–q31 (G.J.B. Van Ommen, personal communication). YAC4 and YAC8 have an abnormal-sized *Not*I fragment (280 and 290 kb, respectively) detected with MCT112 but not with 26P while YACs 3, 5, 6 and 7 contain a 260 kb fragment common for MCT112 and 26P as predicted by the genomic map (Fujita et al. 1991). Mapping with *Bss*HII, *Eag*I and *Sac*II indicated that YAC4 was rearranged on the 26P side, while orientation of YAC8 based solely on the rare cutting enzyme site

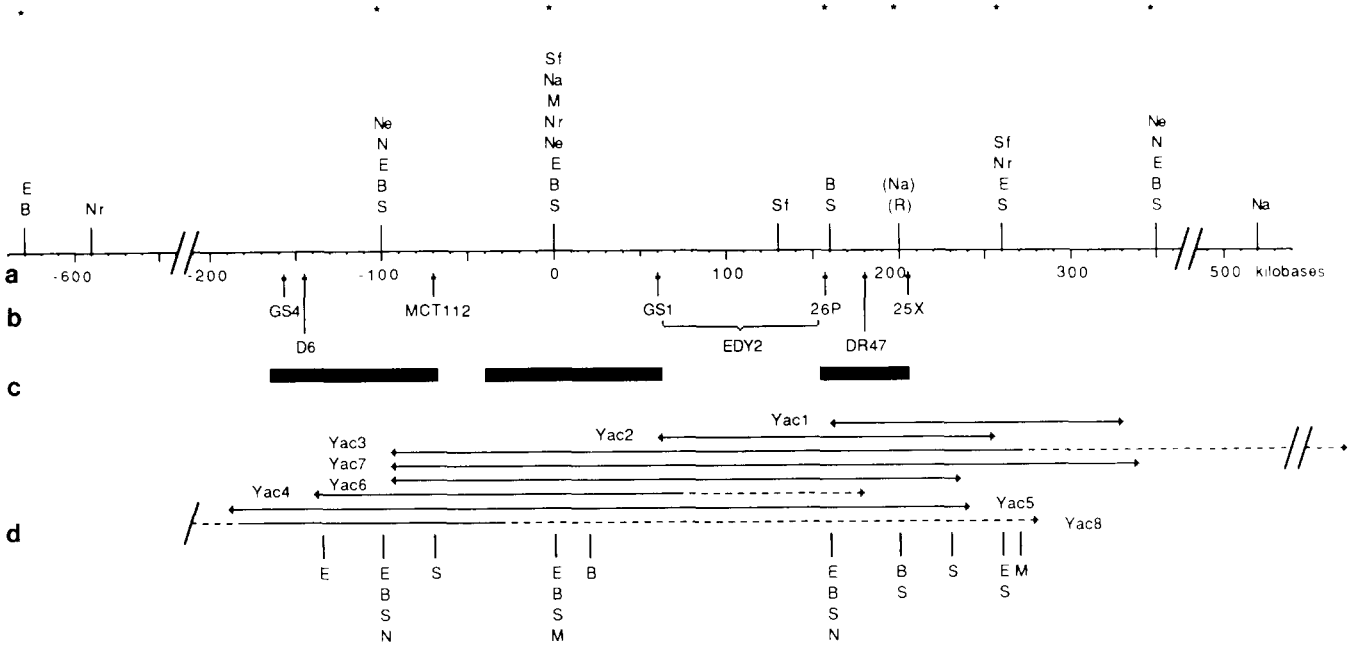

Fig. 2a–d. Map of overlapping YACs spanning D9S15 and D9S5. *Stars* at the top of the figure indicate the position of CpG clusters. **a** Reference restriction map established from cell line genomic DNA (*Nar*I and *Rsr*II sites indicated in *parentheses* were identified on cosmids (Fujita et al. 1991); the CpG island at position −630 was identified in the present study). **b** Position of markers along the genomic map; the exact position of probe EDY2 within the segment indicated by the *brace* is not known. **c** Contigs of cosmids are shown as *thick horizontal bars;* the contig containing probe DR47 is described in Fujita et al. (1991). **d** *Arrows* represent YACs isolated either with MCT112 (YACs 3–8) or with DR47 (YACs 1 and 2); *dashed lines* indicate segments that do not correspond to chromosome 9. Restriction sites shown at bottom of the figure were observed in YAC DNA. The scale is given in kb in **a**. Enzymes abbreviations: N, *Not*I; E, *Eag*I; B, *Bss*HII; S, *Sac*II; Sf, *Sfi*I; Na, *Nar*I; Ne, *Nae*I; Nr, *Nru*I, R, *Rsr*II; M, *Mlu*I

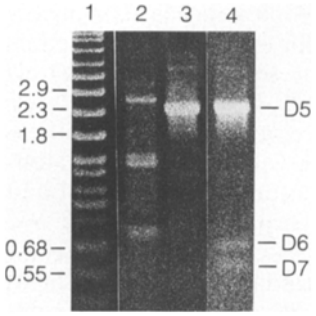

Fig. 3. Inter-Alu and Alu-vector fragments from YAC3 and YAC5. Total yeast DNA was amplified with Alu and YAC vector primer. Products were fractionated on an agarose gel and ethidium bromide stained. *Lane 1* size marker, size of five fragments is indicated in kb on the left; *lane 2* YAC3 amplified with Alu primer NA71; *lane 3* YAC5 amplified with Alu primer NA71; *lane 4* YAC5 amplified with Alu primer NA71 and vector primer NL10. D5 is an inter-Alu product present only in YAC5. D6 and D7 fragments were only amplified in presence of both Alu and vector primers

map was not possible. YAC5 contained the extremity of our YAC contig (YAC8 excluded) on the D9S15 side. In order to establish that this extremity of YAC5 was not the result of a cloning artifact, we sought to isolate the corresponding end fragment and other YAC subfragments that could be used to map YAC8.

Isolation of YAC subfragments by inter-Alu and Alu-vector PCR

We used the method of Breukel et al. (1990) to amplify the right arm extremity of YAC5 and the left arm extremity of YAC4. PFGE YAC mapping indicated that the two extremities were both on the D9S15 side of the contig. YAC DNA was amplified in two separate reactions, one containing Alu primer plus vector primer (Alu-vector) and one with Alu primer alone (inter-Alu). We obtained two fragments specific for the Alu-vector reaction with YAC5 (fragments D6 and D7 in Fig. 3) and one fragment with YAC4 (fragment K8, not shown). Fragments D6, D7 and K8 were isolated and subsequently reamplified in the presence of Alu plus vector primers; the fragments failed to reamplify in the presence of one of the two primers alone (not shown). Other fragments were obtained in inter-Alu reactions from the same YACs (see below).

Fragments D6, D7 and K8 were hybridized to the YAC PFGE blots. D7 gave a complex pattern including hybridization to yeast DNA and was not further analyzed. The D6 and K8 hybridization pattern was identical to the pattern of the corresponding YAC vector probe (left arm for YAC4 and right arm for YAC5). The smallest end-containing fragment was a 55-kb *Eag*I fragment for YAC5 and ≤ 10-kb *Eag*I fragment for YAC4, suggesting that D6 and K8 are indeed at the correct extremity of YACs 5 and 4, respectively. In addition, K8 was present in YAC5 and D6 was absent from YAC4. We demonstrated that D6 maps to chromosome 9 by hybridization to our localization panel of human-rodent hybrid cell lines (not shown).

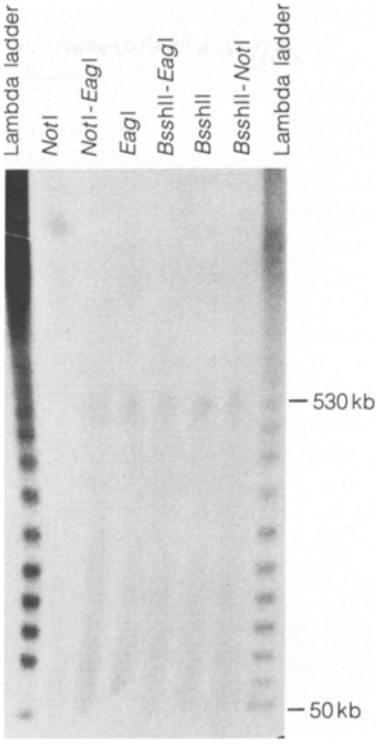

Fig. 4. Detection of a new CpG island with probe D5. D5 was hybridized to a pulsed field gel electrophoresis (PFGE) blot of rare cutter enzyme digested DNA. Enzyme or combination of enzymes used to digest CF11-4 DNA is indicated at the top of each lane. A lambda DNA fragment ladder is shown as size marker in the first and last lanes. Position of the 530-kb band is indicated

The hybridization of probe D6 to YAC8 allowed orientation of the YAC and indicated that it is rearranged on the 26P side. We sought to isolate the YAC8 extremity on the D6 side by Alu-vector PCR (Breukel et al. 1990). We obtained a specific Alu-vector fragment that failed to map at the extremity of YAC8. However, we obtained an inter-Alu fragment, H6, that localized beyond D6 on our YAC8 map. H6 was hybridized to the localization panel and, to our surprise, this fragment did not map to chromosome 9 (not shown), indicating that YAC8 contains non-chromosome 9 sequences at both ends.

Several other inter-Alu fragments were isolated from the different YACs as well as the right Alu-vector fragment of YAC2. This latter fragment, called GS1 and an inter-Alu fragment, EDY2, are localized between 26P and MCT112 and were used to search for polymorphisms (see below). An inter-Alu fragment, D5, was amplified from YAC5 but not from YAC3 (Fig. 3). Further mapping indicated that D5 is not present in YAC4 (not shown). Since probe D5 is separated from MCT112 by a CpG island and gave clear signals on blots of total genomic DNA, it was hybridized to *Bss*HII, *Eag*I and *Not*I single and double digests of CF11.4 hybrid cell line DNA (Fig. 4); this cell line contains an X;9 translocation as the only human chromosome. D5 detects a *Not*I fragment over 1 Mb in size. In single or double digests with *Bss*HII and *Eag*I, D5 detects the same 530-kb fragment, indicating the presence of a second CpG island distant from D5 and MCT112 (Fig. 2).

Subcloning of YACs into cosmids

We chose to subclone some YACs into cosmids in order to facilitate the search for polymorphisms, conserved sequences and transcribed sequences, particularly around the CpG islands. Cosmid libraries were constructed from total yeast DNA of YAC4 and YAC8 clones. These clones were chosen prior to complete knowledge of their rearranged status. YAC4 cosmids were screened with total human genomic DNA and 20 positive clones in the chromosome 9 portion of YAC4 were further analyzed. Since YAC8 turned out to contain two non-chromosome 9 fragments (see above), only six YAC8 cosmids positive for D6 were further analyzed. Overlapping cosmids were ordered on the basis of common *Eco*RI fragments and common hybridizing fragments. We defined two contigs of cosmid (96 and 102 kb in size) that were mapped with rare cutting enzymes and by probe hybridization. The position of the contigs is shown in Fig. 2c. Cosmid mapping showed that, although D6 is within the last 55 kb of the YAC5 end, it is not the extremity of YAC5. To exclude the possibility that parts of the cosmids correspond to non-chromosome 9 sequence of YAC4 or YAC8, we checked that the extremities of the cosmid contig were present on more than one YAC.

Identification of new polymorphisms

To orient the search for the FRDA gene, more genetic markers spread along the YAC contig are needed. We tested a few Alu-PCR fragments for RFLPs. Probe EDY2 detects a biallelic polymorphism on *Pvu*II and *Eco*RI digested DNA (frequency of rarer allele = 0.2). We then sought to find more informative markers based on microsatellite sequences. We looked for microsatellite sequences linked to the 3′ end of the Alu sequences (Orita et al. 1990) of our Alu-PCR fragments. We found a $(TAAA)_6$ repeat next to the Alu sequence of GS1. PCR amplification of this repeat revealed two very rare alleles in addition to the frequent allele (2 occurrences of allele 1 and 1 occurrence of allele 2 on 50 chromosomes tested). We then searched for $(CA)_n$ microsatellites. Blots of *Eco*RI digested cosmids were hybridized with a nick-translated CA repeat probe. Several positive fragments were obtained, one of which, named GS4, was immediately beyond fragment D6. We analyzed this sequence because of its position 80 and 310 kb away from the previously used polymorphic markers MCT112 and 26P, respectively. GS4 sequence was amplified by PCR from 149 independent chromosomes from 40 French families with FA. This sequence was found to be highly polymorphic with 12 alleles so far identified (Fig. 5 and Table 3) and has a PIC value of 0.83. Preliminary data suggest that this polymorphism is in linkage disequilibrium with FA (Table 3) although weaker than that found with MCT112 in the same population (Fujita et al. 1990). Other CA repeats are currently being analyzed. Amplification conditions and sequence of primers used to amplify GS1 and GS4 are given in Materials and methods.

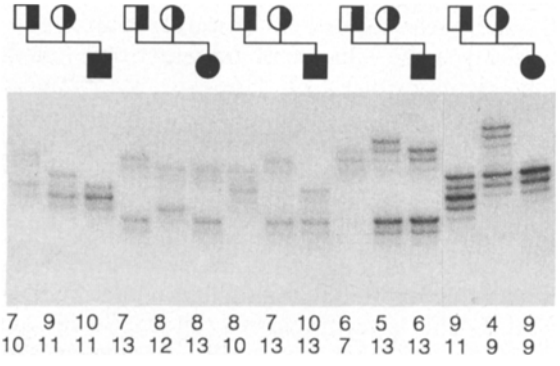

Fig. 5. Segregation of GS4 microsatellite alleles. Pedigree of individuals from Friedreich ataxia (FA) families is indicated at the top of the lanes. Allele identification is given below each lane

Discussion

Eight YAC clones positive for D9S5 or D9S15 probes were obtained from two different libraries (Brownstein et al. 1989; Albertsen et al. 1990). YAC mapping by PFGE indicated that four YAC clones (two from each library screened) were not fully colinear with the other four YACs or with the PFGE map generated from cell line DNA (Wilkes et al. 1991; Fujita et al. 1991). We sought to define the overlap between the YACs by generating YAC probes with Alu-vector and inter-Alu PCR amplification. Alu-vector amplifications performed to isolate YAC end probes (Breukel et al. 1990) were not very effective in our hands since two out of four PCR products that were specific for the Alu-vector reaction did not correspond to the YAC extremity, probably owing to unspecific annealing of the vector primer in human genomic DNA. Several Alu-PCR fragments (D5, D6 and H6) and cosmid derived probes were useful to identify and characterize the chimeric YAC clones. One clone contained two different YACs, one of which was on chro-

Table 3. Frequency of GS4 alleles. Frequencies were derived from analysis of 40 FA families. Frequencies were calculated separately for alleles associated with FA ($n = 80$) and for alleles associated with normal chromosomes ($n = 69$). Allele 3 has not been detected yet. Successive alleles differ by one CA dinucleotide (the longest observed being allele 1). * $P < 0.05$ that frequencies are the same on normal and FA chromosomes

Frequency	Allele												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Normal chromosome	0.00	0.00	0.00	0.01	0.06	0.07*	0.22	0.10	0.15	0.19*	0.03	0.03	0.15
FA chromosome	0.03	0.01	0.00	0.00	0.06	0.19*	0.15	0.11	0.14	0.05*	0.03	0.00	0.24

mosome 9. Two clones contained chimeric YACs with double inserts and one clone contained a triple insert.

The addition of all YAC sequences from chromosome 9 indicates that we have cloned ≈530 kb in overlapping YACs. Rare cutter enzyme mapping of the YACs confirms the PFGE map that we have established from cell line genomic DNA (Fujita et al. 1991). Our long range physical map is also in general agreement with the map published by Wilkes et al. (1991), but with two important differences. We found DR47 to be at position +175 kb on our map (Fig. 2), while Wilkes and colleagues mapped DR47 in a 60 kb fragment at position 290 to 350. Our data are based on the detection of a 260 kb fragment common for MCT112 and 26P both by PFGE of cell line DNA (Fujita et al. 1991) and YAC DNA (this study, Fig. 1) and on the fact that DR47 is 20 kb apart from 26P (Fujita et al. 1990). We also found a CpG island at position −630 kb on our map while Wilkes and colleagues found by partial digestion experiments no CpG islands up to position −2300 kb in the same direction. This difference might be accounted for by differences in the cell lines used as a source of DNA. However, our identification of the CpG island is not based on partial digestions but on the use of a new probe, D5, detecting a novel 530 kb *Eag*I-*Bss*HII fragment.

Five CpG clusters, previously identified within 450 kb around the D9S5 and D9S15 loci (Wilkes et al. 1991; Fujita et al. 1991), are contained in the YAC contig. The YAC clones allowed us to look for rare cutting sites that are completely methylated in cell lines and non-detectable on PFGE of total genomic DNA. New rare cutting sites were detected but they were not associated in clusters. Two methylated CpG clusters (positions 160 and 200 kb in Fig. 2) were previously identified in a cosmid walk around DR47 (Fujita et al. 1991). In the present study, two CpG islands and neighboring regions were subcloned in cosmids (CpG islands at positions 0 and −100 kb in Figure 2). Sequences around the four cloned CpG clusters are currently being used to look for transcribed sequences.

The YAC contig encompasses loci D9S5 and D9S15, which are tightly linked to the FRDA locus (Chamberlain et al. 1988; Fujita et al. 1989), and show linkage disequilibrium with the disease (Hanauer et al. 1990; Fujita et al. 1990; Pandolfo et al. 1990). Lod score calculations give a confidence interval (max lod-1) for the position of the FRDA gene of 0.7 c*M* on each side of the D9S15 marker. This interval should correspond to approximately 1.5 Mb, one-third of which is now cloned in YACs. In order to search for possible recombination and variation in linkage disequilibrium with FRDA, we isolated new highly polymorphic markers. Several $(CA)_n$ microsatellite repeats that are potentially polymorphic have been identified. One of them, GS4, was further characterized because of its position outside of the interval defined by markers MCT112 and 26P. GS4 is highly polymorphic (PIC = 0.83) and was found in linkage equilibrium with the neighboring microsatellite and *Msp*I polymorphisms at MCT112, which are 80 kb away, in both normal and FA chromosomes (G.S., manuscript in preparation). Another polymorphic microsatellite, located between MCT112 and 26P, has been isolated by direct YAC subcloning (G.S., manuscript in preparation). Markers from GS4 to 26P are now used to define haplotypes over a 310 kb region, especially in populations where a founder effect is suspected to account for a high incidence of FA. In a preliminary analysis of Louisiana-Acadian patients, markers described in this paper were used to identify a major haplotype in strong linkage disequilibrium with the disease (Sirugo et al. 1991). It was possible to infer that some minor haplotypes were likely to have derived from the major one by ancient recombination events, suggesting a localization of the FA gene on the D9S15–D9S5 side of GS4. This could fit with the decreased linkage disequilibrium (compared with D9S15; Fujita et al. 1990) that we observed with GS4 in the French FA families.

Acknowledgements. We wish to thank Dr. A. Hanauer for advice and discussion, B. Boulay and J.-M. Lafontaine for photogrpahic work and A. Staub for synthesis of primers. We also wish to thank Drs. H. Dauwerse and G. J. B. van Ommen for in situ hybridization with YAC DNA. This work was supported by grants from the Association Française contre les Myopathies (AFM), the Ministère de la Recherche, the CNRS and INSERM and by fellowships from AFM to R. F. and G. S.

References

Albertsen HM, Le Paslier D, Abderrahim H, Dausset J, Cann H, Cohen D (1989) Improved control of partial DNA restriction in agarose using limiting concentrations of Mg^{++}. Nucleic Acids Res 17:808

Albertsen HM, Abderahim H, Cann H, Dausset J, Le Paslier D, Cohen D (1990) Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents. Proc Natl Acad Sci USA 87:456–460

Anand R, Villasante A, Tyler-Smith C (1989) Construction of yeast artificial chromosome libraries with large inserts using fractionation by pulse-field gel electrophoresis. Nucleic Acids Res 17:3425–3433

Bellis M, Pages M, Royzes G (1987) A simple method for preparing yeast chromosomes for pulse field gel electrophoresis. Nucleic Acids Res 15:6749

Breukel C, Wijnen J, Tops C, Klift H vd, Dauwerse H, Khan M (1990) Vector-Alu PCR: a rapid step in mapping cosmids and YACs. Nucleic Acids Res 18:3097

Brownstein BH, Silverman GA, Little RD, Burke DT, Korsmeyer SJ, Schlessinger D, Olson MV (1989) Isolation of single-copy human genes from a library of yeast artificial chromosome clones. Science 236:806–812

Burke DT, Carle GF, Olson MV (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science 236:806–812

Carlson M, Nakamura Y, Krapcho K, Fujimoto E, O'Connell P, Leppert M, Lathrop GM, Lalouel JM, White R (1987) Isolation and mapping of a polymorphic DNA sequence pMCT112 on chromosome 9q (D9S15). Nucleic Acids Res 15:10614

Chamberlain S, Shaw J, Rowland A, Wallis J, South S, Nakamura Y, von Gabain AS, Farral M, Williamson R (1988) Mapping of mutation causing Friedreich ataxia to human chromosome 9. Nature 334:248–250

Chamberlain S, Shaw J, Wallis J, Rowland A, Chow L, Farral M, Keats B, Richter A, Roy M, Melançon S, Deufel T, Berciano J, Williamson R (1989) Genetic homogeneity at the Friedreich ataxia locus on chromosome 9. Am J Hum Genet 44:518–521

Feinberg A, Vogelstein B (1984) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 137:266–267

Fujita R, Agid Y, Trouillas P, Seck A, Tommasi-Davenas C, Driesel AJ, Olek K, Grzeschik KH, Nakamura Y, Mandel JL, Hanauer A (1989) Confirmation of linkage of Friedreich Ataxia to chromosome 9 and identification of a new closely linked marker. Genomics 4:110–111

Fujita R, Hanauer A, Sirugo G, Heilig R, Mandel JL (1990) Additional polymorphisms at marker loci D9S5 and D9S15 generate extended haplotypes in linkage disequilibrium with Friedreich ataxia. Proc Natl Acad Sci USA 87:1796–1800

Fujita R, Hanauer A, Vincent A, Mandel JL, Kœnig M (1991) Physical mapping of two loci (D9S5 and D9S15) tightly linked to Friedreich ataxia locus (FRDA) and identification of nearby CpG islands by pulse field gel electrophoresis. Genomics 10: 915–920

Hanauer A, Chery M, Fujita R, Driesel AJ, Gilgenkrantz S, Mandel JL (1990) The Friedreich ataxia gene is assigned to chromosome 9q13–q21 by mapping of tightly linked markers and shows linkage disequilibrium with D9S15. Am J Hum Genet 46:133–137

Heilig R, Lemaire C, Mandel JL (1987). A 230 kb cosmid walk in Duchenne muscular dystrophy gene: detection of a conserved sequence and a possible deletion prone region. Nucleic Acids Res 15:9129–9142

Hsu LC, Yoshida A, Mohandas T (1986) Chromosomal assignment of the genes for human aldehyde dehydrogenase-1 and aldehyde dehydrogenase-2. Am J Hum Genet 38:641–648

Nizetic DB, Zehetner G, Monaco AP, Gellen L, Young BD, Lehrach H (1991) Construction, arraying and high density screening of large insert libraries of the human X and 21 chromosomes, their potential use as reference libraries. Proc Natl Acad Sci USA 88:3233–3237

Orita M, Sekiya T, Hayashi K (1990) DNA sequence polymorphisms in Alu repeats. Genomics 8:271–278

Orzechowski HD, Henning J, Winter P, Grzeschik KH, Olek K, Driesel AJ (1987) A human single copy DNA probe (DR47) detects a TaqI RFLP on chromosome 9 (D9S5). Nucleic Acids Res 15:6310

Pandolfo M, Sirugo G, Antonelli A, Weitnauer L, Ferretti L, Leone M, Dones I, Cerino A, Fujita R, Hanauer A, Mandel JL, diDonato S (1990) Friedreich ataxia in Italian families: genetic homogeneity and linkage disequilibrium with the marker loci D9S5 and D9S15. Am J Hum Genet 47:228–235

Sirugo G, Keats B, Fujita R, Duclos F, Purohit K, Kœnig M, Mandel JL (1991) Friedreich ataxia in Louisiana-Acadians: demonstration of a founder effect by analysis of microsatellite generated extended haplotypes. Am J Hum Genet 50:559–566

Southern EM, Anand R, Brown WRA, Fletscher DS (1987) A model for the separation of large DNA molecules by crossed field gel electrophoresis. Nucleic Acids Res 15:5925–5943

Wilkes D, Shaw J, Anand R, Riley J, Winter P, Wallis J, Driesel AG, Williamson R, Chamberlain S (1991) Identification of CpG islands in a physical map encompassing the Friedreich's ataxia locus. Genomics 9:90–95

Zonana J, Roberts SH, Thomas NST, Harper PS (1988) Recognition and reanalysis of a cell line from a manifesting female with linked hypohidrotic ectodermal dysplasia and an X:autosome balanced translocation. J Med Genet 25:383–386