

# Isolation of a GCC repeat showing expansion in FRAXF, a fragile site distal to FRAXA and FRAXE

Julia E. Parrish<sup>1</sup>, Ben A. Oostra<sup>2</sup>, Annemieke J.M.H. Verkerk<sup>2</sup>, C. Sue Richards<sup>1</sup>, James Reynolds<sup>3</sup>, Aimee S. Spikes<sup>1</sup>, Lisa G. Shaffer<sup>1</sup> & David L. Nelson<sup>1</sup>

Three folate-sensitive fragile sites, termed FRAXA, FRAXE and FRAXF, have been identified on the distal end of chromosome Xq. The first two contain expanded, hypermethylated and unstable CGG (or GCC) repeats within CpG islands. We now report the isolation of similar sequences responsible for the third fragile site, FRAXF. A 5-kilobase *EcoRI* fragment derived from a cosmid coincident with the cytogenetic anomaly detects expanded, methylated and unstable sequences in five individuals who exhibit fragile sites in distal Xq; these individuals have normal repeat lengths at both FRAXA and FRAXE. By sequence analysis, the expanded region contains a GCC repeat. PCR and sequence analysis of chromosomes from the general population indicates that the repeat is polymorphic (6 to 29 triplets), and is stable upon transmission.

<sup>1</sup>Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas 77030, USA  
<sup>2</sup>Department of Clinical Genetics, Erasmus University, Rotterdam, The Netherlands  
<sup>3</sup>Shodair Hospital, Helena, Montana 59604, USA

Correspondence should be addressed D.L.N.

Chromosomal fragile sites are relatively common in humans, with over 100 described to date (for review see ref. 1). In nearly every case, the presence of the fragile site is not associated with any phenotypic abnormality. A striking exception is found in fragile X syndrome, where a rare folate-sensitive fragile site in Xq27.3 (FRAXA) is associated with mental retardation (for review see ref. 2). FRAXA has been extensively characterized<sup>3</sup> and was the first of the numerous trinucleotide repeat mutations associated with human genetic disease. The CGG repeat at FRAXA is present in the 5' untranslated region of *FMR1*, a gene whose expression is extinguished by expansion and methylation of the repeat<sup>4,5</sup>, resulting in the disease phenotype. A second fragile site involved with mental retardation, FRAXE, lies ~600 kilobases (kb) distal to FRAXA and was recently characterized<sup>6</sup>. This fragile site was confused with FRAXA until molecular probes became available; the phenotype associated with FRAXE involves mental impairment as well, but it is much milder than that seen with expansions at the FRAXA locus (Hamel *et al.*, manuscript submitted). As with FRAXA, chromosomal fragility appears to be due to CGG (GCC) repeat expansion<sup>6</sup>.

With the advent of DNA-based probes for FRAXA and FRAXE, it became apparent that in some families, fragile sites in Xq27.3–q28 were not associated with expanded repeats at FRAXA or FRAXE. The presence of another fragile site distal to FRAXE was demonstrated by fluorescence *in situ* hybridization (FISH) analysis using cosmid probes close to the iduronate sulphatase gene in Xq28, 600 kb distal to FRAXE<sup>7</sup>. This probe was found to be proximal to the cytogenetic anomaly in the family studied. This third fragile site was designated FRAXF.

Unlike FRAXA and FRAXE, the inheritance of the fragile site at the FRAXF locus did not appear to correlate well with the presence of mental impairment in this pedigree.

We have isolated and characterized a DNA fragment which appears to represent the FRAXF locus. This fragment lies within the region of Xq28 delineated by Hirst *et al.*<sup>7</sup> as containing the FRAXF site, and is found by FISH to coincide with the fragile site. In two families, we demonstrate expansion of a GCC repeat which directly correlates with expression of the fragile site in each cytogenetically positive individual. The sites for three different methylation-sensitive restriction enzymes can no longer be cleaved in DNAs carrying expanded versions of the repeat. We also demonstrate meiotic stability of alleles in the normal size range (6–29 repeats), and one instance of mitotic instability in a 29 repeat allele.

## Cloning of the FRAXF fragile site

In an effort to identify additional *FMR1* cDNA clones, a human fetal brain cDNA library (Clontech) was screened using a fragment containing part of the 5' end of *FMR1*, including the CGG repeat. A clone (H10) containing a 1.4 kb insert was isolated and by restriction map comparison with the *FMR1* cDNA and genomic DNA, found to differ from *FMR1*. FISH analysis showed localization to Xq28, with additional weaker X-specific signals over most of the chromosome (data not shown). In an effort to characterize this cDNA, a cosmid library constructed from flow-sorted X chromosomes was screened, and 13 cosmids identified, seven of which localize to Xq28. Two nearly identical cosmids, 23H10 and 51A8, were assigned to the interval between *IDS* and *GABRA3* by hybridization to somatic cell hybrids (data not shown). Since this is the region to



**Fig. 1a** Restriction map of the 5 kb *EcoRI* fragment (pRI5.0) cloned from cosmid 23H10. The hatched box represents the sequenced region. The GCC repeat lies just 3' of the *SacII* site. The 5 kb *EcoRI* fragment was gel-isolated and ligated into pBluescript SK(-) for restriction mapping. **b**, Sequence through the repeat region. The GCC repeat is boxed. Primer positions are underlined.

**b**

```

CCGTGGTCGG GCAGGGGCTG AAAGCACCG GCACGGGAGG AGGAAGCGGG      50
CGGGCGCCGA GGCCGACTGT TTTGCCTGGG GACCGCTTGC ACCCCGAGGA      100
GGCTCGGGCA GGCGCCGGG TTCCTCGGGC TGCAGCATCT CGCCCGCCGT      150
GCCTCCCCGG AGCCGAACAC CAGCCCGCGC CCGAGCCCGC AGCGCGGACT      200
CCCAGGGGCG CCAACGACGC CGCCTCACCT CGGGTTGAAG TCCTGGAAGA      250
GGCCCTCAG GTTCATGGCG GAGAACTTCA CCGCGGCGTC CTCCTCCTCC      300
TCCCCGCAC CCGTGCTGC A CAGCCTGCG CCTTACAGCG GGTTTCATGGC      350
GCCAGCGCCA GCCCGTCCA CGCTGCTGCT CCGCTACTG CTGCCGTCCC      400
CGCTGCCGTC GCCGTCGGG TC GCCGCGCGC CGCCGCGCC GCCGCC G GA      450
GAAACCTGAG CCACCGCCCC CTGCCCTT CC TTCCGGGCTT CCGTACGAGG      500
GGCCCGCAT GCGTCCGAG CCCC GCCAGA GCGCTCCTCG CTGGGAGGGT      550
CCCCATCCTT GTGTTGCGAC GCGACCCG      578
    
```

which FRAXF had been localized<sup>7</sup>, the H10-positive 5 kb *EcoRI* fragment of cosmid 23H10 was pursued as a possible candidate for FRAXF.

A restriction map of the clone, pRI5.0, is shown in Fig. 1a. The clone contains sites for the methylation-sensitive enzymes *EagI*, *SacI* and *BssHIII*; it also contains several *SmaI* sites (not shown). Hybridization to H10 was localized to the region between the *SacII* and *BssHIII* sites.

**Sequence analysis**

Two subclones were constructed in order to facilitate sequencing, one deleting to the *EagI* site, the other to the *SacII* site from the corresponding sites within the vector polylinker. Sequencing reactions were performed on both clones using the M13 reverse sequencing primer; the two runs overlap, yielding 578 base pairs (bp) of contiguous sequence (Fig. 1b). A triplet repeat of (GCC)<sub>8</sub> is present, and is flanked by a hexamer repeat (GCCGTC)<sub>3</sub>. The GCC-repeat portion of the sequence shows homology to FRAXA and to other CGG repeat-containing clones in GenBank (release 83.0); however, there were no significant similarities found outside the repeat. Comparison of the FRAXF sequence of the H10 cDNA reveals that the cross-hybridization is most likely due to two GCC-repeat motifs in the cDNA; there were no significant similarities beyond these repeats. The FRAXF sequence contains open reading frames in both orientations, one of which extends the

entire length of the sequenced region. The translation of this open reading frame produces a peptide which has similarity to the mouse homeobox protein *engrailed-1*; the engrailed protein has a polyalanine which is responsible for the highest similarity (high score = 80), but there is also significant similarity 5' of this region.

**Confirmation of map location**

PCR primers (see Fig. 1b) were designed to amplify a 156 bp segment including the GCC repeat. In order to confirm that the amplification produced by this primer pair is specific to the appropriate region of Xq28, we assayed several somatic cell hybrids (Fig. 2). Products of the appropriate size were produced from total human DNA, Y751BM1 (a hybrid cell line containing an intact X chromosome as its only human component) and hybrid cell lines Y162Aza and Q1N. Hamster and mouse DNAs did not amplify; hybrids micro 21D and CY34 were also negative. These results confirm that the PCR product is specific to a locus between *IDS* and *DXS305*.

**Fluorescence in situ hybridization**

In order to demonstrate unequivocally that the cloned DNA anomaly represents the cytogenetically observed fragile site, we performed FISH analysis. A cosmid from the *IDS* region, 112G10, was used to probe induced metaphase chromosomes derived from an individual expressing FRAXA, and from the proband in family 1 (see below). This probe clearly hybridized to the portion of the chromosome distal to the FRAXA site, and was clearly proximal to the fragile site expressed in the proband's chromosomes (data not shown). In order to generate further probe material, the ends of the original two cosmids were used to rescreen the cosmid library. Two of the cosmids identified, 13F1 and 82E3, overlapped in the region of the GCC repeat and extended 20–30 kb in each direction (data not shown). These two cosmids were used together to probe induced metaphase chromosomes derived from the proband in family 1. Hybridization signal is visible on both sides of the fragile site (Fig. 3), demonstrating that these cosmids span the fragile site, and suggesting that the GCC repeat is expanded in FRAXF.

**Family 1**

We obtained a DNA sample from a mentally retarded male (III-4, Fig. 4) who expresses a fragile site in Xq28 in 23% of metaphases, without repeat expansion at FRAXA or FRAXE (see Methodology for details). We also obtained samples from the mother of the proband (II-7) and from his full brother (III-5) who is also mentally retarded but exhibits no fragile site. The proband and his brother have opposite maternal alleles at the FRAXA CGG repeat, both of which are within the normal range.

Results of Southern analysis indicate that the proband (III-4) exhibits both expansion and methylation when a fragment containing the GCC repeat is used as a probe.

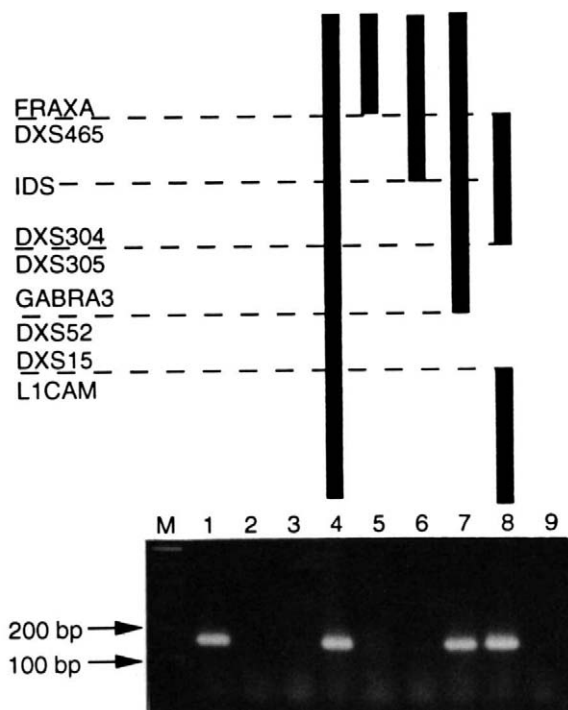


Fig. 2 PCR assay using FRAXF primers on a somatic cell hybrid mapping panel. Template DNAs are: lane 1, total human; lane 2, hamster; lane 3, mouse; lane 4, Y751BM1; lane 5, micro 21D; lane 6, CY34; lane 7, Y162aza; lane 8, Q1N; lane 9, no DNA. Approximate positions of the breakpoints in the hybrid cell lines are shown above with the nearest flanking markers<sup>17</sup>.

Figure 5a demonstrates that the *EagI* and *BssHIII* sites are completely methylated in the expanded alleles of both the proband and his mother. Figure 5b shows a Southern blot designed to test methylation at the *SacII* site and to accentuate the size differences between the normal and expanded alleles. The site is methylated in the expanded alleles; furthermore, the proband's allele is about 200 bp larger than his mother's expanded allele. III-5 shows the normal male digestion pattern, indicating receipt of his mother's normal allele. The brothers also have opposite maternal alleles at *DXS52* (data not shown).

Further analysis of this family revealed that the four maternal uncles of the proband (II-9-12), three of whom are mentally retarded (see Fig. 4), are cytogenetically fragile X negative. The grandmother of the proband (I-3) was also tested, and is cytogenetically negative as well. DNAs from these five individuals were analysed by Southern hybridization, and all show the normal patterns, with no expansion or unusual methylation (data not shown).

### Family 2

We obtained DNA samples from four members of another family described as having a fragile site in Xq27.3-q28 with no phenotypic abnormalities<sup>8</sup>. Four cytogenetically positive individuals were tested, a female, her brother and her two sons. All are phenotypically normal, although one male was described as a "plodder" in school<sup>8</sup>. Southern analysis reveals that individual II-2 (the female) has a normal allele and an expanded allele, while all three male samples (II-3, III-1 and III-2) exhibit methylated alleles

which are expanded by about 1.6 kb (Fig. 5c and data not shown).

### Variability due to changes in GCC repeat length

Southern analysis of 50 unrelated males showed variation in the size of the *SacII/BamHI* fragment (data not shown). PCR was performed on selected samples to determine whether the variation was occurring at the GCC repeat or elsewhere within the 800 bp fragment. In every case, the size of the PCR product directly correlated with the size of the fragment seen on the Southern blot. In order to exclude the possibility of variation at another point within the amplified segment, alleles of several apparent repeat lengths were cloned and sequenced, using the forward PCR primer as a sequencing primer. There was no variation in any of the cloned products outside of the GCC repeat; furthermore, the differences in the number of repeats were sufficient to account for the size differences between PCR products.

### Analysis of the site as a genetic marker

In order to determine the utility of the GCC repeat as a genetic marker, we analysed the repeat length in the parents of the CEPH collection of families (Table 1). The most common allele sizes are six and eight repeats, with a clustering of allele sizes distributed between 15-18 repeats. A total of 37 female samples produced products; of these, 29 produced two distinct alleles. Thus heterozygosity of this polymorphism is ~78%, in spite of the limited number of alleles and the dominance of a few.

Four CEPH families were analysed further to investigate mendelian inheritance of alleles and meiotic stability of the repeat. Family 12 stably segregates three different alleles of average size (8, 13 and 15), and family 1420 shows stability of a large (18 repeat) paternal allele (data not shown). Results from family 1333 show meiotic stability of maternal alleles of 15 and 18 repeats. Results

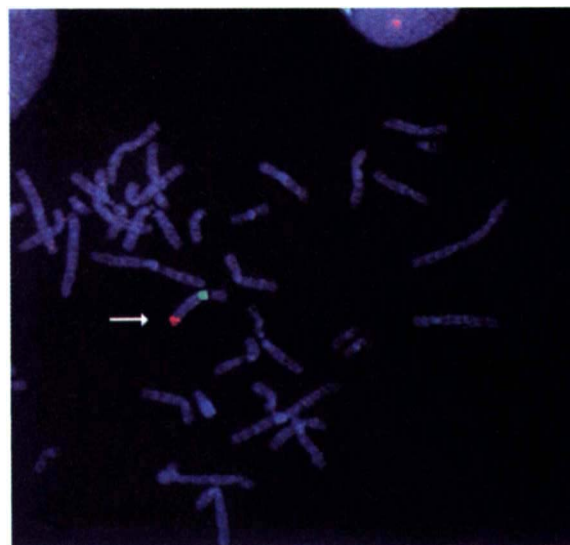


Fig. 3 Two-colour fluorescence *in situ* hybridization. The alpha-satellite for the X chromosome was labelled in biotin and detected with avidin-FITC (green) and cosmids 13F1 and 82E3 were labelled with digoxigenin and detected with anti-digoxigenin-rhodamine (red). Arrow indicates a hybridization pattern consistent with both proximal and distal hybridization of the cosmids with respect to the fragile site.

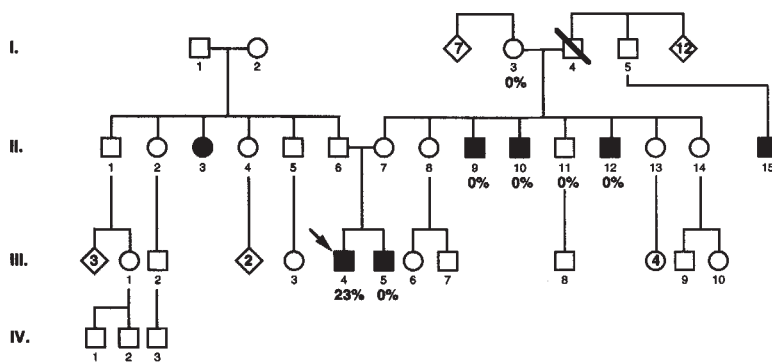


Fig. 4 Pedigree of Family 1. The proband is indicated by an arrow. Filled symbols represent mentally retarded individuals. Individuals tested for cytotenetic fragile X expression are labelled with the percentage expression observed.

from family 1424 are shown in Fig. 6. This family segregates the largest normal sized allele identified so far, a 29 repeat maternal allele. This allele was transmitted stably to five of the children, but one male child demonstrates two alleles (29 and 30 repeats).

**Discussion**

Based on its behaviour in two families segregating an Xq28 fragile site distinct from FRAXA and FRAXE, and on the *in situ* hybridization results, we conclude that the GCC repeat found in cosmid 23H10 is the molecular basis for the FRAXF site. The repeat is similar in all aspects to the trinucleotide repeats associated with the other three folate-sensitive fragile sites described at the molecular level<sup>3,6,9</sup>. It is the same repeat sequence as is found in the other sites (GCC or CGG) and demonstrates similar levels of expansion and methylation on chromosomes exhibiting a fragile site. It is polymorphic in the general population, and subject to unstable inheritance upon transmission of alleles in the size range associated with a fragile site. Due to the limited availability of families showing expansion at this locus, it is as yet unknown whether sizes intermediate between the 6–29 repeat range found in the general population and the 300–500 repeat size found in fragile site-containing chromosomes might show unstable transmission as in FRAXA and FRAXE, or whether expansions to the large version occur in a single large jump and might be constrained to transmissions from one sex or the other<sup>3</sup>.

In apparent contrast to its neighbouring fragile sites, FRAXF does not contribute to an obvious phenotype in males carrying the expanded and methylated allele. The proband in family 1 was diagnosed with fragile X syndrome due to the presence of mental retardation and a cytogenetically visible fragile site at Xq27.3–q28. While the presence of normal alleles at FRAXA and at FRAXE strongly suggested that his mental impairment was not a result of fragile X syndrome, the presence of an expanded and methylated repeat at FRAXF suggested that the mental retardation segregating in an X-linked fashion in his family could be due to effects of this mutation on a nearby gene. However, the finding that his similarly impaired brother has inherited the opposite maternal alleles at DXS548, FRAXA, FRAXE, FRAXF and DXS52 eliminates this region of the X chromosome as a candidate for any shared, X-linked mutation involved in mental impairment

in these brothers. Furthermore, the proband's cognitively impaired maternal uncles have normal alleles at FRAXF and no fragile sites. Family 2 contains three cognitively normal males with cytogenetic expression of the fragile site<sup>8</sup>, expansion of the repeats and complete methylation. Indeed, lack of direct correlation of the cytogenetic fragile site with mental retardation was noted in the original report of FRAXF<sup>7</sup>. Thus, FRAXF is likely to be more typical of chromosomal fragile sites, and like FRA16A<sup>9</sup>, may not disturb gene expression in any phenotypically observable manner. On the other hand, it seems likely that any genes in the vicinity of FRAXF would be transcriptionally repressed in an individual expressing the fragile site; attempts to define nearby genes are ongoing to allow this hypothesis to be tested. Sequence comparison and Southern analysis indicate that the H10 cDNA is not transcribed from this location.

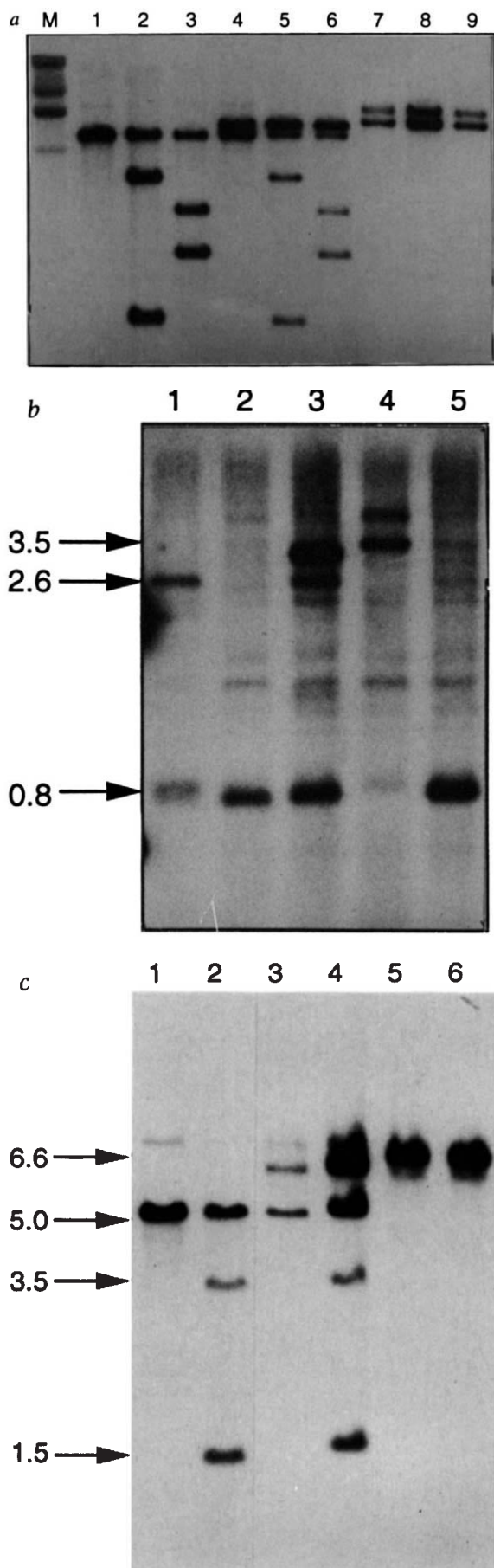
Several of the CEPH parents failed to amplify with the FRAXF primers; these appear to be true null alleles, as they consistently failed to amplify with either *Taq* or *Pfu*

**Table 1 FRAXF allele sizes in CEPH parents**

| Fathers: |             | Mothers: |              |
|----------|-------------|----------|--------------|
| ID#      | allele size | ID#      | allele sizes |
| 201      | 8           | 202      | 6, 8         |
| 1201     | 15          | 1202     | 8, 13        |
| 1701     | 15          | 1702     | 8, 16        |
| 2101     | 15          | 2102     | 6, 8         |
| 2301     | 8           | 2302     | 6, 8         |
| 2801     | 8           | 2802     | 6, 8         |
| 3501     | 8           | 3502     | 8, 8         |
| 3701     | 16          | 3702     | 6, 16        |
| 4501     | NAa         | 4502     | NA           |
| 6601     | 6           | 6602     | NA           |
| 10201    | 8           | 10202    | 8, 8         |
| 10401    | 8           | 10402    | 8, 8         |
| 88401    | 17          | 88402    | 8, 15        |
| 133101   | 8           | 133102   | 6, 16        |
| 133201   | 6           | 133202   | NA           |
| 133301   | 6           | 133302   | 15, 18       |
| 133401   | 6           | 133402   | 6, 8         |
| 134001   | 6           | 134002   | 16, 16       |
| 134101   | 8           | 134102   | 6, 8         |
| 134401   | 15          | 134402   | 8, 15        |
| 134501   | 8           | 134502   | 8, 8         |
| 134601   | 8           | 134602   | 8, 16        |
| 134701   | 16          | 134702   | 6, 8         |
| 134901   | 8           | 134902   | 16, 16       |
| 135001   | NA          | 135002   | 8, 17        |
| 136201   | 8           | 136202   | 8, 18        |
| 137501   | 8           | 137502   | 6, 8         |
| 137701   | 8           | 137702   | 8, 13        |
| 140801   | 13          | 140802   | 6, 8         |
| 141301   | 6           | 141302   | 6, 8         |
| 141601   | 6           | 141602   | 8, 16        |
| 141801   | NDb         | 141802   | 6, 8         |
| 142001   | 18          | 142002   | 16, 16       |
| 142101   | 6           | 142102   | 6, 8         |
| 142301   | 8           | 142302   | 15, 18       |
| 142401   | 8           | 142402   | 8, 29        |
| 1329101  | 6           | 1329102  | 8, 8         |
| 1329201  | 8           | 1329202  | 6, 16        |
| 1329301  | 6           | 1329302  | 8, 12        |
| 1329401  | 8           | 1329402  | 6, 8         |

<sup>a</sup>NA, no amplification.

<sup>b</sup>ND, not done.

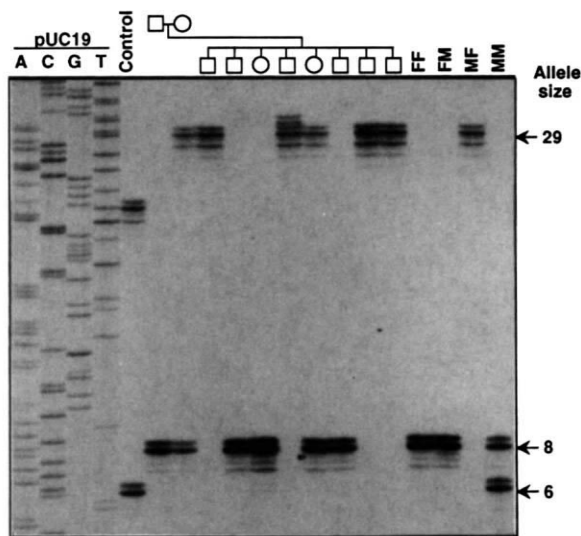


polymerase. These may represent alleles which are too large to amplify across; alternatively, they may represent point mutations in the sequences complementary to the primers, microdeletions at this locus, or low-quality DNA preparations. Southern blot analysis of these individuals should allow these possibilities to be distinguished. Identification of more families with large alleles will be helpful in determining the mechanisms of instability of this repeat, although the frequency of lengths sufficient to generate a fragile site is expected to be very low.

Mitotic instability was found in a 29-repeat allele in CEPH family 1424. This manifested in an increase from 29 to 30 repeats, likely through a replication slippage process<sup>10</sup>. As the DNAs that make up the CEPH collection are derived from long-term cultures of lymphoblastoid cell lines, this alteration may have occurred during propagation of the cell line, rather than having been present in this individual's blood leukocytes. In either case, this finding suggests that alleles of this size are approaching the threshold of instability. This is consistent with the upper limits of the normal ranges of all trinucleotide repeats described so far, and in particular is reminiscent of FRAXA alleles, where marked instability is found when the number of uninterrupted CGG repeats exceeds 33 (ref. 11).

Identification of FRAXF represents the fourth folate-sensitive fragile site to be described at the molecular level.

Fig. 5 Southern analysis of family 1 (a and b) and of family 2 (c). a, DNAs were digested with *EcoRI* alone (lanes 1, 4 and 7) or in combination with either *EagI* (lanes 2, 5 and 8) or *BssHII* (lanes 3, 6 and 9). The filter was probed with the 5 kb *EcoRI* fragment (see Fig. 1a). In normal DNA, digestion with *EcoRI* yields a 5 kb genomic fragment (lane 1). *EagI* digestion reduces this to 1.5 and 3.5 kb fragments (lane 2), while *BssHII* digests it into 2.2 and 2.8 kb fragments (lane 3). Normal female DNA (lanes 1–3) shows all possible fragments due to methylation of the inactive X chromosome. The mother of the proband (lane 4) exhibits the normal 5 kb *EcoRI* fragment as well as a larger fragment. Digestion with *EagI* (lane 5) and *BssHII* (lane 6) produces fragments of the expected sizes. DNA from patient 1 (lanes 7–9) shows an identical pattern with all three digests, indicating no digestion at the *EagI* or *BssHII* sites. He has only an expanded version of this locus, and appears mosaic for additional larger products. Lane M contains  $\lambda$  DNA digested with *HindIII*. b, DNAs were digested with *BamHI*, *EcoRI* and *SacII*. The filter was probed with the 0.8 kb *BamHI/SacII* fragment (see Fig. 1a). Cleavage at the *SacII* site cuts the 2.6 kb *BamHI/EcoRI* fragment into 0.8 and 1.8 kb fragments. In normal female DNA (lane 1), both the 0.8 kb and the 2.6 kb fragments are visible. In normal male DNA (lane 2), only the 0.8 kb fragment is visible, indicating that the *SacII* site is completely unmethylated in normal DNA. DNA from the mother of the proband (lane 3), yields the normal 0.8 kb and 2.6 kb bands due to random inactivation of the normal allele. She also has an expanded allele, indicated by the presence of a 3.3 kb band. The proband (lane 4) shows a 3.5 kb fragment not digested by *SacII*. III-5 (lane 5) shows the normal male digestion pattern. c, DNAs were digested with *EcoRI* alone (lanes 1, 3 and 5) or in combination with *EagI* (lanes 2, 4 and 6). Lanes 1 and 2 show the normal female digestion pattern. DNA from individual II:2, a female (lanes 3 and 4) exhibits both a normal and an expanded allele. DNA from one of her sons, III:1, shows an allele which is expanded to a greater extent than that of his mother; his allele is methylated, as shown by lack of digestion with *EagI* (lane 6).



**Fig. 6** Segregation analysis of FRAXF through CEPH family 1424. This family shows stable segregation of a 29 repeat allele, with an apparent instance of mitotic instability. Allele size is expressed as number of repeats. Control is a female with 6 and 23 repeats. FF, Father's father; FM, father's mother; MF, mother's father; MM, mother's mother.

The features common to all four strongly suggest that other folate-sensitive fragile sites will follow a similar pattern. FRAXF represents the ninth unstable trinucleotide repeat described in the human, and the second (along with FRA16A) without a clear phenotype.

**Methodology**

**Clinical evaluation.** The proband in family 1 was noted to have strabismus in infancy; this was corrected by surgery. As an infant he was borderline slow in his motor development and was hypotonic. He walked at age 14–16 months. He began to speak single words at one year of age and subsequently had difficulty articulating. He was evaluated by a neurologist at 11 years of age and the diagnosis of a partial complex seizure disorder was made. He was in special education and functioning at around a first grade level. At 12 years he had a height at the 50th percentile and a head circumference at the 10th percentile. His ears were anteverted with hypoplasia of the mid-portion of both helices. He had a long narrow face and a mild mandibular prognathism and had a “hypotonic mouth breather” type facial appearance. Neurologic exam showed generalized hypotonia and mild hyperextensibility of joints. A follow-up evaluation at 14 years showed a height at the 75th percentile and a head circumference at the 25th percentile and testes of normal size.

**DNA isolation.** Genomic DNA was prepared from lymphoblastoid cell lines as described<sup>12</sup>. DNA was extracted from blood samples by phenol/chloroform extraction using standard techniques<sup>13</sup>. Plasmid DNAs were prepared using Wizard Minipreps (Promega).

**Sequencing.** Sequencing reactions were carried out in the Baylor Human Genome Center's nucleic acids core using an automated sequencer (ABI), or were carried out manually using a Sequenase Quick-Denature plasmid sequencing kit (USB) on double-stranded template.

**Southern analysis.** For each digest, 5 µg of genomic DNA was digested overnight with 20–40 U of each enzyme in buffers recommended for double digests (New England Biolabs). Electrophoresis through 1% agarose gels and blotting to Hybond N

nylon membrane (Amersham) was performed using standard techniques<sup>13</sup>. Hybridizations were carried out as described<sup>12</sup>. Probes were gel-isolated using the Prep-A-Gene kit (BioRad).

**Cytogenetics.** Fragile X induction was performed using excess 5-fluoro-2'-deoxyuridine (FUdR) (Sigma) or excess thymidine (Sigma) according to standard procedures<sup>14</sup>. Following routine GTG-banding, 15 cells were analysed from the FUdR cultures and 15 cells were analysed from the thymidine cultures. Seven of the 30 cells analysed showed a positive fragile X site (23%).

FISH was performed using cosmids 13F1 and 82E3 simultaneously, or cosmid 112G10 alone. Slides were prepared for FISH and each cosmid was labelled with digoxigenin as previously described<sup>15</sup>. For each experiment, 25 ng µl<sup>-1</sup> of cosmid DNA and 0.5 ng µl<sup>-1</sup> of biotinylated, X chromosome-specific alpha-satellite DNA (Oncor) was used in a 65% formamide hybridization solution. The slides were post-washed using a 50% formamide/2XSSC solution at 43 °C as described<sup>15</sup>. Dual colour probe detection was performed with amplification of the digoxigenin-labelled hybridization signal as described<sup>16</sup>. The slides were counterstained with DAPI and visualized using a Zeiss Axiophot fluorescent microscope equipped with a triple-band-pass filter. Digital images were captured using a PSI Powergene 810 probe system and printed using a Tektronix Color/Monochrome Phasar II SDX printer.

**PCR assay.** Primers 16082 (5'-AGG GGC AGG GGG CGG TGG CTC AGG TTT CTC-3') and 16083 (5'-CAG CCT GCG CCT TAC AGC GGG TTC ATG GCG-3') were synthesized based on the sequence of FRAXF (GenBank, # L35600). PCR reactions were carried out in 1 × Cetus PCR buffer; 200 µM each dATP, dCTP, dTTP; 50 µM dGTP; 150 µM 7-deaza-dGTP; 10% glycerol; 0.05 U ml<sup>-1</sup> AmpliTaq polymerase (Cetus); 1 mM each primer. Template DNA consisted of 100 ng of genomic DNA per 20 µl reaction. Reactions were carried out in a DNA Thermal Cycler 480 (Perkin Elmer), using the following cycle profile: 5 min at 95 °C; 30 cycles of 94 °C (1 min), 72 °C (1.5 min); 4 min at 72 °C. Products were resolved on 2% agarose gels for plus/minus assays. For allele sizing, 0.2 µl of α-32P-dCTP per reaction was added, and reaction volumes were reduced to 15 µl. The number of cycles was reduced to 25. Products were removed from under oil and added to an equal volume of formamide loading dye; 2 µl were loaded into the wells of a 6% polyacrylamide/8M urea/glycerol-tolerant gel. For size standards, a sequencing reaction was carried out on pUC19 DNA using the Quick-Denature Plasmid Sequencing kit (USB) according to the manufacturer's instructions. Gels were run 2–4 hours at 50 °C, fixed for 15 min in 5% acetic acid/15% methanol, dried down under vacuum at 80 °C, and exposed at -80 °C 2 h to overnight.

For large alleles, *Pfu* polymerase was used in the following reaction conditions: 1 × cloned *Pfu* buffer (Stratagene); 200 µM each dNTP; 15% DMSO; 0.05 U µl<sup>-1</sup> *exo(-)* *Pfu* polymerase (Stratagene); 0.5 µM each primer; 100 ng genomic DNA per 20 µl reaction. Cycle profile: 5 min at 98 °C; 30 cycles of 98 °C (1 min), 65 °C (1 min), 73 °C (2 min); 4 min at 73 °C. For allele sizing, label was added and reaction volumes and cycles were reduced as above for *Taq*; products were resolved on acrylamide gels as above.

**Acknowledgements**

We thank S. Tirrell, S. Madu and A. McAughan, E. Eichler and J. Mulley for their help and contribution to this work. The chromosome specific gene library LLOXNC01 used in this work was constructed at the Lawrence Livermore National Laboratory. J.E.P. is supported by the human genome distinguished postdoctoral fellowship program from the OHER of the U.S. Department of Energy, administered by the Oak Ridge Institute for Science and Education. This work is supported in part by grants DE-FG05-92ER61401 and DE-FG03-94ERG1830 from the U.S. Department of Energy (DLN), grant NIH 5R01 HD29256-02 from the NICHD of the NIH, and a Center grant from the NCHGR of the NIH (NIH 5P30 HG00210).

Received 22 August; accepted 20 September 1994.

1. Sutherland, G.R. Chromosomal fragile sites. *Genet. Anal. Tech. Appl.* **8**, 161–166 (1991).
2. Mandel, J.-L. & Heitz, D. Molecular genetics of the fragile-X syndrome: a novel type of unstable mutation. *Curr. Op. Genet. Devel.* **2**, 422–430 (1992).
3. Warren, S.T. & Nelson, D.L. Advances in molecular analysis of Fragile X syndrome. *J Am. med. Assoc.* **271**, 536–542 (1994).
4. Pieretti, M. *et al.* Absence of expression of the FMR-1 gene in Fragile X syndrome. *Cell* **66**, 817–822 (1991).
5. Sutcliffe, J.S. *et al.* DNA methylation represses FMR-1 transcription in fragile X syndrome. *Hum. molec. Genet.* **1**, 397–400 (1992).
6. Knight, S.J.L. *et al.* Trinucleotide repeat amplification and hypermethylation of a CpG island in FRAAXE mental retardation. *Cell* **74**, 127–134 (1993).
7. Hirst, M.C. *et al.* The identification of a third fragile site, FRAXF, in Xq27–q28 distal to both FRAAXA and FRAAXE. *Hum. molec. Genet.* **2**, 197–200 (1993).
8. Romain, D.R. & Chapman, C.J. Fragile site Xq27.3 in a family without mental retardation. *Clin. Genet.* **41**, 33–35 (1992).
9. Nancarrow, J.K. *et al.* Implications of FRA16A structure for the mechanism of chromosomal fragile site genesis. *Science* **264**, 1938–1941 (1994).
10. Wells, R.D. & Sinden, R.R. In Genome analysis: genome rearrangement and stability (eds Davies, K.E. & Warren, S.T.) 107–138 (Cold Spring Harbor Laboratory Press, New York, 1993).
11. Eichler, E.E. *et al.* Length of uninterrupted CGG repeats determines instability in the FMR1 gene. *Nature Genet.* **8**, 88–94 (1994).
12. Parrish, J.E. & D.L. Nelson Regional assignment of 19 X-linked ESTs. *Hum. molec. Genet.* **2**, 1901–1905 (1993).
13. Sambrook, J., Fritsch, E.F. and Maniatis, T. *Molecular cloning: a laboratory manual*. 2nd edn (Cold Spring Harbor Laboratory Press, New York, 1989).
14. Jacky, P. Fragile X and other heritable fragile sites on human chromosomes. In *The ACT Cytogenetics Laboratory Manual*. (ed. Barch, M.J.) 504 (1991).
15. Shaffer L.G. *et al.* Molecular characterization of de novo secondary trisomy 13. *Am. J. hum. Genet.* (in the press) (1994).
16. Han, J.-Y., Choo, K.H.A. & Shaffer, L.G. Molecular cytogenetic characterization of 17 rob(13q14q) by fluorescence *in situ* hybridization, narrowing the region containing the breakpoints. *Am. J. hum. Genet.* (in the press) (1994).
17. Schlessinger, D., Mandel, J.-L., Monaco, A.P., Nelson, D.L. & Willard, H.F. Report of the fourth international workshop on human X chromosome mapping 1993. *Cytogenet. cell Genet.* **64**, 148–170 (1993).