Molecular heterogeneity of the fragile X syndrome

Y.Nakahori, S.J.L.Knight, J.Holland¹, C.Schwartz², A.Roche, J.Tarleton², S.Wong², T.J.Flint, U.Froster-Iskenius³, D.Bentley¹, K.E.Davies* and M.C.Hirst

Institute of Molecular Medicine, John Radcliffe Hospital, Headington, Oxford OX3 9DU, ¹Paediatric Research Unit, The Prince Philip Research Laboratories, Guy's Tower, 8th Floor, London Bridge, London SE1 9RT, UK, ²Greenwood Genetic Center, 1 Gregor Mendel Circle, Greenwood, SC 29646, USA and ³Lubeck Medical University, 2400 Lubeck, FRG

Received July 3, 1991; Revised and Accepted July 30, 1991

ABSTRACT

The fragile X syndrome is an X-linked disorder which has been shown to be associated with the length variation of a DNA fragment containing a CGG trinucleotide repeat element at or close to the fragile site. Phenotypically normal carriers of the disorder generally have a smaller length variation than affected individuals. We have cloned the region in cosmids and defined the area containing the amplified sequence. We have used probes from the region to analyse the mutation in families. We show that the mutation evolves in different ways in different individuals of the same family. In addition we show that not all fragile X positive individuals show this amplification of DNA sequence even though they show expression of the fragile site at levels greater than 25%. One patient has alterations in the region adjacent to the CGG repeat elements. Three patients in fragile X families have the normal fragment with amplification in a small population of their cells. These observations indicate that there is molecular heterogeneity in the fragile X syndrome and that the DNA fragment length variation is not the only sequence responsible for the expression of the fragile site or the disease phenotype.

INTRODUCTION

The fragile X syndrome is a common genetic form of mental retardation affecting 1 in 1,500 males and 1 in 2,500 females (1,2). In addition to the mental retardation, many affected males have long faces, large jaws, protruding ears and macro-orchidism (for review see ref. 3). The disorder is associated with the expression of a fragile site at Xq27.3 in lymphocytes of patients when they are grown under conditions of thymidine stress (4,5). The inheritance of the disease is unusual since although it is X-linked it shows variable penetrance (6,7). Twenty per cent of males carrying the mutation do not express the fragile site and are phenotypically normal (known as normal transmitting males;

NTMs), whereas 30% of females carrying the mutation are affected. Daughters of NTMs are almost never mentally retarded and rarely express the fragile site.

Several theories have been proposed to account for the observed segregation patterns (8,9,10). Most of these suggest the occurrence of a phenotypically neutral premutation which is then converted into a full mutation by a second event occurring in female meiosis. Laird (8) suggested that the inheritance could be explained on the basis of imprinting mediated through DNA methylation. Indeed, our recent studies (11) and those of others (12) have shown that a CpG island at or near the fragile site is specifically methylated in patients whereas it is unmethylated in normal controls. NTMs are not usually methylated at this site. However, we have shown that this CpG island is methylated in a phenotypically normal male expressing the fragile site in 34% of his cells (11).

Using microdissection clone STS nucleation, we have recently constructed a 1.1Mb YAC contig across this region of hypermethylation (13,14). By direct screening of cosmid clones with inter-Alu PCR products, we have isolated the region immediately adjacent to the CpG island which is methylated in patients. DNA probes from this region detect differences in genomic DNA fragments in fragile X patients (15). These DNA sequence changes are identical to those already reported by other investigators (16,17,18) and occur within the coding region of a gene (FMR-1) which is expressed in blood cells and brain (18). The FMR-1 gene does not show any homology to any previously described sequence but has characteristics similar to protamines which are involved in chromatin structure. Thus FMR-1 could play a role in chromosome condensation.

The changes in fragment length adjacent to the CpG island can be readily monitored in Southern blots after digestion of DNA samples with a variety of restriction enzymes (15,16,17,18). In this paper, we present the analysis of the fragile X mutation in several families including the family of the NTM showing methylation at the CpG island. Our data suggest that the DNA fragment length changes in FMR-1 are present in the vast majority of cases of the fragile X syndrome but not all. The implications

* To whom correspondence should be addressed

of these results for the mechanism of the occurrence of the fragile site and the clinical expression of the disorder are discussed.

MATERIALS AND METHODS

Identification of cosmids from the YAC contig

Alu-PCR products were generated from YACs which were shown to cover the fragile X region (14). 27 cycles at 94°C, 1 minute, 50°C, 2 minutes, 72°C, 2 minutes, were carried out on 50ng of total yeast DNA using the Alu consensus primers 938 (5'-CACTGCACTCCAGCCTGGG) and 939 (5'-GTGCTGGGA-TTACAGGCGTG). After ethanol precipitation, the total products were radiolabelled, pre-annealed with sheared human DNA and were used to screen a gridded cosmid library from a 49,XXXXX individual (Holland et al, manuscript in preparation). Among 13 positively hybridising cosmids generated from the Alu PCR product with primer 938 from the YAC141H5, 3 were localised proximal and 6 distal to the CpG island. Cosmid 181/D4.6 maps immediately distal to the CpG island and contains the CGG repeat (see Results). Cosmid restriction fragments were subcloned into pUC18 and maintained in *E. coli* DH1. Ox1.9 is a 1.9kb BamHI-EcoRI fragment and Ox1.4 is a 1.4kb BamHI-PstI fragment which also includes sequences from the cosmid vector. The Ox0.55 probe was generated by PCR using the primers 600 (5'-CTGCAGGAGGCGGCCCGG) and 602 (5'-GCTAGCAG-GGCTGAAGAGAAGATG). Primer 600 is made to the PstI site in the FMR first intron based upon sequence obtained from Ox1.9. Primer 602 lies in the FMR coding region and corresponds to positions 175–198 in the cDNA sequence in ref 18. PCR was carried out for 30 cycles of 94°C, 1 minute; 64°C, 1 minute; 72°C, 1 minute in standard conditions containing 1.5mM MgCl₂.

Southern blot and hybridisation analyses

DNA samples were digested with restriction enzymes under the recommended conditions of the commercial supplier and blotted onto Hybond N (Amersham International) after electrophoresis. Labelling of probes was performed by the hexaprimer method (19). Probes were hybridised under conditions as described previously (13, 14). Filters were washed in 0.1×SSC and 0.1% SDS at 60°C or 65°C and exposed against film for 1–6 days at –70°C.

RESULTS AND DISCUSSION

Mapping of the mutated region

Cosmids from the fragile site region were obtained by the direct screening of gridded cosmid libraries with total inter-Alu PCR products from YACs shown to form a physical contig of the region (see Materials and Methods; 14). Cosmids were analysed for rare-cutting restriction enzyme sites and aligned to the corresponding fragments in the YAC contig. Using this method, cosmid 181/D4.6 was identified which contains sequences immediately adjacent to the hypermethylated CpG island. Hybridisation of fragments from this cosmid were used to construct a genomic restriction map of the region (Fig. 1a). This map is in agreement with that previously reported (16, 17) but lacks one of the EcoRI fragments indicated on the map of Verkerk et al (18). An analysis with the probe Ox1.9 of DNA samples from fragile X patients indicated differences from the normal fragment sizes. These differences could be detected with the enzymes HindIII, EcoRI and BglII but not with PstI. Figure 1b

shows a typical analysis with EcoRI. None of the affected males show defined bands but a faint smear of fragments extending above the position of the normal fragment at 5.1kb. The smear of fragments indicates somatic variation of the mutation and has been noted previously (15,16,17,18). No significant differences were seen in these fragments for lymphoblastoid cell line and lymphocyte samples of the same individual.

Figure 1. a. Genomic map of the region containing the fragile X mutation and the CpG island. After mapping the cosmid 181/D4.6 with EcoRI (E), HindIII (H), XhoI (X), Bam HI (B), BglII and XbaI, the genomic map of the surrounding region was expanded by southern hybridization analyses on genomic DNA. The rare cutter restriction enzyme used were NruI (Nru), EagI (Eag), BssHII (Bs) and SacII (Sac). Only the HindIII sites flanking to the CpG island are represented. The detailed map of the subclone is expanded below the genomic map. To make the detailed map of the subclones, NarI (N), BanI (Ba), NheI (Nh) and PstI (P) were used in addition to the enzymes above. The NarI sites are also BanI sites. Sequence data from analysis of Ox1.4 (numbered according to the cDNA numbering used in ref 18) extends the 5' of the sequence of the cDNA in ref 18. The two nucleotides after the NarI site are GC in our sequence data (marked by **) which alters one amino acid. The CGG_n refers to varying copies of the trinucleotide found in Ox1.4 clones (see text). Sequencing also revealed an exon-intron boundary at position 246 after lys 82. (Ba) represents a polymorphic BanI site (see text). + indicates sites methylated in patients and inactive X chromosomes. **b.** Analysis of DNA samples from fragile X patients and normal male (M) and normal female (F) controls. Samples were digested with EcoRI; the blot was hybridised to a properdin cDNA probe p5/6 as a control (top panel) and Ox1.9 in the bottom panel.

Since no fragment length changes were observed in PstI digests, the variation must be occurring proximal to this site. DNA sequence was generated from the Ox1.9 fragment which allowed us to design PCR primers to screen for variation proximal to this PstI site. The PCR product, Ox0.55 (Fig. 1a), is identical to the PstI probe pfxa3 used by other investigators (16). PCR amplification of this fragment in fragile X individuals was identical to that observed for normal individuals confirming that the DNA fragment variations are occurring in the small region proximal to the Nhe I site. The probe Ox0.55 detects changes in DNA fragment size in fragile X DNA samples digested with PstI as has been reported (16). In general, most DNA variations could be detected in EcoRI digests after hybridisation with Ox1.9 (Fig. 1b) but small insertions (100bp) were more easily visible in PstI digests.

Sequence analysis

Ox1.4, the fragment of the 181/D4.6 cosmid adjacent to the CpG island and including the fragment altered in patients, shows a variable size when grown in bacteria. Sequence analysis indicated that this variation was due to copy number of the CGG trinucleotide (see fig 1a). Two extremes isolated and sequenced had copy numbers of (CGG)₃ and (CGG)₂₆. This indicates an instability of this region in bacteria and suggests that complete sequence data for this region may be difficult to obtain from cloned sources. We also found no copies of the AGG trinucleotide as reported in reference 18. As this region appears to have deleted in some of our clones, we cannot rule out the presence of other unstable DNA elements from within this region. Therefore, a complete analysis of the mutations in this region may rely upon the direct genomic sequencing of patient material.

Further analysis of the sequence of Ox1.4 indicated a divergence at position 246 in the cDNA sequence of FMR-1 (18) and that of Ox1.4. The presence of a consensus splice donor site at this position in Ox1.4 (see fig 1a) indicated that this divergence was due to the presence of an intron in the Ox1.4 genomic fragment.

Family studies of the mutation

Figure 2 shows the inheritance of the DNA sequence changes in a four generation fragile X family showing male transmission. (We have previously analysed some of these individuals using BglII digests where doublets were not seen because of the limit

of resolution of the gel in the range of the 12kb normal BglII fragment;15). The analysis here was carried out by HindIII digestion of DNA samples and hybridisation with probe Ox1.9. It is possible to follow the expansion and contraction of the DNA fragment in this family through the generations. The carrier female in the first generation (lane 2) has a normal 5.1kb allele and a mutant 5.3kb allele. In her affected son (lane 4), the mutant allele is visible only as a smear. Her phenotypically normal son (lane 6) has inherited a mutant allele which has increased in size to 5.6kb. His daughter (lane 9) inherits the mutant allele which has now decreased in size to 5.2kb. In her affected son (lane 3) the mutant allele is visible only as a smear of fragments indicating somatic heterogeneity. This analysis demonstrates how the mutation can vary by both reduction and increase in size of the DNA fragment and that this change can occur where there has been no opportunity for recombination in a female meiosis.

The NTM in figure 2 shows an increase in size of 500bp compared to the normal fragment. In five additional NTMs that we have studied, the fragment size increase was 200–300bp.

Figure 3 shows the analysis of the mutation in another family with several affected sibs. Lanes 1–4 and 7 are affected brothers where an increase in size of the normal BglII fragment is seen in all cases with a large degree of somatic heterogeneity. A smear of fragments can also be seen in the affected sister (lane 5) together with the normal fragment of 12kb. These patterns are in contrast to the abnormal but more defined fragments observed for the other affected males in this pedigree (lanes 10,14). The carrier mother (lane 8) of the affected males (lanes 1–4) shows a tight doublet in this analysis indicating an insertion of 200–300bp of DNA. Interestingly, the affected boy in lane 1 shows a fragment in the normal position in addition to a faint smear. We have seen this type of pattern in 3 out of 111 (2.7%)

Figure 2. Analysis of the evolution of the fragile X mutation in a family. The numbers on the pedigree (b) correspond to the numbers on the lanes on the autoradiogram in (a). Levels of fragile site expression are given as %. Samples were digested with HindIII; the blot was hybridised to Ox1.9. □, ○, normal male and female respectively; ■, ● affected male and female respectively. Dots in symbols indicate carrier status.

Figure 3. Analysis of the fragile X mutation in a family after digestion with BglII and hybridisation to Ox1.9. The numbers on the pedigree (b) correspond to the lane numbers in (a). Levels of fragile site expression are given as %.

Figure 4. BanI analysis in individuals with Ox1.9. Lanes 1, 5 and 8 are normal controls, lanes 2, 6, and 7 affected males and lanes 3 and 4 NTMs (see text). The rare BanI RFLP is present in the individual in lane 2.

fragile X males. A similar case was also reported by Oberle et al (17). Cases such as these could be misinterpreted as normal on gels which do not resolve the smear of fragments.

Molecular heterogeneity

We have studied five exceptional families in which expression of the fragile site is not always associated with the typical clinical phenotype of the fragile X syndrome or where the clinical features are seen but no expression of the fragile site is observed.

The first family studied had an isolated case who was fragile X positive (3/60) in an initial test and fragile X negative on re-testing. He first presented with clinical features suggesting the Martin-Bell syndrome but he now has an IQ in the normal range and has autistic features. His mother was fragile X negative. He showed no detectable fragment change in EcoRI digests with Ox1.9 (data not shown). However, in assays of the adjacent region with BanI digests, this individual showed a small shift from the normal sized fragment of 2.8kb (lane 6, Fig 4). This shift, which was observed in three independent analyses, could be explained either by a small insertion of DNA or a polymorphism. This shift was not observed in 32 other individuals.

The amplification of sequence seen in EcoRI analyses of affected males is not seen in BanI digests (e.g. lanes 2 and 7). This is consistent with the map given in figure 1a where there are BanI sites between the CGG repeat and the sequence Ox1.9 used as probe. This map has one extra BanI site compared with the genomic map of Oberle et al (17). The 3.0kb BanI fragment was reported to be increased in size by 500bp in patients due to methylation (17). We only observed methylation of this fragment in 3 out of 28 fragile X males. However, the adjacent proximal BanI site (also an NarI site; see figure 1a) is methylated in patients (unpublished observations). As noted by others (17) we also detect a polymorphic BanI site in this region (labelled as (Ba) in Figure 1a) which generates a polymorphism with allele sizes of approximately 3.0kb and 2.5kb/0.5kb. The patient in lane 2 has the rare allele and a lower fragment at 1.3kb due to partial methylation.

A second family analysed also had an isolated case. He showed the fragile site in 7% of his cells and had clinical features typical of the syndrome. We did not observe any increase in size of DNA fragments from the region in EcoRI or PstI analyses of this

Figure 5. Analysis of individuals from pedigrees in (5b, c and d) after EcoRI digestion and hybridisation with Ox1.9. Lanes 17 and 18 are normal controls.

individual. At present, we do not know whether this male has a mutation elsewhere in Xq27.3 or whether his phenotype is due to a completely different genetic disease.

In the third family (figure 5b), the proband is a fragile X positive dull male. He is at present too young to determine whether he has the typical clinical features of the fragile X syndrome. His mother is fragile X negative whereas his grandfather, who is phenotypically normal, also expresses the fragile site. We have shown previously that this grandfather is methylated at the CpG island. The grandfather, his daughter and the proband show normal sized fragments (lanes 4, 5, and 6, Fig.5a). These individuals do not show insertions in the BanI fragment (the grandfather is lane 4 of Fig.4).

The partial pedigree of the fourth family where the fragile X syndrome was found not to be associated with a DNA fragment size variation is also shown in Figure 5c. The analysis of these samples is given in lanes 8–12 of Figure 5a. No evidence of extra fragments or a smear is seen in the affected female in the second generation (lane 10). Her brother who expresses the fragile site at high levels and who is phenotypically normal also shows a normal pattern (lane 9). The mother of these individuals shows a high level of expression of the fragile site. No evidence for a mutated X chromosome is visible (lane 11). We tested the individuals in these two pedigrees in PstI digests with the probe Ox0.55 to test for very small amplifications but none were found (data not shown).

The fifth exceptional family studied (figure 5d) was chosen because the mentally retarded brother of a fragile X positive, mentally retarded female was found to be fragile X negative in two laboratories on repeated occasions (20). This male was found to show a smear of fragments (lane 15) with the probe Ox1.9. His mentally retarded sister (lane 16) showed the normal fragment and abnormal fragments of higher molecular weight. The obligate carrier mother gave a normal pattern (lane 14). Since this female is the daughter of a normal transmitting male, she may have a fragment size increase below the resolution of this gel.

The individuals expressing the fragile site with apparently normal fragments presented in figure 5 must have either very small amplifications of DNA sequence (< 50bp) or they express the fragile site at high levels because of a DNA sequence change elsewhere in the region. Whether this CGG repeat is the molecular basis of the fragile site is still unclear as predictions based upon the biochemical affects of the agents used to induce fragility suggested the presence of an AT rich element at the fragile site (5). The relationship of these elements with respect to the fragile site is currently under investigation by *in situ* hybridisation. If the mutation at the fragile site alters the folding of the chromatin in the region, it may be that there are several sites in Xq27.3 which might be mutated with a similar effect. We are currently analysing these patients with probes from our cosmid contig of the region to investigate this further.

In conclusion, we find evidence for the amplification of a DNA sequence in fragile X patients and in NTMs in the region of the FMR-1 gene containing the CGG repeats. These amplification events are occurring in the 500bp containing the CGG trinucleotide repeat immediately adjacent to the hypermethylated CpG island. This may involve a direct amplification of this repeat or the variable insertion of another element within it. The effects of this amplification upon the stability of the FMR-1 mRNA and any effects upon the production or activity of the FMR-1 protein are currently under investigation. However, as some fragile X positive individuals do not show any change in the size of the fragment at this position, more subtle alterations may also be occurring within this region. If the amplification of the CGG region is the primary event in the mutation process in the fragile X syndrome, our data suggest that this may be the most common mechanism but that there must be alternative mechanisms which have the same phenotypic consequences.

ACKNOWLEDGEMENTS

We are particularly grateful to Patricia Jacobs and James Macpherson for the families shown in figure 5 and for critical reading of the manuscript. We are also grateful to all clinicians and cytogeneticists who contributed samples for this project. We thank Helen Blaber for assistance in the preparation of the manuscript and Bryan Bolton (PHLS, Porton Down) for cell lines. We thank the Medical Research Council of Great Britain, Action Research for the Crippled Child, South Carolina Department of Mental Retardation and NIH Grant No: MH45916 for financial support.

REFERENCES

- Webb, T., Bunday, S., Thake, J. and Todd, A. (1986) *J. Med. Genet.*, **23**, 396–399.
- Kahkonen, M., Alitalo, T., Airaksinen, E., Matilainen, A., Launiala, R., Autio, S. and Leisti, J. (1987) *Hum. Genet.*, **77**, 85–87.
- Fryns, J.-P. (1990) In Davies, K.E. (ed.), *The Fragile X Syndrome*, OUP, Oxford, pp. 1–39.
- Lubs, H.A. (1969) *Am. J. Hum. Genet.*, **21**, 231–244.
- Sutherland, G.R. (1977) *Science*, **197**, 256–266.
- Sherman, S.L., Morton, N.E., Jacobs, P.A. and Turner, G. (1984) *Ann. Hum. Genet.*, **48**, 21–37.
- Sherman, S.L., Jacobs, P.A., Morton, N.E., Froster-Iskenius, U., Howard-Peebles, P.N., Nielsen, K.B., Partington, M.W., Sutherland, G.R., Turner, G. and Watson, M. (1985) *Hum. Genet.*, **69**, 289–299.
- Laird, C., Jaffe, E., Karpen, G., Lamb, M. and Nelson, R. (1987) *Trends in Genet.*, **3**, 274–281.
- Pembrey, M.E., Winter, R.M. and Davies, K.E. (1985) *Am. J. Med. Genet.*, **21**, 709–717.
- Nussbaum, R.L., Airhart, S.D. and Ledbetter, D.H. (1986) *Am. J. Med. Genet.*, **23**, 715–722.
- Bell, M.V., Hirst, M.C., Nakahori, Y., MacKinnon, R.N., Roche, A., Flint, T.J., Jacobs, P.A., Tommerup, N., Tranebjærg, L., Froster-Iskenius, U., Kerr, B., Turner, G., Lindenbaum, R.H., Winter, R., Pembrey, M., Thibodeau, S. and Davies, K.E. (1991) *Cell*, **68**, 861–866.
- Vincent, A., Heitz, D., Petit, C., Kretz, C., Oberle, I. and Mandel, J.-L. (1991) *Nature*, **329**, 634–626.
- Hirst, M.C., Roche, A., Flint, T.J., MacKinnon, R.N., Bassett, J.H.D., Nakahori, Y., Watson, J.E.V., Bell, M.V., Patterson, M.N., Boyd, Y., Thomas, N.S.T., Knight, S.J.L., Warren, S.T., Hors-Cayla, M., Schmidt, M., Sutherland, G.R. and Davies, K.E. (1991) *Genomics*, **10**, 243–249.
- Hirst, M.C., Rack, K., Nakahori, Y., Roche, A., Bell, M.V., Flynn, G., Christadoulou, Z., MacKinnon, R.N., Francis, M., Littler, A.J., Anand, R., Poustka, A.-M., Lehrach, H., Schlessinger, D., D'Urso, M., Buckle, V.J. and Davies, K.E. (1991) *Nucleic Acid Res.*, **19**, 3283–3288.
- Hirst, M.C., Nakahori, Y., Knight, S.J.L., Schwartz, C., Thibodeau, S.N., Connor, J.M., Fryns, J.-P., Roche, A., Flint, T.J. and Davies, K.E. (1991) *J. Med. Genet.*, in Press.
- Yu, S., Pritchard, M., Kremer, E., Lynch, M., Nancarrow, J., Baker, E., Holman, K., Mulley, J.C., Warren, S.T., Schlessinger, D., Sutherland, G.R. and Richards, R.I. (1991) *Science*, **252**, 1179–1181.
- Oberle, I., Rousseau, F., Heitz, D., Kretz, C., Devys, D., Hanauer, A., Boue, J., Bertheas, M. and Mandel, J.-L. (1991) *Science*, **252**, 1097–1102.
- Verkerk, A.J.M.H., Pieretti, M., Sutcliffe, J.S., Fu, Y.-H., Kuhl, D.P.A., Pizzuti, A., Reiner, O., Richards, S., Victoria, M.F., Zhang, F., Eussen, B.E., van Ommen, G.-J.B., Blonden, L.A.J., Riggins, G.J., Chastain, J.L., Kunst, C.B., Galjaard, H., Caskey, C.T., Nelson, D.L., Oostra, B.A. and Warren, S.T. (1991) *Cell*, **65**, 905–914.
- Feinberg, A.P. and Vogelstein, B. (1983) *Anal Biochem*, **132**, 6–13.
- Temple, I.K., Barattner, M., Pembrey, M.E., Butler, L., Jacobs, P. and Davies, K.E. (1990) *Lancet*, **336**, 1131.