Bgl II subclone of Lambda 5R containing THE-1 sequences found in the discrete extrachromosomal circles of HeLa cells, from C. Schmid [R. Misra, A. G. Matera, C. W. Schmid, M. G. Rush, *Nucleic Acids Res.* **17**, 8327 (1989)]; the *c-myc* clone HSR-1 [K. Alitano, M. Schwab, C. C. Lin, H. E. Varmus, J. M. Bishop, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1707 (1983)] from J. Bishop. Cosmids containing the entire p53 genomic region were from Y. Nakamura; cosmids including the entire dihydrofolate reductase amplicon from hamster cells [B. Anachkova and J. L. Hamlin, *Mol. Cell. Biol.* **9**, 532 (1989)] were from J. Hamlin.

13. Binding site definition by PCR: One primer for each reaction was labeled with ^{32}P at the 5′ end with T4 polynucleotide kinase in a 5-μl reaction, and the kinase inactivated at 70°C for 5 min. PCR was done with 350 ng of each of the appropriate primers and approximately 50 ng plasmid template in a 50-μl volume; 25 cycles were used with the PCR conditions specified in S. J. Baker *et al.* [*Cancer Res.* **50**, 7717 (1990)]. The products were extracted with phenol and chloroform, ethanol-precipitated, and dissolved in 3 mM tris, 0.2 mM EDTA prior to binding. Subfragment 1 contained bp 1 to 425 of subclone 10d of fragment A (Fig. 3A); subfragments 1a, 1b, 1c, 1d, and 1e were generated by digestion of subfragment 1 with Bam HI, Mbo I, Hind III, Hind III, and Bam HI, respectively. Subfragment 2 contained bp 283 to 425. Subfragment 3a was generated by digestion of subfragment 3 (bp 106 to 294) with Hae III. Subfragment 4a was produced from subfragment 4 (bp 1 to 141) by Hind III digestion. Subfragments 5a and 5b were products of the Hae III digestion of subfragment 5 (bp 87 to 141). Mutant subfragments 5mut1 and 5mut2 were produced with primers P3m1 (5′-GAAAGAAAAGGCAAGGCCAGGAAAGT-3′) and P3mut2 (5′-GAAAGAAAAGGCAAGGCCA*TTA*-AAGT-3′) and were identical to subfragment 5 except for the positions italicized in the primers. Subfragment 6 contained bp 106 to 138, and the insert was excized by digestion with Hind III and Bam HI to generate 6a or with Hind III and Eco RI to generate 6b.
14. Methylation interference assay: PCR products labeled at one end were generated with primers labeled with T4 polynucleotide kinase (U.S. Biochemicals). The product of each PCR reaction was purified by polyacrylamide gel electrophoresis, eluted from a crushed gel slice in 500 mM ammonium acetate, extracted with phenol and chloroform, and precipitated with ethanol. DNA (2×10^6 dpm) was methylated at G residues by means of dimethylsulfate as described [T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982), p. 477], ethanol-precipitated and dissolved in 10 μl of 3 mM tris, 0.2 mM EDTA. A volume of 0.5 μl was removed as the DNA control. A total of 4.5 μl was added to a binding reaction containing baculovirus-produced p53 or vaccinia-infected cell lysates. The immunoprecipitated DNA was purified by SDS–proteinase K digestion, extracted with phenol and chloroform, and ethanol-precipitated. The control DNA and precipitates of bound DNA were cleaved with piperidine at the methylated sites. Equivalent amounts of labeled DNA were loaded and separated on a denaturing polyacrylamide (14.5%) sequencing gel, which was fixed and dried for autoradiography.
15. J. Huberman, personal communication.
16. W. R. Jelinek *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1398 (1980).
17. Fragments A and B were precipitated more efficiently when a combination of the two monoclonal antibodies (*11*) were used rather than one antibody alone. While this is attributable in part to an increased effectiveness of immunoprecipitation of p53, a potential role for antibodies in stabilization of the DNA-protein interaction cannot be excluded. [V. Zimarino, S. Wilson, C. Wu, *Science* **249**, 546 (1990)].
18. Supported by The Clayton Fund, the McAshan Fund, and NIH grants CA06973, CA43460, and CA33620. We thank K. Johnson for construction of recombinant plasmids and confirmation of p53 expression in the vaccinia system and E. Fearon for reading the manuscript.

4 April 1991; accepted 24 May 1991

Mapping of DNA Instability at the Fragile X to a Trinucleotide Repeat Sequence p(CCG)*n*

E. J. KREMER, M. PRITCHARD, M. LYNCH, S. YU, K. HOLMAN, E. BAKER, S. T. WARREN, D. SCHLESSINGER, G. R. SUTHERLAND,* R. I. RICHARDS

The sequence of a Pst I restriction fragment was determined that demonstrates instability in fragile X syndrome pedigrees. The region of instability was localized to a trinucleotide repeat p(CCG)*n*. The sequences flanking this repeat were identical in normal and affected individuals. The breakpoints in two somatic cell hybrids constructed to break at the fragile site also mapped to this repeat sequence. The repeat exhibits instability both when cloned in a nonhomologous host and after amplification by the polymerase chain reaction. These results suggest variation in the trinucleotide repeat copy number as the molecular basis for the instability and possibly the fragile site. This would account for the observed properties of this region in vivo and in vitro.

FRAGILE X SYNDROME IS THE MOST common form of familial mental retardation (*1*). In addition to characteristic clinical symptoms the disorder has an associated chromosomal abnormality at Xq27.3 in the form of a rare fragile site. We and others (*2, 3*) have isolated large fragments of human DNA in the form of yeast artificial chromosomes (YACs) that span the fragile X region. Previously we localized the fragile X region within this cloned DNA by in situ hybridization and identified the location of breakpoints of two somatic cell hybrids, constructed to break at the fragile X (*4*), to a common 5-kb Eco RI restriction fragment (pfxa1). This restriction fragment was found to be unstable in fragile X pedigrees, the instability segregating with genotype (*5*). The unstable sequences were further localized to a 1-kb Pst I restriction fragment (pfxa2) (Fig. 1A). A search through the GenBank DNA sequence collection revealed limited identity with a functionally heterogenous group of proteins encoded by genes which all contain $p(CCG)_n$ sequences. This identity ranged from 60% for a 433-nucleotide sequence of the chicken protamine gene to 85% for a 105-nucleotide sequence of the human androgen receptor. Other notable proteins within this range were several fish antifreeze proteins and mammalian keratins. The functional significance, if any, of these sequence identities is not clear.

The pfxa2 sequence contains two regions of particular note. One is a CpG-rich region that contains seven recognition sites for CpG-dependent, infrequently cutting restriction endonucleases (positions 1 to 357). Three of these sites have been demonstrated (*6, 7*) to be targets for methylation in fragile X–affected individuals. CpG-rich regions are also noteworthy in that they frequently identify the promoter regions of eukaryotic genes (*8*). The second outstanding feature of pfxa2 is the $p(CCG)_n$ repeat (position 358 to 476). This sequence also contains CpG and may itself be the subject of methylation.

To ascertain which region of pfxa2 gave rise to the observed instability, oligodeoxyribonucleotide primers suitable for polymerase chain reaction (PCR) analysis of the 1-kb Pst I restriction fragment were designed (Fig. 2B). Lymphocyte DNA from normal and affected individuals was used as template for the reaction. While the reactions primed by the oligos 201 and 204 and 209 and 214 gave constant products from both normal and fragile X–genotype individuals, the reaction primed by oligos 203 and 213 did not produce product in any reaction except for one that contained high molar concentrations of the subcloned XTY26 YAC from a fragile X genotype individual (*5*) as template (*9*). The composition of the sequence between the 203 and 213 primers in XTY26 is 92% GC, which presents a formidable problem to the PCR. A variety of PCR conditions and additives (including trimethylammonium chloride, dimethyl sulfoxide, formamide, and glycerol) have been utilized unsuccessfully in attempts to facilitate amplification across this region. The products generated from XTY26 were also noted to be heterogeneous (*9*).

In attempts to obtain sequence data for this region from normal individuals and additional fragile X genotype individuals, two approaches were undertaken. The first used two-stage PCR. Starting material was either total chromosomal DNA or, in one

E. J. Kremer, M. Pritchard, M. Lynch, S. Yu, K. Holman, E. Baker, G. R. Sutherland, R. I. Richards, Department of Cytogenetics and Molecular Genetics, Adelaide Children's Hospital, North Adelaide, South Australia, Australia.
S. T. Warren, Departments of Biochemistry and Pediatrics, Emory University School of Medicine, Atlanta, GA.
D. Schlessinger, Center for Genetics in Medicine, Washington University School of Medicine, St. Louis, MO.

*To whom correspondence should be addressed.

case, Eco RI–digested DNA from a normal individual; the DNA was fractionated by agarose gel electrophoresis to enrich for the 5-kb Eco RI fragment which contains the p(CCG)$_n$ repeat. In the first-stage PCR the 201 and 214 primers (Fig. 2) were utilized. The products of this reaction were used as template for the second-stage 203- and 213-primed PCR. The products of these reactions were then subcloned into M13 for sequence analysis. This analysis revealed that only the length of the repeat sequences varied—the flanking sequences between the PCR primers and the repeat remaining constant (Fig. 1B). In all cases the cloned PCR products were substantially shorter than anticipated, particularly since the fragile X individuals had large insertions or amplifications of sequences in this region, indicating deletion of the unstable sequences in vitro.

In a second approach the probes pfxa3 and pfxa5 (Fig. 2A) were used to screen human chromosomal DNA libraries. One library (LAOXNL01) purchased from the American Type Culture Collection had been constructed from flow-sorted human X chromosomes and contained 4.25×10^5 independent recombinant clones (8.5 genomic equivalents). Despite screening 7.5×10^5 clones of the amplified library, neither unique probe gave positive clones which spanned the trinucleotide p(CCG)$_n$ repeat. A second library was constructed in lambda ZAPII (Stratagene) from a normal individual's DNA that had been digested with Eco RI and fractionated by agarose gel electrophoresis to enrich for the sequence containing the 5-kb p(CCG)$_n$ repeat. Screening of five genome equivalents of the primary plating of independent recombinant clones with either probe again failed to reveal any clones which spanned the repeat. We conclude that this sequence is underrepresented in chromosomal DNA libraries, presumably because of its composition. Support for this conclusion was previously noted (*5*) in generating the λ contig from XTY26 that spanned the repeat sequence. The library from which the contig was established contained five unique sequence equivalents, yet only one clone, λ5, was obtained that spanned the p(CCG)$_n$ repeat.

To further substantiate the localization of the instability of the p(CCG)$_n$ repeat and to more accurately ascertain its length in normal and affected individuals, additional higher resolution Southern blot mapping was undertaken (Fig. 3). Restriction sites revealed by sequence analysis of the Pst I fragment were used (Fig. 2C) with the PCR product generated from the 201 plus 204 oligos (pfxa4) as probe. This analysis (Fig. 2D) confirmed that the region of instability is located between the Sau 3AI (position 222) and Nhe I (position 509) recognition sites that flank the p(CCG)$_n$ repeat. Furthermore, a series of somatic cell hybrids have been constructed from a fragile X–expressing parent cell line in such a way as to select for breakage at the fragile site (*4*). Two of these hybrids, micro21D and Q1X, which have proximal and distal parts of the X chromosome from the fragile X were subjected to Southern blot analysis. DNA from micro21D hybridized to pfxa4 but not pfxa3, while QIX showed the opposite pattern of hybridization. PCR primed by oligos 201 and 204 gave a product of the correct size for DNA from micro21D but gave no product for QIX DNA, whereas PCR primed by oligos 209 and 214 gave the expected product for QIX but no product for micro21D.

These analyses demonstrate that both hybrids had breakpoints that localized to the p(CCG)$_n$ repeat, further supporting the localization of the fragile site to the p(CCG)$_n$

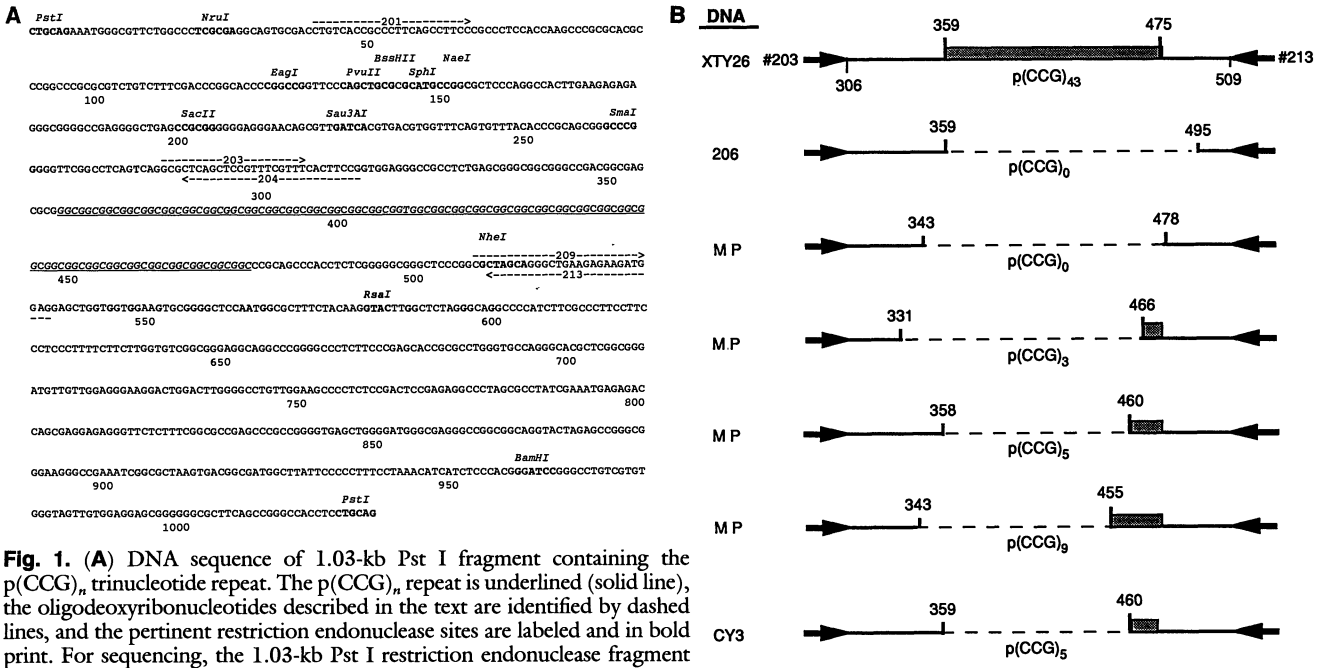

Fig. 1. (**A**) DNA sequence of 1.03-kb Pst I fragment containing the p(CCG)$_n$ trinucleotide repeat. The p(CCG)$_n$ repeat is underlined (solid line), the oligodeoxyribonucleotides described in the text are identified by dashed lines, and the pertinent restriction endonuclease sites are labeled and in bold print. For sequencing, the 1.03-kb Pst I restriction endonuclease fragment was isolated from pfxa1 and subcloned into the Pst I site in M13 mp18. Only clones in the forward orientation could be isolated (~30 clones screened). The 530-bp Nhe I to Pst I restriction endonucleases fragment was also isolated from pfxa1 and subcloned in both orientations into the Xba 1–Pst I sites in M13 mp18 and 19. The difficulties in isolating M13 clones that spanned the p(CCG)$_n$ repeat in the reverse direction led us to use double-stranded sequencing of pfxa2 using oligodeoxyribonucleotide primers #201, 203, 204, 209, and 213. All sequencing was performed with Sanger's dideoxy method and with TAQuence sequencing kit (U.S. Biochemical Corp.). Because of high GC content of the template DNA, samples were routinely prepared with 7-deaza-dGTP, denatured in a final concentration of 50% formamide at 90°C for 5 min, and loaded onto sequencing gels immediately without allowing to cool. (**B**) Deletion of sequence from two normals (CY3 and MP) and a fragile X patient (206). The four deletions for MP represent four independently cloned PCR products. PCR products spanning the p(CCG)$_n$ repeat were generated as described in the text and separated on a 1.5% low melting point agarose (IBI) gel. The band was excised, extracted with phenol and phenol:chloroform, treated with kinase, and subcloned into the Sma I site of M13 mp18. The sequence of each sample was identical to that of XTY26 DNA except at points of deletions as indicated. The precise sites of deletion could not always be determined because of sequence redundancy flanking the deletion site. The shaded region (bases 359 to 475) indicates the perfect repeat p(CCG)$_n$. Other copies of the p(CCG)$_n$ repeat were found in both flanking sequences and on occasion also appeared to be sites of deletion end points (for example, 338 to 343 and 494 to 497).

repeat itself (Fig. 2E). The PCR results and the sequencing of the PCR products indicated that only the length of the p(CCG)$_n$ repeat region varied, presumably due to variation in the copy number of the repeat. Unexpectedly, the two normal males in Fig. 3 showed consistent minor differences in mobility of the repeat region. To further characterize this apparent polymorphism, DNA from 12 normal males was digested with Sau 3AI and Nhe I and found to exhibit minor-length polymorphism (*9*). To determine the mode of inheritance of this polymorphism, DNA from mothers in the Centre d'Etude du Polymorphisme Humain (CEPH) pedigree was screened for heterozygosity of this band and family number 1349 genotyped (Fig. 4A). The inheritance was found to be stable codominant Mendelian segregation as opposed to that observed in fragile X syndrome pedigrees where the unstable band frequently changes in size when passed from carrier parent to offspring.

The p(CCG)$_n$ repeat appears to consist of about 40 ± 25 copies in normal individuals as determined by the length of the Sau 3AI (position 222) to Nhe I (position 509) fragment. More than 300 normal X chromosomes have been analyzed for the size of their pfxa3 hybridizing Pst I fragment. Of these the largest and smallest were analyzed for the size of their Sau 3AI (222) to Nhe I (509) fragment (*13*) which corresponded to 65 and 15 copies of the repeat, respectively. The repeat length for XTY26 of 43 copies therefore falls within the normal range for the p(CCG)$_n$ repeat, yet XTY26 was cloned from a fragile X–expressing hybrid cell line, X.3000.1 (*5*). We therefore undertook a similar blotting analysis of X.3000.1 DNA with the pfxa3 probe in order to ascertain the basis for this apparent discrepancy. This analysis showed that the p(CCG)$_n$-containing sequence is about 900 bp larger in X.3000.1 than its cloned counterparts (XTY26, λ5, and pfxa1) (Fig. 4B). This further suggests the p(CCG)$_n$ repeat is unstable during cloning as well as in fragile X pedigrees. The nature of the repeat region may therefore only be determined beyond doubt by direct genomic sequencing, if indeed this is technically feasible on such a GC-rich region.

In situ hybridization studies further demonstrated that the p(CCG)$_n$ repeat and the fragile site coincide. Five probes were labeled with biotin and used for fluorescent in situ hybridization to metaphase chromosomes expressing the fragile X. Two of the probes (λ5 and pfxa1) contain the p(CCG)$_n$ repeat and its flanking DNA. Of the other three, pfxa5 is proximal to the fragile site, and pfxa6 and pfxa7 are distal to it (Fig. 2A). The results (Table 1) show that the probes that flank the repeat also flank the fragile site. These in situ hybridization results were a little surprising since the level of resolution achieved here was greater than had been obtained previously (*5*) where unequivocal localization of whole λ clone probes over a region of 15 kb around the fragile site could not be obtained. The resons for achieving high resolution are unclear but could relate to the smaller sizes of the probes used. The in situ hybridization studies were carried out blindly.

We have concluded from all the experimental evidence that the unstable DNA sequence which characterizes the fragile X genotype maps to the p(CCG)$_n$ trinucleotide repeat. We have demonstrated that normal X chromosomes have about 40 ± 25 copies of p(CCG)$_n$ and that within these limits the sequence is a stable DNA polymorphism. The fragile X genotype is characterized by an increased amount of unstable DNA that maps to the repeat. Most of this unstable DNA and indeed most of the repeat in normal X chromosomes is lost during cloning and DNA amplification by PCR; thus, its exact nature must remain

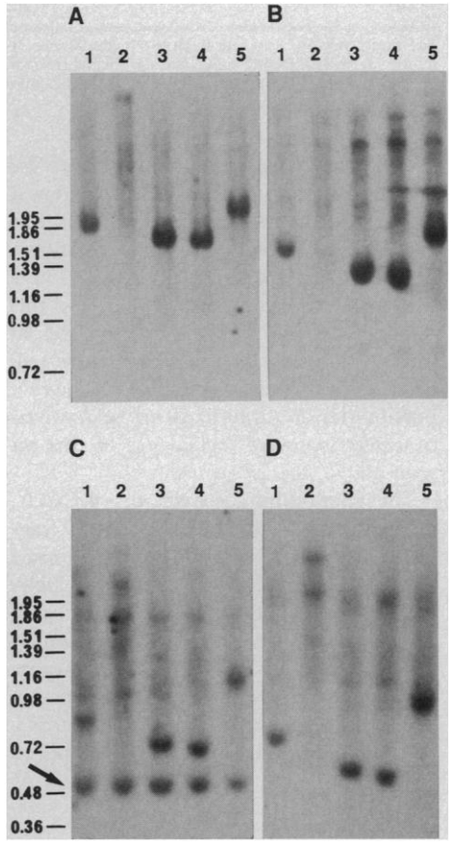

Fig. 3. Southern blot analysis with fragile X–affected and normal males. Total genomic DNA from lymphocytes was extracted and purified. A portion of each sample (10 μg) was digested to completion with (**A**) Pvu II and Bam HI, (**B**) Pvu II and Nhe I, (**C**) Sau 3AI, (**D**) Pst I and Rsa I. Lanes 1, 2, and 5 are from affected males. Lanes 3 and 4 are from normal males. Samples were separated by electrophoresis on a 1.3% agarose gel and transferred to Hybond N$^+$ blotting membrane (Amersham). The probe pfxa4 was ^{32}P-labeled by random priming and hybridized to the blots for 12 hours at 65°C in 7% SDS and 0.5M phosphate buffer (pH 7.0). The blots were washed at 70°C for 20 min in 0.2× SSC containing 0.1% SDS and exposed to Xomat XK-1 film (Kodak) for 12 hours at −80°C in the presence of two intensifying screens. The probe pfxa4 has an internal Sau 3AI site, therefore this probe hybridizes to two Sau 3AI fragments, one containing the unstable p(CCG)$_n$ sequence and the other a constant band that is proximal to the repeat. This band indicates that the size variation in the variable and unstable band is not due to electrophoretic conditions, sample variation, or incomplete digestion. The band of hybridization at 500 bp identified by an arrow in (C) served as an internal control. Molecular size marker is Eco RI–digested Spp 1 phage.

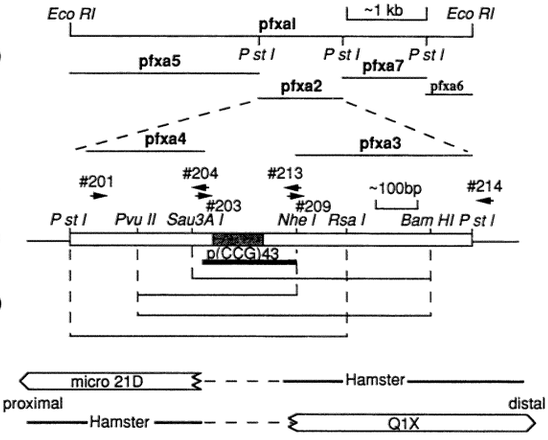

Fig. 2. Schematic diagram demonstrating localization of the unstable and variable region to the p(CCG)$_n$ repeat. (**A**) Location and identification of pfxa probes used for Southern blot analysis and in situ hybridization; (**B**) location of oligodeoxyribonucleotide primers used in PCR to identify the unstable and variable region and to obtain PCR product for sequence analysis; (**C**) informative restriction endonuclease sites flanking the trinucleotide repeat; (**D**) data from Southern blot analysis localizing the unstable and variable region between the Sau 3AI and the Nhe I restriction enzyme sites. The horizontal lines represent DNA fragments from restriction endonuclease digests that hybridized to pfxa4 and showed size variability outside the normal range in fragile X genotype lymphocyte DNA samples. The heavy line below the shaded p(CCG)$_{43}$ repeat is the area of instability defined by PCR and Southern blot analysis. (**E**) Schematic diagram representing the hybrid cell lines micro21D and QIX and their approximate break points identified by Southern blot and PCR analysis.

Table 1. Location of signal for various probes in relation to the fragile site at Xq27.3.

Probe	Position of signal in relation to fragile site (number of chromosomes)			
	Proximal	Central	Distal	Proximal and distal
λ5	10	2	30	3
pfxa1	5	0	5	0
pfxa5	12	1	2	1
pfxa7	4	3	11	0
pfxa6	1	1	8	2

Sequential metaphase spreads from two fragile X males were examined until at least ten X chromosomes expressing the fragile site and exhibiting signal from probe hybridization had been scored. The position of the signal was scored as proximal, central (that is, overlying the gap in the chromosome), or distal to the fragile site.

speculative. There are two possibilities. We favor the interpretation that the repeat sequence becomes amplified; however, it is also possible that some other unknown sequence is inserted into the repeat and is constantly lost during manipulation of this region. For the sake of convenience, in further discussion the increase in size will be referred to as an amplification. The two possibilities are only likely to be resolved by direct genomic sequencing of this region.

The relationship between the $p(CCG)_n$ repeat and cytological expression of the fragile site is not clear. Proof that this sequence can produce a fragile site will require DNA transfer experiments of either the cloned region or of synthetic $p(CCG)_n$ polymer. The fragile site is expressed in chromosomes prepared from cells cultured in media in which there is either an excess or a deficiency of thymidine and appears to be due to under-replication of the DNA at the fragile site (*10*). Excess thymidine inhibits ribonucleotide reductase and results in a deficiency of dCTP for DNA synthesis. It is not difficult to see how $p(CCG)_n$ might have difficulty replicating under such conditions. A deficiency of thymidine was previously considered to be another primary nucleotide pool perturbation directly responsible for fragile site expression but, if the fragile site is truly amplified $p(CCG)_n$, this is unlikely. However, thymidine depletion is accompanied by dGTP depletion: there is some evidence that addition of guanosine to cultures deficient in thymidine inhibits fragile site expression (*11*).

The composition of the repeat sequence would render its replication sensitive to depletion of either dCTP or dGTP, resulting in single-stranded DNA which would fail to package for mitosis and thus appear as a fragile site. A further correlation between the length of the repeat and the likelihood of under-replication is also evident. In normal pedigrees we have observed minor variations in the length of the repeat corresponding to ~40 ± 25 copies. These length polymorphisms are inherited in a stable, co-dominant, Mendelian manner. In fragile X–syndrome pedigrees the $p(CCG)_n$ repeat is amplified well above the normal range in both carriers and affected individuals. In addition, the repeat sequences exhibit instability and are generally larger in affected members of a pedigree than their unaffected carrier relatives who usually do not express the fragile site. The increase in repeat length in fragile X individuals compared with normals, together with the observed effects of nucleotides on fragile site expression, lends support to the idea that fragile site expression is a consequence of under-replication of the amplified $p(CCG)_n$ repeat sequence.

Copies of this repeat are found elsewhere in the genome—of particular note are the breakpoint cluster region (BCR) gene promoter and the androgen receptor gene—however, the reported lengths of these sequences are all less than ten copies (*12*), whereas the sequence in XTY26 contains 43 copies.

These studies do not explain why such an unstable sequence would be maintained, let alone further amplified in fragile X pedigrees. Neither do they address the issue of methylation of the region in fragile X syndrome individuals (*6, 7*). The composition of the unstable sequence, which contains many targets for methylation, provides a link between instability seen in fragile X genotype and the methylation of this region associated with the fragile X syndrome phenotype.

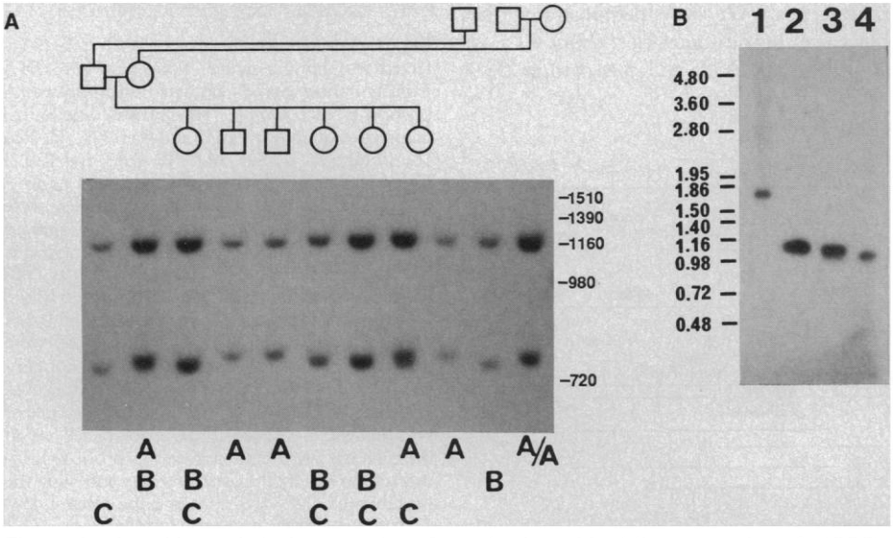

Fig. 4. Southern blot analysis demonstrating (**A**) co-dominant Mendelian segregation of $p(CCG)_n$ repeat polymorphism in CEPH family 1349; DNA was digested to completion with Sau 3AI and treated as for Fig. 2 except that pfxa3 was used as the probe. (**B**) Instability of $p(CCG)_n$ repeat region during cloning. Approximately equimolar amounts of each sample were digested to completion with the restriction endonuclease Pst 1, separated by electrophoresis on a 1% agarose gel, and treated as for Fig. 2. Lane 1, X3000.1; lane 2, XTY26; lane 3, 5; and lane 4, pfxa1. Size markers are indicated in base pairs.

REFERENCES AND NOTES

1. G. R. Sutherland and F. Hecht, *Fragile Sites on Human Chromosomes* (Oxford Univ. Press, New York, 1985).
2. E. Kremer *et al.*, *Am. J. Hum. Genet.*, in press.
3. D. Heitz *et al.*, *Science* **251**, 1236 (1991).
4. S. Warren *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 3856 (1990).
5. S. Yu *et al.*, *Science* **252**, 1179 (1991).
6. A. Vincent *et al.*, *Nature* **349**, 624 (1991).
7. M. V. Bell *et al.*, *Cell* **64**, 861 (1991).
8. S. Lindsay and A. P. Bird, *Nature* **327**, 336 (1987).
9. E. Kremer *et al.*, unpublished data.
10. G. R. Sutherland, *Mutat. Res.* **200**, 207 (1988).
11. _____ and E. Baker, *Am. J. Med. Genet.* **23**, 409 (1986).
12. R. Williamson *et al.*, *Cytogenet. Cell Genet.* **55**, 457 (1990).
13. S. Yu, unpublished data.

14. We thank H. Scott for assistance with sequence analysis, H. Kozman and L. Z. Chen for help in figure design, and A. Naion and J. Westley for manuscript preparation. Supported by grants from the Australian National Health and Medical Research Council and the Adelaide Children's Hospital Foundation.

29 April 1991; accepted 22 May 1991