Am. J. Hum. Genet. 53:1217–1228, 1993

Analysis of a CGG Sequence at the FMR-1 Locus in Fragile X Families and in the General Population

K. Snow,* L. K. Doud,* R. Hagerman,† R. G. Pergolizzi,‡ S. H. Erster,‡ and S. N. Thibodeau*

*Mayo Clinic and Foundation, Rochester, MN; †Denver Children's Hospital, Denver; ‡North Shore University Hospital—Cornell University Medical College, Manhasset, NY

Summary

In this study, we have characterized a CGG repeat at the FMR-1 locus in more than 100 families (more than 500 individuals) presenting for fragile X testing and in 247 individuals from the general population. Both Southern blot and PCR-based assays were evaluated for their ability to detect premutations, full mutations, and variability in normal allele sizes. Among the Southern blot assays, the probes Ox1.9 or StB12.3 with a double restriction-enzyme digest were the most sensitive in detecting both small and large amplifications and, in addition, provided information on methylation of an adjacent CpG island. In the PCR-based assays, analysis of PCR products on denaturing DNA sequencing gels allowed the most accurate determination of CGG repeat number up to approximately 130 repeats. A combination of a Southern blot assay with a double digest and the PCR–sequencing-gel assay detected the spectrum of amplification-type mutations at the FMR-1 locus. In the patient population, a CGG repeat of 51 was the largest to be stably inherited, and a repeat of 57 was the smallest size of premutation to be unstably inherited. When premutations were transmitted by females, the size of repeat correlated with risk of expansion to a full mutation in the next generation. Full mutations (large repeats typically associated with an abnormal methylation pattern and mitotic instability) were associated with clinical and cytogenetic manifestations in males but not necessarily in females. In the control population, the CGG repeat ranged from 13 to 61, but 94% of alleles had fewer than 40 repeats. The most frequent allele (34%) was a repeat of 30. One female had an allele (61 repeats) within a range consistent with fragile X premutations, while two other individuals each had a repeat of 52. This suggests that the frequency of unstable alleles in the general population may be approximately 1%.

Introduction

Recently, a number of investigators have identified the gene (i.e., the FMR-1 gene) and gene mutations responsible for fragile X syndrome (Bell et al. 1991; Kremer et al. 1991; Oberlé et al. 1991; Verkerk et al. 1991; Yu et al. 1991). As a result of these studies, there appear to be two areas of the gene that have particular clinical relevance: a CGG repeat sequence at the 5′ end of the FMR-1 gene and a CpG island upstream of the CGG repeat. In normal individuals, the number of CGG repeats is variable, with a range of approximately 5 to 50 (Fu et al. 1991; Kremer et al. 1991). In contrast, individuals affected with fragile X syndrome have DNA amplification consistent with several hundred to several thousand copies of the CGG repeat, while phenotypically normal carriers have up to approximately 200 copies (Fu et al. 1991; Oberlé et al. 1991). In addition, the CpG island upstream from the CGG repeat is abnormally methylated in affected individuals, the effect of which appears to be reduced transcription of the FMR-1 gene (Pieretti et al. 1991; Sutcliffe et al. 1992).

The discovery of the causative defect for fragile X syndrome has quickly led to the development of a number of DNA-based assays, including both Southern and PCR-based approaches, that can detect the FMR-1 amplification-type mutation (Fu et al. 1991; Hirst et al. 1991; Kremer et al. 1991; Rousseau et al. 1991; Brown et al. 1992; Erster et al. 1992; Pergolizzi et al. 1992). These assays are particularly useful for the identifica-

Received June 18, 1993; revision received August 3, 1993.

Address for correspondence and reprints: Dr. S. N. Thibodeau, Laboratory Genetics, Mayo Clinic, 200 First Street S.W., Rochester, MN 55905.

0002-9297/93/5306-0008$02.00

tion of normal transmitting males (NTMs) and carrier females who are frequently negative for Xq27.3 fragility (fra(X)). In addition, molecular studies offer the possibility of performing more reliable prenatal testing (Sutherland et al. 1991).

As an increasing number of cases are being analyzed for fragile X syndrome by direct mutation testing, a number of questions have arisen relating to the interpretation of results. For example, several reports have described an apparently normal CGG repeat in the setting of classical fragile X syndrome, including the expression of fra(X) (Nakahori et al. 1991; Dennis et al. 1992; Oberlé et al. 1992; De Boulle et al. 1993). Do these exceptional cases have an alternative FMR-1 mutation or possibly a mutation at another locus? Second, it has been shown that the numbers of CGG repeats that constitute a premutation versus a normal allele are very close (Fu et al. 1991; Macpherson et al. 1992). What is the upper limit of normal, the lower limit of a premutation, and how should repeat numbers in a borderline zone be interpreted? Third, in prenatal testing, can DNA-based studies be used for both carrier testing and prenatal diagnosis? Finally, what is the frequency of mutations in the general population, and how frequently do we expect to find a mutation in a spouse of a carrier of fragile X syndrome?

In this study, 116 families referred for fragile X testing were examined by a number of DNA-based assays. The aim of the study was to identify the best approach for molecular analysis of fragile X families and to address some of the questions and issues that have arisen over the use of the molecular assays. To accomplish this, the (CGG)n repeat number and the methylation status of the 5′ CpG island of the FMR-1 gene were studied in affected and unaffected carriers of fragile X syndrome and in individuals from the general population.

Subjects and Methods

Subjects

The patient population included 116 families (530 individuals). Of these families, 40 had been previously studied by linkage analysis and by direct mutation analysis using Southern blotting and the probe Ox1.9 (Snow et al. 1992). Individuals in the previously studied families were defined as learning disabled (LD) if they had received special education services from their school, while mental retardation (MR) was defined as an IQ score of 69 or less. Of the other 76 pedigrees, approximately two-thirds were referred for testing to rule out fragile X syndrome. Many rule-out cases represented sporadic incidences of developmental delay, mental impairment, or autistic behavior with negative, borderline, or untested fra(X). The remainder of cases were more suggestive for fragile X syndrome. This was based on documentation of X-linked MR, other clinical findings, and/or fra(X).

Control samples were obtained from a blood-donor population. Samples were labeled as male or female at the time of collection but were otherwise unmarked, so that it would not be possible to trace the identity of donors. Samples from 50 males and 197 females were included as controls.

Methods

Southern blot analyses.—High-molecular-weight DNA was isolated from peripheral blood leukocytes by using an Applied Biosystem 340A DNA extractor (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions. Purified DNA (2.5 μg) was digested to completion with the appropriate restriction endonucleases. For the double digests, DNA was digested for 2 h with *Hin*dIII alone, before addition of the second enzyme. Digested DNA was fractionated on 0.8% agarose gels and then was transferred to nylon membrane by the method of Southern. Probes Ox1.9 (from Dr. K. E. Davies, Oxford) and StB12.3 (from Dr. J.-L. Mandel, Strasbourg) were radiolabeled with ^{32}P-dCTP by the random-primed oligolabeling method (Amersham), while fxa 241 was supplied prelabeled with ^{32}P (Oncor). For each probe, membranes were hybridized overnight (in formamide buffer) and then were washed to 0.1 × SSC at 60°C. Typical autoradiography times at −70°C were 1 to 3 d.

PCR-based methods.—PCR products were analyzed either by Southern blotting (using a 5′$[CGG]_5$3′ probe) or on denaturing DNA sequencing gels. PCR primers and reaction conditions were as described elsewhere (Pergolizzi et al. 1992), except that reactions also included 10% glycerol and thermocycler parameters were 95°C for 5 min, then 30 cycles of denaturation at 97°C for 30 s, annealing at 55°C for 60 s, and elongation at 72°C for 60 s, followed by a 10-min final extension at 72°C. The presence of glycerol increased the yield of PCR products, particularly for alleles with large premutations or full mutations. For the analysis of products on a sequencing gel, PCR reactions also included ^{32}P-dCTP, and radiolabeled products were diluted 1:1 with 95% formamide loading buffer and then were separated on 6% denaturing sequencing gels (Fu et al. 1991). An M13 sequencing ladder (G+A reactions mixed together) was used as a size marker, and, for a CGG repeat of 30, the product size was 203 bp. Autoradiography of

dried sequencing gels was for 1 to 2 d at −70°C. For the analysis of products by Southern blotting, PCR products were separated on 1.8% agarose and then were transferred to nylon membrane. It was discovered that it was most important to perform electrophoresis of PCR products in the absence of ethidium bromide, since its presence (when included to visualize DNA size markers) inhibited the migration of DNA. Labeling and hybridization of the oligonucleotide probe (5′[CGG]$_5$3′) were performed as described elsewhere (Pergolizzi et al. 1992), and membranes were washed in 6 × SSC, 0.2% SDS for 15 min at 42°C and then in 2 × SSC, 0.2% SDS for 30 min at 56°C.

Cytogenetic testing.—According to established guidelines for cytogenetic testing of fragile X syndrome (Jacky et al. 1991), cytogenetic analyses were considered positive when 4% or more of the cells were shown to contain a fra(X) chromosome. Fragility between 1% and 4% was considered borderline, while values less than or equal to 1% were considered negative.

Results

Comparison of Methodologies Used to Examine the CGG Repeat

In this study, several methodologies were used to detect amplification within the CGG repeat, including several Southern blot approaches and two PCR-based methods. The Southern blot analyses utilized a variety of probe/enzyme combinations, including (*a*) Ox1.9 or StB12.3 along with either a single-enzyme digest (*Hin*dIII) or a double-enzyme digest (*Hin*dIII+*Eag*I or *Hin*dIII+*Nru*I) and (*b*) the DNA probe fxa241 along with a single-enzyme digest (*Pst*I). The major difference between the two PCR-based assays was essentially the method of product detection: in one the product was detected directly on a sequencing gel, while in the other the product was detected after blotting and hybridization with a radiolabeled (CGG)$_5$ probe. Figure 1 depicts representative results obtained by various methodologies in the analysis of both small (CGG)n amplifications (i.e., premutations) and large (CGG)n amplifications (i.e., full mutations). In the evaluation of the relative efficiencies of each of the assays for detecting the spectrum of CGG amplifications encountered in fragile X families, many individuals were examined by three or more different assays. As described in table 1, each of the assays had limitations in the range of mutations that it could detect.

Summary of Results

Table 2 presents the results of analysis of the CGG repeat in 458 individuals (from 74 families) referred for fragile X testing. This population encompasses the 40 families previously studied with Ox1.9/*Hin*dIII (Snow et al. 1992). This earlier study demonstrated that Southern blot analysis using the probe Ox1.9 and *Hin*dIII digestion could detect molecular alterations associated with fragile X syndrome, since a mutation was detected in 66 (99%) of 67 clinically affected males, in 12 (92%) of 13 transmitting males, and in 95 (85%) of 112 carrier females. However, it was questioned whether a lack of assay sensitivity contributed toward the inability to detect mutations in *all* affected individuals, obligate carriers, and individuals predicted, by linkage, to be carriers. Therefore, many individuals included in the previous study were reanalyzed using one or more of the alternative assays described above. As can be seen from table 2, several results were still discrepant from those predicted either by the clinical phenotype or by linkage analysis. These discrepant cases, several of which were discussed in our previous report (Snow et al. 1992), are summarized in table 3 and in the Discussion of the present report. In addition to the cases described in tables 2 and 3, a mutation was not detected for 42 pedigrees (encompassing 72 individuals) referred for testing to rule out fragile X syndrome. These sporadic cases of developmental delay, LD, or MR with negative fra(X) will not be discussed further in the present report, since they most likely do not represent fragile X syndrome.

A comparison between cytogenetic and molecular analyses demonstrated that 74 of 75 fra(X)-positive males and 43 of 44 fra(X)-positive females had a CGG amplification. The two exceptional cases are described in table 3. On the other hand, of individuals who were found to have a CGG amplification and who had been tested for fra(X), 13 (15%) of 87 males and 81 (65%) of 124 females had either zero or less than 4% fra(X). Most of these individuals carried a premutation, but six males and seven females had a full mutation and mental impairment with negative or borderline fragility. As shown in figure 2 (lanes 3, 5, and 6), amplification in these cases ranged from approximately 0.6 kb to 2 kb and was associated with incomplete methylation of the CpG island.

Correlation of Genotype versus Phenotype

In this study, features of the full mutation included a CGG amplification of more than approximately 0.6 kb of DNA, abnormal methylation of the adjacent CpG island, and somatic instability of the (CGG)n region. Of those males who had a full mutation, 68 (96%) of 71 had MR, and the remaining three had LD. In contrast,

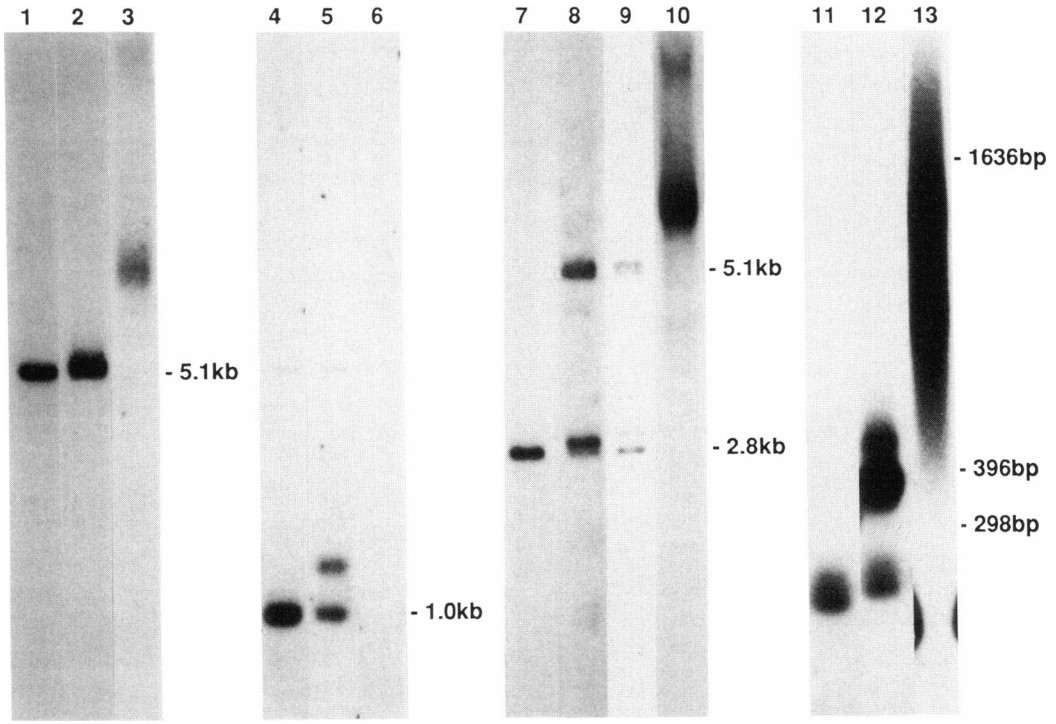

Figure 1 Analysis of similar sizes of (CGG)n by different assays. Lanes 1 to 3, Ox1.9/*Hin*dIII Southern blot. Lanes 4 to 6, fxa241/*Pst*I Southern blot. Lanes 7 to 10, Ox1.9/*Hin*dIII+*Eag*I Southern blot. Lanes 11 to 13, PCR analyzed by Southern blot. Lanes 1, 4, 7, and 11 represent analysis of a normal male who has 37 CGG. Lanes 2, 5, 9, and 12 represent analysis of females who have a premutation that is larger than their normal allele by 167 bp, 200 bp, 164 bp, and 164 bp, respectively. Lane 8 shows a result obtained for a female who has a difference of 93 bp between her two normal alleles (i.e., 20 and 51 CGG). Lanes 3, 6, 10, and 13 show results obtained for males who carry a full mutation. All (CGG)n numbers were determined by a PCR–sequencing-gel assay.

only 3 (8%) of 36 females with a full mutation had MR, 29 (81%) of 36 had LD, and 4 (11%) of 36 had no mental impairment. A full mutation was also documented for an additional 20 males and 30 females for whom there was inadequate information on mental status.

Premutations identified in this study had amplification of less than 0.5 kb, and the CpG island did not show abnormal methylation. Of those individuals found to have a premutation, the vast majority were clinically and cytogenetically normal. Nevertheless, in agreement with previous findings (Rousseau et al. 1991), LD was documented for 5 (7%) of 74 females with a premutation and in 1 (7%) of 14 males with a premutation. However, these frequencies are similar to those expected for mild mental impairment in the general population, on the basis of a mean IQ score of 100 and an SD of 15. Therefore, it is unlikely that the FMR-1 premutation is a cause of LD in individuals included in this study.

In 20 (25%) of 91 males and 6 (9%) of 66 females who had a full mutation, a premutation was also detected; that is, these individuals had a mosaic pattern. Furthermore, an additional three males had a normal size allele as well as a full mutation. Results of karyotype analysis of these individuals were unavailable; however, molecular studies suggested that at least two of these individuals were not likely to be 47,XXY. First, Southern blot analysis using a *Hin*dIII+*Eag*I double digest did not detect any inactive normal X fragment, as expected for an XXY individual. Second, analysis of one of the mosaic individuals and his parents by the PCR–sequencing-gel assay showed that his normal size allele was smaller than either of the normal alleles carried by his parents. For the third mosaic male with a normal fragment, Southern blot analysis with a double digest detected faint bands (relative to the intensity of the full mutation) at positions expected for normal active and normal inactive alleles, suggesting that he could be mos47,XXY. The PCR–sequencing-gel assay indicated

Table 1

Limitations of Assays Used to Analyze the FMR-1 CGG Repeat

Assay	Limitations	No. of Individuals Studied
Ox1.9/*Hin*dIII	Lower limit of detection is a difference of approximately 150 bp between alleles	476
fxa241/*Pst*I	Full mutations may be difficult to detect Females who have 2 normal alleles that differ in size by 25 or more repeats may appear to have a premutation	234
Ox1.9/*Hin*dIII+*Eag*I or StB12.3/*Hin*dIII+*Nru*I	Lower limit of detection is a difference of approximately 90 bp between alleles Females who have two normal alleles that differ in size by 25 or more repeats may appear to have a premutation	222
PCR–Southern blot	Heterozygous females often demonstrate three bands (possibly heteroduplex formation) Inefficient PCR amplification of full mutations causes misleading results for some full mutation females and for some mosaic individuals	154
PCR–sequencing gel	Upper limit of detection is an amplification of approximately 300 bp Individuals mosaic for a premutation plus a full mutation may show amplification of only the premutation	385

that the third mosaic male's normal size allele was the same size (i.e., 30 repeats) as his mother's normal allele but different than his father's allele. The phenotype in individuals with a mosaic mutation was similar to that detected in individuals with a full mutation: each of the mosaic males and one mosaic female had MR, and the other five mosaic females had LD. The levels of fra(X) for individuals with a mosaic pattern were also not significantly different ($P=.4$ for males; $P=.2$ for females) from those for individuals with a full mutation pattern. DNA from mosaic individuals was analyzed by the PCR–sequencing-gel assay, and a premutation (or normal) size allele was detected in 9 of 29 cases. The premutation in the other 20 cases was probably too large to be efficiently amplified by PCR, and/or the PCR product did not enter the sequencing gel matrix. Mothers of 25 mosaic individuals were studied to determine their mutation type: 19 had a premutation, 5 had a full mutation, and 1 had a mosaic pattern.

Progression of Premutation to Full Mutation

The likelihood that a premutation will progress to a full mutation, from one generation to the next, was studied for both carrier females and NTMs. Figure 3 shows the frequency of conversion to a full mutation in 106 offspring of females who have premutations with a size between 57 and 130 repeats. The smallest repeat size to be converted to a full mutation in the next generation was 61, and the largest repeat size to remain as a premutation in the next generation was 90 (115 in the offspring). Figure 3 attempts to correct for ascertainment bias by removing data for the affected proband in each family. Without correction for ascertainment bias, risks were altered both for transmission of repeat sizes 71 to 80 (10 [43%] of 23 offspring had a full mutation) and for transmission of repeat sizes 81 to 90 (32 [86%] of 37 offspring had a full mutation).

In contrast to transmission of premutations by females, transmission of premutations by males resulted in little change to the size of CGG amplification. Of 15 daughters of NTMs, 4 had a smaller premutation compared with that in their father, 5 had the same size premutation, and 6 had a larger premutation (data not shown). The largest decrease was from 150 to 105 repeats, and the largest increase was from 72 to 95 repeats.

Characterization of a Normal CGG Repeat

In families presented for fragile X testing, meiotic stability was observed for CGG repeats of up to 51. For

Table 2

Results of Molecular Testing Performed on Fragile X Syndrome Family Members

	No. of Individuals, by CGG Repeat Amplification Status	
	Present	Absent
Males:		
Affected	86	5[a]
Carrier, by linkage	12	2[a]
Normal, by linkage	0	36
Not at risk	0	48
Unknown[b]	6	14
Females:		
Obligate carrier	60	0
Affected	42	1[a]
Carrier, by linkage	22	2[a]
Normal, by linkage	0	63
Not at risk	0	2
Unknown[b]	37	20

[a] Described in table 3.

[b] Linkage uninformative or no clinical/cytogenetic information.

repeats of 43, 47, 50, and 51, the number of meioses in which the CGG repeat was stably inherited was 3, 2, 1, and 5, respectively. The alleles with a repeat of 51 were from one family that underwent fragile X testing but in which neither instability of the CGG repeat nor abnormal methylation of the CpG island was detected (fig. 4). The proband in this family had MR and dysmorphic features and did not show fra(X) (0/200). Other family history included LD in a sister (who was also fra(X) negative) and LD in a maternal uncle. Analysis of the CGG repeat suggested that the clinical phenotype was not associated with an abnormality of the FMR-1 gene: first, a clinically normal brother and a normal maternal uncle of the proband also carried the allele with 51 repeats; and, second, the LD sister did not inherit the 51 repeat allele.

The distribution of allele sizes was also studied in a control population of blood donors. As shown in figure 5, alleles from 50 male and 197 female control individuals ranged in size from 13 to 61 repeats, with 30 being the most common allele (35%). Alleles with 45 or more repeats were observed for eight individuals, and the highest repeat sizes observed were 52 and 61 in females and 52 in males. Among females, the frequency of homozygosity was 18%. Observed homozygosity did not significantly differ from the expected level, 16.8%, calculated from allele frequencies in the control population ($\chi^2 = 0.1$; $P=.7$).

Discussion

The mutation associated with fragile X syndrome has been characterized and involves variable amplification of a CGG repeat within the FMR-1 gene (Bell et al. 1991; Kremer et al. 1991; Oberlé et al. 1991; Verkerk et al. 1991; Yu et al. 1991). In a previous study, which used the probe Ox1.9 and *Hin*dIII-digested DNA to analyze the mutation in 40 fragile X families, 15% of obligate carriers of fragile X syndrome were shown to have negative or equivocal results (Snow et al. 1992). This was despite the presence of a (CGG)n amplification mutation in an affected family member. Therefore, in the present study, additional methods for analyzing the CGG repeat were evaluated.

The most likely explanation for our previous negative results for obligate carrier females was inability of the assay to resolve very small amplifications (i.e., small premutations). Thus, we paid particular attention to the analysis of such alleles. With Southern blot methodology, it was found that each of the double-digest assays and the fxa241/*Pst*I combination could detect almost all of the premutations. However, in some cases these assays were unable to distinguish between a female who had alleles at the lower and upper extremes of normal and a female who had a small premutation. This is illustrated in lane 8 of figure 1, where a female who had a difference of 93 bp between her two normal alleles (CGG repeats of 20 and 51) could, by Southern blot assays, be mistaken to have a premutation. This example and others (Macpherson et al. 1992) therefore highlight one of the needs for an accurate assessment of CGG repeat number in the normal to premutation range.

It is likely that some overlap exists between the sizes of CGG that are meiotically stable (i.e., "normal") and the sizes that are meiotically unstable (i.e., premutations). For example, meiotic stability has been demonstrated for alleles with up to 51 repeats (fig. 4), whereas alleles with repeats of 52 or more appear to be unstable (Fu et al. 1991). Thus, the critical region of overlap between normal alleles and premutations seems to be approximately 50, but, so far, only a small number of meioses have been studied in this vicinity, so the range of overlap is uncertain. To determine the likely carrier status of individuals who carry a borderline repeat (e.g., between 45 and 55), it is necessary to study multiple family members to document the meiotic stability (e.g.,

Table 3

Discrepancies between Results of DNA Testing and Phenotype or Linkage Studies

Category and Pedigree No.	Pertinent Data
Affected male with normal CGG:	
FX 113	Described previously by Snow et al. (1992), 4% fra(X), MR male
FX 112	Described previously by Snow et al. (1992), 1/270 fra(X), typical Martin Bell phenotype for male
FX 132	fra(X) negative, MR, macroorchidism, large ears
FX 161	4/100, 0/100 fra(X), LD, large ears, hand flapping, hand biting, hyperextensible joints
FX 200	fra(X) negative, LD, strong family history of XLMR, large ears, long face, hyperextensible joints
Affected female with normal CGG:	
FX 138	5% fra(X), MR, hyperactive, hand flapping
Carrier, by linkage with normal CGG:	
FX 112	Sister and nephew of affected male, CGG normal in affected male
FX 113	Brother of affected male, CGG normal in affected male
FX 25	Daughter of a female with a CGG amplification, carrier risk 98% with RN1A and 1A1

fig. 4). To accomplish this, the assay method needs to be able to resolve a difference of even a single CGG repeat.

Two PCR-based assays were evaluated for their ability to accurately determine allele sizes in the high normal to low premutation range. In one PCR assay, PCR products were separated under nondenaturing conditions and were detected by Southern blotting. While this procedure gave strong signals for all premutations, female carriers often demonstrated the presence of three bands, which precluded accurate determination of allele sizes. In a controlled experiment, a third band was also observed when PCR products from a normal male and from a male with a premutation were mixed, heated to denature, and allowed to slowly cool (data not shown). Therefore, it is likely that the extra band represents an alternative DNA structure (e.g., heteroduplex). In contrast, separation of products on denaturing sequencing gels prevented formation of secondary DNA structures, and allele sizes could be determined by comparison with an M13 sequencing ladder. The ability of the PCR–denaturing-polyacrylamide-gel assay to resolve alleles differing by a single repeat was supported by agreement, in the number of CGG repeats, between observed and expected heterozygosity. In addition, several alleles in the normal to low premutation range have been directly sequenced, and results confirm the number of repeats estimated by the PCR–sequencing-gel assay (data not shown). Thus, the PCR–sequencing-gel assay was preferred for accurate sizing of alleles in the normal to premutation range. However, both of the PCR-based assays had limitations in the detection of full mutations, probably because of both inefficient PCR amplification of very large sequences of DNA and, in the case of the PCR–sequencing-gel assay, inefficient entry into the gel matrix. This was a particular problem for those individuals who had a mosaic pattern by Southern blot analysis. For example, 6 of 23 mosaic males tested by the PCR–sequencing-gel assay gave results indistinguishable from those obtained for premutation males. Moreover, an additional two mosaic males demonstrated the presence of a normal allele by the PCR–sequencing-gel assay and were indistinguishable from normal males.

Together, results obtained from both the Southern blot assays and the PCR-based assays suggest that neither of these methodologies alone can detect the spectrum of CGG repeat sizes encountered in fragile X testing. However, when both Southern blot analysis using StB12.3 or Ox1.9 along with a double digest (e.g., *Hin*dIII+*Nru*I) and the PCR–sequencing-gel method are used together, the two assays complement each

other to provide reliable results for routine analysis of specimens referred for fragile X testing (for confirmation of diagnosis or determination of carrier status).

Three additional points regarding the PCR-based assays are discussed below. First, a novel finding in using Southern blotting to detect PCR products was that inclusion of ethidium bromide in the agarose gel led to almost complete inhibition of migration of PCR-amplified DNA in the agarose gel matrix. This is possibly related to an altered interaction between ethidium bromide and DNA containing 7-deaza GTP. Evidence for an altered interaction is the lack of ethidium bromide staining of DNA containing the modified base (Latimer and Lee 1991). Second, sequences of the PCR primers used in the present study matched the published sequence data for FMR-1 as described by Verkerk et al. (1991), but the upstream primer contained a 2-bp mismatch vis-à-vis other published data (Fu et al. 1991). Because of this discrepancy, another oligonucleotide was synthesized to match the sequence described by Fu et al. (1991). Both primers gave PCR products of the same size, but greater yield was obtained using the upstream primer originally described by Pergolizzi et al. (1992) (data not shown). This result could be explained

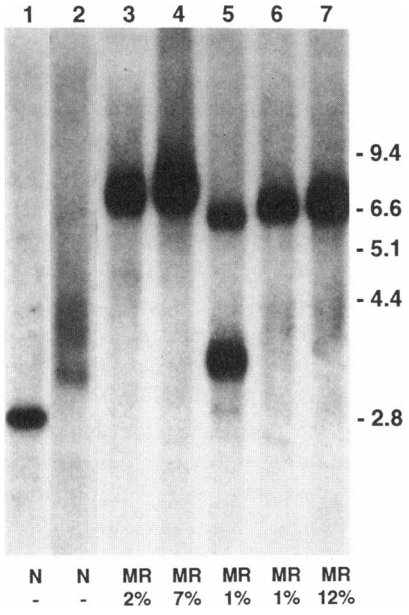

Figure 2 Variable patterns of methylation of the CpG island 5′ to FMR-1. All results were obtained using the StB12.3/*Hin*dIII+*Nru*I Southern blot assay. Lane 1, Normal control male. Lane 2, NTM. Lanes 3 to 7, Affected males. Mental status (N = no mental impairment; and MR = mental retardation) and levels of fragility at Xq27.3 are indicated below each lane.

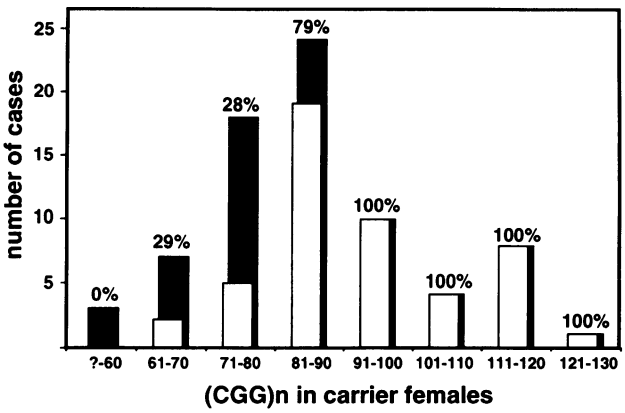

Figure 3 Risk for expansion of premutations to full mutations when transmission is by females. Blackened bars indicate the number of offspring who were found to have inherited an unstable allele from females who carry a premutation; and unblackened bars indicate the fraction of those offspring who inherited the unstable allele as a full mutation. Results are shown for size categories of the (CGG)n determined by a PCR–sequencing-gel assay.

by the effect of a mismatch in reducing the T_m of the upstream primer, which was calculated to be 76°C, compared with 70°C for the downstream primer (estimated $T_m = 2°C \times [A+T] + 4°C \times [G+C]$). Third, this study identified a repeat of 30 as being the most frequent allele in the general population. However, a previous study (Fu et al. 1991) identified 29 as being the most common repeat. This discrepancy is most likely due to differences in interpretation of sizes of PCR products, rather than to a real difference between the populations studied. It can be seen that the patterns of "shadow bands" generated in the two studies differ (compare fig. 4 in the present study with fig. 3 in Fu et al. 1991). This variability, which may lead to interpretation differences, can probably be attributed to the use of different PCR primers. The shadow bands are artifacts commonly seen in PCR amplification of microsatellites and probably arise because of replication slippage (Hauge and Litt 1993). This assay-dependent variability in repeat sizes needs to be recognized when results for related individuals are obtained from different laboratories. To eliminate size differences caused by assay variability (either interlab or intralab), it is recommended that analyses of related individuals be performed at the same time, with PCR products loaded into adjacent lanes in the gel.

One of the most important applications of these molecular methods is in carrier testing of individuals who have a family history of fragile X syndrome. Whereas

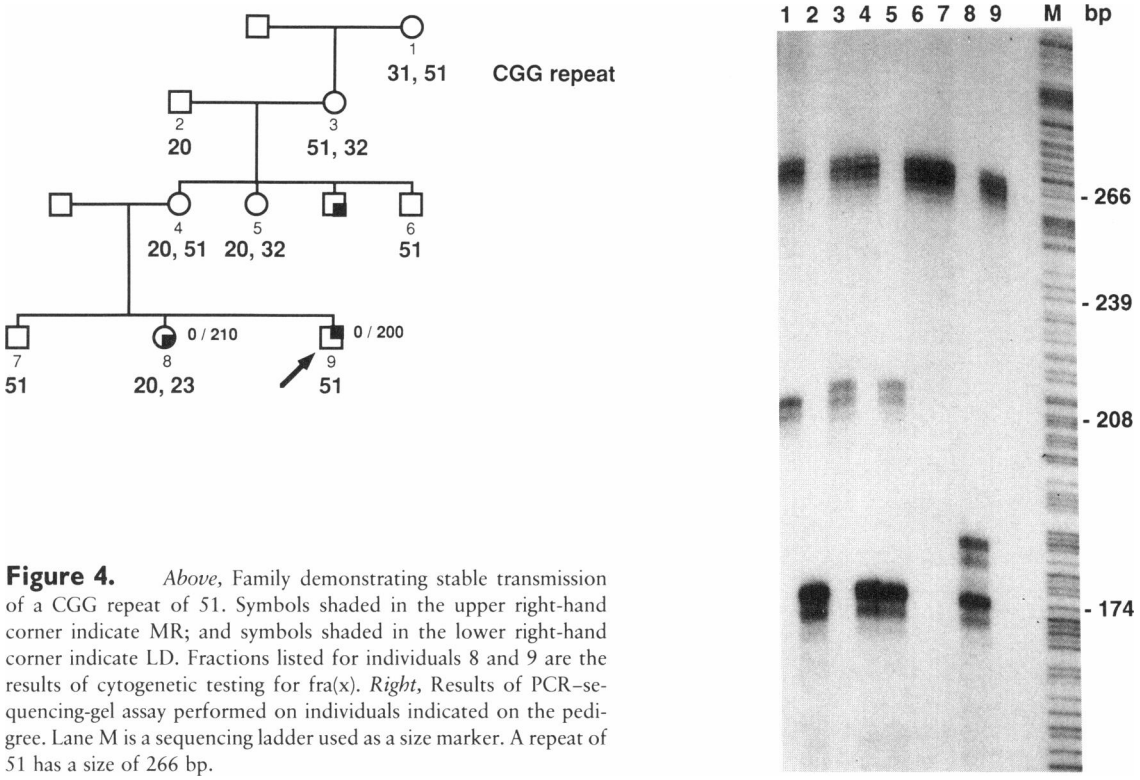

Figure 4. *Above*, Family demonstrating stable transmission of a CGG repeat of 51. Symbols shaded in the upper right-hand corner indicate MR; and symbols shaded in the lower right-hand corner indicate LD. Fractions listed for individuals 8 and 9 are the results of cytogenetic testing for fra(x). *Right*, Results of PCR–sequencing-gel assay performed on individuals indicated on the pedigree. Lane M is a sequencing ladder used as a size marker. A repeat of 51 has a size of 266 bp.

carrier females and NTMs are frequently negative for fra(X), these individuals do demonstrate the presence of a CGG amplification. Nonetheless, in performing carrier testing, it is important to initially document the presence of a CGG amplification in an affected family member or obligate carrier, because of the possibility of genetic heterogeneity. For example, in this study, a CGG amplification could not be demonstrated in the index case of 6 of 73 families referred for fragile X testing (tables 2 and 3). These individuals may have another type of FMR-1 mutation, such as a deletion (Gedeon et al. 1992; Wöhrle et al. 1992) or a point mutation (De Boulle et al. 1993). Alternatively, it is possible that a mutation at another locus could be responsible for the observed phenotype. The observation that some amplification-negative cases expressed either low levels of fragility or no fragility does not rule out a diagnosis of fragile X syndrome. For example, it has been shown that deletion at the FMR-1 locus may be associated with absence of fragility in some cases of fragile X syndrome (Gedeon et al. 1992; Wöhrle et al. 1992). In addition, this study identified six affected males and seven affected females who were fra(X) negative but who had a (CGG)n amplification. On the other hand, several recent reports have described individuals who express a high frequency of fragility at Xq27.3 but who do not demonstrate an abnormal CGG repeat (Nakahori et al. 1991; Dennis et al. 1992; Oberlé et al. 1992). These cases may involve fragility at nearby fragile sites —e.g., FRAXE or FRAXF—that can be distinguished from FRAXA by molecular cytogenetics (Sutherland and Baker 1992; Hirst et al. 1993).

The individual in pedigree Fx25 differs from the other discrepant cases in table 3, in that (*a*) a CGG amplification was identified in the family and (*b*) this individual has been predicted to be a carrier by linkage analysis. The pedigree was illustrated in our previous report (Snow et al. 1992), and at that time it was suggested that insensitivity of the Ox1.9/*Hin*dIII assay, a double-recombination event, or mutation reversion could account for the discrepancy. A preliminary analysis of closely linked flanking (CA)n repeats suggests that mutation reversion is the most probable explanation (data not shown).

Of importance in genetic counseling of carriers of fragile X syndrome is the question of whether testing

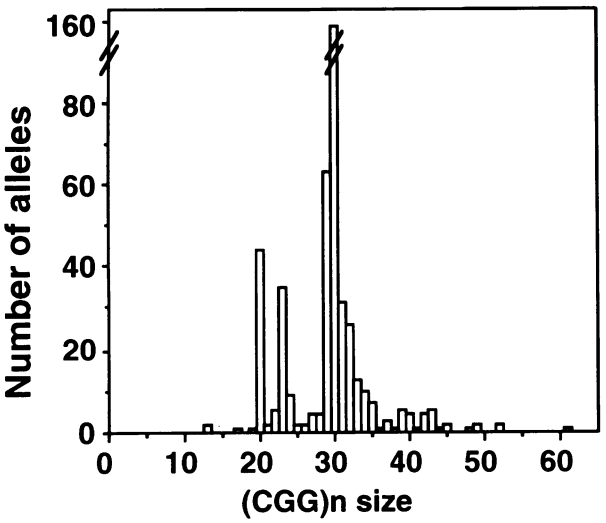

Figure 5 Frequency of FMR-1 (CGG)n sizes in the general population. The graph summarizes results obtained for 50 males and 197 females. (CGG)n sizes were determined by a PCR–sequencing-gel assay.

can help to provide an estimation of the risk that a carrier will have an affected child. When this question is rephrased in terms of the dynamics of the fragile X mutation, two issues need to be addressed: first, for a given size of premutation, what is the risk for expansion to a full mutation in the next generation? and, second, how does the CGG repeat size and/or the presence of abnormal methylation correlate with phenotype? For male carriers of a premutation (i.e., NTMs), expansion to a full mutation in daughters has not been observed, and thus the risk for this occurrence appears to be very small. Even for male carriers of relatively large premutations—i.e., between 150 and 200 repeats—the mutation remains as a premutation in daughters. However, for female carriers, the risk for expansion to a full mutation in offspring is dependent on the size of premutation being transmitted, as shown in figure 3. In predicting risks for expansion to a full mutation, our results are in good agreement with previous studies (Fu et al. 1991; Heitz et al. 1992), in demonstrating 100% risk for more than 90 repeats, less than 30% risk for fewer than 70 repeats, and an intermediate risk for 70 to 90 repeats.

For individuals who have a full mutation, the present study suggests that there is a high risk of mental impairment (100% for males and 89% for females). However, the severity of mental impairment is sex dependent: 96% of males showed MR, compared with less severe LD in 91% of the affected females. The finding of substantial methylation differences between males who have MR suggests that there may be a threshold level of CpG methylation above which mental impairment occurs. This is supported by comparison of Southern blot patterns between an affected male with MR (fig. 2, lane 5) and an NTM with no mental impairment (fig. 2, lane 2). This example shows little difference in CGG size of unmethylated fragments, but perhaps a more significant difference is that a subpopulation of apparently methylated DNA is present in the sample from the male with MR but is not present in the sample from the transmitting male. Possibly, both length of the (CGG)n and methylation contribute toward the phenotype by affecting expression of the FMR-1 gene product in a complex manner. Quantitative analysis of FMR-1 transcripts may help to define the relationships between abnormal methylation, regulation of FMR-1 transcription, and phenotypic expression. In the meantime, imprecise or yet-to-be-defined correlation between parameters of the mutation (number of repeats and extent of abnormal methylation) and the presence or severity of an affected phenotype needs to be considered when molecular testing is used, particularly in prenatal cases.

Of all the fragile X families now studied for the (CGG)n amplification, no cases of a new mutation have been reported. This suggests that the frequency of nonpenetrant carriers of fragile X syndrome might be higher than originally postulated (Smits et al. 1992). Indeed, in the present study, analysis of 247 individuals from the general population estimates the premutation allele frequency to be 0.8% (when repeats of 52 and 61 are considered as premutations). However, this estimate uses the ranges for premutation and meiotically stable normal (CGG)n observed for fragile X families. Clearly, these ranges are subject to ascertainment bias, and alleles in the general population may have a greater degree of meiotic stability for the same number of CGG repeats. A possible model is that a (CGG)n region is meiotically stable until it sustains one or more base substitutions that trigger an instability mechanism perhaps involving replication slippage. The recent evidence for a founder effect in fragile X syndrome (Richards et al. 1992; Oudet et al. 1993) may therefore involve an allele that, to become the unstable sequence, needs to undergo fewer base changes. Alternatively, founder alleles may be associated with sequences within or adjacent to the (CGG)n that promote transposition or sequence-specific recombination as proposed for minisatellites (Rogers 1985; Wahls et al. 1990). In any event, the study of transmission of alleles not ascertained through affected individuals may help to deter-

mine the true frequency of premutation alleles in the general population. Furthermore, models for inheritance patterns at the FMR-1 locus (Morton and Macpherson 1992) may possibly be refined by additional characterization of differences between stable alleles and unstable alleles.

In summary, the molecular changes associated with fragile X syndrome—i.e., expansion of a CGG repeat and methylation of a CpG island—can be reliably detected when a combination of Southern blot and PCR-based assays is used. An expansion of more than approximately 1 kb with an abnormal methylation pattern (full mutation) generally correlates with MR in males and with a lesser degree of mental impairment in females. For nonpenetrant females with a smaller amplification and a normal methylation pattern (premutation), determination of the number of CGG repeats can be used to predict a risk that the premutation will expand to a full mutation in the next generation. For individuals who carry a repeat size close to the lower limit shown to be unstable, analysis of other family members is recommended, to determine stability of the allele through female meiosis. Further studies of unstable versus stable alleles in this borderline zone may define another measurable parameter that more reliably predicts stability. Studies are also required for determination of the frequency of unstable alleles in the general population. However, when a lower limit of instability found in fragile X syndrome is applied to repeat sizes observed in the general population, the premutation allele frequency is estimated to be approximately 1%.

Acknowledgments

We acknowledge Dr. Kay Davies and Dr. Jean-Louis Mandel, who provided us with probes used in this study. We are grateful to Dr. Ted Brown for both reviewing the manuscript and providing helpful suggestions. We thank Karen Erwin for expert help in preparation of the manuscript. This work was funded in part by NIMH grant MH-45916.

References

Bell MV, Hirst MC, Nakahori Y, MacKinnon RN, Roche A, Flint TJ, Jacobs PA, et al (1991) Physical mapping across the fragile X: hypermethylation and clinical expression of the fragile X syndrome. Cell 64:861–866

Brown WT, Houck G, Jeziorowska A, Ding X, Dobkin C, Levinson FN, Sklower-Brooks S, et al (1992) Development of a PCR protocol for fragile X carrier screening: application to population screening and prenatal diagnosis. Am J Hum Genet Suppl 51:A20

De Boulle K, Verkerk AJMH, Reyniers E, Vits L, Hendrickx J, Van Roy B, Van Den Bos F, et al (1993) A point mutation in the *FMR-1* gene associated with fragile X mental retardation. Nature Genet 3:31–35

Dennis NR, Curtis G, Macpherson JN, Jacobs PA (1992) Two families with Xq27.3 fragility, no detectable insert in the FMR-1 gene, mild mental impairment, and absence of the Martin-Bell phenotype. Am J Med Genet 43:232–236

Erster SH, Brown WT, Goonewardena P, Dobkin CS, Jenkins EC, Pergolizzi RG (1992) Polymerase chain reaction analysis of fragile X-mutations. Hum Genet 90:55–61

Fu Y-H, Kuhl DPA, Pizzuti A, Pieretti M, Sutcliffe JS, Richards S, Verkerk AJMH, et al (1991) Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. Cell 67:1047–1058

Gedeon AK, Baker E, Robinson H, Partington MW, Gross B, Manca A, Korn B, et al (1992) Fragile X syndrome without CGG amplification has an *FMR1* deletion. Nature Genet 1:341–344

Hauge XY, Litt M (1993) A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphisms by the PCR. Hum Mol Genet 2:411–415

Heitz D, Devys D, Imbert G, Kretz C, Mandel J-L (1992) Inheritance of the fragile X syndrome: size of the fragile X premutation is a major determinant of the transition to full mutation. J Med Genet 29:794–801

Hirst MC, Barnicoat A, Flynn G, Wang Q, Daker M, Buckle VJ, Davies KE, et al (1993) The identification of a third fragile site, FRAXF, in Xq27-q28 distal to both FRAXA and FRAXE. Hum Mol Genet 2:197–200

Hirst MC, Nakahori Y, Knight SJL, Schwartz C, Thibodeau SN, Roche A, Flint TJ, et al (1991) Genotype prediction in the fragile X syndrome. J Med Genet 28:824–829

Jacky PB, Ahuja YR, Anyane-Yeboa K, Breg WR, Carpenter NJ, Froster-Iskenius UG, Fryns J-P, et al (1991) Guidelines for the preparation and analysis of the fragile X chromosome in lymphocytes. Am J Med Genet 38:400–403

Kremer EJ, Pritchard M, Lynch M, Yu S, Holman K, Baker E, Warren ST, et al (1991) Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence $p(CCG)_n$. Science 252:1711–1714

Latimer LJP, Lee JS (1991) Ethidium bromide does not fluoresce when intercalated adjacent to 7-deazaguanine in duplex DNA. J Biol Chem 266:13849–13851

Macpherson JN, Nelson DL, Jacobs PA (1992) Frequent small amplifications in the FMR-1 gene in fra(X) families: limits to the diagnosis of 'premutations.' J Med Genet 29:802–806

Morton NE, Macpherson JN (1992) Population genetics of the fragile-X syndrome: multiallelic model for the *FMR1* locus. Proc Natl Acad Sci USA 89:4215–4217

Nakahori Y, Knight SJL, Holland J, Schwartz C, Roche A, Tarleton J, Wong S, et al (1991) Molecular heterogeneity of the fragile X syndrome. Nucleic Acids Res 19:4355–4359

Oberlé I, Boué J, Croquette MF, Voelckel MA, Mattei MG, Mandel JL (1992) Three families with high expression of a fragile site at Xq27.3, lack of anomalies at the FMR-1 CpG island, and no clear phenotypic association. Am J Med Genet 43:224–231

Oberlé I, Rousseau F, Heitz D, Kretz C, Devys D, Hanauer A, Boué J, et al (1991) Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. Science 252:1097–1102

Oudet C, Mornet E, Serre JL, Thomas F, Lentes-Zengerling S, Kretz C, Deluchat C, et al (1993) Linkage disequilibrium between the fragile X mutation and two closely linked CA repeats suggests that fragile X chromosomes are derived from a small number of founder chromosomes. Am J Hum Genet 52:297–304

Pergolizzi RG, Erster SH, Goonewardena P, Brown WT (1992) Detection of full fragile X mutation. Lancet 339:271–272

Pieretti M, Zhang F, Fu Y-H, Warren ST, Oostra BA, Caskey CT, Nelson DL (1991) Absence of expression of the *FMR-1* gene in fragile X syndrome. Cell 66:817–822

Richards RI, Holman K, Friend K, Kremer E, Hillen D, Staples A, Brown WT, et al (1992) Evidence of founder chromosomes in fragile X syndrome. Nature Genet 1:257–260

Rogers JH (1985) The origin and evolution of retroposons. Int Rev Cytol 93:187–279

Rousseau F, Heitz D, Biancalana V, Blumenfeld S, Kretz C, Boué J, Tommerup N, et al (1991) Direct diagnosis by DNA analysis of the fragile X syndrome of mental retardation. N Engl J Med 325:1673–1681

Smits A, Smeets D, Hamel B, Dreesen J, van Oost B (1992) High prevalence of the fra(X) syndrome cannot be explained by a high mutation rate. Am J Med Genet 43:345–352

Snow K, Doud L, Hagerman R, Hull C, Hirst MC, Davies KE, Thibodeau SN (1992) Analysis of mutations at the fragile X locus using the DNA probe Ox1.9. Am J Med Genet 43:244–254

Sutcliffe JS, Nelson DL, Zhang F, Pieretti M, Caskey CT, Saxe D, Warren ST (1992) DNA methylation represses *FMR-1* transcription in fragile X syndrome. Hum Mol Genet 1:397–400

Sutherland GR, Baker E (1992) Characterization of a new fragile site easily confused with the fragile X. Hum Mol Genet 1:111–113

Sutherland GR, Gedeon A, Kornman L, Donnelly A, Byard RW, Mulley JC, Kremer E, et al (1991) Prenatal diagnosis of fragile X syndrome by direct detection of the unstable DNA sequence. N Engl J Med 325:1720–1722

Verkerk AJMH, Pieretti M, Sutcliffe JS, Fu Y-H, Kuhl DPA, Pizzuti A, Reiner O, et al (1991) Identification of a gene (*FMR-1*) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. Cell 65:905–914

Wahls WP, Wallace LJ, Moore PD (1990) Hypervariable minisatellite DNA is a hotspot for homologous recombination in human cells. Cell 60:95–103

Wöhrle D, Kotzot D, Hirst MC, Manca A, Korn B, Schmidt A, Barbi G, et al (1992) A microdeletion of less than 250 kb, including the proximal part of the FMR-1 gene and the fragile-X site, in a male with the clinical phenotype of fragile-X syndrome. Am J Hum Genet 51:299–306

Yu S, Pritchard M, Kremer E, Lynch M, Nancarrow J, Baker E, Holman K, et al (1991) Fragile X genotype characterized by an unstable region of DNA. Science 252:1179–1181