Research Article

Instability of a 550–Base Pair DNA Segment and Abnormal Methylation in Fragile X Syndrome

I. OBERLÉ, F. ROUSSEAU, D. HEITZ, C. KRETZ, D. DEVYS, A. HANAUER, J. BOUÉ, M. F. BERTHEAS, J. L. MANDEL*

The fragile X syndrome, a common cause of inherited mental retardation, is characterized by an unusual mode of inheritance. Phenotypic expression has been linked to abnormal cytosine methylation of a single CpG island, at or very near the fragile site. Probes adjacent to this island detected very localized DNA rearrangements that constituted the fragile X mutations, and whose target was a 550–base pair GC-rich fragment. Normal transmitting males had a 150– to 400–base pair insertion that was inherited by their daughers either unchanged, or with small differences in size. Fragile X–positive individuals in the next generation had much larger fragments that differed among siblings and showed a generally heterogeneous pattern indicating somatic mutation. The mutated allele appeared unmethylated in normal transmitting males, methylated only on the inactive X chromosome in their daughters, and totally methylated in most fragile X males. However, some males had a mosaic pattern. Expression of the fragile X syndrome thus appears to result from a two-step mutation as well as a highly localized methylation. Carriers of the fragile X mutation can easily be detected regardless of sex or phenotypic expression, and rare apparent false negatives may result from genetic heterogeneity or misdiagnosis.

THE FRAGILE X SYNDROME MAY BE THE MOST FREQUENT cause of inherited mental retardation; the incidence is about 1 in 1500 males and 1 in 2500 females (*1, 2*). The diagnosis is based on the presence in a variable proportion of cells (5 to 60 percent), of a fragile site on the X chromosome, at Xq27.3, which is induced to appear as a gap or break in vitro by culture conditions that affect deoxynucleotide synthesis (*3, 4*). Genetic linkage studies (*5*), as well as analysis of somatic cell hybrids that carry a fragile X chromosome on a rodent background (*6, 7*), have shown that the mutation is cis-acting and located at or very close to the fragile site. Segregation studies have shown that about 20 percent of males who carry the mutation have no clinical or cytogenetic expression (normal transmitting males; NTM's). Of the carrier females 30 percent show some mental retardation that is usually associated with a higher level of fragile site expression than in mentally normal carriers. Furthermore, penetrance appears variable in different sibships (even within the same family). Daughters of normal transmitting males rarely express any symptom while penetrance is high in sons and daughters of these carrier women (*8*). A very high mutation frequency at the fragile X locus has been assumed to account for the high incidence of the disease (the highest of all X-linked diseases) in view of the low reproductive fitness of affected males (*8, 9*).

Several hypotheses have been proposed to account for the cytogenetic expression of the fragile site and the peculiarities in the inheritance of the syndrome, in particular the apparent necessity for the mutation to be passed through a female for phenotypic expression in later generations. Pembrey *et al.* (*10*) have proposed that NTM's carry a premutation which would be converted to full mutation in germ cells of their daughters by a meiotic crossing over event, leading for instance to amplification of a pyrimidine-rich sequence (*1*). The use of closely linked polymorphic markers flanking the fragile X locus did not, however, detect such meiotic recombination events (*5*). Laird (*11*) has proposed that the fragile X mutation locally blocks the reactivation, prior to oogenesis, of a previously inactive X chromosome. This would create an imprinting on the mutant chromosome and inhibit expression of nearby genes.

Detection of fragile X–specific patterns by probes around the CpG island. We have recently shown by pulsed field gel electrophoresis that Bss HII, Sac II, and Eag I sites in the vicinity of probe Do33 (DXS465) are resistant to digestion in fragile X males but not in normal males, or in the four NTM's analyzed (*12*). These observations were confirmed (*14*) and could be best explained by abnormal methylation of a CpG island, consistent with an imprinting mechanism. Our previous results indicated that a single CpG island within the yeast artificial chromosomes (YAC's) 209G4 or 141H5 was methylated specifically in males with the fragile X syndrome (*13*). We subcloned the sequences around the CpG island and isolated three clones that cover 9 kb of that region: StB12, StX21, and StA22 (Fig. 1). The cloned region was devoid of repetitive sequences, with the exception of a 900-bp Eag I–Pst I fragment (StB12.5).

Probe StB12.3 detected the same rare cutter fragments as Do33 (DXS465) in normal or affected males (*12*), while StX21E gave the same pattern at probe St677 (DXS463), confirming that the sites methylated in fragile X patients are indeed contained in the subcloned region (*15*). Double digests were done on DNA samples from a series of normal, fragile X carriers or affected individuals with

I. Oberlé, F. Rousseau, D. Heitz, C. Kretz, D. Devys, A. Hanauer, J. L. Mandel, Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Institut de Chimie Biologique, Faculté de Médecine, 67085 Strasbourg Cédex, France. J. Boué, INSERM U73, Château de Longchamp, 75016 Paris, France. M. F. Bertheas, Hôpital Nord, 42277 St. Priest en Jarez, France.

*To whom correspondence should be addressed.

Bss HII, Sac II, or Eag I and with a restriction enzyme insensitive to methylation (Bgl II, Bam HI, or Pvu II). The Eag I site detected in Eag I plus Bgl II digests with probes StB12.3 or StA22 is not methylated in a normal male (Fig. 2, lane 2) or in an NTM (Fig. 2, lane 5); it is about 50 percent methylated in a normal female (Fig. 2, lane 9) and is totally methylated in two affected males (Fig. 2, lanes 4 and 7). Abnormal, larger fragments were detected in most individuals carrying a fragile X mutation, either in the Bgl II digest (Fig. 2, lanes 3, 4, 7, and 8) or in the Eag I–Bgl II fragments detected by StB12.3 (Fig. 2, same individuals as above and in the NTM and his daughter, lanes 5 and 6). Similar results were obtained on other individuals for digests with Bss HII and Sac II (*15*). The patient analyzed in Fig. 2, lane 3, had a complex pattern with an abnormal Bgl II fragment resistant to Eag I digestion, and an apparently normal one totally digested by this enzyme (the Eag I–Bgl II fragment revealed by StB12.3 appears, however, slightly larger than normal). The more telomeric probe F33 (Fig. 1A) detected an Eag I site that was totally digested in males, and half methylated in females, regardless of their clinical status (Fig. 2) (*13*).

These initial results suggested that, in addition to the methylation abnormalities, the 7-kb Eag I–Bgl II fragment detected by StB12.3 on the telomeric side of the CpG island may contain fragile X–specific DNA rearrangements.

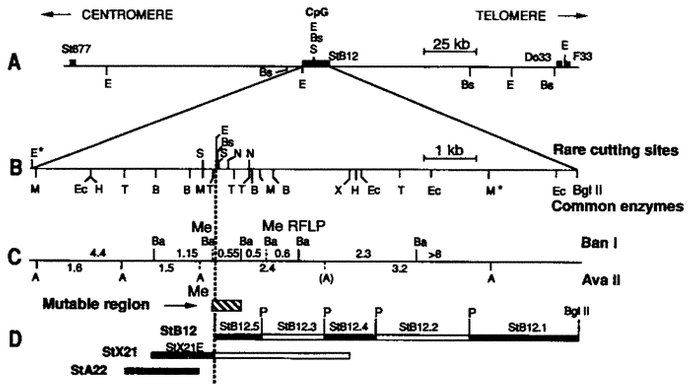

Fig. 1. Map of the fragile X locus. (**A**) Large-scale map of the fragile X region including the flanking probes St677, Do33, and F33 and the rare cutter restriction sites Eag I (E), Bss HII (Bs), and Sac II (S). The sites above the line are those cut on the active X chromosome in genomic DNA. Sites under the line have been mapped in yeast artificial chromosomes (YAC's) 141H5 and 209G4 (*13*). The box around the CpG island corresponds to the 9- kb region subcloned from the YAC's. (**B**) Restriction maps of the subcloned region. Mapping was done on cloned DNA and checked on genomic DNA with the following enzymes: Bst YI sites closest to the CpG island (B), Eco RI (Ec), Hind III (H), Msp I (M), Nar I (N), Taq I (T), and Xmn I (X). The sites marked by an asterisk are not cut in genomic DNA. The 1.1-kb interval between Msp I sites flanking the CpG island contains several other Msp I sites. (**C**) Restriction map of the Ban I (Ba) and Ava II (A) sites. Sites that are methylatable (Me) or polymorphic (RFLP) in genomic DNA are specified. An additional methylatable Ban I site has been found by sequencing, in the middle of the 550-bp fragment that contains the CpG island. The Ava II site between brackets showed inefficient and variable digestion, probably due to methylation, in males and females. Sizes of the restriction fragments are indicated in kilobases. (**D**) Localization of subcloned probes. Eag I–Bgl II fragments were cloned in pBluescript KS$^-$ and screened with total human DNA. The cloned fragments were mapped by hybridization to pulsed field blots of YAC 141H5 that had been digested with Eag I and Bss HII. A 7-kb clone (StB12) was mapped to the telomeric side of the CpG island. This fragment was further subcloned after Pst I (P) digestion and a nonrepetitive probe (StB12.3) close to the Eag I site was used to further screen an Xmn I library of the YAC. This screening yielded the 3.8-kb probe StX21. This probe contains rare cutter sites specific for the CpG island and clustered within 1 kb. The probe StX21E contains the segment centromeric to the Eag I site. Further screening of an Ava II library yielded the 1.5-kb probe StA22. The vertical dashed line refers to the Eag I site within the CpG island. The region that is the target for fragile X mutations is indicated by the hatched box.

Methylation abnormalities. Among many restriction digests tested with probe StB12.3, Ban I yielded a clear picture, which distinguished normal males from normal females, and normal males from affected males. In normal males, a polymorphic fragment was detected. Allele 1 (frequency = 0.85) corresponds to a 2.9-kb fragment (Fig. 3, lane 1) while allele 2 is defined by 2.3- and 0.6-kb fragments (Fig. 3, lane 5). In normal females homozygous for allele 1, two fragments are seen at 2.9 and 3.4 kb, that were of similar intensity in leukocyte DNA (Fig. 3, lane 13). In females homozygous for allele 2, an additional fragment was observed at 1.1 kb (Fig. 3, lanes 23 and 24). The adjacent probe, StB12.4, did not detect the 0.6 and 1.1-kb bands but detected the larger bands (*15*). The polymorphic site is thus located on the centromeric side of StB12.4, and the additional bands observed in females are best explained by methylation on the inactive X of a Ban I site located near to the centromeric end of fragment StB12.3 (Fig. 1C). The recognition site of Ban I is GGPyPuCC and may overlap with a methylatable CpG dinucleotide. Apart from this methylation-specific pattern, StB12.3 did not detect abnormal size Ban I fragments in fragile X patients or carriers.

Since the pattern is very simple to analyze, it was tested extensively in fragile X families (Table 1). The Ban I site appeared unmethylated in all normal males tested, totally methylated in 80 percent of the mentally retarded fragile X males, and at least 50 percent methylated

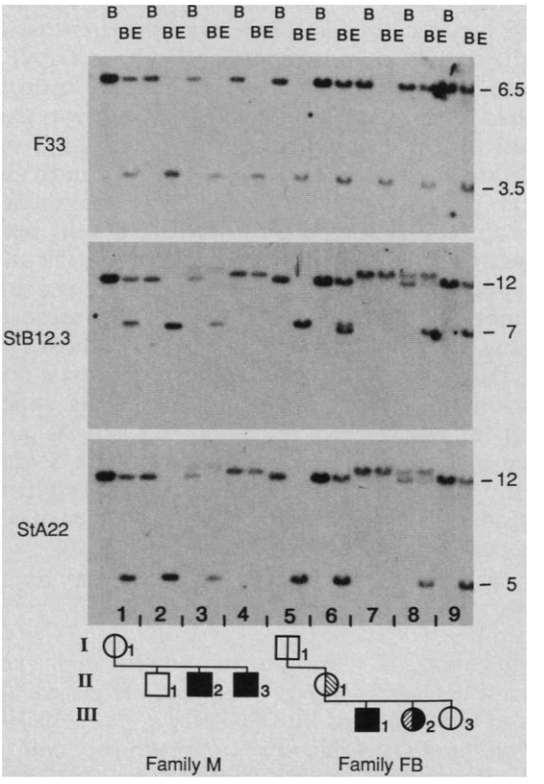

Fig. 2. Analysis of the CpG island in fragile X families; digestibility of the Eag I site and presence of abnormal fragments. Leukocyte DNA from members of fragile X families was digested with Bgl II (B) or Bgl II + Eag I (BE), separated by electrophoresis in 1 percent agarose, blotted onto diazobenzyloxymethyl paper (*28*), and hybridized to the probes indicated on the left. Sizes of the restriction fragments (in kilobases) are indicated on the right. Symbols describing each individual are shown at the bottom. Squares represent males; circles represent females. Undivided symbols indicate individuals who appeared normal and were not tested cytogenetically. For divided symbols, the right half represents the fragile X frequency, the left half the degree of mental retardation. (Open divided symbols) normal, no fragile X; (hatched symbols) mild mental retardation ("dull") (left), <4 percent fragile X (right); (closed symbols) mentally retarded (left), ≥4 percent fragile X (right). Generations are indicated by Roman numerals. Individual II of family FB (lane 5) was shown by segregation analysis to be a normal transmitting male (*29*).

in the remaining ones with a single exception (in family PC). In the four NTM's tested the site was also totally unmethylated (Fig. 3, lane 19). In women, detection of methylation abnormalities relies on dosage. The 25 daughters of NTM's analyzed were indistinguishable from normal females (Fig. 3, lanes 9, 10, 12, 14, and 16) while a clear excess of the methylated fragment was seen in 8 of 18 fragile X–expressing females (Fig. 3, lane 17).

Other methylatable sites followed the same sex– and fragile X–specific pattern and were all located in the immediate vicinity of the CpG island. There was a second Ban I site about 50 bp centromeric to the Eag I site, which was extensively analyzed with probes StA22 or StX21E. Other sites studied on a smaller number of patients and controls are one Ava II site in which the recognition site GG$\frac{A}{T}$CC may overlap with a CpG dinucleotide (Fig. 4) and several Hpa II (Msp I) sites within 1 kb of the Eag I site. In contrast in an Eag I plus Bgl II double digest, StA22 did not detect the expected 3.5-kb Eag I–Eag I fragment (Fig. 2). The centromeric Eag I site, marked with an asterisk in Fig. 1B, is thus probably methylated on both the active and inactive X chromosome, regardless of the presence of a fragile X mutation. Eco RI plus Hpa II double digests showed that the Hpa II site located at 4.8 kb from the CpG island on the telomeric side was also always methylated (Eco RI plus Msp I digests confirmed that the site indeed exists) (*15*). Furthermore, the presence of relatively large Msp I fragments (4.0 and 3.2 kb) immediately flanking the CpG island indicates that the density of methylatable sites is high only in the 1.1-kb segment that includes the Eag I, Bss HII, and Sac II sites. Hence, the region specifically methylated in patients appears less than 7 kb and may be as small as 1 or 2 kb.

Probes adjacent to the CpG island detect highly variable patterns. When the Ban I blots were hybridized to probes StX21E or StA22, a striking pattern was revealed in fragile X families (Fig. 3). In normal males a 1.15-kb fragment was detected by both probes. In normal females, a 1.7-kb band was also detected, which was about as intense as the smaller band. This again may be accounted for by a Ban site that is methylated on the inactive X, and is located within the CpG island, about 50 bp centromeric to the Eag I site. Probe StA22 hybridized to an additional 4.4-kb Ban I fragment, invariable in all 224 samples tested (fragile X or normals).

Out of 46 fragile X males, 44 showed an abnormal pattern (Table 1). The 1.15-kb fragment was absent in 33 individuals with an abnormal pattern (as in Fig. 3, lane 3) and much reduced in 9 others (as in Fig. 3, lane 2). Two of the 44 had the normal 1.15-kb fragment but also higher size bands (*15*). Those bands of larger size (1.9 to ≈6 kb) appeared in affected males who either had a very heterogeneous pattern (Fig. 3, lanes 2 and 3) or a single major abnormal band (Fig. 3, lanes 7 and 22). The pattern observed was different in males from the same family or even in affected brothers (Fig. 3, lanes 2 and 3). In these Ban I digests, NTM's showed the same pattern as normal males (Fig. 3, lane 19). In almost all carrier females, an abnormal pattern was observed regardless of their ability to express the fragile site. In all daughters of NTM's, an abnormal band was present at 1.9 to 2.2 kb. This band was generally of the same size between sisters (family MV) although differences of 100 to 200 bp were observed in some sibships (Fig. 3, family LR). In females with fragile X expression, the abnormal band was much larger (Fig. 3, lane 17) and could even be replaced by a heterogeneous pattern of bands (Fig. 3, lane 18) as seen in affected males. In these expressing females the normal methylated 1.7-kb band showed a decrease in intensity or, more rarely, disappeared (Fig. 3, lanes 17 and 18). The 1.15-kb unmethylated band appeared, in general, with a normal dosage.

Samples from some of the fragile X families were analyzed with additional enzymes. Abnormal patterns were also observed in Eco RI, Hind III, Ava II, and Xmn I digests with probes StB12.3, StX21E, and StA22. For a given individual, the size differences

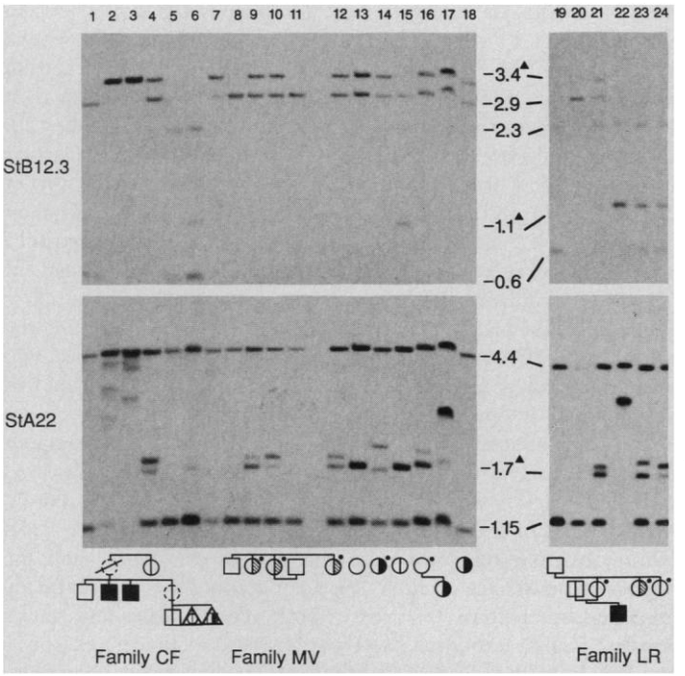

Fig. 3. Fragile X–specific methylation and mutations detected in Ban I digests by probes StB12.3 and StA22. Leukocyte DNA was digested with Ban I, separated by electrophoresis in 1 percent agarose, and blotted onto Hybond N$^+$ membranes. Hybridization was at 42°C in a buffer containing 40 percent formamide, and blots were washed at 65°C in 0.5× standard saline citrate, 0.1 percent SDS (*28*). Sizes of restriction fragments are indicated in kilobases. ▲, Fragments corresponding to sites methylated on the inactive X. Symbols are as in Fig. 2. Daughters of NTM's are marked with an asterisk. Symbols in triangles represent chorionic villi of a male and female fetus. Familial relationships are indicated and dashed symbols refer to individuals that were not analyzed on this blot.

between normal and abnormal bands appeared the same (within the precision of the size determination) with the different enzymes (Fig. 4). This indicates that the abnormalities are due to real length differences within the fragments analyzed, and not to the presence of abnormal structures (such as single-stranded DNA or conformation changes). A normal pattern was observed with the above probes in Taq I, Pst I, and Msp I digests.

Abnormal fragments were observed in Eco RI digests with StB12.3 for all nine NTM's analyzed; the normal 5.2-kb fragment was replaced by a single fragment of 5.35 to 5.6 kb. The same difference in size was present in Ava II or Xmn I digests, and could be best measured in Bst YI digests because of the smaller size (1.35 kb) of the target fragment (Fig. 1B). We conclude that abnormal fragments can be observed with probes StA22 or StX21E in Ban I digests only if the Ban I site centromeric to the CpG island is methylated, and are thus not detected in NTM's. In females with an abnormal Ban I fragment in the 1.9- to 2.2-kb range, who show 50 percent methylation with probe StB12.3, comparison of the Eco RI and Ban I digests indicated that the abnormal fragment is not methylated on the active X, but methylated on the inactive one. The variable intensity of the abnormal Ban I band compared to the normal, methylated, 1.7-kb band indicates, in leukocyte samples, the relative proportion of the mutation on the active and inactive X.

These results indicate that the region where the mutations take place is the 550-bp segment between the two Ban I sites and that includes the Eag I, Bss HII, and one Sac II site (Fig. 1, hatched box).

Mutation changes through generations. We have analyzed several large families to determine whether there were consistent patterns related to the parental origin of the mutation in females, the

position within the pedigree, and the phenotypic expression. In family K (Fig. 4), two female carriers and an NTM are present in generation I. The Eco RI–digested DNA from carrier I2 had an additional abnormal band with a 1.3-kb increase (defined as Δ = 1.3) above the normal 5.2-kb fragment. Her two affected sons II6 and II7 had a larger abnormal band (Δ = 2.5 or 2.3), similar to their affected nephew in generation III (Δ = 2.2). The two carrier daughters (II2 and II4) had a smaller abnormal band than their mother (Δ = 1.0 and 0.8). The other carrier female in generation I had a slightly shorter and less intense band than the normal fragment (Δ = −0.2) and, in addition, a faint and fuzzy larger band ($\Delta \simeq$ 2.2). The NTM had an abnormal fragment (Δ = 0.3) replacing the normal one. His daughter, who is an obligatory carrier, had two abnormal bands at Δ = 0.4 and 0.5, while his affected grandson had a very large fragment (Δ = 2.2). In females, the sum of the intensities of abnormal bands appeared equal to that of the fragment derived from the normal chromosome. The size differences observed in Eco RI digests were consistent with those observed in Ava II, Xmn I, or Ban I digests of the same individuals (Fig. 4) (*15*). Lymphoblastoid cell lines were analyzed in this family and the oligoclonal character of such cells may account for the unusually homogeneous pattern observed in both affected males and carrier females; a single abnormal band was observed in all cases except I1 and II8. Individual I1 was also the only case where we observed a fragment smaller than the normal one.

A general pattern has emerged from the analysis of 49 families. The results from seven of the larger families are summarized in Fig. 5. NTM's as well as their daughters always showed insertions in the 150- to 500-bp range. The mutation was passed from an NTM to his daughters either unchanged, within the limits of the analysis, or with a size increase of up to 200 bp, and sisters may thus appear different (family LR in Fig. 3). In the next generation, about 80 percent (40 of 48) of males or females analyzed who had inherited the mutated chromosome, as deduced from cytogenetic testing or from linkage study, showed an average increase of 1.5 to 2.5 kb in the size of the abnormal fragment that was most often associated with heterogeneity (as indicated by either multiple discrete bands or even a smear). In rare cases, abnormal bands could not be clearly seen in males lacking the normal 1.15-kb Ban I band. This might be explained by the high degree of heterogeneity, which resulted in an extended and faint smear (Fig. 5, two brothers in family TB). Coexistence of homogenous or heterogeneous patterns were observed within the same family, and even within a sibship (family TB). In some cases, however, mutations with a small change in size could be transmitted to daughters through one (family CF) or even two (family TB) generations. An affected mother with large fragments may have children with smaller abnormal fragments (family TB). The appearance of fragments with only a small increase in size was associated with lack of clinical expression and absence or low levels of fragile site expression (≤5 percent). The only exceptions are "mosaic" individuals in whom a near-normal fragment coexisted with larger ones (Fig. 2, lane 3, and female I1 from family K in Fig. 4).

Application to diagnosis. Our analysis shows that the detection of abnormal fragments is an excellent diagnostic indicator for the fragile X trait, and that the size of such fragments and the presence

Table 1. Diagnostic performances of StB12.3 and StA22 in Ban I digests. Methylation of the Ban I site detected by StB12.3 was scored from the relative intensity of the methylated and unmethylated bands. Absent and complete correspond to the absence of the methylated or unmethylated band, respectively. "Medium to high" refers to cases in which the intensity of the methylated band was equal or superior to that of the unmethylated one. The pattern detected by StA22 was considered as abnormal if the 1.15-kb band was absent or if abnormal size fragments were present (or both). Results from all fragile X families were included except for the three families with uncertain diagnosis cited. Results from chorion villus biopsies were also excluded. *N*, number of individuals.

Fragile X status	Pattern detected by StA22 (*N*)		Methylation pattern detected by StB12.3 (*N*)		
	Normal	Abnormal	Absent	Medium to high	Complete
	Males				
Normal	59	0	50	0	0
NTM	9	0	4	0	0
Affected	2†‡	44	1†	5	27
	Females				
Normal	41§	0			
Positive carriers	1†	43§			
Negative carriers*	0	17			
Cytogenetically untested carriers*	0	17			
Unknown*	21	7			

*Carrier status as deduced from geneology or segregation of polymorphic markers (or both). †The count includes one member of the PC family (see text). ‡One case showed methylation abnormalities with StB12.3. §In 16 of the 84 women in these classes, a very low percentage of fragile site was found (≤3%). They were classified as carrier (*14*) or noncarrier (*2*) by segregation analysis, and this was confirmed with the StA22 probe.

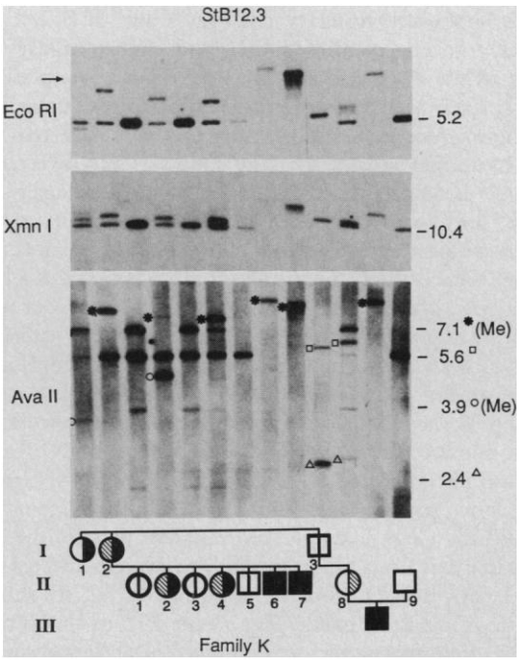

Fig. 4. Comparison of patterns observed in restriction digests in a fragile X family over three generations. Family K, which includes an NTM (I3), was previously reported (*30*). DNA's from lymphoblastoid cell lines were digested by Eco RI, Xmn I, or Ava II and hybridized to probe StB12.3. Sizes (in kilobases) of the restriction fragments detected in normal individuals are indicated on the right. The Ava II pattern is complex because one site was methylated on the inactive X and another site was, in general, very inefficiently digested (probably due to methylation). Thus a normal female may show 2.4-kb (faint) and 5.6-kb bands, and the methylated (Me) fragments on the inactive X, at 3.9 and 7.1 kb. The abnormal fragments are indicated by symbols on the gel that correspond to the symbols next to the sizes of the normal fragments from which they are derived (shown at the right). For instance in I2 and II4, the main abnormal fragment was derived from the 7.1-kb fragment, while in II2 it was mostly derived from the 3.9-kb fragment. The carrier female I1 showed an abnormal band smaller than the normal one (visible in the Eco RI and Ava II digests, but not in the Xmn I digest, because of incomplete digestion) and a faint and fuzzy larger one (indicated by an arrow in the Eco RI digest).

of abnormal methylation is associated with cytogenetic and clinical expression (Table 1 and Fig. 5). For example, we could detect all 17 carrier females who did not express the fragile X and found a fragile X mutation in 7 of 21 females at risk with a previously unknown status (in whom cytogenetic test and linkage analysis were inconclusive). We have observed similar abnormal patterns in chorionic villi (see lanes 6 and 7 in Fig. 3), but our limited data indicate that methylation may not be complete, in accordance with the lesser stability of X inactivation observed by Migeon *et al.* (*16*) in this tissue.

No false positives were observed in our analyses. However, some rare patterns and apparent false negatives should be noted. In two males we found a "mosaic status," with coexistence of a large methylated fragment and a single smaller unmethylated one typical of NTM (Fig. 2, lane 3). Mosaicism may account for some of the affected males with partial methylation and may also have occurred in a patient who showed very incomplete methylation in pulsed field electrophoretic analysis [patient AZ8 in (*12*)]. In one family (PC) with four members (three females and a 7-year-old boy) showing high fragile X expression (14 to 40 percent), we did not see any abnormal pattern. We propose that this represents genetic heterogeneity and may be similar to the family described by Voelckel *et al.* (*17*), where high fragile X expression was unlinked to mental retardation. Three other families in which we failed to find any abnormality were excluded from our analysis because of a combination of low fragile X expression in the single affected male (two families), very incomplete cytogenetic and clinical documentation (one family), or presence of double recombinations with close flanking markers in patients (two families). These families should be reinvestigated because misdiagnosis appears the most likely explanation. The analysis of 224 individuals confirms the total diagnostic specificity of the combined use of StA22 and StB12.3 (100 percent) and its very high sensitivity (98 percent).

The cloning of sequences flanking rare cutter restriction sites between markers DXS465 and DXS463 has allowed us to confirm that methylation at a single CpG island is specifically associated to expression of the fragile X syndrome in males. The abnormal methylated region may extend over only a few kilobases. This CpG island is also methylated on the inactive X in normal females, as is usual for X-linked islands (*18*). The use of probes adjacent to this island has uncovered an extraordinary instability that affects a 550-bp CpG-rich segment also observed by Yu *et al.* (*31*). By DNA blot analysis we have been able to detect abnormal patterns in almost all carriers of the fragile X mutation, regardless of sex or their ability to express a clinical or cytogenetic phenotype. A few families did not show an abnormal size fragment or the specific methylation pattern in affected males, and we believe that this reflects either genetic heterogeneity or inaccurate diagnosis.

Our observations have immediate diagnostic applications since it will be now possible to detect with probably total reliability and specificity all male or female carriers of a fragile X mutation. This is important for carrier diagnosis in females because cytogenetic analysis has a false negative rate of about 50 percent (even higher for daughters of normal transmitter males) (*8*) and false positives are possible when a very low frequency of fragile site expression is observed. Furthermore the cytogenetic test is, especially in females, tedious and expensive since a large number of metaphases need to be analyzed (*19*) and the frequency of expression is sensitive to many variables (*4*). The DNA test should allow a more complete ascertainment of the disease and better epidemiological studies of the association of fragile X with, for example, autism or mild mental retardation. It should also be possible to estimate the true frequency of NTM's, which has been predicted to be equal to that of affected males according to Laird's hypothesis (*20*).

Our initial analysis shows that abnormal patterns are also seen in chorionic villi (Fig. 3, lane 7). This should simplify prenatal diagnosis, which previously had relied on a combination of complicated culturing of chorionic villus cells and either linkage analysis or further cytogenetic analysis on fetal blood (*21*). Furthermore, we believe that some prediction of the clinical outcome will be possible. Our present data show that there is little or no mental impairment in individuals (males or females) who carry an abnormal fragment that differs by ≤500 bp from the normal, while the presence of a fragment with a larger insertion (or a smear) and the associated methylation are tightly linked to high expression of the fragile site and to mental retardation in males. Mental retardation in females with a large abnormal fragment is probably linked to preferential inactivation of the normal X chromosome in cells or tissues (still unknown) implicated in this phenotype. It is not certain that this can be predicted from analysis of chorionic villi (or leukocytes).

Possible mechanisms. Two types of mutated regions can be distinguished at the molecular level. Normal transmitting males and some females, including all daughters of NTM's, have a small insertion of 150 to 500 bp. The adjacent CpG island is unmethylated on the active X chromosome, but it is methylated on the inactive X as in normal females. In these individuals there was little or no clinical expression of disease and only a low level (if any) of

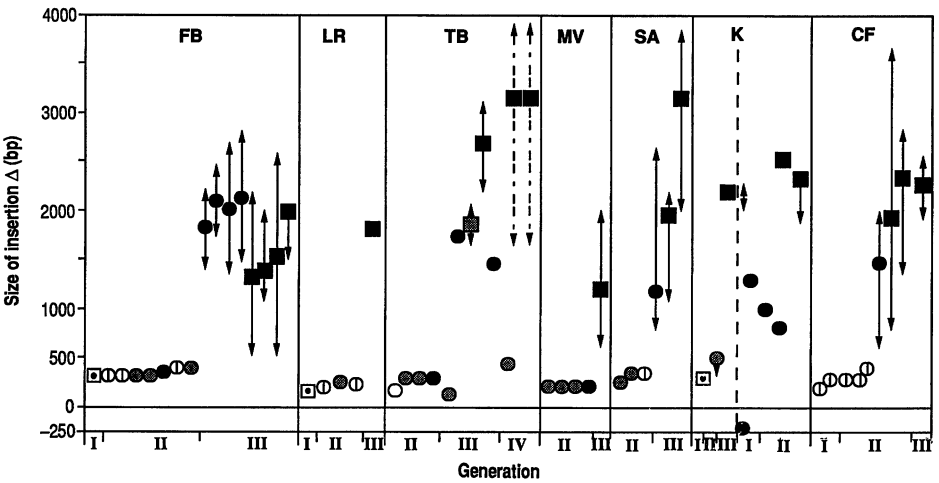

Fig. 5. Variation in the size of mutated fragments in successive generations of fragile X families. The differences between the sizes of abnormal fragments and the corresponding normal one (Δ) were measured in Ban I, Eco RI, or Ava II digests and are represented on the ordinate. Each individual is represented by a symbol defining his or her cytogenetic status. Only carriers of fragile X mutations are represented. Squares with center dots, NTM's; open circles, cytogenetically untested carrier females; divided circles, female with no fragile X detected; gray symbols, <4 percent fragile X; dark symbols, ≥4 percent fragile X. All males with ≥4 percent fragile X were mentally retarded, but not all females. The three females with changes in fragment size of 400 bp and noted as dark symbols had frequencies of fragile X ≤6 percent. Arrows approximate the extent of size heterogeneity (discrete bands or smear) when present. Dashed arrows (family TB) indicate a widely spread and very faint smear. Generations are indicated by Roman numerals. The first six families have an NTM in generation one, whose fragment size was calculated in Eco RI or Ava II digests. In family CF, the two carriers in generation I may be daughters of an NTM. Family K has one branch with an NTM (see Fig. 4). The only woman with a fragment smaller than normal also shows a smear with a high size change (I1 in Fig. 4). In family K, DNA was derived from lymphoblastoid cell lines and showed little heterogeneity.

fragile X. In affected males and in females with high expression of fragile X, the insertion is much larger (1000 to 3000 bp or more) and is often heterogeneous within an individual; the CpG island is methylated on the active X, which is most easily seen in males. The first pattern may be considered as a premutation; while the second is the full mutation that results in phenotypic expression, and its instability causes further somatic mutations. As postulated by Pembrey *et al.* (*10*) and Nussbaum and Ledbetter (*1*), the transition from premutation to mutation occurs only after passage through a female, and could result from sequence amplification. This event is not obligatory, but occurs at high frequency; we observed 80 percent in our sample although the percentage was subject to ascertainment bias in favor of phenotypic expression. To account for the segregation data (*8*), it would be necessary to assume that the probability of this event reoccurring may vary in different sibships.

Two main questions can be raised. When does the transition occur, and what mechanism dictates whether a premutation will lead to a full mutation? The most likely explanation would be that the full mutation is produced during oogenesis. Adapting Laird's hypothesis (*11, 20*), the transition would occur on a previously inactive X chromosome, where the premutation has locally blocked reactivation prior to oogenesis. The X inactivation is linked to methylation of the CpG island, and in this scheme abnormal methylation would precede and possibly facilitate the size increase (by interfering with replication). However, this may be difficult to reconcile with the males who appear as mosaics carrying both the unmethylated premutation and the methylated full mutation because this would imply that a secondary somatic mutation resulted both in a return to the premutation status and to loss of methylation. Laird's hypothesis predicts that NTM's should be as frequent as affected males, the 20 percent figure being due to observation biases (*20*) and this should now easily be testable by analysis of all sibships in fragile X families.

An alternative possibility is that the transition occurs at random, with a frequency that may depend on the exact structure of the premutation. The premutation is not completely stable within a family in that small size variations may occur during transmission to the next generation. These variations could affect the probability of subsequent amplification. The amplified sequence would interfere with normal control of methylation. Methylation abnormalities would in such a scheme be a consequence of the presence of the full mutation. To account for the mosaic cases, we must assume that transition can also take place during early embryogenesis, and that the mosaics inherited a premutation from their mother. It is then difficult to explain why such somatic amplification does not occur early in embryogenesis when the premutated X is transmitted from the father. One may invoke differential imprinting on the male and female germ cells; such imprinting accounts for preferential inactivation of the paternal X in extraembryonic membranes in mouse (*22*), or for differential methylation of CpG sites outside CpG islands in sperm and oocytes (*23*). Regardless of the model, our data indicate that once the full mutation is established it often generates somatic instability. The occurrence of somatic mutations at different times during early embryogenesis would account for the observation of homogeneous or heterogeneous band patterns. The mutation does appear to be stable in lymphoblastoid lines.

Sequence analysis of mutant alleles should clarify the mechanism of their generation. The premutation could be due to unequal sister chromatid exchange or to unequal meiotic crossing over. The mutation rate for this event would appear extremely high given the small size (550 bp) of the target. From segregation data it was estimated that the mutation rate for fragile X should be greater than or equal to the rate for Duchenne muscular dystrophy, in which the target gene is 2.10^6 bp (*8, 9*). A simple methylatable sequence of the form $(GCC)_{10}$ TCC $(GCC)_9$ is present in the normal fragment (*15*) and may be involved in the mutation mechanism, which in turn could lead to amplification of this sequence. This sequence might assume an abnormal conformation, especially when methylated; for example, methylation of a $(CG)_n$ sequence facilitates formation of Z-DNA (*24*). The presence of the premutation could lead to impairment of replication during oogenesis or early embryogenesis, a time when cell division is very rapid, and facilitate the further amplification to the full mutation. Amplification would occur either by unscheduled replication (the onion skin model) proposed for other types of amplification (*25*) or by unequal sister chromatid exchanges (SCE). High frequencies of SCE's at the fragile sites have indeed been observed in patients (*26*). Molecular analysis of premutations and mutations, as well as a systematic study of their transmission within families, including atypical cases (*17, 27*), should lead to a better understanding of this unusual disease.

REFERENCES AND NOTES

1. R. L. Nussbaum and D. H. Ledbetter, *Annu. Rev. Genet.* **20**, 109 (1986).
2. T. Webbs, in *The Fragile X Syndrome*, K. E. Davies, Ed. (Oxford Univ. Press, New York, 1989), pp. 40–50.
3. G. R. Sutherland, *Am. J. Hum. Genet.* **31**, 125 (1979); _____, E. Baker, A. Fratini, *Am. J. Med. Genet.* **22**, 433 (1985).
4. N. Tommerup, in *The Fragile X Syndrome*, K. E. Davies, Ed. (Oxford Univ. Press, New York, 1989), pp. 102–122.
5. G. K. Suthers *et al.*, *Am. J. Hum. Genet.* **48**, 460 (1991); F. Rousseau *et al.*, *ibid.*, p. 108.
6. R. L. Nussbaum, *Am. J. Med. Genet.* **23**, 457 (1986).
7. S. T. Warren *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 3856 (1990).
8. S. L. Sherman *et al.*, *Hum. Genet.* **69**, 289 (1985).
9. F. Vogel, W. E. Crusio, C. Kovac, J. P. Fryns, M. Freund, *ibid.* **86**, 25 (1990).
10. M. E. Pembrey, R. M. Winter, K. E. Davies, *Am. J. Med. Genet.* **21**, 709 (1985).
11. C. D. Laird, *Genetics* **117**, 587 (1987).
12. A. Vincent *et al.*, *Nature* **349**, 624 (1991).
13. D. Heitz *et al.*, *Science* **251**, 1236 (1991).
14. M. V. Bell *et al.*, *Cell* **64**, 861 (1991).
15. F. Rousseau *et al.*, unpublished data.
16. B. R. Migeon, M. Schmidt, J. Axelman, C. Ruta Cullen, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 2182 (1986).
17. M. A. Voelckel *et al.*, *Hum. Genet.* **81**, 353 (1989).
18. S. F. Wolf *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2806 (1984); D. Toniolo *et al.*, *EMBO J.* **7**, 401 (1988).
19. P. B. Jacky *et al.*, *Am. J. Med. Genet.* **38**, 400 (1991).
20. J. A. Sved and C. D. Laird, *Am. J. Hum. Genet.* **46**, 443 (1990); C. D. Laird, M. M. Lamb, J. L. Thorne, *ibid.*, p. 696.
21. I. Kennerknecht, G. Barbi, N. Dahl, P. Steinbach, *Am. J. Med. Genet.* **38**, 467 (1991); E. C. Jenkins *et al.*, *ibid.*, p. 447.
22. S. G. Grant and V. M. Chapman, *Annu. Rev. Genet.* **22**, 199 (1988).
23. D. J. Driscoll and B. R. Migeon, *Somatic Cell Mol. Genet.* **16**, 267 (1990).
24. M. Behe and G. Felsenfeld, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 1619 (1981).
25. G. Stark and G. M. Wahl, *Annu. Rev. Biochem.* **53**, 447 (1984).
26. N. Tommerup, *Hum. Genet.* **81**, 377 (1989); S. L. Wenger, J. C. Hennessey, M. W. Steele, *Am. J. Med. Genet.* **26**, 909 (1987).
27. C. D. Laird, *Am. J. Med. Genet.* **38**, 391 (1991).
28. I. Oberlé *et al.*, *Hum. Genet.* **72**, 43 (1986).
29. G. Camerino, M. G. Mattei, J. F. Mattei, M. Jaye, J. L. Mandel, *Nature* **306**, 701 (1983).
30. F. A. Rhoads, A. C. Oglesby, M. Mayer, P. A. Jacobs, *Am. J. Med. Genet.* **12**, 205 (1982).
31. S. Yu *et al.*, *Science* **252**, 1179 (1991).

32. We thank V. Biancalana and S. Blumenfeld for RFLP analysis in fragile X families and E. Engel, S. Gilgenkrantz, P. Jacobs, M. G. Mattei, and N. Tommerup for referring families to us. We thank B. Boulay and D. Massel for help in preparation of the manuscript. Supported by grants from Ministère de la Recherche et de la Technologie (88.C.0178), Caisse Nationale de l'Assurance Maladie des Travailleurs Salariés (CNAMTS) and Association Française contre les Myopathies (J.L.M.), and a postdoctoral fellowship from the Medical Research Council of Canada (F.R.).

1 April 1991; accepted 26 April 1991