Am. J. Hum. Genet. 57:1006–1018, 1995

Prevalence of Carriers of Premutation-Size Alleles of the *FMR1* Gene—and Implications for the Population Genetics of the Fragile X Syndrome

François Rousseau,[1] Patricia Rouillard,[1] Marie-Lou Morel,[1] Edward W. Khandjian,[1] and Kenneth Morgan[2]

[1]Unité de Recherche en Génétique Humaine et Moléculaire, Centre de Recherche de l'Hôpital St-François-d'Assise and Département de Biochimie, Faculté de Médecine, Université Laval, Quebec City; and [2]Departments of Human Genetics and Medicine, McGill University, and the Montreal General Hospital Research Institute, Montreal

Summary

The fragile X syndrome is the second leading cause of mental retardation after Down syndrome. Fragile X premutations are not associated with any clinical phenotype but are at high risk of expanding to full mutations causing the disease when they are transmitted by a carrier woman. There is no reliable estimate of the prevalence of women who are carriers of fragile X premutations. We have screened 10,624 unselected women by Southern blot for the presence of *FMR1* premutation alleles and have confirmed their size by PCR analysis. We found 41 carriers of alleles with 55–101 CGG repeats, a prevalence of 1/259 women (95% confidence interval 1/373–1/198). Thirty percent of these alleles carry an inferred haplotype that corresponds to the most frequent haplotype found in fragile X males and may indeed constitute premutations associated with a significant risk of expansion on transmission by carrier women. We identified another inferred haplotype that is rare in both normal and fragile X chromosomes but that is present on 13 (57%) of 23 chromosomes carrying *FMR1* alleles with 53–64 CGG repeats. This suggests either (1) that this haplotype may be stable or (2) that the associated premutation-size alleles have not yet reached equilibrium in this population and that the incidence of fragile X syndrome may increase in the future.

Introduction

The fragile X syndrome of mental retardation is a leading cause of mental retardation (Nussbaum and Ledbetter 1989; Sherman 1991). It was the first disease shown to be associated with unstable or "dynamic" mutations (Oberlé et al. 1991). The disease is associated with an expansion of a CGG repeat (Oberlé et al. 1991; Verkerk et al. 1991; Yu et al. 1991) located in the CpG island (Heitz et al. 1991; Oberlé et al. 1991) of the *FMR1* gene (Verkerk et al. 1991). The CpG island shows abnormal DNA methylation in all affected individuals (Heitz et al. 1991; Oberlé et al. 1991; Rousseau et al. 1991; Vincent et al. 1991), which is associated with absence of expression of the *FMR1* gene (Pieretti et al. 1991). Normal alleles have a variable number of CGG repeats (range 6–54), with a mode of 29 (Fu et al. 1991). Fragile X mutations are classified into premutations and full mutations, on the basis of both their size and CpG-island methylation on the active X chromosome (Oberlé et al. 1991; Rousseau et al. 1991). Fragile X premutations have been defined as having a size of ~55–230 CGG repeats and are not associated with any abnormal methylation (Oberlé et al. 1991). In fragile X families, they are found in all normal transmitting males and in a majority of unaffected carrier females. Premutations do not seem to be associated with any abnormal phenotype (Oberlé et al. 1991; Rousseau et al. 1991, 1994*a*). When transmitted by a carrier female, fragile X premutations have an average risk of 80% (range 10%–100%), depending on allele size (Fu et al. 1991; Heitz et al. 1992; Yu et al. 1992), of expanding to full mutations, which are defined as being >~230 CGGs and associated with abnormal methylation of the *FMR1* CpG island. In males, fragile X full mutations are associated with mental impairment in >99% of the cases, whereas only ~55% of female carriers of full mutations are affected (Rousseau et al. 1991, 1994*a*). Up to 15% of carriers of full mutations also have premutations in a proportion of their cells; they have been termed "mosaics" and appear, in general, as affected as are carriers of only full mutations (Rousseau et al. 1994*a*). No single-step transition from a normal allele to a premutation or to a full mutation has been reported. Given the multistep process of transition from small premutations of most 54–60 trinucleotide repeats to full mutations (Oberlé et al. 1991), new mutations have not been observed in families

Received May 3, 1995; accepted for publication July 27, 1995.
Address for correspondence and reprints: Dr. François Rousseau, Unité de Recherche en Génétique Humaine et Moléculaire, Centre de Recherche de l'Hôpital-St-François-d'Assise, 10, rue de l'Espinay, Quebec G1L 3L5, Canada.
© 1995 by The American Society of Human Genetics. All rights reserved.
0002-9297/95/5705-0005$02.00

ascertained through an affected individual, because they are rare and/or because it has not been possible to study more than three or four generations.

The incidence of the fragile X syndrome has been reported to be roughly the same in several Caucasian populations and is ~1/2,000 male births and ~1/4,000 female births (Sherman et al. 1985; Brown et al. 1986). Linkage disequilibrium between fragile X mutations and closely linked markers has been reported as evidence for founder effect and, at the same time, multiple origins of these mutations (e.g., see Richards et al. 1992; Oudet et al. 1993*a*, 1993*b*; Macpherson et al. 1994). Since affected males and penetrant females have a reduced reproductive fitness, the similar incidences of the disease in various countries, combined with the evidence for a limited number of initial mutations, suggest that the mutation rate from normal alleles to premutation alleles is very low and is followed by a rate of mutation increasing with the size of the expansion but small enough to generate a large pool of premutations (Oudet et al. 1993*a*).

Because fragile X premutations are not associated with any abnormal phenotypic feature, carriers of premutations can only be detected by direct genotypic analysis (Oberlé et al. 1991; Rousseau et al. 1991). In fragile X families, the prevalence of premutations is higher than that of full mutations (Rousseau et al. 1991, 1994*a*). Occasionally, in fragile X families, spouses (who have married in) have been found to be carriers of premutation-size alleles (Rousseau et al. 1991; Macpherson et al. 1992). On the basis of this occasional finding, some authors (Rousseau et al. 1994*a*) have proposed that the prevalence of women who would be carriers (premutation or full mutation) could be as high as 1/250. There is significant variation in the estimates of the prevalence of carriers of fragile X premutations: 1/163–1/1,538 (Fu et al. 1991; Reiss et al. 1994; Rousseau et al. 1994*a*). Because of the small sample sizes, accurate estimates of this prevalence have been difficult to obtain.

Models have been proposed to explain the transmission of fragile X mutations (e.g., see Pembrey et al. 1985; Laird 1987), as well as to explain the linkage disequilibrium observed between fragile X mutations and specific alleles of flanking markers (Morton and Macpherson 1992; Oudet et al. 1993*a*; Kolehmainen 1994). One model predicts a pool of premutated alleles that are generated by a mutation rate, from normal to premutated alleles, of $\sim 2.5 \times 10^{-4}$ and that mutate toward larger premutations at a rate proportional to their size, until they become full mutations (Oudet et al. 1993*a*). With a nine-allele model (Kolehmainen 1994), an equilibrium frequency of .00138 for alleles of 50–59 trinucleotide repeats was predicted.

Direct DNA testing of the *FMR1* locus (i.e., determination of both size of the expansion and methylation status of the CpG island) allows identification of carriers of mutations in fragile X families (Rousseau et al. 1991; Sutherland et al. 1991). Carriers can be offered genetic counseling and, when appropriate, prenatal diagnosis (Rousseau et al. 1991; Sutherland et al. 1991). Because the fragile X syndrome has been estimated to have a high incidence and is a major cause of mental retardation, some authors have considered screening for fragile X premutation carriers in the general population or in women of childbearing age (Bonthron and Strain 1993; Bundey and Norman 1993; Howard-Peebles et al. 1993; Palomaki and Haddow 1993; Rousseau 1994). However, most authors agree that accurate estimates of the prevalence of fragile X mutations are required before implementation of such programs is considered. Also, there would be technical and cost considerations in implementing such programs. For example, PCR-based techniques may not be sensitive enough to detect all premutations, full mutations, and mosaic cases; and Southern-blot techniques have not been adapted for cost-effective screening. However, PCR protocols have been proposed for screening of fragile X mutations (Brown et al. 1993); but they require both transfer of the PCR products to a nylon membrane and hybridization to a probe, to detect the very faint bands amplified.

We screened, by Southern blot, a large cohort of unselected women to obtain a reliable estimate of the prevalence of fragile X premutation-size alleles in the general population. We found a carrier prevalence of 1/259 women. We analyzed haplotypes of normal size alleles, premutation-size alleles, and full mutations, to better define the likelihood, in the next generation, that the premutation-size alleles identified by screening will expand to full mutations.

Material and Methods

Samples

A total of 11,860 consecutive hemoglobin/hematocrit or complete blood count (CBC) leftover samples from female outpatients were collected from the hematology laboratory of a general hospital in Quebec during a period of 18 mo. Each sample was assigned a unique number and was handled in an unlinked anonymous protocol precluding the association of DNA results with personal identifying information. Some 1,236 duplicates were identified because of file-number matches prior to DNA analysis, and the anonymous sample numbers of the replicates were tagged as duplicates to be excluded from the subsequent study. In this study, procedures for sample and information collection and coding, analysis of samples, and recording and reporting of results were approved by the appropriate institutional ethical review board.

DNA Purification and Southern Blot Analysis

Samples were grouped sequentially in pools of five, and 40 μl of whole blood was transferred to a 1.5-ml Eppendorf tube for a total of 200 μl of whole blood per pool. The remainder of the CBC sample was labeled with the sample number and was kept frozen at −20°C until further study. DNA was purified from the whole blood pools by a miniprep method developed in our laboratory (Rousseau et al. 1994*b*), was digested in the same initial Eppendorf tube with the restriction endonuclease *Bcl*I, as reported previously (Rousseau et al. 1992), was precipitated with ammonium acetate 7.5 M and ethanol, was resuspended in loading buffer, and was loaded on a 6% polyacrylamide gel (1:30 bis-acrylamide) in 1 × TBE (Tris-borate EDTA), followed by migration for 16 h at 80 V, as described elsewhere (Rousseau et al. 1994*b*). After an ethidium bromide staining, the DNA was electroblotted on a Hybond N+ membrane (Amersham) at 80 mA for 4 h. The blot was then hybridized to probe StB12XX according to standard protocols (Rousseau et al. 1992). After the washing steps, the blot was exposed overnight and until slight overexposure, before premutation-size bands were searched for.

Screening Strategy

Positive pools were identified by the presence of a *Bcl*I band of *FMR1* sequence that hybridized to probe StB12XX and that was significantly larger than bands of the other alleles in the pool (fig. 2). Practically, this corresponded to a difference of migration of ⩾3 mm between the mean size of the other alleles and the outlier band. Pools with discrete individual bands that were <3 mm from the bulk of the other alleles were classified as containing "large" alleles. For pools where the presence of a large or premutation-size band was suspected, the five different samples were individually reanalyzed after repurification of DNA according to the protocol described above (fig. 2). DNA from true positive samples was extracted for determination of the number of triplet repeats and for haplotype analyses (see below). The sensitivity of this screening strategy was estimated by reanalyzing individually 1,000 samples that constituted 200 negative pools. No premutation-size allele was found, suggesting that few, if any, alleles in the premutation range were missed.

Typing of the *FMR1* Repeats and Flanking Markers

The number of trinucleotide repeats at the *FRAXA* locus was determined, for all positive samples, by PCR according to the protocol of Fu et al. (1991) and by blotting of the sequencing gel (6% polyacrylamide denaturing gel), followed by hybridization to probe StB12.5 (Oberlé et al. 1991) to reveal the PCR products. Sizing of the trinucleotide repeat was done by comparison with a sequence ladder loaded on the gel. *DXS548* (Verkerk et al. 1991) and *FRAXAC2* (Richards et al. 1992) microsatellites were analyzed according to published protocols (Oudet et al. 1993*a*). Two other polymorphic markers were used, namely, *FRAXAC1* (Richards et al. 1992) and a polymorphic *Ban*I site (Oberlé et al. 1991) analyzed by PCR using two novel primers (*FMR1Ban*F: 5′ TAGCCAAACGTGTCCTGTCTG 3′ and *FMR1Ban*R: 5′ GGCCGAAATCGGCGCTAAGTGACG 3′; GenBank accession no. G10729), followed by digestion with *Ban*I and agarose gel electrophoresis. The PCR product contains also one constant *Ban*I site, which is useful as an internal control for completeness of digestion.

Allele Frequencies

The sample also included fragile X chromosomes observed in 35 fragile X families and 232 non–fragile X, independent chromosomes from 68 women selected randomly from the sample of 10,624 women included in the study and from 96 consecutive blood leftovers from men who were not analyzed prior to their inclusion in the control group. Linkage disequilibrium was tested by analysis of allelic association between each marker locus and CGG repeats coded into size classes as described below. Marker-allele frequencies were estimated conditional on the CGG-repeat-size class, by maximum likelihood. Maximum-likelihood analysis was accomplished by the EM algorithm (see below). The following CGG-repeat-size classes based on the antimodes of the frequency distribution of repeat alleles were used in the maximum-likelihood estimation of conditional marker-allele frequencies: (1) 0–26 triplet repeats, (2) 27–33 triplet repeats, (3) 34–40 triplet repeats, (4) 41–52 triplet repeats, (5) 53–64 triplet repeats, (6) 65–101 triplet repeats, and (7) fragile X full mutations. Eleven alleles were observed for *DXS548*, five for *FRAXAC1*, two for the *Ban*I RFLP and, 13 for *FRAXAC2*. Maximum-likelihood estimates of conditional allele frequencies showed that for each of *DXS548*, *FRAXAC1*, and *FRAXAC2* there was a predominant allele in CGG-repeat-size class 5 (53–64 repeats), an allele that was not the predominant allele in any of the other CGG-repeat-size classes. Because of these observations and the sparse distribution of alleles, we reanalyzed the data according to the following scheme, where we chose the most frequent allele at a given locus in the polymorphic normal range, the most frequent allele in males with the fragile X syndrome, and the most frequent allele in class 5 CGG repeats and grouped together all other alleles as a single allele. We estimated marker-allele frequencies for the seven size classes of CGG repeats, assuming that allele frequencies were either independent of the number of CGG repeats or conditional on the number of repeats. A test of linkage disequilibrium for a marker locus was based on the likelihood-ratio test comparing the two

hypotheses. Maximum-likelihood estimates of marker-allele frequencies conditional on CGG-repeat-size class were obtained by the EM algorithm (Dempster et al. 1977; Weir 1990). For maximum-likelihood estimation using the EM algorithm, it was assumed that the combinations of haplotypes of a phase-ambiguous genotype were in relative multinomial proportions. Haplotypes comprised alleles of a specific marker locus and size class of CGG repeats. Under the model of linkage equilibrium (i.e., no association), allele frequencies of a specific marker locus are the same for each CGG-repeat-size class. Estimation of the conditional marker-allele frequencies was carried out iteratively until convergence to a specified precision was reached.

The most frequent haplotype found in fragile X syndrome males in Quebec we call the "major" fragile X haplotype. The alternative major haplotype inferred for premutation carriers was determined as that comprising the most frequent allele at each of four marker loci estimated for *FMR1* alleles with CGG repeats in the range of 54–64.

Results

Screening for Premutation-Size FMR1 Alleles

We screened 10,624 unselected women for premutation-size alleles in an unlinked-anonymous design. These women came from a population of >600,000 individuals in the Quebec City metropolitan area. The design and analysis of the samples, using a high-throughput Southern-blot method (Rousseau et al. 1994*b*), are depicted in figure 1 and have been described above. Samples constituting positive pools were reanalyzed individually (fig. 2), and those with large alleles were typed by PCR to determine the number of *FMR1* CGG repeats (Fu et al. 1991). We found 41 carriers of an allele with >54 CGG repeats, i.e., in the fragile X premutation-size range. Therefore, the estimate of the prevalence of premutation carriers in the study population was 1/259 (95% confidence interval 1/373–1/198). The number of CGG repeats of these alleles was 55–101 (appendix A).

Among individuals with no premutation-size alleles (a random sample and samples from positive pools with distinct bands on Southern blot), we observed 83 women with 34–53 CGG repeats on one of their two *FMR1* alleles. We do not consider that we ascertained all alleles of 34–53 CGG repeats.

Strong Linkage Disequilibrium for Fragile X Syndrome in Quebec

Linkage disequilibrium between fragile X syndrome and flanking polymorphic markers has been reported in several countries (e.g., see Richards et al. 1992; Oudet et al. 1993*a*, 1993*b*; Macpherson et al. 1994). We determined haplotypes of the fragile X chromosomes in 35

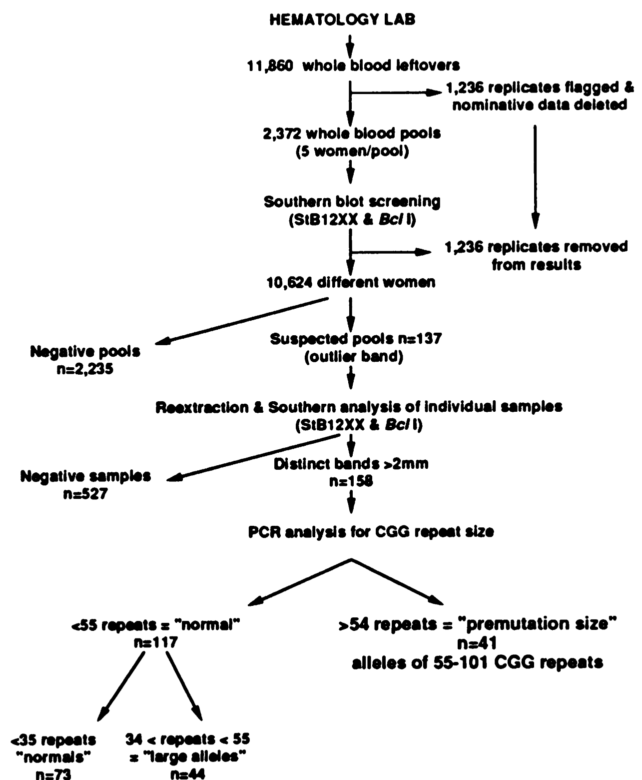

Figure 1 Study design for the identification of carriers of a large normal or premutation-size *FMR1* alleles in the general population (see text).

independent families in the Quebec population, using four flanking or intragenic polymorphic markers. These markers span a 160-kb region starting 150 kb 5′ of the *FMR1* gene and including the first 10 kb of the gene. Their positions are as follows: *DXS548* (Verkerk et al. 1991)–(140 kb)–*FRAXAC1* (Richards et al. 1992)–(10 kb)–CGG repeats–(0.7 kb)–*Ban*I RFLP (Oberlé et al. 1991)–(10 kb)–*FRAXAC2* (Richards et al. 1992).

We found evidence for very strong linkage disequilibrium between marker haplotypes and fragile X mutations in affected males in Quebec. A total of 16/35 (46%) of the fragile X chromosomes harbored the same 204-112-1-154 haplotype, compared with only 3/75 (4%) of control male chromosomes (appendixes B and C) ($\chi^2 = 29.1$, $P < .0001$). The three chromosomes with the major fragile X haplotype had >40 trinucleotide repeats at the *FMR1* locus (appendix C).

This is one of the strongest linkage disequilibria reported for the fragile X syndrome and is consistent with a founder effect in the Quebec population. Interestingly, 14 of the 16 fragile X males with the major fragile X haplotype (204-112-1-154) were most probably of French-Canadian descent, as determined by the names

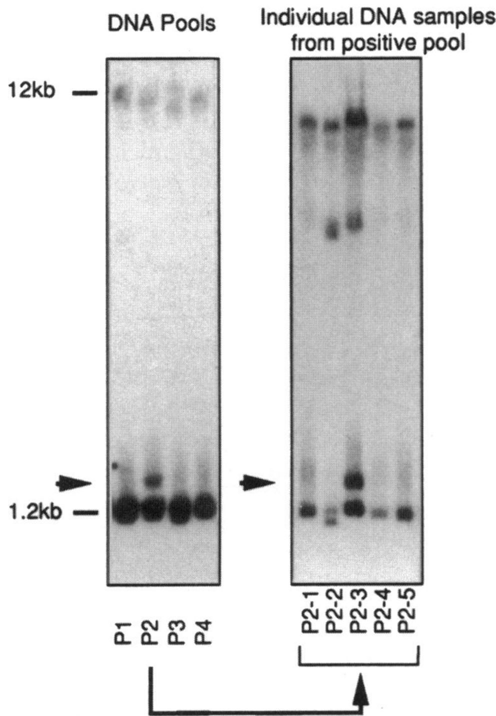

Figure 2 Southern blot of analysis of four pools of five samples each (*left*) and individual samples of a positive pool (*right*), by *Bcl*I digestion followed by hybridization to probe StB12XX (see Material and Methods). Individual sample P2-3 from pool P2 had an allele of 61 CGG repeats (or a Δ of 93 bp).

of both parents of the carrier mother. If the most distant marker (*DXS548*) is not taken into account, then 23/35 (66%) of the fragile X chromosomes have the same haplotype (112-1-154). Only four of the mothers of these males do not have French-Canadian family names for both of their parents. Only in Finland has a stronger founder effect been reported (Oudet et al. 1993*b*); there 19/26 (73%) of fragile X chromosomes harbored a particular *DXS548-FRAXAC2* haplotype.

Allelic Association of FMR1 *Premutation-Size Alleles Identified by Screening*

If the premutation-size alleles identified by screening have the same risk of expanding to full mutations as do those in fragile X families, then we expect to find a pattern of allelic association similar to that of fragile X haplotypes. We estimated conditional allele frequencies of the polymorphisms for various size classes of CGG repeats. The sample comprised data from the following individuals: a random sample of control females ($n = 68$), control males ($n = 96$), carriers of premutation-size alleles ($n = 41$ females and 1 male), fragile X males ($n = 35$), carrier women of large normal alleles (35–54 CGG repeats; $n = 44$), and carrier women of alleles of >35 repeats who were from analyzed pools ($n = 73$). We divided CGG-repeat alleles into seven size classes, according to the antimodes observed in the allele distribution of samples. Likelihood-ratio tests of linkage disequilibrium between CGG-repeat size and *DXS548*, *FRAXAC1*, and *FRAXAC2* were all highly significant (fig. 3; LR statistic, df = 18, 165.7, 136.4, and 126.4, respectively; all $P < .0001$), with the exception of the *Ban*I RFLP (data not shown; LR statistic, df = 6, 14.9; $P < .025$).

The results of the analysis for the three most informative polymorphic markers (*DXS548*, *FRAXAC1*, and *FRAXAC2*) are shown in figure 3. The most frequent

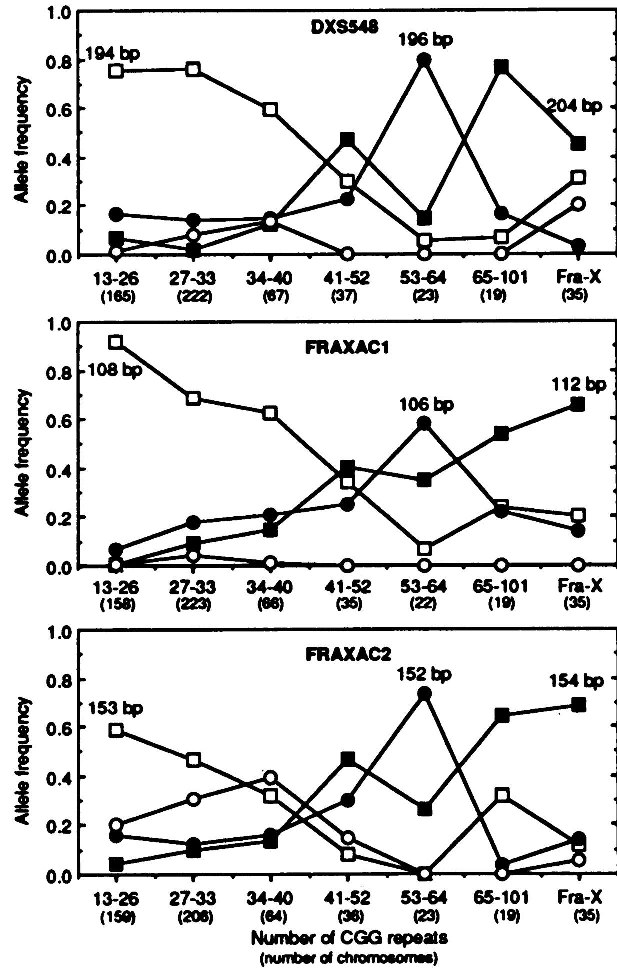

Figure 3 Conditional allele frequencies of three markers (*DXS548*, *FRAXAC1*, and *FRAXAC2*) closely linked to the *FMR1* locus for CGG-repeat-allele size classes. The relative frequencies of the most common allele of each marker for the fragile X males (*blackened squares*), the most common allele in control males and females (*unblackened squares*), the allele estimated to be most frequent in the 53–64-repeat-size class (*black dots*), and all other alleles combined (*unblackened circles*) are given for each CGG-repeat-allele size class. The sizes (in bp) of the three major alleles of each marker are indicated, and the number of chromosomes in each size class is given in parentheses.

allele of each marker in fragile X males is also the one most frequently associated with *FMR1* alleles containing 65–101 CGG repeats in the 41 carriers of premutation-size alleles. The conditional allele frequencies for the 65–101-repeat-size class are .75, .54, and .6 for *DXS548*-204, *FRAXAC1*-112, and *FRAXAC2*-154, respectively. Interestingly, the most frequent fragile X–associated allele of each of these markers is also relatively more frequent in the "grey zone" size class of 41–52 CGG repeats. However, unexpectedly, this is not the case for *FMR1* alleles in the 53–64-CGG-repeat class, the lower range of premutation-size alleles. The most frequent alleles in this size class are *DXS548*-196, *FRAXAC1*-106, and *FRAXAC2*-152, which have relative frequencies of .75, .58, and .74, respectively. These alleles are neither those of the major fragile X haplotype nor those of the most frequent haplotype of CGG-repeat alleles in the normal size range. If the most common haplotype in the lower-size class of premutation-size alleles (53–64 CGG repeats) is composed of the most frequent alleles of each marker (196-106-152), then this inferred alternative haplotype is rare (1/35) in fragile X males as well as in control males (2/77) (appendixes A–C).

It appears that 12/41 (29%) of premutation carriers have the four alleles of the major fragile X haplotype (204-112-1-154), whereas 15/41 (37%) have the four alleles that might define an alternative major premutation haplotype (196-106-1-152) (appendix A). Interestingly, there are only 5 of the total 12 possible carriers of the major fragile X haplotype, compared with 13 of the total 15 possible carriers of the inferred alternative major haplotype, among the 23 women who carry a premutation-size allele in the size class of 53–64 CGG repeats. Therefore, the 18 women who are carriers of larger premutation-size alleles (65–101 repeats) could more likely be carrying the major fragile X haplotype than the alternative haplotype. Among 78 fully genotyped women who carried an allele of 34–52 CGG repeats, the four marker alleles constituting the major fragile X haplotype could be found in 13 individuals, and the four marker alleles of the putative alternative major haplotype could be found in 8 other individuals (data not shown).

Discussion

We used an efficient screening design for large sample sizes (Rousseau et al. 1994*b*). Pooling of five samples at a time and the use of *Bcl*I restriction enzyme in Southern blot analysis is adequate for premutation detection, but it does not allow for a sensitive detection of those fragile X full mutations that generate widespread smears (Rousseau et al. 1992). The cost of this Southern blot technique, including DNA purification and technician time, was ~$12 US for individual sample analyses (Rousseau et al. 1994*b*). Other enzyme-probe combinations should detect, with great sensitivity, premutations as well as full mutations or mosaics and should allow the determination of the methylation status of the *FMR1* CpG island, an important parameter in the clinical expression of fragile X mutations (Rousseau et al. 1994*a*).

Prior to the present study, there were observations suggesting that the prevalence of fragile X premutation carriers could be relatively high, i.e., ~1/500 women (Fu et al. 1991; Rousseau et al. 1991, 1994*a*). However, the estimates were not precise, because sample sizes were small, resulting in very wide confidence intervals. Recently, the finding of one premutation-size allele of 75 CGG repeats in a screening of 1,538 independent chromosomes was reported (Reiss et al. 1994), and the investigators suggested that alleles of 46–54 CGG repeats might also be meiotically unstable.

We undertook a population-based study to obtain a reliable estimate of the prevalence of carriers of premutation-size *FMR1* alleles. Among the 10,624 women who were initially screened by Southern blot, we found 41 carriers of premutation-size alleles in the 55–101 range of trinucleotide repeats. It was not possible to analyze transmission of these premutation-size alleles, because the samples were assessed in an unlinked-anonymous design. The sample population was derived from sequential samples that were left over after routine hematology laboratory analyses of outpatients of a general hospital. No significant sampling bias of carriers of fragile X premutations is expected. We have estimated that 1/259 women (95% confidence interval 1/373–1/198) are carriers of an *FMR1* allele of >54 CGG repeats. In this sample, we observed that the prevalence of female carriers of alleles of ⩾46 CGG repeats was 1/177 (95% confidence interval 1/237–1/141); but this may be an underestimate, because we may not have identified all samples with such allele sizes (because of limitations of the Southern blot screening protocol).

In fragile X families, it has been shown (Fu et al. 1991; Heitz et al. 1992; Yu et al. 1992) that the probability of expansion to a full mutation increases with the size of a premutation. For premutations of 54–71 CGG repeats, the probability of expansion to a full mutation has been estimated to be ~10%, when a partial adjustment for ascertainment through fragile X males is used (Heitz et al. 1992). In the case of the premutation-size alleles identified in the present study, the question remains as to whether they have the same probability of expansion to a full mutation on transmission as do premutations of the same size in fragile X families (see below).

The current understanding of the sequence of events leading from a normal size CGG-repeat array to fragile X premutations and, subsequently, to full mutations

suggests a limited number of initial mutations from normal alleles with ~30 repeats to slightly larger alleles that become increasingly unstable on transmission to the next generation as the trinucleotide-repeat allele increases in size (Morton and Macpherson 1992; Oudet et al. 1993*a*). It has been shown that the instability is correlated not only with the size of the repeat array but also with the length of uninterrupted CGG repeats (Eichler et al. 1994; Hirst et al. 1994; Kunst and Warren 1994; Snow et al. 1994). Linkage disequilibrium with flanking markers would be the result of a very small number of initial events leading to progressive instability in the following generations. Whether this linkage disequilibrium reflects the presence of a haplotype predisposing to instability or a few random mutational events having generated slightly expanded alleles with intrinsic instability remains to be determined. However, the latter hypothesis appears to be favored by the analysis of the CGG-repeat-array interruptions.

We have demonstrated strong linkage disequilibrium between fragile X mutations and closely linked polymorphisms. Linkage disequilibrium for the fragile X syndrome has been reported in other populations (e.g., see Richards et al. 1992; Jacobs et al. 1993; Oudet et al. 1993*a*, 1993*b*; Macpherson et al. 1994). In the present study, there is also significant allelic association between the CGG-repeat-size classes and the flanking polymorphic markers among chromosomes from the general population. The 41 premutation-size alleles detected by screening also show linkage disequilibrium with closely linked markers when compared with normal size *FMR1* alleles, and 30% of premutation-size alleles can be inferred to be associated with the major fragile X haplotype found in Quebec fragile X males. This suggests that their probability of expanding to full mutations associated with the disease is similar to that of premutations of the same size that are found in fragile X families.

We can attempt to estimate the contribution of the female carriers of premutation-size alleles to the frequency of full mutations in the next generation of the Quebec population, using (*a*) estimated probabilities of expansion of a premutation to a full mutation conditional on CGG-repeat size in fragile X families (Heitz et al. 1992), with the assumption that these probabilities apply to all *FMR1* premutation-size alleles identified by screening; and (*b*) the estimated prevalence of the carriers of premutation-size alleles of different sizes, with the assumprion that these carriers do not have reduced fitness. With these assumptions, a frequency of 42 carriers of fragile X full mutations (carriers born to women who carry a premutation-size allele) per 100,000 births would be predicted (table 1). Estimated incidences of fragile X cases are 40/100,000–67/100,000 births (or 1/2,500–1/1,500). Since we do not have an accurate estimate of the frequency of fragile X mental retardation in the Quebec population, we cannot determine with precision (1) whether the predicted frequency of new carriers of full mutations due to conversion of maternal premutation-size alleles in the next generation is an overestimate and (2) what proportion of fragile X cases is attributable to this source. Since all fragile X cases are not born from mothers carrying a premutation—i.e., some are children of carriers of full mutations (Oberlé et al. 1991; Rousseau et al. 1991) that were not detected in this study—it is unlikely that all the premutation-size alleles detected by population screening are at high risk of expanding directly to a full mutation in the next generation.

On the basis of our study, there appear to exist in the Quebec population three general classes of *FMR1* premutation-size alleles: (1) alleles on the fragile X–associated haplotype (likely to be highly unstable), (2) alleles on the most frequent normal haplotype (which might be, on average, more stable but that still account for 10%–20% of the fragile X chromosomes), and (3) alleles on a haplotype that is infrequent in both normal and fragile X chromosomes, the origin and stability of which merit further consideration.

The distributions of the marker alleles among the CGG-repeat classes suggest the existence of an alternative major haplotype (196-106-1-152) in the 53–64-repeat class. This haplotype (inferred from the conditional allele frequencies in this class) could be found in, at most, 13/23 (57%) of alleles in the 53–64-repeat class; in, at most, 8/64 (13%) of women with large alleles (34–52 repeats); in none of 54 randomly selected women (108 chromosomes); and in only 1/35 of fragile X chromosomes. This suggests that there may be a proportion of the premutation-size alleles that have a very small risk of further expansion to fragile X mutations. Alternatively, and by analogy with the hypothesis of Kunst and Warren (1994), the CGG-repeat alleles of this haplotype may be unstable but have not yet reached equilibrium and may contribute to an increase in the frequency of fragile X syndrome in future generations. In our sample, we estimate the proportion of premutation-size alleles associated with the alternative haplotype to be approximately one-third. Hence, the risks of expansion of premutation-size alleles may be heterogeneous but likely could be recognized by either use of the associated haplotype or, better, by AGG interspersion analysis.

The origin of the alternative major haplotype associated with premutation-size alleles (196-106-1-152) is presently unclear. However, after adjustment is made for differences in marker-allele designations among laboratories, by matching allele frequencies in British (Macpherson et al. 1994), Australian (Richards et al. 1992), American (Richards et al. 1992), and French (Oudet et al. 1993*a*) populations, the alternative haplotype associated

Table 1

Expected Frequency of Fragile X Full Mutations from Premutation-Size Allele

No. of CGG Repeats	Probability of Expansion of Premutation Allele to Full Mutation (r)[a]	Prevalence/100,000 Women (p)[b]	Expected No. of Carriers of a Full Mutation/ 100,000 Births ($pr/2$)[c]
55–71	.10	301	15
72–88	.56	66	19
89–105	.90	19	8
Total		386	42

[a] SOURCE.—Heitz et al. (1992).
[b] SOURCE.—Present study.
[c] Calculated under the assumption that the risk of expansion of the CGG-repeat allele was the same as that in fragile X families.

with premutation-size alleles in the Quebec population may be seen to correspond to (1) the most frequent haplotype in Australian (*FRAXAC1*-106 and *FRAXAC2*-152; 38%) and British (*DXS548*-196, *FRAXAC1*-106, and *FRAXAC2*-152; 16%) fragile X families and (2) a haplotype found in 11% of French (*DXS548*-196 and *FRAXAC2*-152) and in 34% of American (*DXS548*-196) fragile X families. If the founder chromosome bearing the major alternative haplotype in Quebec is the same as that found in Australia, Europe, and the United States, then its *FMR1* CGG-repeat array appears to be more stable in the Quebec population, since it is infrequent in fragile X families. The Quebec population arose mostly from a limited population of settlers (Charbonneau and Robert 1987), and it is possible that the founder alleles that correspond to this subgroup of premutation-size alleles carried a stabilizing AGG interspersion.

The frequency of the alternative major "premutation" haplotype (196-106-1-152) can be determined by screening for male carriers (authors' unpublished data). The study of additional fragile X families might provide confirmation of the low frequency of this third haplotype among fragile X chromosomes in the Quebec population.

The high frequency of premutation-size alleles found by screening in this study, combined with the observation that approximately one-third of them are associated with the major fragile X haplotype found in fragile X patients, raises the question of the appropriateness of population screening for carriers of *FMR1* premutation-size alleles. Because of the dynamics of *FMR1* gene mutation, the fragile X syndrome often appears in a family with no prior history of mental retardation (Nussbaum and Ledbetter 1989), which therefore limits the sensitivity of carrier detection by a proband-based approach. If population screening was to be considered, the means (population, strategy, and technique) would have to be carefully studied and validated. The risk of being a carrier and of having an affected child would have to be reliably determined for small premutation-size alleles, and the support and genetic-counseling resources would have to be available. A low-cost technique for detecting all fragile X mutations (premutations, full mutations, and mosaics) with high sensitivity and specificity is needed to identify all unaffected carrier women, since ~45% of women with full mutations are asymptomatic but have a high risk of having affected children (Rousseau et al. 1991). Current screening protocols that use Southern blot analysis or CGG-repeat PCR analysis may lack sensitivity to detect mosaic genotypes and full mutations in females. Also needed would be a comparative evaluation of a prospective population-based screening approach and a proband-based strategy aimed at improving ascertainment of fragile X syndrome cases, with follow-up screening of the families of carrier females. This latter approach would take advantage, in genetic counseling, of the use of published rates of expansion to full mutations and would be more economical. However, a proband-based approach would not be expected to identify carriers in families with no prior history of the disease; but these carriers may represent only a small (yet unknown) proportion of the high-risk premutation carriers. Furthermore, the effectiveness of this approach would be reduced by difficulties in contacting all relatives at risk, because of ethical considerations and lack of communication within families.

Acknowledgments

We thank Prof. Jean-Louis Mandel for helpful discussion and comments. We are indebted to Dr. Jacques Boulay and to the personnel of the hematology laboratory of Hospital St-François-d'Assise for their contribution and to Drs. Bernard Lemieux, Louis Dallaire, Vazken Der Kaloustian, Bertrand Fouquette, and Rachel Laframboise for samples from fragile X cases. We thank Jacqueline Dionne, Richard Réhel, Mario Vaillancourt, and Pierre DeGrand-

pré for their expert technical assistance, and we thank Richard Couture for the photographs. This work was supported by a grant from the Medical Research Council of Canada (MRC) to F.R., by a grant from the Canadian Genetic Diseases Network (Federal Networks of Centres of Excellence program) to F.R. and K.M., and by a grant from IREP for the Fonds pour la Formation des Chercheurs et l'Avancement de la Recherche (Programme de Coopération Scientifique Internationale, Québec) to F.R. F.R. is a research scholar of the MRC. E.W.K. and F.R. are also supported by La Fondation de l'Hôpital St-François-d'Assise and indirectly by Le Fonds de Recherche en Santé du Québec. P.R. holds an MRC studentship, and M.-L.M. holds an NSERC studentship.

Appendix A

Genotypes of 41 Premutation-Size Alleles Identified by Screening

DXS548	*FRAXAC1*	No. of CGG Repeats	*Ban*I	*FRAXAC2*
204/194	112/108	55/33	1/1	154/153
204/196	112/108	55/34	1/1	154/153
196/194	108/108	56/21	1/1	153/152
*196/194	*106/108	*56/31	*1/1	*152/155
*196/178	*106/108	*57/31	*1/1	*152/153
*196/194	*106/108	*57/33	*1/1	*152/153
196/196	. . ./. . .	58/24	. . ./. . .	152/155
204/196	112/108	58/24	1/1	154/150
*196/196	*106/112	*58/22	*1/1	*152/150
*196/194	*106/108	*59/32	*1/1	*152/150
*196/196	*106/106	*59/34	*1/1	*152/149
*196/194	*106/108	*60/23	*1/2	*152/153
196/194	106/108	60/33	1/1	154/153
204/194	112/106	61/22	1/1	154/153
194/194	112/108	61/34	1/2	153/152
*196/194	*106/108	*61/22	*1/2	*152/150
*196/194	*106/108	*61/33	*1/1	*152/153
*196/194	*106/108	*61/34	*1/1	*152/153
*196/194	*106/108	*61/33	*1/1	*152/155
196/196	112/108	62/26	1/1	153/152
*196/194	*106/108	*62/33	*1/1	*152/153
204/194	112/108	63/34	1/1	154/153
*196/194	*106/108	*63/31	*1/2	*152/150
204/196	112/108	66/36	1/1	154/150
204/194	112/108	67/22	1/2	154/153
204/194	108/108	67/22	1/1	154/153
194/194	108/108	68/33	1/1	153/153
204/204	112/108	70/31	1/1	154/149
204/204	108/106	70/34	1/1	153/149
204/204	112/108	71/20	1/2	154/149
204/196	112/106	71/36	1/1	154/152
204/194	108/108	71/32	1/1	154/153
204/194	112/108	72/24	1/2	154/152
196/194	112/108	74/33	1/1	153/151
204/196	108/106	74/30	1/1	154/152
204/194	108/106	76/36	1/1	154/153
*196/194	*106/112	*79/32	*1/1	*152/153
204/194	108/106	80/36	1/1	154/153
*196/204	*106/108	*86/33	*1/1	*152/153
196/194	112/106	99/15	1/1	153/151
204/194	112/108	101/35	1/1	154/153

Underlining indicates a marker allele of the major fragile X haplotype in Quebec (this allele is on the left side of each typing). These alleles constitute the inferred haplotype associated with the premutation-size trinucleotide-repeat alleles. The alternative inferred haplotype predominant among 53–64 CGG-repeat alleles is denoted by an asterisk.

Appendix B

Genotype of 35 Independent Male Carriers of Fragile X Mutations

DXS548	*FRAXAC1*	*FMR1*	*Ban*I	*FRAXAC2*	No. (Non-French Canadians)
206	112	FULL	1	154	1
204	112	FULL	1	154	16 (2)
194	112	FULL	1	154	5 (2)
192	112	FULL	1	154	1
202	108	FULL	1	152	1
198	108	FULL	2	152	1
194	108	FULL	1	154	1
194	108	FULL	1	153	3 (1)
194	108	FULL	2	153	1
202	106	FULL	1	150	1
198	106	FULL	1	152	2 (1)
*196	*106	*FULL	*1	*152	1
194	106	FULL	1	149	1

Marker alleles of 35 carrier males from unrelated fragile X families in Quebec. Underlining indicates a marker allele of the major fragile X haplotype in Quebec. The alternative inferred haplotype predominant among 53–64 CGG-repeat alleles is denoted by an asterisk. In addition to the 35 individuals indicated, a young carrier of a premutation-size allele of 80 CGG repeats, who was referred for mental retardation, was also tested; this individual was found to harbor the major fragile X haplotype.

Appendix C

Genotypes of 96 Independent Normal Males

DXS548	*FRAXAC1*	No. of CGG Repeats	*Ban*I	*FRAXAC2*	No.
194	108	13	1	152	1
196	108	16	...	153	1
194	108	16	1	153	1
194	108	18	1	153	3
204	108	20	1	153	1
194	108	20	1	153	2
194	108	20	2	152	1
196	106	20	1	150	1
...	108	21	1	...	1
194	108	21	1	154	1
196	108	22	1	151	1
194	108	22	2	153	1
204	108	23	1	152	1
196	108	23	1	...	1
194	108	23	2	152	1
194	106	24	1	155	1
202	112	27	1	154	1
194	108	27	2	...	1
194	106	27	1	149	2

210	112	28	1	153	1
194	108	28	1	152	1
194	108	28	2	152	1
...	106	28	1	149	1
*196	*106	*28	*1	*152	1
194	106	28	1	...	1
206	112	29	1	154	1
196	112	29	1	154	1
206	110	29	1	...	1
196	108	29	1	153	1
196	108	29	2	151	1
194	108	29	1	...	1
194	108	29	1	153	2
194	108	29	2	151	1
206	112	30	1	...	1
204	112	30	1	...	1
194	110	30	1	153	1
194	108	30	1	155	1
194	108	30	1	154	2
194	108	30	1	153	5
194	108	30	1	152	3
196	106	30	1	149	1
206	112	31	1	...	1
206	112	31	1	153	1
194	110	31	1	...	1
194	108	31	1	...	1
194	108	31	1	155	1
194	108	31	1	154	1
194	108	31	1	153	6
194	108	31	2	153	1
194	108	31	1	152	4
204	106	31	...	149	1
196	106	31	1	154	1
196	106	31	1	151	1
196	106	31	1	150	1
194	108	32	1	...	1
194	108	32	1	153	6
194	108	32	1	152	1
...	108	33	1	153	1
194	108	33	1	...	2
...	106	33	1	149	1
196	106	33	1	...	1
202	112	34	1	153	1
194	108	35	1	152	1
198	106	36	1	150	1
196	112	38	1	...	1
196	108	39	2	152	1
204	112	42	1	154	2
*196	*106	*45	*1	*152	1
204	112	48	1	154	1

Underlining indicates a marker allele of the major fragile X haplotype in Quebec. Three of the four *FMR1* alleles with >40 CGG repeats have this haplotype. The alternative inferred haplotype predominant among 53–64 CGG-repeat alleles is denoted by an asterisk.

References

Bonthron D, Strain L (1993) Population screening for fragile-X syndrome. Lancet 341:769–770

Brown WT, Houck GE, Jeziorowska A, Levinson FN, Ding X, Dobkin C, Zhong N, et al (1993) Rapid fragile X carrier screening and prenatal diagnosis using a nonradioactive PCR test. JAMA 270:1569–1575

Brown WT, Jenkins ED, Cohen IL, Fisch GS, Wolf-Schein EG, Gross A, Waterhouse L, et al (1986) Fragile X and autism: a multicenter survey. Am J Med Genet 23:341–350

Bundey S, Norman E (1993) Population screening for fragile-X syndrome. Lancet 341:770

Charbonneau H, Robert N (1987) The French origins of the Canadian population, 1608–1759. In: Harris RC (ed) Historical atlas of Canada. Vol 1: From the beginning to 1800. University of Toronto Press, Toronto, plate 45

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the *EM* algorithm. J R Stat Soc [B] 39:1–38

Eichler EE, Holden JJA, Bradley W, Popovich BW, Reiss AL, Snow K, Thibodeau SN, et al (1994) Length of uninterrupted CGG repeats determines instability in the FMR1 gene. Nat Genet 8:88–94

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

Heitz D, Devys D, Imbert G, Kretz C, Mandel JL (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

Heitz D, Rousseau F, Devys D, Saccone S, Abderrahim H, LePaslier D, Cohen D, et al (1991) Isolation of sequences that span the fragile X and identification of a fragile-X-related CpG island. Science 251:1236–1239

Hirst MC, Grewal PK, Davies KE (1994) Precursor arrays for triplet repeat expansion at the fragile X locus. Hum Mol Genet 9:1553–1560

Howard-Peebles PN, Maddalena A, Black SH, Schulman JD (1993) Population screening for fragile-X syndrome. Lancet 341:770

Jacobs PA, Bullman H, Macpherson J, Youings S, Rooney V, Watson A, Dennis NP (1993) Population studies of the fragile(X): a molecular approach. J Med Genet 30:454–459

Kolehmainen K (1994) Population genetics of fragile X: a multiple allele model with variable risk of CGG repeat expansion. Am J Med Genet 51:428–435

Kunst CB, Warren ST (1994) Cryptic and polar variation of the fragile X repeat could result in predisposing normal alleles. Cell 77:853–861

Laird CD (1987) Proposed mechanism of inheritance and expression of the human fragile-X syndrome of mental retardation. Genetics 117:587–599

Macpherson JN, Bullman H, Youings SA, Jacobs PA (1994) Insert size and flanking haplotype in fragile X and normal populations: possible multiple origins for the fragile X mutation. Hum Mol Genet 3:399–405

Macpherson J, Harvey J, Curtis G, Webb T, Heitz D, Rousseau F, Jacobs P (1992) A reinvestigation of thirty three fragile(X) families using probe StB12.3. Am J Med Genet 43:905–912

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

Nussbaum RL, Ledbetter DH (1989) The fragile X syndrome. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) The metabolic basis of inherited disease, 6th ed. McGraw-Hill, New York, pp 327–341

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*a*) 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

Oudet C, von Koskull H, Nordström AM, Peippo M, Mandel JL (1993*b*) Striking founder effect for the fragile X syndrome in Finland. Eur J Hum Genet 1:181–189

Palomaki GE, Haddow JE (1993) Is it time for population-based prenatal screening for fragile-X? Lancet 341:373–374

Pembrey ME, Winter RM, Davies KE (1985) A premutation that generates a defect at crossing-over explains the inheritance of fragile X mental retardation. Am J Med Genet 21:709–711

Pieretti M, Zhang F, Fu YH, 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

Reiss AL, Kazazian HH Jr, Krebs CM, McAughan A, Boehm CD, Abrams MT, Nelson DL (1994) Frequency and stability of the fragile X premutation. Hum Mol Genet 3:393–398

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. Nat Genet 1:257–260

Rousseau F (1994) The fragile X syndrome: implications of molecular genetics for the clinical syndrome. Eur J Clin Invest 24:1–10

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

Rousseau F, Heitz D, Biancalana V, Oberlé I, Mandel JL (1992) On some technical aspects of direct DNA diagnosis of the fragile X syndrome. Am J Med Genet 43:197–207

Rousseau F, Heitz D, Tarleton J, MacPherson J, Malmgren H, Dahl N, Barricoat A, et al (1994*a*) A multicenter study on genotype-phenotype correlations in the fragile X syndrome, using direct diagnosis with probe StB12.3: the first 2,253 cases. Am J Hum Genet 55:225–237

Rousseau F, Réhel R, Rouillard P, DeGrandpré P, Khandjian EW (1994*b*) High throughput and economical mutation detection and RFLP analysis using a mini-method for DNA preparation from whole blood and acrylamide gel electrophoresis. Hum Mutat 4:51–54

Sherman S (1991) Epidemiology in fragile X syndrome. In: Hagerman RJ, Sylverman AC (eds) Diagnosis, treatment,

Am. J. Hum. Genet. 57:1006–1018, 1995

and research. Johns Hopkins University Press, Baltimore, pp 69–86

Sherman SL, Jacobs PA, Morton NE, Froster-Iskenius U, Howard-Peebles PN, Nielsen KB, Partington MW, et al (1985) Further segregation analysis of the fragile X syndrome with special reference to transmitting males Hum Genet 69:289–299 (erratum, Hum Genet 71:184–186 [1985])

Snow K, Tester DJ, Kruckeberg KE, Schaid DJ, Thibodeau SN (1994) Sequence analysis of the fragile X trinucleotide repeat: implications for the origin of the fragile X mutation. Hum Mol Genet 3:1543–1551

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 YH, 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

Vincent A, Heitz D, Petit C, Kretz C, Oberlé I, Mandel JL (1991) Abnormal pattern detected in fragile-X patients by pulsed-field gel electrophoresis. Nature 349:624–626

Weir BS (1990) Genetic data analysis: methods for discrete population genetic data. Sinauer, Sunderland, MA

Yu S, Mulley J, Loesch D, Turner G, Donnelly A, Gedeon A, Hileen D, et al (1992) Fragile-X syndrome: unique genetics of the heritable unstable element. Am J Hum Genet 50:968–980

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