American Journal of Medical Genetics 64:220–225 (1996)

Genetic Variation and Evolutionary Stability of the FMR1 CGG Repeat in Six Closed Human Populations

Evan E. Eichler and David L. Nelson

Department of Molecular and Human Genetics, Human Genome Center, Baylor College of Medicine, Houston, Texas

In an attempt to understand the allelic diversity and mutability of the human FMR1 CGG repeat, we have analyzed the AGG substructure of this locus within six genetically-closed populations (Mbuti pygmy, Baka pygmy, R. surui, Karitiana, Mayan, and Hutterite). Most alleles (61/92 or 66%) possessed two AGG interspersions occurring with a periodicity of one AGG every nine or ten CGG repeats, indicating that this pattern is highly conserved in all human populations. Significant differences in allele distribution were observed among the populations for rare variants possessing fewer or more AGG interruptions than the canonical FMR1 CGG repeat sequence. Comparisons of expected heterozygosity of the FMR1 CGG repeat locus with 30 other microsatellite loci, demonstrated remarkably similar levels of polymorphism within each population, suggesting that most FMR1 CGG repeat alleles mutate at rates indistinguishable from other microsatellite loci. A single allele (1 out of 92) was identified with a large uninterrupted tract of pure repeats (42 pure CGG triplets). Retrospective pedigree analysis indicated that this allele had been transmitted unstably. Although such alleles mutate rapidly and likely represent evolving premutations, our analysis suggests that in spite of the estimated frequency of their occurrence, these unstable alleles do not significantly alter the expected heterozygosity of the FMR1 CGG repeat in the human population.

Received for publication September 11, 1995; revision received December 28, 1995.

Evan E. Eichler is now at Human Genome Center, L-452, Biology and Biotechnology Research Program, Lawrence Livermore National Laboratory, Livermore, CA 94551.

Address reprint requests to David L. Nelson, Department of Molecular and Human Genetics, Baylor College of Medicine, 902E, One Baylor Plaza, Houston, TX 77030.

KEY WORDS: fragile X syndrome, X-linked mental retardation, population genetics, mutation rates, CGG repeat

INTRODUCTION

The fragile X syndrome is the most common form of inherited mental retardation in man. Although a considerable range in the incidence of this disease has been reported [Turner et al., 1986; Turner et al., 1992; Webb and Bundey, 1991; Webb et al., 1986; Webb et al., 1986], epidemiological investigations considering cytological, molecular and clinical aspects of the disease indicate a frequency of approximately 1/2,000 [Morton et al., 1995]. Cases of the fragile X syndrome, furthermore, have been observed among diverse human populations, suggesting no particular ethnic predilection to the development of the disease [Howard-Peebles and Stoddard, 1980; Rhoads, 1984; Rivera et al., 1981; Soysa et al., 1982; Venter et al., 1981].

The disease is caused primarily by the expansion of a CGG repeat, located in the first exon of the FMR1 gene [Fu et al., 1991; Kremer et al., 1991; Oberle et al., 1991; Yu et al., 1991]. Expansion of the repeat beyond 200 repeat units results in hypermethylation of the putative promoter region, abolition of FMR1 transcription and concomitant mental retardation [Pieretti et al., 1991; Sutcliffe et al., 1992]. FMR1 CGG repeat alleles have, in general, been assigned to one of three categories based on total repeat length [Fu et al., 1991; Snow et al., 1993] . Normal (stable) alleles length vary from 5–55 repeat units in length; premutation (unstable) alleles are generally larger (60–200) and progress rapidly, within a few generations to full mutations (>200–2,000 repeats) associated with disease. However, considerable overlap has been reported between normal and premutation alleles, suggesting that total repeat length may not be a sufficient indicator of stability at this locus [Fu et al., 1991; Reiss et al., 1994; Richards et al., 1992; Zhong et al., 1993]. Recent investigations into the AGG substructure of the FMR1 CGG repeat have provided the molecular basis for this overlap [Eichler et al., 1994; Hirst et al., 1994; Kunst and Warren, 1994; Snow et al., 1994]. Most alleles among Caucasian populations possess AGG interruptions occurring with a periodicity of once every nine or ten CGG

repeats [Eichler et al., 1994; Hirst et al., 1994; Kunst and Warren, 1994; Snow et al., 1994]. Infrequently (~2.0%), rare alleles have been observed in the normal population which are deficient in AGG interruptions and also possess a longer tract of uninterrupted repeats (>20 pure repeats). Intergenerational transmission studies of premutation and unstable alleles in human pedigrees show a correlation between the length of this tract of pure repeats and the propensity for instability, with unstable transmissions being observed around lengths of 35 pure repeats. The scarcity of AGG interruptions among premutation and unstable alleles, furthermore, has suggested that the loss of AGG interruptions may be a pivotal event in predisposing normal human alleles to disease.

The FMR1 CGG repeat belongs to a class of short tandem repetitive DNA sequences (STRs) known as microsatellites. Microsatellites are abundant in the human genome and are frequently polymorphic [Weissenbach et al., 1992], with the informativeness of these markers dependent on the purity of the repeat [Riggins et al., 1992; Weber, 1990]. Relative to nonrepetitive DNA, microsatellites have been shown to mutate rapidly with rates ranging from 1×10^{-4} to 1×10^{-5} [Bell and Torney, 1993; Edwards et al., 1992; Hästbacka et al., 1992; Kwiatkowski et al., 1991]. The propensity for microsatellites to mutate is property of their repetitive DNA structure and is likely mediated by processes which involve the formation of mispaired slip-strand intermediates [Bell and Torney, 1993; Levinson and Gutman, 1986; Schlötterer and Tautz, 1992]. In addition, to normal microsatellite polymorphic variability, it has been suggested based on the FMR1 CGG microsatellite might, also, incur mutations (rate 2.5×10^{-4}) which are predisposed to the development of disease [Morton and Macpherson, 1992]. Such "rare" new mutant alleles increase in size at a rate of 1×10^{-2}/generation, requiring an estimated 90 generations to progress to the full mutation or disease state [Chakravarti, 1992]. The existence of these intermediates has been used to explain the high incidence of the fragile X syndrome, despite strong selection pressure against the fully mutated allele [Morton and Macpherson, 1992]. Owing to the longevity of predisposed alleles in the population and their increased propensity to mutate, one might expect such alleles to contribute significantly to the polymorphism of the FMR1 CGG repeat [Morris et al., 1995; Morton and Macpherson, 1992].

Genetically closed populations provide a unique experimental context to assess microsatellite instability and to investigate the mutation rates of alleles predisposed to genetic disease [Bowcock et al., 1994; Edwards et al., 1992; Hästbacka et al., 1992]. We have examined both sequence content and length variation of the FMR1 CGG repeat in six diverse and genetically closed populations, namely, the Mbuti pygmy, Baka pygmy, Surui, Karitiana, Mayans, and Hutterites. Despite extensive genetic isolation and distant (200,000 years) [Bowcock et al., 1994; Dorit et al., 1995; Horai et al., 1995] genetic separation among these various populations, the substructure of the FMR1 CGG repeat is remarkably conserved, with the AGG interruptions occurring consistently with a periodicity of once every nine or ten CGG repeats. Although ethnic differences exist, these observation suggest that the canonical symmetrical array structure (CGG)9AGG(CGG)9AGG(CGG)9 likely represents one of the ancestral states in the human population. We have calculated the expected heterozygosity based on repeat length and sequence content for the FMR1 CGG repeat locus in each population. Our measures of polymorphic heterogeneity of the FMR1 CGG repeat closely approximate previous estimates of average heterozygosity for these populations based on 30 microsatellite loci [Bowcock et al., 1994; Kidd et al., 1991], suggesting that the FMR1 CGG repeat mutates at a frequency similar to other microsatellites. In our analysis, a single allele (1/92) was identified with a long uninterrupted tract of pure CGG repeats (42 repeats). Retrospective pedigree analysis indicated that this allele was unstably transmitted consistent with previous models that an instability threshold exists beyond 35 repeats. Based on the observed frequency of this unstable evolving premutation, the size of the founding population and the occurrence of the last genetic bottleneck in this group, we calculate that predisposing mutations occur at a rate of 2.4×10^{-4}. Such mutations, we estimate, generate a small proportion (~1%) of all FMR1 CGG repeat alleles and do not contribute significantly to the heterozygosity of this locus, suggesting a rapid progression to the disease state.

MATERIALS AND METHODS

DNA Samples

DNA samples for the Surui (Rondônia, Brazil), Mayan (Campeche, Mexico), Karitiana (Brazil), Baka pygmy (Cameroon), and Mbuti pygmy (Zaire) were obtained from 41 unrelated male individuals from the Human Genome Diversity Project as described previously [Bowcock et al., 1994; Kidd et al., 1991]. Samples from the Hutterite population (North America) were obtained from unrelated male individuals representing all three branches (Lehrerleut, Dariusleut, and Schmiedeleut) from collections maintained at Integrated Genetics. Initially, a total of 15 Hutterite samples were analyzed for repeat length and CGG repeat substructure (Table I). An additional 26 samples (22 unrelated and four from pedigrees) were analyzed subsequently in order to ascertain stability of uninterrupted pure CGG repeat alleles in this population.

AGG Interspersion Analysis

FMR1 CGG repeat length and AGG substructure were determined as previously described [Eichler et al., 1994; Fu et al., 1991]. Interspersion configuration is symbolized as follows: a "+" sign denotes the position of an AGG interruption, while a number refers to the length of contiguous uninterrupted CGG repeats. A "9+9+9" interspersion pattern, for example corresponds to the sequence (CGG)9AGG(CGG)9AGG(CGG)9 relative to the direction of FMR1 transcription.

Statistical Analysis

Significant differences in the proportion of different allele variants among the populations were assessed

TABLE I. FMR1 CGG Repeat Substructure Among Six Closed Human Populations*

AGG configuration	Number of Alleles Within Each Population						
	Baka	Hutterite	Karitiana	Mayan	Mbuti	R. Surui	Total
9+7	1						1
9+9					1		1
10+9		9					9
13+9		1					1
9+17					2		2
30.00		1					1
10+9+9	1	21	6	7		8	43
9+9+10	1				2		3
9+9+9	5	2	1	3	4	4	19
10+9+10		1					1
11+9+9		1					1
10+10+9	1	2					3
10+9+12		1					1
10+9+9+9		1					1
9+9+10+9			1		1		2
10+9+10+9			2				2
42.00		1					1
Total	9	41	10	10	10	12	92

*The distribution of 92 FMR1 CGG repeat alleles among the six different populations is summarized. AGG configuration (see Materials and Methods) is arranged in increasing order of repeat length.

using probabilities obtained from the one-tailed Fisher exact test [Zar, 1984]. Nei's unbiased estimate of expected heterozygosity was employed to assess the polymorphism of the FMR1 CGG repeat locus [Arinami et al., 1993]. These values were compared to the average heterozygosity and standard errors obtained from these populations for 30 previously described microsatellite loci [Bowcock et al., 1994]. In order to estimate the frequency of new mutations predisposing to instability and fragile X disease in the Hutterite population, a set of previously reported formulae was used [Hästbacka et al., 1992]. The following assumptions were made in this analysis: i) the unstably transmitted allele (IG500) represents an evolving premutation, ii) the predisposing event has occurred since the founding of the Hutterite population, and iii) the frequency of unstable alleles observed in the Hutterite sample (1/42 or 0.024) represents the proportion in the total population. In order to calculate the growth rate of the population (d), the following equation was solved:

$$n = n_o e^{gd}$$

where n represents the current total population size; n_o is the number of founder chromosomes and g indicates the number of generations since the founding of the current population. Owing to extensive genealogical records extending back to the most recent genetic bottleneck of the current Hutterite population in 1745, the following values were used; g = 9; n_o = 32; and n = 10,000 [Ziglschmid, 1947]. Solving for d and using the proportion of unstable alleles in the population (π_{obs} = 0.024), we estimated the rate (a) of a new mutation generating a lineage giving rise to 500 (N) unstable alleles in the existing population as:

$$\pi_{obs} = \alpha dN$$

with a standard deviation of $\pm 2\ \alpha d$.

RESULTS AND DISCUSSION

Table I summarizes the distribution of length and AGG interspersion substructure for the FMR1 CGG repeat for 92 chromosomes derived from six genetically closed populations (Baka, Mbuti, Mayan, Surui, Karitiana, and Hutterite). Most alleles (61/92 or 66.3%) possess substructures of the type 9+9+9 or 10+9+9 (see Materials and Methods for nomenclature). The predominance of these repeat substructures suggests that they existed in the human population, prior to the divergence of the races (200,000 years ago) [Bowcock et al., 1994; Dorit et al., 1995; Horai et al., 1995] and/or that there has been a strong selection pressure to maintain AGG interruptions with a periodicity of one AGG every 9/10 CGG repeats. Despite this homogeneity of repeat substructure, ethnic-specific allele variants have been observed. For example, the Amerindian tribe, Karitiana, is significantly enriched ($P < .01$; Fisher's one-tailed exact test) for highly interspersed alleles (>2 AGG interruptions). Conversely, small alleles with a single interruption (of the type 10+9), occur almost exclusively among the caucasian Hutterite population. The significant ($P < 0.0001$, Fisher's test) absence of these alleles among the Amerindians and African groups confirms previous open population studies [Eichler et al., 1995] that the 10+9 allele variant is largely restricted to Caucasian populations.

Expected heterozygosities for each population were calculated based on total repeat length and FMR1 CGG repeat substructure (Table II). An average heterozygosity of 64.1% was calculated from these six groups. The heterozygosity of this locus is slightly less than previously reported estimates [Arinami et al., 1993; Fu et al., 1991; Snow et al., 1993]. This is to be expected since earlier determinations of heterozygosity were based on relatively open, pan-mictic populations. The populations under investigation in this study have been subject to

TABLE II. Comparisons of FMR1 CGG Repeat Heterozygosities*

	Comparison of FMR1 CGG repeat heterozygosities			
	FMR1 microsatellite		Other microsatellites	
Population	%H with AGG	%H without AGG	Average %H	2sd
Baka pygmy	77.7	75.6	71.2	6.01
Mbuti pygmy	82.2	82.2	77.9	5.24
African average	80.0	78.9	74.5	5.63
Surui	49.9	49.9	59.7	4.69
Karitiana	60.0	55.6	47.4	5.70
Mayan	43.8	43.8	65.4	5.88
Amerindian average	51.2	49.8	57.5	5.42
Hutterite	71.1	70.7	—	—

*The expected heterozygosity (%H) for each population was calculated using Nei's unbiased estimate. Heterozygosities were determined based on repeat length differences (%H without AGG) and on the substructure of the repeat (%H with AGG). The values obtained were compared against previously described average heterozygosities and standard deviations (2 sd) [Kidd et al., 1991] for 30 other autosomal microsatellite loci in these populations.

more extensive founder effects and genetic bottlenecks due to their geographic and cultural isolation. As a result, greater genetic homogeneity is demonstrated [Bowcock et al., 1994; Kidd et al., 1991]. Surprisingly, only a relatively small proportion of the genetic variability of the FMR1 CGG repeat can be attributed to differences in the AGG substructure (approximately 1.1%; 64.1 vs. 63.0 average % H, with and without considering AGG interruptions, respectively). This suggests, once again, that the positions of the AGG interruptions have been largely invariant over the passage of time.

We have compared our values of heterozygosity of the FMR1 CGG repeat for each population with previous averaged estimates of heterozygosity from 30 additional microsatellite loci [Bowcock et al., 1994; Kidd et al., 1991]. Remarkably, these comparisons show that the FMR1 CGG repeat has been as polymorphic (within 2 standard deviations of the mean; see African pygmies) or less polymorphic (below 2 standard deviations of the mean; see Amerindian tribes) than the average of 30 microsatellite loci in the human genome. The average heterozygosity of the FMR1 CGG repeat for the Amerindian tribes (H = 49.2%) is dramatically lower than the heterozygosity for the African pygmies (H = 78.9%) (Table II). This parallels previous observations with other microsatellites which determined Amerindian and African heterozygosities of 57.5% and 74.5%, respectively [Bowcock et al., 1994]. These differences have been used to suggest that the founding populations of the African continent were genetically more heterogeneous and/or more ancient, supporting the out-of Africa hypothesis for the origin of man [Bowcock et al., 1994; Kidd et al., 1991]. The lower estimates of heterozygosity of the FMR1 CGG repeat locus when compared to other microsatellites among the Amerindians may support recent founder effects. In this regard, it is interesting that the FMR1 CGG repeat heterozygosity values are not consistently lower among all three Amerindian tribes (Table II). Among the Karitiana, the heterozygosity is slightly larger than the average of other microsatellite loci. Owing to independent assortment of chromosomes, recombination, and genetic drift, it is conceivable that the FRAXA locus has experienced different founder effect and drift phenomenon from the other microsatellite loci. However, the general concordance of % heterozygosity between the FMR1 CGG repeat and other independent STR markers among genetically distinct populations clearly suggests that most FMR1 CGG repeat alleles mutate at frequencies similar to other microsatellite loci.

In our survey, a small proportion (2/92 or 2.2%) of FMR1 alleles possessed long uninterrupted tracts of pure CGG repeats (>20). Initially, we surveyed only 15 chromosomes of the Hutterite population and observed an unusual FMR1 CGG repeat structure with 42 pure repeats (Table I). An additional 26 alleles were then analyzed from this population identifying a second allele with 30 pure repeats (Table I). Owing to the extensive genealogical data for the Hutterites and the existence of pedigree DNA material which could be obtained, intergenerational transmission studies of repeat length stability were performed. Our analysis showed that the 30 pure CGG repeat allele did not segregate to the two available offspring in the pedigree, while the 42 pure repeat demonstrated one stable and one unstable transmission (42 to 43 transition) (see Fig. 1). This result confirms previous studies which indicate that an instability threshold exists for the FMR1 CGG repeat at lengths of 34–37 uninterrupted CGG repeats [Eichler et al., 1994]. Such unstable alleles in the human population, it has been postulated, likely represent evolving premutation alleles predisposed to disease [Eichler et al., 1994; Snow et al., 1994]. Since the characteristics of the Hutterite founding population are well documented, we attempted to calculate the mutation rate for predisposing alleles to instability and disease. Using previously defined formula for genetically closed populations (see Materials and Methods), we estimate that the conversion rate of normal (polymorphic) FMR1 CGG repeat alleles to "normal" alleles predisposed to eventual disease is approximately $2.4 \times 10^{-4} \pm 3.0 \times 10^{-4}$. Although the volatility of this estimate is relatively high due to the observation of a single occurrence (demonstrated by the relatively large standard error), our value is remarkably similar to the conversion rate (2.5 × 10-4) of N to S alleles reported by

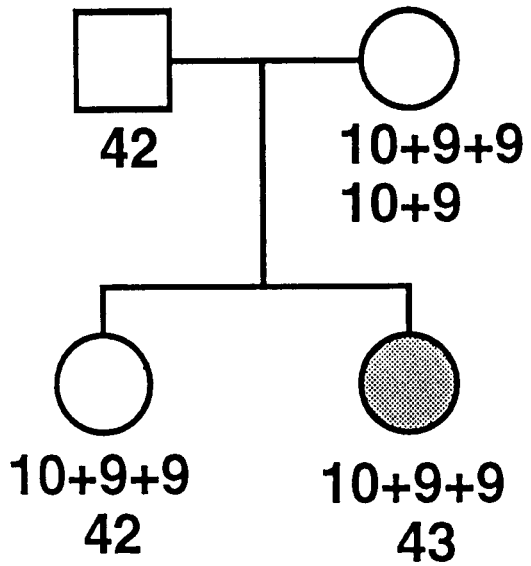

Fig. 1. Unstable Transmission of an FMR1 CGG repeat allele devoid of interruptions. Two intergenerational transmissions of an FMR1 CGG repeat allele of 42 pure repeats are shown. The unstable transmission (42 to 43 repeats) is shaded. The FMR1 CGG repeat genotype is symbolized below each individual in the pedigree.

Morton and Macpherson [Morton and Macpherson, 1992]. The substructure of the unstable allele in this population which is completely devoid of AGG interruptions, further confirms that the predisposing mutational event may be the loss of one or both AGG interruptions [Eichler et al., 1994; Hirst et al., 1994; Kunst and Warren, 1994; Snow et al., 1994]. Although further intergenerational stability and interspersion analysis of genetically closed populations will be necessary to confirm this mutation rate, our data suggests that the loss of AGG interruptions may occur as frequently as 1/2,000 in the human population.

In conclusion, we have surveyed the FMR1 CGG repeat substructure of 92 chromosomes distributed among six genetically-closed populations. Although significant ethnic differences exist, the 10+9+9 and 9+9+9 account for most interspersion patterns at this locus. Comparisons of expected heterozygosity of the FMR1 CGG repeat substructure with 30 other microsatellite loci, indicates similar levels of polymorphism, suggesting that most FMR1 CGG repeats mutate at rates equivalent to other microsatellite loci. A single large uninterrupted allele (42 repeats) was observed to be unstably transmitted in a human pedigree. Based on the frequency of such unstable alleles and the characteristics of the founder population, we estimate that the mutation rate predisposing alleles to instability and eventual disease is $2.4 \times 10\text{-}4$. The substructure of this unstable, evolving premutation, which is devoid of AGG interruptions further supports the notion that the loss of AGG interruptions may be the pivotal event in predisposing alleles to instability. However, such unstable alleles represent a small proportion (>2%) of all FMR1 CGG interspersion variants and likely do not contribute significantly to the total polymorphism of this locus.

ACKNOWLEDGMENTS

We would like to thank Dr. J. Kidd and Dr. K. Klinger for providing many of the samples used in this analysis. This work was supported in part by 24a grant from the U.S. N.I.H. (HD29256 to DLN).

REFERENCES

Arinam T, Asano M, Kobayashi K, Yanag, H, Hamaguchi H (1993): Data on the CGG repeat at the fragile X site in the non-retarded Japanese population and family suggest the presence of a subgroup of normal alleles predisposing to mutate. Hum Genet 92: 431–436.

Bell GI, Torney DC (1993): Repetitive DNA sequences: Some considerations for simple sequence repeats. Computers Chem 17:185–190.

Bowcock AM, Ruiz-Linares A, Minch E, Kidd JR, Cavalli-Sforza LL (1994): High resolution of human evolutionary trees with polymorphic microsatellites. Nature 368:455–457.

Chakravarti A (1992): Fragile X founder effect? Nat Genet 1:237–238.

Dorit RL, Akashi H, Gilbert W (1995): Absence of polymorphim at the ZFY locus on the human Y chromosome. Science 268:1183–1185.

Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R (1992): Genetic variation at five trimeric and tetrameric tandem repeat loci loci in four human population groups. Genomics 12:241–253.

Eichler EE, Hammond HA, Macpherson JN, Ward PA, Nelson, DL (1995): Population survey of the human FMR1 CGG repeat substructure suggests biased polarity for the loss of AGG interruptions. Hum Molec Genet 4:2199–2208.

Eichler EE, Holden JJA, Popovich BW, Reiss AL, Snow K, Thibodeau SN, Richards CS, Ward PA, Nelson DL (1994): Length of uninterrupted CGG repeats determines stability in the FMR1 gene. Nat Genet 8:88–94.

Fu YH, Kuhl DPA, Pizzuti A, Pieretti M, Sutcliffe JS, Richards S, Verkerk AJMH, Holden JJA, Fenwick RG, Jr., Warren ST, Oostra BA, Nelson DL, Caskey CT (1991): Variation of the CGG repeat at the fragile X site results in genetic instability: Resolution of the Sherman paradox. Cell 67:1047–1058.

Hästbacka J, de la Chapelle A, Kaitila I, Sistonen P, Weaver A, Lander E (1992): Linkage disequilibrium mapping in isolated founder populations: Diastrophic dysplasia in inland. Nature Genet 2:204–211.

Hirst MC, Grewal PK, Davies KE (1994): Precursor arrays for triplet repeat expansion at the fragile X locus. Hum Molec Genet 3: 1553–1560.

Horai S, Hayasaka K, Kondo R, Tsugane K, Takahata N (1995): Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. Proc Natl Acad Sci USA 92:532–536.

Howard-Peebles PN, Stoddard GR (1980): Race distribution in X-linked mental retardation with macro-orchidism and fragile site in Xq. Am J Hum Genet 32:629–630.

Kidd JR, Black FL, Weiss KM, Balazs I, Kidd KK (1991): Studies of three Amerindian populations using nuclear DNA polymorphisms. Hum. Biology 63:775– 794.

Kremer EJ, Yu S, Pritchard M, Nagaraja R, Heitz D, Lynch M, Baker E, Hyland VJ, Little RD, Wada M, Toniolo D, Vincent A, Rousseau F, Schlessinger D, Sutherland GR, Richards RI (1991): Isolation of a human DNA sequence which spans the fragile X. Am J Hum Genet 49:656–661.

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

Kwiatkowski DJ, Henske EP, Weimer K, Ozelius L, Gusella JF, Haines J (1991): Construction of a GT polymorphism map of human 9q. Genomics 12:229–240.

Levinson G, Gutman GA (1986): Slipped-strand mispairing: A major mechanism for DNA sequence evolution. Molec Biol Evol 4:203–221.

Morris A, Morton NE, Collins A, Macpherson J, Nelson DL, Sherman S (1995): An n-allele model for progressive amplifcation in the FMR1 locus. Proc Natl Acad Sci USA 92:4833–4837.

Morton JE, Rindl PM, Bullock S, Bundey S, Webb T (1995): Fragile X syndrome is less common than previously estimated. J Med Genet 32:144–145.

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

Oberle I, Rousseau F, Heitz D, Kretz C, Devys D, Hanauer A, Boue J, Bertheas MF, Mandel JL (1991): Instability of a 550-base pair DNA segment and abnormal ethylation in fragile X syndrome. Science 252:1097–1102.

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 Molec Genet 3:393–398.

Rhoads FA (1984): Fragile-X syndrome in Hawaii: A summary of clinical experience. Am J Med Genet 17:209–214.

Richards RI, Holman K, Friend K, Kremer E, Hillen D, Staples A, Brown WT, Goonewardena P, Tarleton J, Schwartz C, Sutherland GR (1992): Evidence of founder chromosomes in fragile X syndrome. Nat Genet 1:257–260.

Riggins GJ, Lokey LK, Chastain JL, Leiner HA, Sherman SL, Wilkinson KD, Warren ST (1992): Human genes containing polymorphic trinucleotide repeats. Nature Genet 2:186–191.

Rivera H, Hernandez L, Plascenia J, Sanchez-Corona J, Garcia-Cruz D, Cantu JM (1981): Some observations on the mental deficiency of normo-functional testicular hyperplasia and fra(X)(q28) chromosome syndrome. Ann Genet 24:220–222.

Schlötterer C, Tautz D (1992): Slippage synthesis of simple sequence DNA. Nuc Acids Res 20:211–215.

Snow K, Doud LK, Hagerman R, Pergolizzi RG, Erster SH, Thibodeau SN (1993): Analysis of a CGG sequence at the FMR-1 locus in the fragile X families and in the general population. Am J Hum Genet 53:1217–1228.

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

Soysa P, Senanayahe M, Mikkelsen M, Poulsen H (1982): Martin-Bell syndrome fra(X) (q28) in a Sri Lankan family. J Ment Defic Res 26:251–257.

Sutcliffe JS, Zhang F, Caskey CT, Nelson DL, Warren ST (1992): PCR amplification and analysis of yeast artificial chromosomes. Genomics 13:1303–1306.

Turner G, Robinson H, Laing S, Purvis-Smith S (1986): Preventive screening for the fragile X syndrome. N Engl J Med 315:607–699.

Turner G, Robinson H, Laing S, van den Berk M, Colley A, Goddard A, Sherman S, Partington M (1992): Population screening for fragile X. Lancet 339:1210–1213.

Venter PA, Gericke GS, Dawson B, Op't Hof J (1981): A marker X chromosome associated with nonspecific male mental retardation. S Afr Med J 21:807–811.

Webb T, Bundey S (1991): Prevalence of fragile X syndrome (letter). J Med Genet 28:358.

Webb TP, Bundey S, Thake A, Todd J (1986): The frequency of the fragile X chromosome among schoolchildren in Coventry. J Med Genet 23:396–369.

Webb TP, Bundey SE, Thake AI, Todd J (1986): Population incidence and segregation ratios in the Martin Bell syndrome. Am J Hum Genet 23:573–580.

Weber JL (1990): Informativeness of human $(dC\text{-}dA)_n.(cG\text{-}dT)_n$ polymorphisms. Genomics 7:524–530.

Weissenbach J, Gyapay G, Dib C, Vignal A, Morisette J, Millasseau P, Vaysseix G, Lathrop M (1992): A second-generation linkage map of the human genome. Nature 359:794–801.

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

Zar JH (1984). "Biostatistical Analysis," second Edition. New Jersey: Prentice Hall.

Zhong N, Dobkin C, Brown WT (1993): A complex mutable polymrophism located within the fragile X gene. Nat Genet 5:248–253.

Ziglschmid AJF (1947). "Das Kleine Geschichtsbuch der Hutterischen Bruder." Ithaka: Carl-Schurz Foundation.