Am. J. Hum. Genet. 64:354–359, 1999

HUMAN GENETICS '99: TRINUCLEOTIDE REPEATS
Fragile Sites—Cytogenetic Similarity with Molecular Diversity

Grant R. Sutherland and Robert I. Richards

Department of Cytogenetics and Molecular Genetics, Women's and Children's Hospital, Adelaide, Australia

When examined in metaphase chromosome preparations, all fragile sites appear as a gap or discontinuity in chromosome structure. These gaps, which are induced by several specific treatments of cultured cells, are of variable width and promote chromosome breakage with variable efficiency. Within a single metaphase, it is not possible to distinguish between a fragile site and random chromosomal damage. Only the statistically significant recurrence of a lesion at the same locus and under the same culture conditions delineates fragile sites. Several classes of fragile sites have now been characterized at the molecular level. The "rare" fragile sites contain tandemly repeated sequences of varying complexity, which undergo expansions or, occasionally, contractions. The common fragile sites remain as enigmatic at the molecular level as the rare ones once were at the genetic level (Sutherland 1985).

The cytogenetic expression of a fragile site is but one manifestation of genomic instability that is generated by the DNA sequences at fragile-site loci. For instance, the rare folate-sensitive fragile sites are associated with transcriptional silencing of genes, and some of the common fragile sites may also affect gene expression by creating local regions of chromosomal instability. The current challenge is to understand the mechanisms of this instability and its biological significance.

Fragile sites are classified according to their frequency and the conditions or agents that induce their expression (Sutherland et al. 1996). Common fragile sites appear to be part of normal chromosomal structure and are probably present at all common fragile-site loci in all individuals. Rare fragile sites vary in frequency from only a handful of case reports up to 1 in 40 chromosomes. A classification of fragile sites and some details of their cytogenetic and molecular properties are given in table 1.

Fragile sites are identifiable as gaps or chromosomal breaks in only a fraction of metaphase spreads from a given individual. At one extreme is *FRA16B,* which, when induced by berenil, may be found in >90% of metaphases (Schmid et al. 1986). At the opposite end of the spectrum are the common aphidicolin fragile sites. Even the most conspicuous of these, *FRA3B,* is rarely seen in >10% of metaphases (Smeets et al. 1986), and many of the other common fragile sites are seen in <5% of metaphases. Whatever the mechanisms are that result in fragile-site expression, they usually operate successfully in only a minority of cells.

Most of the treatments that induce fragile sites (e.g., perturbed nucleotide pools and aphidicolin) result in slowing DNA replication, particularly at fragile-site loci (Hansen et al. 1997; Le Beau et al. 1998). It is thus possible that the common fragile sites, and perhaps the rare ones, arise because of incompletely replicated DNA sequences that do not package completely for mitosis. Whereas such packaging is completed in most cells, in a varying, often small proportion of cells it may not be completed before the end of G_2, and it is in these cells that fragile sites manifest themselves as gaps or breaks.

Sequences of Fragile Sites

Only three of the different classes of rare fragile sites have been studied at the molecular level. Most is known about the rare folate-sensitive fragile sites, of which five members have been cloned. This group includes *FRAXA,* the site of the molecular lesion in the fragile X syndrome gene *FMR1,* which underlies the most common familial form of mental retardation.

The normal alleles at the *FRAXA* locus vary in size from ~6–55 CCG repeats, mostly with interspersed CCT units after every 9–10 CCG units (Hirst et al. 1994), although there is variation in this pattern, with some alleles having no or only one CCT unit. In Tunisian Jews, there is an increased incidence of normal alleles without CCT units, and this is likely to be the basis of an apparently high incidence of fragile X syndrome in this ethnic group (Falik-Zaccai et al. 1997). The increased length of premutation and mutation alleles appears to

Received November 27, 1998; accepted for publication December 18, 1998; electronically published February 1, 1999.

Address for correspondence and reprints: Dr. Grant R. Sutherland, Department of Cytogenetics and Molecular Genetics, Women's and Children's Hospital, 72 King William Road, Adelaide 5006, Australia. E-mail: gsutherland@medicine.adelaide.edu.au

0002-9297/99/6402-0005$02.00

Table1

Classification and Some Properties of Fragile Sites

Class	Induced by	Number Recognized	Number Cloned	Repeat Motif
Rare:				
1. Folate sensitive	Low folate and thymidine, FUdR, MTX, high thymidine	23	5	CCG
2. Distamycin A inducible	a. Distamycin A, other minor groove-binding oligopeptide antibiotics, and BrdU	2	1	33 bp AT-rich
	b. As above, but not BrdU (recorded only in the Japanese)	3	0	...
3. BrdU inducible	BrdU or BrdC	2	1	~42 bp variable AT-rich
Common:				
4. Aphidicolin inducible	Aphidicolin for at least 24 h	76	3	None obvious
5. 5-Azacytidine inducible	5-Azacytidine for no more than 8 h	4	0	...
6. BrdU inducible	BrdU for no more than 8 h	6	0	...
7. Adenovirus 12 inducible	Adenovirus 12 in U1 small nuclear RNA genes in p53-expressing cells	4	0	...

be due to increased numbers of CCG units alone. Furthermore, the alleles that expand to give pre- and full mutations are in linkage disequilibrium with microsatellite (Richards et al. 1992) and single nucleotide polymorphisms (SNPs; Gunter et al. 1998) within the gene or close to its 5′ end. The presence of one of these SNPs, ATL1, can be used to predict which alleles are likely to proceed to expansion (Gunter et al. 1998), although it is unlikely that the ATL1 polymorphism, in the first intron of *FMR1,* has functional significance in relation to expansion. More likely, it is in linkage disequilibrium with some CCG repeat structure that promotes expansion.

The only other folate-sensitive fragile site for which there is information on allele structure is *FRA16A* (Nancarrow et al. 1995). As with *FRAXA,* some of the longer alleles that lack CCT interruption might be expected to undergo expansion more readily than interrupted alleles. However, too few expanded alleles are available for study to test this prediction rigorously.

The cloning of *FRA16B* showed that repeat sequences other than trinucleotides could undergo expansion and result in a fragile site (Yu et al. 1997). At this locus, a 33-bp AT-rich repeat, which is highly polymorphic on normal chromosomes, expands up to several thousand copies to yield a fragile site. Such large expansions can be viewed only on pulse-field gels, where variation in size by a relatively small number of copies would not be resolved. Hence, although the expansions in different families with *FRA16B* are clearly distinct, neither size variation within families nor somatic instability in carriers of this fragile site has been observed.

Fragile site *FRA10B* differs from *FRA16B* in that, whereas the former is induced only by nucleotide analogues such as BrdU or BrdC, the latter is also (indeed, more strongly) induced by DNA minor groove-binding agents such as distamycin A and berenil. *FRA10B* is the most complex fragile site yet studied. It is an expansion of 5–100 kb, consisting of an AT-rich minisatellite repeat unit of ~42 bp in length, which varies in size and composition between families (Hewett et al. 1998). The expanded sequence shows both somatic and intergenerational instability.

Three common fragile sites (*FRA3B*, *FRA7G,* and *FRA7H*) have also been studied in some detail (Mishmar et al. 1998). In contrast to the rare fragile sites, here the fragility occurs over a region of tens to hundreds of kilobases rather than at a single point. Sequence analysis of these regions reveals no striking features that could account for the fragile site, although there are regions of decreased stability and increased flexibility of the DNA that might be significant (Mishmar et al. 1998; Palin et al. 1998). *FRA3B* occurs within the *FHIT* gene, but the other two fragile sites are not known to be within genes. These common fragile sites are late-replicating regions, and it is of interest that aphidicolin, which induces these fragile sites, further delays their replication (Le Beau et al. 1998). This finding supports the model that these fragile sites may simply represent very late–replicating regions of DNA, which, in a minority of cells, fail to complete replication before mitosis begins and, for this reason, do not package well into chromosomes.

A common aphidicolin type fragile site in the Chinese hamster has been cloned on the basis of its increased sensitivity for the incorporation of a linearized plasmid containing a selectable drug-resistance marker. Sequence analysis showed the region of the fragile site to be AT-rich, with a number of other features, including homologies to yeast autonomous replicating sequences and a consensus sequence for replication origins (Palin et al. 1998). Fragile sites inducible by adenovirus 12 occur in

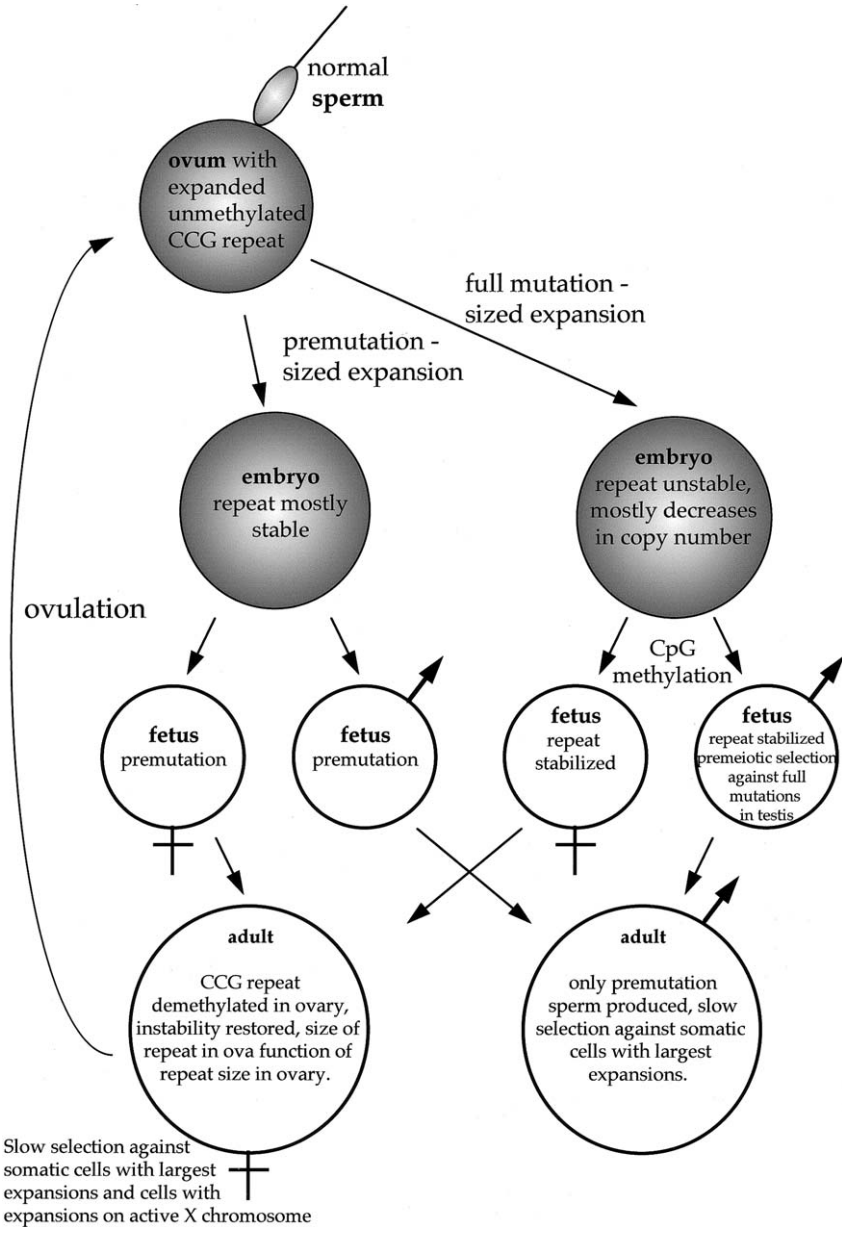

Figure 1 CCG repeat instability and methylation at the *FRAXA* locus during male and female development. FRAXA is the best-studied such site, but some of the observations depicted here probably apply at other fragile sites as well.

the small nuclear RNA gene clusters but only in cells expressing p53. It has been postulated that a viral protein causes p53 to undergo a gain of function, leading either to perturbed transcription by RNA polymerase II or to inefficient chromatin condensation. As Li et al. (1998) note, continued replication or transcription beyond G_2 might be expected to interfere with chromatin compaction.

The Induction of Instability at Fragile Sites

Although fragile sites are themselves visual manifestations of genome instability, there is considerable evidence that genome instability at these loci occurs at other levels. Because most or all of the rare fragile sites are prone to intergenerational and somatic repeat expansion, it is worth considering how this instability comes

about. In the case of *FRAXA,* the best-known example, certain normal alleles can increase in size from within the normal range of copy numbers into the premutation range. Mechanisms to account for instability include unequal crossing over, gene conversion, and replication slippage (see Sinden 1999 [in this issue]). These events are apparently rare, and the rate at which they occur is probably little different from the mutation rate of other simple tandem repeats ($\sim 10^{-3}$ per locus/generation). Once the number of uninterrupted repeats exceeds ~80, the length of an Okazaki fragment, replication slippage is likely to be the predominant mechanism by which increasingly larger size changes occur, the process of dynamic mutation (Richards and Sutherland 1994; Sutherland et al. 1998). This model has recently received experimental support (Freudenreich et al. 1998; Sarkar et al. 1998).

Superimposed on this process are a number of factors such as the age and sex of the individual in which the changes occur. The most likely series of events in the genesis of an individual with a *FRAXA* full mutation is shown in figure 1. An ovum with a full mutation is produced (by either a pre- or a full mutation carrier) and fertilized. This full mutation is unmethylated (Malter et al. 1997), and it remains so throughout early embryogenesis. The unmethylated mutation is somatically unstable and gives rise to cell lineages with different numbers of copies of the repeating unit. Most of these are probably breakdown products with fewer copies of the repeat than are found in the ovum, but some may contain additional copies. Once the expanded repeat undergoes CpG methylation, further changes in repeat copy number are rare. Methylation is probably not complete until 12–14 weeks of gestation. At least in the case of the *FMR1* gene, methylation also leads to transcriptional silencing of the allele on the fragile X chromosome. In the male fetus, there is premeiotic selection against germ cells that contain the full mutation in favor of cells with a premutation (Ashley-Koch et al. 1998). Eventually spermatogenesis results in sperm with premutations only. However, no selection occurs in the ovary of a female fetus, so full mutations persist in the female germline. During oogenesis, when CpG methylation is erased, instability is restored to the CCG repeat, and ova with different-sized mutations are generated. Finally, as individuals age, there is very slow selection against somatic cells with the largest alleles (Ashley-Koch et al. 1998). In females, there is slow selection against cells in which the fragile site is on the active X chromosome (Rousseau et al. 1991). It is likely that this series of events is also applicable to *FRAXE* but not entirely to the autosomal sites. Thus, full mutations in *FRA16A* do not appear to be subject to selection in male germ cells (Nancarrow et al. 1994).

Although intergenerational and somatic instability can occur at *FRA10B* and probably also at *FRA16B,* the rates at which these changes occur are much lower than at folate-sensitive fragile sites, perhaps because replication slippage is less likely with longer-repeat motifs. Other mechanisms, such as gene conversion, may account for changes in allele size at *FRA10B* and *FRA16B* (Jeffreys et al. 1994). Genomic instability at the common fragile sites has been documented extensively for *FRA3B* and to a lesser extent for *FRA7G* (Huang et al. 1998). Instability at *FRA3B* occurs by recombination between long L1 sequences (Inoue et al. 1997) and generates deletions in the *FHIT* gene in a range of solid tumors. *FRA7G* is within an area of instability in prostate cancers, but the mechanism of its instability has not been determined.

Other manifestations of genomic instability at *FRA3B* include its function as a target for viral integration, specifically of HPV16. There is also evidence for the common fragile sites being involved in gene amplification by being points of breakage in a breakage-fusion bridge mechanism (Coquelle et al. 1998).

The Biological Significance of Fragile Sites

The biological significance of chromosomal fragility is not easily generalized among fragile loci. Two of the rare fragile sites on the X chromosome are manifestations of disease (mental retardation)–producing dynamic mutations, whereas the third rare fragile site on this chromosome, *FRAXF,* is not found within any known gene and does not lead to any obvious phenotypic abnormality. The autosomal fragile sites of this type may be associated with transcriptional silencing of genes in which they might be located. Breakage at fragile sites may be capable of producing chromosomal deletion syndromes, as perhaps occurs when *FRA11B* breakage leads to Jacobsen syndrome (Jones et al. 1995). It is of interest that none of the rare folate-sensitive fragile sites has been identified in species other than humans, yet repeats that are potentially capable of expansion to fragile sites have been identified in other species.

Fragile sites are probably all associated with localized genomic instability; although this instability might lead only to the "gap" seen in cytogenetic preparations, it is also possible that the consequences are more significant. There is growing evidence that some common fragile sites predispose their surrounding region to the localized chromosomal instability seen in certain cancers. *FRA3B* is one such region of instability, and abnormal transcripts of the *FHIT* gene in which it is located are found in a range of tumor and normal cell types (Siprashvili et al. 1997; Carapeti et al. 1998; Otterson et al. 1998). The human caveolin genes *CAV1* and *CAV2* are located in the vicinity of *FRA7G,* which is frequently deleted in human cancers (Engelman et al. 1998*b*). Caveolin-1 has

been shown to have a role in the anchorage-dependent inhibition of growth in NIH 3T3 cells (Galbiati et al. 1998). The caveolins are therefore candidates for the tumor-suppressor gene presumed to be located in the *FRA7G* region (Engelman et al. 1998*a*). By similar mechanisms, fragile sites may also lead to gene amplification in tumors. The noncompacted state of the DNA at fragile sites may also facilitate the integration of viruses into the genome.

References

Ashley-Koch AE, Robinson H, Glicksman AE, Nolin SL, Schwartz CE, Brown WT, Turner G, et al (1998) Examination of factors associated with instability of the FMR1 CGG repeat. Am J Hum Genet 63:776–785

Carapeti M, Aguiar RC, Sill H, Goldman JM, Cross NC (1998) Aberrant transcripts of the FHIT gene are expressed in normal and leukaemic haemopoietic cells. Br J Cancer 78: 601–605

Coquelle A, Toledo F, Stern S, Bieth A, Debatisse M (1998) A new role for hypoxia in tumor progression: induction of fragile site triggering genomic rearrangements and formation of complex DMs and HSRs. Mol Cell 2:259–265

Engelman JA, Zhang X, Galbiati F, Volonte D, Sotgia F, Pestell RG, Minetti C, et al (1998*a*) Molecular genetics of the caveolin gene family: implications for human cancers, diabetes, Alzheimer disease, and muscular dystrophy. Am J Hum Genet 63:1578–1587

Engelman JA, Zhang XL, Lisanti MP (1998*b*) Genes encoding human caveolin-1 and -2 are co-localized to the D7S-522 locus (7q31.1), a known fragile site (FRA7G) that is frequently deleted in human cancers. FEBS Lett 436:403–410

Falik-Zaccai TC, Shachak E, Yalon M, Lis Z, Borochowitz Z, Macpherson JN, Nelson DL, et al (1997) Predisposition to the fragile X syndrome in Jews of Tunisian descent is due to the absence of AGG interruptions on a rare Mediterranean haplotype. Am J Hum Genet 60:103–112

Freudenreich CH, Kantrow SM, Zakian VA (1998) Expansion and length-dependent fragility of CTG repeats in yeast. Science 279:853–856

Galbiati F, Volonte D, Engelman JA, Watanabe G, Burk R, Pestell RG, Lisanti MP (1998) Targeted downregulation of caveolin-1 is sufficient to drive cell transformation and hyperactivate the p42/44 MAP kinase cascade. EMBO J 17: 6633–6648

Gunter C, Paradee W, Crawford DC, Meadows KA, Newman J, Kunst CB, Nelson DL, et al (1998) Re-examination of factors associated with expansion of CGG repeats using a single nucleotide polymorphism in *FMR1*. Hum Mol Genet 7:1935–1946

Hansen RS, Canfield TK, Fjeld AD, Mumm S, Laird CD, Gartler SM (1997) A variable domain of delayed replication in *FRAXA* fragile X chromosomes: X inactivation–like spread of late replication. Proc Natl Acad Sci USA 94:4587–4592

Hewett DR, Handt O, Hobson L, Mangelsdorf M, Eyre H, Baker E, Sutherland GR, et al (1998) FRA10B structure reveals common elements in repeat expansion and chromosomal fragile site genesis. Mol Cell 1:773–781

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

Huang H, Qian C, Jenkins RB, Smith DI (1998) Fish mapping of YAC clones at human chromosomal band 7q31.2: identification of YACs spanning FRA7G within the common region of LOH in breast and prostate cancer. Genes Chromosom Cancer 21:152–159

Inoue H, Ishii H, Alder H, Snyder E, Druck T, Huebner K, Croce CM (1997) Sequence of the *FRA3B* common fragile region: implications for the mechanism of *FHIT* deletion. Proc Natl Acad Sci USA 94:14584–14589

Jeffreys A, Tamaki K, MacLeod A, Monckton DG, Neil DL, Armour JAL (1994) Complex gene conversion events in germline mutation at minisatellites. Nat Genet 6:136–145

Jones C, Penny L, Mattina T, Yu S, Baker E, Voullaire L, Langdon WY, et al (1995) Association of a chromosome deletion syndrome with a fragile site within the proto-oncogene CBL2. Nature 376:145–149

Le Beau MM, Rassool FV, Neilly ME, Espinosa R, III, Glover TW, Smith DI, McKeithan TW (1998) Replication of a common fragile site, FRA3B, occurs late in S phase and is delayed further upon induction: implications for the mechanism of fragile site induction. Hum Mol Genet 7:755–761

Li Z, Bailey AD, Buchowski J, Weiner AM (1998) A tandem array of minimal U1 small nuclear RNA genes is sufficient to generate a new adenovirus type 12–inducible chromosome fragile site. J Virol 72:4205–4211

Malter HE, Iber JC, Willemsen R, de Graaff E, Tarleton JC, Leisti J, Warren ST, et al (1997) Characterization of the full fragile X syndrome mutation in fetal gametes. Nat Genet 15:165–169

Mishmar D, Rahat A, Scherer SW, Nyakatura G, Hinzmann B, Kohwi Y, Mandel-Gutfroind Y, et al (1998) Molecular characterization of a common fragile site (FRA7H) on human chromosome 7 by the cloning of a simian virus 40 integration site. Proc Natl Acad Sci USA 14:8141–8146

Nancarrow JK, Holman K, Mangelsdorf M, Hori T, Denton M, Sutherland GR, Richards RI (1995) Molecular basis of p(CCG)n repeat instability at the *FRA16A* fragile site locus. Hum Mol Genet 4:367–372

Nancarrow JK, Kremer E, Holman K, Eyre H, Doggett NA, Le Paslier D, Callen DF, et al (1994) Implications of FRA16A structure for the mechanism of chromosomal fragile site genesis. Science 264:1938–1941

Otterson GA, Xiao GH, Geradts J, Jin F, Chen WD, Niklinska W, Kaye FJ, et al (1998) Protein expression and functional analysis of the FHIT gene in human tumor cells. J Natl Cancer Inst 18:426–432

Palin AH, Critcher R, Fitzgerald DJ, Anderson JN, Farr CJ (1998) Direct cloning and analysis of DNA sequences from a region of the Chinese hamster genome associated with aphidicolin-sensitive fragility. J Cell Sci 111:1623–1634

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

Richards RI, Sutherland GR (1994) Simple repeat DNA is not replicated simply. Nat Genet 6:114–116

Rousseau F, Heitz D, Oberlé I, Mandel J-L (1991) Selection

in blood cells from female carriers of the fragile X syndrome: inverse correlation between age and proportion of active X chromosomes carrying the full mutation. J Med Genet 28: 830–836

Sarkar PS, Chang HC, Boudi FB, Reddy S (1998) CTG repeats show bimodal amplification in E. coli. Cell 95:531–540

Schmid M, Feichtinger W, Jessberger A, Köhler J, Lange R (1986) The fragile site (16)(q22). I. Induction by AT-specific DNA-ligands and population frequency. Hum Genet 74: 67–73

Sinden RR (1999) Biological implications of the DNA structures associated with disease-causing triplet repeats. Am J Hum Genet 64:346–353 (in this issue)

Siprashvili Z, Sozzi G, Barnes LD, McCue P, Robinson AK, Eryomin V, Sard L, et al (1997) Replacement of FHIT in cancer cells suppresses tumorigenicity. Proc Natl Acad Sci USA 94:13771–13776

Smeets DFCM, Scheres JMJC, Hustinx TWJ (1986) The most common fragile site in man is 3p14. Hum Genet 72: 215–220

Sutherland GR (1985) The enigma of the fragile X chromosome. Trends Genet 1:108–112

Sutherland GR, Baker E, Richards RI (1996) Fragile sites. Encyclopedia Molec Biol Molec Med 2:313–318

———(1998) Fragile sites still breaking. Trends Genet 14: 501–506

Yu S, Mangelsdorf M, Hewett D, Hobson L, Baker E, Eyre H, Lapsys N, et al (1997) Human chromosomal fragile site FRA16B is an amplified AT-rich minisatellite repeat. Cell 88:367–374