ORIGIN AND REACTIVATION OF THE FRAGILE X GENE

OORSPRONG EN REACTIVERING VAN HET FRAGIELE X GEN

PROEFSCHRIFT

ter verkrijging van de graad van doctor

aan de Erasmus Universiteit Rotterdam

op gezag van de Rector Magnificus

Prof. dr. ir. J.H. van Bemmel

en volgens besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op

woensdag 24 oktober 2001 om 13.45

door

Pietro Chiurazzi

geboren te Rome (Italië)

Promotiecommissie

Promotoren: Prof. Dr. B.A. Oostra

Prof. Dr. G. Neri

Overige leden: Prof. Dr. J.A. Grootegoed

Prof. Dr. H. Galjaard

Dr. J.N.J. Philipsen

Omslag: Daniela Mariani, Pietro Chiurazzi, Peter J. van der Spek

To all those who smile
and make others smile

CONTENTS

Scope of this thesis

This thesis reports on my last 5 years of research on two different aspects of the fragile X syndrome of mental retardation, namely the haplotypes associated with fragile X full mutations and the mechanisms of transcriptional inactivation of the expanded alleles. The experiments were carried out both in the Department of Clinical Genetics at the Erasmus University (Rotterdam) and in the Institute of Medical Genetics at the Catholic University (Rome) with the supervision of both Prof. Ben A. Oostra and Prof. Giovanni Neri.

Chapter 1 starts with a review article on the genetics of mental retardation, followed by a recent update of the catalogue of the X-linked forms of mental retardation (XLMR). The last section of the Introduction summarizes our current understanding of the fragile X syndrome.

Chapter 2.1 presents an Editorial on the significance of founder effects observed in the fragile X syndrome. The experiments described in Chapter 2.2 led to the identification of fragile X founder chromosomes in Italy, three of which account for over 70% of the total. These results support the notion that only few mutational events generated the unstable "protomutations" which could eventually develop into a "premutation", then becoming "full" fragile X mutations over a relatively long (historical) period of time. Haplotypes of a sub-Saharan African population were also determined, although no fragile X samples were available, to get an idea of the genetic background at the fragile X locus in Africa (Chapter 2.3). This small group of people living in Cameroun displays an extensive genetic diversity when compared to any other European population tested.

Chapter 3 presents the results of the other line of research concentrating on the reactivation of the inactive *FMR1* gene in lymphoblastoid cell lines harbouring a full mutation by treating them with 5-azadeoxycytidine (a compound capable of inducing passive DNA demethylation), histone hyperacetylating drugs, or a combination of both. Recent results on the analysis of the methylation status of the *FMR1* promoter before and after 5-azadC treatments are also presented.

The general discussion in Chapter 4 recapitulates the two main lines of research presented in this thesis and underlines the (many) uncertainties remaining in our present understanding of the molecular pathogenesis of the fragile X syndrome, speculating on future research challenges.

Chapter 1

Introduction

.

1.1

Genetics of mental retardation

Aim of this review is to present the latest advances in the identification of genetic determinants of intellectual deficiency. Mental retardation (MR) is often associated with other neurologic symptoms, metabolic disorders or malformation syndromes. The purpose of the review is to subdivide the large field of MR into categories that may help professionals in making a diagnosis. Nonspecific MR can also segregate in families and the mapping and cloning of corresponding mutant genes will eventually advance our understanding of normal and abnormal brain functioning. Several genes responsible for nonspecific X-linked mental retardation have been identified in the last 12-24 months and are being intensively investigated. This will hopefully lead to possibilities of either genetic or pharmacological therapy.

Definition

Mental retardation (MR) can be defined as a failure to develop cognitive abilities and achieve an intelligence level that would be appropriate for the age group. In most cases this will cause a deficit in the "adaptive behavior" of the subject. Intellectual functioning must be significantly below average, i.e. the IQ should be about 70 or lower in patients who can take an IQ test. Conventionally, MR is subdivided into mild (IQ between 50 and 70), moderate (IQ between 35 and 50), severe (IQ between 20 and 35) and profound (IQ below 20). The "adaptive" functioning of the patient (i.e. his or her social abilities) is sometimes relatively good and may mask the extent of the general intellectual functioning. On the contrary, an autistic behavior can make a formal IQ testing impossible and could lead to an underestimation of the real intelligence of some patients.

Prevalence and etiology

About 2-3% of the total population is reported to be functioning two standard deviations below the mean IQ of the general population i.e. below 70. Mild MR (IQ between 50 and 70) is most frequent (up to 80-85 % of all MR) and is often associated with lower socioeconomic status, whereas the more severe forms occur in all social groups and in families of all educational levels. These estimates can vary a lot between different countries; for example in Finland and Sweden the frequency of MR is reported to be approximately 0.5 % [1,2]. This is because malnutrition is probably the most common cause of mild MR worldwide, in conjunction with sociocultural deprivation and other health problems related to poverty. In fact, malnutrition in a pregnant woman may affect fetal brain development, resulting in mental retardation. Even

polygenic predisposition to a low IQ can be counteracted by growing up in a rich and stimulating environment. A recent report on the effects of environmental neurotoxicants on the developing brain helps to appreciate how much exposure of pregnant women to a polluted environment will impact on the IQ of the baby [3]. More severe forms of MR (IQ <50) are generally associated with other manifestations (either neurologic or not) and their frequency cannot be easily reduced. Generally, the more severe the MR, the more likely it is: 1) that MR will be associated with other manifestations and 2) that a precise etiologic diagnosis will be made. Table 1 briefly summarizes the frequency estimates and the possible causes of MR of various severity.

Table 1 – Frequency estimates, presentation and possible causes of mental retardation

	Relative frequency	Absolute frequency	Presentation	Genetic factors	Environmental factors
Mild (IQ>50)	85%	0.2-2%	often familial	polygenic "predisposition"	(maternal) malnutrition placental insufficiency sociocultural deprivation
Moderate (IQ>35)	10%			monogenic inheritance	maternal PKU pollutants and chemicals alcohol, nicotine, drugs prematurity
Severe (IQ>20)	3%	0.3-0.4%	both familial and sporadic; often associated to other anomalies	(mostly recessive or X linked)	fetal infections (CMV, rubella, toxoplasmosis)
Profound (IQ<20)	1%			chromosomal abnormalities	peri- and postnatal trauma vascular accidents, asphyxia

Genetic causes of MR and pathogenesis

Development of a functional brain depends on precise sequence of proliferative, migratory and maturational events. A careful genetic control is required to induce the formation of the neural tube, which later forms the three vesicles of prosencephalon, mesencephalon and rhombencephalon in its rostral extremity. Then the lateral ventricles will bud from the prosencephalon and will eventually support the formation of cerebral hemispheres with its 6-layered neocortex, the specialised tissue few millimeters-thick where sensations are interpreted, memories checked, emotions integrated and actions planned i.e. the place where our (un)conscious thinking goes on. Any interference with the development or functioning of the neocortex (i.e. positioning, viability and connectivity of cortical neurons) is likely to cause MR and/or possibly epilepsy. Cortical neurons are actually formed and proliferate in the

periventricular region from where they migrate to the surface of the cerebral hemispheres. Even if cortical neurons have reached their place and are normally connected, they can be damaged by toxic compounds accumulating in metabolic disorders (toxic neuro-degeneration). Finally, it is possible that neurons reach their right place and are viable but their cross-talk is not working properly, either because of reduced number of mature synaptic connections or because of inefficient synaptic transmission. This appears to be the case of many nonspecific MR forms like those due to mutant X-linked genes involved in the control of the cytoskeleton and neurite outgrowth or in synaptic vesicle cycling (reviewed by Chelly [4]). Synapses are formed at a very rapid rate during the early months of life achieving maximum density between six and twelve months after birth. They decrease after that due to a sort of natural selection ("pruning"), which appears to spare synapses that are frequently used. This process is probably defective in patients with the fragile X syndrome [5], which is the second most common cause of genetic MR after Down syndrome. Early experiences are thus vital to the retention of synapses and even in the case of MR of unknown causes (possibly with the exception of the neurodegenerative forms), the best improvement is obtained when early stimulation programs are instated. Table 2 presents a schematic list of genetic defects that can cause MR, while Table 3 attempts a classification of the clinical presentation of the various MR conditions, correlated with their pathophysiology.

Table 2 – Genetic causes of mental retardation

Chromosome aneusomies
     trisomy 21 (Down syndrome)
     aneusomies of the X chromosome (e.g. Turner, Klinefelter, tetra/pentasomy X)

Partial chromosomal aberrations
     partial trisomies (e.g. 4p, 9p)
     deletions (e.g. 5p- /cri du chat, 4p- /Wolf-Hirschorn)
     translocations

Subchromosomal abnormalities
     contiguous gene syndromes
     cryptic (subtelomeric) translocations

Monogenic disorders
     autosomal recessive (familial)
     autosomal dominant (mainly sporadic)
     X-linked
     mitochondrial disorders

Polygenic predisposition

Table 3 – Clinical classification and pathogenesis of genetic mental retardation

Multiple congenital anomalies (MCA) syndromes

facial dysmorphisms, multiple malformations	abnormal connections ?
craniosynostoses, microcephaly	cranial constriction, brain hypoplasia

Central nervous system (CNS) malformations

schizencephaly, porencephaly	brain clefts
holoprosencephaly, agenesis of corpus callosum, hydrocephalus	abnormal prosencephalon development
neuronal migration defects, lissencephaly	malpositioning of cortical neurons
neoplasms (tuberous sclerosis)	compression from benign neoplasms

Neuromuscular disorders

Muscular dystrophies	abnormal connections ?
Epilepsy	excessive convulsions ?

Metabolic disorders

Salla disease	
peroxisomal disorders	neurotoxicity and degeneration
ceroid lipofuscinoses	
mitochondrial defects	

Isolated/nonspecific MR

MRX conditions	abnormal connectivity and plasticity

Examples are given for the different classes of genetic MR under the headings in boldface.

Chromosome abnormalities

The most common chromosome abnormality in live newborns is trisomy 21 (Down syndrome) which is invariably associated with MR. The (almost) complete sequence of chromosome 21 is now known [6] but still much effort will be needed to understand what is the effect of the many genes present in 3 instead of 2 copies. Chromosomal aneusomies cause a gene dosage difference for large number of genes and the phenotypic effect is pleiotropic, therefore they always cause syndromes of multiple congenital anomalies (MCA) and MR. However, some genes probably play a major role in determining the phenotype. In the case of the 5p- (cri-du-chat) syndrome, the breakpoints in patients with 5p terminal deletions were characterized with respect to the

severity of MR and the physical location of the delta-catenin gene. A strong correlation was found between the hemizygous loss of delta-catenin and severe MR [7].

Subchromosomal deletions (undetectable by standard cytogenetic analysis) still involve a few genes and cause "contiguous gene" syndromes, as monosomy 1p36 [8], which include MR. Tentler et al. [9] described one such microdeletion on the long arm of chromosome 19 causing MR, skeletal malformations and anemia.

These small deletions are easily seen with fluorescent in situ hybridization (FISH). Petrij et al. [10] describe a reliable FISH method to look for deletions of the *CBP* gene localized in 16p13.3, whose haploinsufficiency cause the Rubinstein-Taybi syndrome.

(Cryptic) Translocations

Translocations can produce less devastating effects than chromosome aneusomies, when they are "balanced" i.e. there is no net loss or gain of genetic material. However, the breakpoint can interrupt an important gene and cause a phenotypic effect.

In some cases only MR will be observed as reported by Villard et al. [11] and Gecz et al. [12], who describe 2 inversions of the X chromosome with a breakpoint in Xq13.1 and one X:autosome balanced translocation interrupting the glutamate receptor 3 (*GRIA3*) gene in Xq24. This latter gene is an attractive candidate for causing nonspecific MR, but no point mutations have been found yet in the coding region.

Complex translocations can involve numerous chromosomes as in the case of a child with mild MR, short stature and microcephaly where 4 chromosomes are rearranged with 9 breakpoints [13].

Combining FISH and DNA analysis of polymorphic markers, a new interesting class of translocations has been recently identified: cryptic subtelomeric translocations. Apparently the gene-rich euchromatic portions close to the end (telomere) of many chromosomes contain homologous sequences that allow translocations of similar length [14-16].

These cryptic translocations are often transmitted in an unbalanced way i.e. with loss or duplication of a small subtelomeric portion, mimicking autosomal recessive inheritance, as in a family with MR and alpha-thalassemia [17] which turned out to have a cryptic translocation between the telomeres of 3q and 16p.

Knight et al. [15] tested 284 and 182 children with unexplained moderate-severe MR and mild MR, respectively, and found subtle chromosomal abnormalities in 7.4% in the children with moderate to severe MR, and in 0.5% in the children with mild retardation. They suggest that,

once recognizable syndromes have been excluded, subtelomeric abnormalities are the commonest cause of undiagnosed moderate-severe MR.

Monogenic disorders

Mutations in single genes can cause a variety of clinical conditions associated with MR or MR alone. While chromosomal unbalances determine complex syndromes, monogenic disorders can present themselves in any of the forms described in Table 3. As far as inheritance pattern is concerned, autosomal dominant (AD) forms are often sporadic if reproductive fitness is reduced, but a family with AD cerebellar ataxia and mild MR was described [18].

Autosomal recessive (AR) transmission is often observed in the case of metabolic disorders where loss of function of both alleles must occur to reduce enzyme levels enough to cause a disease.

X-linked MR (XLMR) families have been extensively studied over the last 30 years, partly because of the greater ease of genetic mapping on the X chromosome but also because of the high frequency of XLMR (reviewed by Neri & Chiurazzi [19]). In fact, hemizygous males have no possibility of compensating deleterious mutations present on their X chromosome, and therefore an excess of males with MR has often been reported (see Herbst & Miller [20]).

An XLMR Database is available on the Internet (http://xlmr.interfree.it/home.htm) and lists over 200 entries, 31 of which correspond to cloned genes mutated in either specific (MRXS) or nonspecific (MRX) mental retardation conditions. Figure 1 depicts the localization of presently cloned XLMR genes.

(Polygenic) predisposition

A challenging aspect of research on the genetic determinants of MR will be to understand the contribution of genes causing mild MR i.e. genes that interfere slightly with cortical functioning. An example of such predisposing genes is represented by the *KCNQ2* and *KCNQ3* genes encoding for voltage-dependent potassium channels which were found mutated in families with benign neonatal and infantile seizures, respectively [21-22]. Missense mutations probably alter partially the channel properties, influencing the threshold for seizures in carrier individuals. Closer collaboration between neuroscientists and molecular geneticists will be necessary to evaluate such subtle differences in intellectual abilities and to correlate individual neuropsychological profiles with genetic variants that are not obviously pathologic.
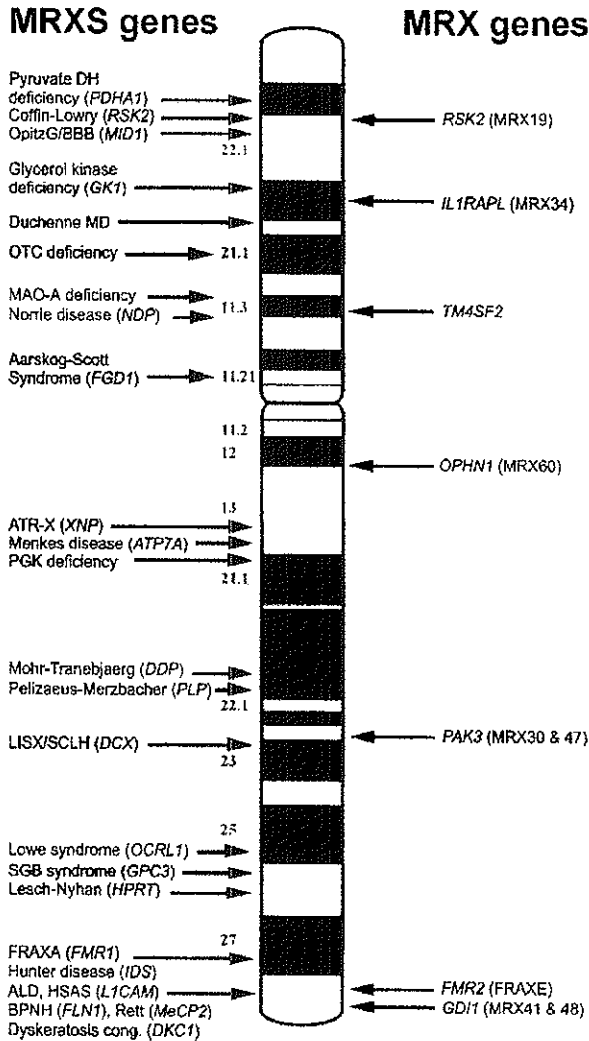
Figure 1 – Map of the X chromosome (G banding) with localization of cloned genes responsible for specific (MRXS, to the left) and nonspecific (MRX, to the right) MR transmitted in an X-linked fashion.

Multiple congenital anomalies (MCA)/MR syndromes

Facial dysmorphism is almost always present in MCA/MR syndromes and suggests a global alteration of the developmental program. When single genes are responsible for a syndrome they will often code for regulatory proteins necessary to control transcription of other genes or signal transduction cascades.

Craniosynostoses are not necessarily associated with MR, but a recent report by Gripp et al. [23] signals MR in the Saethre-Chotzen syndrome with large deletions of the *TWIST* gene. Two genes for primary microcephaly and mild MR have been mapped to chromosome 9q34 and 19q13. In these cases neuroimaging shows a small but structurally normal cerebral cortex [24-25]; it is likely that the head is smaller because the underlying brain grew less. Relatively rare conditions have been observed again: the AD blepharo-naso-facial syndrome [26] and the AR neonatal progeroid (Wiedemann-Rautenstrauch) syndrome [27], both of them associated with moderate MR. A possibly X-linked variant of spondylo-epimetaphyseal dysplasia with severe MR has been reported for the first time [28]. Finally, one locus for the more common Bardet-Biedl syndrome (MR, atypical retinitis pigmentosa, obesity, polydactily) has been further delimitated in 11q13 [29], while the neuropsychological profile of the Cohen syndrome (facial dysmorphisms, MR, myopia, granulocytopenia) has been evaluated [30].

Central nervous system (CNS) malformations

CNS malformations are sometimes associated with facial dysmorphisms, but can also be phenotypically unremarkable. Hydrocephalus due to stenosis of the aqueduct of Sylvius is often due to mutations in the X-linked gene *L1CAM* that encodes a protein of the immunoglobulin superfamily involved in neuron-neuron contact. A recent review on the spectrum and detection rate of *L1CAM* mutations was published [32]. Defects of neuronal migration occur when postmitotic neurons trying to reach the cortex from the inner periventrivular region are arrested at some stage. Lissencephaly ('smooth brain') results from migrational arrest of virtually all cortical neurons short of their normal destination and causes profound MR and seizures. One such gene, *LIS1*, is located on 17p13. In 1998 a second gene (*DCX*) was cloned on the X chromosome [33-34]. Mutations of *DCX* cause an X-linked dominant form of classic lissencephaly with severe MR and epilepsy in males and subcortical laminar heterotopia (SCLH), associated with milder MR and epilepsy in carrier females. In women about half of the neurons reach the cortex while another half (due to random X-inactivation) stop beneath, generating an image of "double cortex" (DC). The *DCX* gene is expressed in migrating neurons

and probably directs migration by regulating the organisation and stability of microtubules. Defects in another X-linked gene, filamin 1, completely prevent migration of cortical neurons causing bilateral periventricular neuronal heterotopia (BPNH) in females [36]. This condition is almost always lethal in males.

Neuromuscular disorders

MR is often associated with other neuromuscular manifestations when neurons other than cortical ones are also defective. Epilepsy is caused by an abnormal firing of cortical neurons and can be genetically determined. When seizures are frequent and uncontrolled, especially during the first months of life, they may result in cognitive deterioration. We mention a study describing a family with AD nocturnal frontal lobe epilepsy and mild MR [37] and another suggesting that also uncontrolled partial seizures may cause MR [38]. Another family with the X-linked West syndrome of infantile spasms and mild MR has been described in Norway [39]. Epilepsy and progressive MR are associated in neurodegenerative metabolic disorders as in the case of ceroid lipofuscinoses. The latest cloned gene is *CLN8* [40], whose mutations cause accumulation of autofluorescent lipopigment.

Ataxia, spastic paraplegia and athetosis associated with MR indicate involvement of cerebellum, upper motor neurons and basal ganglia, respectively. Rett syndrome is a severe X-linked condition, mostly lethal in males, characterised by progressive loss of acquired intellectual abilities, ataxia, spastic paraparesis and seizures. Mutations in the *MeCP2* gene have been found in Rett patients [41]. Muscular dystrophies are not always associated with MR, though one-third of patients with Duchenne muscular dystrophy have mild to moderate MR. Mehler [42] reviewed the role of brain dystrophin, which is expressed in the neuronal soma and dendrites and seems to modulate synaptic integrity, plasticity and signal integration. Finally, mitochondrial disorders always cause neurologic involvement (abnormal tone, seizures, extrapyramidal movements, autonomic dysfunction). Nissenkorn et al. [43] reviewed the charts of 37 children and found that over 50% had MR at the onset and less than 20% had myopathy.

Metabolic disorders

Metabolic disorders are recessively inherited and are caused by deficits of enzymes that lead to accumulation of certain substrates and depletion of their products. Intracellular storage of abnormal levels of unprocessed macromolecules (e.g. glycogen, complex lipids), the release of excessive amounts of small metabolites (e.g. ammonia groups) or the impairment of cellular

energy production produce dramatic effects particularly on neurons, causing lethargy, vomiting, seizures and hypotonia. A new gene, *SLC17A5* was found mutated in patients with a rare Finnish sialic acid storage disorder (Salla disease), characterised by hypotonia, cerebellar ataxia and MR; visceromegaly and coarse features are present in infantile cases and the gene encodes a lysosomal protein [44]. Mutations in the dolichol phosphate mannose synthase (*DPM1*) have been found in subclass Ie of the congenital disorder of glycosylation (CDG-Ie), a metabolic deficiency of glycoprotein biosynthesis with severe MR [45]. Finally, in 1998 mutations in the delta-7-reductase gene formally linked cholesterol biogenesis with the Smith-Lemli-Opitz (SLO) syndrome [46-47]. This is probably the only metabolic disorder which causes a defect of morphogenesis (craniofacial anomalies, polydactyly and urogenital malformations). In fact, the coarse facial features and visceromegaly found e.g. in storage disorders are only secondary deformation events. As discussed in a review [48], the severity depends on the low values of blood cholesterol that have a direct impact on all cellular systems, though malformations may be due to lack of covalent addition of cholesterol to Sonic hedgehog (Shh) signaling protein involved in patterning of the ventral forebrain and limb buds. Encouraging reports on the extremely beneficial results of oral supplementation therapy in SLO patients [49] remind us of the next challenging task: using our knowledge to devise effective therapies.

Isolated/nonspecific MR

If cortical neurons have reached the right place during development and are viable, it does not mean that the synaptic connections will be made in the right number or will mature and function in the right way. In this case it is plausible that "isolated" or nonspecific MR will be observed. Chelly reviewed the role of the first four genes involved in the pathogenesis of nonspecific X-linked MR (MRX), namely *FMR2, OPHN1, GDI1* and *PAK3* [4]. The last 3 of these genes code for proteins involved in the control of the cytoskeleton and neurite outgrowth or in synaptic vesicle cycling. Two more MRX genes have been identified, namely *IL1RAPL* [50] and *TM4SF2* [51]. The first encodes a protein homologous to interleukin-1 receptor accessory proteins (IL1RAP), and its high level of expression in the hippocampal memory system suggests a role in the physiologic processes underlying memory and learning abilities. The second gene, *TM4SF2*, is highly expressed in the cerebral cortex and hippocampus. It encodes a member of the tetraspanin family, found in molecular complexes that include beta-1 integrins, and may have a role in the control of neurite outgrowth. It is interesting to note that in both cases the chromosomal localization was not obtained by linkage analysis of a large pedigree with MR, but

thanks to the observation of patients with a microdeletion and an X;autosome translocation, respectively. This underlines the importance of investigating even sporadic patients with cytogenetic anomalies that provide valuable positional information.

Conclusion

Many genes involved in mental retardation have been mapped and cloned in recent years. The almost complete sequencing of the human genome will be of great help in identifying and cloning of new genes involved in mental retardation. The next challenge will be the unraveling of the actual molecular pathways leading to mental retardation.

References

1. Rantakallio P, von Wendt L: Mental retardation and subnormality in a birth cohort of 12,000 children in Northern Finland. Am J Ment Defic 1986, 90:380-387.

2. Hagberg B, Hagberg G, Lewerth A, Lindberg U: Mild mental retardation in Swedish school children. I. Prevalence. Acta Paediatr Scand 1981, 70:441-444.

3. Weiss B, Landrigan PJ: The Developing Brain and the Environment: An Introduction.
Environ Health Perspect 2000, 108 Suppl 3:373-374.

4. Chelly J: Breakthroughs in molecular and cellular mechanisms underlying X-linked mental retardation. Hum Mol Genet 1999, 8:1833-1838.

5. Weiler IJ, Greenough WT: Synaptic synthesis of the Fragile X protein: possible involvement in synapse maturation and elimination. Am J Med Genet 1999, 83:248-252.

6. Hattori M, Fujiyama A, Taylor TD, Watanabe H, Yada T, Park H-S, Toyoda A, Ishii K, Totoki Y, Choi D-K, Soeda E, Ohki M, Takagi T, Sakaki Y, Taudien S, Blechschmidt K, Polley A, Menzel U, Delabar J, Kumpf K, Lehmann R, Patterson D, Reichwald K, Rump A, Schillhabel M, Schudy A, Zimmermann W, Rosenthal A, Kudoh J, Shibuya K, Kawasaki K, Asakawa S, Shintani A, Sasaki T, Nagamine K, Mitsuyama S, Antonarakis SE, Minoshima S, Shimizu N, Nordsick G, Hornischer K, Brandt P, Scharfe M, Schön O, Desario A, Reichelt J, Kauer G, Blöcker H, Ramser J, Beck A, Klages S, Hennig S, Riesselmann L, Dagand E, Haaf T, Wehrmeyer S, Borzym K, Gardiner K, Nizetic D, Francis F, Lehrach H, Reinhardt R & Yaspo M-L: The DNA sequence of human chromosome 21. Nature 2000, 405:311-319.

7. Medina M, Marinescu RC, Overhauser J, Kosik KS: Hemizygosity of delta-catenin (CTNND2) is associated with severe mental retardation in cri-du-chat syndrome. Genomics 2000, 63:157-164.

8. Slavotinek A, Shaffer LG, Shapira SK: Monosomy 1p36. J Med Genet 1999, 36:657-663.

9. Tentler D, Gustavsson P, Elinder G, Eklof O, Gordon L, Mandel A, Dahl N: A microdeletion in 19q13.2 associated with mental retardation, skeletal malformations, and Diamond-Blackfan anaemia suggests a novel contiguous gene syndrome. J Med Genet 2000, 37:128-131.

10. Petrij F, Dauwerse HG, Blough RI, Giles RH, van der Smagt JJ, Wallerstein R, Maaswinkel-Mooy PD, van Karnebeek CD, van Ommen GJB, van Haeringen A, Rubinstein JH, Saal HM, Hennekam RCM, Peters DJM, Breuning MH: Diagnostic analysis of the Rubinstein-Taybi syndrome: five cosmids should be used for microdeletion detection and low number of protein truncating mutations. J Med Genet 2000, 37:168-176.

11. Villard L, Briault S, Lossi AM, Paringaux C, Belougne J, Colleaux L, Pincus DR, Woollatt E, Lespinasse J, Munnich A, Moraine C, Fontes M, Geoz J: Two unrelated patients with inversions of the X chromosome and non-specific mental retardation: physical and transcriptional mapping of their common breakpoint region in Xq13.1. J Med Genet 1999, 36:754-758.

12. Geoz J, Barnett S, Liu J, Hollway G, Donnelly A, Eyre H, Eshkevari HS, Baltazar R, Grunn A, Nagaraja R, Gilliam C, Peltonen L, Sutherland GR, Baron M, Mulley JC: Characterization of the human glutamate receptor subunit 3 gene (GRIA3), a candidate for bipolar disorder and nonspecific X-linked mental retardation. Genomics 1999, 62:356-368.

13. Joyce CA, Cabral de Almeida JC, Santa Rose AA, Correia P, Moraes L, Bastos E, Llerena J Jr: A de novo complex chromosomal rearrangement with nine breakpoints characterized by FISH in a boy with mild mental retardation, developmental delay, short stature and microcephaly. Clin Genet 1999, 56:86-92.

14. Knight SJ, Flint J: Perfect endings: a review of subtelomeric probes and their use in clinical diagnosis. J Med Genet 2000, 37:401-409.

15. Knight SJ, Regan R, Nicod A, Horsley SW, Kearney L, Homfray T, Winter RM, Bolton P, Flint J: Subtle chromosomal rearrangements in children with unexplained mental retardation. Lancet 1999, 354:1676-1681.

16. Slavotinek A, Rosenberg M, Knight S, Gaunt L, Fergusson W, Killoran C, Clayton-Smith J, Kingston H, Campbell RH, Flint J, Donnai D, Biesecker L: Screening for submicroscopic chromosome rearrangements in children with idiopathic mental retardation using microsatellite markers for the chromosome telomeres. J Med Genet 1999, 36:405-411.

17. Holinski-Feder E, Reyniers E, Uhrig S, Golla A, Wauters J, Kroisel P, Bossuyt P, Rost I, Jedele K, Zierler H, Schwab S, Wildenauer D, Speicher MR, Willems PJ, Meitinger T, Kooy RF: Familial mental retardation syndrome ATR-16 due to an inherited cryptic subtelomeric translocation, t(3;16)(q29;p13.3). Am J Hum Genet 2000, 66:16-25.

18. Herman-Bert A, Stevanin G, Netter JC, Rascol O, Brassat D, Calvas P, Camuzat A, Yuan Qp, Schalling M, Durr A, Brice A: Mapping of Spinocerebellar Ataxia 13 to Chromosome 19q13.3-q13.4 in a Family with Autosomal Dominant Cerebellar Ataxia and Mental Retardation. Am J Hum Genet 2000, 67:229-235.

19. Neri G, Chiurazzi P: X-linked mental retardation. Adv Genet 1999, 41:55-94.

20. Herbst DS, Miller JR: Nonspecific X-linked mental retardation II: the frequency in British Columbia. Am J Med Genet 1980, 7:461-469.

21. Singh NA, Charlier C, Stauffer D, DuPont BR, Leach RJ, Melis R, Ronen GM, Bjerre I, Quattlebaum T, Murphy JV, McHarg ML, Gagnon D, Rosales TO, Peiffer A, Anderson VE, Leppert M: A novel potassium channel gene, KCNQ2, is mutated in an inherited epilepsy of newborns. Nature Genet 1998, 18:25-29.

22. Charlier C, Singh NA, Ryan SG, Lewis TB, Reus BE, Leach RJ, Leppert M: A pore mutation in a novel KQT-like potassium channel gene in an idiopathic epilepsy family. Nat Genet 1998, 18:53-55.

23. Gripp KW, Zackai EH, Stolle CA: Mutations in the human TWIST gene. Hum Mutat 2000, 15:150-155 (erratum in 15:479).

24. Moynihan L, Jackson AP, Roberts E, Karbani G, Lewis I, Corry P, Turner G, Mueller RF, Lench NJ, Woods CG: A third novel locus for primary autosomal recessive microcephaly maps to chromosome 9q34. Am J Hum Genet 2000, 66:724-727.

25. Roberts E, Jackson AP, Carradice AC, Deeble VJ, Mannan J, Rashid Y, Jafri H, McHale DP, Markham AF, Lench NJ, Woods CG: The second locus for autosomal recessive primary microcephaly (MCPH2) maps to chromosome 19q13.1-13.2. Eur J Hum Genet 1999, 7:815-820.

26. Stoll C, Terzic J, Fischbach M: A three generations family with blepharo-naso-facial malformations suggestive of Pashayan syndrome. Genet Couns 1999, 10:337-343.

27. Pivnick EK, Angle B, Kaufman RA, Hall BD, Pitukcheewanont P, Hersh JH, Fowlkes JL, Sanders LP, O'Brien JM, Carroll GS, Gunther WM, Morrow HG, Burghen GA, Ward JC: Neonatal progeroid (Wiedemann-Rautenstrauch) syndrome: report of five new cases and review. Am J Med Genet 2000, 90:131-140.

28. Bieganski T, Dawydzik B, Kozlowski K: Spondylo-epimetaphyseal dysplasia: a new X-linked variant with mental retardation. Eur J Pediatr 1999, 158:809-814.

29. Katsanis N, Lewis RA, Stockton DW, Mai PM, Baird L, Beales PL, Leppert M, Lupski JR: Delineation of the critical interval of Bardet-Biedl syndrome 1 (BBS1) to a small region of 11q13, through linkage and haplotype analysis of 91 pedigrees. Am J Hum Genet 1999, 65:1672-1679.

30. Kivitie-Kallio S, Larsen A, Kajasto K, Norio R: Neurological and psychological findings in patients with Cohen syndrome: a study of 18 patients aged 11 months to 57 years. Neuropediatrics 1999, 30:181-189.

31. Brunelli S, Faiella A, Capra V, Nigro V, Simeone A, Cama A, Boncinelli E: Germline mutations in the homeobox gene EMX2 in patients with severe schizencephaly. Nat Genet 1996, 12:94-96.

32. Finckh U, Schroder J, Ressler B, Veske A, Gal A: Spectrum and detection rate of L1CAM mutations in isolated and familial cases with clinically suspected L1-disease. Am J Med Genet 2000, 92:40-46.

33. des Portes V, Pinard JM, Billuart P, Vinet MC, Koulakoff A, Carrie A, Gelot A, Dupuis E, Motte J, Berwald-Netter Y, Catala M, Kahn A, Beldjord C, Chelly J: A novel CNS gene required for neuronal migration and involved in X-linked subcortical laminar heterotopia and lissencephaly syndrome. Cell 1998, 92:51-61.

34. Gleeson JG, Allen KM, Fox J W, Lamperti ED, Berkovic S, Scheffer I, Cooper EC, Dobyns WB, Minnerath SR, Ross ME, Walsh CA : Doublecortin, a brain-specific gene mutated in human X-linked lissencephaly and double cortex syndrome, encodes a putative signaling protein. Cell 1998, 92: 63-72.

35. Gleeson JG, Lin PT, Flanagan LA, Walsh CA: Doublecortin is a microtubule-associated protein and is expressed widely by migrating neurons. Neuron 1999, 23:257-271.

36. Fox JW, Lamperti ED, Eksioglu YZ, Hong SE, Feng Y, Graham DA, Scheffer IE, Dobyns WB, Hirsch BA, Radtke RA, Berkovic SF, Huttenlocher PR, Walsh CA: Mutations in filamin 1 prevent migration of cerebral cortical neurons in human periventricular heterotopia. Neuron 1998, 21: 1315-1325.

37. Ito M, Kobayashi K, Fujii T, Okuno T, Hirose S, Iwata H, Mitsudome A, Kaneko S: Electroclinical picture of autosomal dominant nocturnal frontal lobe epilepsy in a Japanese family. Epilepsia 2000, 41:52-58.

38. Kramer U, Fattal A, Nevo Y, Leitner Y, Harel S: Mental retardation subsequent to refractory partial seizures in infancy. Brain Dev 2000, 22:31-34.

39. Stromme P, Sundet K, Mork C, Cassiman JJ, Fryns JP, Claes S: X linked mental retardation and infantile spasms in a family: new clinical data and linkage to Xp11.4-Xp22.11. J Med Genet 1999, 36:374-378.

40. Ranta S, Zhang Y, Ross B, Lonka L, Takkunen E, Messer A, Sharp J, Wheeler R, Kusumi K, Mole S, Liu W, Soares MB, Bonaldo MF, Hirvasniemi A, de la Chapelle A, Gilliam TC, Lehesjoki AE: The neuronal ceroid lipofuscinoses in human EPMR and mnd mutant mice are associated with mutations in CLN8. Nat Genet 1999, 23:233-236.

41. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY: Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nat Genet 1999, 23:185-188.

42. Mehler MF: Brain dystrophin, neurogenetics and mental retardation. Brain Res Brain Res Rev 2000, 32:277-307.

43. Nissenkorn A, Zeharia A, Lev D, Watemberg N, Fattal-Valevski A, Barash V, Gutman A, Harel S, Lerman-Sagie T: Neurologic presentations of mitochondrial disorders. J Child Neurol 2000, 15:44-48.

44. Verheijen FW, Verbeek E, Aula N, Beerens CE, Havelaar AC, Joosse M, Peltonen L, Aula P, Galjaard H, van der Spek PJ, Mancini GM: A new gene, encoding an anion transporter, is mutated in sialic acid storage diseases. Nat Genet 1999, 23:462-465.

45. Kim S, Westphal V, Srikrishna G, Mehta DP, Peterson S, Filiano J, Karnes PS, Patterson MC, Freeze HH: Dolichol phosphate mannose synthase (DPM1) mutations define congenital disorder of glycosylation Ie (CDG-Ie). J Clin Invest 2000, 105:191-198.

46. Wassif CA, Maslen C, Kachilele-Linjewile S, Lin D, Linck LM, Connor WE, Steiner R, Porter FD: Mutations in the human sterol delta-7-reductase gene at 11q12-13 cause Smith-Lemli-Opitz syndrome. Am J Hum Genet 1998, 63:55-62.

47. Yu H, Lee MH, Starck L, Elias ER, Irons M, Salen G, Patel SB, Tint GS: Spectrum of Delta(7)-dehydrocholesterol reductase mutations in patients with the Smith-Lemli-Opitz (RSH) syndrome. Hum Mol Genet 2000, 9:1385-1391.

48. Kelley RI: RSH/Smith-Lemli-Opitz syndrome: mutations and metabolic morphogenesis. Am J Hum Genet 1998, 63:322-326.

49. Starck L, Bjorkhem I, Ritzen EM, Nilsson BY, von Dobeln U: Beneficial effects of dietary supplementation in a disorder with defective synthesis of cholesterol. A case report of a girl with Smith-Lemli-Opitz syndrome, polyneuropathy and precocious puberty. Acta Paediatr 1999, 88:729-733.

50. Carrié A, Jun L, Bienvenu T, Vinet M-C, McDonell N, Couvert P, Zemni R, Cardona A, Van Buggenhout G, Frints S, Hamel B, Moraine C, Ropers HH, Strom T, Howell GR, Whittaker A, Ross MT, Kahn A, Fryns JP, Beldjord C, Marynen P & Chelly J: A new member of the IL-1receptor family highly expressed in hippocampus and involved in X-linked mental retardation. Nature Genet 1999, 23: 25-31.

51. Zemni R, Bienvenu T, Vinet MC, Sefiani A, Carrié A, Billuart P, McDonell N, Couvert P, Francis F, Chafey P, Fauchereau F, Friocourt G, des Portes V, Cardona A, Frints S, Meindl A, Brandau O, Ronce N, Moraine C, van Bokhoven H, Ropers HH, Sudbrak R, Kahn A, Fryns JP, Beldjord C & Chelly J: A new gene involved in X-linked mental retardation identified by analysis of an X;2 balanced translocation. Nature Genet 2000, 24: 167-170.

1.2

X-linked mental retardation

Adapted from

Chiurazzi P, Hamel BCJ, Neri G (2001)
XLMR Genes: Update 2000. Eur J Hum Genet 9:71-81

This is the sixth edition of the catalogue of X-linked mental retardation (XLMR) genes, i.e. X-linked genes whose malfunctioning causes mental retardation. The cloning era is not yet concluded, actually much remains to be done to account for the 202 XLMR conditions listed in this update. Many of these may eventually prove to be due to mutations in the same gene but the present number of 33 cloned genes falls surely short of the actual total count. It is now clear that even small families or individual patients with cytogenetic rearrangements can be instrumental in pinning down the remaining genes. DNA chip technology will hopefully allow (re)screening large numbers of patients for mutations in candidate genes or testing the expression levels of many candidate genes in informative families. Slowly, our knowledge of the structure and functioning of the proteins encoded by these genes is beginning to cast some light on the biological pathways required for the normal development of intelligence. Correlations between the molecular defects and the phenotypic manifestations are also being established. In order to facilitate the exchange of existing information and to allow its timely update, we prepared the first edition of the XLMR database (available at http://xlmr.interfree.it/home.htm) and invite all colleagues, expert in the field, to contribute with their experience.

Introduction

Mental retardation (MR) can be defined as a failure to develop cognitive abilities and achieve a level of intelligence that would be appropriate for the age group. In most cases MR causes a deficit in the "adaptive behavior". A percentage of the general population, variably estimated between 0.5 % and 2 %, is reported to be functioning two standard deviations below the average (i.e. to have an IQ of less than 70). Genetic determinants underlie many of these conditions[1] and an excess of affected males has often been reported[2-3], especially in the mild-to-moderate MR range (IQ between 70 and 35). This phenomenon was eventually interpreted as due to X-linked mutations, whose effects become more apparent in the hemizygous males who cannot compensate for deleterious mutations present on their X chromosome[2,4]. Herbst and Miller estimated the frequency of X-linked mental retardation (XLMR) at 1.8/1000 males in British Columbia[5]. Recently the prevalence of the fragile X syndrome, probably the commonest of XLMR conditions, has been reestimated at 1/4000 males at most[6] and may represent up to 15-20 % of the total XLMR. Neri et al.[7] prepared the first XLMR update in 1990 and tried to include all X-linked forms of mental retardation in their listing, which steadily grew to the total number of 179 entries in the previous update[8]. This XLMR update 2000 now lists 202 conditions subdivided in different nosological

classes that will be illustrated in the next section.

Information sources and classification of XLMR conditions

The information sources employed to compile this listing have been many and diverse. First of all, we have regularly scanned the published literature i.e. articles in scientific journals and abstracts presented at the American Society of Human Genetics, the European Society of Human Genetics and the biannual International Workshop on Fragile X and XLMR. The On-line Mendelian Inheritance in Man (OMIM) catalogue was often checked and provided valuable information and references (available at http://www.ncbi.nlm.nih.gov/omim/). For approximately 560 entries listed in the OMIM catalogue, MR is mentioned in the clinical synopsis and 90 (16%) of these are X-linked (as per June 30, 2000). Lastly, some inclusions and exclusions were suggested to us by many colleagues, whose advice was extremely helpful. It is inevitable that some items may have been missed or erroneously included in this listing. Any suggestions or corrections will be gratefully accepted and it is now possible to do so in real time by contributing to the on-line XLMR database. We excluded small pedigrees with uncertain X-linkage (e.g. a mother with two affected sons), but such "uncertain" families may still be extremely valuable for genetic studies and could be included in a dedicated section of the on-line XLMR database. Two affected half-brothers or an affected nephew and uncle have been considered sufficient proof of X linkage.

In the nosology of XLMR, a major distinction has been made between "nonspecific" and "specific" or "syndromal" conditions. According to the HUGO Nomenclature Committee (http://www.gene.ucl.ac.uk/nomenclature/), the former are indicated by the acronym MRX and the latter by the acronym MRXS, where "S" stands for "syndromal". The terms "syndromal" and "specific" are currently used interchangeably to indicate conditions that are clinically recognizable because of a specific pattern of physical, neurological, or metabolic abnormalities[4]. A dedicated survey of MRXS conditions, compiled by some of the Authors of this Update 2000, is presently in press[9]. Sometimes MRX numbers were assigned to pedigrees that were eventually described as "syndromal" and this has created some confusion. The MRX designation is actually intended for families whose only consistent clinical manifestation is X-linked MR. MRXS conditions have been somewhat arbitrarily subdivided into four classes: malformation syndromes (Table 1), neuromuscular disorders (Table 2), metabolic (Table 3) and dominant (Table 4) conditions. By "malformation syndrome" we intend a condition characterized by MR and multiple congenital anomalies. A "neuromuscular disorder" is one with a major involvement of the nervous system and/or muscles. "Metabolic" conditions are

considered separately because their pathophysiology is known and due to the abnormal functioning of specific enzymes. "Dominant" conditions have been set apart because of their peculiar inheritance, with near absence of affected males (males die before birth, with the notable exception of one form with epilepsy and MR restricted to females) and presence of affected females. The distinction between specific and nonspecific conditions should intuitively have a molecular correlate. Actually, one can speculate that MRXS genes code for proteins with a broad range of molecular targets (e.g. transcriptional regulators such as XNP/ATR-X or protein kinases such as RSK2), while most of the MRX genes now cloned seem to produce proteins with more limited and specific tasks[10,11] (e.g. regulating the shuttling of synaptic vesicles or modulating the establishment of synaptic contacts between neurons). This hypothesis will need further experimental support, also to explain such discrepancies as observed for the *RSK2* gene mutated in the Coffin-Lowry syndrome and in the nonspecific MRX19 family[12], or for the *MeCP2* gene responsible for the Rett syndrome but also mutated in a small MRX family[13].

Listing and Maps

In tables 1-4, which list the specific conditions, the name, the OMIM number or reference and a brief description are provided for every entry, while the gene localization and the name of the gene are obviously indicated only where available. Table 1 contains several additions of syndromes described in the last two years.

Table 1 – Malformation syndromes

MIM No.	Name	Locus	Gene	Description
*305400	Aarskog-Scott	Xp11.21	*FGDY*	Hypertelorism, short stature, downslanting palpebral fissures, anteverted nostrils, shawl scrotum, joint hyperlaxity
304200	Akesson			Cutis verticis gyrata, thyroid aplasia
#301040	ATR-X	Xq13	*XH2/XNP* *300032	Microcephaly, "coarse" face, genital and skeletal anomalies, alpha-thalassemia. It includes Juberg-Marsidi (#309590), Carpenter-Waziri [15], Holmes-Gang [16], a family with spastic diplegia [17] and possibly [18] Smith-Fineman-Myers (309580)
*301900	Borjeson-Forssman-Lehmann	Xq26-q27		Obesity, hypogonadism, round face, narrow palpebral fissures, epilepsy
301950	Branchial arch			Short stature, downslanting eyes, lowset ears, webbed neck, highly arched palate
308830	Cantu'			Macrocephaly, dwarfism, keratosis follicularis
*309620	Christian	Xq27-q28		Skeletal dysplasia, VI nerve palsy
*300243	Christianson	Xq23-q27.3		Profound MR, mutism despite normal hearing, craniofacial dysmorphism, grand-mal epilepsy, ophthalmoplegia, cerebellar atrophy
309490	Chudley-Lowry			Short stature, obesity, small genitalia
#303600	Coffin-Lowry	Xp22.2-p22.1	*RSK2* *300075	"Coarse" face, drumstick phalanges, skeletal anomalies. It includes MRX19 [12]

#305000	Dyskeratosis congenita	Xq28	*DKC1* *300126	Skin pigmentation, nail dystrophy, leukoplakia of oral mucosa
*305450	FG	Xq12-q21.3		Macrocephaly, agenesis of corpus callosum, gastrointestinal anomalies, deafness
*309550	Fragile X	Xq27.3	*FMR1*	Macrocephaly, long face and ears, macroorchidism
301590	Graham			Anophthalmos, ankyloblepharon, orbital underdevelopment
*302000	Hereditary bullous dysfunction	Xq27.3-q28		Short stature, microcephaly, alopecia, bullous dystrophy, hypogenitalism
300064	Hyde-Forster			Craniofacial anomalies with plagiocephaly, flattened occiput
307010	Hydrocephalus with cerebellar agenesis			Hydrocephalus, cerebellar agenesis, absence of Magendie and Luschka's foramina
*309800	Lenz	Xq27-q28		Microphthalmia, thumb and skeletal anomalies, urogenital and cardiovascular anomalies. It includes Siber et al. [21]
*309520	Lujan-Fryns			Marfanoid habitus, triangular face, narrow palate, hypernasal voice
*309605	Miles /MRXS4	Xp11.2-q23		Microcephalus, asymmetric face, hypogonadism, joint hypermobility, 10 digital arches
*302350	Nance-Horan	Xp22.31 -p22.13		Cataract, microcornea, cone-shaped incisors, supernumerary teeth
*300000	Opitz G/BBB	Xp22	*MID1*	Hypertelorism, midline abnormalities, heart defects, hypospadias
*311300	Oto-palato-digital	Xq27-q28		Short stature, hearing loss, cleft palate, characteristic face
*309510	Partington /MRXS1	Xp22.2-p22.1		Dysarthria, dystonic hand movements, ataxia, seizures
*304340	Pettigrew /MRXS5	Xq26-q27.1		Long "coarse" face, hydrocephalus, hypotonia, spasticity, ataxia, seizures, iron accumulation in basal ganglia, Dandy-Walker anomaly
300055	PPM-X	Xq28		Psychosis, pyramidal signs, macroorchidism
*309610	Prieto /MRXS2	Xp21.1-p11.3		Peculiar face, dental anomalies, sacral dimple, joint dysplasia, epilepsy
300004	Proud			Microcephaly, agenesis of corpus callosum, arthrogryposis, renal dysplasia, hypospadias
*309500	Renpenning /MRXS8	Xp11.4-p11.2		Microcephaly, short stature
308200	Rud	Xp22		Ichthyosis, epilepsy, nystagmus, hypogonadism
314320	Say-Meyer			Trigonocephaly, short stature
312840	Schimke			Early onset choreoathetosis with later spasticity, microcephaly, growth failure, external ophthalmoplegia, variable deafness
#312870	Simpson-Golabi-Behmel	Xq26	*GPC3* *300037	Macrosomia, "coarse" face, polydactyly, extra nipples, heart defects
309580	Smith-Fineman-Myers			Peculiar face, microcephaly, short stature, seizures
*309583	Snyder-Robinson	Xp22.2-p21.2		Macrocephaly, long thin face, high narrow/cleft palate, asthenic body build, scoliosis
*309470	Sutherland /MRXS3	Xp11.3-q12		Microcephaly, short stature, small testes, spastic diplegia
309480	Tranebjaerg I			Epilepsy, psoriasis
*314390	VACTERL with hydrocephalus			Vertebral, anal, tracheo-esophageal, renal and radial defects, hydrocephalus
*314500	Van den Bosch			Choroideremia, acrokeratosis verruciformis, anhydrosis, skeletal deformities
311450	W syndrome			Characteristic face, clefting, subluxed elbow, camptodactily
308400	Warkany			Intrauterine growth retardation, microcephaly
*309545	Wilson	(Xp11-Xq27)		Aphasia, growth failure, brachycephaly, large mouth with thick lips, seizures, frequent infections
*309585	Wilson /MRXS6	Xp21.1-q22		Obesity, gynecomastia, tapering fingers, emotional lability
	Abidi [22]	Xq12-q21		Short stature, small head, sloping forehead, hearing loss
	Ahmad /MRXS7 [23]	Xp11.3-q22		Obesity, hypogonadism, tapered fingers
	Aldred [24]			Retinitis pigmentosa, microcephaly
	Armfield [25]	Xq28		Macrocephaly, glaucoma, cleft palate, seizures, short stature, small hands and feet
	Atkin-Flaitz [26]			Macrocephaly, "coarse" face, short stature, macroorchidism
	Baraitser [27]			Macrocephaly, large ears, broad nasal tip, thick lower lip, teeth anomalies, obesity, macroorchidism

Brzustowicz [14]	Xp22	early lethal, multiple congenital anomalies, hydrops fetalis, Simpson-Golabi-Behmel like
Carpenter [28]	Xq23-q24	Congenital hip dislocation, microcephaly, hypertelorism and dysmorphic facial features, short neck and sternum
Chudley [29]	Xq21.2-q23	Prognathism, synophrys, hirsutism, seizures, abnormal gait and weakness
Golabi-Ito-Hall [30]		Short stature, triangular face, epicanthic folds, microcephaly, brittle hair
Hall [31]		Cleft lip/palate, facial dysmorphism, inguinal hernia, digital defects
Hamel [32]		Congenital heart defect, cleft palate, short stature, facial anomalies
Hockey [33]		Precocious puberty, progressive IQ deterioration (mild to moderate)
Homfray [34]		Coarse facial features, epilepsy, progressive joint contractures
Johnson [35]	Xq12-q21	Macrocephaly, macroorchidism, midface hypoplasia, triangular face
Kang [36]		Microcephaly, dysgenesis of corpus call., hydrocephalus, spasticity, short broad hands, facial anomalies
Lubs [37]	Xq28	Hypertelorism, short nose, seizures, hearing loss, early demise, cardiomegaly at autopsy
MEHMO [38]	Xp22.13-p21.1	Epilepsy, hypogonadism and hypogenitalism, microcephaly, obesity
Oosterwijk [39]	(Xp11.4-q24) excluded	A/symphalangism of hands and feet, hearing loss, verrucosis and hypertrichosis, immunodeficiency
Porteous [40]	Xp11.4-q13	Short stature, high-pitched voice, high forehead, receding hairline
Reish [41]		Multiple congenital anomalies, growth retardation, ectodermal dysplasia
Seemanova [42]		Microcephaly, microphthalmia, growth retardation
Shashi [43]	Xq26-q27	Coarse facial features, puffy eyelids, obesity, large ears and testes
Shrimpton/MRXS9 [44]	Xq12-q21.31	Microcephaly, variably short stature
Siderius-Hamel [45]	Xp11.3-q21.3	Cleft lip and palate, broad nasal tip, large hands
Sklower-Brooks [46]		Peculiar face, growth retardation, optic atrophy, spastic diplegia, atrophic hydrocephalus
Stevenson [47]	Xq12-q21.2	Hypotonia, areflexia, tapered fingers, arches increased, genu valgum
Stocco dos Santos [48]		Short stature, hip luxation, precocious puberty
Stoll [49]		Short stature, prominent forehead, hypertelorism, broad nasal tip, anteverted nares
Tariverdian [50]		Acromegaly, CNS anomalies, macroorchidism
Turner [51]	Xp21.2-q13	Macrocephaly, heterozygote expression
Vasquez [52]		Hypogonadism, gynecomastia, short stature, obesity
Vitale [53]	Xq24	Short stature, brachydactyly, narrow downslanted palpebral fissures, large bulbous nose, macrostomia
Vles [54]		Corpus callosum agenesis, spastic quadriparesis, irregular lining of lateral ventrioles
Wittwer [55]		Square face, high broad forehead, frontal bossing, hypertelorism, downslanting palpebral fiss., anteverted nares
Young-Hughes [56]		Short stature, obesity, hypogonadism

The presence of a second Simpson-Golabi-Behmel locus in Xp22 should be noted[14], as it underscores the possibility of genetic heterogeneity of clinically similar/identical conditions. On the other hand, a clear example of clinical variability is represented by the ATR-X syndrome. Mutations in the *XNP* gene have now been found in at least four other conditions, namely Juberg-Marsidi (#309590), Carpenter-Waziri[15], Holmes-Gang[16], a family with spastic diplegia, microcephaly and short stature[17]. An *XNP* mutation was also found in a family reported as Smith-Fineman-Myers[18], though a separate entry for this latter condition is still included in

Table 1. In this case "lumping" of clinically distinguishable conditions has been possible after the gene cloning and rescreening of patients belonging to families with either a resembling phenotype or linkage to the same chromosomal interval. Until mutational screening allows more lumping, we are convinced that the wisest approach is "splitting", i.e. considering as different any similar but not quite identical conditions, until proven otherwise. Linkage exclusion is sometimes very useful for splitting two conditions segregating in two small families that will not reach the threshold of LOD score significance. In Table 2, another example of lumping is offered by the spectrum of *L1CAM* mutations listed under HSAS (#307000), but the most extreme example of clinical heterogeneity is presently represented by the *RSK2* mutations causing either the Coffin-Lowry syndrome or a nonspecific MRX[12]. It is instructive to note that the Zollino syndrome of pachygyria and MR[19], which was present in the update 1998, has been excluded from this update because it was found to be due to a cryptic subtelomeric translocation.

Table 2 - Neuromuscular disorders

MIM No.	Name	Locus	Gene	Description
*309600	Allan-Herndon-Dudley	Xp11.4-q22		Severe hypotonia, joint contractures, muscular atrophy
*302500	Apak			Spinocerebellar ataxia, nystagmus, dysarthria
*301835	Arts	Xq21.33-q24		Early death, hypotonia, ataxia, deafness, loss of vision, recurrent infections
301840	Ataxia-dementia			Ataxia, pyramidal tract signs, adult-onset dementia
312890	Baar-Gabriel			Athetotic spastic paraplegia
309660	Bergia			Cardiomyopathy (lethal), scapuloperoneal muscular dystrophy, myopia
*302801	CMTX2	Xp22.3-1		Motor-sensory neuropathy, formerly listed as Ionasescu (Family 1)
*310490	Cowchock-Fishbeck	Xq24-q26.1		Motor-sensory neuropathy and deafness **It is possibly allelic to CMTX3 [*302802]**
*310200	Duchenne muscular dystrophy	Xp21.3-1	DMD	Pseudohypertrofic muscular dystrophy
309560	Fitzsimmons			Spastic paraplegia, pes cavus, palmoplantar hyperkeratosis
#312920	Goldblatt	Xq13-q21.1		Complicated spastic paraplegia with nystagmus and optic atrophy
*309555	Gustavson	Xq25-q26		Optic atrophy, hearing loss, epilepsy, spasticity, restricted joint mobility, early death
#307000	HSAS	Xq28	L1CAM *308840	Hydrocephalus with stenosis of the aqueduct of Silvius. **It includes spastic paraplegia I [#312900] and MASA [#303350]**
*304100	Menkes-Kaplan			Partial agenesis of corpus callosum, seizures
*304700	Mohr-Tranebjaerg	Xq22	DDP	Hearing loss, visual impairment, ataxia, spastic paraplegia **It includes Jensen [#311150]**
*310600	Norrie	Xp11.3	NDP	Blindness, hearing loss
*311050	OPA-2	Xp11.4-p11.21		Optic atrophy, abnormal reflexes, dysarthria, tremor
311400	Paine & Seemanova			Spastic diplegia, myoclonic seizures, cerebellar hypoplasia
*312080	Pelizaeus-Merzbacher	Xq21.33-q22	PLP	Spasticity, cerebellar ataxia, parkinsonism. **It includes spastic paraplegia II [#312920]**
308850	Plott			Laryngeal abductor paralysis
*300220	Reyniers/MRXS10	Xp22.1-q21.3		Choreoathetosis, dysarthria, psychosis
301790	Schmidley			Hypotonia, ataxia, sensorineural deafness, optic atrophy, early demise

#300067	SCLH/XLIS	Xq22.3-q23	*DCX* *300121*	Subcortical laminar heterotopia in females, lissencephaly and epilepsy in males
*311510	Waisman-Laxova	Xq27.2-qter		Parkinsonism, seizures, apparent basal ganglia degeneration
*308350	West	Xp21.3-p22.1		Infantile spasms, hypsarrythmia, early death
*314580	Wieacker-Wolff	Xp11.3-q13		Contractures, distal muscular atrophy, dyspraxia of ocular and facial muscles
	Arena [57]			Spastic paraplegia, ataxin, titubation, iron deposits in basal ganglia
	Berry-Kravis (XLAG) [58]			Lissencephaly with frontal pachygiria and posterior agyria, agenesis of corpus callosum, neonatal epilepsy and hypotonia, hypothermia, ambiguous genitalia
	Bertini [59]	Xp22.3		Ataxia, hypotonia, recurrent infections
	Cabezas [60]	Xq22.3-q25		Short stature, small testes, muscle wasting, tremor
	Claes [61]			Slowly progressive spastic paraplegia, maxillary hypoplasia, facial hypotonia
	Fried [62]	Xp22		Hydrocephaly, spastic diplegia, calcification of basal ganglia
	Garcia [63]	Xp11-q13		Learning difficulty, epilepsy, intermittent aggressive behaviour
	Hamel BCD [64]	Xp11.3-q21.3		Blindness, convulsions, hypomyelination, spasticity, early death
	SPG7 [65]	Xq11.2-q23		Quadriplegia, motor aphasia, reduced vision, dysfunction of bowel and bladder
	Tranebjaerg II [66]			Dyspraxia, ataxia, seizures, pes equinovarus, macroorchidism
	XMRE [67]	Xp21.1-p11.4		Generalized tonic-clonic and atonic seizures, moderate MR, normal electromyography and nerve conduction

Table 3 lists metabolic conditions whose genes have almost all been cloned, while Table 4 summarizes the information on eight X-linked dominant disorders; genes for three of these were recently cloned. Only one of these conditions, EFMR (*300088), is not lethal in males: actually it affects exclusively females and it has been hypothesized that the responsible gene may have a homologue on the Y chromosome.

Table 3 - Metabolic conditions

MIM No.	Name	Locus	Gene	Description
*300100	Adrenoleukodystrophy	Xq28	*ALD*	Spastic quadriparesis, impaired vision, ataxia, dementia
*300123	MRGH	Xq24-q27.3		Isolated GH deficiency, short stature, small sella turcica, overlap with *312000 ?
*307030	Hyperglycerolemia	Xp21.3	*GK1*	Glyceroluria, poor growth, esotropia, osteoporosis
*309900	Hunter disease	Xq28	*IDS*	"Coarse" face, dysostosis multiplex, dwarfism, hepatosplenomegaly, heart involvement
*308000	Lesch-Nyhan	Xq26	*HPRT*	Cerebral palsy, choreoathetosis, self-distructive biting
*309000	Lowe	Xq25-q26.1	*OCRL1*	Hydrophthalmia, cataract, vitamin D-resistant rickets
*309850	MAO-A deficiency	Xp11.3	*MAOA*	Aggressive behaviour, disturbance in monoamine Metabolism
#309400	Menkes	Xq13	*ATP7A* *300011*	Growth retardation, peculiar hair, focal cerebral & cerebellar degeneration. It includes occipital horn syndrome or cutis laxa [#304150]
*311250	OTC deficiency	Xp21.1	*OTC*	Hyperammonemia
*312000	Panhypopituitarism	Xq25-q26		Combined deficiency of pituitary hormones, overlap with *300123 ?
*311800	PGK1 deficiency	Xq21.1	*PGK1*	Myoglobinuria, epilepsy, hemolytic anemia
*312170	Pyruvate DH complex E1a subunit deficiency	Xp22.1	*PDHA1*	Lactic acidosis, ataxia

Table 4 - Dominant conditions

MIM No.	Name	Locus	Gene	Description
*304050	Aicardi	Xp22		Agenesis of corpus callosum,chorioretinopathy, microphthalmia, seizures
#300049	BPNH	Xq28	*FLN1* *300017	Epilepsy, periventricular nodular heterotopia, mild hypoplasia of corpus callosum and/or cerebellum in females; syndactyly and severe MR in rare affected males
*300088	EFMR	Xq21.3-q22.2		MR and epilepsy in females only, males spared
*305600	Goltz			Focal dermal hypoplasia, short missing digits, polysyndactyly, microphthalmia
#308310	Incontinentia pigmenti Type II	Xq28	*NEMO* *300248	Incontinentia pigmenti, incomplete dentition, retinal abnormalities
*309801	MIDAS	Xp22		Microphthalmia, dermal aplasia, sclerocornea
*311200	OFD1	Xp22.2-3		Midline clefting of face, tongue nodules, syndactyly
#312750	Rett	Xq28	*MeCP2* *300005	Ataxia, autism, dementia

A brief note should be made here to remember that a skewed X inactivation is often observed in carrier females of these dominant mutations. One may hypothesize that cells which inactivate the normal allele die preferentially during embryogenesis or divide less than those with the normal allele on the active X chromosome. Sometimes skewage of X-inactivation is observed also in female carriers of conditions which are not lethal in males (such as ATR-X). However, in many cases there is random X-inactivation, probably because no selection occurs during early development, and female carriers may have different degrees of involvement depending on the fraction of mutant alleles on the active X chromosome. In heterozygous females two distinct cell populations will exist side by side, as nicely demonstrated, for instance, by the formation of a "double cortex" in female carriers of the *XLIS* (#300067) mutant gene (Table 2).

Table 5 - MRX cloned genes

MIM No.	Name	Locus	Gene	Includes	Excludes
*309548	FRAXE	Xq28	*FMR2*		
*300104	RABGDIA	Xq28	*GDI1*	MRX41 & 48	MRX3,25,28 & 16
*300127	Oligophrenin-1	Xq12	*OPHN1*	MRX60	MRX14 & 52
*300142	p21 Act. Kinase 3	Xq22	*PAK3*	MRX30 & 47	
*300206	IL1 Rec acc. Protein	Xp22.1-p21.3	*IL1RAPL*	MRX34 (deletion at DXS1218)	MRX2, 36, 43, 54
*300096	Tetraspanin	Xp11.4	*TM4SF2*		MRX15 & 55
*300267	αPIX	Xq26	*ARHGEF6*	MRX46	

For reasons of space table 5 lists only the cloned MRX genes; the complete listing of MRX conditions with their localization can be retrieved from the on-line XLMR database. Several of these genes encode proteins involved in the control of the cytoskeleton and neurite outgrowth or in synaptic vesicle cycling[10], i.e. proteins that have a specific and limited task in the functioning

of the central nervous system[20] and will not disturb development nor impair cell viability. Table 6 has the total counts.
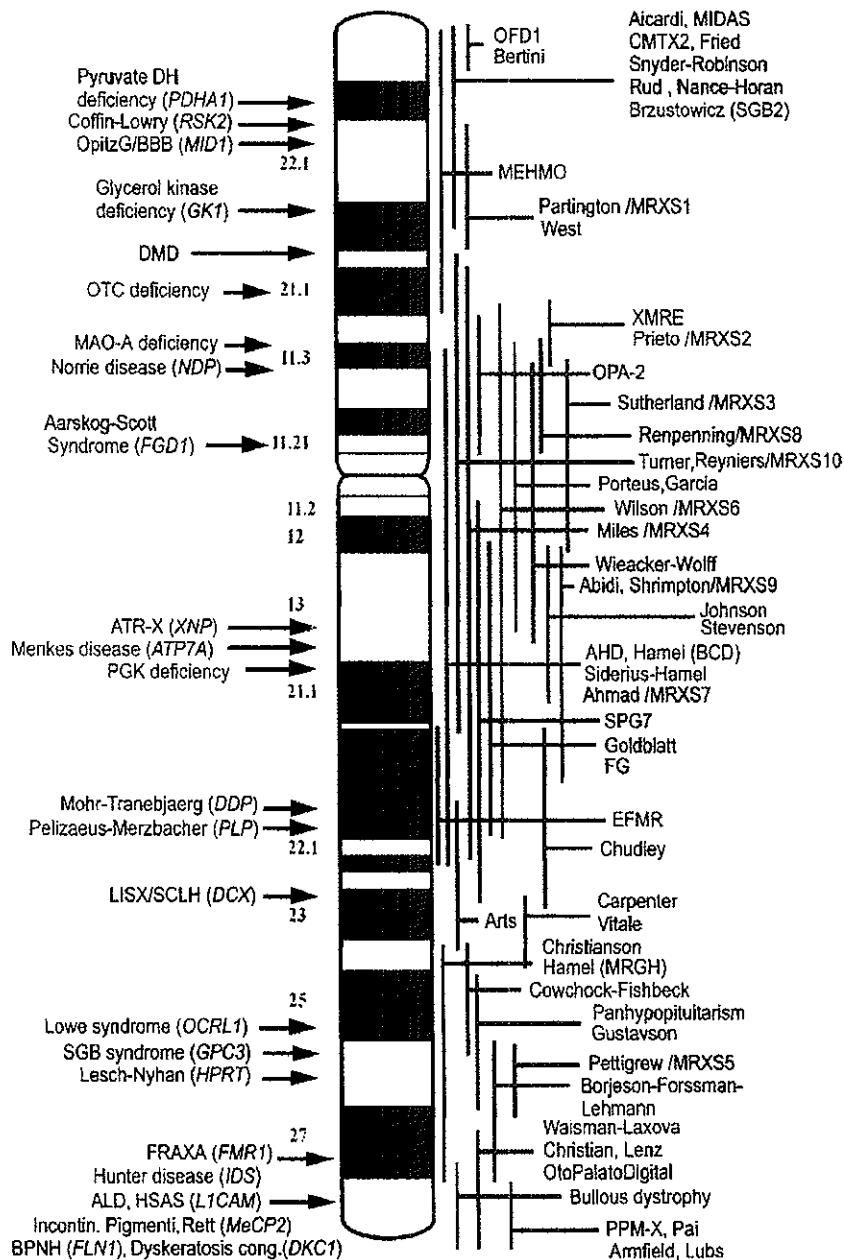
Table 6 - XLMR Genes: Total counts

	Total	Mapped	Cloned
Malformations syndromes	79	34	7
Neuromuscular disorders	37	18	6
Metabolic conditions	12	2	10
Dominant conditions	8	4	3
Total MRXS	**136**	**58**	**26**
MRX	**66**	**59**	**7**
Total entries	**202**	**117**	**33**

A total of 202 conditions (136 MRXS and 66 MRX) have been included in the update (October 19, 2000). The genes of 117 of these (58 MRXS and 59 MRX) have been regionally mapped and 33 (26 MRXS and 7 MRX) cloned. Figures 1 and 2 depict the cloned/mapped MRXS and MRX genes, respectively.

The on-line XLMR database

To integrate the published XLMR update we designed a simple on-line catalogue, the XLMR database, which is now available on the Internet at http://xlmr.interfree.it/home.htm.

This is obviously no substitute for information deriving from journals, books, expert systems and clinical experience, but it could become a useful tool if most of the fellow researchers will help to improve it and to keep it updated. Potentially, it offers a number of advantages over any published document: it can be updated as frequently as needed and there is much more space for additional information that may never be published. Presently, the on-line catalogue contains an expanded version of Tables 1-6 of this update (extra colums are available and the OMIM numbers and references are hyperlinked). A colour-coded version of Figures 1 and 2 is also available and we plan to hyperlink the conditions mentioned there with the corresponding Tables. An individual record for every condition would be highly desirable, possibly with more information on the cloned genes (with a link to Locuslink of NCBI, at least). Also thanks to the space available, the complete linkage data might be attached and information on all tested loci (and not only recombinant ones) could be compared between overlapping conditions.

Pyruvate DH
deficiency (*PDHA1*)
Coffin-Lowry (*RSK2*)
OpitzG/BBB (*MID1*)
22.1

Glycerol kinase
deficiency (*GK1*)

DMD

OTC deficiency   21.1

MAO-A deficiency
Norrie disease (*NDP*)   11.3

Aarskog-Scott
Syndrome (*FGD1*)   11.21

11.2
12

13

ATR-X (*XNP*)
Menkes disease (*ATP7A*)
PGK deficiency
21.1

Mohr-Tranebjaerg (*DDP*)
Pelizaeus-Merzbacher (*PLP*)
22.1

LISX/SCLH (*DCX*)
23

25

Lowe syndrome (*OCRL1*)
SGB syndrome (*GPC3*)
Lesch-Nyhan (*HPRT*)

27

FRAXA (*FMR1*)
Hunter disease (*IDS*)
ALD, HSAS (*L1CAM*)
Incontin. Pigmenti,Rett (*MeCP2*)
BPNH (*FLN1*), Dyskeratosis cong.(*DKC1*)

OFD1
Bertini

Aicardi, MIDAS
CMTX2, Fried
Snyder-Robinson
Rud , Nance-Horan
Brzustowicz (SGB2)

MEHMO

Partington /MRXS1
West

XMRE
Prieto /MRXS2

OPA-2

Sutherland /MRXS3
Renpenning/MRXS8
Turner,Reyniers/MRXS10
Porteus,Garcia
Wilson /MRXS6
Miles /MRXS4
Wieacker-Wolff
Abidi, Shrimpton/MRXS9
Johnson
Stevenson
AHD, Hamel (BCD)
Siderius-Hamel
Ahmad /MRXS7
SPG7
Goldblatt
FG

EFMR
Chudley

Carpenter
Vitale
Christianson
Hamel (MRGH)
Cowchock-Fishbeck
Panhypopituitarism
Gustavson
Pettigrew /MRXS5
Borjeson-Forssman-
Lehmann
Waisman-Laxova
Christian, Lenz
OtoPalatoDigital
Bullous dystrophy
PPM-X, Pai
Arnfield, Lubs

Arts

Figure 1 – Ideogram of the X chromosome (G-banding) with the localization of the 26 cloned (arrows) and 58 mapped (bars) MRXS genes.
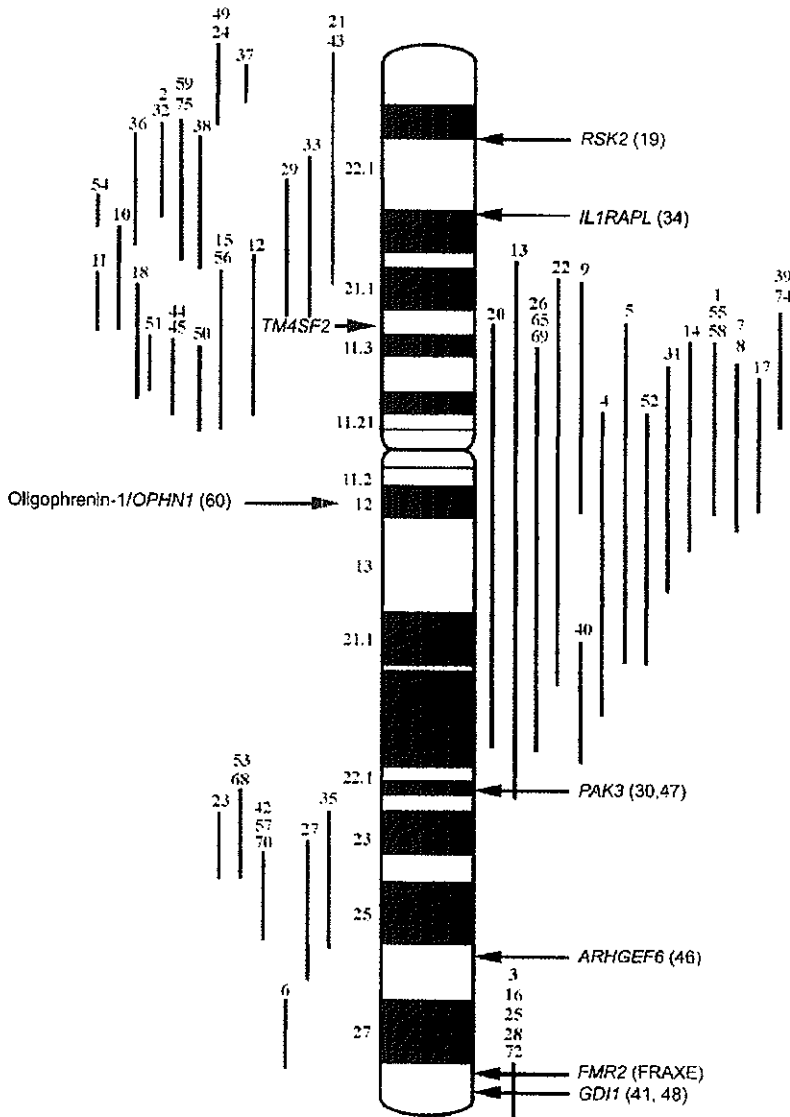
Figure 2 - Ideogram of the X chromosome (G-banding) with the localization of the 7 cloned MRX genes (arrows) and of the 59 mapped MRX families (bars). The extra cloned gene indicated is *RSK2*, which is also depicted in Figure 1 and most often mutated in the Coffin-Lowry syndrome (Table 1).

Negative linkage data, which are not easily publishable but still very relevant and useful, may be stored. Eventually, images could be attached to the card describing each specific XLMR condition. Finally, a listing of available DNA from small families or individual cases with cytogenetic rearrangements in the various laboratories researching on XLMR may be envisaged, which would certainly boost the opportunity for collaboration. In brief, this informal Web site could become a "meeting point" for the XLMR community.

Future research

Although the primary sequence of the human genome will soon be completed, the cloning era is far from being concluded and much "sequence mining" and "traditional" cloning remains to be done to account for the 202 XLMR conditions listed in this update. Many of these may eventually prove to be due to mutations in the same gene but the present number of 33 cloned XLMR genes falls definitely short of the actual one. Unexpectedly, mutations in already cloned genes were not found to account for mental retardation in families mapping to the same narrow intervals where these genes are located. Therefore, new strategies will have to be designed to exploit even small families and individual patients for the identification of new XLMR genes. Some recent observations are encouraging. We now appreciate the value that even small families or individual patients with cytogenetic rearrangements can have in helping to pin down the remaining genes. For example, the MRX genes *ILIRAPL* and *TM4SF2* were identified thanks to the observation of patients with a microdeletion and an X;autosome translocation, respectively, and confirmatory mutations were found in small families that could not reach a linkage significance. DNA chip technology will hopefully allow (re)screening of large number of patient DNAs for mutations in several candidate genes at the same time as well as testing the expression of these genes in informative families. Finally, the next challenge awaiting us will be the unraveling of the structure and function of the proteins encoded by the XLMR genes. This will be a daunting task belonging to the new era of "proteomics" but will surely lead us to a better understanding of the biological pathways required for the normal development of intelligence.

Acknowledgements

References

1. Chiurazzi P, Oostra BA. Genetics of mental retardation. *Curr Op Pediat* 2000; **12**: 529-535.

2. Lehrke R. Theory of X-linkage of major intellectual traits. *Am J Ment Defic* 1972; 76: 611-619.

3. Turner G, Turner B. X-linked mental retardation. *J Med Genet* 1974; **11**: 109-113.

4. Neri G, Chiurazzi P. X-linked mental retardation. *Adv Genet* 1999; **41**: 55-94.

5. Herbst DS, Miller JR. Nonspecific X-linked mental retardation II: the frequency in British Columbia. *Am J Med Genet* 1980; 7: 461-469.

6. Turner G, Webb T, Wake S, Robinson H. Prevalence of fragile X syndrome. *Am J Med Genet* 1996; **64**: 196-197.

7. Neri G, Gurrieri F, Gal A, Lubs HA. XLMR genes: update 1990. *Am J Med Genet* 1991; **38**:186-189.

8. Lubs H, Chiurazzi P, Arena J, Schwartz C, Tranebjaerg L, Neri G. XLMR genes: update 1998. *Am J Med Genet* 1999; **83**: 237-247.

9. Hamel B, Chiurazzi P, Lubs H. Syndromic XLMR genes (MRXS): update 2000. *Am J Med Genet*, in press.

10. Chelly J. Breakthroughs in molecular and cellular mechanisms underlying X-linked mental retardation. *Hum Mol Genet* 1999; **8**:1833-1838.

11. Toniolo D, D'Adamo P. X-linked non-specific mental retardation. *Curr Opin Genet Dev* 2000; **10**: 280-285.

12. Merienne K, Jacquot S, Pannetier S et al. A missense mutation in RPS6KA3 (RSK2) responsible for non-specific mental retardation. *Nat Genet* 1999; **22**: 13-14.

13. Orrico A, Lam CW, Galli L et al. MeCP2 mutation in male patients with non-specific X-linked mental retardation. *FEBS Lett* 2000; **24106**: 1-4.

14. Brzustowicz LM, Farrell S, Khan MB, Weksberg R. Mapping of a new SGBS locus to chromosome Xp22 in a family with a severe form of Simpson-Golabi-Behmel syndrome. *Am J Hum Genet* 1999; **65**: 779-783.

15. Abidi F, Schwartz CE, Carpenter NJ, Villard L, Fontes M, Curtis M. Carpenter-Waziri syndrome results from a mutation in XNP. *Am J Med Genet* 1999; **85**: 249-251.

16. Schwartz C, Abidi A, Lubs H, Stevenson R. A mutation in *XNP* results in the Holmes-Gang syndrome: further evidence that mutations in *XNP* may account for a relatively large fraction of mental retardation in males. *Eur J Hum Genet* 1999; 7 (S1): 37.

17. Lossi AM, Millan JM, Villard L et al. Mutation of the XNP/ATR-X gene in a family with severe mental retardation, spastic paraplegia and skewed pattern of X inactivation: demonstration that the mutation is involved in the inactivation bias. *Am J Hum Genet* 1999; **65**: 558-562.

18. Villard L, Fontes M, Ades LC, Gecz J. Identification of a mutation in the XNP/ATR-X gene in a family reported as Smith-Fineman-Myers syndrome. *Am J Med Genet* 2000; **91**: 83-85.

19. Zollino M, Mastroiacovo P, Zampino G, Mariotti P, Neri G. A new XLMR syndrome with characteristic face, hypogenitalism, congenital hypotonia and pachygyria. *Am J Med Genet* 1992; **43**: 452-457.

20. Antonarakis SE, Van Aelst L. Mind the GAP, Rho, Rab and GDI. *Nat Genet* 1998; **19**: 106-108.

21. Siber M. X-linked recessive microcencephaly, microphthalmia with corneal opacities, spastic quadriplegia, hypospadias and cryptorchidism. *Clin Genet* 1984; **26**: 453-456.

22. Abidi F, Hall BD, Cadle RG et al. X-linked mental retardation with variable stature, head circumference, and testicular volume linked to Xq12-q21. *Am J Med Genet* 1999; **85**: 223-229.

23. Ahmad W, De Fusco M, ul Haque MF et al. Linkage mapping of a new syndromic form of X-linked mental retardation, MRXS7, associated with obesity. *Eur J Hum Genet* 1999; 7: 828-832.

24. Aldred MA, Dry KL, Knight-Jones EB et al. Genetic analysis of a kindred with X-linked mental handicap and retinitis pigmentosa. *Am J Hum Genet* 1994; **55**: 916-922.

25. Armfield K, Nelson R, Lubs HA et al. X-linked mental retardation syndrome with short stature, small hands and feet, seizures, cleft palate, and glaucoma is linked to Xq28. *Am J Med Genet* 1999; **85**: 236-242.

26. Atkin JF, Flaitz K, Patil S, Smith W. A new X-linked mental retardation syndrome. *Am J Med Genet* 1985; **21**: 697-705.

27. Baraitser M, Reardon W, Vijeratnam S. Nonspecific X-linked mental retardation with macrocephaly and obesity: a further family. *Am J Med Genet* 1995; **57**: 380-384.

28. Carpenter NJ, Givens H, Randell L, Lutz R, Miles JH. Clinical characterization and gene mapping of a family with X-linked mental retardation, facial dysmorphism, congenital hip dislocation and skewed pattern of X-inactivation. *Am J Hum Genet* 2000; 67(Suppl.2):315.

29. Chudley AE, Tackels DC, Lubs HA et al. X-linked mental retardation syndrome with seizures, hypogammaglobulinemia, and progressive gait disturbance is regionally mapped between xq21.33 and Xq23. *Am J Med Genet* 1999; **85**: 255-262.

30. Golabi M, Ito M, Hall BD. A new X-linked multiple congenital anomalies/mental retardation syndrome. *Am J Med Genet* 1984; **17**: 367-374.

31. Hall BD, Robl JM. New X-linked MR/MCA syndrome associated with cleft lip/palate, upslanted/short palpebral fissures, high nasal bridge, prominent nasal tip, inguinal hernia and minor digital defects. *Am J Hum Genet* 1999; **65**: A151.

32. Hamel BC, Mariman EC, van Beersum SE, Schoonbrood-Lenssen AM, Ropers HH. Mental retardation, congenital heart defect, cleft palate, short stature, and facial anomalies: a new X-linked multiple congenital anomalies/mental retardation syndrome: clinical description and molecular studies. *Am J Med Genet* 1994; **51**: 591-597.

33. Hockey A. X-linked intellectual handicap and precocious puberty with obesity in carrier females. *Am J Med Genet* 1986; **23**: 127-137.

34. Homfray T, Holland T, Patton M. A new X-linked mental retardation syndrome. *Clin Dysmorphol* 1995; **4**: 289-293.

35. Johnson JP, Nelson R, Schwartz CE. A family with mental retardation, variable macrocephaly and macro-orchidism, and linkage to Xq12-q21. *J Med Genet* 1998; **35**: 1026-1030.

36. Kang WM, Huang CC, Lin SJ. X-linked recessive inheritance of dysgenesis of corpus callosum in a Chinese family. *Am J Med Genet* 1992; **44**: 619-623.

37. Lubs H, Abidi F, Bier JA et al. XLMR syndrome characterized by multiple respiratory infections, hypertelorism, severe CNS deterioration and early death localizes to distal Xq28. *Am J Med Genet* 1999; **85**: 243-248.

38. Steinmuller R, Steinberger D, Muller U. MEHMO (mental retardation, epileptic seizures, hypogonadism and -genitalism, microcephaly, obesity), a novel syndrome: assignment of disease locus to xp21.1-p22.13. *Eur J Hum Genet* 1998; **6**: 201-206.

39. Oosterwijk JC, Wischmeijer A, Losekoot M, Haraldson A, Theunissen G, van Gelderen I. A new X-linked mental retardation syndrome with distal limb defects, hearing impairment, verrucosis and immunodeficiency. *Am J Hum Genet* 1999; **65**: A337.

40. Porteous MEM, Johnson H, Burn J, Curtis A, Lindsay S, Bhattacharya SS, Goodship JA. A new mental retardation syndrome mapping to the pericentromeric region of the X chromosome. *Am J Hum Genet* 1992; **51**: A106.

41. Reish O, Gorlin RJ, Hordinsky M, Rest EB, Burke B, Berry SA. Brain anomalies, retardation of mentality and growth, ectodermal dysplasia, skeletal malformations, Hirschsprung disease, ear deformity and deafness, eye hypoplasia, cleft palate, cryptorchidism, and kidney dysplasia/hypoplasia (BRESEK/BRESHECK): new X-linked syndrome? *Am J Med Genet* 1997; **68**: 386-390.

42. Seemanova E, Lesny I. X-linked microcephaly, microphthalmia, microcornea, congenital cataract, hypogenitalism, mental deficiency, growth retardation, spasticity: possible new syndrome. *Am J Med Genet* 1996; **66**: 179-183.

43. Shashi V, Berry MN, Shoaf S, Sciote JJ, Goldstein D, Hart TC. A unique form of mental retardation with a distinctive phenotype maps to Xq26-q27. *Am J Hum Genet* 2000; **66**: 469-479.

44. Shrimpton AE, Daly KM, Hoo JJ. Mapping of a gene (MRXS9) for X-linked mental retardation, microcephaly, and variably short stature to Xq12-q21.31. *Am J Med Genet* 1999; **84**: 293-299.

45. Siderius LE, Hamel BC, van Bokhoven H et al. X-linked mental retardation associated with cleft lip/palate maps to Xp11.3-q21.3. *Am J Med Genet* 1999; **85**: 216-220.

46. Brooks SS, Wisniewski K, Brown WT. New X-linked mental retardation (XLMR) syndrome with distinct facial appearance and growth retardation. *Am J Med Genet* 1994; **51**: 586-590.

47. Stevenson RE, Hane B, Arena JF et al. Arch fingerprints, hypotonia, and areflexia associated with X linked mental retardation. *J Med Genet* 1997; **34**: 465-469.

48. dos Santos RC, Barretto OC, Nonoyama K et al. X-linked syndrome: mental retardation, hip luxation, and G6PD variant. *Am J Med Genet* 1991; **39**: 133-136.

49. Stoll C, Geraudel A, Chauvin A. New X-linked syndrome of mental retardation, short stature, and hypertelorism. *Am J Med Genet* 1991; **39**: 474-478.

50. Tariverdian G, Froster-Iskenius U, Deuschl G, Wolff G. Mental retardation, acromegalic face, and megalotestes in two half-brothers: a specific form of X-linked mental retardation without fra(X)(q)? *Am J Med Genet* 1991; **38**: 208-211.

51. Turner G, Gedeon A, Mulley J. X-linked mental retardation with heterozygous expression and macrocephaly: pericentromeric gene localization. *Am J Med Genet* 1994; **51**: 575-580.

52. Vasquez SB, Hurst DL, Sotos JF. X-linked hypogonadism, gynecomastia, mental retardation, short stature, and obesity--a new syndrome. *J Pediatr* 1979; **94**: 56-60.

53. Vitale E, Specchia C, Devoto M et al. A novel X-linked mental retardation syndrome with short stature maps to Xq24. *Am J Hum Genet* 2000; **67**(Suppl.2):332.

42

54. Vles JS, Fryns JP, Folmer K et al. Corpus callosum agenesis, spastic quadriparesis and irregular lining of the lateral ventricles on CT-scan. A distinct X-linked mental retardation syndrome? *Genet Couns* 1990; 1: 97-102.

55. Wittwer B, Kircheisen R, Leutelt J, Orth U, Gal A. New X-linked mental retardation syndrome with the gene mapped tentatively in Xp22.3. *Am J Med Genet* 1996; 64: 42-49.

56. Young ID, Hughes HE. Sex-linked mental retardation, short stature, obesity and hypogonadism: report of a family. *J Ment Defic Res* 1982; 26: 153-162.

57. Arena JF, Schwartz C, Stevenson R et al. Spastic paraplegia with iron deposits in the basal ganglia: a new X-linked mental retardation syndrome. *Am J Med Genet* 1992; 43: 479-490.

58. Dobyns WB, Berry-Kravis E, Havernick NJ, Holden KR, Viskochil D. X-linked lissencephaly with absent corpus callosum and ambiguous genitalia. *Am J Med Genet* 1999; 86: 331-337.

59. des Portes V, Bachner L, Bruls T et al. X-linked neurodegenerative syndrome with congenital ataxia, late-onset progressive myoclonic encephalopathy and selective macular degeneration, linked to Xp22.33-pter. *Am J Med Genet* 1996; 64: 69-72.

60. Cabezas DA, Slaugh R, Abidi F et al. A new X linked mental retardation (XLMR) syndrome with short stature, small testes, muscle wasting, and tremor localises to Xq24-q25. *J Med Genet* 1997; 37: 663-668.

61. Claes S, Devriendt K, Van Goethem G et al. Novel syndromic form of X-linked complicated spastic paraplegia. *Am J Med Genet* 94:1-4.

62. Strain L, Wright AF, Bonthron DT. Fried syndrome is a distinct X linked mental retardation syndrome mapping to Xp22. *J Med Genet* 1997; 34: 535-540.

63. Garcia CC, Buddles MRH, Seager M, Goodship JA. A family with X-linked epilepsy mapping to Xp11-Xq13. *Am J Hum Genet* 1999; 65: A250.

64. Hamel BC, Wesseling P, Renier WO et al. A new X linked neurodegenerative syndrome with mental retardation, blindness, convulsions, spasticity, mild hypomyelination, and early death maps to the pericentromeric region. *J Med Genet* 1999; 36: 140-143.

65. Steinmuller R, Lantigua-Cruz A, Garcia-Garcia R, Kostrzewa M, Steinberger D, Muller U. Evidence of a third locus in X-linked recessive spastic paraplegia. *Hum Genet* 1997; 100: 287-289.

66. Tranebjaerg L, Lou H, Andresen J. New X-linked syndrome with apraxia, ataxia, and mental deficiency: clinical, cytogenetic and neuropsychological studies in two Danish families. *Am J Med Genet* 1992; 43: 498-504.

67. Hedera P, Alvarado D, Beydoun A, Fink JK. Novel mental retardation-epilepsy syndrome linked to Xp21.1-p11.4. *Am J Hum Genet* 2000; 67(Suppl.2):332.

1.3

Fragile X syndrome

The fragile X syndrome is probably the most frequent form of inherited mental retardation. It is almost always caused by the expansion of a (CGG)ₙ sequence in the promoter region of the *FMR1* gene. The expanded sequence becomes almost invariably methylated and the gene is silenced. The resulting lack of the FMR1 protein (FMRP) causes the clinical phenotype.

The fragile X syndrome [Kooy et al., 2000] is the prototype of a growing list of disorders known to be caused by the so-called "dynamic mutations", resulting from the instability and expansion of trinucleotide repeats [Djian, 1998]. In the case of the fragile X syndrome the mutant gene *FMR1* is located at Xq27.3 and contains a repeated CGG triplet in its 5' untranslated region [Verkerk et al., 1991]. Patients with fragile X syndrome harbour an expanded (>200 repeats) and abnormally methylated CGG sequence that cannot be transcribed into mRNA, causing a lack of FMRP. The syndrome derives its name from a fragile site on the X chromosome, FRAXA, which was first observed by Lubs [1969] in four mentally retarded males and three obligate carrier females of the same family. FRAXA was later shown to colocalize with the CGG repeat in Xq27.3. The expression of the fragile X site is best induced when cells are cultured with low folate concentration and either fluorodeoxyuridine or an excess of thymidine is added [Jacky et al., 1991]. The fragile site is usually expressed in 30 to 50% of the cells examined. However, a lower expression is not unusual and in fact it can be sometimes as low as 4 or 5%, especially in carrier females. It appears as a decondensed chromatin gap between Xq28 and the rest of the X chromosome which sometimes can be actually broken; indeed it has been shown that DNA replication is delayed well after the S phase in the region containing the expanded CGG repeat and could be incomplete at mitosis, thus determining the chromosomal "fragility" [Hansen et al., 1993]. The first large family with mental retardation and macroorchidism in males transmitted in an X-linked fashion, and later confirmed to have fragile X syndrome, has been described over 50 years ago by Martin and Bell [1943] and their names have been often used as eponym for the syndrome.

Clinical phenotype

The clinical phenotype of the fragile X syndrome can be quite variable. In typical cases there is tall stature and relative macrocephaly, long and narrow face with prominent forehead and mandible, and midface hypoplasia with hypoteloric, sunken eyes. The ears are large and the palate is highly arched. Testes are generally large with volumes up to 100 ml. Generalized muscular hypotonia is a virtually constant finding and is usually accompanied by joint laxity.

These latter findings might be caused by an underlying connective tissue dysplasia, which could also be responsible for the frequently observed mitral valve prolapse. Mental retardation is usually of moderate degree and the behaviour tends to be introverted, with poor eye contact and avoidance of new and unexpected situations. In extreme cases this behaviour can be described as autistic. The phenotype is usually more subtle in newborns and children, where facial traits tend to be less pronounced and macroorchidism is less obvious. Increased birthweight and generalized congenital hypotonia may be the only significant findings. Autistic behaviour, hyperactivity and attention deficit have been described in children. Seizures may also occur during infancy and a characteristic EEG pattern has been reported, which is characterized by trains of medium-high voltage spikes, discharging from the temporal regions during sleep [Sanfilippo et al., 1986; Musumeci et al., 1999]. Epileptic seizures, if present, generally disappear before puberty and appear to respond well to treatment. Brain MRI of patients shows overall volume conservation of brain tissue with a diminished white-to-gray matter ratio and relatively enlarged caudate nucleus and hippocampus, while cerebrospinal fluid is increased especially in the lateral ventricles [Reiss et al., 1994; Reiss et al., 1995]. The fourth ventricle is also enlarged in correspondence to a smaller posterior cerebellar vermis [Reiss et al., 1991a; Reiss et al., 1991b]. Among non typical cases of the syndrome, a subgroup was identified which, because of obesity and short stature, bore some resemblance to the Prader-Willi syndrome [de Vries et al., 1993]. Although this is nothing more than a superficial similarity, it is a good reminder of the pitfalls of a purely clinical diagnosis and justifies the view that every mentally retarded person should be tested for fragile X syndrome, in absence of another reasonable diagnosis. The affected females, who represent about one third of all females carrying a full mutation, usually do not demonstrate a characteristic physical phenotype. They are mildly retarded and may present only a learning disability, with a shy and introverted personality.

Gene structure and protein isoforms

The *FMR1* gene structure has been determined in detail, and this revealed that the seventeen exons of the gene are embedded in 38 kb of genomic sequence in Xq27.3 [Eichler et al., 1993]. The polymorphic CGG repeat is located in the 5' untranslated region of exon 1 and is included in all *FMR1* transcripts [Verkerk et al., 1993]. *FMR1* was shown to be ubiquitously transcribed during murine and human embryogenesis [Hinds et al., 1993] with the highest level of expression in differentiated neurons of the hippocampus and of basal ganglia [Abitbol et al., 1993], while in adult mice *Fmr1* is mainly expressed in neurons and in spermatogonia. FMR1

protein (FMRP) has been detected in synapses, in dendritic spines and soma of rat neurons, but not in their nucleus or axon, and active FMRP production has been demonstrated near synapses in response to neurotransmitter activation [Weiler et al., 1997]. It is probable that FMRP action is required for normal maturation of synaptic connections, which indeed appear immature and are reduced in number in fragile X brains [Hinton et al., 1991], though neurons are neither lost nor degenerating.

The 4.4 kb full length mRNA can code for a protein with a maximum length of 632 amino acids and an apparent molecular mass of 70-80 kDa [Verheij et al., 1993; Devys et al., 1993] and although twenty different transcripts might be produced by alternative splicing [Verkerk et al., 1993; Ashley et al., 1993a], only 4-5 of them and their corresponding protein products are actually detected in various tissues [Verheij et al., 1995]. Isoform 7 (ISO7), which lacks only the 21 amino acids of exon 12, represents the most prominent of all the FMR1 proteins and corresponds to the highest band on Western blotting with an approximate molecular mass of 80 kDa [Sittler et al., 1996]. Two KH domains (KH1 and KH2) and one RGG box, common to several RNA-binding proteins, have been identified in exons 8, 10 and 15, respectively [Siomi et al., 1993; Ashley et al., 1993b]. It was shown that FMRP could bind synthetic RNAs *in vitro*, and the importance of KH domains was underscored by the description of a severely retarded fragile X patient with a point mutation (Ile304Asn) in the KH2 domain [De Boulle et al., 1993] that strongly impaired the RNA binding activity of FMRP [Siomi et al., 1994], abolishing the binding to polyribosomes [Feng et al., 1997]. It seems that FMRP binds mRNAs and participates in mRNP (ribonucleoprotein) particle formation [Corbin et al., 1997; Feng et al., 1997], but is also found associated with the 60S large ribosomal subunit [Khandjian et al., 1996; Tamanini et al., 1996; Siomi et al., 1996], either directly (binding the rRNA with its KH domains and RGG box) or indirectly (via other ribosomal proteins [Siomi et al., 1996]). ISO7 is localized in the cytoplasm [Verheij et al., 1993; Devys et al., 1993] and actually wild-type FMRP has been detected at the electron microscopical level associated with ribosomes attached to the endoplasmatic reticulum and free in the cytoplasm, but also some molecules are present in the nucleoli where the ribosomal precursor particles are assembled [Willemsen et al., 1996a]. Studies with deletion constructs suggested that the N-terminus of the protein encoded by exons 1-5, is able to direct FMRP to the nucleus. A nuclear export signal (NES) is encoded by exon 14 [Sittler et al., 1996] and was shown to act as the NES of the HIV-1 Rev regulatory protein [Fridell et al., 1996], suggesting that FMRP shuttles between cytoplasm and nucleus [Tamanini

et al., 1999a]. Anyhow, it is possible that only a minor fraction of wild-type FMRP is actually entering the nucleus, while most of it could be active directly in the cytoplasm.

The serendipitous discovery of a fragile X related protein, FXR1P [Siomi et al., 1995], led to the search of other FXR proteins possibly interacting with FMRP or complementing its functions. FXR2P was identified with the yeast two-hybrid system [Zhang et al., 1995] and can bind FXR1P. Both FXRPs and FMRP are highly homologous in the N-terminal portion including the KH domains, the RGG box and the first half of exon 14 (the ribosome binding site coinciding with the NES) [Zhang et al., 1995] and their genes probably evolved from a common ancestral one. An FXR1P isoform is highly expressed in skeletal muscle and postmeiotic spermatids, and absent in differentiated neurons and in spermatogonia [Coy et al., 1995]. More information on the differential expression of FMR1, FXR1 and FXR2 proteins in human brain and testis has been reported by Tamanini et al. [1997].

All three proteins can interact with each other and form hetero- as well as homodimers *in vitro* [Zhang et al., 1995]. FMRP-FXR2P complexes *in vitro* are formed by the interaction of the respective coiled-coil domains encoded by exon 7 [Siomi et al., 1996] and could be functionally relevant. Finally, gel filtration experiments showed that FMRP is part of a macromolecular complex with an apparent molecular weight of over 600 kDa, which is reduced to a 100-300 kDa complex with RNAse treatment or medium salt conditions, while the FMR1 monomer is freed only with high salt concentrations [Tamanini et al., 1996].

It is possible that monomers of FMRP, FXR2P and/or other proteins associate to constitute the 100-300 kDa complex, though recent data suggest that these complexes are mainly formed by homo-multimers [Tamanini et al., 1999b]. However, the formation of the >600 kDa complex seems to require the constitution of a ribonucleoprotein complex with mRNAs and at least three novel FMRP-interacting proteins: NUFIP [Bardoni et al., 1999], CYFIP1 and CYFIP2 [Schenck et al., 2001].

In conclusion, FMR1 protein is an RNA-binding protein associated with polyribosomes, not essential for cell survival and expressed in actively dividing cells and growth-arrested neurons alike, probably required for optimal translation of a subset of mRNAs in cells with a high synthetic activity. Very recently it has been proposed that FMRP may act as a negative regulator of translation [Laggerbauer et al., 2001].

Trinucleotide amplification

In the vast majority of cases (>95%) the fragile X syndrome is caused by a unique type of "dynamic" mutation due to the expansion and hypermethylation of a potentially unstable CGG trinucleotide repeat in the 5' untranslated region (UTR) of the *FMR1* gene. Abnormal amplification of trinucleotide repeats in transcribed sequences has been shown to be responsible for an increasing number of genetic disorders [Djian, 1998]. However, both the mechanisms of repeat amplification and the timing of expansion of "dynamic" mutations are still poorly understood.

It is useful to make a first distinction between expanded repeats that lie inside the coding sequence of a gene (and are translated into amino acids) and repeats which reside in introns or in the 5' or 3' UTR [for review see Cummings and Zoghbi, 2000].

In many neurodegenerative disorders (Huntington disease, DRPLA and several spinocerebellar ataxias), an expanded CAG tract in the coding region of the gene leads to the production of proteins with a variable polyglutamine tract which is somehow toxic to the cells already when the number of glutamines doubles (up to 100 repeats from the 10-20 of the normal range). Some selection mechanism may eliminate cells with larger expansions (>200 repeats).

On the contrary, if the repeat is transcribed but not translated, the number of triplets can be as high as 1000 or more. This is the case in the fragile X syndrome (CGG repeat in the 5' UTR of the *FMR1* gene), myotonic dystrophy (CTG repeat in the 3' UTR of the *DMPK* gene) and in Friedreich ataxia (a GAA repeat in the first intron of the *FRDA* gene). Recently, a pentanucleotide repeat ATTCT has been found in intron 9 of the *SCA10* gene, which also is hyperexpanded in patients with spinocerebellar ataxia linked to chromosome 22 [Matsuura et al., 2000].

In the case of fragile X, as in every "dynamic" mutation, we generally observe a multistep process of expansion that takes place over many generations. As illustrated in Table 1, three classes of alleles are observed in the *FMR1* gene according to the length of the CGG repeat: wild-type alleles (approx. 5-50 repeats, mostly 29-30), "premutations" (approx. 50-200 repeats) and "full mutations" (200-1000 repeats and more). A fourth class of alleles, largely overlapping with the wild-type in size (approx. 40-55 repeats) and called by many 'gray zone' alleles, have been listed in Table 1 as "protomutations". Protomutations are only slightly unstable over many generations and never expand to full mutation. They are transmitted silently across generations, sometimes increasing (or decreasing) of just few repeats, creating a 'pool' of at-risk alleles in the population that must not, but can eventually become premutations.

"Premutations" are often found in carrier mothers of affected fragile X boys, who themselves harbour a "full" mutation. Carrier mothers may have a full mutation and be clinically normal or slightly affected, depending on the random X-inactivation [Willemsen et al., 2000]. Both female and male premutation carriers are normal, although higher levels of *FMR1* mRNA (up to 2-3 more times) have been detected in their cells [Tassone et al., 2000a], possibly because the mutated mRNA is translated less efficiently. However, premutations above 70 CGG repeats are very unstable and have a high risk of expansion to full mutation when transmitted by a female [Heitz et al., 1992]. Therefore, in a family with several carrier women, a sudden 'burst' of fragile X children is eventually observed: this variable penetrance of the fragile X syndrome - the so-called 'Sherman paradox' - is explained by the sequential and progressive nature of "dynamic" mutations [Fu et al., 1991].

Table 1 – Types of fragile X alleles

	approximate size (CGGs)	relative (in)stability	molecular phenotype	clinical phenotype
Wild-type	5-50	stable		normal individuals
Protomutation (gray zone)	±40-55	slightly unstable over many generations, only rare expansion to premutation range		normal individuals
Premutation	50-200	relatively stable in somatic cells; unstable when transmitted to next generation, but never expands to full mutation when transmitted by a male	increased mRNA levels	normal carriers; possibly premature menopause
Full mutation	>200 up to 1000-	somatically unstable (mosaicism) and unstable when transmitted though affected males only transmit premutations (only premutated sperm)	promoter methylation no mRNA no FMR1 protein	fragile X males slightly affected or normal carrier females

It is likely that different amplification mechanisms cause the initial limited and progressive lengthening of the repeat (transforming a wild-type into a proto- and then premutation allele) and the subsequent massive expansion which transforms a premutation into a full mutation. Full mutations thus appear to be generated by a multistep process requiring the sequential action of different mechanisms [Morton and Macpherson,1992; Kolehmainen,1994; Richards and Sutherland,1994; Chiurazzi et al.,1996]. In fact, no direct conversion of a wild-type to a full

mutation allele has ever been observed in fragile X families; all mothers of affected individuals were found to be carriers of either a pre- or a full mutation.

Initial instability and founder effects

The initial events leading to the instability of a wild-type *FMR1* allele are apparently much rarer than those determining the final transition from premutation to full mutation [Chiurazzi et al., 1996]. Slipped-strand mispairing, an intrastrand process occurring during DNA replication [Levinson and Gutman,1987; Schloetterer and Tautz,1992] or possibly repair [Sinden, 2001], is assumed to be one of the mechanisms leading to the initial expansion of trinucleotide "microsatellite" sequences. *In vitro* replication studies of expanded CTG and CGG repeats demonstrated that DNA polymerase pauses after copying 29-31 pure repeat units [Kang et al., 1995a]. This is likely to allow the formation of secondary structures on the nascent (lagging) strand, including unimolecular hairpins [Darlow and Leach, 1998], which could results in more substantial (up to 10 CGGs or so) increases in repeat length [Wells, 1996] if not correctly excised.

On the contrary, if hairpins should be formed on the template strand and incorrectly excised by the repair machinery, a smaller allele would result [Hirst and White, 1998]. Initially a gradual accumulation of one or few repeats per generation is thought to occur, depending on the total length of the CGG repeat, but also on the repeat configuration [Hirst et al., 1995]. In fact, sequencing analysis of wild-type *FMR1* alleles revealed that the CGG repeat stretch is commonly interrupted by AGG triplets, often two, occurring every 9-10 CGGs [Eichler et al., 1994; Hirst et al., 1994; Snow et al., 1994; Zhong et al., 1995; Kunst et al., 1996], which apparently have a stabilizing effect by preventing replication slippage [Heale and Petes, 1995; Weisman-Shomer et al., 2000]. Loss or lack of the most distal 3' [Kunst and Warren,1994; Eichler et al., 1994; Eichler et al., 1995; Eichler et al., 1996] or proximal 5' [Crawford et al.,2000a; Crawford et al., 2000b] AGG interruption and subsequent purity (>24 CGGs) of either end of the CGG array is an important determinant of the CGG repeat instability. However, slippage of an entire AGG(CGG)$_9$ tract caused a 10-unit increase of a normal allele with 2 AGG interruptions has also been reported [Macpherson et al.,1995].

In vitro studies using constructs introduced in bacteria [Kang et al.,1995b; Shimizu et al., 1996; Hirst and White, 1998] showed that size changes (mostly reductions) occurred in an orientation-dependent fashion.

Using yeast as eukaryotic model, White et al. [1999] were able to observe also substantial expansions (+40/+50 CGGs), again depending on the orientation: in fact they observed that it is always the CGG (not the CCG) strand which has the highest propensity for expansion (as a displaced Okazaki fragment) and contraction (as a lagging-strand template). White et al. [1999] also pinpointed one of the possible *trans*-acting factors that influence repeat (in)stability: *RAD27* (homologue to human *FEN1*) is necessary for the correct processing of Okazaki fragments and its absence strongly increased array instability. Interestingly, White et al. [1999] did not observe any increase in the frequency of large expansion events when studying various CGG arrays in *mlh*1, *msh*2, *msh*3, and *msh*6 mutant backgrounds. In fact, it has been surprisingly concluded that repeat instability requires mismatch repair proteins [Manley et al., 1999; Kovtun and McMurray, 2001].

In vivo instability of normal [Crawford et al., 2000c] and premutated [Nolin et al.,1999] *FMR1* alleles has been monitored by either small pool or single cell PCR on sperm and lymphocytes. Reductions (-10 to -20) and amplifications (+10 to + 20) have been found to have a similar frequency in normal males, both in sperm and blood at the respective rates of approximately 8 and 5×10^{-4} of detectable variants [Crawford et al., 2000c]. In the case of premutated males a much higher instability was observed: more than 80% of the alleles were unstable with a striking predominance of expansions (+10 to +30) over contractions [Nolin et al., 1999]. Mornet et al. [1996] studied a male carrying a 55 CGG repeat allele, a 'gray zone' or protomutation allele (Table 1), and found again more variability in sperm than in blood and a more than double rate of contractions (up to -30) compared to expansions (up to +10) with a frequency of approx. 20×10^{-4} and 8×10^{-4}. This observation underscores the concept that only very few unstable protomutation alleles will not revert to wild-type size but reach instead the premutation size, thus committing to becoming full mutations. These rare protomutations will be linked to a limited number of ancestral haplotypes. These haplotypes define the fragile X founder chromosomes observed in different world populations [Chiurazzi et al., 1996] despite the relatively high prevalence and low fitness of affected individuals (see Chapter 2).

Transition from pre- to full mutations

The dramatic expansion - the "big jump" - that transforms a premutation into a full mutation must be caused by mechanisms different from those considered in the previous paragraph. It is possible that when a so-called 'expansion threshold' (about 70 pure CGGs) is reached [Eichler et al., 1994], multiple hairpins and/or stem-and-loop structures form on the nascent lagging

strand, because Okazaki fragments comprised exclusively of CGG repeats may slip at both ends [Richards and Sutherland, 1994]. Such structures would be extremely unstable and, after inappropriate repair, may result in a variety of expanded full mutations frequently accompanied by contracted or even deleted alleles. In fact, fragile X patients are often mosaics for full mutations of different sizes, for full mutations and premutations [Chiurazzi et al., 1994a], or alleles of normal size [van den Ouweland et al., 1994] or deletions of the entire CGG stretch and part of its flanking sequences [de Graaff et al., 1995; Mila et al., 1996; Garcia-Arocena et al. 2000].

We now must address the question of timing: when (and where) does the pre- to full expansion take place? Let us consider a first scenario: the "big jump" is limited to gametogenesis and it cannot happen during embryogenesis, though a zygotic full mutation will increase and reduce its size during development (somatic mosaicism), also producing cell lineages with either a premutation, a wild-type allele or even a deletion [Chiurazzi et al., 1994]. In the second scenario the zygote harbours a premutation and the "big jump" is made during early embryogenesis: in this case we must explain why premutations transmitted from a carrier male to his daughters never expand to full mutations. In fact, until now, all daughters of carrier males are premutation carriers, sometimes with a smaller allele compared to their father [Rousseau et al.,1991]. This is explained by the observation that sperm of male premutation carriers, as well as that of fragile X patients [Reyniers et al.,1993], harbour only premutations. Thus, according to the first scenario, premutations never expand during embryogenesis, sperm only contains premutations and obviously all daughters of carrier (and even affected) males will be premutated [Willems et al., 1992; Smeets et al., 1995].

Otherwise, if pre- to full expansion happens postzygotically (second scenario), some kind of imprinting mechanism should tag maternal premutations only as likely candidates for expansion. Malter et al. [1997] presented evidence of a selection process going on during spermatogenesis: full mutation alleles can be detected in oocytes and in fetal spermatogonia, but only premutations seem to survive in testes at a later stage. Such selection cannot be due to the lack of the FMR1 protein, because *Fmr1* knockout mice (see below) are normally fertile [Bakker et al., 1994]. This observation again supports the first scenario, in which the pre- to full expansion is probably limited to gametogenesis and postzygotical instability generates mosaicism in the full mutation range, often accompanied by reductions in the premutation (or even normal) range [Chiurazzi et al.,1994a]. Large contractions of expanded repeats in sperm have also been reported in Friedreich ataxia [Pianese et al., 1997] and myotonic dystrophy [Ashizawa et al.,

1994], supporting the hypothesis of Nolin et al. [1999] that long repeat expansions are generally unstable over numerous mitotic divisions and tend to a dramatic size reduction. In fact, oogonia undergo only few division (approx. 20) before fertilization, while spermatogonia may divide hundreds of times before giving rise to mature sperm cells. Contraction during spermatogenesis also argues against a major contribution of mitotic events to the "big jump", though replication slippage is likely to have an important role in the initial expansion (see previous paragraph). In this light, great attention should be paid to recent evidence, obtained from animal models for Huntington and myotonic dystrophy, indicating that repeat instability occurs to a greater degree in cells that do not divide [Kovtun and McMurray, 2001; Seznec et al., 2000; Fortune et al., 2000], thus in the absence of DNA replication [Sinden, 2001]. Erroneous repair could therefore be responsible for dramatic expansions such as that reported by Kennedy and Shelbourne [2000] in the striatum of a Huntington murine model. Double strand break repair by gene conversion with several round of unwinding and re-invasion between sister chromatids has been put forward as a possible mechanism of replication-independent expansion [Richard and Paques, 2000]. Finally, sex of the offspring also influences the degree of CGG instability [Loesch et al., 1995], as it appears that repeat expansion is enhanced in mother-to-son transmissions compared with mother-to-daughter transmissions.

Inactivation of full mutations

Data of Malter et al. [1997] on a female 13-week fragile X fetus indicated that full mutations were unmethylated in oocytes, though they were completely methylated in all somatic tissues. Thus, inactivation and methylation of the *FMR1* promoter appear to take place after fertilization during early embryogenesis. The only (partial) exception to the rule occurs in chorionic villi, in which inactivation of full mutations is observed in spite of a (relative) hypomethylation of the *FMR1* promoter [Carola Bontekoe, PhD thesis]. This suggests that cytosine methylation is simply one link in the chain of events leading to inactivation, though an important one. However, in somatic tissues all cytosine residues in the repeat itself and in the upstream CpG island become completely methylated as if they were on the inactive X chromosome [Stoeger et al., 1997; Luo et al., 1993; Hornstra et al., 1993; Hansen et al., 1992].

It is possible that the abnormal structures formed by the full mutation [Darlow and Leach, 1998; Mitas, 1997; Gacy et al., 1995] or even trinucleotide repeat binding proteins [Richards et al., 1993; Deissler et al., 1996] would attract *de novo* DNA methyltransferases, that eventually methylate the whole CpG island in the effort of stopping the expansion of the CGG repeat

[Bestor and Tycko, 1996; Kho et al., 1998]. The unfortunate side effects are the transcriptional silencing of the *FMR1* gene [Pieretti et al.,1991] and the consequent absence of its protein [Verheij et al.,1993], that causes mental retardation.

How does methylation lead to transcriptional inactivation? A first pathway is an indirect one: methylcytosine binding proteins such as MeCP2 and/or other MBDs [Hendrich and Bird, 1998] are involved in recruiting multiprotein complexes to methyl CpG-enriched regions in the genome to repress transcription. Again cytosine methylation emerges as a tag for potentially dangerous (unstable) DNA sequences [Bestor and Tycko, 1996]. Several histone deacetylases belong to the multiprotein complex assembled by methylcytosine binding proteins, which remove critical acetyl groups from histone H3 and H4 tails causing a local compaction of chromatin and preventing access of transcription factors [Razin, 1998]. Such a locally inactive chromatin state is most likely responsible for both the transcriptional silencing [Pieretti et al., 1991; Sutcliffe et al., 1992; Hwu et al., 1993] and the delayed replication timing of the *FMR1* gene harbouring a full mutation [Hansen et al.,1993; Hansen et al.,1996; Samadashwily et al.,1997], which supposedly causes the fragile site [Laird et al.,1987].

However, a second more direct effect of methylation consists of preventing binding of some transcription factors, even when chromatin is locally open: Kumari and Usdin [2001] further characterized the *FMR1* promoter and the transcription factors that physiologically interact with it, showing that when the action of α-Pal/Nrf-1 is inhibited by methylation of its binding site, 70% of the promoter activity is lost. Therefore, in a direct or indirect way, CpG methylation is one of the most important factors determining transcriptional repression of local chromatin domains.

The key role of cytosine methylation is well appreciated when one considers that we obtained *in vitro* reactivation of *FMR1* expression after inducing DNA demethylation with 5-aza-2-deoxycytidine (5-azadC) in fragile X lymphoblastoid cells (Chapter 3.1 - Chiurazzi et al. [1998]). In the following set of experiments (Chapter 3.2 - Chiurazzi et al.[1999]), we observed a strong synergistic effect of 5-azadC and histone hyperacetylating drugs, such as sodium- and 4-phenyl-butyrate, in reactivating the fully mutated *FMR1* gene. These results support the hypothesis of a sequential cascade of events (first CpG island methylation, then histone deacetylation) acting as synergistic layers for the silencing of genes, although DNA methylation is dominant over histone deacetylation [Cameron et al., 1999]. In fact, histone hyperacetylating drugs alone, in the presence of methylation, were almost not capable of reactivating the fully mutated *FMR1* gene [Chiurazzi et al., 1999]. Chapter 3.3 reports on the analysis of the

methylation status of every CpG site in the *FMR1* promoter of fragile X cell lines before and after various treatments with 5-azadC for different time periods. Our results suggest that the larger the full mutation, the longer it takes to achieve demethylation of the upstream CpG island. Various reports on rare intellectually normal [Rousseau et al.,1994; Smeets et al.,1995; de Vries et al.,1996] or minimally affected [McConkie-Rosell et al.,1993; Hagerman et al.,1994; Wang et al.,1996] males with an unmethylated full mutation confirm that the abnormally amplified CGG tract *per se* can still be transcribed and translated. Although *FMR1* mRNA levels seem to be even more elevated than in premutation carriers [Tassone et al., 2000b] in an effort to compensate the less efficient translation [Feng et al.,1995], these rare male carriers with unmethylated full mutations are often intellectually normal. This means that enough, though less, FMR1 protein is made in their cells. Of course, somatic mosaicism is commonplace in fragile X patients and undetected premutation alleles instead of unmethylated full mutations may be present in the brain: to answer this question careful postmortem studies on the brains of such individuals should be done. Only one such study has been reported [Taylor et al., 1999] and confirmed by immunocytochemistry the presence of the FMR1 protein in regions of the brain with the unmethylated mutation, while it was absent only in the parietal lobe where the full mutation was methylated.

In closing this paragraph, it is interesting to remember that the *FMR2* gene, located less than a megabase distally in Xq28 [Knight et al., 1993], also becomes inactivated in the same way, leading to a less specific and milder form of mental retardation. This gene is unrelated to *FMR1*, but also contains a CCG repeat in its 5' CpG island and is associated with the folate-sensitive fragile site FRAXE. Other CGG tracts have been identified which can undergo expansion and hypermethylation and are associated with fragile sites: FRAXF [Parrish et al., 1994], FRA16A [Nancarrow et al.,1994] and FRA11B [Jones et al.,1995]. However, all these repeats reside in untranscribed regions of the genome and have not been associated with any genetic disorder when "fully" expanded and hypermethylated.

Animal models

A transgenic mouse with a fusion gene consisting of a *E.coli* beta-galactosidase reporter gene (*lacZ*) linked to the *FMR1* promoter region was established by Hergersberg et al. [1995]. The fusion gene showed an expression pattern closely resembling the endogenous one, indicating that the 2.8 kb fragment 5' of the CGG repeat, at the start of the *FMR1* gene, contains most *cis*-acting elements regulating its transcription.

Fmr1 knockout mice have been generated by homologous recombination of a targeting vector interrupting exon 5 in embryonic stem (ES) cells [Bakker et al., 1994]. It is important to note that no reduced fertility of mutants of either sex has been observed and heterozygous females had normal litter size with the expected distribution of offspring with the mutant allele. Thus, *Fmr1* is not necessary for spermatogenesis or oogenesis in mice, nor for normal embryonic development or postnatal viability. These observations are in accordance with our knowledge of human fragile X syndrome, although the discovery that sperm of affected males with only a full mutation in blood contains exclusively a premutation suggested that *FMR1* expression may be useful in male spermatogenesis [Reyniers et al., 1993]. In fact *FMR1* is actively transcribed in the continuously dividing spermatogonia [Devys et al., 1993] and, in mosaic individuals with a premutation, full mutation alleles are likely to be counterselected. *Fmr1* expression is also high in both murine fetal testes and ovary [Bachner et al., 1993] while it declines in adult gonads though it disappears only in mature ovary.

Fmr1 knockout mice show no overt anatomical or histological abnormalities but do have macroorchidism and exhibit hyperactivity and learning deficits [Bakker et al., 1994; Bakker et al., 2000]. Apparently an increased Sertoli cell proliferation during testicular development is responsible for the macroorchidism, though this increase does not appear to be the result of major changes in FSH signal transduction in knockout mice [Slegtenhorst-Eegdeman et al., 1998]. *Fxr2* knockout mice have also been realized [Bontekoe et al., 2001a] and show a behavioural phenotype overlapping in part with the *Fmr1* knockout. *Fxr1* knockout embryos do not survive [H. Siomi, personal communication], and therefore a new generation of conditional *Fxr1* knockout mice is being developed.

On the *Fmr1* knockout background, transgenic copies of *FMR1* (minigenes, cosmids and even YACs) have been introduced in the attempt of "rescuing" the abnormal phenotype. Peier et al. [2000] used an YAC containing the entire human *FMR1* gene as transgene and obtained mice expressing protein at a higher level (10 to 15 times) in a cell- and tissue-specific manner. Macro-orchidism was absent in these mice indicating functional rescue by the human protein, but behavioural abnormalities suggest that overexpression of the FMR1 protein may be just as deleterious as its absence.

Finally, in order to construct a model suited to study the CGG instability, a transgenic mouse with an 81-repeat array was first realized: the array seems stable both somatically and upon transmission [Bontekoe et al., 1997], though it contained 2 interruptions. Very recently, a knock-in mouse was obtained with a pure 98 CGG repeat inserted through homologous

recombination at the murine *Fmr1* locus: this model does show a limited instability (plus or minus few repeats) [Bontekoe et al., 2001b].

Mutations other than CGG expansion

The identification of mutations other than the expansion of the CGG repeat, although in a minority of patients, has also been important in confirming that fragile X syndrome is a single gene disorder. Thus, point mutations [De Boulle et al., 1993; Lugenbeel et al., 1995] or small intragenic deletions [Meijer et al., 1994; Hart et al., 1995] ruled out the possibility that the abnormal hypermethylation, associated with the full mutation, might not be restricted to the *FMR1* promoter but could affect the expression of other genes in that chromosomal region. The deletion described by Meijer et al. [1994], was transmitted by a male to his daughter, thus indirectly confirming that FMRP is not necessary for male fertility (although it is highly expressed in spermatogonia). Several larger deletions, even encompassing the entire *FMR1* gene [Tarleton et al., 1993], have also been reported. A review of the deletion cases has been published by Hammond et al. [1997] along with the description of an additional patient. Careful analysis of small deletions limited to the promoter region helped defining the essential regulatory sequences governing *FMR1* transcription [Gronskov et al., 1997]. Finally, we note that some rare patients with all the phenotypic manifestations of the fragile X syndrome show no detectable alteration of the *FMR1* gene, which is apparently not involved in the pathogenesis of their condition [Chiurazzi et al., 1994b]. Considering the interactions between FMRP and the FXR proteins, all present in neurons, it may be possible that these fragile X patients have a mutation in either the *FXR1* or *FXR2* gene.

Diagnosis and prevalence

Molecular diagnosis of the CGG amplification, which constitutes >95% of the fragile X mutations, is available since the cloning of the *FMR1* gene in 1991, and fundamentally relies on Southern blotting and hybridization of probes specific for the promoter region, while PCR is employed to accurately determine the length of the CGG repeat tract up to the premutation range. It is possible to perform a modified PCR across the CGG repeat but, given the low product yield for full mutations, blotting of the products and hybridization with a (CGG)n probe [Brown et al.,1993] is still required. Perkin-Elmer Biosystems has marketed a fragile X size polymorphism assay kit that can easily detect the size of normal and premutated alleles with an ABI sequencer, but its reliability for sizing full mutations remains questionable.

Screening for full and large premutations should thus combine both PCR and Southern blotting, possibly using a pooling/reanalysis strategy as in Rousseau et al. [1995]. Cytogenetic testing can still be considered when looking for full mutations in males, although positive cases should be checked with DNA analysis and may lead to the identification of few FRAXE individuals [Knight et al.,1993].

A rapid method based on antibody detection of the FMR1 protein in cells of blood smears has been described and validated by Willemsen et al. [1995], and is useful when screening affected males [Willemsen et al.,1997a]. Furthermore, the protein test has also been optimized for hair roots [Willemsen et al., 1999], thus avoiding the need for blood drawing and also allowing the analysis of the hair bulb cells which are of ectodermal origin and possibly less subject to selection effects in mosaic individuals than when blood cells are used (of mesodermal origin).

Prenatal diagnosis on amniocytes still depends on the availability of sufficient DNA to perform a Southern blotting, after double digestion including a methylation-sensitive enzyme (usually EagI or BssHII), while the sex of the fetus can be determined with a standard karyotype or by Y-specific PCR analysis. The methylation-sensitive enzyme is not necessary on CVS material because methylation is almost always absent in chorionic villi. False positives, due to sub-optimal amplification, and false negatives, due to the possible presence of 'contracted' alleles in the wild-type range, can both occur when performing PCR alone on a sample from a male fetus.

Furthermore, only direct DNA analysis after digestion with methylation-sensitive enzyme can demonstrate the actual methylation status of the *FMR1* CpG island, especially in the presence of a full mutation. Given the occurrence of unmethylated full mutations in unaffected transmitting males [Rousseau et al.,1994; Smeets et al.,1995; de Vries et al.,1996] and the evidence that in extraembryonic tissues, such as chorionic villi, a full mutation may remain largely undermethylated until 10-11 weeks of gestation [Sutcliffe et al.,1992; Devys et al.,1992; Luo et al.,1993; Suzumori et al.,1993; Iida et al.,1994; Castellvi-Bel et al.,1995], CVS might not display the hypermethylation already present in the embryonic tissues and may need confirmation with amniocentesis. Detection of the FMR1 protein is also possible in amniocytes [Willemsen et al.,1997b] and chorionic villi [Willemsen et al.,1996b], but given the statistical interpretation of the results required and, if villi are tested, the possibility of false negative results due to the possible hypomethylation of the CpG island, the role of a protein test can only be viewed as confirmatory of the DNA analysis.

Though the fragile X syndrome is still believed to account for the majority of XLMR cases, it seems that its prevalence is not as high as initially estimated. A recent re-evaluation of the same population which yielded the much quoted figure of 1:1300 males [Webb et al.,1986] led to the conclusion that 1:4000 males is probably a more realistic figure [Turner et al.,1996]. This apparent discrepancy can be explained by the use in the re-study of molecular diagnostic tools, which are more accurate and specific than the cytogenetic test previously available. A survey in the Netherlands suggested the even lower figure of 1:6000 males [de Vries et al., 1997]. No general population screening has been done on unselected population samples as newborns, but surveys have concentrated on children with mental retardation or learning disabilities or institutionalized patients. Some data on the prevalence of healthy female carriers have been provided by a French-Canadian study which, screening 10,624 unselected women with the reliable molecular test, found 41 (1:259) carriers of *FMR1* premutated alleles with 55-101 CGG repeats [Rousseau et al.,1995]. Further similar studies are needed to establish whether this unexpectedly high prevalence of premutation carriers is unique to the specific population studied or it applies to other populations as well, as it seems more likely [Sherman et al.,1995]. Evidence that expansion to full mutation upon transmission from a premutated mother is more likely in male than female fetuses has been provided by Loesch and co-workers [Loesch et al.,1995] and may explain a relative lack of premutated males in the general population [Rousseau et al.,1996]. Large population studies on unselected series of newborns would be useful to settle the question of the true prevalence of affected (fully mutated) and normal transmitting (premutated) males, and of full mutation and premutation carrier females.

Though very few fragile X cases have been reported without amplification of the CGG repeat and with either point mutations or deletions in other parts of the *FMR1* gene [Gronskov et al.,1998], it is worth considering that the prevalence of these "non-dynamic" mutations might be underestimated because most molecular diagnostic strategies only check the status of the CGG repeat and its flanking sequences.

Treatment

Useful guidelines for health supervision of fragile X children have been published by the American Academy of Pediatrics [1996] and include advice for both physical and behavioural components of the syndrome. After confirmation of the diagnosis with the molecular test and appropriate genetic counselling to the parents for subsequent pregnancies, a series of medical examinations can be envisaged depending on the age of the child. Development during the first

year of life may be normal, though hypotonia and irritability may be apparent, but in early childhood it is important to have an ophtalmologic examination (strabismus, myopia), to perform an echocardiogram if a murmur or click is present (mitral valve prolapse) and to check for orthopedic problems (flat feet, scoliosis and loose joints). Inguinal hernia should also be excluded. History for seizures or staring episodes should be reviewed and electroencephalogram can be obtained if appropriate, though antiepileptic medication after a single seizure is not advisable, given the self-limiting course of epileptic manifestations in adolescence [Musumeci et al.,1999]. It is also important to check for hyperactive behaviour (head banging, hand biting, etc.) and severe attention deficit, which are a major concern in the school age and can be treated also pharmacologically [Hagerman,1997].

Presently, a large trial on the efficacy of L-acetylcarnitine in diminishing hyperactive behaviour and improving attention span is underway, following the pilot study of Torrioli et al. [1999]. However, socialization and school integration, possibly within a mainstream program with individual support, are extremely important to help overcome the relative shyness. A visual presentation of information is preferred in order to maximize attention. Sport and regular physical activity (e.g. swimming) are important for counteracting the hypotonic posture and improve motor coordination. Macroorchidism becomes apparent in late childhood and is not a sign of precocious puberty. Speech, language and occupational therapy should be goal-oriented and help the adolescent and young adults to attain as much autonomy as possible. Support from family organizations is extremely important especially for the parents and sibs, because it eases the sense of isolation and helplessness that often follows diagnosis. Treatments specifically aimed at recovering the function of the *FMR1* gene have been attempted with folic acid because of its action on the cytogenetic expression of fragile site, and although a few reports indicated some effect on the behaviour [Wells and Madison,1986; Gillberg et al.,1986; Hagerman et al.,1986], others did not confirm these observations [Froster-Iskenius et al.,1986; Madison et al.,1986; Webb et al.,1990]. It appears that, folate supplementation, although reducing the levels of cytogenetic expression of the fragile site, is not able to reactivate the *FMR1* gene.

References

Abitbol M, Menini C, Delezoide AL, Rhyner T, Vekemans M, Mallet J (1993): Nucleus basalis magnocellularis and hippocampus are the major sites of FMR1 expression in the human fetal brain. Nat Genet 4:147-153

American Academy of Pediatrics, Committee on Genetics (1996): Health supervision for children with fragile X syndrome. Pediatrics 98:297-300

Ashizawa T, Dunne PW, Ward MS, Seltzer WK, Richards CS (1994): Effects of the sex of myotonic dystrophy patients on the unstable triplet repeat in their affected offspring. Neurology 44:120-122.

Ashley CT, Sutcliffe JS, Kunst CB, Leiner HA, Eichler EE, Nelson DL, Warren ST (1993a): Human and murine FMR1: alternative splicing and translational initiation downstream of the CGG repeat. Nat Genet 4:244-251

Ashley CT, Wilkinson KD, Reines D, Warren ST (1993b): FMR1 protein: conserved RNP family domains and selective RNA binding. Science 262:563-566

Bachner D, Manca A, Steinbach P, Woehrle D, Just W, Vogel W, Hameister H, Poustka A (1993): Enhanced expression of the murine FMR1 gene during germ cell proliferation suggests a special function in both the male and female gonad. Hum Mol Genet 2:2043-2050

Bardoni, B., A. Schenck and J. L. Mandel (1999): A novel RNA-binding nuclear protein that interacts with the fragile X mental retardation (FMR1) protein. Hum Mol Genet 8: 2557-2566.

Bakker CE, Verheij C, Willemsen R, van der Helm R, Oerlemans F, Vermey M, Bygrave A, Hoogeveen AT, Oostra BA, Reyniers E, De Boulle K, D'Hooge R, Cras P, van Velzen D, Nagels G, Martin JJ, De Deyn PP, Darby JK, Willems PJ (1994): Fmr1 knockout mice: a model to study fragile X mental retardation. Cell 78:23-33

Bakker CE, de Diego Otero Y, Bontekoe C, Raghoe P, Luteijn T, Hoogeveen AT, Oostra BA, Willemsen R (2000): Immunocytochemical and biochemical characterization of FMRP, FXR1P, and FXR2P in the mouse. Exp Cell Res 258:162-170.

Bestor TH, Tycko B (1996): Creation of genomic methylation patterns. Nat Genet 12:363-367

Bontekoe CJ, de Graaff E, Nieuwenhuizen IM, Willemsen R, Oostra BA (1997): FMR1 premutation allele (CGG)81 is stable in mice. Eur J Hum Genet 5:293-298.

Bontekoe CJM, McIlwain KL, Nieuwenhuizen IM, Yuva-Paylor LA, Nellis A, Willemsen R, Kirkpatrick L, Fang Z, Bakker CE, McAninch R, Cheng N, Merriweather M, Hoogeveen AT, Nelson D, Oostra BA, Paylor R (2001a): Fxr2 knockout mouse: a model for mental retardation. Submitted

Bontekoe CJ, Bakker CE, Nieuwenhuizen IM, van der Linde H, Lans H, de Lange D, Hirst MC, Oostra BA (2001b): Instability of a (CGG)₉₈ repeat in the Fmr1 promoter. Hum Mol Genet 10:1693-1699.

Brown WT, Houck GE, Jeziorowska A, Levinson FN, Ding X, Dobkin C, Zhong N, Henderson J, Brooks SS, Jenkins EC (1993): Rapid fragile X carrier screening and prenatal diagnosis using a nonradioactive PCR test. JAMA 270:1569-1575

Cameron EE, Bachman KE, Myohanen S, Herman JG, Baylin SB (1999): Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. Nat Genet 21:103-107.

Castellvi-Bel S, Mila M, Soler A, Carrio A, Sanchez A, Villa M, Jimenez MD, Estivill X (1995): Prenatal diagnosis of fragile X syndrome: (CGG)ₙ expansion and methylation of chorionic villus samples. Prenat Diagn 15:801-807

Chiurazzi P, Kozak L, Neri G (1994a):Unstable triplets and their mutational mechanism: Size reduction of the CGG repeat vs. germline mosaicism in the Fragile X syndrome. Am J Med Genet 51:517-521

Chiurazzi P, de Graaff E, Ng J, Verkerk AJMH, Wolfson S, Fisch JS, Kozak L, Neri G, Oostra BA (1994b): No apparent involvement of the FMR1 gene in five patients with phenotypic manifestations of the fragile X syndrome. Am J Med Genet 51:309-314

Chiurazzi P, Macpherson J, Sherman S, Neri G (1996): Editorial Comment: Significance of linkage disequilibrium between the fragile X locus and its flanking markers. Am J Med Genet 64:203-208

Chiurazzi P, Pomponi MG, Willemsen R, Oostra BA, Neri G (1998): In vitro reactivation of the FMR1 gene involved in fragile X syndrome. Hum Mol Genet 7:109-113

Chiurazzi P, Pomponi MG, Pietrobono R, Bakker CE, Neri G, Oostra BA (1999): Synergistic effect of histone hyperacetylation and DNA demethylation in the reactivation of the FMR1 gene. Hum Mol Genet 8:2317-2323.

Corbin F, BouillonM, Fortin A, Morin S, Rousseau F, Khandjian EW (1997): The fragile X mental retardation protein is associated with poly(A)+ mRNA in actively translating polyribosomes. Hum Mol Genet 6:1465-1472

Coy JF, Sedlacek Z, Baehner D, Hameister H, Joos S, Lichter P, Delius H, Poustka A (1995): Highly conserved 3'UTR and expression pattern of FXR1 points to a divergent gene regulation of FXR1 and FMR1. Hum Mol Genet 4:2209-2218

Crawford DC, Schwartz CE, Meadows KL, Newman JL, Taft LF, Gunter C, Brown WT, Carpenter NJ, Howard-Peebles PN, Monaghan KG, Nolin SL, Reiss AL, Feldman GL, Rohlfs EM, Warren ST, Sherman SL (2000a): Survey of the fragile X syndrome CGG repeat and the short-tandem-repeat and single-nucleotide-polymorphism haplotypes in an African American population. Am J Hum Genet 66:480-493.

Crawford DC, Zhang F, Wilson B, Warren ST, Sherman SL (2000b): Fragile X CGG repeat structures among African-Americans: identification of a novel factor responsible for repeat instability. Hum Mol Genet 9:1759-1769.

Crawford DC, Wilson B, Sherman SL (2000c): Factors involved in the initial mutation of the fragile X CGG repeat as determined by sperm small pool PCR. Hum Mol Genet 9:2909-2918.

Cummings CJ, Zoghbi HY (2000): Fourteen and counting: unraveling trinucleotide repeat diseases. Hum Mol Genet 9:909-916.

Darlow JM, Leach DRF (1998): Secondary Structures in d(CGG) and d(CCG) Repeat Tracts. J Mol Biol 275:3-16

De Boulle K, Verkerk AJMH, Reyniers E, Vits L, Hendrickx J, Van Roy B, van den Bos F, de Graaff E, Oostra BA, Willems PJ (1993): A point mutation in the FMR1 gene associated with fragile X mental retardation. Nat Genet 3:31-35

de Graaff E, Rouillard P, Willems PJ, Smits APT, Rousseau F, Oostra BA (1995): Hotspot for deletions in the CGG repeat region of FMR1 in fragile X patients. Hum Mol Genet 4:45-49

Deissler H, Behn-Krappa A, Doerfler W (1996): Purification of nuclear proteins from human HeLa cells that bind specifically to the unstable tandem repeat $(CGG)_n$ in the human FMR1 gene. J Biol Chem 271:4327-4334

de Vries BB, Fryns JP, Butler MG, Canziani F, Wesby-van-Swaay E, van-Hemel JO, Oostra BA, Halley DJ, Niermeijer MF (1993): Clinical and molecular studies in fragile X patients with a Prader-Willi-like phenotype. J Med Genet 30:761-766

de Vries BBA, Jansen CAM, Duits AA, Verheij C, Willemsen R, van Hemel JO, van den Ouweland AMW, Niermeijer MF, Oostra BA, Halley DJJ (1996): Variable FMR1 gene methylation leads to variable phenotype in 3 males from one fragile X family. J Med Genet 33:1007-1010

de Vries BB, van den Ouweland AM, Mohkamsing S, Duivenvoorden HJ, Mol E, Gelsema K, van Rijn M, Halley DJ, Sandkuijl LA, Oostra BA, Tibben A, Niermeijer MF (1997): Screening and diagnosis for the fragile X syndrome among the mentally retarded: an epidemiological and psychological survey. Collaborative Fragile X Study Group. Am J Hum Genet 61:660-667

Devys D, Biancalana V, Rousseau F, Boue' J, Mandel JL, Oberle' I (1992): Analysis of full fragile X mutations in fetal tissues and monozygotic twins indicate that abnormal methylation and somatic heterogeneity are established early in development. Am J Med Genet 43:208-216

Devys D, Lutz Y, Rouyer N, Bellocq JP, Mandel JL (1993): The FMR1 protein is cytoplasmic, most abundant in neurons and appears normal in carriers of a fragile X premutation. Nat Genet 4:335-340

Djian P (1998): Evolution of simple repeats in DNA and their relation to human disease. Cell 94:155-160

Eichler EE, Richards S, Gibbs RA, Nelson DL (1993): Fine structure of the human FMR1 gene. Hum Mol Genet 2:1147-1153

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 instability in the FMR1 gene. Nat Genet 8:88-94

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 interruption. Hum Mol Genet 4:2199-2208

Eichler EE, Macpherson JN, Murray A, Jacobs PA, Chakravarti A, Nelson DL (1996): Haplotype and interspersion analysis of the FMR1 CGG repeat identifies two different mutational pathways for the origin of the fragile X syndrome. Hum Mol Genet 5:319-330

Feng Y, Zhang F, Lokey LK, Chastain JL, Lakkis L, Eberhart D, Warren ST (1995): Translational suppression by trinucleotide repeat expansion at FMR1. Science 268:731-734

Feng Y, Absher D, Eberhart DE, Brown V, Malter HE, Warren ST (1997): FMRP associates with polyribosomes as an mRNP and the I304N mutation of severe fragile X syndrome abolishes this association. Mol Cell 1:109-118

Fortune MT, Vassilopoulos C, Coolbaugh MI, Siciliano MJ, Monckton DG (2000): Dramatic, expansion-biased, age-dependent, tissue-specific somatic mosaicism in a transgenic mouse model of triplet repeat instability. Hum Mol Genet 9:439-445.

Fridell RA, Benson RE, Hua J, Bogerd HP, Cullen BR (1996): A nuclear role for the fragile X mental retardation protein. EMBO J 15:5408-5414

Froster-Iskenius U, Bodeker K, Oepen T, Matthes R, Piper U, Schwinger E (1986):Folic acid treatment in males and females with fragile-(X)-syndrome. Am J Med Genet 23:273-289

Fu YH, Kuhl DP, Pizzuti A, Pieretti M, Sutcliffe JS, Richards S, Verkerk AJ, Holden JJ, 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.

Gacy AM, Goellner G, Juranic N, Macura S, McMurray CT (1995): Trinucleotide repeats that expand in in human disease form hairpin structures in vitro. Cell 81:533-540

Garcia Arocena D, de Diego Y, Oostra BA, Willemsen R, Mirta Rodriguez M (2000): A fragile X case with an amplification/deletion mosaic pattern. Hum Genet 106:366-369.

Gillberg C, Wahlstrom J, Johansson R, Tornblom M, Albertsson-Wikland K (1986): Folic acid as an adjunct in the treatment of children with the autism fragile-X syndrome (AFRAX). Dev Med Child Neurol 28:624-627

Gronskov K, Hjalgrim H, Bjerager MO, Brondum-Nielsen K (1997): Deletion of all CGG repeats plus flanking sequence in FMR1 does not abolish gene expression. Am J Hum Genet 61:961-967

Gronskov K, Hallberg A, Brondum-Nielsen K (1998): Mutational analysis of the FMR1 gene in 118 mentally retarded males suspected of fragile X syndrome: absence of prevalent mutations. Hum Genet 102:440-445

Hagerman RJ, Jackson AW, Levitas A, Braden M, McBogg P, Kemper M, McGavran L, Berry R, Matus I, Hagerman PJ (1986):Oral folic acid versus placebo in the treatment of males with the fragile X syndrome. Am J Med Genet 23:241-262

Hagerman RJ, Hull CE, Safanda JF, Carpenter I, Staley LW, O'Connor RA, Seydel C, Mazzocco M, Snow K, Thibodeau SN, Kuhl D, Nelson DL, Caskey CT, Taylor AK (1994): High functioning fragile X males: demonstration of an unmethylated fully expanded FMR1 mutation associated with protein expression. Am J Med Genet 51:298-308

Hagerman R (1997): Fragile X: treatment of hyperactivity. Pediatrics 99:753.

Hammond LS, Macias MM, Tarleton JC, Shashidhar Pai G(1997):Fragile X syndrome and deletions in FMR1: new case and review of the literature. Am J Med Genet 72:430-434

Hansen RS, Gartler SM, Scott CR, Chen SH, Laird CD (1992): Methylation analysis of CGG sites in the CpG island of the human FMR1 gene. Hum Mol Genet 1:571-578

Hansen RS, Canfield TK, Lamb MM, Gartler SM, Laird CD (1993): Association of fragile X syndrome with delayed replication of the FMR1 gene. Cell 73:1403-1409

Hansen RS, Canfield TK, Fjeld AD, Gartler SM (1996): Role of late replication timing in the silencing of X-linked genes. Hum Mol Genet 5:1345-1353

Hart PS, Olson SM, Crandall K, Tarleton J (1995): Fragile X syndrome resulting from a 400 bp deletion within the FMR1 gene. Am J Hum Genet 57:A1395

Heale SM, Petes TD (1995): The stabilization of repetitive tracts of DNA by variant repeats requires a functional DNA mismatch repair system. Cell 83:539-545

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.

Hendrich B, Bird A (1998): Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol Cell Biol 18:6538-6547.

Hergersberg M, Matsuo K, Gassmann M, Schaffner W, Luescher B, Ruelicke T, Aguzzi A (1995): Tissue-specific expression of a FMR1/betagalactosidase fusion gene in transgenic mice. Hum Molec Genet 4:359-366

Hinds HL, Ashley CT, Sutcliffe JS, Nelson DL, Warren ST, Housman DE, Schalling M (1993): Tissue specific expression of FMR1 provides evidence for a functional role in fragile X syndrome. Nat Genet 3:36-43

Hinton VJ, Brown WT, Wisniewski K, Rudelli RD (1991) :Analysis of neocortex in three males with the fragile X syndrome. Am J Med Genet 41:289-294

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

Hirst MC (1995): FMR1 triplet arrays: paying the price for perfection. J Med Genet 32:761-763

Hirst MC, White PJ (1998): Cloned human FMR1 trinucleotide repeats exhibit a length- and orientation-dependent instability suggestive of in vivo lagging strand secondary structure. Nucl Acid Res 26:2353-2358

Hornstra IK, Nelson DL, Warren ST, Yang TP (1993): High resolution methylation analysis of the FMR1 gene trinucleotide repeat region in fragile X syndrome. Hum Mol Genet 2:1659-1665

Hwu WL, Lee YM, Lee SC, Wang TR (1993): In vitro DNA methylation inhibits FMR1 promoter. Biochem Biophys Res Comm 193:324-329

Iida T, Nakahori Y, Tsutsumi O, Taketani Y, Nakagome Y (1994): The CpG island of the FMR1 gene is methylated differently among embryonic tissues: implication for prenatal diagnosis. Hum Reprod 9:1471-1473

Jacky PB, Ahuja YR, Anyane-Yeboa K, Breg WR, Carpenter NJ, Froster-Iskenius UG, Fryns JP, Glover TW, Gustavson KH, Hoegerman SF, Holmgren G, Howard-Peebles PN, Jenkins EC, Krawczun MS, Neri G, Pettigrew A, Schaap T, Schonberg SA, Shapiro LR, Spinner N, Steinbach P, Vianna-Morgante AM, Watson MS, Wilmot PL (1991): Guidelines for the preparation and analysis of the Fragile X chromosome in lymphocytes. Am J Med Genet 38:400-403

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

Kang SM, Ohshima K, Shimizu M, Amirhaeri S, Wells RD (1995a): Pausing of DNA synthesis in vitro at specific loci in CTG and CGG triplet repeats from human hereditary disease genes. J Biol Chem 270:27014-27021

Kang S, Jaworski A, Ohshima K, Wells RD (1995b): Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in E.coli. Nat Genet 10:213-218

Kennedy L, Shelbourne PF (2000): Dramatic mutation instability in HD mouse striatum: does polyglutamine load contribute to cell-specific vulnerability in Huntington's disease? Hum Mol Genet 9:2539-2544.

Khandjian EW, Corbin F, Woerly S, Rousseau F (1996): The fragile X mental retardation protein is associated with ribosomes. Nat Genet 12:91-93

Kho MR, Baker DJ, Laayoun A, Smith SS (1998): Stalling of human DNA (cytosine-5) methyltransferase at single-strand conformers from a site of dynamic mutation. J Mol Biol 275:67-79

Knight SJL, Flannery AV, Hirst MC, Campbell L, Christodoulou Z, Phelps SR, Pointon J, Middleton-Price HR, Barnicoat A, Pembrey ME, Holland J, Oostra BA, Bobrow M, Davies KE (1993): Trinucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation. Cell 74:127-134

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

Kooy RF, Willemsen R, Oostra BA (2000): Fragile X syndrome at the turn of the century. Mol Med Today 6:193-198.

Kovtun IV, McMurray CT (2001): Trinucleotide expansion in haploid germ cells by gap repair. Nat Genet 27:407-411.

Kumari D, Usdin K (2001): Interaction of the transcription factors USF1, USF2, and alpha-Pal/Nrf-1 with the FMR1 promoter. Implications for Fragile X mental retardation syndrome. J Biol Chem 276:4357-4364.

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

Kunst CB, Zerylnick C, Karickhoff L, Eichler EE, Bullard J, Maryse C, Holden JJJ, Torroni A, Nelson DL, Warren ST (1996): FMR1 in global populations. Am J Hum Genet 58:513-522

Laggerbauer B, Ostareck D, Keidel EM, Ostareck-Lederer A, Fischer U (2001): Evidence that fragile X mental retardation protein is a negative regulator of translation. Hum Mol Genet 10:329-338.

Laird C, Jaffe E, Karpen G, Lamb M, Nelson R (1987): Fragile sites in human chromosomes as regions of late-replicating DNA. Trends Genet 3:274-281

Levinson G, Gutman GA (1987): Slipped strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol 4:203-221

Loesch DZ, Huggins R, Petrovio L, Slater H (1995): Expansion of the CGG repeat in fragile X in the FMR1 gene depends on the sex of the offspring. Am J Hum Genet 57:1408-1413

Lubs HA (1969): A marker X-chromosome. Am J Hum Genet 21:231-244

Lugenbeel KA, Peier AM, Carson NL, Chudley AE, Nelson DL (1995): Intragenic loss of function mutations demonstrate the primary role of FMR1 in fragile X syndrome. Nat Genet 10:483-485

Luo S, Courtland Robinson J, Reiss AL, Migeon BR (1993): DNA methylation of the fragile X locus in somatic and germ cells during fetal development: relevance to the fragile X syndrome and X inactivation. Somat Cell Mol Genet 19:393-404

Macpherson JN, Curtis G, Crolla JA, Dennis N, Migeon B, Grewal PK, Hirst MC, Davies KE, Jacobs PA (1995): Unusual (CGG)n expansion and recombination in a family with fragile X and DiGeorge syndrome. J Med Genet 32:236-239

Madison LS, Wells TE, Fristo TE, Benesch CG (1986): A controlled study of folic acid treatment in three fragile X syndrome males. J Dev Behav Pediatr 7:253-256

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

Mandel JL (1994): Trinucleotide diseases on the rise. Nat Genet 7:453-455

Manley K, Shirley TL, Flaherty L, Messer A (1999): Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice. Nat Genet 23:471-473.

Martin JP, Bell J (1943): A pedigree of mental defect showing sex linkage. J Neurol Neurosurg Psychiatry 6:154-157

Matsuura T, Yamagata T, Burgess DL, Rasmussen A, Grewal RP, Watase K, Khajavi M, McCall AE, Davis CF, Zu L, Achari M, Pulst SM, Alonso E, Noebels JL, Nelson DL, Zoghbi HY, Ashizawa T (2000): Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. Nat Genet 26:191-194.

McConkie-Rosell A, Lachiewicz A, Spiridigliozzi GA, Tarleton J, Shoenwald S, Phelan MC, Goonewardena P, Ding X, Brown WT (1993): Evidence that methylation of the FMR1 locus is responsible for variant phenotypic expression of the fragile X syndrome. Am J Hum Genet 53:800-809

Meijer H, de Graaff E, Merckx DML, Jongbloed RJE, de Die-Smulders CEM, Engelen JJM, Fryns JP, Curfs PMG, Oostra BA (1994): A deletion of 1.6 kb proximal to the CGG repeat of the FMR1 gene causes the clinical phenotype of the fragile X syndrome. Hum Mol Genet 3:615-620

Mila M, Castellvi-Bel S, Sanchez A, Lazaro C, Villa M, Estivill X (1996): Mosaicism for the fragile X syndrome full mutation and deletions within the CGG repeat of the FMR1 gene. J Med Genet 33:338-340

Mitas M (1997): Trinucleotide repeats associated with human disease. Nucleic Acids Res 25:2245-2254

Mornet E, Chateau C, Hirst MC, Thepot F, Taillandier A, Cibois O, Serre JL (1996): Analysis of germline variation at the FMR1 CGG repeat shows variation in the normal-premutated borderline range. Hum Mol Genet 5:821-825

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

Musumeci SA, Hagerman RJ, Ferri R, Bosco P, Dalla Bernardina B, Tassinari CA, De Sarro GB, Elia M (1999): Epilepsy and EEG findings in males with fragile X syndrome. Epilepsia 40:1092-1099.

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

Nolin SL, Houck GE Jr, Gargano AD, Blumstein H, Dobkin CS, Brown WT (1999): FMR1 CGG-repeat instability in single sperm and lymphocytes of fragile-X premutation males. Am J Hum Genet 65:680-688.

Parrish JE, Oostra BA, Verkerk AJ, Richards CS, Reynolds J, Spikes AS, Shaffer LG, Nelson DL (1994): Isolation of a GCC repeat showing expansion in FRAXF, a fragile site distal to FRAXA and FRAXE. Nat Genet 8:229-235

Peier AM, McIlwain KL, Kenneson A, Warren ST, Paylor R, Nelson DL (2000): (Over)correction of FMR1 deficiency with YAC transgenics: behavioral and physical features. Hum Mol Genet 9:1145-1159.

Pianese L, Cavalcanti F, De Michele G, Filla A, Campanella G, Calabrese O, Castaldo I, et al (1997) The effect of parental gender on the GAA dynamic mutation in the FRDA gene. Am J Hum Genet 60:460-463.

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

Razin A (1998): CpG methylation, chromatin structure and gene silencing-a three-way connection. EMBO J 17:4905-4908

Reiss AL, Aylward E, Freund LS, Joshi PK, Bryan RN (1991a): Neurosnatomy of fragile X syndrome: the posterior fossa. Ann Neurol 29:26-32

Reiss AL, Freund L, Tseng JE, Joshi PK (1991b): Neuroanatomy in fragile X females: the posterior fossa. Am J Hum Genet 49:279-288

Reiss AL, Lee J, Freund L (1994): Neuroanatomy of fragile X syndrome: The temporal lobe. Neurology 44:1317-1324

Reiss AL, Abrams MT, Greenlaw R, Freund L, Denckla MB (1995): Neurodevelopmental effects of the FMR1 full mutation in humans. Nat Medic 1:159-167

Reyniers E, Vits L, De Boulle K, Van Roy B, van Velzen D, de Graaff E, Verkerk AJMH, Jorens HZ, Darby JK, Oostra BA, Willems PJ (1993): The full mutation in the FMR1 gene of male fragile X patients is absent in their sperm. Nat Genet 4:143-146

Richard GF, Paques F (2000): Mini- and microsatellite expansions: the recombination connection. EMBO Rep 1:122-126.

Richards RI, Holman K, Yu S, Sutherland GR (1993): Fragile X syndrome unstable element, p(CCG)$_n$, and other simple tandem repeat sequences are binding sites for specific nuclear proteins. Hum Mol Genet 2:1429-1435

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

Rousseau F, Heitz D, Biancalana V, Blumenfeld S, Kretz C, Boué J, Tommerup N, Van der Hagen C, DeLozier-Blanchet C, Croquette MF, Gilgenkrantz S, Jalbert P, Voelckel MA, Oberlé I, Mandel JL (1991): Direct diagnosis by DNA analysis of the fragile X syndrome of mental retardation. N Eng J Med 325:1673-1681

Rousseau F, Robb LJ, Rouillard P, Der Kaloustian VM (1994): No mental retardation in a man with 40% abnormal methylation at the FMR1 locus and transmission of sperm cell mutations as premutations. Hum Mol Genet 3:927-930

Rousseau F, Rouillard P, Morel ML, Khandjian EW, Morgan K (1995): Prevalence of carriers of premutation-size alleles of the FMR1 gene and implications for the population genetics of the fragile X syndrome. Am J Hum Genet 57:1006-1018

Rousseau F, Morel ML, Rouillard P, Khandjian EW, Morgan K (1996): Surprisingly low prevalence of the FMR1 premutations among males from the general population. Am J Hum Genet 59:A188

Samadashwily GM, Raca G, Mirkin SM (1997): Trinucleotide repeats affect DNA replication in vivo. Nat Genet 17:298-304

Sanfilippo S, Ragusa R, Musumeci S, Neri G (1986): Fragile X mental retardation: prevalence in a group of institutionalized patients in Italy and description of a novel EEG pattern. Am J Med Genet 23:589-595

Schenck A, Bardoni B, Moro A, Bagni C, Mandel JL (2001): A highly conserved protein family interacting with the fragile X mental retardation protein (FMRP) and displaying selective interactions with FMRP-related proteins FXR1P and FXR2P. Proc Natl Acad Sci USA, in press

Schloetterer C, Tautz D (1992): Slippage synthesis of simple sequence DNA. Nucl Ac Res 20:211-215

Seznec H, Lia-Baldini AS, Duros C, Fouquet C, Lacroix C, Hofmann-Radvanyi H, Junien C, Gourdon G (2000): Transgenic mice carrying large human genomic sequences with expanded CTG repeat mimic closely the DM CTG repeat intergenerational and somatic instability. Hum Mol Genet 9:1185-1194.

Sherman SL (1995): The high prevalence of fragile X premutation carrier females: is this frequency unique to the French Canadian population? Am J Hum Genet 57:991-993

Shimizu M, Gellibolian R, Oostra BA, Wells RD (1996): Cloning and characterization, and properties of plasmids containing CGG triplet repeats from the FMR1 gene. J Mol Biol 258:614-626

Sinden RR (2001): Origins of instability. Nature 411:757-758.

Siomi H, Siomi MC, Nussbaum RL, Dreyfuss G (1993): The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein. Cell 74:291-298

Siomi H, Choi M, Siomi MC, Nussbaum RL, Dreyfuss G (1994): Essential role for KH domains in RNA binding: impaired RNA binding by a mutation in the KH domain of FMR1 that causes fragile X syndrome. Cell 77:33-39

Siomi MC, Siomi H, Sauer WH, Srinivasan S, Nussbaum RL, Dreyfuss G (1995): FXR1, an autosomal homolog of the fragile X mental retardation gene. EMBO J 14:2401-2408

Siomi MC, Zhang Y, Siomi H, Dreyfuss G (1996): Specific sequences in the fragile X syndrome protein FMR1 and the FXR proteins mediate their binding to 60S ribosomal subunits and the interactions among them. Mol Cell Biol 16:3825-3832

Sittler A, Devys D, Weber C, Mandel JL (1996): Alternative splicing of exon 14 determines nuclear or cytoplasmic localisation of FMR1 protein isoforms. Hum Molec Genet 5:95-102

Slegtenhorst-Eegdeman KE, de Rooij DG, Verhoef-Post M, van de Kant HJ, Bakker CE, Oostra BA, Grootegoed JA, Themmen AP (1998): Macroorchidism in FMR1 knockout mice is caused by increased Sertoli cell proliferation during testicular development. Endocrinology 139:156-162

Smeets HJM, Smits APT, Verheij C, Theelen JPG, van de Burgt I, Hoogeveen AT, Oosterwijk JC, Oostra BA (1995): Normal phenotype in two brothers with a full FMR1 mutation. Hum Mol Genet 4:2103-2108

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

Stoger R, Kajimura TM, Brown WT, Laird CD (1997): Epigenetic variation illustrated by DNA methylation patterns of the fragile-X gene FMR1. Hum Mol Genet 6:1791-1801

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

Suzumori K, Yamauchi M, Seki N, Kondo I, Hori T (1993): Prenatal diagnosis of a hypermethylated full fragile X mutation in chorionic villi of a male fetus. J Med Genet 30:785-787

Tamanini F, Meijer N, Verheij C, Willems PJ, Galjaard H, Oostra BA, Hoogeveen AT (1996): FMRP is associated to the ribosomes via RNA. Hum Mol Genet 5:809-813

Tamanini F, Willemsen R, van Unen L, Bontekoe C, Galjaard H, Oostra BA, Hoogeveen A (1997): Differential expression of FMR1, FXR1 and FXR2 proteins in human brain and testis. Hum Mol Genet 6:1315-1322.

Tamanini F, Bontekoe C, Bakker CE, van Unen L, Anar B, Willemsen R, Yoshida M, Galjaard H, Oostra BA, Hoogeveen AT (1999a): Different targets for the fragile X-related proteins revealed by their distinct nuclear localizations. Hum Mol Genet 8:863-869.

Tamanini F, van Unen L, Bakker C, Sacchi N, Galjaard H, Oostra BA, Hoogeveen AT (1999b): Oligomerization properties of fragile-X mental-retardation protein (FMRP) and the fragile-X-related proteins FXR1P and FXR2P. Biochem J 343:517-523.

Tarleton J, Richie R, Schwartz C, Rao K, Aylsworth AS, Lachiewicz A (1993): An extensive de novo deletion removing FMR1 in a patient with mental retardation and the fragile X phenotype. Hum Mol Genet 2:1973-1974

Tassone F, Hagerman RJ, Taylor AK, Gane LW, Godfrey TE, Hagerman PJ (2000a): Elevated levels of FMR1 mRNA in carrier males: a new mechanism of involvement in the fragile-X syndrome. Am J Hum Genet 66:6-15.

Tassone F, Hagerman RJ, Loesch DZ, Lachiewicz A, Taylor AK, Hagerman PJ (2000b): Fragile X males with unmethylated, full mutation trinucleotide repeat expansions have elevated levels of FMR1 messenger RNA. Am J Med Genet 94:232-236.

Taylor AK, Tassone F, Dyer PN, Hersch SM, Harris JB, Greenough WT, Hagerman RJ (1999): Tissue heterogeneity of the FMR1 mutation in a high-functioning male with fragile X syndrome. Am J Med Genet 84:233-239.

Torrioli MG, Vernacotola S, Mariotti P, Bianchi E, Calvani M, De Gaetano A, Chiurazzi P, Neri G (1999): Double-blind, placebo-controlled study of L-acetylcarnitine for the treatment of hyperactive behavior in fragile X syndrome. Am J Med Genet 87:366-368.

Turner G, Webb T, Wake S, Robinson H (1996): Prevalence of fragile X syndrome. Am J Med Genet 64:196-197

van den Ouweland AMW, De Vries BBA, Bakker PLG, Deelen WH, de Graaff E, van Hemel JO,Oostra BA, Niermeijer MF, Halley DJJ (1994): DNA diagnosis of the fragile X syndrome in a series of 236 mentally retarded subjects and evidence for a reversal of mutation in the FMR1 gene. Am J Med Genet 51:482-485

Verheij C, Bakker CE, de Graaff E, Keulemans J, Willemsen R, Verkerk AJMH, Galjaard H, Reuser AJJ, Hoogeveen AT, Oostra BA (1993): Characterization and localization of the FMR1 gene product associated with fragile X syndrome. Nature 363:722-724

Verheij C, de Graaff E, Bakker CE, Willemsen R, Willems PJ, Meijer N, Galjaard H, Reuser AJJ, Oostra BA, Hoogeveen AT (1995): Characterization of FMR1 proteins isolated from different tissues. Hum Mol Genet 4:895-901

Verkerk AJMH, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DPA, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang F, Eussen BE, van Ommen GJB, Blonden LAJ, Riggins GJ, Chastain JL, Kunst CB, Galjaard H, Caskey CT, Nelson DL DL, Oostra BA, Warren ST (1991): Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting lenght variation in fragile X syndrome. Cell 65:905-914

Verkerk AJMH, de Graaff E, De Boulle K, Eichler EE, Konecki DS, Reyniers E, Manca A, Poustka A, Willems PJ, Nelson DL, Oostra BA (1993): Alternative splicing in the fragile X gene FMR1. Hum Mol Genet 2:399-404

Wang Z, Taylor AK, Bridge JA (1996): FMR1 fully expanded mutation with minimal methylation in a high functioning fragile X male. J Med Genet 33:376-378

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

Webb T, Crawley P, Bundey S (1990): Folate treatment of a boy with fragile-X syndrome. J Ment Defic Res 34:67-73

Weiler IJ, Irwin SA, Klintsova AY, Spencer CM, Brazelton AD, Miyashiro K, Comery TA, Patel B, Eberwine J, Greenough WT (1997): Fragile X mental retardation protein is translated near synapses in response to neurotransmitter activation. Proc Natl Acad Sci USA 94:5395-5400

Weisman-Shomer P, Cohen E, Fry M (2000): Interruption of the fragile X syndrome expanded sequence d(CGG)(n) by interspersed d(AGG) trinucleotides diminishes the formation and stability of d(CGG)(n) tetrahelical structures. Nucleic Acids Res 28:1535-1541.

Wells RD (1996): Molecular basis of genetic instability of triplet repeats. J Biol Chem 271:2875-2878

Wells TE, Madison LS (1986): Assessment of behavior change in a fragile-X syndrome male treated with folic acid. Am J Med Genet 23:291-296

White PJ, Borts RH, Hirst MC (1999): Stability of the human fragile X (CGG)(n) triplet repeat array in Saccharomyces cerevisiae deficient in aspects of DNA metabolism. Mol Cell Biol 19:5675-5684.

Willems PJ, Van Roy B, De Boulle K, Vits L, Reyniers E, Beck O, Dumon JE, Verkerk A, Oostra B (1992): Segregation of the fragile X mutation from an affected male to his normal daughter. Hum Mol Genet 1:511-515.

Willemsen R, Mohkamsing S, De Vries B, Devys D, van den Ouweland A, Mandel JL, Galjaard H, Oostra BA (1995): Rapid antibody test for fragile X syndrome. Lancet 345:1147-1148

Willemsen R, Bontekoe C, Tamanini F, Galjaard H, Hoogeveen AT, Oostra BA (1996a): Association of FMRP with ribosomal precursor particles in the nucleolus. Biochem Biophys Res Comm 225:27-33

Willemsen R, Oosterwijk JC, Los FJ, Galjaard H, Oostra BA (1996b): Prenatal diagnosis of fragile X syndrome. Lancet 348:967-968

Willemsen R, Smits A, Mohkamsing S, van Beerendonk H, de Haan A, de Vries B, van den Ouweland A, Sistermans E, Galjaard H, Oostra BA (1997a): Rapid antibody test for diagnosing fragile X syndrome: a validation of the technique. Hum Genet 99:308-311

Willemsen R, Los F, Mohkamsing S, van den Ouweland A, Deelen W, Galjaard H, Oostra BA (1997b): Rapid antibody test for prenatal diagnosis of fragile X syndrome on amniotic fluid cells: a new appraisal. J Med Genet 34:250-251

Willemsen R, Anar B, De Diego Otero Y, de Vries BB, Hilhorst-Hofstee Y, Smits A, van Looveren E, Willems PJ, Galjaard H, Oostra BA (1999): Noninvasive test for fragile X syndrome, using hair root analysis. Am J Hum Genet 65:98-103.

Willemsen R, Olmer R, De Diego Otero Y, Oostra BA (2000): Twin sisters, monozygotic with the fragile X mutation, but with a different phenotype. J Med Genet 37:603-604.

Zhang Y, O'Connor JP, Siomi MC, Srinivasan S, Dutra A, Nussbaum RL, Dreyfuss G (1995): The fragile X mental retardation syndrome protein interacts with novel homologs FXR1 and FXR2. EMBO J 14:5358-5366

Zhong N, Yang W, Dobkin C, Brown WT (1995): Fragile X gene instability: Anchoring AGGs and linked microsatellites. Am J Hum Genet 57:351-361

Chapter 2

Origin of fragile X mutations

2.1

Founder effect in the fragile X syndrome

The identification of several microsatellite markers flanking the FRAXA locus [Richards et al., 1991a; Richards et al., 1991b; Riggins et al., 1992] has been instrumental in the positional cloning of the *FMR1* gene. These markers can still be valuable in family studies, e.g. as additional evidence in prenatal diagnosis.

Additionally, they were employed to verify the presence of any significant gametic disequilibrium between the fragile X mutation and some haplotypes, although the high mutation rate predicted from the early segregation studies [Sherman et al., 1985] implied that new mutants would arise on almost every chromosomal background. Thus, the discovery of linkage disequilibrium encompassing the fragile X locus had been somewhat surprising. Our aim is to review the available evidence of such gametic association and to underline their implications for the mutational mechanism.

1) Fragile X-associated haplotypes

Tables I and II summarize the allelic frequency at two of the most investigated loci, DXS548 and FRAXAC1, in various fragile X and control populations, mainly of European descent [Richards et al., 1992; Hirst et al., 1993; Jacobs et al., 1993; Buyle et al., 1993; Oudet et al., 1993a; Haataja et al., 1994; Malmgren et al., 1994; Macpherson et al., 1994; Zhong et al., 1994a; Chiurazzi et al., 1996a] and of Asiatic origin [Richards et al., 1994a]. In all cases it is apparent that, although fragile X patients can display several haplotypes, only a few of them account for almost 70-80% of the total, with a distribution significantly different from that of the controls. Moreover, isolated populations like the Finns [Oudet el al., 1993b; Haataja et al., 1994; Zhong et al., 1996] and Swedes [Malmgren et al., 1994] show an even more pronounced effect with one single dominant haplotype shared by the majority of patients. We can also note that the main fragile X haplotypes are common to the different European nations and seem distributed along geographical gradients. For example, the FRAXAC1-A/DXS548-2 haplotype, which is prevalent in Italy, and the FRAXAC1-D/DXS548-6 haplotype, prevalent in the UK, are almost equally frequent in France. Preliminary studies have also investigated the haplotypes of normal Chinese [Zhong et al., 1994b] and black Africans [Chiurazzi et al., 1996b]. In accordance with genome-wide microsatellite surveys [Deka et al., 1995], Africans show the highest heterozygosity and mean allele number, possibly suggesting that they have had the longest evolutionary history, while Asiatic populations have the least genetic diversity and lack the European founder haplotype FRAXAC1-A/DXS548-2.

Study	Allele	Chromos. Number	1	2	3	4	5	6	6.5	7	8	9
Zhong [1994a]	fraxa	125	1.6	16.8	8.8	4.8	0.8	20.8	-	39.2	7.2	-
USA	contr.	190	1.6	7.9	3.7	-	1.1	9.4	-	73.1	3.2	-
Macpherson [1994]	fraxa	44	2.3	18.2	-	-	-	36.3	-	34.1	9.1	-
UK	contr.	188	4.3	6.4	-	-	-	14.9	-	73.3	1.1	-
Buyle [1993]	fraxa	68	-	36.8	-	1.5	-	20.6	-	39.6	1.5	-
BE,NL	contr.	134	3	10.4	0.8	1.5	-	10.4	-	73.1	-	0.8
Oudet [1993a]	fraxa	106	0.9	27.5	0.9	0.9	-	30.2	-	39.6	-	-
France	contr.	162	1.2	9.3	0.6	0.6	-	14.2	-	72.3	1.8	-
Chiurazzi [1996a]	fraxa	125	5.6	28.8	9.6	1.6	-	19.2	0.8	28.8	5.6	-
Italy	contr.	215	1.9	4.7	-	-	1.4	13.9	-	75.3	1.9	0.9
Malmgren [1994]	fraxa	28	-	3.6	-	-	-	46.4	-	50	-	-
Sweden	contr.	28	-	7.1	-	-	-	14.3	-	78.6	-	-
Haataja [1994]	fraxa	60	-	8	-	-	-	90	-	2	-	-
Finland	contr.	283	2	10	3	-	-	16	-	69	-	-
Zhong [94b] China	contr.	227	-	-	-	0.4	1.3	13.2	-	80.7	4.4	-

Table 1 - Allele distribution at locus DXS548 in different fragile X and control populations. Frequencies are given in order to facilitate comparison, while absolute numbers of chromosomes tested are indicated in the first column. Allele numbers are as in Macpherson et al.[1994].

Reference	Allele	F-6	E-5	D-4	C-3	B-2	A-1	Z-0	Chromos. Number
Zhong [1994a]	**fraxa**	-	-	34.4	33.6	3.2	28.8	-	125
USA	contr.	-	1.6	22.1	66.8	5.8	3.7	-	190
Richards [1992]	**fraxa**	-	-	45	32	2	21	-	47
Australia, UK, Ceph	contr.	-	-	17	74	2	7	-	202
Hirst [1993]	**fraxa**	-	-	28.9	38.4	5.5	27.4	-	73
UK,Ccph	contr.	-	0.7	16.2	78.5	-	4.6	-	130
Macpherson [1994]	**fraxa**	-	-	34	29.6	9.1	27.3	-	44
UK	contr.	-	-	19.2	75.5	-	5.3	0.7	188
Jacobs [1993] UK	contr.	-	0.3	17.3	74	0.7	7	-	304
Chiurazzi [1996a]	**fraxa**	0.7	0.7	27	39.5	0.7	31.4	-	137
Italy	contr.	-	-	23	68.9	2.1	6	-	235
Richards [1994a]	**fraxa**	-		40	60	-	-	-	40
Japan	contr.	-		69	31	-	-	-	142
Zhong [94b] China	contr.	-	-	70.8	29.2	-	-	-	216

Table II - Allele distribution at locus FRAXAC1 in different fragile X and control populations. Frequencies are given in order to facilitate comparison, while absolute numbers of chromosomes tested are indicated in the first column. Allele names and numbers are as in Richards et al.[1992], while numbers are as in Macpherson et al.[1994].

2) CGG repeat structure

Recent surveys [Kunst and Warren, 1994; Hirst et al., 1994; Eichler et al., 1994; Snow et al., 1994; Zhong et al., 1995; Rousseau et al., 1995; Eichler et al., 1996] have focused on the normal control population and looked at the correlation between haplotypes and CGG length and internal structure. The presence and position of intercalating AGGs seem to be a major determinant of the triplet repeat stability [Hirst, 1995]. In controls, it was shown that one principal fragile X haplotype (FRAXAC1-A/DXS548-2 or 2-1-3 of Eichler et al. [1996]) is associated with longer than average arrays which have a variable 3' uninterrupted tract while retaining the two most 5' AGG interruptions (e.g. 9A9Ax, where x>20), and seem to constitute a pool of more unstable triplets. The 8-generation family reported by Smits et al.[1993] with DXS548-2 allele in five probands may be an example of this instability. On the other hand, the other most prevalent fragile X haplotypes (FRAXAC1-D/DXS548-6 or 6-4-4 and 6-4-5 of Eichler et al.[1996]) are not associated with repeats of longer total size in the control population, but they either lack the second AGG interruption (e.g. 9A22 or 9A26) or have a slightly larger middle CGG tract (e.g. 9A11A9 or 9A12A9). In this case a more rapid transition to the premutation (S-to-Z, according to the Morton and Macpherson [1992]) appears to take place, possibly after the loss of the second stabilizing AGG.

3) Gametic association and multi-step mutational pathways

It is evident that a correct interpretation of linkage disequilibrium and CGG structure data depends on the understanding of the mutational mechanisms responsible for the instability of the FMR1 CGG repeat. Almost two thirds of the fragile X full mutations can be found associated to few (4-5) haplotypes: this is surprising, given the estimate of the overall mutation rate of 0.8×10^{-4}. This rate of mutation from the normal stable to abnormal unstable state is based on the recently re-evaluated prevalence of affected males (1/4000, Turner et al.[1996]) and the assumption of equilibrium. Such an estimate appears too high for founder chromosomes to be detectable after several generations. As correctly pointed out by Chakravarti [1992] and Kolehmainen [1994], it is difficult to reconcile this high frequency of new mutations with the suggestion that most fragile X mutations arose on a small number of chromosomes. Actually, hundreds of "founder" alleles per million meioses generated at random would have produced a fragile X haplotype distribution almost identical to that of the controls. A better explanation would be that of a preferential occurrence of mutations "on particular chromosomal backgrounds" [Kolehmainen, 1994].

The far-reaching hypothesis of a multi-step mutational pathway (Fig. 1a), leading to the constitution of pools of "intermediate" alleles (e.g. N, S, Z, L and possibly other alleles) with a distinct identity and connected by different mechanisms, each operating at its peculiar rate, was formalized by Morton and Macpherson [1992] and then incorporated in all other models [Kolehmainen, 1994; Morris et al., 1995a and b; Ashley and Sherman, 1995].
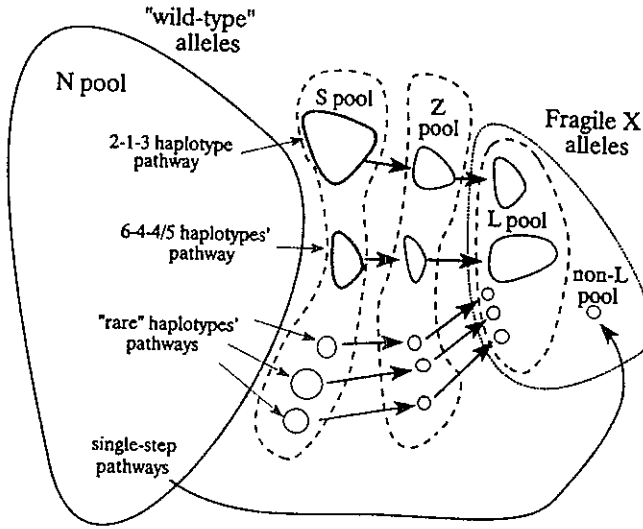
Fig. 1a - Multiple mutational pathways leading from the "wild-type" alleles (N pool) to the fragile X alleles (L pool). Single-step events (traditional mutations as deletions and point mutations) lead to non-full mutation (non-L) fragile X, while multi-step processes lead to the more prevalent full mutations (L pool). The S and Z intermediate pools are also indicated.

A useful analogy can be drawn with a biochemical pathway (Fig. 1b) and we shall note that the *absolute* number of alleles transferred from one to the next pool is the same "at equilibrium" (i.e. at steady flow), but the *relative* frequency of transition is different because the size of the originating pool is different.
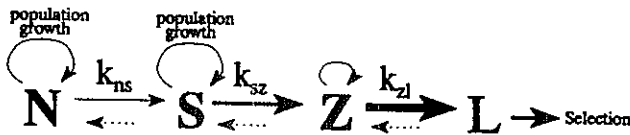
Fig. 1b - One multi-step mutational pathway represented as a chain of biochemical reactions, each one corresponding to a different mechanism with its peculiar transition rates. Note that reverse reactions have almost undetectable rates and the pathway is in a state of constant forward flow (i.e. the biochemical steady-state equivalent to the mutation-selection "equilibrium").

The analogy also readily explains that the S "pool" is larger than the Z "pool" because the S-to-Z transition rate is lower than the Z-to-L rate (infact Z alleles are readily converted into L alleles). Thus, the different size of the pools of the intermediate "products" (i.e. their particular prevalence in the population) make up for the differences between the individual transitional rates between one pool and the other. This allows the coexistence of low-frequency initial events (accounting for the gametic association) and the prevalence of fragile X. One way to reconcile the relatively high mutation rate and the observation of haplotype association is also to consider several different mutational pathways (Fig. 1a). For example, one pathway could include the loss of an AGG which is probably due to a point mutation. This point mutation may only occur at a rate of 10^{-5} to 10^{-7}, therefore it may be associated with a few founder chromosomes. Another pathway may be replication slippage leading to an increase of perfect repeats (e.g. on 2-1-3 chromosomes as suggested by Eichler et al.[1996]). Thus, although the overall mutation rate of $0.8*10^{-4}$ is too high to result in founder effects, some steps of each underlying pathway should occur at much lower rates and can be expected to be associated with specific backgrounds. Genetic drift of the neutral S alleles would play a large role in creating "reservoir" pools. Eventually, other mechanisms (e.g. Okazaki fragment slippage at both ends) may account for the transition to larger premutations or full mutations   [Richards and Sutherland, 1994] at a much faster rate, estimated between $1*10^{-2}$ and almost 0.8 per generation, at least when the Z allele is transmitted from the mother.

4) Origin of fragile X chromosomes
As shown in Fig. 1a, the fragile X prevalence of 1 in 4000 males is probably an endpoint obtained from different "pathways". A small percentage of non-L (non-full mutation) fragile X alleles are due to traditional single-step mechanisms (e.g. deletions or point mutations) which are necessarily short lived and not preferentially associated with any haplotype.
Roughly two thirds of the full mutation (L) fragile X chromosomes are in linkage disequilibrium with specific haplotypes for the reasons discussed above. A single ancestral chromosome with haplotype 2-1-3, associated with a 9A9Ax triplet array, might account for 15-25 % of the fragile X chromosomes. The N-to-S transition event in this case may have been the growth by few units of the 3' CGG tract by slipped-strand mispair mutation mechanism [Levinson and Gutman, 1987; Weber and Wong, 1993; Di Rienzo et al., 1994], additionally constrained by the conservation of the first two AGGs. Alternatively, it may have been a 29-to-39 "jump" (from

9A9A9 to 9A9A9A9 like in Macpherson et al.[1995]) followed by the loss of an AGG (9A9A9A9 to 9A9A19).

In the case of haplotypes 6-4-4 and 6-4-5 the N-to-S event is probably a particular "symmetric" interspersion pattern (9AxA9, with x>10), as suggested by Eichler et al.[1996]. This might predispose to the rapid loss of the second AGG (the S-to-Z transition of this pathway) and we note that no longer-than-average CGG alleles are associated with these haplotypes in the control population (which contains both N and S alleles, while Z and L alleles can be distinguished from their size). Actually, this may be because the S-to-Z transition rate on this pathway is higher than that on the 2-1-3 haplotype pathway and does not allow the formation of a large detectable "pool" of S intermediates (Fig. 1a).

We must note that some low-frequency haplotypes are likely to derive from these few "major" fragile X-associated haplotypes by recombination and/or mutation at the microsatellite themselves. In fact, during the several generations after the initial mutation, the relatively more mutable marker microsatellites may transform a "major" haplotype into a set of apparently "new" ones, clustered next to it [Chiurazzi et al., 1996a]; this could also partially account for the high heterozygosity observed in Fragile X samples. The analysis of the less informative but more stable single nucleotide polymorphisms FMRa and FMRb [Kunst and Warren, 1994] might be appropriate in resolving this particular issue.

A local instability mechanism has been suggested by Zhong et al.[1994a] to explain linkage disequilibrium: here the initial amplification steps at the CGG repeat would predispose the flanking markers to instability, but this should happen in a "directional" way (possibly in the sense of a size increase), otherwise no preferential association would be detected. Anyhow, at this time, no such mechanism has been identified.

The remaining third of fragile X full mutations is not showing any significant linkage disequilibrium: these mutations are mainly associated with less frequent haplotypes, which may be recent mutations that did not yet attain high frequencies, or sometimes with unique "rare" haplotypes, not observed on any control chromosome [Macpherson et al., 1994]. It is possible that some external factor, perhaps a mutation at a DNA repair locus, gives rise to simultaneous instability in FMR1 and its flanking loci. If this happened in an individual homozygous for an instability mutation [Brown et al., 1996], an unpredicted "rare" haplotype would be stabilized in the heterozygous offspring. As yet, no "mutator" phenotype has ever been observed "in vivo" or "in vitro" in fragile X families (normal life span, no specific familial cancers), although exceptional pedigrees might, by coincidence, happen to show a genome-wide instability. Such a

family was described by Zhong et al.[1993], where 3 FRAXAC2 mutations were observed, one of which was significantly not on the fragile X chromosome [Richards et al., 1994b; Mornet et al., 1994; Zhong et al., 1994c]. Surveys on isolated normal populations [Eichler and Nelson, 1996; Chiurazzi et al., 1996b; Kunst et al., 1996] indicate that the FMR1 triplet array attained its average size before the spread of Homo sapiens, thus making gross differences in fragile X prevalence among the various ethnic groups unlikely. Nevertheless, the particular AGG interspersion pattern in small and isolated populations, like some Native Americans [Kunst et al., 1996], might indeed stabilize the CGG triplet over hundreds of generations. Other population surveys investigating the size and purity of the FMR1 triplet arrays will be needed to refine our understanding of the first amplification steps, also bearing in mind the possibility of size reductions [Vits et al., 1994; Chiurazzi et al., 1994], and to understand the relative roles of hairpin formation and replication slippage on the leading and/or lagging strand. Extended family analysis will be necessary to clarify the behaviour of "gray zone" alleles.

5) Note on complex microsatellites and on nomenclature

Detection of linkage disequilibrium can be less efficient when inadvertently using a complex marker microsatellite i.e. a marker including more simple repeats in the same amplification product; in this case alleles cannot be unequivocally distinguished by their size and exact typing would require laborious sequencing, like in the case of FRAXAC2 whose complex structure $(GT)xC(TA)y(T)z$ has been elucidated by Zhong et al.[1993]. A much less polymorphic secondary sequence might be also present in the DXS548 PCR product [Chiurazzi et al., 1996a and b] but in this case it was possible to realize a new 5' primer that dissects the primary CA repeats and allows a straightforward PCR size typing [Chiurazzi et al., 1996a]. The involvement of many independent research groups has created discrepancies in nomenclature for the microsatellite alleles. An attempt to draw together the various systems currently in use was reported by Macpherson et al.[1994] in their Table 1. Systems based on absolute size of PCR products should be avoided as they suffer from the variability of DNA migration at electrophoresis and become confusing when new primers are used to amplify the same sequencer. On the other hand, systems based on arbitrary letter or number codes, which can be standardized by circulating reference DNAs, are easier to employ although they do not refer to any actual sequence information. In theory, systems based on the number of repeats would be most satisfying and possibly a two-digit format with a reference point 00 for zero repeats could allocate both even and odd microsatellite alleles. Such a format would be easy to convert to for

future publications, given only access to some original sequence information and the availability of a standard set of reference DNAs which could be circulated between the laboratories. In the absence of any consensus the increasing body of data may eventually become impossible to collate and compare.

Acknowledgments

References

Ashley AE, Sherman SL (1995): Population dynamics of a meiotic/mitotic expansion model for the Fragile X syndrome. Am J Hum Genet 57:1414-1425

Brown WT, Zhong N, Dobkin C (1996): Positive fragile X microsatellite associations point to a common mechanism of dynamic mutation evolution. Am J Hum Genet 58:641-643

Buyle S, Reyniers E, Vits L, De Boulle K, Handig I, Wuyts FLE, Deelen W, Halley DJJ, Oostra BA, Willems PJ (1993): Founder effect in a Belgian-Dutch fragile X population. Hum Genet 92:269-272

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

Chiurazzi P, Kozak L, Neri G (1994): Unstable triplets and their mutational mechanism: size reduction of the CGG repeat vs. germline mosaicism in the fragile X syndrome. Am J Med Genet 51:517-521

Chiurazzi P, Genuardi M, Kozak L, Giovannucci-Uzielli ML, Bussani C, Dagna-Bricarelli F, Grasso M, Perrone L, Sebastio G, Sperandeo MP, Oostra BA, Neri G (1996a): Fragile X founder chromosomes in Italy: few initial events and possible explanation for their heterogeneity. Am J Med Genet 64:209-215

Chiurazzi P, Destro-Bisol G, Genuardi M, Oostra BA, Spedini G, Neri G (1996b): Extended gene diversity at the FMR1 locus and neighbouring CA repeats in a sub-Saharan population. Am J Med Genet 64:216-219

Deka R, Jin L, Shriver MD, Yu LM, DeCroo S, Hundrieser J, Bunker CH, Ferrell RE, Chakraborty R (1995): Population genetics of dinucleotide (CA)(GT) polymorphism in world populations. Am J Hum Genet 56:461-474

Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB (1994): Mutational processes of simple-sequence repeat loci in human populations. Proc Natl Acad Sci USA 91:3166-3170

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 instability in the FMR1 gene. Nat Genet 8:88-94

Eichler EE, Nelson DL (1996): Genetic variation and evolutionary stability of the FMR1 CGG repeat in six closed human populations. Am J Med Genet 64:220-225

Eichler EE, Macpherson JN, Murray A, Jacobs PA, Chakravarti A, Nelson DL (1996): Haplotype and interspersion analysis of the FMR1 CGG repeat identifies two different mutational pathways for the origin of the fragile X syndrome. Hum Mol Genet 5:319-330

Haataja R, Vaisanen ML, Li M, Ryynanen M, Leisti J (1994): The fragile X syndrome in Finland: demonstration of a founder effect by analysis of microsatellite haplotypes. Hum Genet 94:479-483

Hirst MC, Knight SJL, Christodoulou Z, Grewal PK, Fryns JP, Davies KE (1993): Origins of the fragile X syndrome mutation. J Med Genet 30:647-650

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

Hirst MC (1995): FMR1 triplet arrays: paying the price for perfection. J Med Genet 32:761-763

Jacobs PA, Bullman H, Macpherson J, Youings S, Rooney V, Watson A, Dennis NR (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 polar variation of the fragile X repeat could result in predisposing normal alleles. Cell 77:853-861

Kunst CB, Zerylnick C, Karickhoff L, Eichler E, Bullard J, Chalifoux M, Holden JJA, Torroni A, Nelson DL, Warren ST (1996): FMR1 in global populations. Am J Hum Genet 58:513-522

Levinson G, Gutman GA (1987): Slipped strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol 4:203-221

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 JN, Curtis G, Crolla JA, Dennis N, Migeon B, Grewal PK, Hirst MC, Davies KE, Jacobs PA (1995): Unusual (CGG)n expansion and recombination in a family with fragile X and DiGeorge syndrome. J Med Genet 32:236-239

Malmgren H, Gustavson KH, Oudet C, Holmgren G, Pettersson U, Dahl N (1994): Strong founder effect for the fragile X syndrome in Sweden. Eur J Hum Genet 2:103-109

Mornet E, Chateau C, Taillandier A, Montagnon M, Simon-Bouy B, Serre JL, Boue' A (1994): FRAXAC2 instability. Nat Genet 7:122-123

Morris A, Morton NE, Collins A, Lawrence S, Macpherson JN (1995a): Evolutionary dynamics of the FMR1 locus. Ann Hum Genet 59:283-289

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

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

Oudet C, Mornet E, Serre JL, Thomas F, Lentes-Zengerling S, Kretz C, Deluchat C, Tejada I, Boue' J, Boue' A, Mandel JL (1993a): Linkage disequilibrium between the fragile X mutation and two closely linked CA repeats suggest 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, Nordstroem AM, Peippo M, Mandel JL (1993b): Striking founder effect for the fragile X syndrome in Finland. Eur J Hum Genet 1:181-189

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

Richards RI, Shen Y, Holman K, Kozman H, Hyland VJ, Mulley JC, Sutherland GR (1991a): Fragile X syndrome: diagnosis using highly polymorphic microsatellite markers. Am J Hum Genet 48:1051-1057

Richards RI, Holman K, Kozman H, Kremer E, Lynch M, Pritchard M, Yu S, Mulley J, Sutherland GR (1991b): Fragile X syndrome: genetic localization by linkage mapping of two microsatellite repeats FRAXAC1 and FRAXAC2 which immediately flank the fragile site. J Med Genet 28:818-823

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 fragileX syndrome. Nat Genet 1:257-260

Richards RI, Kondo I, Holman K, Yamauchi M, Seki N, Kishi K, Staples A, Sutherland GR, Hori T (1994a): Haplotype analysis at the FRAXA locus in the Japanese population. Am J Med Genet 51:412-416

Richards RI, Holman K, Friend K, Staples A, Sutherland GR, Oudet C, Biancalana V, Mandel JL (1994b): FRAXAC2 instability. (letter) Nat Genet 7:122

Riggins GJ, Sherman SL, Oostra BA, Sutcliffe JS, Feitell D, Nelson DL, van Oost BA, Smits APT, Ramos FJ, Pfender E, Kuhl DPA, Caskey CT, Warren ST (1992): Characterization of a highly polymorphic dinucleotide repeat 150 kb proximal to the fragile X site. Am J Med Genet 43:237-243

Rousseau F, Rouillard P, Morel ML, Khandjian EW, Morgan K (1995): Prevalence of carriers of premutation-size alleles of the FMR1 gene- and implications for the population genetics of the fragile X syndrome. Am J Hum Genet 57:1006-1018

Sherman SL, Jacobs PA, Morton NE, Froster-Iskenius U, Howard-Peebles PN, Nielsen KB, Partington MW, Sutherland GR, Turner G, Watson M (1985): Further segregation analysis of the fragile X syndrome with special reference to transmitting males. Hum Genet 69:289-299

Smits APT, Dreesen JCFM, Post JG, Smeets DFCM, de Die-Smulders C, Spaans-van der Bijl T, Govaerts LCP, Warren ST, Oostra BA, van Oost BA (1993): The fragile X syndrome: no evidence for any recent mutations. J Med Genet 30:94-96

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

Turner G, Webb T, Wake S, Robinson H (1996): The prevalence of the fragile X syndrome. Am J Med Genet 64:196-197.

Vits L, Deboulle K, Reyniers E, Handig I, Darby JK, Oostra BA, Willems PJ (1994): Apparent regression of the CGG repeat in FMR1 to an allele of normal size. Hum Genet 94:523-526

Weber JL, Wong C (1993): Mutation of human short tandem repeats. Hum Mol Genet 2:1123-1128

Zhong N, Dobkin C, Brown WT (1993): A complex mutable polymorphism located within the fragile X gene. Nat Genet 5:248-252

Zhong N, Ye L, Dobkin C, Brown WT (1994a): Fragile X founder chromosome effects: linkage disequilibrium or microsatellite heterogeneity? Am J Med Genet 51:405-411

Zhong N, Liu X, Gou S, Houck GE,Jr, Li S, Dobkin C, Brown WT (1994b): Distribution of FMR1 and associated microsatellite alleles in a normal chinese population. Am J Med Genet 51:417-422

Zhong N, Dobkin C, Brown WT (1994c): FRAXAC2 instability. (letter) Nat Genet 7:123

Zhong N, Yang W, Dobkin C, Brown WT (1995): Fragile X gene instability: Anchoring AGGs and linked microsatellites. Am J Hum Genet 57:351-361

Zhong N, Kajanoja E, Smits B, Pietrofesa J, Curley D, Wang D, Ju W, Nolin S, Dobkin C, Ryynanen M, Brown T (1996): Fragile X founder effects and new mutations in Finland. Am J Med Genet 64:226-233.

Fragile X founder chromosomes in Italy

A total of 137 fragile X and 235 control chromosomes from various regions of Italy were haplotyped by analizing two neighbouring marker microsatellites, FRAXAC1 and DXS548. The number of CGG repeats at the 5' end of the FMR1 gene was also assessed in 141 control chromosomes and correlated with their haplotypes. Significant linkage disequilibrium between some "major" haplotypes and fragile X was observed, while other "minor" haplotypes may have originated by subsequent mutation at the marker microsatellite loci and/or recombination between them. Recent evidence suggests that the initial mechanism leading to CGG instability might consist of rare ($10^{-6/-7}$) CGG repeat slippage events and/or loss of a stabilizing AGG via A-to-C transversion. Also, the apparently high variety of fragile X chromosomes may be partly due to the relatively high mutation rate ($10^{-4/-5}$) of the microsatellite markers used in haplotyping.

Our fragile X sample showed a higher than expected heterozygosity when compared to the control sample and we suggest that this might be explained by the chance occurrence of the few founding events on different chromosomes, irrespective of their actual frequency in the population. Alternatively, a local mechanism could enhance the microsatellite mutation rate only on fragile X chromosomes, or fragile X mutations might occur more frequently on certain background haplotypes.

Short tandem repeats (STRs), or microsatellites, are polymorphic head-to-tail arrangements of DNA sequences up to about 6 bp [Weber and Wong, 1993]. They are extremely diffused in the human genome and are almost evenly distributed between coding and uncoding regions [Tautz and Renz, 1984], thus becoming useful in establishing high density linkage maps [Weissenbach et al., 1992; Gyapay et al., 1994]. They appear to be 5 to 10 times more frequent than it would be expected by chance [Tautz et al., 1986], possibly as a consequence of a slipped-strand mispairing mechanism [Levinson and Gutman, 1987] particularly active in eukaryotic genomes [Charlesworth et al., 1994]. This mechanism has been shown to operate both *in vitro* [Schloetterer and Tautz, 1992; Hauge and Litt, 1993] and *in vivo* with mean mutation rates estimated at 10^{-4} to almost 10^{-2} per locus per gamete per generation, although the frequency of slippage seems to be higher in the male compared to the female germline [Weissenbach et al., 1992; Weber and Wong, 1993]. Mutation rates of tetranucleotide repeats 4 times higher than that of dinucleotide repeats has been reported [Mahtani and Willard, 1993; Weber and Wong, 1993]. Trinucleotide repeats can be tolerated also inside transcribed and translated sequences without affecting the reading frame but their potential instability can result in gene malfunction if

excessive amplification occurs [Kuhl and Caskey, 1993]. This is the case of the fragile X syndrome and several neurodegenerative disorders [Willems, 1994]. Although "mutator" genes, like MSH2, have been identified that can affect the overall genomic slippage rate by altering the efficiency of the mismatch repair systems [Kunkel, 1993], all presently identified "dynamic mutations" seem to be local phenomena, apparently determined by the particular structure of their trinucleotide repeat. Considering the expectedly high mutation rate of the CGG repeat at the 5' untranslated region of the FMR1 gene, whose amplification is responsible for most cases of fragile X syndrome, the finding of a linkage disequilibrium between the mutation and certain haplotypes defined by flanking microsatellite markers was unpredicted [Richards et al., 1992; Oudet et al., 1993a and b; Buyle et al., 1993; Hirst et al., 1993; Jacobs et al., 1993; Macpherson et al., 1994; Malmgren et al., 1994; Montagnon et al., 1994; Zhong et al., 1994a]. A founder effect suggests a lengthy multistep process leading to the pathogenetic amplification with different transition rates between each mutational stage. In order to search for the existence of founder chromosomes in Italy, 137 unrelated fragile X and 235 control chromosomes were haplotyped with two neighbouring microsatellite markers, DXS548 and FRAXAC1, and the CGG repeat number was assessed in 141 of the control chromosomes to check for the possible existence of a correlation between total repeat length and "at risk" haplotypes.

Patients and Methods
DNA aliquots of 137 unrelated fragile X male patients were collected at four genetic centers (Rome, Genua, Naples and Florence) thus providing a representative sample of most Italian regions. DNA from 141 healthy males and 47 healthy females was also collected from the same centers and constituted our pool of 235 control chromosomes. The microsatellites employed for haplotyping were FRAXAC1 [Richards et al., 1991] and DXS548 [Riggins et al., 1992] which are located 7 kb and 150 kb, respectively, proximal to the CGG repeat at the 5' end of the FMR1 gene. A duplex PCR protocol has been developed to allow simultaneous amplification in a single reaction. Thirty cycles (94°C/1min - 55°C/1min - 72°C/1min) were performed in a hot bonnet Peltier effect thermal cycler and reactions were carried out in 10 μl with 200 ng genomic DNA, 0.3 units of Taq polymerase, 1.5 mM MgCl$_2$, 10% DMSO, 0.2 mM each dNTP, except dCTP which was 2 μM, 0.1 μl of alpha-^{32}P-dCTP (3000 Ci/ml), 10 pmoles of each FRAXAC1 [Zhong et al., 1993] and DXS548 [Verkerk et al., 1991] primer. Primer sequences were respectively: FRAXAC1-F (GAT CTA ATC AAC ATC TAT AGA CTT TAT T), FRAXAC1-R (GAT GAG AGT CAC TTG AAG CTG G), DXS548-F (GTA CAT TAG AGT

CAC CTG TGG TGC) and DXS548-R (AGA GCT TCA CTA TGC AAT GGA ATC). A new forward DXS548 primer (GAA TAG TCT CTG GGG TGG ATC TC) has been designed which produces a shorter PCR product including only the main GT repeat (see Results). The DNA of the 141 control males was also amplified with primers "c" and "f" [Fu et al.,1991] to assess the CGG repeat number with 200 ng genomic DNA, 0.65 units of Taq polymerase, 2.0 mM $MgCl_2$, 10% DMSO, 0.2 mM each dNTP, except dGTP which was 0.05 mM, 0.15 mM 7-deaza-dGTP, 0.13 μl of alpha-^{32}P-dCTP (3000 Ci/ml) and 2 pmoles of each primer. Thirty cycles were again used in a two-step PCR (94°C/45sec - 68°C/2min30sec). A 10 μl aliquot of stop solution (95% formamide and dyes) was added to each tube, PCR products were denatured for 10min at 95°C and 4 μl were loaded on a 6% polyacrylamide/ 7M urea denaturing gel. Samples were run at 70 watts for 2 to 4 hours depending on the size of the amplified products (approximately 200 and 150 bp for the duplex PCR and 300 bp for the 30 CGG repeat allele). Gels were dried without fixation and exposed for 12 to 48 hrs at -80°C with intensifying screens. Reference DNAs were consistently re-loaded at the sides and in the middle of every gel to rule out gel distortion artifacts and samples were always run a second time in increasing size order to cross-check their relative size. Some samples were reamplified if they failed to give any signal the first time or if they showed ambiguous bands. CGG repeat size was evaluated next to an M13 sequence assuming the thickest band to be the original product, as 2-3 faster shadow bands were always present [Hauge and Litt, 1993]. FRAXAC1 alleles were named with letters (A to F) as in Richards et al. [1992] and DXS548 alleles were named as in MacPherson et al. [1994], although the exact size of allele 8 calculated by sequencing (187 bp with 16 GT repeats) did not correspond to that (192 bp) quoted by Riggins et al. [1992] and Oudet et al. [1993a]. The chi-square test with Yate's correction is well suited to analyze the results and has been employed to detect significant differences between the fragile X and control samples. Heterozygosity and its variance were estimated using formulas [8.4] and [8.12] described in Nei [1987].

Results

FRAXAC1 and DXS548 allele distributions in both fragile X subjects and controls are shown in Fig. 1a and b, respectively. As in all previous studies on Caucasian populations, the most frequent alleles in the controls are FRAXAC1-C [Richards et al., 1992; Hirst et al., 1993; Jacobs et al., 1993; Macpherson et al., 1994; Zhong et al., 1994a] and DXS548-7 [Oudet et al., 1993a and b; Buyle et al., 1993; Malmgren et al., 1994; Montagnon et al., 1994; Macpherson et al., 1994; Zhong et al.,1994a]. The frequency of both alleles decreases significantly in the fragile X

population (chi-square 29.8 and 68.5 respectively, p<0.001), matched by an increase in the frequency of other alleles, expecially FRAXACl-A (chi-square 41.2, p<0.001) and DXS548-2 (chi-square 37.4, p<0.001). These values demonstrate the existence of a significant linkage disequilibrium between the fragile X mutation and some alleles of the neighbouring microsatellites.
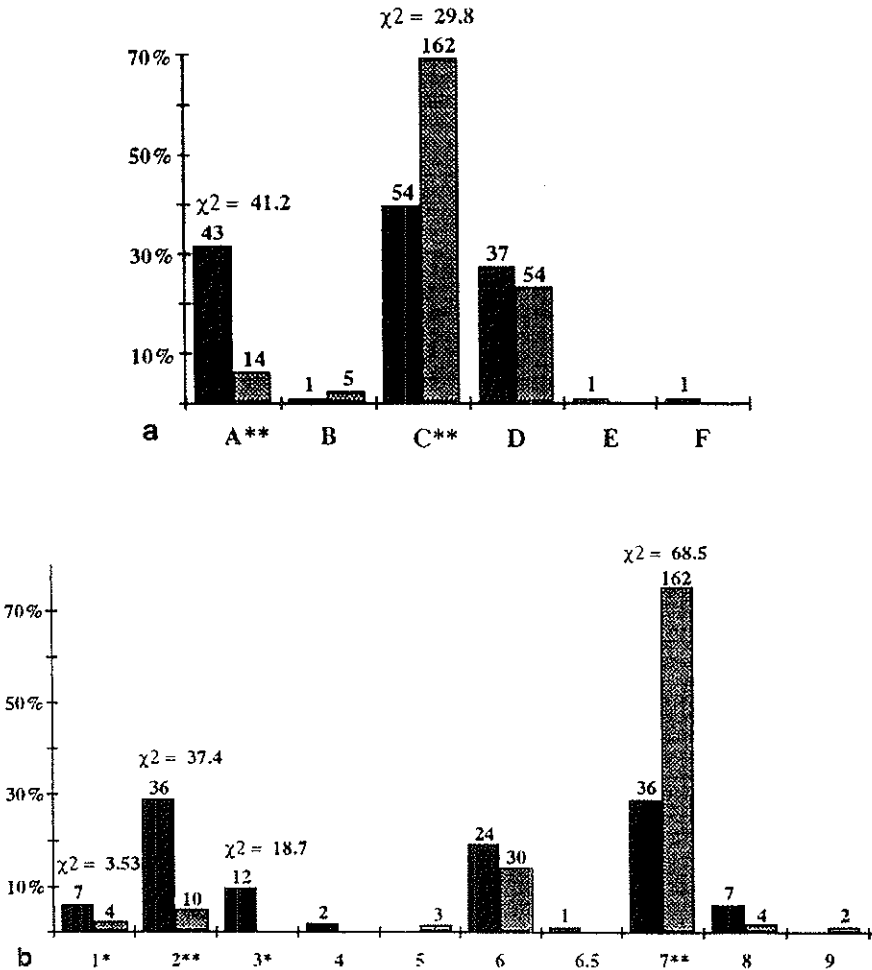
Figure 1 - a) allele distribution at FRAXAC1 locus of 137 fragile X chromosomes (black bars) and 235 controls (gray bars); **b)** allele distribution at DXS548 locus of 125 fragile X chromosomes and 215 controls. Absolute chromosome numbers are reported on the top of every bar. Statistically significant differences are indicated by asterisks next to the allele symbol and chi-square values are written on the top of corresponding bars.

DXS548 intermediate allele 6.5, observed only once, positions itself between alleles 6 and 7. PCR has been repeated several times and the product was also run with an automatic sequencer to confirm the intermediate size of this allele. Examining the DXS548 sequence [S.T. Warren, unpublished] approximately 50 bp proximal to the GT repeat, we could notice a $(C)_4G(C)_{11}$ sequence whose longer poly-C tract might be slightly polymorphic [Weber, 1990]. In fact, when we directly sequenced three clones of a DXS548-8 allele we found a $(C)_4G(C)_{10}$ motif; this 1 bp insertion/deletion polymorphism could account for these intermediate alleles, although its rare occurence in European populations does not blur the information obtained with the original primers. Anyhow, more of these intermediate alleles have been observed in a normal black African population (Chiurazzi et al., 1996), therefore a new forward DXS548 primer has been designed which gives a shorter (120 bp) product including only the main GT repeat.

Fig. 2 represents the FRAXAC1-DXS548 haplotypes in decreasing order of frequency in the control population, and the corresponding frequencies found in the fragile X sample.
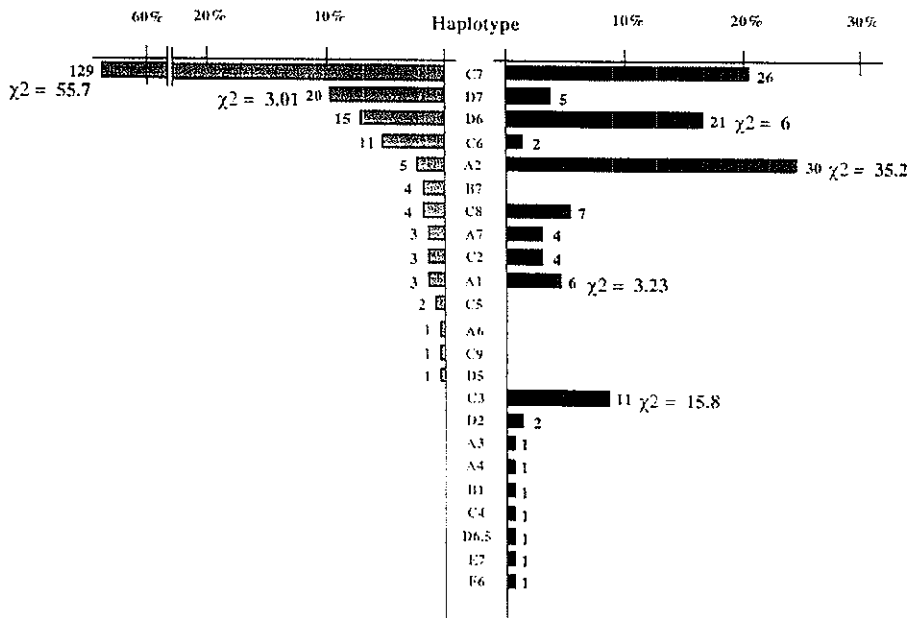
Figure 2 - Haplotype distribution of 125 fragile X chromosomes (black bars) and 202 controls (gray bars) presented in decreasing order of frequency in the normal control population. Absolute chromosome numbers are reported next to every bar. Statistically significant differences are indicated by writing chi-square values aside the corresponding bars.

Again as in previous studies, haplotype C-7 is by far the most common among the controls, while in the fragile X sample its significant decrease (chi-square 55.7, $p<0.001$) is mainly compensated for by the rise of haplotype A-2 (chi-square 35.2, $p<0.001$). Also clearly increased in the patients' population are haplotypes C-3 (chi-square 15.8, $p<0.001$) and D-6 (chi square 6, $p<0.015$), while few other haplotypes give a minor contribution to the fragile X pool.

A comparison with the data of Macpherson et al. [1994] and, relatively to DXS548 only, with those of Oudet et al. [1993a,b], Buyle et al. [1993] and Malmgren et al. [1994] suggests a striking similarity of fragile X haplotypes distribution, with a South-to-North gradient noted for haplotypes A-2 and D-6, the latter being more prevalent in Scandinavia and Britain, while the former is dominant in Italy and Southern Europe. Actually, A-2 was also observed in 2 out of 3 patients from Iran and in one Singhalese boy that were referred to us [Chiurazzi, unpublished].
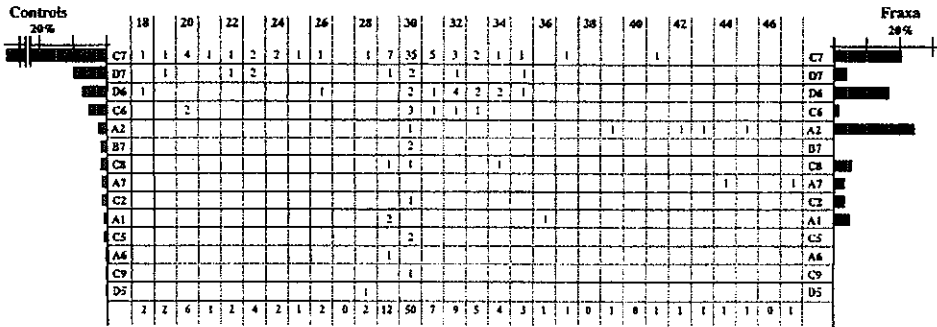
Figure 3 - CGG repeat number distribution of 141 male controls, subdivided according to haplotype.

The CGG repeat distribution of our control population, subdivided according to haplotype, is shown in Fig.3 in order to verify if the total repeat length is a major factor in determining a higher instability of some haplotypes. The few control chromosomes with haplotype A-2 have indeed a higher mean CGG repeat number, as also found by Kunst and Warren [1994]. However, this is not the case for D-6 and the other haplotypes that are more prevalent among the fragile X population, underlining the role of factors other than total repeat length in causing triplet instability.

Fig.4 presents the fragile X and control haplotype distribution as a two-dimensional array of the DXS548 and FRAXAC1 alleles. Its simple inspection allows two important observations: 1) the

relatively large number of different fragile X haplotypes, although 4 of them (A-2, C-7, D-6 and C-3), that we shall designate major haplotypes, make up for 70% of the total; 2) the higher heterozygosity (86%±1.6%) of the fragile X sample compared to that of controls (57.4%±3.9%), which is not unexpected if we consider that the fragile X chromosomes are more evenly distributed over the 4 major haplotypes (A-2,24%; C-7,20.8%; D-6,16.8% and C-3, 8.8%), while the control distribution has a single dominant peak (C-7,64%).

FraX / wt	1	2	3	4	5	6	6.5	7	8	9	Total
A	4.80* / 1.49	24* / 2.48	0.8	0.8		/ 0.5		3.2 / 1.49			33.6 / 5.94
B	0.8							/ 1.98			0.8 / 1.98
C		3.2 / 1.49	8.8*	0.8	/ 0.99	1.6 / 5.45		20.8 / 63.86*	5.6 / 1.98	/ 0.5	40.8 / 74.26
D		1.6			/ 0.5	16.8* / 7.43	0.8	4.0 / 9.9*			23.2 / 17.82
E								0.8			0.8
F						0.8					0.8
Total	5.6 / 1.49	28.8 / 3.96	9.6	1.6	/ 1.49	19.2 / 13.37	0.8	28.8 / 77.23	5.6 / 1.98	/ 0.5	100 / 100

Figure 4 - Haplotype distribution of 125 fragile X chromosomes (bold type) and 202 controls (wt) presented as a bidimensional array. FRAXAC1 alleles correspond to rows and DXS548 alleles to columns. Frequencies rather than numbers are given in order to facilitate comparison and statistically significant differences are indicated by asterisks. Haplotype expected heterozigosity (± standard error) was estimated for the fragile X (86%±1.6%) and for the control samples (57.4%±3.9%).

Discussion

The study of fragile X population genetics has great relevance to test the various hypotheses made on the nature of the "dynamic" mutational mechanism affecting the FMR1 gene. In fact, the first evidence of linkage disequilibrium between FMR1 mutations and flanking marker loci prompted Morton and Macpherson [1992] to propose a multistep model with the frequency of transition increasing from one mutational stage to the next (N-to-S < S-to-Z < Z-to-L). This basic property is common to other models [Kolehmainen, 1994; Ashley and Sherman, 1994] and allows a few protomutations to generate a relatively high number of full mutations [Chakravarti, 1992]. Thus, founder chromosomes can be identified, although the strongly reduced fitness of

the affected males would imply many new mutations on almost every chromosomal background [Sherman et al., 1985] if a single-step mutational mechanism is assumed.

A complex mechanism also emerges from recent observations on slipped-strand mispairing and hairpin formation, delineating a double-stage scenario. A first phase of low-grade instability would include the eventual loss of stabilizing imperfections as intercalated AGGs [Kunst and Warren, 1994; Eichler et al., 1994; Hirst et al., 1994; Snow et al., 1994; Zhong et al., 1995] and slippage of a few or, more rarely, even 10 repeat units [Macpherson et al., 1995]. Such a complex stepwise mutational mechanism (SMM), possibly common to all microsatellites, would allow both expansions and reductions to occur even in discrete "jumps" and seems to fit better the experimental data [Di Rienzo et al., 1994; Deka et al., 1995; Watkins et al., 1995]. The determinant factors for this initial instability would be both the CGG repeat internal structure and its length, as also suggested by our data presented in Fig.3. Then, a second phase of more pronounced instability, leading to larger expansions, would be possible only for microsatellite with specific repeat sequences because, above a certain length threshold, formation of stable hairpin structures is favoured [Gacy et al., 1995] and, in the case of CGG repeats, hairpins would facilitate and be further stabilized by cytosine methylation [Fry and Loeb, 1994]. Moreover, homogenous Okazaki fragments composed almost exclusively of repeats would be even more prone to mispairing and hairpin formation [Richards and Sutherland, 1994]; while the location and distance of the closest origin of replication might favour expansion over deletion events [Kang et al., 1995] and account for the amplification polarity [Kunst and Warren, 1994]. Kolehmainen [1994] correctly pointed out that an estimate of 2×10^{-4} for the initial transition rate (wild-type to protomutation) would result in too many founder alleles to accept the suggestion that all fragile X mutations originated on a small number of founder chromosomes. We suggest that the above mentioned mechanistic evidence is compatible with a downward revision of the estimate of the initial (N-to-S) transition rate to values comparable to those of most point mutations ($10^{-6/-7}$), as in the case of the A-to-C transversion leading to the loss of a stabilizing AGG [Eichler et al., 1994]. Furthermore, the striking similarity among Caucasian fragile X haplotypes, the South-North gradient of haplotype A-2 possibly originated on an ancient Indoeuropean chromosome, and the analysis of fragile X pedigrees extending over several generations [Smits et al., 1993], all argue in favour of an extremely long history of fragile X mutations. Indeed, recent studies on a black African population [Chiurazzi et al., 1996] and on primates [Deelen et al., 1994; Rubinsztein et al., 1995; Zhong et al., 1995] show that the CGG repeat at the 5' end of the FMR1 gene is evolutionarily conserved and already

attained the mean size of 30 repeats with stabilizing AGGs in the primates. Of course, some founder chromosomes, as D-6, could be of more recent origin and may have not yet contributed fully to the pool of fragile X patients yet [Kunst and Warren, 1994; Mandel, 1994]. Based on the above evidence, we may now assume a multistep mutational mechanism with initial transition rate of $10^{-6/-7}$ and reinterpretate the first of the observations derived from Fig.4. The relative large number of different fragile X haplotypes can be reconciled with a strong linkage disequilibrium; in fact only the major haplotypes (A-2, C-7, D-6, C-3 and possibly A-1 and C-8) probably reflect a mutational event and represent the original founder chromosomes, while the other fragile X haplotypes may have originated from the major ones via slippage at one or both flanking markers, given the relatively higher ($10^{-3/-5}$) mutation rate of microsatellites [Weber and Wong, 1993]. In less frequent cases, recombination between microsatellite markers might have created a minor haplotype, one such case possibly being haplotype A-7 that could originate from a crossing-over between a founder A-2 and a common control C-7 haplotype. In fact, the only two A-7 control chromosomes have a high CGG repeat number, as it is the case for most A-2 chromosomes from which they would be derived. The identification and use of less polymorphic markers (e.g. RFLPs) in the FMR1 gene will be crucial for the future assessment of this problem.

The second observation we made from Fig.4 was that of the higher heterozygosity ($86\% \pm 1.6\%$) of the fragile X sample compared to that of controls ($57.4\% \pm 3.9\%$). This is not an obvious finding because in the absence of a founder effect, when new mutations are frequent, the haplotype distributions of patients and controls would be almost identical, while in presence of a strong founder effect the majority of mutations would be expected to occur on a more limited haplotype background. However, the observed high heterozygosity could still be compatible with few founding events if they occured on different and distant haplotypes, irrespective of their relative frequency in the control population. Of course, it would have been more likely for the founding chromosomes to cluster on and around the common control C-7 haplotype, but fragile X mutations during history might have generated more dispersed clusters of major and minor founder chromosomes by chance. This "casual" hypothesis has the advantage of not implying any particular mechanism linking the stability of the FMR1 CGG repeat with that of its neighbouring marker loci, but in alternative, as suggested by Morris et al. [1995], we have to consider three other possible models accounting for the association of some haplotypes with the fragile X mutations. In the "concurrent" model another factor, perhaps a mutation at a DNA repair locus, causes a simultaneous instability of the fragile X locus and of its flanking markers.

Although there has been no observation of increased microsatellite instability at other genomic loci in most fragile X families, this could be the case of a minority of families, possibly including that with multiple FRAXAC2 mutations reported by Zhong et al. [1993]. The "causal" model supposes that some haplotypes might predispose to mutation the FMR1 CGG repeat, but this would imply the presence of yet another cis-element acting on the triplet while it seems that length and purity of the CGG repeat itself are the main determinants of its stability. Finally, the "consequent" model suggests that the fragile X mutation might enhance the neighbouring microsatellite mutation rate [Zhong et al., 1994a and c; Mornet et al., 1994; Richards et al., 1994b; Zhong et al., 1995], and this deserves careful verification although it would tend to reduce any linkage disequilibrium by definition and there is no clear evidence of a plausible mechanism. However, a thorough analysis of larger sets of data from Asiatic populations might prove helpful, considering that Japanese [Richards et al., 1994a] and possibly Chinese populations [Zhong et al., 1994b] are less heterogenous than European ones and present very similar fragile X and control haplotype distributions; this should allow an easier verification of the possible existence of an increased mutability of flanking markers on fragile X chromosomes.

Acknowledgments

References

Ashley AE, Sherman SL (1995): Population dynamics of a meiotic/mitotic expansion model for the Fragile X syndrome. Am J Hum Genet 57:1414-1425

Buyle S, Reyniers E, Vits L, De Boulle K, Handig I, Wuyts FLE, Deelen W, Halley DJJ, Oostra BA, Willems PJ (1993): Founder effect in a Belgian-Dutch fragile X population. Hum Genet 92:269-272

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

Charlesworth B, Sniegowski P, Stephan W (1994): The evolutionary dynamics of repetitive DNA in eukaryotes. Nature 371:215-220

Chiurazzi P, Destro-Bisol G, Genuardi M, Oostra BA, Spedini G, Neri G (1996): Extended gene diversity at the FMR1 locus and neighbouring CA repeats in a sub-Saharan population. Am J Med Genet 64:216-219

Deelen W, Bakker C, Halley DJJ, Oostra BA (1994): Conservation of CGG region in FMR1 gene in mammals. Am J Med Genet 51:513-516

Deka R, Jin L, Shriver MD, Yu LM, DeCroo S, Hundrieser J, Bunker CH, Ferrell RE, Chakraborty R (1995) Population genetics of dinucleotide (CA)(GT) polymorphism in world populations. Am J Hum Genet 56:461-474

Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB (1994): Mutational processes of simple-sequence repeat loci in human populations. Proc Natl Acad Sci USA 91:3166-3170

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

Fry M, Loeb LA (1994): The fragile X syndrome d(CGG)$_n$ nucleotide repeats form a stable tetrahelical structure. Proc Natl Acad Sci USA 91:4950-4954

Fu YH, Kuhl DPA, Pizzuti A, Pieretti M, Sutcliffe JS, Richards S, Verkerk AJMH, Holden JJA, Fenwick RG, 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

Gacy AM, Goellner G, Juranic N, Macoura S, McMurray CT (1995): Trinucleotide repeats that expand in human disease form hairpin structures in vitro. Cell 81:533-540

Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, Marc S, Bernardi G, Lathrop M, Weissenbach J (1994): The 1993-1994 Genethon human genetic linkage map. Nat Genet 7:246-249

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

Hirst MC, Knight SJL, Christodoulou Z, Grewal PK, Fryns JP, Davies KE (1993): Origins of the fragile X syndrome mutation. J Med Genet 30:647-650

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

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

Kang S, Jaworski A, Ohshima K, Wells RD (1995): Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in E.coli. Nat Genet 10:213-218

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

Kuhl DPA, Caskey CT (1993): Trinucleotide repeats and genome variation. Curr Opin Genet Develop 3:404-407

Kunkel TA (1993): Slippery DNA and diseases. Nature 365: 207-208

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

Levinson G, Gutman GA (1987): Slipped strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol 4:203-221

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 JN, Curtis G, Crolla JA, Dennis N, Migeon B, Grewal PK, Hirst MC, Davies KE, Jacobs PA (1995): Unusual (CGG)n expansion and recombination in a family with fragile X and DiGeorge syndrome. J Med Genet 32:236-239

Mahtani MM, Willard HF (1993): A polymorphic X-linked tetranucleotide repeat locus displaying a high rate of new mutation: implications for mechanisms of mutation at short tandem repeat loci. Hum Mol Genet 2:431-437

Malmgren H, Gustavson KH, Oudet C, Holmgren G, Pettersson U, Dahl N (1994): Strong founder effect for the fragile X syndrome in Sweden. Eur J Hum Genet 2:103-109

Mandel JL (1994): Trinucleotide diseases on the rise. Nat Genet 7:453-455

Montagnon M, Boygo A, Deluchat C, Jokic M, Chateau C, Taillandier A, Thomas F, Simon-Bouy B, Boue' J, Serre JL, Boue' A, Mornet E (1994): Transition from normal to premutated alleles in fragile X syndrome results from a multistep process. Eur J Hum Genet 2:125-131

Mornet E, Chateau C, Taillandier A, Montagnon M, Simon-Bouy B, Serre JL, Boue' A (1994): FRAXAC2 instability. Nat Genet 7:122-123

Morris A, Morton NE, Collins A, Lawrence S, Macpherson JN (1995): Evolutionary dynamics of the FMR1 locus. Ann Hum Genet 59:283-289

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

Nei M (1987): Molecular evolutionary genetics. Columbia University Press, New York

Oudet C, Mornet E, Serre JL, Thomas F, Lentes-Zengerling S, Kretz C, Deluchat C, Tejada I, Boue' J, Boue' A, Mandel JL (1993a): Linkage disequilibrium between the fragile X mutation and two closely linked CA repeats suggest 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, Nordstroem AM, Peippo M, Mandel JL (1993b): Striking founder effect for the fragile X syndrome in Finland. Eur J Hum Genet 1:181-189

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

Richards RI, Holman K, Kozman H, Kremer E, Lynch M, Pritchard M, Yu S, Mulley J, Sutherland GR (1991): Fragile X syndrome: genetic localization by linkage mapping of two microsatellite repeats FRAXAC1 and FRAXAC2 which immediately flank the fragile site. J Med Genet 28:818-823

Richards RI, Holman K, Friendendend 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

Richards RI, Kondo I, Holman K, Yamauchi M, Seki N, Kishi K, Staples A, Sutherland GR, Hori T (1994a): Haplotype analysis at the FRAXA locus in the Japanese population. Am J Med Genet 51:412-416

Richards RI, Holman K, Friend K, Staples A, Sutherland GR, Oudet C, Bianoulana V, Mandel JL (1994b): FRAXAC2 instability. Nat Genet 7:122

Riggins GJ, Sherman SL, Oostra BA, Sutcliffe JS, Feitell D, Nelson DL, van Oost BA, Smits APT, Ramos FJ, Pfender E, Kuhl DPA, Caskey CT, Warren ST (1992): Characterization of a highly polymorphic dinucleotide repeat 150 kb proximal to the fragile X site. Am J Med Genet 43:237-243

Rubinsztein DC, Amos W, Leggo J, Goodburn S, Jain S, Li SH, Margolis RL, Ross CA, Ferguson-Smith MA (1995): Microsatellite evolution-evidence for directionality and variation in rate between species. Nat Genet 10:337-343

Schloetterer C, Tautz D (1992): Slippage synthesis of simple sequence DNA. Nucl Ac Res 20:211-215

Sherman SL, Jacobs PA, Morton NE, Froster-Iskenius U, Howard-Peebles PN, Nielsen KB, Partington MW, Sutherland GR, Turner G, Watson M (1985): Further segregation analysis of the fragile X syndrome with special reference to transmitting males. Hum Genet 69:289-299

Smits APT, Dreesen JCFM, Post JG, Smeets DFCM, de Die-Smulders C, Spaans-van der Bijl T, Govaerts LCP, Warren ST, Oostra BA, van Oost BA (1993): The fragile X syndrome: no evidence for any recent mutations. J Med Genet 30:94-96

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

Tautz D, Renz M (1984): Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucl Ac Res 12:4127-4138

Tautz D, Trick M, Dover GA (1986): Cryptic simplicity in DNA is a major source of genetic variation. Nature 322:652-656

Verkerk AJMH, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DPA, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang F, Eussen BE, van Ommen GJB, Blonden LAJ, Riggins GJ, Chastain JL, Kunst CB, Galjaard H, Caskey CT, Nelson DL DL, Oostra BA, Warren ST (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

Watkins WS, Bamshad M, Jorde LB (1995): Population genetics of trinucleotide repeat polymorphism. Hum Mol Genet 4:1485-1491

Weber JL (1990): Informativeness of human $(dC-dA)_n(dG-dT)_n$ polymorphism. Genomics 7:524-530

Weber JL, Wong C (1993): Mutation of human short tandem repeats. Hum Mol Genet 2:1123-1128

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

Willems PJ (1994): Dynamic mutations hit double figures. Nat Genet 8:213-215

Zhong N, Dobkin C, Brown WT (1993): A complex mutable polymorphism located within the fragile X gene. Nat Genet 5:248-252

Zhong N, Ye L, Dobkin C, Brown WT (1994a): Fragile X founder chromosome effects: linkage disequilibrium or microsatellite heterogeneity? Am J Med Genet 51:405-411

Zhong N, Liu X, Gou S, Houck GE,Jr, Li S, Dobkin C, Brown WT (1994b): Distribution of FMR1 and associated microsatellite alleles in a normal chinese population. Am J Med Genet 51:417-422

Zhong N, Dobkin C, Brown WT (1994c): FRAXAC2 instability. Nat Genet 7:123

Zhong N, Yang W, Dobkin C, Brown WT (1995): Fragile X gene instability: Anchoring AGGs and linked microsatellites. Am J Hum Genet 57:351-361

2.3

African haplotypes at the fragile X locus

We report on the allele distributions in a normal black African population at two microsatellite loci neighbouring the FRAXA locus and at the CGG repeat in the 5' end of the *FMR1* gene, which causes the fragile X syndrome. The CGG repeat distribution was found to be similar to that of other ethnic groups, as well as to that of non-human primates, possibly predicting a comparable prevalence of fragile X in Africa.

Significant linkage disequilibrium has been observed between fragile X mutations and alleles of the DXS548 and FRAXAC1 loci in European and Asian populations, and some founder chromosomes may be extremely old. Those associated with FRAXAC1-A and DXS548-2 alleles are not present in the Asian fragile X samples. We searched for these alleles and their frequency in the well defined Bamileke population of Cameroon. All previously described alleles and some new ones were found in this sample, supporting the hypothesis of their pre-existance and subsequent loss in Asian populations. Finally, the heterozygosity of the Bamileke sample was significantly higher at both marker loci and comparable to that of Europeans at the CGG repeat, confirming the notion that genetic diversity is greater in Africans than in other groups and supporting the view that evolution of modern man started in Africa.

Only few reports have dealt with fragile X in sub-Saharan Africa [Sutherland and Hecht, 1985; Venter et al., 1981; Howard-Peebles and Stoddard, 1980] and, although the prevalence estimate in whites of European origin has been recently lowered to approximately 1 in 4000 males [Turner et al., 1996], fragile X syndrome seems either to be less frequent or to have been underascertained in African populations. Watkins et al. [1995] surveyed five trinucleotide repeat loci other than fragile X and suggested that the presence of alleles at the upper end of the normal range might correlate with an increased prevalence of the corresponding diseases, as in the case of dentato-rubro-pallido-luysian atrophy (DRPLA) in the Japanese population. Thus, we determined the allele distribution of the CGG repeat at the 5' end of the FMR1 gene in a normal sample from the Bamileke population of Cameroon, in order to test if a possible lack of larger sized alleles might explain a low incidence of fragile X in Africa, as it was proposed by Goldman et al. [1994] in the case of myotonic dystrophy. We also tested this population with flanking markers FRAXAC1 and DXS548, to verify whether alleles FRAXAC1-A and DXS548-2, which are frequently found on European FMR1 founder chromosomes [Chiurazzi et al., 1996; Macpherson et al., 1994], are present in Africa. These alleles are completely missing in the Japanese and Chinese patients populations

[Richards et al., 1994; Arinami et al., 1993; Zhong et al., 1994] and we hypothesize their preexistence in Africa, their subsequent spread to Indoeuropeans and loss in the Asian populations.

Sample and Methods

Our sub-Saharan African sample is comprised of 85 independent chromosomes derived from 9 male and 38 female unrelated Bamileke individuals. Bamilekes are considered the major ethnic group of Cameroon, accounting for 11% of the country population [Spedini and Destro-Bisol, 1988]. An additional 8 male Bororo and 8 male Sanga individuals from the Central African Republic were also tested. Bororos, considered the most unadmixed group of the larger Fulbe population, are nomadic shepherds, while the Sanga people derive their name from the river on which they base their subsistence. According to Greenberg [1980], the languages spoken by these populations belong to the same linguistic philum (Niger-Kordofan), although the linguistic family and group shared by Bamilekes and Sangas (Bantoid-Bantu) differ from those of Bororos (West Atlantic-Fulfulde).

The microsatellites studied were FRAXAC1 [Richards et al., 1991] and DXS548 [Riggins et al., 1992] which are located 7 kb and 150 kb, respectively, proximal to the CGG repeat at the 5' end of the FMR1 gene. A duplex PCR protocol has been developed to allow simultaneous amplification in a single reaction [Chiurazzi et al., 1996]. Primers "c" and "f" [Fu et al., 1991] were employed to assess the CGG repeat number with 200 ng genomic DNA, 0.65 units of Taq polymerase, 2.0 mM MgCl, 10% DMSO, 0.2 mM each dNTP, except dGTP which was 0.05 mM, 0.15 mM 7-deaza-dGTP, 0.13 µl of alpha-^{32}P-dCTP (3000 Ci/ml) and 2 pmoles of each primer. Thirty cycles were again used in a two-step PCR (94°C/45" - 68°C /2'30"). Samples were run at 70 watts for 2 to 4 hours depending on the size of the amplified products (approximately 200 and 150 bp for the duplex PCR and 300 bp for the 30 CGG repeat allele). Gels were dried without fixation and exposed for 12 to 48 hrs at -80°C with intensifying screens. Reference DNAs were consistently re-loaded at the sides and in the middle of every gel to rule out gel distortion artifacts and samples were reamplified if they failed to give any signal the first time or if they showed ambiguous bands. CGG repeat size was evaluated next to an M13 sequence assuming the thickest band to be the original product, as 2-3 faster shadow bands were always present. FRAXAC1 alleles were named with letters (A to F) as in Richards et al. [1992] and DXS548 alleles were named as in MacPherson et al.

[1994]. Heterozygosity and its variance were estimated for all three ethnic groups using formulas [8.4] and [8.12] described in Nei [1987].

Results

Because of the low number of Sanga and Bororo individuals tested, only data from the Bamileke sample have been thoroughly analyzed and are presented in Fig.1 and Table 1. Italian [Chiurazzi et al., 1996] and Chinese [Zhong et al., 1994] control groups are also displayed to provide comparison with Caucasian and Asian populations, respectively.

Fig. 1 represents the allele distribution at the CGG repeat of the FMR1 gene and clearly shows that the African distribution has two modes, with a substantial proportion of alleles at 30 repeats (27%) and at 29 repeats (27%), coinciding with the European and Asian modes, respectively. A minor peak is present at 22 repeats (7%) and the range spans from allele 22 to 41 (Italian range, 18-47; Chinese range, 21-45), while 15.3% of the sample is above 31 repeats (24.1% in Italians; 19.5% in Chinese).
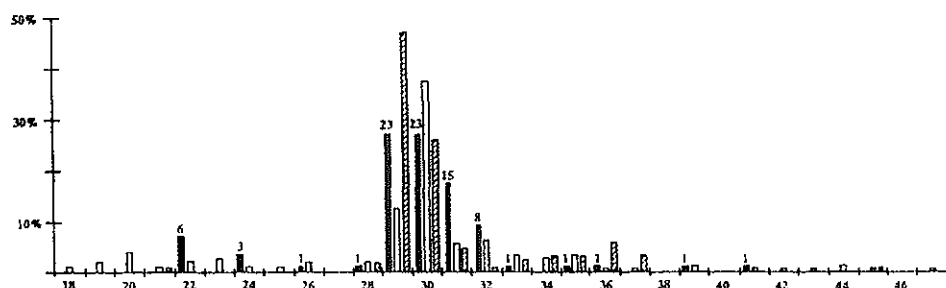
Figure 1 – CGG repeat number distribution at the *FMR1* locus of 85 unrelated normal chromosomes from the Bamileke ethnic group of Cameroon (black bars). Absolute numbers are noted on top of every bar. For comparison, a European [Chiurazzi et al., 1996] and an Asian [Zhong et al., 1994] control populations are also represented by the white and shaded bars, respectively.

Fig. 2 and 3 show the distributions of FRAXAC1 and DXS548 alleles, which are rather spread out in the African population.

At the FRAXAC1 locus, again the Bamileke chromosomes are almost equally distributed between the European mode (allele C, 53.6%) and the Asian one (allele D, 37.8%) and the range comprises alleles A to F, this latter being described before only in one fragile X patient from Southern Italy [Chiurazzi et al., 1996].
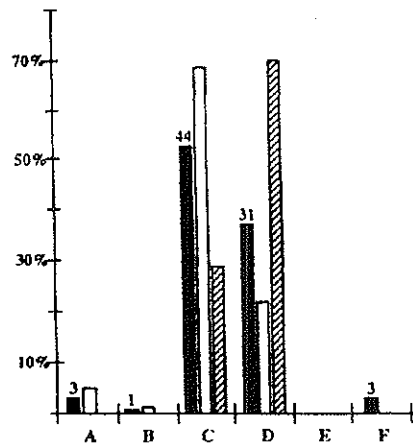
Figure 2 – Allele distribution at the FRAXAC1 locus of 82 unrelated normal chromosomes from the Bamileke ethnic group of Cameroon (black bars). Absolute numbers are noted on each bar. European [Chiurazzi et al., 1996] and Asian [Zhong et al., 1994] control populations are also represented by the white and shaded bars, respectively.

At the DXS548 locus one can notice rare alleles too as DXS548-3 (11%), which we found in almost 9% of our fragile X patients but not in controls [Chiurazzi et al., 1996], and DXS548-0 that we observed in a Sanga male, which has never been reported before. As we previously noted [Chiurazzi et al., 1996], intermediate-size alleles 1.5, 3.5 and 5.5 (1, 4 and 3 cases, respectively) could be due to a 1-bp insertion/deletion polymorphism within a variable $(C)4G(C)11$ sequence, approx. 50 bp upstream from the GT repeat.
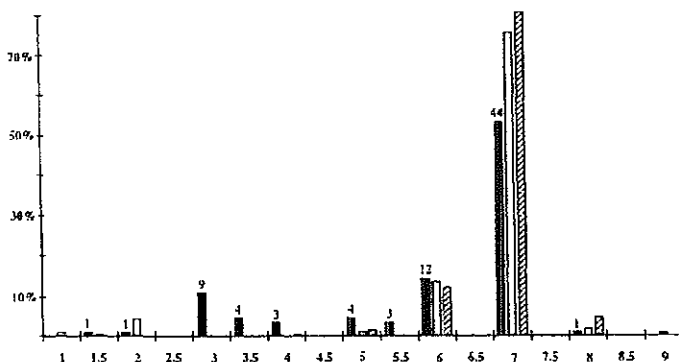
Figure 3 – Allele distribution at the DXS548 locus of 82 unrelated normal chromosomes from the Bamileke ethnic group of Cameroon (black bars). Absolute numbers are noted on each bar. European [Chiurazzi et al., 1996] and Asian [Zhong et al., 1994] control populations are also represented by the white and shaded bars, respectively.

Table I compares the expected heterozygosities at every locus, demonstrating that the Bamilekes have the highest gene diversity at both marker loci FRAXAC1 and DXS548, while the estimated value intervals (± standard error) of the Bamilekes and Italians extensively overlap at the FMR1 CGG repeat locus.

	Bamileke	**Italians**	**Chinese**
FMR1-**CGG**	0.816 ± 0.021 (85)	0.832 ± 0.027 (141)	0.706 ± 0.033 (123)
FRAXAC1	0.573 ± 0.032 (82)	0.467 ± 0.031 (231)	0.419 ± 0.026 (206)
DXS548	0.679 ± 0.050 (82)	0.405 ± 0.039 (211)	0.327 ± 0.039 (206)

Table I – Expected heterozygosity values. Heterozygosity ± standard error at the three analyzed loci were calculated for the Bamileke ethnic group of Cameroon [present study], the Italian control sample that we recently reported [Chiurazzi et al., 1996] and the Chinese studied by Zhong et al. [1994]. Total chromosome numbers are indicated in brackets.

Finally, we will also note that DXS548 intermediate-size alleles are more frequent in Bamilekes (9.7%) than in Italians, where they can be considered extremely rare (1 out of 215, i.e. <0.5%); thus we may assume that Africans show an increased heterozygosity also at the $(C)4G(C)11$ contained in the DXS548 PCR product. Therefore, we are currently investigating the variability of this other polymorphic element by using a different 5' primer. This will help to exploit the full informative potential of the main GT repeat without blurring effects.

Discussion

Surveys of allele distribution at various microsatellite loci in different human populations and groups of non-human primates have helped to elucidate their mutational mechanism, while at the same time shedding some light on the evolutionary history of our species [Cavalli-Sforza, 1991; Cavalli-Sforza and Piazza, 1993]. A complex stepwise mutational mechanism (SMM), allowing both expansion and reduction to occur mainly by single bp slippage but also in discrete "jumps", seems to fit best the experimental data [Di Rienzo et al., 1994; Deka et al., 1995] accounting for the multimodal distributions often observed at several loci [Mandel, 1994]. If constant mutation rates and no selective constraints are assumed for microsatellite variation, gene diversity can be correlated with time elapsed since foundation, although a small

effective population size could reduce the observed variability [Paabo, 1995]. We observed the highest heterozygosity in the ethnically well defined Bamilekes (Table I), despite the larger size of the two non-African samples and their presumably greater extent of recent admixture. This observation is consistent with other evidence based on PCR analysis of mini- and microsatellite loci in various human populations [Bowcock et al., 1994; Destro-Bisol et al., 1994; Deka et al., 1995]. Although the evolutionary significance of a greater genetic diversity is controversial [Rogers and Jorde, 1995], our data might support the hypothesis of a monocentric origin of modern man, migrating out of Africa [Gibbons, 1995].

The substantially similar distribution of Bamileke, Italian and Chinese alleles at the FMR1 CGG repeat locus (Fig. 1a), with the highest observed African allele at 41 repeats (vs. 47 in Italians and 45 in Chinese), would suggest a detectable prevalence of the fragile X syndrome in Africa. On the other hand, Rubinsztein et al. [1995] have proposed the existence of a general trend toward size increase of microsatellite repeats during evolution, based on comparison between human and primate samples, and indeed the lack of high-normal range alleles at the CAG repeat of the myotonin locus might explain why the prevalence of myotonic distrophy in Africa appears low [Goldman et al., 1994; Watkins et al., 1995]. Deelen et al. [1994] first reported CGG counting at the FMR1 locus in an african green monkey that showed two alleles of 26 and 31 repeats, while Zhong et al. [1995] tested 8 chimpanzees and 6 orangutans, finding a mean repeat length of 34 and 30 repeats respectively and also confirming the presence of intercalated AGGs. Actually, the same data of Rubinsztein et al. [1995] demonstrate a complete overlap of human FMR1 CGG allele distribution with that of gorillas, orangutans and baboons, while their sample of 19 chimpanzees showed an even higher mean length. So it seems more likely to assume that microsatellite sizes fluctuate during evolution in balance between expansion and reduction trends, and the FMR1 CGG repeat itself seems to have reached its present size already before the other primates diverged from the human lineage.

As further proof, we should consider that our African sample included two previously undescribed alleles, one being the smallest (FRAXAC1-F) at its locus and the other being the largest (DXS548-0); gene diversity in Africans is then increased both in the sense of increasing and decreasing size. From Fig. 1b and c we verify the presence in the Bamilekes of both alleles FRAXAC1-A and DXS548-2 (Fig. 1b and c), which are frequently associated on fragile X chromosomes in Caucasians while they are almost absent in the Far East Asians. They

were also found in two Iranian and one Singhalese patients that were referred to us [Chiurazzi et al., 1996]. Thus we speculate that Indoeuropean fragile X A-2 chromosomes could derive from an ancient African founder while they might have been lost in Asian populations, possibly because of a "bottleneck" effect.

Acknowledgments

References

Arinami T, Asano M, Kobayashi K, Yanagi 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

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

Cavalli-Sforza LL (1991): Genes, peoples and languages. Sci Am 265:72-78

Cavalli-Sforza LL, Piazza A (1993): Human genomic diversity in Europe: a summary of recent research and prospects for the future. Eur J Hum Genet 1:3-18

Chiurazzi P, Genuardi M, Kozak L, Giovannucci-Uzielli ML, Bussani C, Dagna-Bricarelli F, Grasso M, Perrone L, Sebastio G, Sperandeo MP, Oostra BA, Neri G (1996): Fragile X founder chromosomes in Italy: few initial events as possible explanation for their heterogeneity. Am J Med Genet 64:209-215

Deelen W, Bakker C, Halley DJJ, Oostra BA (1994): Conservation of CGG region in FMR1 gene in mammals. Am J Med Genet 51:513-516

Deka R, Jin L, Shriver MD, Yu LM, DeCroo S, Hundrieser J, Bunker CH, Ferrell RE, Chakraborty R (1995): Population genetics of dinucleotide (CA)(GT) polymorphism in world populations. Am J Hum Genet 56:461-474

Destro-Bisol G, Presciuttini S, d'Aloja E, Dobosz M, Spedini G, Pascali VL (1994): Genetic variation at the ApoB 3' HVR, D2S44, and D7S21 loci in the Ewondo ethnic group of Cameroon. Am J Hum Genet 55:168-174

Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB (1994): Mutational processes of simple-sequence repeat loci in human populations. Proc Natl Acad Sci 91:3166-3170

Fu YH, Kuhl DPA, Pizzuti A, Pieretti M, Sutcliffe JS, Richards S, Verkerk AJMH, Holden JJA, Fenwick RG, 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

Gibbons A (1995): Out of Africa- at last? Science 267:1272-1273

Goldman A, Ramsay M, Jenkins T (1994): Absence of myotonic dystrophy in southern african negroids is associated with a significantly lower number of CTG trinucleotcolcolcolcotide repeats. J Med Genet 31:37-40

Greenberg JH (1980): Classification des langues d'Afrique. In J. Ki Zerbo (ed.), "Histoire generale d'Afrique" tome I, Jeune Afrique/Unesco, Paris, pp.321-338

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

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

Mandel JL (1994): Trinucleotide diseases on the rise. Nat Genet 7:453-455

Nei M (1987): Molecular evolutionary genetics. Columbia University Press. New York

Paabo S (1995): The Y chromosome and the origin of all of us (men). Science 268:1141-1142

Richards RI, Holman K, Kozman H, Kremer E, Lynch M, Pritchard M, Yu S, Mulley J, Sutherland GR (1991): Fragile X syndrome: genetic localization by linkage mapping of two microsatellite repeats FRAXAC1 and FRAXAC2 which immediately flank the fragile site. J Med Genet 28:818-823

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

Richards RI, Kondo I, Holman Kn K, Yamauchi M, Seki N, Kishi K, Staples A, Sutherland GR, Hori T (1994): Haplotype analysis at the FRAXA locus in the Japanese population. Am J Med Genet 51:412-41-416

Riggins GJ, Sherman SL, Oostra BA, Sutcliffe JS, Feitell D, Nelson DL, van Oost BA, Smits APT, Ramos FJ, Pfender E, Kuhl DPA, Caskey CT, Warren ST (1992): Characterization of a highly polymorphic dinucleotide repeat 150 kb proximal to the fragile X site. Am J Med Genet 43:237-243

Rogers AR, Jorde LB (1995): Genetic evidence of modern human origins. Human Biology 67:1-36

Rubinsztein DC, Amos W, Leggo J, Goodburn S, Jain S, Li SH, Margolis RL, Ross CA, Ferguson-Smith MA (1995): Microsatellite evolution- evidence for directionality and variation in rate between species. Nat Genet 10:337-343

Spedini G, Destro-Bisol G (1988): Etude anthropologique au Cameroun. Le peuplement ancien. Donnes archeologiques, linguistiques at anthropologiques. Antropologia Contemporanea 11:93-112

Sutherland GR, Hecht F (1985): Fragile sites on human chromosomes. Oxford University Press, Chapter 11, page 177 "Racial distribution"

Turner G, Webb T, Wake S, Robinson H (1996): Prevalence of fragile X syndrome. Am Am J Med Genet 64:196-197

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 60:807-811

Watkins WS, Bamshad M, Jorde LB (1995): Population genetics of trinucleotide repeat polymorphism. Hum Mol Genet 4:1485-1491

Zhong N, Liu X, Gou S, Houck GE,Jr, Li S, Dobkin C, Brown WT (1994): Distribution of FMR1 and associated microsatellite alleles in a normal chinese population. Am J Med Genet 51:417-422

Zhong N, Yang W, Dobkin C, Brown WT (1995): Fragile X gene instability: Anchoring AGGs and linked microsatellites. Am J Hum Genet 57:351-361

Zhang Y, O'Connor JP, Siomi MC, Srinivasan S, Dutra A, Nussbaum RL, Dreyfuss G (1995): The fragile X mental retardation syndrome protein interacts with novel homologs FXR1 and FXR2. EMBO J 14:5358-5366

Zhong N, Yang W, Dobkin C, Brown WT (1995): Fragile X gene instability: Anchoring AGGs and linked microsatellites. Am J Hum Genet 57:351-361

Chapter 3

Reactivation of the inactive *FMR1* gene

3.1

In vitro reactivation of the *FMR1* gene

Fragile X syndrome is the most frequent cause of heritable mental retardation. Most patients have a mutation in the 5' untranslated region of the *FMR1* gene, consisting of the amplification of a polymorphic $(CGG)_n$ repeat sequence, and cytogenetically express the folate-sensitive fragile site FRAXA in Xq27.3. Fragile X patients harbour an expanded sequence with more than 200 CGG repeats (full mutation), accompanied by methylation of most cytosines of the sequence itself and of the upstream CpG island. This abnormal hypermethylation of the promoter suppresses gene transcription, resulting in the absence of the FMR1 protein. Rare individuals of normal intelligence were shown to carry a completely or partially unmethylated full mutation and to express the FMR1 protein. Given this observation and knowing that the open reading frame of the mutated *FMR1* gene is intact, we decided to investigate whether its activity could be restored *in vitro* by inducing DNA demethylation with 5-azadeoxycytidine (5-azadC) in fragile X patients' lymphoblastoid cells. We report that treatment with 5-azadC causes reactivation of fully mutated *FMR1* genes with 300 to 800 repeats, as shown by the restoration of specific mRNA and protein production. This effect correlates with the extent of promoter demethylation, determined by restriction analysis with methylation-sensitive enzymes. These results confirm the critical role of the *FMR1* promoter hypermethylation in the pathogenesis of the fragile X syndrome, provide additional explanation for the normal IQ of the rare males with unmethylated full mutations and pave the way to future attempts at pharmacologically restoring mutant *FMR1* gene activity *in vivo*.

Fragile X syndrome is the most frequent cause of heritable mental retardation[1]. Over 95 % of patients have a mutation in exon 1 of the *FMR1* gene, consisting of the amplification of a $(CGG)_n$ sequence[2]. This sequence, which is polymorphic in the normal population with a mean copy number of 30 CGGs, is transcribed but not translated[3]. It becomes meiotically and mitotically unstable in premutation carriers with 50 to 200 CGG repeats. Patients with fragile X syndrome have an expansion of more than 200 repeats (full mutation), accompanied by methylation of most cytosines of the CGG stretch and of the upstream CpG island[4,5]. It has been suggested that the abnormal hypermethylation of the *FMR1* promoter suppresses gene transcription, resulting in absence of specific mRNA and protein[6-10], in spite of the fact that the open reading frame of the gene is intact. The observation of rare individuals of normal intelligence, carrying a completely or partially unmethylated full mutation and expressing the FMR1 protein[11-17], also supports this hypothesis. However, since patients are often mosaics

harbouring several amplified alleles of different lengths, as well as premutations which are still transcriptionally active, genotype-phenotype correlations *in vivo* are of limited use.

Therefore we decided to try to restore the activity of the *FMR1* gene *in vitro* by inducing DNA demethylation with 5-azadeoxycytidine (5-azadC) in lymphoblastoid cell lines established from fragile X patients with a "pure" full mutation, which did not have any detectable level of *FMR1* mRNA and protein expression.

Results

Lymphoblastoid cell lines from 5 male fragile X patients and 2 male controls were used in the experiments. The size of the amplification in the patients' cell lines ranged from 300 to 800 CGG repeats. Both peripheral lymphocytes and EBV-transformed lymphoblasts of control individuals abundantly express FMR1 protein, while the patients' cell lines did not produce any detectable amounts of either specific mRNA or protein, with the partial exception of patient 5 as discussed below. All cell lines were cultured in duplicate as described in the Methods and one aliquot was treated with 5-azadC for 7 days while tissue culture medium was added as mock treatment to the other aliquot. Cell counts at the end of each experiment indicated that cell viability was not significantly different in treated and untreated cell lines.

FMR1 gene expression

The effect of 5-azadC on transcription was tested by RT-PCR (Fig.1) using primers 4924-4925 for *FMR1* (145 bp product) and, as an internal control, primers 243-244 for the housekeeping *HPRT* gene (370 bp product). No *FMR1* product is seen in lanes 2,4,6,8 and 10, corresponding to the untreated patients' samples (patients 1 to 5), while the specific band is clearly present in lanes 3,5,7,9 and 11 that contain samples of the same patients treated with 5-azadC. Some *FMR1* mRNA is actually detectable in the untreated sample from patient 5 (lane 10), in whom DNA methylation analysis showed that the full mutation was partially demethylated (data not shown). In the untreated and treated control males' cells the *FMR1* gene is fully expressed (lanes 12 and 13) and the 5-azadC treatment does not appear to affect the *FMR1/HPRT* ratio. A modified radioactive RT-PCR protocol with reduced number of cycles allowed semi-quantitation of the *FMR1* transcripts detectable in the treated patients' lines (Fig.2), which range from 10 to over 50% of the normal mean value, as evaluated by densitometric analysis of the autoradiograms.
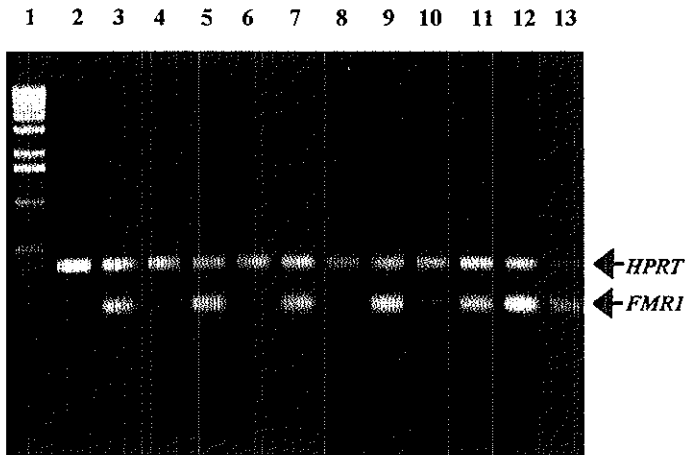
Figure 1. RT-PCR products (32 cycles) of *FMR1* (lower band, 145 bp) and *HPRT* (upper band, 370 bp) from all patients' cell lines with and without 5-azadC treatment. Fragments were separated on a 1.2% agarose gel and stained with ethidium bromide. Samples corresponding to untreated cell lines of patients 1 to 5 are shown in lanes 2,4,6,8 and 10, while treated samples were loaded next, in lanes 3,5,7,9 and 11. Samples of a control male are shown in lanes 12 (untreated) and 13 (treated). A 1-kb molecular weight marker was loaded in lane 1.
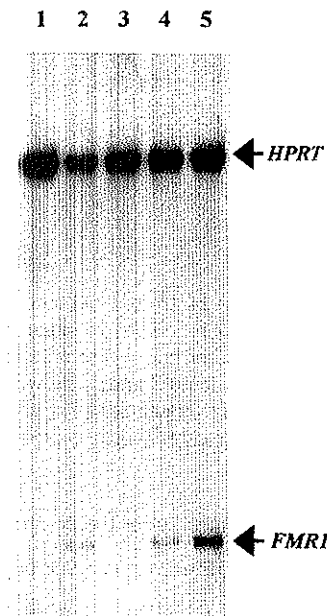
Figure 2. Radioactive RT-PCR products (21 cycles) of *FMR1* and *HPRT* from two patients' cell lines with and without 5-azadC treatment. Samples corresponding to untreated cell lines of patients 1 and 3 and of one control male are shown in lanes 1,3 and 5, respectively, while treated samples of patients 1 and 3 were loaded in lanes 2 and 4.

CpG island demethylation

DNA methylation analysis using methylation-sensitive enzymes showed loss of the usual hypermethylation of patients' full mutations. A representative experiment is shown in Fig.3 for patients 3 and 4, who harbour full mutations of 700 and 500 CGG repeats, respectively. Partial demethylation of the *Sst*II (equivalent to *Sac*II) site of the *FMR1* promoter is clearly present in 5-azadC treated cells (lanes 4 and 6), while hypermethylation is complete in the corresponding untreated samples (lanes 3 and 5). Again, there was no evidence of any premutation or normal sized fragments after treatment.
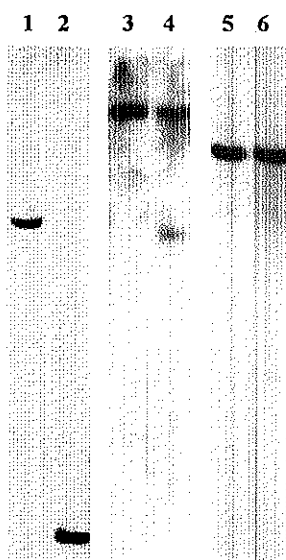
Figure 3. Southern blot of genomic DNA of treated and untreated patient and control cell lines. Lane 1: normal control digested with *Hind*III only; lane 2: same sample digested with *Hind*III and *Sst*II (isoschizomer of *Sac*II). Samples in all other lanes were digested with both *Hind*III and *Sst*II. Untreated samples of patients 3 and 4 are loaded in lanes 3 and 5, showing only methylated bands of 700 and 500 repeats, respectively. The corresponding treated samples are loaded in lanes 4 and 6, showing the appearance of demethylated bands. The promoter-specific probe Ox1.9 was used for detection.

Methylation analysis was also performed with the enzymes *Eag*I and *Bss*HII (data not shown). However, these enzymes were not as effective as *Sst*II, suggesting an uneven distribution of the demethylated sites across the *FMR1* CpG island. It is worth noting that expression of the *FMR1* gene has also been demonstrated in a tumor harbouring a large premutation with both *Eag*I and *Bss*HII sites completely methylated[18]. Thus it seems that the methylation density of the 5' CpG

island as a whole may be more important than the methylation of specific sites in determining the transcriptional status of the *FMR1* gene.

Protein expression

Reappearance of the FMR1 protein in the cytoplasm of the treated patients' cell lines was demonstrated by immunocytochemistry, as shown in Fig. 4. Cells from an untreated patient show no specific fuchsin red staining because no protein is recognized by the specific mouse monoclonal antibody (Fig.4b), as opposed to cells from a normal control (Fig.4a), whose cytoplasm stains positive. After 5-azadC treatment a proportion of patients' cells show specific staining (Fig. 4c), though sometimes of reduced intensity. The number of cells expressing FMR1 protein is in line with the observed level of promoter demethylation.

a – control male   **b – untreated patient**   **c – treated patient**
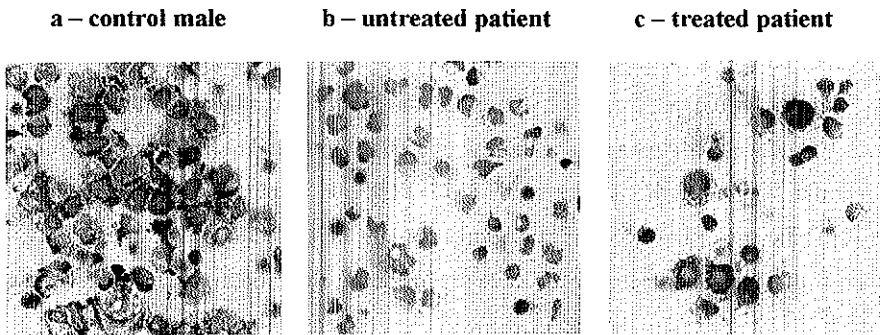
Figure 4. Immunocytochemistry of cytospins prepared from treated and untreated cell lines (for the colors, please refer to the original paper). **a,** control male. **b,** Patient 3 without treatment. **c,** Patient 3 after 5-azadC treatment.

Discussion

We have shown that mutant *FMR1* genes with expansions of 300-800 CGG triplet repeats are able to resume transcription after 5-azadC treatment (Fig. 1 and 2), concomitant with demethylation of the promoter (Fig. 3). The efficiency of transcription appears to correlate with the degree of demethylation, as also indirectly suggested by previous studies in which the activity of a recombinant construct containing the *FMR1* promoter was inhibited by hypermethylation[9-10]. The extent of promoter demethylation attained in our experiments is probably related to the fraction of cycling cells which incorporated 5-azadC in culture. Thus, it appears that the amount of specific *FMR1* mRNA produced by the cells depends on the degree

of demethylation of the promoter, while the amplification itself does not seem to block transcription of hypomethylated alleles. Similar effects of both 5-azadC and 5-azacytidine (5-azaC) have been reported on the transcriptional reactivation of several genes, like PGK1 and HPRT[19-21], which are also normally silenced by the hypermethylation of their respective 5' CpG islands on the inactive X chromosome.

Moreover, we showed that mutant *FMR1* transcripts with up to 800 CGGs can be translated into protein (Fig.4), while it had been suggested that, above a threshold of 280 triplets, translation of abnormal mRNAs is completely abolished because of stalling of the repeated sequence on the small ribosomal subunit[22]. In other words, we provide direct evidence that hypermethylation of the *FMR1* promoter and not amplification of the CGG repeat is the major determinant in abolishing FMR1 protein production. The difference between our results and those of Feng *et al.*[22], who tested fibroblast-like cell lines, could be due to different cell specificity in the expression of the FMR1 protein. Our experiments are in line with the observations of Smeets *et al.*[11] on lymphoblastoid cell lines of two normal males with unmethylated full mutation and we were able to induce a comparable situation with a DNA demethylation protocol. Thus, even if the abnormal length of the amplified *FMR1* mRNA is likely to diminish somewhat the traslational efficiency, fragile X full mutations can still be transcribed and translated when the promoter of the gene is demethylated. Reduced levels of the protein appear to be sufficient to prevent mental retardation in the few males with unmethylated full mutations[11].

In perspective, we think that our experiments pave the way to future attempts at pharmacologically restoring gene activity in the fragile X patients. End-stage patients with β-thalassemia were treated with intravenous azacytidine, which was able to boost gamma globin expression up to therapeutic levels[23]. While long term treatment with such agents may have serious adverse effects, use of safer DNA demethylating drugs, after *in vitro* testing for efficacy and lack of toxicity, may represent a not so remote therapeutic possibility for fragile X patients. Eventually, one might even consider the possibility of a preventive treatment. In fact, there is evidence that methylation of fully mutated *FMR1* gene takes place only during a short time period before week 10 of fetal development[7]. If this early methylation event could be either prevented or reversed, it might be expected that the gene will stay demethylated during further development and after birth.

Materials and Methods

Cell culture and DNA demethylating treatment.

EBV-transformed lymphoblastoid cell lines were established from blood samples of fragile X patients and normal controls using the procedure of Neitzel[24]. Cells were grown in RPMI1640 medium with 10% fetal calf serum and penicillin/ streptomycin at 37°C, in 5% CO_2 atmosphere. Cells were counted and seeded at the initial concentration of $3x10^5$ /ml and synchronized with two 8-hour blocks of 1 mM thymidine (Sigma) spaced by a 10-hour interval. After each thymidine block the cells were centrifuged and resuspended in fresh medium. Treatment with the DNA demethylating agent 5-azadeoxycytidine (5-azadC, Sigma, 0.5-1 μM) was started after the first thymidine block. Total cell culture volume was tripled by the end of the treatment (from 20 to 60 ml) and fresh 5-azadC was added every 24-48 hours, to keep its concentration constant. A control flask for each cell line was also grown without addition of 5-azadC. At day 7 cell concentration was measured and both treated and untreated cultures were recovered and subdivided.

RT-PCR analysis.

Total RNA was extracted with the single-step acid phenol method[25]. cDNA synthesis was carried out with 200 Units of MoMLV-RT (Gibco/BRL) incubating 5 μg of total RNA and 1.2 μg of random hexamers (Pharmacia) at 37° C for 60-90 min. Expression of specific mRNA was determined by RT-PCR using primers 4925-4924 located in exons 3 and 4 of the *FMR1* gene[6]. As internal control of amplification we employed primers 244-243 of the *HPRT* gene which is constitutively expressed in all cells[6]. 32 PCR cycles (94° 1'- 55° 1'- 72° 2') were employed to detect the amplification products with ethidium bromide after separation on a 1.2% agarose gel. Cycle number was lowered to 21 when α-^{32}P dCTP was added and PCR products were separated on a denaturing 6% polyacrylamide gel (7M urea). Densitometric analysis of the autoradiographic film was carried out with the NIH Image software (version 1.61).

DNA methylation analysis.

Genomic DNA was extracted according to a routine salt-chloroform protocol[26]. Southern blots were prepared digesting 10 μg genomic DNA with 24 Units of methylation-sensitive enzyme *Sst*II (Gibco/BRL), an isoschizomer of *Sac*II, and 50 Units of *Hind*III (MBI/Fermentas), incubated overnight at 37° C in the *Hind*III buffer. Samples were run on a 0.8% agarose gel and blotted on Hybond N+ nylon membrane (Amersham). DNA methylation status of the *FMR1* promoter was evaluated by hybridization of probe Ox1.9 [27]. Filters were hybridized overnight at 65°C in 0.5 M sodium phosphate/ 7% SDS buffer and subsequently washed for 30 min with SSC 2x/SDS 0.1% at room temperature and for 10 min with SSC 0.5x/SDS 0.1% at 65°C, then exposed for 3-7 days at -80°C to Kodak Ektamat G films with intensifying screens.

Protein immunocytochemistry.

Cytospins of treated and untreated cell suspensions were prepared on microscope slides and the presence of the FMR1 protein was tested with a previously described antibody test[28] that employs the specific mouse monoclonal antibody 1A1 (Euromedex). Cells were fixed for 10 min with a drop of 4% paraformaldehyde and permeabilized with 100% methanol for 20 min. Slides were then washed in 0.1 M PBS+ (0.15% glycine, 0.5% BSA) and incubated 1 hour at room temperature with 1:200 mouse monoclonal antibody. After washing and another 1-hour incubation with 1:200 biotinylated goat anti-mouse immunoglobulins, a final 45 min incubation followed, with 1:100 streptavidin-biotinylated alkaline phosphatase (StreptABComplex/alkaline phosphatase, DAKO). Revelation was obtained with the New Fuchsin substrate-chromogen system (DAKO) for 15-30 min after extensive washing with 0.1 TRIS. Nuclei were counterstained with Gill's haematoxylin and slides mounted.

Acknowledgements

References

1. Oostra, B.A. and Willems P.J. (1995) A fragile gene. *Bioessays* 17,941-947.

2. Verkerk, A.J.M.H., Pieretti, M., Sutcliffe, J.S., Fu, Y.H., Kuhl, D.P.A., Pizzuti, A., Reiner, O., Richards, S., Victoria, M.F., Zhang, F., Eussen, B.E., van Ommen, G.J.B., Blonden, L.A.J., Riggins, G.J., Chastain, J.L., Kunst, C.B., Galjaard, H., Caskey, C.T., Nelson, D.L., Oostra, B.A., Warren, S.T. (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.

3. Fu, Y.H., Kuhl, D.P.A., Pizzuti, A., Pieretti, M., Sutcliffe, J.S., Richards, S., Verkerk, A.J.M.H., Holden, J.J.A., Fenwick, R.G., Warren, S.T., Oostra, B.A., Nelson, D.L., Caskey, C.T. (1991) Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* 67:1047-1058

4. Luo, S., Robinson, J.C., Reiss, A.L. and Migeon, B.R. (1993) DNA methylation of the fragile X locus in somatic and germ cells during fetal development: relevance to the fragile X syndrome and X inactivation. *Som. Cell. Mol. Genet.* 19,393-404.

5. Hansen, R.S., Gartler, S.M., Scott, C.R., Chen, S.H. and Laird, C.D. (1992) Methylation analysis of CGG sites in the CpG island of the human FMR1 gene. *Hum. Mol. Genet.* 1,571-578.

6. Pieretti, M., Zhang, F., Fu, Y.H., Warren, S.T., Oostra, B.A., Caskey, C.T., Nelson, D.L. (1991) Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell* 66,817-822.

7. Sutcliffe, J.S., Nelson, D.L., Zhang, F., Pieretti, M., Caskey, C.T., Saxe, D., Warren, S.T. (1992) DNA methylation represses FMR-1 transcription in fragile X syndrome. *Hum. Mol. Genet.* 1,397-400.

8. Verheij, C., Bakker, C.E., de Graaff, E., Keulemans, J., Willemsen, R., Verkerk, A.J., Galjaard, H., Reuser, A.J., Hoogeveen, A.T., Oostra, B.A. (1993) Characterization and localization of the *FMR1* gene product associated with fragile X syndrome. *Nature* 363,722-724.

9. Hwu, W.L., Lee, Y.M., Lee, S.C. and Wang, T.R. (1993) In vitro DNA methylation inhibits FMR1 promoter. *Biochem. Biophys. Res. Comm.* 193,324-329.

10. Sandberg, G. and Schalling, M. (1997) Effect of in vitro promoter methylation and CGG repeat expansion on FMR-1 expression. *Nucleic Acids Res.* 25,2883-2887.

11. Smeets, H.J.M., Smits, A.P.T., Verheij, C.E., Theelen, J.P.G., Willemsen, R., van de Burgt, I., Hoogeveen, A.T., Oosterwijk, J.C., Oostra, B.A. (1995) Normal phenotype in two brothers with a full FMR1 mutation. *Hum. Mol. Genet.* 4,2103-2108.

12. de Vries, B.B.A., Jansen, C.C.A.M., Duits, A.A., Verheij, C., Willemsen, R., van Hemel, J.O., van den Ouweland, A.M.W., Niermeijer, M.F., Oostra, B.A., Halley, D.J.J. (1996) Variable FMR1 gene methylation of large expansions leads to variable phenotype in three males from one fragile X family. *J. Med. Genet.* 33,1007-1010.

13. Wang, Z., Taylor, A.K. and Bridge, J.A. (1996) FMR1 fully expanded mutation with minimal methylation in a high functioning fragile X male. *J. Med. Genet.* 33,376-378.

14. Hagerman, R.J., Hull, C.E., Safanda, J.F., Carpenter, I., Staley, L.W., O'Connor, R.A., Seydel, C., Mazzocco, M.M., Snow, K., Thibodeau, S.N., Kuhl, D., Nelson, D.L., Caskey, C.T., Taylor, A.K. (1994) High functioning fragile X males: demonstration of an unmethylated fully expanded FMR1 mutation associated with protein expression. *Am. J. Med. Genet.* 51,298-308.

15. McConkie-Rossell, A., Lachiewicz, A.M., Spiridigliozzi, G.A., Tarleton, J., Schoenwald, S., Phelan, M.C., Goonewardena, P., Ding, X., Brown, W.T. (1993) Evidence that methylation of the FMR-1 locus is responsible for variable phenotypic expression of fragile X syndrome. *Am. J. Hum. Genet.* 53,800-809.

16. Rousseau, F., Robb, L.J., Rouillard, P. and Der Kaloustian, V.M. (1994) No mental retardation in a man with 40% abnormal methylation at the FMR1 locus and transmission of sperm cell mutations as premutations. *Hum. Mol. Genet.* 3,927-930.

17. Loesch, D.Z., Huggins, R., Hay, D.A., Gedeon, A.K., Mulley, J.C., Sutherland, G.R. (1993) Genotype-phenotype relationships in fragile X syndrome: a family study. *Am. J. Hum. Genet.* 53,1064-1073.

18. de Graaff, E., Willemsen, R., Zhong, N., de Die-Smulders, C.E.M., Brown, W.T., Freling, G., Oostra, B. (1995) Instability of the CGG repeat and expression of the FMR1 protein in a male fragile X patient with a lung tumor. *Am. J. Hum. Genet.* 57,609-618

19. Sasaki, T., Hansen, R.S., Gartler, S.M. (1992) Hemimethylation and hypersensitivity are early events in transcriptional reactivation of human inactive X-linked genes in a hamster x human somatic cell hybrid. *Mol. Cell. Biol.* **12**,3819-3826

20. Hansen, R.S., Gartler, S.M. (1990) 5-azacytidine induced reactivation of the human X chromosome-linked PGK1 gene is associated with a large region of cytosine demethylation in the 5' CpG island. *Proc. Natl. Acad. Sci. USA* **87**,4171-4178

21. Hansen, R.S., Canfield, T.K., Fjeld, A.D., Gartler, S.M. (1996) Role of late replication timing in the silencing of X-linked genes. *Hum. Mol. Genet.* **5**,1345-1353

22. Feng, Y., Zhang, F., Lokey, L.K., Chastain, J.L., Lakkis, L., Eberhart, D., Warren, S.T. (1995) Translational suppression by trinucleotide repeat expansion at FMR1. *Science* **268**,731-734.

23. Lowrey, C.H. and Nienhuis, A.W. (1993) Treatment with azacitidine of patients with end stage β-thalassemia. *New Engl. J. Med.* **329**,845-848.

24. Neitzel H. (1986) A routine method for the establishment of permanent growing lymphoblastoid cell lines. *Hum. Genet.* **73**,320-326.

25. Chomczynski P. and Sacchi N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**,156-159.

26. Muellenbach, R., Lagoda, P.J.L. and Welter, C. (1989) An efficient salt-chloroform extraction of DNA from blood and tissues. *Trends Genet.* **5**,391.

27. Nakahori, Y., Knight, S.J., Holland, J., Schwartz, C., Roche, A., Tarleton, J., Wong, S., Flint, T.J., Froster-Iskenius, U., Bentley, D., Davies, K.E., Hirst, M.C. (1991) Molecular heterogeneity of the fragile X syndrome. *Nucleic Acids Res.* **19**,4355-4359.

28. Willemsen, R., Mohkamsing, S., de Vries, B., Devys, D., van den Ouweland, A., Mandel, J.L., Galjaard, H., Oostra, B.A. (1995) Rapid antibody test for fragile X syndrome. *Lancet* **345**,1147-1148.

Synergy of histone hyperacetylation and DNA demethylation in reactivating *FMR1*

Most fragile X syndrome patients have an expansion of a $(CGG)_n$ sequence with more than 200 repeats (full mutation) in the *FMR1* gene, responsible for this condition. Hypermethylation of the expanded repeat and of the *FMR1* promoter is almost always present and apparently suppresses transcription, resulting in the absence of the FMR1 protein. We recently showed that transcriptional reactivation of *FMR1* full mutations can be achieved by inducing DNA demethylation with 5-azadeoxycytidine. The level of histone acetylation is another important factor in regulating gene expression, therefore we treated lymphoblastoid cell lines of nonmosaic full mutation patients with three drugs capable of inducing histone hyperacetylation. We observed a consistent, although modest, reactivation of the *FMR1* gene with 4-phenylbutyrate, sodium butyrate and trichostatin A, as shown by RT-PCR. However, we report that combining these drugs with 5-azadC results in a 2- to 5-fold increase of *FMR1* mRNA levels obtained with 5-azadC alone, thus showing a marked synergistic effect of histone hyperacetylation and DNA demethylation in the reactivation of *FMR1* full mutations.

Fragile X syndrome is a common cause of inherited mental retardation with minor physical manifestations, due to mutations in the *FMR1* gene located in Xq27.3 (1). Most patients have a mutation in the 5' untranslated region of this gene, consisting of an expanded $(CGG)_n$ sequence with more than 200 repeats (full mutation). Hypermethylation of most cytosines in the CGG repeat and in the upstream CpG island is observed and this correlates with gene silencing and absence of the FMR1 protein.

Rare individuals of normal intelligence were shown to carry a completely or partially unmethylated full mutation and to express the FMR1 protein (2-4). This clearly shows that the absence of the FMR1 protein is the cause of the disease. Given this observation and knowing that the open reading frame of the *FMR1* gene is intact in patients with a full mutation, we investigated whether the gene's activity could be restored *in vitro* by inducing DNA demethylation with 5-azadeoxycytidine (5-azadC) in fragile X lymphoblastoid cell lines. In fact, we found that transcriptional reactivation of *FMR1* full mutations and restoration of protein production can be achieved with 5-azadC treatment (5).

It was recently shown that DNA methylation can cause transcriptional silencing of genes through local deacetylation of histonic proteins (6-8). Actually, transcriptionally active chromatin appears to contain more acetylated core histones than tightly packaged heterochromatin (9-11) and treatment with inhibitors of histone deacetylases can increase gene

expression (12). We had previously shown that these drugs inhibit the cytogenetic expression of the fragile site FRAXA in cultured lymphocytes of patients (13).

Therefore we decided to test whether treatment with histone hyperacetylating drugs could also lead to transcriptional reactivation of the *FMR1* gene in lymphoblastoid cell lines of fragile X patients. Since fragile X patients are often mosaics harbouring several amplified alleles of different lengths, as well as premutations which are still transcriptionally active, we made sure to work with cell lines harbouring only fully expanded alleles. We have used 4-phenylbutyrate (4-PBA) and sodium butyrate (BA), which are reversible inhibitors of histone deacetylases (HDACs) active at millimolar concentrations (14), and trichostatin A (TSA) which is also a reversible HDAC inhibitor already at micromolar concentration (15,16). We also extended our initial observation on 5-azadC by employing different concentrations of the drug, for variable periods of time. Most importantly, we tested if the combined treatment with 5-azadC and hyperacetylating drugs would result in a synergistic or simply additive effect.

Results

Lymphoblastoid cell lines were established from peripheral blood lymphocytes of several male fragile X patients and the size of the $(CGG)_n$ expansion was determined by Southern blot analysis after restriction with HindIII and EagI (data not shown). Only cell lines harbouring a full mutation (i.e. with more than 200 CGG repeats) and without constitutive *FMR1* expression were used in these experiments. More precisely, cell line E3 contains approx. 270-370 CGG repeats, line E4 contains 370-520 repeats, while lines E6 and E7 have between 590 and 710 CGG repeats. All the full mutations were completely methylated at the EagI site tested with the Southern blot. Reverse transcription (RT)-PCR, employing primers for *FMR1* mRNA and *HPRT* mRNA as internal control, was performed in order to confirm the silencing of the *FMR1* gene in the patients' untreated cell lines.

Lymphoblasts were cultured and treated as described below with either a histone deacetylase inhibitor alone or in combination with the demethylating agent 5-azadC. After RNA extraction, its concentration was measured and the same amount was used in an optimized semiquantitative RT-PCR assay to estimate the relative abundance of the *FMR1* and *HPRT* transcripts. It should be immediately noted that, although the same amount of total RNA was employed in each RT reaction, the absolute RNA yield and the number of harvested cells decreased with increasing concentrations of the drugs used in the treatment. This effect on cell viability was much more pronounced with TSA (up to 90-95 % decrease) than with either BA or 4-PBA, depending on

the drug concentration (data not shown). Only a slight decrease was observed with 5-azadC alone (up to 50 %), as previously reported (5). A typical result obtained after performing a radioactive RT-PCR is presented in Figure 1. The 195-base pair (bp) band corresponding to the *FMR1* mRNA is visible in the normal (male) control line (lane 11) and in the treated patients' cell lines (lanes 2-5 and 7-10) but not in the untreated ones (lanes 1 and 6). The 386-bp band specific for the *HPRT* transcript serves as an internal control for the quantification of input RNA. After densitometric analysis the results were normalized for the intensity of the *HPRT* band and ratios were calculated in order to express the *FMR1* reactivation as a fraction of *FMR1* mRNA in the control cell line.
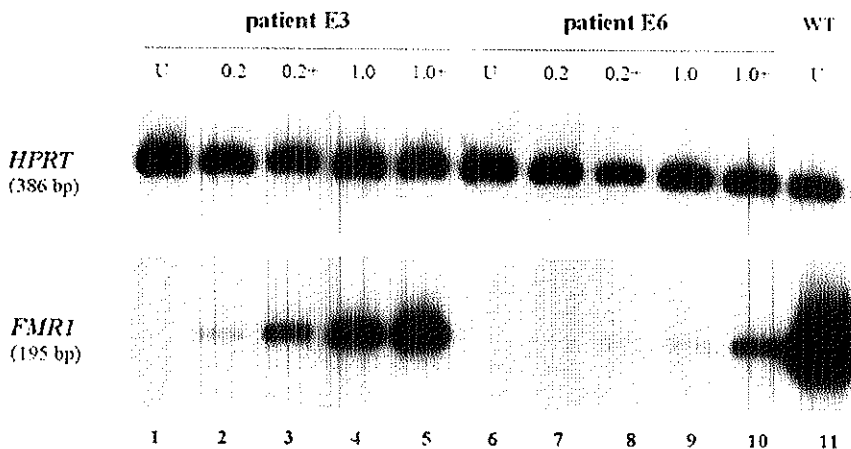
Fig. 1 – Typical RT-PCR experiment illustrating the results of a combined treatment with 5-azadC and 4-PBA. The upper band (386 bp) corresponds to the transcript of the *HPRT* housekeeping gene used as internal control. The lower band (195 bp) is specific for the *FMR1* mRNA and is absent in both patients' untreated cell lines (lanes 1 and 6). "U" stands for "untreated"; "0.2" or "1.0" means treated with either 0.2 or 1.0 µM of 5-azadC; "0.2+" or "1.0+" means treated with either 0.2 or 1.0 µM of 5-azadC and 5 mM 4-PBA. "WT" indicates the lane of the normal control sample.

A modest but consistent reactivation of the *FMR1* gene was obtained by treating the cell lines for 48 hours with various concentrations of 4-PBA (Fig. 2a) and TSA (Fig. 2b) in the millimolar and micromolar range, respectively. Similar results were obtained when the same cell lines were treated with BA in place of 4-PBA (data not shown). 4-PBA was effective at and above 5mM concentration, while TSA-induced reactivation peaked at 300-500 nM and decreased above 1 µM, possibly because of its adverse effect on cell viability. As noted above, the number of cells harvested sharply decreased after treatments with HDAC inhibitors, which are known to have

powerful growth-arresting effects (17,18). Therefore we repeated the treatment with 4-PBA over a 48 or 96-hour period, but washing the cells free of the drug for either 4 or 8 hours every day. Although the *FMR1* to *HPRT* ratio did not change, we observed an increase in cell number and viability (data not shown).
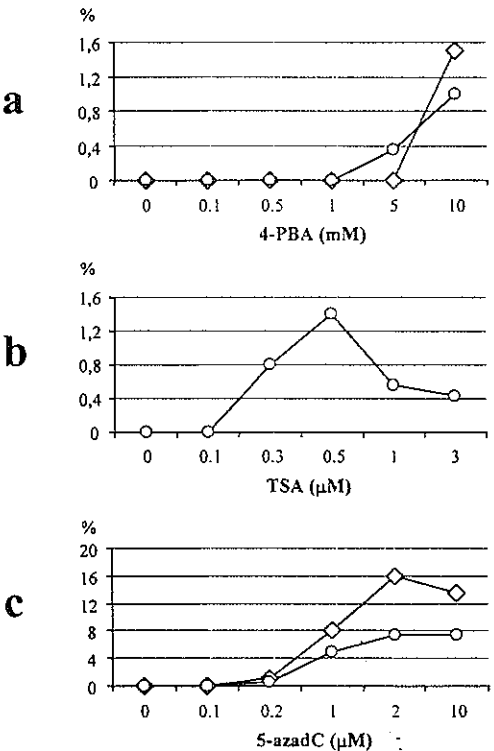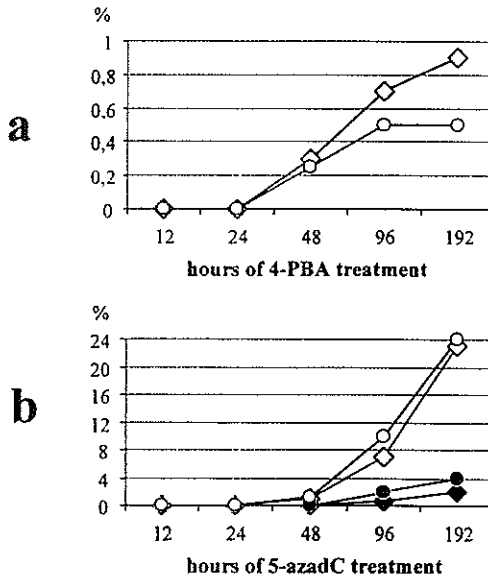

Fig. 2 – Relative levels of *FMR1* reactivation after treatment with various concentrations of 4-PBA (a), TSA (b) or 5-azadC (c). A single dose of 4-PBA and TSA was added for 48 hours, while a daily dose of 5-azadC was added for 72 hours before cell harvesting. Open circles and diamonds correspond to the cell lines of patient E4 and E3, respectively. Reactivation levels are indicated as percentage of wild-type expression. The experiments were replicated with minimal differences for each cell line.

Figure 2c illustrates the results of a similar experiment performed over a 72-hour interval with different concentrations of 5-azadC alone, which shows that 5-azadC has a much greater effect (sometimes up to 10-fold) on *FMR1* expression, compared to the small reactivation obtained with HDAC inhibitors. It is also clear that the effect becomes apparent above 200 nM and reaches a plateau above 2μM concentration. Furthermore, a different reactivation response was

observed between the two patients' cell lines employed in this experiment, which suggests that some full mutations may be reactivated more easily than others.

In order to estimate an optimal incubation time for the drugs' action, the same cell lines were treated with either 4-PBA or 5-azadC for variable time intervals comprised between 12 hours and 8 days (Fig.3).
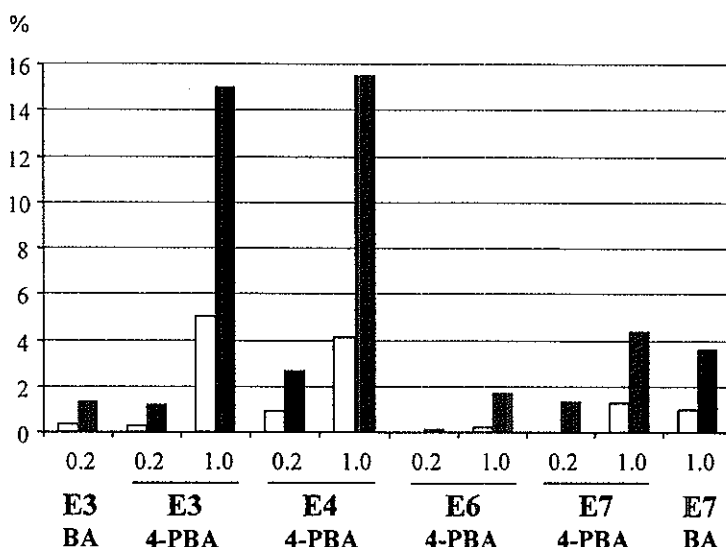

Fig. 3 – Relative levels of *FMRI* reactivation after treatment with 5 mM 4-PBA (a) and 1 μM (open symbols) or 0.2 μM (filled symbols) 5-azadC (b) for 12, 24, 48, 96 and 192 hours, respectively. Circles and diamonds correspond to the cell lines of patient E4 and E3, respectively. 4-PBA was added every 48 hours and 5-azadC daily, while cell culture medium was changed every 48 hours. Reactivation levels are indicated as percentage of wild-type expression. The experiments were replicated with minimal differences for each cell line.

With 4-PBA treatment (5 mM), adding the drug every 48 hours after the medium had been refreshed, the *FMRI* to *HPRT* ratio started to increase after 24 hours and tended to plateau during the last 4 days (Fig. 3a). The effect of the daily addition of 5-azadC was tested at 1 μM or 0.2 μM concentration (Fig. 3b). 5-azadC at 1 μM had again a much stronger effect than 4-PBA (5 mM) and the *FMRI* to *HPRT* ratio clearly increased after 48 hours up to the end of the treatment. The lower concentration (0.2 μM) was suboptimal at all time points.

Combined treatments with 5-azadC and HDAC inhibitors were then performed in order to test whether the respective effects would be simply additive or possibly synergistic. 5-azadC was

added daily for a total of 72 hours at two concentrations (0.2 and 1 µM) with the aim of facilitating the detection of any synergy of treatment when using the suboptimal dose of 0.2 µM. Optimal concentrations of 4-PBA (5 mM), BA (5 mM) and TSA (0.3 µM) were then added for the last 24 hours to one culture while another received just a mock treatment (with adequate volumes of PBS or DMSO). Figure 1 visually depicts the results for two of the cell lines corresponding to patient E3 and E6 treated with 5-azadC and 4-PBA. The complete results of the combined treatments are presented in graphical form in Figure 4.

Fig. 4 – Relative levels of *FMR1* reactivation after combined treatment with 5-azadC (0.2 or 1.0 µM) for a total of 72 hours and 4-PBA or BA (5 mM) for the last 24 hours. Open bars correspond to cell cultures treated with 5-azadC alone, while filled bars represent the cell lines which received the combined treatment. The cell line name and the secondary drug employed (4-PBA or BA) are indicated in boldface beneath the 5-azadC concentration (0.2 or 1 µM). Reactivation levels are indicated as percentage of wild-type expression.

It is readily appreciated that a synergistic increase (2- to 5-fold) in the levels of *FMR1* reactivation is obtained when 4-PBA or BA are added. There is again some variation in the level of *FMR1* reactivation between the different cell lines treated with either 5-azadC alone (open bars) or in combination with the deacetylase inhibitors (filled bars); nevertheless the effect of the combined treatment with 4-PBA and BA is always more than additive. Although TSA did slightly reactivate the *FMR1* gene when used alone (Fig. 2b), it is not clear if the *FMR1* to *HPRT* ratio increased when TSA was added for the last 24 hours to the 5-azadC treated cells

(data not shown). In fact, the extremely low yield of total RNA from the cell cultures treated with TSA did not allow to perform a sufficiently reliable RT-PCR.

Discussion

Though it was known for long time that cytosine methylation in the promoter region of genes often correlates with their transcriptional silencing, it is not clear if just the binding of transcription factors is affected by this epigenetic modification or if first the chromatin structure is altered and then transcription is inhibited (19). Although the first possibility was not excluded, it was recently reported that binding of proteins such as MeCP2 to methylated DNA can recruit a multiprotein complex including HDACs which determines an increased packaging of chromatin (6-8). Therefore it appears that DNA methylation is effectively modifying the local chromatin structure and consequently reducing the potential access of transcription factors to gene promoter regions (20,21).

The amino terminal tails of histones H4, H3 and H2B extend beyond the nucleosomal core (22) and acetylation of lysine K5, K8, K12 and K16 of histone H4, which neutralizes its positive charge, possibly weakens the interaction with DNA and destabilizes internucleosomal binding, thereby facilitating the unfolding of chromatin fibers and the access of the transcriptional machinery (10). HDACs remove precisely those acetyl groups which are added by specific histone acetyltransferases (HATs) to the lysine residues in the amino termini of core histones H3 and H4. Antibodies against specific acetylation sites in histone H4 have been used to show that potentially active euchromatin can be modified at all acetylable lysines, whereas H4 is hypoacetylated in heterochromatin (23,24).

It is worth remembering that the action of HATs and HDACs is directly modulated by a variety of transcriptional activators or repressors interacting in multiprotein complexes, including the Mad-Max heterodimers and members of the nuclear receptor superfamily (11,18,25). DNA methylation seems to be but one of the many triggers of local histone deacetylation and chromatin inactivation; in fact a selective distribution of acetylated histones is also observed in organisms such as yeast (26) and *Drosophila* (27), both of which lack DNA methylation.

As far as the *FMR1* gene is concerned, we have shown that DNA hypermethylation of its promoter associated to a large $(CGG)_n$ expansion (full mutation) is actually causing the transcriptional silencing of the gene, because demethylation with 5-azadC can reactivate its expression (5). We have now extended our previous observations on the effects of 5-azadC and found that concentrations between 1 μM and 10 μM are most effective, although the extent of

FMR1 reactivation is quite variable in different patients' cell lines (Fig. 2c). It is obvious that the reactivation obtained by 5-azadC is only partial and we showed, by immunocytochemistry, that it correlates with the proportion of cells (10-20 %) actually reexpressing the protein (5). In fact, although individual reactivated cells have almost normal levels of FMR1 protein, many cells do not respond to the treatment either because of a low efficiency of 5-azadC incorporation and subsequent passive DNA demethylation or because of the toxicity of the drug (above 10 μM), which causes the formation of DNA adducts by irreversibly binding to the catalytic core of DNA methyltransferases (28). It is tempting to speculate that the difference between individual cell lines in *FMR1* reactivation levels after 5-azadC treatment may correlate with the extent of their promoter hypermethylation. To elucidate this point further investigations such as those of Stoeger *et al.* (29), assessing the methylation of individual CpG sites in the *FMR1* promoter before and after various drug treatments, will be required. A schematic representation of the chromatin modification possibly caused by the *FMR1* promoter hypermethylation is shown in Figure 5, which adapts for the fragile X locus the model recently proposed by several groups (6-8).
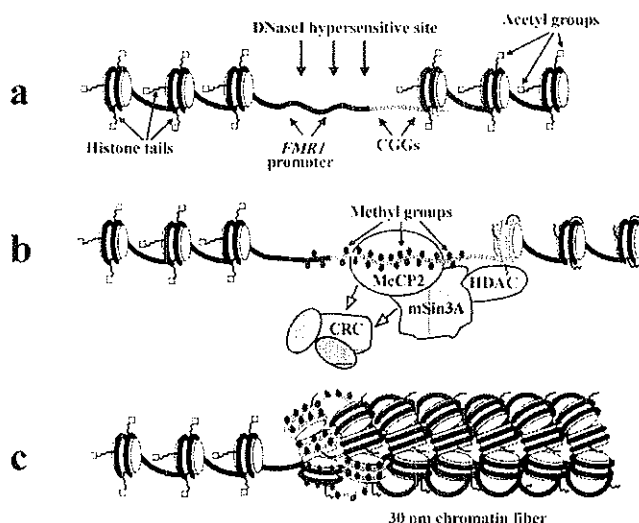
Fig. 5 – Schematic representation of the chromatin structure at the *FMR1* promoter. **Upper panel** - In the normal situation the active gene has an open chromatin with spaced nucleosomes and acetylated histone tails. The promoter is free of nucleosomes and the DNaseI hypersensitive site is shown. **Middle panel** - When the CGG repeat is expanded (shaded in gray), it becomes hypermethylated and then the binding of MeCP2 and the recruitment of a multiprotein histone deacetylating complex rapidly leads to a more packaged and less accessible chromatin structure (**lower panel**), causing the inactivation of the *FMR1* gene. DNA hypermethylation could also lead to chromatin inactivation without directly affecting histone acetylation, possibly through a chromatin remodelling complex (CRC) interacting with the adaptor protein mSin3A, with MeCP2 (open arrowheads) or yet other proteins binding methylated DNA.

It is clear from the diagram that inhibition of HDACs should allow the unfolding of the chromatinic fiber and the repositioning of nucleosomes, possibly determining the reappearance of the DNaseI hypersensitive site, which is normally present in the promoter region (30). In fact, we have observed reactivation of *FMR1* expression after treatment with 4-PBA, BA and TSA alone, although at levels less than 2 % of those observed in normal cells (Fig. 2a and 2b). This effect is presumably due to a global hyperacetylation of histones H4 and H3 that may increase the transcription of several genes, including the fully mutated *FMR1*. Coffee *et al.* (31) confirmed that the 5' region of the *FMR1* gene is associated with acetylated histones H3 and H4 in cells from normal individuals, but acetylation is reduced in cells from fragile X patients. Treatment of fragile X lymphoblasts with 5-azadC resulted in reassociation of acetylated histones H3 and H4 with the *FMR1* promoter and transcriptional reactivation. However treatment with TSA (0.3 µM for 24 hours), which led to almost complete acetylation of histone H4, but little acetylation of histone H3, did not cause any reactivation of *FMR1* expression (31). Different experimental and cell culture conditions may explain why we were able to detect *FMR1* reactivation after TSA treatment, although the extreme antiproliferative effect of TSA in particular may severely interfere with *FMR1* expression, as discussed below.

There could be several explanations for the modest level of *FMR1* reactivation by HDAC inhibitors, the first one being the persistence of DNA methylation which triggers the chromatin remodeling cascade shown in Fig. 5. Whereas treatment with 5-azadC actually prevents the assembly of the multisubunit repressor complex, HDAC inhibitors are blocking the chromatin inactivating machinery at a later step. This possibility is in line with the results of Cameron *et al.* (32) who have combined 5-azadC and TSA treatment in cancer cell lines with aberrant promoter hypermethylation of some tumour-suppressor genes and concluded that DNA methylation is dominant over histone acetylation in determining the silencing at those loci.

Secondly, 4-PBA, BA and especially TSA arrest the cell cycle at multiple points (17,33,34) and induce cell differentiation (35), while *FMR1* expression is increased in dividing cells (36) with the notable exception of neurons. Therefore, the potential *FMR1* reactivation may be contrasted by the growth-arresting properties of these drugs. In fact, it appears that a pulsed treatment with 4-PBA is more effective in our *in vitro* model (data not shown) and also *in vivo*, when boosting the levels of gamma globin chain expression (37).

Finally, it now seems likely that chromatin structure and nucleosome positioning can be modified by several redundant chromatin remodeling complexes (38), some of which do not affect histone acetylation and harbour a SWI2/SNF2 ATPase family protein (39). If the

promoter hypermethylation could increase chromatin packaging and cause *FMR1* gene silencing without directly acting on histone acetylation (21,40,41), as indicated by the open arrows in Fig. 5, then the HDAC inhibitors would only partially relieve the silencing effects of DNA methylation.

The time curve experiment illustrated in Fig. 3 demonstrates that a continued treatment over several days can increase the effect of 5-azadC, possibly by allowing a more complete DNA demethylation. This result retrospectively supports our previous choice for a 7-day treatment with 5-azadC (5) and suggests that a passive demethylation process is taking place in the treated cells. On the other hand, 4-PBA action is apparent already after 24 hours, while in the last 96 hours the growth arrest induced by the treatment may contrast the *FMR1* reactivation, as discussed above. TSA has been shown by others (32) to be active already within 6 hours from its addition and its antiproliferative effect is apparent after 12 hours.

The combined effect of the 5-azadC and 4-PBA or BA treatment is clearly synergistic and not simply additive (Fig.4). The fact that the effects of the DNA demethylating and histone hyperacetylating treatments reinforce each other confirms that DNA methylation and histone deacetylation occur in sequence in the same pathway which leads to the silencing of the fully mutated *FMR1* gene. It is worth pointing out that the extent of *FMR1* reactivation seems to correlate with the size of the full mutation, in fact cell lines E6 and E7 harbour approx. 600 to 700 CGG repeats and respond less to the treatments than lines E3 (270-370 repeats) and E4 (370-530 repeats). Indeed, a larger CGG expansion will also contain more methylated cytosines triggering the local inactivation of chromatin and a longer time might be required to achieve a thorough demethylation with the 5-azadC treatment.

Our results are also in agreement with the recent observation of a synergistic interaction of 5-azadC and TSA in regulating the re-expression of silenced tumor-suppressor genes (32). Although we were not able to unequivocally prove that TSA can also reinforce the reactivation of the *FMR1* gene induced by 5-azadC, it is possible that exposing the cell cultures for shorter times (less than 24 hours) to this potent drug may eventually uncover its synergic effect. However, McCaffrey *et al.* (42) have also noted that TSA was less effective than arginine-butyrate in increasing the expression of the gamma globin gene *in vitro*, probably because of its pronounced cytotoxic effect. In fact TSA does not appear to be a suitable drug for *in vivo* applications because of its extremely strong side effects which confine its use to *in vitro* experiments.

On the other hand, both BA and 4-PBA have been used *in vivo* to boost gamma globin expression in patients with either sickle cell disease or beta thalassemia (37,43). 4-PBA, which is a new FDA approved drug, was also employed to increase the expression of the variant $\Delta F508$ CFTR protein in patients with cystic fibrosis (44) and was effective in reactivating an analog of the *ALD* gene in KO mice (45).

The synergistic effect of 5-azadC and 4-PBA or BA suggests the possibility of using lower dosages of both drugs in order to obtain appreciable *FMR1* reactivation levels also *in vivo*. 5-azadC has already been employed *in vivo* (46), again in order to increase the expression of gamma globin gene expression in beta thalassemia patients, although its potential mutagenic effect limits the duration of a therapeutic application. It remains to be established if after a short term treatment with 5-azadC leading to partial demethylation and reactivation of the *FMR1* gene, the patients' full mutation can remain stably active as it happens for the rare cases of normal males with an unmethylated full mutation (3), or if the abnormal $(CGG)_n$ expansion would inevitably cause remethylation of the promoter and gene silencing. Further *in vitro* investigations will hopefully indicate whether a long term treatment with histone hyperacetylating drugs as 4-PBA, may help not only in obtaining the initial *FMR1* reactivation but also in maintaining such an active status.

Materials and Methods

Cell culture and stock solutions

Lymphoblastoid cell lines were established by Epstein-Barr virus (EBV) transformation from peripheral blood lymphocytes of male fragile X patients and normal male controls. Cells were grown in RPMI1640 medium with 10% fetal calf serum and penicillin/streptomycin at 37°C, in closed flasks without CO_2. Sodium butyrate (BA, Sigma) and 4-phenylbutyric acid (4-PBA, Fluka) were resuspended in sterile water to a concentration of 500 mM and stored at -80°C in aliquots. Trichostatin A (TSA, Sigma) was resuspended in DMSO to a concentration of 10 mM and stored at -20°C in small aliquots and a working solution of 1mM was prepared just before use. Finally, a 10 mM stock solution of 5-azadeoxycytidine (5-azadC, Sigma) was prepared in sterile water and stored at -80°C in aliquots.

Treatments with histone hyperacetylating drugs

Cells were counted, split and seeded at the initial concentration of $2.5-3 \times 10^5$ cells/ml in a total volume of 10 ml per flask. A single dose of either 4-PBA, BA or TSA was added to the flasks and was thoroughly resuspended. A control flask for each patient cell line was left untreated or received a mock treatment with a comparable volume of DMSO in the case of the TSA treatment. Cells were harvested after 48 hours and the RNA extracted. When a time curve was performed for 4-PBA, the optimal concentration of the drug (5 mM) was added to the cells and RNA was extracted from a treated flask at each time point (12, 24, 48, 96 and 192 hours, respectively). Cells were washed and resuspended in fresh medium with the appropriate drug concentration every 48 hours.

Treatments with DNA demethylating drug
Cells were counted, split and seeded at the initial concentration of $2.5-3 \times 10^5$ cells/ml in a total volume of 10 ml per flask. A daily dose of 5-azadC was added to the flasks and was thoroughly resuspended, while a control flask was left untreated. Cells were harvested after 72 hours of treatment when various micromolar concentrations were tested, while the time curve was performed by adding 1 μM of 5-azadC every day and extracting the RNA at the same time points used with 4-PBA. The medium was changed every 48 hours.

Combined treatments
Cells were counted and seeded as described above. 5-azadC (200 nM or 1 μM) was added daily for 72 hours, while an optimal dose of either 4-PBA (5 mM), BA (5 mM) or TSA (300 nM) was added during the last 24 hours. Control cultures were grown with either 5-azadC alone or without any treatment for the total duration of the experiment. After 72 hours, RNA was extracted from all flasks at the same time.

RT-PCR analysis
Total RNA was extracted with the single-step acid phenol method, using RNAzol B (Tel-Test, Inc.). cDNA synthesis was carried out at 37° C for 120 min in a total volume of 40 μl with 180 Units of MoMLV-RT and its buffer (Gibco-BRL), 1 mM DTT, 10 units of RNase inhibitor (Promega), 0.8 mM each dNTP, and preincubating 5 μg of total RNA with 0.6 μg of random hexamers (Pharmacia) at 65° C for 10 min. Expression of *FMR1* specific mRNA was determined by RT-PCR using primers K9 (GTA TGG TAC CAT TTG TTT TTG TG, exon 3) and K6 (CAT CAT CAG TCA CAT AGC TTT TTT C, exon 4) that yield a 195 bp product. As internal control for the amplification and to normalize the quantity of input RNA, we employed primers 244 (AAT TAT GGA CAG GAC TGA ACG TC, exon 2 and 3) and 243 (CGT GGG TCC TTT TCA CCA GCA AG, exon 7) of the housekeeping gene *HPRT*, which yield a 386 bp product. Test reactions were initially performed with 18, 21 and 24 cycles and also with variable amounts of input cDNA from a control cell line to ensure that the PCR reaction would be still in the linear range (data not shown). 24 cycles of amplification were further employed (94° 1'- 55° 1'- 72° 2') in a total volume of 10 μl with 1 mM $MgCl_2$, 250 μM each dNTP, 0.5 mM spermidine, 0.75 Units Taq polymerase (Gibco-BRL), 0.1 μl of α-^{32}P dCTP (Amersham-Pharmacia, 10 mCi/ml) and 0.7 μl of cDNA. A large excess of primers (5 picomoles each) was added to the reaction in order to maintain it in the exponential phase at 24 cycles. PCR products were separated on a denaturing 6% polyacrylamide gel with 7 M urea run at 70 Watts for 2 hours on a S2 apparatus (BRL). The gel was fixed with 10 % methanol/ 10% acetic acid solution, dried for at least 40 minutes at 80°C with a vacuum pump and exposed to autoradiographic films (Fuji New RX) for a variable time interval of 3 hours to 6 days, in order to have several exposures to scan for densitometric analysis.

Densitometric analysis
Autoradiographic films were scanned at high resolution (300 dpi) and densitometric analysis was carried out with the Windows version of the NIH Image software available from Scion Corporation (version Beta 3b, July 1998). Band intensity of the *FMR1* transcript was normalized to the intensity of the *HPRT* signal in order to account for differences in input RNA or RT-PCR efficiency. Ratios between the intensity of the *FMR1*-specific signals in the treated and control cell lines were then calculated in order to express the reactivation levels as a fraction of the wild-type expression.

References

1. Oostra, B.A. and Willems, P.J. (1995) A fragile gene. *Bioessays,* 17, 941-947.

2. Hagerman, R.J., Hull, C.E., Safanda, J.F., Carpenter, I., Staley, L.W., O'Connor, R.A., Seydel, C., Mazzocco, M.M., Snow, K., Thibodeau, S.N., Kuhl, D., Nelson, D.L., Caskey, C.T. and Taylor, A.K. (1994) High functioning fragile X males: demonstration of an unmethylated fully expanded FMR-1 mutation associated with protein expression. *Am. J. Med. Genet.,* 51, 298-308.

3. Smeets, H.J., Smits, A.P., Verheij, C.E., Theelen, J.P., Willemsen, R., van de Burgt, I., Hoogeveen, A.T., Oosterwijk, J.C. and Oostra, B.A. (1995) Normal phenotype in two brothers with a full FMR1 mutation. *Hum. Mol. Genet.,* 4, 2103-2108.

4. de Vries, B.B., Jansen, C.C., Duits, A.A., Verheij, C., Willemsen, R., van Hemel, J.O., van den Ouweland, A.M., Niermeijer, M.F., Oostra, B.A. and Halley, D.J. (1996) Variable FMR1 gene methylation of large expansions leads to variable phenotype in three males from one fragile X family. *J. Med. Genet.,* 33, 1007-1010.

5. Chiurazzi, P., Pomponi, M.G., Willemsen, R., Oostra, B.A. and Neri, G. (1998) In vitro reactivation of the FMR1 gene involved in fragile X syndrome. *Hum. Mol. Genet.,* 7, 109-113.

6. Eden, S., Hashimshony, T., Keshet, I., Cedar, H. and Thorne, A.W. (1998) DNA methylation models histone acetylation. *Nature,* 394, 842.

7. Jones, P.L., Veenstra, G.J., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J. and Wolffe, A.P. (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature Genet.,* 19, 187-191.

8. Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N. and Bird, A. (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature,* 393, 386-389.

9. Davie, J.R. and Candido, E.P. (1978) Acetylated histone H4 is preferentially associated with template-active chromatin. *Proc. Natl. Acad. Sci. USA,* 75, 3574-3577.

10. Grunstein, M. (1997) Histone acetylation in chromatin structure and transcription. *Nature,* 389, 349-352.

11. Kuo, M.H. and Allis, C.D. (1998) Roles of histone acetyltransferases and deacetylases in gene regulation. *Bioessays,* 20, 615-626.

12. Lea, M.A. and Randolph, V.M. (1998) Induction of reporter gene expression by inhibitors of histone deacetylase. *Anticancer Res.,* 18, 2717-2722.

13. Pomponi, M.G. and Neri, G. (1994) Butyrate and acetyl-carnitine inhibit the cytogenetic expression of the fragile X in vitro. *Am. J. Med. Genet.,* 51, 447-450.

14. Cousens, L.S., Gallwitz, D. and Alberts, B.M. (1979) Different accessibilities in chromatin to histone acetylase. *J. Biol. Chem.,* 254, 1716-1723.

15. Yoshida, M., Kijima, M., Akita, M. and Beppu, T. (1990) Potent and specific inhibition of mammalian histone deacetylase both in vivo and in vitro by trichostatin A. *J. Biol. Chem.,* 265, 17174-17179.

16. Yoshida, M., Horinouchi, S. and Beppu, T. (1995) Trichostatin A and trapoxin: novel chemical probes for the role of histone acetylation in chromatin structure and function. *Bioessays,* 17, 423-430.

17. Yoshida, M. and Beppu, T. (1988) Reversible arrest of proliferation of rat 3Y1 fibroblasts in both the G1 and G2 phases by trichostatin A. *Exp. Cell. Res.,* 177, 122-131.

18. Kouzarides, T. (1999) Histone acetylases and deacetylases in cell proliferation. *Curr. Opin. Genet. Dev.,* 9, 40-48.

19. Kass, S.U., Landsberger, N. and Wolffe, A.P. (1997) DNA methylation directs a time-dependent repression of transcription initiation. *Curr. Biol.,* 7, 157-165.

20. Razin, A. (1998) CpG methylation, chromatin structure and gene silencing-a three-way connection. *Embo J.,* 17, 4905-4908.

21. Ng, H.H. and Bird, A. (1999) DNA methylation and chromatin modification. *Curr. Opin. Genet. Dev.,* 9, 158-163.

22. Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. and Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature,* 389, 251-260.

23. Jeppesen, P. and Turner, B.M. (1993) The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell,* 74, 281-289.

24. Johnson, C.A., O'Neill, L.P., Mitchell, A. and Turner, B.M. (1998) Distinctive patterns of histone H4 acetylation are associated with defined sequence elements within both heterochromatic and euchromatic regions of the human genome. *Nucleic Acids Res.,* 26, 994-1001.

25. David, G., Alland, L., Hong, S.H., Wong, C.W., DePinho, R.A. and Dejean, A. (1998) Histone deacetylase associated with mSin3A mediates repression by the acute promyelocytic leukemia-associated PLZF protein. *Oncogene*, 16, 2549-2556.

26. Clarke, D.J., O'Neill, L.P. and Turner, B.M. (1993) Selective use of H4 acetylation sites in the yeast Saccharomyces cerevisiae. *Biochem. J.*, 294, 557-561.

27. Turner, B.M., Birley, A.J. and Lavender, J. (1992) Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in Drosophila polytene nuclei. *Cell*, 69, 375-384.

28. Juttermann, R., Li, E. and Jaenisch, R. (1994) Toxicity of 5-aza-2'-deoxycytidine to mammalian cells is mediated primarily by covalent trapping of DNA methyltransferase rather than DNA demethylation. *Proc. Natl. Acad. Sci. U S A*, 91, 11797-11801.

29. Stoger, R., Kajimura, T.M., Brown, W.T. and Laird, C.D. (1997) Epigenetic variation illustrated by DNA methylation patterns of the fragile-X gene FMR1. *Hum. Mol. Genet.*, 6, 1791-1801.

30. Luo, S., Robinson, J.C., Reiss, A.L. and Migeon, B.R. (1993) DNA methylation of the fragile X locus in somatic and germ cells during fetal development: relevance to the fragile X syndrome and X inactivation. *Somat. Cell. Mol. Genet.*, 19, 393-404.

31. Coffee, B., Zhang, F., Warren, S.T. and Reines, D. (1999) Acetylated histones are associated with FMR1 in normal but not fragile X-syndrome cells. *Nature Genet.*, 22, 98-101.

32. Cameron, E.E., Bachman, K.E., Myohanen, S., Herman, J.G. and Baylin, S.B. (1999) Synergy of demethylation and histone deacetylase inhibition in the re- expression of genes silenced in cancer. *Nature Genet.*, 21, 103-107.

33. Charollais, R.H., Buquet, C. and Mester, J. (1990) Butyrate blocks the accumulation of CDC2 mRNA in late G1 phase but inhibits both the early and late G1 progression in chemically transformed mouse fibroblasts BP-A31. *J. Cell. Physiol.*, 145, 46-52.

34. Vaziri, C., Stice, L. and Faller, D.V. (1998) Butyrate-induced G1 arrest results from p21-independent disruption of retinoblastoma protein-mediated signals. *Cell Growth Differ.*, 9, 465-474.

35. Newmark, H.L., Lupton, J.R. and Young, C.W. (1994) Butyrate as a differentiating agent: pharmacokinetics, analogues and current status. *Cancer Lett.*, 78, 1-5.

36. Devys, D., Lutz, Y., Rouyer, N., Bellocq, J.P. and Mandel, J.L. (1993) The FMR-1 protein is cytoplasmic, most abundant in neurons and appears normal in carriers of a fragile X premutation. *Nature Genet.*, 4, 335-340.

37. Atweh, G.F., Sutton, M., Nassif, I., Boosalis, V., Dover, G.J., Wallenstein, S., Wright, E., McMahon, L., Stamatoyannopoulos, G., Faller, D.V. and Perrine, S.P. (1999) Sustained induction of fetal hemoglobin by pulse butyrate therapy in sickle cell disease. *Blood*, 93, 1790-1797.

38. Kornberg, R.D. and Lorch, Y. (1999) Chromatin-modifying and -remodeling complexes. *Curr. Opin. Genet. Dev.*, 9, 148-151.

39. Pollard, K.J. and Peterson, C.L. (1998) Chromatin remodeling: a marriage between two families? *Bioessays*, 20, 771-780.

40. Wong, C.W. and Privalsky, M.L. (1998) Transcriptional repression by the SMRT-mSin3 corepressor: multiple interactions, multiple mechanisms, and a potential role for TFIIB. *Mol. Cell. Biol.*, 18, 5500-5510.

41. Jeddeloh, J.A., Stokes, T.L. and Richards, E.J. (1999) Maintenance of genomic methylation requires a SWI2/SNF2-like protein. *Nature Genet.*, 22, 94-97.

42. McCaffrey, P.G., Newsome, D.A., Fibach, E., Yoshida, M. and Su, M.S. (1997) Induction of gamma-globin by histone deacetylase inhibitors. *Blood*, 90, 2075-2083.

43. Collins, A.F., Pearson, H.A., Giardina, P., McDonagh, K.T., Brusilow, S.W. and Dover, G.J. (1995) Oral sodium phenylbutyrate therapy in homozygous beta thalassemia: a clinical trial. *Blood*, 85, 43-49.

44. Rubenstein, R.C. and Zeitlin, P.L. (1998) A pilot clinical trial of oral sodium 4-phenylbutyrate (Buphenyl) in deltaF508-homozygous cystic fibrosis patients: partial restoration of nasal epithelial CFTR function. *Am. J. Respir. Crit. Care Med.*, 157, 484-490.

45. Kemp, S., Wei, H.M., Lu, J.F., Braiterman, L.T., McGuinness, M.C., Moser, A.B., Watkins, P.A. and Smith, K.D. (1998) Gene redundancy and pharmacological gene therapy: implications for X- linked adrenoleukodystrophy. *Nature Med.*, 4, 1261-1268.

46. Lowrey, C.H. and Nienhuis, A.W. (1993) Brief report: treatment with azacitidine of patients with end-stage beta-thalassemia. *N. Engl. J. Med.*, 329, 845-848.

FMR1 promoter demethylation and transcriptional reactivation

In fragile X syndrome, hypermethylation of an expanded CGG repeat and the *FMR1* promoter leads to transcriptional silencing of the gene. Demethylation of the *FMR1* promoter by treatment with 5-azadeoxycytidine results in reactivation of the gene. We employed bisulphite sequencing to detect the methylation status of individual CpG sites in the promoter of the *FMR1* gene. The entire promoter region upstream of the CGG repeat was tested in lymphoblastoid cell lines of fragile X males with full mutations of different sizes, before and after DNA demethylating treatment with 5-azadeoxycytidine.

The demethylation dynamics was explored by testing the same cell line at various time points and we observed that individual clones are either completely demethylated or not, with few relevant exceptions. Also, it appears that the *FMR1* promoter is demethylated more rapidly when the adjacent CGG expansion (full mutation) is relatively smaller. We have indirect evidence that the full mutation itself remains at least partially methylated in many clones with a demethylated promoter. This may explain the apparent discrepancy between the extent of promoter demethylation and the levels of *FMR1* transcriptional reactivation obtained with 5-azadeoxycytidine.

The highly polymorphic CGG trinucleotide repeat which is located in the 5' untranslated region (UTR) of the fragile X mental retardation gene *FMR1*, is associated with the disease phenotype when the allele carries more than 2000 triplets (full mutation) [Verkerk et al., 1991; Fu et al., 1991]. The abnormally expanded CGG repeat of fragile X chromosomes is almost invariably transcriptionally silent [Pieretti et al., 1991] and its inactive state has been correlated with the abnormal methylation of the CpG sites of the expanded repeat and – more importantly – of the upstream promoter region of the *FMR1* gene [Hansen et al., 1992; Sutcliffe et al., 1992]. Rare individuals of normal intelligence were shown to carry a completely or partially unmethylated full mutation and to express the FMR1 protein (FMRP), clearly indicating that the absence of FMRP is the cause of the disease [Hagerman et al., 1994; Smeets et al., 1995]. We sought to reactivate the fully mutated *FMR1* gene *in vitro* by demethylating the DNA of lymphoblastoid cell lines established from fragile X males [Chiurazzi et al., 1998]. Treatment with micromolar concentrations of 5-aza-2-deoxycytidine (5-azadC) for 7 days allowed reprise of transcriptional activity, as judged by RT-PCR, and also FMR1 protein could be observed in a proportion of cells by immunocytochemistry [Chiurazzi et al., 1998]. DNA demethylation of the *FMR1* promoter was assayed by Southern blotting after digestion with the methylation-sensitive restriction enzymes *Eag*I, *Bss*HII and *Sac*II, and only with *Sac*II we could confirm that a partial

demethylation had occurred. Furthermore, the total amount of *FMR1* mRNA (estimated by semiquantitative RT-PCR) did not exceed 20% of wild-type levels, and the percentage of cells expressing FMRP was even lower [Chiurazzi et al., 1998]. It seemed that only a proportion of the cells responded to the treatment, demethylating the *FMR1* promoter and resuming transcription. To confirm this hypothesis we needed a technique that allowed testing the methylation status of individual cells. More recently, experiments were carried out to potentiate 5-azadC-induced reactivation with the addition of the histone hyperacetylating drugs sodium butyrate (BA) and 4-phenylbutyrate (4-PBA) and we observed a strong synergistic effect of these compounds [Chiurazzi et al., 1999], thus confirming that CpG cytosine methylation and histone deacetylation cooperate in silencing chromatinic domains [Cameron et al., 1999; Razin, 1998]. In that set of experiments, it was also noted that cell lines harbouring a shorter CGG expansion (still in the full mutation range e.g. between 250 and 400 repeats) could be reactivated more strongly than those with larger full mutations (above 600 repeats) and it was hypothesized that it may be more difficult to demethylate the *FMR1* promoter in the presence of longer CGG repeat tracts. In order to test if only a proportion or the majority of treated cells undergo DNA demethylation and whether the extent of demethylation is related to the length of the CGG expansion, we set up a bisulphite-sequencing protocol similar to that of Stoeger et al.[1997], but with a new reverse primer that allowed testing of an extra 200 bp. After treating fragile X cell lines with 1 μM 5-azadC for different periods of time, we harvested cells and modified the extracted DNA with sodium bisulphite, as described by Panagopoulos et al. [1999]. More than 10 independent PCR amplifications were then performed with primers specific for the bisulphite-modified upper strand of the *FMR1* promoter, in a fashion similar to Stoeger et al. [1997]. After cloning the pooled PCR products and sequencing of several clones in both directions we were able to reconstruct the 'epigenotype' of individual cells for the whole *FMR1* promoter, including all the *in vivo* footprints first reported by Schwemmle et al. [1997] and the transcription start site until the CGG repeat.

Results

The entire *FMR1* promoter region is illustrated in Figure 1 from position – 641 to + 79, relative to the A of the first codon ATG (base numbering corresponding to Genbank sequence L29074 is given between square brackets in the text). The start of transcription (– 264 [13698]) and the beginning of intron 1 (+52 [14013]) are indicated below the sequence, while the CGG repeat lies in the 5' UTR roughly 140 bp after the transcription start site and 70 bp before the first ATG.

1F

13321 ccaaaccaaa ccaaaccaga ccagacaccc cctcccgcgg aatcccagag aggccgaact
SacII

2F   1

13381 gggataaccg gatgcatttg atttcccacg ccactgagtg cacct ctgca gaa atg g gcg

2 3   4   5   6

13441 ttctggccct cgcgaggcag tgcgacctgt caccgccctt cagccttccc gccctccacc
NruI

7 8   9   10   11 12 13   14   15   16   17

13501 aagcccgcgc acgcccggcc cgcgcgtctg tctttcgacc cggcaccc cggccg gttccc
EagI

IV (α-PAL)   **III (Sp1)**
18 19   20 21 22   23   24

13561 agcagcgcgc at gcgcgcgc tcccaggcca ctt gaagaga gaggcgggg ccgaggggct
BssHII   BssHII   3R

II (Sp1-like)   **I (USF1&2)**
25 26   27   28   29   30

13621 gagcccgcgg ggggagggaa cagcgttgat cacgtgacgt ggtttcagtg tttacacccg
SacII

31   32   33   34   35   36   37

13681 cagcgggccg ggggttc   ggc ctcagtcagg cgctcagctc cgtttcggtt tcacttccgg
mRNA...

38   39   40 41   42 43 44   45 46   47 48   49   50

13741 tggagggccg cctctgagcg ggcggcgggc cgacggcgag cgcgggcggc ggcggtgacg

51 52

13801 gaggcgccgc tgccaggggg cgtgcggcag cg [cgg cgg cgg cgg cgg cgg cgg cgg cgg c
5-aR

13861 gg agg cgg cgg cgg cgg cgg cgg cgg cgg cgg] ctgggcct cgagcgcccg cagcccacct

13921 cttggggggcg ggctcccggc gctagcaggg ctgaagagaa g ATG gag gag ctg gtg gtg g
exon 1...

13981 aa gtg cgg ggc tcc aat ggc gct ttc tac aag gtacttgg ctctagggca ggccccatct
intron 1...

Figure 1 – Promoter region of the *FMR1* gene (sequence numbering from GenBank L29074). Primers employed by us and others in bisulphite-PCR are indicated above or below the sequence for forward (1-F, 2-F) or reverse (3-R, 5-aR), respectively. Individual CpG sites between primers 2-F and 5-aR are numbered from 1 to 52. Restriction sites of methylation sensitive enzymes *SacII, NruI, EagI* and *BssHII* are indicated below the sequence, as well as the start of transcription (mRNA), translation (ATG) and of intron 1. The CGG repeat is within square brackets, downstream of primer 5-aR. The four footprints reported by Schwemmle et al. [1997] and Drouin et al. [1997] are highlighted in italics on the sequence with their putative binding transcription factors indicated in brackets [Kumari and Usdin, 2001].

Primers employed by us and by Stoeger et al.[1997] are underlined in the corresponding positions (1-F, 2-F, 3-R and 5-aR) and their sequences, specific for the bisulphite-modified upper strand, are reported in the Materials and Methods. We employed a semi-nested PCR protocol with a first round of amplification using primers 1-F and 5-aR, followed by a second round using primers 2-F and 5-aR, which gives a final PCR product of 413 bp that was cloned and eventually sequenced in both directions.

As indicated in Figure 1, this PCR product contains 52 potentially methylated CpG sites spanning the four *in vivo* footprints (I to IV) reported by Schwemmle et al. [1997], as well as the recognition sites of the *Nru*I, *Eag*I, *Bss*HII and *Sac*II restriction enzymes, commonly employed to assay the methylation status of the *FMR1* promoter. It is worth pointing out that Stoeger et al. [1997] could analyze only the first 22 CpG sites, comprised within primers 2-F and 3-R, that span the *Eag*I (CpG 16-17) as well as the two *Bss*HII sites (CpG 18-19 and 20-21). The two *Bss*HII sites overlap with footprint IV, which corresponds to the α-PAL/NRF1 protein binding site. The binding of α-PAL transcription factor has been recently shown to contribute to at least 50% of the total *FMR1* promoter activity [Kumari and Usdin, 2001]. The rest of the activity appears to be driven by the binding of transcription factors USF1 and USF2 to footprint I, coinciding with an E-box sequence (sometimes referred to as the c-*MYC* binding site) [Kumari and Usdin, 2001]. On the contrary, binding of Sp1 to recognition sequences in footprint III and II does not seem to contribute in a relevant manner to *FMR1* transcription. Again, primers employed by Stoeger et al. [1997] did not allow testing this portion of promoter sequence including footprints III, II and I. When the DNA of a normal male was tested after bisulphite conversion, all the cytosines were deaminated into uracyl and showed up as thymines in the sequence (data not shown), confirming the efficiency of our bisulphite transformation protocol as well as the observation of Stoeger et al.[1997] that the CpG island spanning the *FMR1* promoter is not methylated on the active X chromosome.

Figures 2, 3 and 4 illustrate the results of our bisulphite sequencing experiments on three cell lines from different fragile X boys with approximately 250 (line E3), 500 (line S1) and between 330 and 700 (line S5) CGG repeats, respectively. On top of each diagram three black bars mark positions 19-21, 23 and 28 that correspond to CpG sites included in footprint IV, III and I (from left to right). RNA was also extracted at the same time points and RT-PCR performed in order to verify the transcriptional status of *FMR1*. RT-PCR results are indicated in brackets next to each experiment diagram. Figure 2a shows the sequences of nine clones from the untreated line E3 (approx. 250 CGG repeats), which indicate the almost complete methylation of the *FMR1*
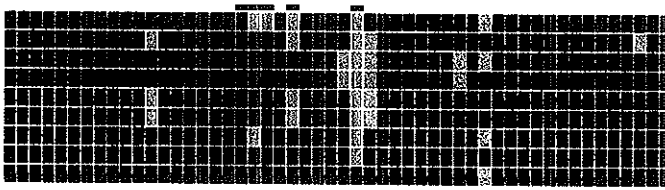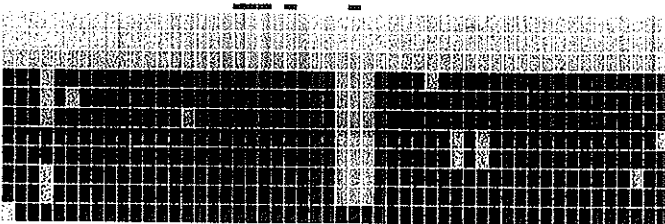
a) E3 before treatment (negative RT-PCR).
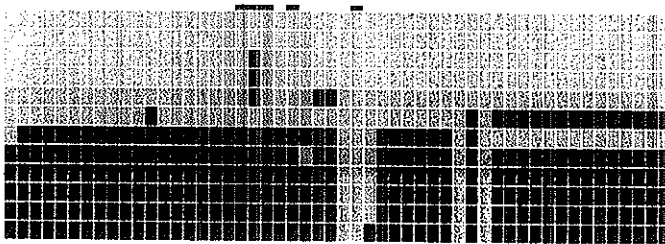
**b)** E3 after 3 days of treatment with 1 µM 5-azadC (RT-PCR positive).
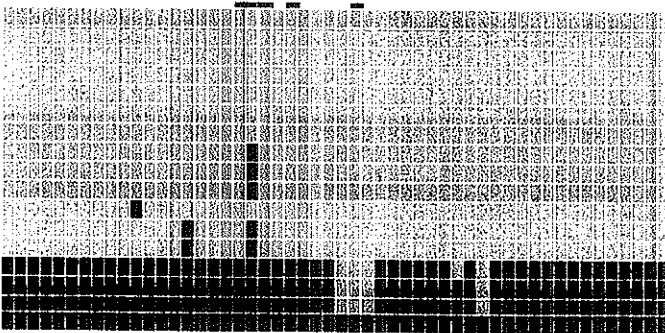
c) E3 after 8 days of treatment with 1 µM 5-azadC (RT-PCR positive).

d) E3 after 5 weeks of suspending the 5-azadC treatment (negative RT-PCR).

Figure 2 – *FMR1* promoter region of line E3 (full mutation of approx. 250 CGGs). Every line corresponds to bisulphite sequencing of an individual clone. From left to right, black or gray positions correspond to the 52 methylated or unmethylated CpG sites, respectively. Sites 19-21, 23 and 28 are marked by a black bar on the top of the diagram and are included in footprints IV, III and I, respectively.

a) S1 before treatment (negative RT-PCR).

b) S1 after 3 days of treatment with 1 μM 5-azadC (RT-PCR positive).

c) S1 after 8 days of treatment with 1 μM 5-azadC (RT-PCR positive).
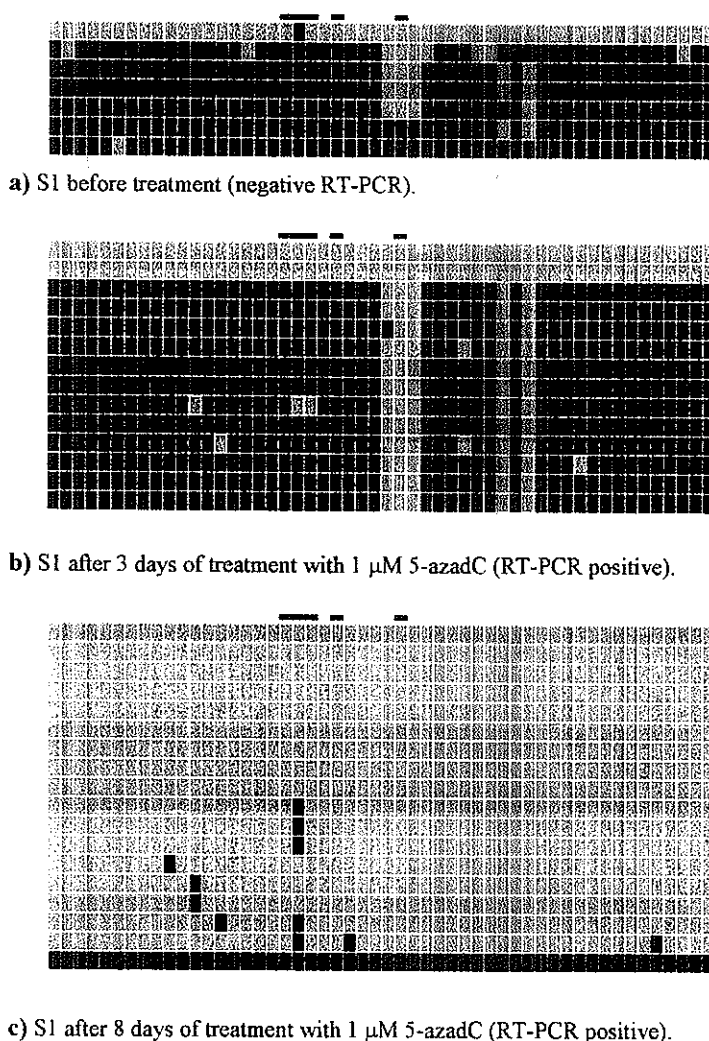
Figure 3 – *FMR1* promoter region of line S1 (full mutation of approx. 500 CGGs). Every line corresponds to bisulphite sequencing of an individual clone. From left to right, black or gray positions correspond to the 52 methylated or unmethylated CpG sites, respectively. Sites 19-21, 23 and 28 are marked by a black bar on the top of the diagram and are included in footprints IV, III and I, respectively.

promoter with the notable exception of position 28, that is unmethylated in all but one clone. RT-PCR was negative before 5-azadC treatment. After 3 days of 1 μM 5-azadC treatment (Fig. 2b), RT-PCR became positive and almost all 52 CpGs tested were now unmethylated in half of the clones (6 out of 12), with the partial exception of position 20. Surprisingly, in the rest of the

otherwise methylated clones, positions 27-29 as well as positions 36 and 38 were also unmethylated. After 8 days of 1 μM 5-azadC treatment (Fig. 2c), an even larger proportion (13 out of 17) clones were unmethylated, again with the above mentioned exceptions of position 20 (relatively methylated) and positions 27-29, 36 and 38 (relatively unmethylated). 5-azadC treatment was stopped after day 8, but we continued to grow the cells and to collect RNA (and DNA) every week until the RT-PCR became negative again after 5 weeks. Figure 2d illustrates the *FMR1* promoter methylation status at that time point: 3 out of 11 clones were unmethylated, the other 8 being almost completely methylated again with the exception of positions 27-29.

Figure 3a shows that, with the exception of one clone, also the *FMR1* promoter of line S1 (approx. 500 CGGs) was methylated before 5-azadC treatment (again positions 27-29 being relatively unmethylated). After 3 days of 1 μM 5-azadC treatment (Fig. 3b), RT-PCR was positive but, in contrast to what we observed with line E3 (± 250 CGGs), only 2 out 14 clones were demethylated. Positions 27-29, as well as 36 and 38, again were the exception. After 8 days of 5-azadC treatment (Fig. 3c), all but one clone were demethylated (only position 20 was still methylated in 5 out of 17 clones). Figure 4a illustrates five methylated clones of cell line S5 (between 330 and 700 CGG repeats) before treatment. RT-PCR was negative, although CpG sites 27-29 are unmethylated. After 8 days of 5-azadC (Fig. 4b), 15 out of 21 clones were almost completely demethylated (with the exception of position 20 in two clones). The other 6 clones remained methylated with the 'usual' exception of CpG sites 27-29 and 38.

In order to test the methylation status of the CGG repeat itself at the various time points, we could not perform bisulphite sequencing across the full mutation, therefore we had to go back to Southern blotting after digestion with *Pst*I and methylation-sensitive enzymes *Mcr*BC or *Fnu*4HI (data not shown). As probe we employed the 562 bp *Xho*I-*Pst*I fragment of the *FMR1* promoter, located immediately downstream of the CGG repeat. *Mcr*BC endonuclease cleaves DNA containing two half-sites $(G/A)^mC$ with a methylated cytosine, separated by 40-80 base pairs, and will not act upon unmethylated DNA. Therefore, *Mcr*BC will cut the *Pst*I fragments containing a methylated full mutation and will not cut normal unmethylated samples (see Burman et al. [1999]). After the various 5-azadC treatments, we observed that *Mcr*BC continues to cleave the DNA, indicating that the CGG repeat is still at least partially methylated in many cells. The same indication was obtained when using *Fnu*4HI in place of *Mcr*BC (data not shown). *Fnu*4HI cleaves only unmethylated DNA and serves as complementary test to *Mcr*BC (see Hansen et al. [1992]). Unlike bisulphite sequencing, digestion with methylation-sensitive enzymes and Southern blotting cannot provide information on individual cells; however, we
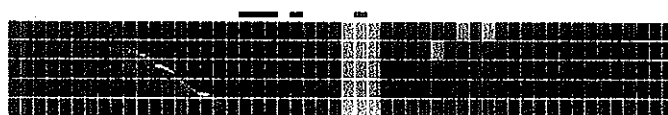
a) S5 before treatment (negative RT-PCR).
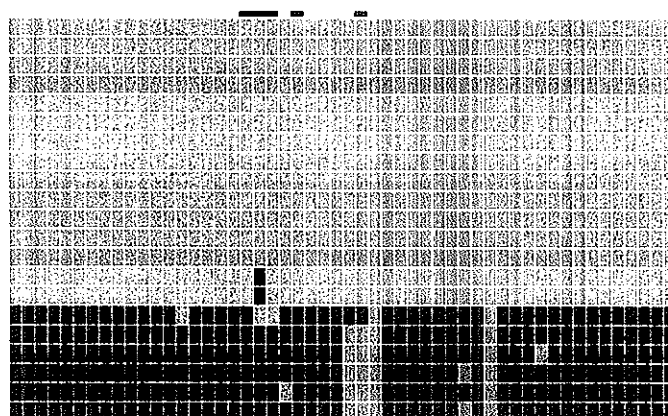
b) S5 after 8 days of treatment with 1 μM 5-azadC (RT-PCR positive).

Figure 4 – *FMR1* promoter region of line S5 (full mutation with multiple bands of approx. 330-700 CGGs). Every line corresponds to bisulphite sequencing of an individual clone. From left to right, black or gray positions correspond to the 52 methylated or unmethylated CpG sites, respectively. Sites 19-21, 23 and 28 are marked by a black bar on the top of the diagram and are included in footprints IV, III and I, respectively.

deduce from our observations that the CGG expansion is still partly methylated in many cells with a demethylated promoter. This may explain the apparent discrepancy between the large extent of promoter demethylation detected e.g. after 8 days of 5-azadC treatment, and the relatively lower levels of *FMR1* transcriptional reactivation.

Discussion

The structural features of the *FMR1* promoter – high GC content, presence of numerous Sp1 sites, lack of a TATA box – are typical of an housekeeping gene [Drouin et al., 1997], as confirmed by the ubiquitous distribution of FMR1 protein in most adult and fetal tissues [Devys et al., 1993; Khandjian et al., 1995]. Genomic footprint analysis of the *FMR1* promoter revealed four sites of protein-DNA interaction [Schwemmle et al., 1997; Drouin et al., 1997], where

different transcription factors bind (see Figure 1). The footprints are usually absent in fragile X males, as the *FMR1* gene is inactive.

Kumari and Usdin [2001] were able to show that in murine brain α-PAL/NRF1 binds footprint IV (CpG sites 18-22), while USF1 and USF2 (but not c-MYC, nor CREB) bind footprint I (CpG site 28). Footprints III and II, corresponding to GC-boxes of Sp1 (CpG site 23) and Sp1-like type (no CpG sites), are probably bound by Sp3 (but not Sp1) in lymphoid cells and contribute less to the activity of the *FMR1* promoter [Kumari and Usdin, 2001]. In transient transfection experiments conducted in rat PC12 cells, constructs with a mutation in either the α-PAL or the E-box site had just 20% of promoter activity left, while constructs with both mutated sites had no residual activity at all. Constructs with a mutation in the GC-box of footprint III and II had 55% and 80% residual promoter activity, respectively [Kumari and Usdin, 2001].

The abnormal methylation of the CGG repeat and of the upstream promoter was soon associated with silencing of the *FMR1* gene in fragile X patients [Oberle' et al., 1991; Pieretti et al., 1991; Sutcliffe et al., 1992], as well as on the inactive X chromosome in normal females [Hansen et al., 1992]. Transfection of reporter constructs confirmed that methylation abolishes the activity of the *FMR1* promoter [Hwu et al., 1993; Sandberg and Schalling, 1997]. Furthermore, we demonstrated *in vitro* the possibility of reactivating methylated fragile X mutations after a demethylating treatment with 5-azadC [Chiurazzi et al., 1998]. 5-azadC is incorporated in DNA and becomes covalently bound with the maintenance DNA methyltransferase (DNMT1) as this enzyme tries to methylate the trivalent nitrogen in position 5 of the pyrimidine ring [Juttermann et al., 1994]. DNMT1 molecules are irreversibly blocked as adducts on the DNA [Jackson-Grusby et al., 1997], thus allowing passive demethylation to take place as cells divide. We also showed that addition of histone hyperacetylating drugs strongly potentiates the effect of 5-azadC in reactivating the *FMR1* gene [Chiurazzi et al., 1999], as it happens for several cancer genes [Cameron et al., 1999]. Our results support the notion that DNA methylation leads to histone deacetylation [Coffee et al., 1999] and chromatin silencing, via binding of methylcytosine binding proteins such as MeCP2 and recruiting of a multiprotein complex [Razin, 1998].

However, histone hyperacetylating drugs alone had almost no effect, indicating that methylation must be silencing the *FMR1* gene also in alternative ways. In fact, Kumari and Usdin [2001] proved that methylation of CpG sites 18-20 (Fig. 1) reduces binding of α-PAL to 45%, while methylation of CpG site 28 in the E-box reduces binding of USF1 and USF2 by about 20% (to 80%). Therefore, we now have evidence that the fully mutated *FMR1* gene is silenced in at least two ways: 1) by indirectly causing a local chromatin modification that reduces the access of the

transcriptional machinery to the promoter region; 2) by direct interference with the binding of transcription factors.

De novo methylation of fragile X full mutations is probably established during early embryogenesis [Malter et al., 1997], possibly in an effort to stabilize the expanded CGG repeat that acts like a parasitic sequence element from a genomic perspective [Bestor and Tycko, 1996]. Rare individuals with an unmethylated full mutation that presumably escaped the *de novo* methylation have been described [Smeets et al., 1995; de Vries et al., 1996]. They have no global impairment of DNA methylation [Burman et al., 1999] and are intellectually normal or 'high functioning' males. Schwemmle [1999] confirmed the presence of footprints in one such individual, suggesting normal transcription. In fact, it seems that the unmethylated CGG expansion does not impede transcription *per se*, neither in the premutation [Tassone et al., 2000a] nor in the full mutation range [Tassone et al., 2000b]. Actually, it seems that *FMR1* mRNA levels are directly correlated with the size of the CGG expansion in the effort of compensating a relative translational deficiency [Feng et al., 1995; Kenneson et al, 2001]. Enough FMR1 protein must be anyhow produced to account for the normal intellect of males with unmethylated full mutations.

With the present series of experiments we intended to refine our understanding of the reactivation process by analyzing the methylation status of each individual CpG site in the *FMR1* promoter before and after 5-azadC treatment. We optimized a technique of bisulphite sequencing similar to that of Stoeger et al. [1997] and of Genc et al. [2000], though we used different primer sets allowing detection of more than 400 bp of sequence immediately upstream of the CGG repeat (Figure 1), including all four footprints of transcription factors. Figures 2a, 3a and 4a show that the various fragile X cell lines tested had an almost completely methylated promoter without any treatment. A relevant exception are CpG sites 27, 28 and 29, that overlap with footprint I and are often unmethylated. We propose that this reflects a continued binding of transcription factors USF1 and USF2, that are relatively unaffected by methylation anyhow [Kumari and Usdin, 2001] and seem to have still access to their recognition sequence, though they are unable to drive transcription alone. Furthermore, we also suggest that the methylation of the expanded CGG repeat and the binding of proteins such as the p20 CGGBP described by Muller-Hartmann et al. [2000] or MeCP2 itself, may be very important in determining the transcriptional status of the *FMR1* gene. However, under other circumstances e.g. in the case of a lung tumor with a premutation described by de Graaff et al. [1995], abundant FMRP was

produced and although the promoter seemed partly methylated, the premutation itself may well have been unmethylated.

After 3 days of 5-azadC treatment (Fig. 2b and 3b) we can observe that a proportion of clones (i.e. of cells) have a completely demethylated promoter, while others are still methylated though positions 27-29 are now almost always demethylated too. Therefore, 5-azadC does not induce a partial or random demethylation of just some CpG sites, but a sequential and thorough demethylation of the entire promoter, explaining why we observed a clear reappearance of FMRP in a proportion of cells with immunocytochemistry [Chiurazzi et al., 1998]. Comparing the proportion of demethylated clones between cell lines E3 and S1 (Fig. 2b and 3b) - with approx. 250 and 500 CGG repeats, respectively - we confirm our previous suggestion that cell lines with a longer full mutation are more resistant to 5-azadC induced demethylation [Chiurazzi et al., 1999].

After 8 days of 5-azadC treatment (Fig. 2c, 3c and 4b) a large proportion (70% to 90%) of the clones are demethylated, though position 20 is often still methylated and may thus prevent α-PAL from binding footprint IV, reducing transcriptional efficiency of more than half [Kumari and Usdin, 2001]. We previously reported *FMR1* mRNA levels of approx. 20-25% relative to a wild-type cell line after 8 days of 1 µM 5-azadC treatment [Chiurazzi et al., 1999]. Taking into account the observations of Tassone et al. [2000b], that unmethylated full mutations are actually being transcribed at even 5- to 8-fold levels relative to normal cells, we have to suspect that actually only 5% of our treated cells were actually reactivated. This fits well with the limited number of cells positive for FMRP that we observed [Chiurazzi et al., 1998], but leaves us with the problem of explaining the discrepancy between the exceeding number of clones with a demethylated promoter illustrated in Fig. 2c, 3c and 4b and the lower number of reactivated cells, as judged by mRNA and protein levels. Considering our experiments with *Mcr*BC and *Fnu*4HI, we propose that a large proportion of the cells with a demethylated promoter still harbour a methylated CGG repeat that somehow still makes transcription very difficult, though it may allow transcription factors binding. Footprinting, quantitative RT-PCR and RNA-FISH studies (this latter allowing the study of individual cells) may help to settle this issue.

However, we now suspect that our demethylating protocol does not reproduce a situation similar to that of males with an unmethylated full mutation. In fact, most of our cells may succeed in demethylating the upstream promoter, while the expanded CGG repeat itself may 'resist' demethylation by attracting the few DNMT1 molecules still active in the nucleus.

Finally, we must consider the methylation status of line E3 after 5 weeks without 5-azadC (Fig. 2d) when the RT-PCR became again negative. Although most of the clones are methylated again (with the exception of positions 27-29), 3 out of 11 are still completely demethylated. We suggest that in these clones the CGG repeat may be methylated and – as stated above – effectively suppress transcription. The reactivated clones may have been remethylated by a *de novo* methyltransferase, starting from the CGG repeat and eventually spreading to the promoter. However, there is no evidence that *de novo* methyltransferases are expressed in adult cells after embryogenesis and we suspect that many of the reactivated cells eventually died by apoptosis because of the many DNMT1-DNA adducts in the first 2-3 weeks after suspending 5-azadC treatment.

In conclusion, a complex picture emerges from our present study: CGG expansion triggers methylation, that dynamically represses transcription by directly interfering with the binding of some transcription factors or indirectly modifying chromatin structure, though probably not causing a dramatic chromatin compaction – with the possible exception of the CGG repeat itself.

Materials and Methods

Cell culture and stock solution

Lymphoblastoid cell lines were established by Epstein-Barr virus (EBV) transformation from peripheral blood lymphocytes of male fragile X patients and normal male controls. Cells were grown in RPMI1640 medium with 10% fetal calf serum and penicillin/streptomycin at 37°C with 5% CO_2.
A 10 mM stock solution of 5-azadeoxycytidine (5-azadC, Sigma) was prepared in sterile water and stored at -80°C in aliquots.

Treatments with DNA demethylating drug

Cells were counted, split and seeded at the initial concentration of $2.5-3 \times 10^5$ cells/ml in a total volume of 30 ml per flask. Immediately before use, 3 µl of the 10 mM 5-azadC stock solution was thawed and added daily to the flasks and thoroughly resuspended (final concentration 1 µM), while a control flask was left untreated. Cells were harvested after 3 and 8 days from the start of treatment. In the case of line E3, cells were grown for a further 5 weeks after discontinuing 5-azadC treatment and pellets were prepared every week. Cell pellets for DNA and RNA extraction were prepared after every harvesting. The medium was changed every 48 hours.

RT-PCR analysis

Total RNA was extracted with the single-step acid phenol method, using RNAzol B (Tel-Test, Inc.). cDNA synthesis was carried out at 37° C for 120 min in a total volume of 40 µl with 180 Units of MoMLV-RT and its buffer (Gibco-BRL), 1 mM DTT, 10 units of RNase inhibitor (Promega), 0.8 mM each dNTP, and preincubating 5 µg of total RNA with 0.6 µg of random hexamers (Pharmacia) at 65° C for 10 min. Expression of *FMR1* specific mRNA was determined by RT-PCR using primers K9 (GTA TGG TAC CAT TTG TTT TTG TG, exon 3) and K6 (CAT CAT CAG TCA CAT AGC TTT TTT C, exon 4) that yield a 195 bp product. As internal control for the amplification, we employed primers 244 (AAT TAT GGA CAG GAC TGA ACG TC, exon 2 and 3) and 243 (CGT GGG TCC TTT TCA CCA GCA AG, exon 7) of the housekeeping gene *HPRT*, which yield a 386 bp product. PCR products were separated on a 1.2 % agarose gel and detected with ethidium bromide under UV illumination.

Bisulphite treatment

A sodium bisulphite 5M solution (pH 5) was prepared dissolving 9.5 grams of powder in 12 ml of distilled water, 3.5 ml NaOH 2N and 2.5 ml hydrochinone 1M. After bringing to pH 5 with NaOH, water is added to the final volume of 20 ml. 5 μg of genomic DNA are diluted in 50 μl of water, denatured at 95°C 10 min, then incubated 30 min at 37°C with 1.5 μl of NaOH 10N (final conc. 0.3N). Then 310 μl of sodium bisulphite 5M are added, 2.5 μl of hydrochinone 0.1M and 136 μl of water (final vol. 500 μl).

A layer of mineral oil is added and the mixture incubates overnight at 55°C. The bisulphite treatment will deaminate only unmethylated cytosines into uracil, therefore discriminating them from the methylated ones which resist deamination. The day after DNA is purified with Promega columns, again denatured with NaOH 0.3N for 15 min at 37°C, precipitated with ammonium acetate pH 7 (final conc. 3M) and 4 volumes of cold ethanol. After 30 min at -80°C, the sample is centrifuged and the pellet washed with 70% ethanol and resuspended in 50 μl of water.

PCR amplification, cloning and sequencing

0.1 μl out of 50μl bisulphite treated DNA are added to each PCR reaction. 12 independent reactions of 25 μl were set up in parallel from each sample to minimize the effects of eventual PCR artifacts. 30 cycles (30 sec 95°C – 30 sec 63°C – 30 sec 72°C) of PCR were made with 10% DMSO, 200 μM dNTPs, 1 unit of Taq polymerase, 2.5 mM MgCl$_2$, and 1 μl of 10 μM stock of primers 1F - 5' GGA ATT TTA GAG AGG TC/TG AAT TGG G 3' and 5aR - 5' CAC ACC CCC TAA CAA C 3'.

One microliter of the first reaction is then used to start a seminested PCR with 35 cycles (30 sec 95°C – 30 sec 60°C – 1 min 72°C), without DMSO, with 3 mM MgCl$_2$, and 1 μl of 10 μM stock of primers 2F - 5' GTT ATT GAG TGT ATT TTT GTA GAA ATG GG 3' and 5aR - 5' CAC ACC CCC TAA CAA C 3'. All 12 reactions of each sample were then pooled, partly evaporated, separated on a 1.2 % agarose gel, and the bands recovered with the GIBCO-BRL Concert Rapid Gel extraction system (11456-019).

Then the PCR products were ligated with the TOPO TA cloning kit by Invitrogen (K460001), bacterial cells included in the kit transformed and plated. After overnight incubation and growth, colonies were picked and eventually minipreps done with GIBCO-BRL Concert rapid plasmid miniprep system (11453-016). A first PCR screen with primers 2F and 5aR, PCR products cleaned on spin columns and 3 to 5 μl out of 30 were used for the sequencing reaction. We used the Amersham-Pharmacia Thermosequenase Dye Terminator kit (US79765) with primers M13F and M13R of the TOPO TA vector, and sequenced on an ABI 373 machine.

References

Bestor TH, Tycko B (1996): Creation of genomic methylation patterns. Nat Genet 12:363-367.

Burman RW, Yates PA, Green LD, Jacky PB, Turker MS, Popovich BW (1999): Hypomethylation of an expanded FMR1 allele is not associated with a global DNA methylation defect. Am J Hum Genet 65:1375-1386.

Cameron EE, Bachman KE, Myohanen S, Herman JG, Baylin SB (1999): Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. Nature Genet 21:103-107.

Chiurazzi P, Pomponi MG, Pietrobono R, Bakker CE, Neri G, Oostra BA (1999): Synergistic effect of histone hyperacetylation and DNA demethylation in the reactivation of the FMR1 gene. Hum Mol Genet 8:2317-2323.

Chiurazzi P, Pomponi MG, Willemsen R, Oostra BA, Neri G (1998): In vitro reactivation of the FMR1 gene involved in fragile X syndrome. Hum Mol Genet 7:109-113.

Coffee B, Zhang F, Warren ST, Reines D (1999): Acetylated histones are associated with FMR1 in normal but not fragile X-syndrome cells. Nature Genet 22: 98-101.

de Graaff E, Willemsen R, Zhong N, de Die-Smulders CE, Brown WT, Freling G, Oostra B (1995): Instability of the CGG repeat and expression of the FMR1 protein in a male fragile X patient with a lung tumor. Am J Hum Genet 57:609-618.

de Vries BB, Jansen CC, Duits AA, Verheij C, Willemsen R, van Hemel JO, van den Ouweland AM, Niermeijer MF, Oostra BA, Halley DJ (1996): Variable FMR1 gene methylation of large expansions leads to variable phenotype in three males from one fragile X family. J Med Genet 33:1007-1010.

Devys D, Lutz Y, Rouyer N, Belloq JP, Mandel JL (1993): The FMR-1 protein is cytoplasmic, most abundant in neurons and appears normal in carriers of a fragile X premutation. Nat Genet 4:335-340.

Drouin R, Angers M, Dallaire N, Rose TM, Khandjian W, Rousseau F (1997): Structural and functional characterization of the human FMR1 promoter reveals similarities with the hnRNP-A2 promoter region. Hum Mol Genet 6:2051-2560.

Feng Y, Zhang F, Lokey LK, Chastain JL, Lakkis L, Eberhart D, Warren ST (1995): Translational suppression by trinucleotide repeat expansion at FMR1. Science 268:731-734.

Fu YH, Kuhl DP, Pizzuti A, Pieretti M, Sutcliffe JS, Richards S, Verkerk AJ, Holden JJ, 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.

Gene B, Muller-Hartmann H, Zeschnigk M, Deissler H, Schmitz B, Majewski F, von Gontard A, Doerfler W (2000): Methylation mosaicism of 5'-(CGG)(n)-3' repeats in fragile X, premutation and normal individuals. Nucleic Acids Res 28:2141-2152.

Hagerman RJ, Hull CE, Safanda JF, Carpenter I, Staley LW, O'Connor RA, Seydel C, Mazzocco M, Snow K, Thibodeau SN, Kuhl D, Nelson DL, Caskey CT, Taylor AK (1994): High functioning fragile X males: demonstration of an unmethylated fully expanded FMR1 mutation associated with protein expression. Am J Med Genet 51:298-308.

Hansen RS, Gartler SM, Scott CR, Chen SH, Laird CD (1992): Methylation analysis of CGG sites in the CpG island of the human FMR1 gene. Hum Mol Genet 1:571-578.

Hwu WL, Lee YM, Lee SC, Wang TR (1993): In vitro DNA methylation inhibits FMR-1 promoter. Biochem Biophys Res Commun 193:324-329.

Jackson-Grusby L, Laird PW, Magge SN, Moeller BJ, Jaenisch R (1997): Mutagenicity of 5-aza-2'-deoxycytidine is mediated by the mammalian DNA methyltransferase. Proc Natl Acad Sci U S A 94:4681-4685.

Juttermann R, Li E, Jaenisch R (1994): Toxicity of 5-aza-2'-deoxycytidine to mammalian cells is mediated primarily by covalent trapping of DNA methyltransferase rather than DNA demethylation. Proc Natl Acad Sci U S A 91:11797-11801.

Kenneson A, Zhang F, Hagedorn K, Warren ST (2001): Reduced FMRP and increased *FMR1* transcription is proportionally associated with CGG repeat number in intermediate-length and premutation carriers. Hum Mol Genet 10:1449-1454.

Khandjian EW, Fortin A, Thibodeau A, Tremblay S, Cote F, Devys D, Mandel JL, Rousseau F (1995): A heterogeneous set of FMR1 proteins is widely distributed in mouse tissues and is modulated in cell culture. Hum Mol Genet 4:783-789.

Kumari D, Usdin K (2001): Interaction of the transcription factors USF1, USF2, and α-Pal/Nrf-1 with the FMR1 promoter. Implications for Fragile X mental retardation syndrome. J Biol Chem 276:4357-4364.

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

Muller-Hartmann H, Deissler H, Naumann F, Schmitz B, Schroer J, Doerfler W (2000): The human 20-kDa 5'-(CGG)(n)-3'-binding protein is targeted to the nucleus and affects the activity of the FMR1 promoter. J Biol Chem 275:6447-6452.

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 methylation in fragile X syndrome. Science 252:1097-1102.

Panagopoulos I, Lassen C, Kristoffersson U, Aman P (1999): A methylation PCR approach for detection of fragile X syndrome. Hum Mutat 14:71-79.

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.

Razin A (1998): CpG methylation, chromatin structure and gene silencing-a three-way connection. Embo J 17:4905-4908.

Sandberg G, Schalling M (1997): Effect of in vitro promoter methylation and CGG repeat expansion on FMR-1 expression. Nucleic Acids Res 25:2883-2887.

Schwemmle S (1999): In vivo footprinting analysis of the FMR1 gene: proposals concerning gene regulation in high-functioning males. Am J Med Genet 84:266-267.

Schwemmle S, de Graaff E, Deissler H, Glaser D, Wohrle D, Kennerknecht I, Just W, Oostra BA, Dorfler W, Vogel W, Steinbach P (1997): Characterization of FMR1 promoter elements by in vivo-footprinting analysis. Am J Hum Genet 60:1354-1362.

Smeets HJ, Smits AP, Verheij CE, Theelen JP, Willemsen R, van de Burgt I, Hoogeveen AT, Oosterwijk JC, Oostra BA (1995): Normal phenotype in two brothers with a full FMR1 mutation. Hum Mol Genet 4:2103-2108.

Stoeger R, Kajimura TM, Brown WT, Laird CD (1997): Epigenetic variation illustrated by DNA methylation patterns of the fragile-X gene FMR1. Hum Mol Genet 6:1791-801.

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

Tassone F, Hagerman RJ, Taylor AK, Gane LW, Godfrey TE, Hagerman PJ (2000a): Elevated levels of FMR1 mRNA in carrier males: a new mechanism of involvement in the fragile-X syndrome. Am J Hum Genet 66:6-15.

Tassone F, Hagerman RJ, Loesch DZ, Luchiewicz A, Taylor AK, Hagerman PJ (2000b): Fragile X males with unmethylated, full mutation trinucleotide repeat expansions have elevated levels of FMR1 messenger RNA. Am J Med Genet 94:232-236.

Verkerk AJMH, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DPA, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang F, Eussen BE, van Ommen GJB, Blonden LAJ, Riggins GJ, Chastain JL, Kunst CB, Galjaard H, Caskey CT, Nelson DL DL, Oostra BA, Warren ST (1991): Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting lenght variation in fragile X syndrome. Cell 65:905-914.

Chapter 4

General Discussion

：

4.1 - Origin of fragile X mutations

The three papers of Chapter 2 report on our methodological and experimental efforts in the characterization of fragile X chromosomes. Since 1992, when Richards et al. [1992] first reported on the detection of founder chromosomes in fragile X syndrome, it was confirmed that unstable repeats behave in a non-Mendelian fashion. Fragile X syndrome does not affect the life span of patients, but it drastically reduces their chances of reproduction, thus leading to a low fitness indeed. Since the prevalence of fragile X is relatively high (1:4000 to 1:6000 males [Turner et al., 1996; de Vries et al., 1997]), one would expect a high mutation rate in order to maintain that prevalence in spite of the continuous loss of fragile X chromosomes through affected males. We could have expected to find new mutations on almost every haplotype background found in the normal population: however, the data showed us that most fragile X mutations are linked to only 3 to 4 haplotypes. Understanding how this could be, meant speculating on the multistep nature of the fragile X mutational mechanism. Fu et al. [1991], shortly after the cloning of the *FMR1* gene, suggested that the unstable nature of the mutation could account for the so-called Sherman paradox of increasing penetrance: in fragile X pedigrees there was a clear increase in the number of affected children from carrier mothers in subsequent generations.

It was realized that the gradual lengthening of the CGG array in the premutation range also increased the risk of 'transition' to the pathological full mutation. What had not been immediately appreciated was that the initial events of destabilization of a wild-type allele were relatively rare, definitely more rare than the subsequent steps (see Fig. 1b, Chapter 2.1).

A multistep mutational pathway is characteristic of all trinucleotide repeat disorders [Cummings & Zoghbi, 2000] and can be pictured like a snowball coming down the mountains: first it is light and rolls very slowly (10^{-6}-10^{-7}), as it becomes heavier it gains in speed, becoming faster and faster (10^{-1}-10^{-2}). The average transition rate of the whole process - in the case of fragile X syndrome – is approx. 10^{-4}, accounting for the relatively high prevalence of fragile X [Chiurazzi et al., 1996a]. The rare initial destabilization events will happen on just a few chromosomes and this will result in few haplotypes (those of the original founder chromosomes) being associated with a much larger final number of affected patients (see Chapter 2.2). Also intrinsic in the hypothesis of the multistep pathway is the presence of 'pools' or 'reservoirs' of 'protomutations' (i.e. already destabilized at-risk alleles, but not yet premutations) in the normal population from which more unstable chromosomes can originate. Size of the CGG array alone is not a good predictor of future instability, but when the presence and position of AGG interruptions is taken

into account (imperfections generally stabilizing the array [Hirst et al.,1999]), a better guess can be made [Eichler et al., 1994].

We tested over 130 independent fragile X chromosomes from people living in different parts of Italy [Chiurazzi et al., 1996b] and found that only three DXS548-FRAXAC1 'major' haplotypes (2-A, 7-C and 6-D) are associated with 60% of fragile X chromosomes (Fig. 2, Chapter 2.2). These same 'major' haplotypes are found in different proportions in all fragile X populations of European descent tested to date, with approx. 20-30% of haplotype 7-C and then either more 2-A in Mediterranean people (including two Iranian and one Singhalese patients) and Western Europe, or more 6-D in Scandinavia and Central Europe [Chiurazzi et al., 1996a]. In the Italian fragile X population, another 26% of patients' chromosomes are associated with 5 'minor' haplotypes (3-C, 8-C, 1-A, 7-D, 7-A, 2-C). The distribution of control alleles is very different with over 70% of the normal chromosomes concentrated with haplotype 7-C and another 20% with three minor haplotypes (7-D, 6-D, 6-C).

We also tested the CGG repeat length in 141 control males and found an association between alleles with more than 36 CGGs and haplotype 2-A (the most common fragile X founder chromosome in Italy as well as in Southern Spain [Yolanda de Diego, pers. comm.]), while most normal chromosomes with haplotype 6-D had between 30 and 35 CGGs (Fig. 3, Chapter 2.2). Our observation is in line with the proposal made by Eichler et al. [1996] on the existence of at least two major mutational pathways leading to destabilization of the *FMR1* array: 1) a slow but continuous increase in size in a 'pool' of protomutations linked to chromosome 2-A (or 7-A); 2) a near average size of CGG arrays in the 'reservoirs' of 6-D and 7-C protomutations that expand more abruptly or 'leap-frog' wise [Macpherson et al., 1995], sometimes after losing an AGG interruption.

Let us now briefly consider why fragile X haplotypes are more numerous than those found on wild-type chromosomes i.e. how can we interpret the higher heterozygosity of fragile X vs. wild-type associated haplotypes. First of all, the three 'major' fragile X haplotypes described in Europeans originated on chromosomes also present in the control population: 7-C being the most frequent normal haplotype, 6-D and 2-A being third and fifth most common (Fig.2, Chapter 2.2): most of the heterozygosity excess is due to the presence of the 2-A haplotype, which is absent, for example, in Asian (normal and fragile X) populations. Three protomutations probably just arose on these 'major' founder chromosomes by chance and not in recent times, because they can be found in all European ethnic groups (the more isolated Finns have almost only the 6-D haplotype [Haataja et al., 1994] and the Basques may have no founder fragile X

chromosome at all [Arrieta et al., 1999]). Furthermore, some of the 'minor' haplotypes may actually derive from the major ones because of subsequent recombination between markers (e.g. 7-A may come from 2-A recombined with a 7-C common normal haplotype) or be due to mutation at the microsatellite marker itself (e.g. 1-A may derive from 2-A because of a mutation in DXS548). However, it is also possible that some 'rare' haplotypes (absent in the normal population and only associated with fragile X chromosomes) may reflect a moment of 'concurrent' instability of both the CGG repeat and microsatellite markers [Morris et al., 1995], due to a defect in DNA mismatch repair that caused the initial destabilization of the CGG array and generated the unusual 'rare' haplotype. With the next generation, however, the "mutator" phenotype would disappear in the heterozygous offspring and would not be detectable in the fragile X patients with that 'rare' haplotype. It's also worth noting that some of the rare fragile X haplotypes may originate from very small population 'pools' and may be lost over time, either because of genetic drift or because protomutations do not always increase in length but can also decrease and revert to the normal range.

In Chapter 2.3 we reported on the distribution of DXS548-FRAXAC1-CGG haplotypes at the *FMR1* locus in a sample of 85 normal chromosomes from the Bamileke ethnic group of Cameroon [Chiurazzi et al., 1996c] and compared them with European (Italian) and Asian (Chinese) control populations. Africans display a strikingly higher heterozygosity at all tested loci, in spite of the relatively small size of the sample, with a higher number of alleles (including many DXS548 'odd' alleles) and a more 'spread out' distribution than either sample from Europe or Asia (Fig. 1, Chapter 2.3). This observation is compatible with a monocentric origin of modern man, migrating out of Africa [Gibbons, 1995] and is explained by the longer evolutionary history of our species in the African continent. On the contrary, Asian populations are the least variable of all world populations (e.g. only alleles 4-8 and C-D are found at the DXS548 and FRAXAC1 loci, respectively). Crawford et al. [2000] reported on a large cohort of 63 African American fragile X patients compared with 102 Caucasian fragile X chromosomes, thus characterizing the first significant sample of African fragile X patients. They also confirm our observation of a higher heterozygosity i.e. genetic diversity of African peoples and found haplotypes 7-C (7-3), 2-A (2-3) and 6-D (6-4) associated with fragile X in the 14%, 5% and 5% of cases, respectively. However, the major African fragile X haplotype is 4-4-5 (i.e. 4-D), which is not found in patients of European descent, but from which European haplotype 6-D (6-4) may have derived. In conclusion, data from Crawford et al. [2000] support

our initial hypothesis that some fragile X chromosomes may have come with men all the way out of Africa.

As more and more groups reported haplotype data at the fragile X locus using different systems to name alleles, we wrote a note on nomenclature [Chiurazzi et al., 1999a] in order to try to invite different authors to test a few reference DNA samples and submit the results together with their own naming of the haplotypes to a Web server that will make available these comparison tables. We hope that in this way, without limiting the Authors' choice as to the method they prefer to use to define alleles, it will be easier to compare the results from different studies. In the future it is likely that more studies with single-nucleotide polymorphisms (SNPs) will integrate the older ones that have been made with the more informative but more mutable microsatellites [Mathews et al., 2001].

Concluding this overview on the mutational history of fragile X syndrome, it is important to underline that founder effects – due to the multistep nature of the trinucleotide repeat disorder – have been documented also e.g. in Friedreich's Ataxia [Labuda et al., 2000] and Myotonic Dystrophy [Goldman et al., 1996]. Both these latter disorders are due to extreme expansions (up to 1000 units) of a GAA and a CTG repeat outside the coding region of the affected gene. In the CAG repeat-polyglutamine disorders like Huntington disease (HD), Kennedy disease (SBMA), and the various forms of spinocerebellar ataxias (SCA), the pathological expansion never exceeds the range of a fragile X premutation (probably thanks to a selection mechanism that prevents excessive expansions inside a coding region). In these conditions the mutational pathway is shorter than in fragile X (or e.g Friedreich's Ataxia), however a founder effect (though possibly less pronounced) has been reported for HD [Dode et al., 1993; Barron et al., 1994] and e.g. SCA2 [Saleem et al., 2000].

4.2 – Inactivation and reactivation of fragile X full mutations

Several generations after the initial destabilization, progressively or abruptly, some protomutations will emerge from the 'gray zone' of at-risk allele pools and become premutations. The two- to six-fold increase in size (relative to wild type range) is not enough to cause pathology, though the *FMR1* mRNA levels are increased [Tassone et al., 2000a; Kenneson et al., 2001], probably in order to compensate for their reduced translational efficiency [Feng et al., 1995]. When, possibly through different mechanisms (slippage of entire Okazaki fragments during replication, erroneous DNA repair or abnormal intrastrand recombination), the premutation expands to full mutation, this latter is inactivated and methylation of the CGG repeat and of the upstream CpG island is detected. It is difficult to know if methylation is the first event in inactivation of a full mutation, but it is surely a necessary step for transcriptional repression. In fact, we were able to *in vitro* reactivate the fully mutated *FMR1* gene by treating patients' cell lines with 5-azadeoxycytidine (5-azadC, see Chapter 3.1) and effectively poisoning the maintenance DNA methyltransferase (DNMT1). As shown, not only is the mRNA re-expressed, but a proportion of cells also show positive immunostaining for FMRP after 5-azadC treatment [Chiurazzi et al., 1998].

What is the normal function of CpG methylation and why does the *FMR1* full mutation become abnormally methylated ? Normally, methylation of dispersed CpG sites across the genome is thought to be performed by *de novo* DNA methyltransferase enzymes (at least DNMT3a and/or DNMT3b [Bestor, 2000]) during pregastrulation in the embryo proper [Razin and Shemer, 1995]. CpG methylation seems to reduce the background transcriptional 'noise' of transposons, of endogenous retroviruses and in general of the many 'parasitic' DNA sequences that litter the genome of vertebrates and are capable of transcription [Bestor and Tycko, 1996; Walsh and Bestor, 1999]. In a paradoxical way, so called "CpG islands" – regions with a high density of CpG sites mainly coincident with promoters of housekeeping genes and origins of replication – are always devoid of methylation. This is probably because of the continuous presence of binding proteins (e.g. transcription factors) that protect them from the activity of *de novo* DNMTs. The unmethylated state of cytosines actually allowed the evolutionary conservation of CpG islands, because methylated cytosines frequently undergo spontaneous deamination to thymine and produce G-T mismatches. Although enzymes for repairing G-T mismatches exist (e.g. MBD4/MED1 [Bellacosa, 2001]), such mismatches are not always efficiently repaired, thus causing C-to-T point mutations and leading to genome-wide loss of CpG sites with the exception of the unmethylated CpG islands.

The expanded CGG repeat of fragile X full mutations represents a 'parasitic' element from a genomic perspective and its methylation - possibly triggered by abnormal secondary structures formed by the CGG expansion [Darlow and Leach, 1998] and somehow spreading to the adjacent promoter and CpG island – is likely to contribute to its stabilization [Woehrle et al., 1996; Woehrle et al., 1998; Woehrle et al., 2001]. The *FMR1* full mutation is not the only CGG repeat that becomes methylated after expansion: the other folate-sensitive fragile sites FRAXE [Knight et al., 1993], FRAXF [Parrish et al., 1994], FRA16A [Nancarrow et al., 1994] and possibly FRA11B [Jones et al., 1995] also behave in a similar way and coincide with regions of delayed DNA replication [Hansen et al., 1993].

Is CpG methylation the only way to repress transcription on a local chromatin domain?

Actually, methylation alone can directly inhibit the binding of transcription factors [Kumari and Usdin, 2001], but it has also been shown to act by recruiting multiprotein complexes containing, among others, histone deacetylases [Razin, 1998]. Histonic protein modifications include not only (de)acetylation but also (de)phosphorylation and (de)methylation, and they somehow affect the chromatin structure in such a way that genes can be turned off and on [Berger, 2000]. Some organisms such as *S. cerevisiae* and *C. elegans* don't methylate their DNA at all, and in *Drosophila* methylation probably occurs during early embryonic development only [Lyko et al., 2000]. However, when methylation is present, it is dominant over histone deacetylation, as it was shown for several antioncogenes [Cameron et al, 1999] as well as for the *FMR1* full mutation [Chiurazzi et al., 1999]. Histone hyperacetylating drugs alone cannot reactivate transcription, as we also reported in Chapter 3.2, but they strongly synergize with the effect of 5-azadC i.e. after demethylation is obtained, histone hyperacetylation further increases the levels of transcription in a geometric fashion (see Fig. 4, Chapter 3.2). However, in the case of genes without a CpG island, histone hyperacetylating drugs can substantially increase the expression level, as it was shown also *in vivo* for the gamma-globin gene in sickle-cell disease [Atweh et al., 1999] and in a mouse model of spinal muscular atrophy (SMA) [Chang et al., 2001].

Our studies on *FMR1* promoter demethylation dynamics after 5-azadC treatment (Chapter 3.3) show that a sequential and thorough demethylation of the entire promoter takes place in a proportion of cells; however, some CpG sites can become (or sometimes already are) demethylated before reactivation of gene transcription (e.g. sites 27-29 spanning the E-box in footprint I). As the USF1 and USF2 transcription factors can bind their recognition sequence in the *FMR1* promoter even if CpG site 28 is methylated [Kumari and Usdin, 2001], it's likely that they allow its passive demethylation by protecting it from maintenance methylation by DNMT1

[Chiurazzi et al., 1999]. We suspect that the dynamics of 5-azadC induced demethylation are not reciprocal to those of the initial *de novo* methylation: this latter probably starts from the expanded CGG repeat and extends into the 5' promoter, while 5-azadC frequently achieves a thorough promoter demethylation although we have indirect evidence that the CGG repeat remains often – at least partly – methylated (experiments with *Mcr*BS and *Fnu*4HI).

We had previously observed that the *FMR1* reactivation levels achieved with 5-azadC are inversely correlated with the size of the CGG expansion [Chiurazzi et al., 1999]. Also, in Chapter 3.3 we found promoter demethylation after 3 days of 5-azadC treatment in a lower proportion of cells when the full mutation was larger (compare Fig. 3b with Fig. 2b). However, after 8 days of treatment, the majority of cells eventually seem to have a demethylated promoter. Therefore we propose that, even when the *FMR1* promoter is completely demethylated, transcription is still not possible, when the CGG repeat is still substantially methylated. Conversely, transcription should be possible if the promoter is methylated but the CGG repeat is not (as it may have been in the unique case of a lung tumor with a premutation and a partly methylated promoter, reported by de Graaff et al. [1995]). In conclusion, it appears that the inactive *FMR1* promoter is not as inaccessible as we previously thought, because at least some transcription factors can still bind to it and in some cases drive transcription in a few cells, as suggested by Tassone et al. [2001]. These latter results are interesting because they show that in peripheral blood leukocytes of fragile X patients with an (apparently) methylated full mutation, a variable amount of *FMR1* mRNA can be detected. In most cases mRNA levels are below 10% of wild-type, corresponding to 1-2% of total cells transcribing (unmethylated full mutations also have 5- to 10-fold higher mRNA levels [Tassone et al., 2000b]), though in few cases up to 10% of the cells may be transcribing. However, we suspect that the *Nru*I enzymatic assay - that Tassone et al. [2001] used to verify the complete methylation of the *FMR1* promoter - may not be sensitive enough (it tests only CpG sites 2 and 3, Fig.1 of Chapter 3.3). A bisulphite sequencing analysis, such as the one we describe in Chapter 3.3, may have uncovered a partial promoter demethylation in a proportion of cells.

4.3 - Future challenges in fragile X research

Ten years ago the *FMRI* gene was cloned and an unprecedented mutational mechanism due to the instability of triplet repeats was discovered [Verkerk et al., 1991]. In 2001, still several questions are waiting for answers: 1) What are the exact mechanisms that allow the initial small-scale instability over many generations and when are they active? 2) What (different) mechanisms can make a premutation expand so dramatically to a full mutation in just one generation? 3) When and in which cells is the pre- to full transition taking place? 4) How and when does the full mutation become methylated? And if we succeed in managing to demethylate it, will it remethylate in somatic cells?

Partial answers to these questions have been given, but much experimental work remains to be done. With respect to the mechanisms of small-scale instability, replication slippage has been often proposed as explanation [Richards and Sutherland, 1994] and may depend both on the length and AGG interspersion pattern of the CGG array [Nolin et al., 1999]. However, recent evidence from murine models of neurodegenerative disease transgenic for CAG repeat arrays showed that repeat variation can occur in the absence of DNA replication [Sinden, 2001] in non-dividing cells, in connection with DNA repair (maybe transcription-coupled repair). A murine knock-in model with a pure 98 CGG repeat in the *Fmr1* gene has been also recently obtained showing repeat instability [Bontekoe et al., 2001] and this model will be instrumental to these investigations. No such large expansion from pre- to full mutation has been, however, observed in murine models to date, and the mechanisms involved in these dramatic amplification are likely to be different from those responsible for the small-scale instability. Formation of large hairpin structures on newly synthesized Okazaki fragments may favor large expansions: White et al. [1999] showed that in yeast a 10-fold increase of expansions in a transgenic CGG array resulted from suppression of *RAD27*, the yeast homologue of the FEN1 nuclease, which has the ability to remove the 5' flap of Okazaki fragments. Another interesting possibility is that repair after a double strand break leads to successive rounds of abnormal unwinding and re-invasion of the donor sequence by the newly synthesized strand, allowing DNA synthesis to proceed more than once within the repeats [Richard and Paques, 2000]. If repair by gene conversion needed the homologue copy on the other X chromosome, this could also explain why only female carriers of a premutation can have children with a full mutation and male premutation carriers have only premutated daughters. Another explanation to this sex-bias was suggested by Malter et al. [1997]: the pre- to full transition may take place only in the germ line, but a selection mechanism only active during male gametogenesis would eliminate spermatogonia with CGG

arrays larger than a premutation [Reyniers et al., 1993]. During embryogenesis then, only variation in the full mutation range and eventually reductions to the pre- or wild-type size range would occur [Chiurazzi et al., 1994].

Knowing exactly when and why the full mutation does become methylated, could be also relevant for finding ways to interfere with this process and to prevent it. Further studies such as that of Burman et al. [1999] on the rare normal males with unmethylated full mutations might help to understand if *de novo* methylation of the full mutation can actually take place only in a restricted time window during early embryonic development [Schwemmle, 1999]. Careful post-mortem examinations of such individuals will also confirm whether or not a premutation allele was 'hiding' in their brain, thus explaining their normal functioning.

We also need more information about the DNA methylation and demethylation machinery: one maintenance (DNMT1) and two *de novo* methyltranferases (DNMT3A and B) have been characterized [Bestor, 2000], but we still are not sure whether *de novo* methylation is absent in adult somatic cells. This last information is critical in attempts to demethylate a full mutation: no remethylation should occur in order for a treatment to be effective. Woehrle et al. [2001] provided evidence that the somatic murine cell line A9 does not have the potential to remethylate an unmethylated full mutation.

An active DNA demethylase enzyme exists that works as repair methylcytosine glycosylase; however, it is still not clear whether another demethylase activity exists that can cleave out only the methyl group from the pyrimidine ring [Wolffe et al., 1999]. Harnessing the potential of such a demethylase could allow us to reactivate the full mutation also in quiescent cells.

Finally, when considering a pharmacological treatment based on DNA demethylation, the impact of such a treatment on the rest of the genome must be considered. Other genes would probably not be affected as cytosine methylation normally does not affect CpG islands [Walsh and Bestor, 1999], but we might worry about diminishing the transcriptional silencing of parasitic sequence elements (transposons and endogenous retroviruses).

Reactivation of the endogenous fully mutated gene remains an attractive alternative to gene therapy based on the addition of a fully functional copy of the *FMR1* gene to brain neurons.

However, supplementing *in vitro* produced FMRP to patients might become feasible if a peptide tag was added following a strategy such as that reported by Schwarze et al. [1999].

Of course, much information is still needed about the cellular and physiologic role of the FMR1 protein: knowing how its deficiency causes e.g. dendritic spine immaturity [Hinton et al., 1991; Weiler et al., 1997; Irwin et al., 2001] could give us clues about alternative treatments that may

compensate FMRP absence. For example, psychopharmacology is likely soon to provide physicians with powerful drugs for counteracting the absence of FMRP: a double-blind clinical trial on more than 100 fragile X boys is testing the efficacy of acetyl-L-carnitine in diminishing hyperactive behaviour and increasing attention span [Torrioli et al., 1999] and very promising compounds (AMPAkines) – that stimulate glutamatergic receptors involved in hippocampal LTP response and memory [Arai et al.,2000; Lauterborn et al., 2000] – are presently being tested in Alzheimer disease and could be also useful in treating fragile X mental deficit.

A final note should be made on screening for fragile X: a recent study from Israel [Toledano-Alhadef et al., 2001] reports on offering the fragile X test to pregnant women or to women of reproductive age. While cautioning against screening during pregnancy is warranted [Neri and Chiurazzi, 2000], neonatal screening, when accompanied by adequate genetic and psychological counselling, would not only provide reliable data on the prevalence of fragile X, but – most importantly – it will give the parental couple the option of early intervention for their child.

References

Arai AC, Kessler M, Rogers G, Lynch G (2000): Effects of the potent ampakine CX614 on hippocampal and recombinant AMPA receptors: interactions with cyclothiazide and GYKI 52466. Mol Pharmacol 58:802-813

Arrieta I, Gil A, Nunez T, Telez M, Martinez B, Criado B, Lostao C (1999): Stability of the FMR1 CGG repeat in a Basque sample. Hum Biol 71:55-68

Atweh GF, Sutton M, Nassif I, Boosalis V, Dover GJ, Wallenstein S, Wright E, McMahon L, Stamatoyannopoulos G, Faller DV, Perrine SP (1999): Sustained induction of fetal hemoglobin by pulse butyrate therapy in sickle cell disease. Blood 93:1790-1797

Barron LH, Rae A, Holloway S, Brock DJ, Warner JP (1994): A single allele from the polymorphic CCG rich sequence immediately 3' to the unstable CAG trinucleotide in the IT15 cDNA shows almost complete disequilibrium with Huntington's disease chromosomes in the Scottish population. Hum Mol Genet 3:173-175

Bellacosa A (2001): Role of MED1 (MBD4) Gene in DNA repair and human cancer. J Cell Physiol 187:137-144

Berger SL (2000): Local or global ? Nature 408:412-414

Bestor TH (2000): The DNA methyltransferases of mammals. Hum Mol genet 9:2395-2402

Bestor TH, Tycko B (1996): Creation of genomic methylation patterns. Nat Genet 12:363-367.

Bontekoe CJ, Bakker CE, Nieuwenhuizen IM, van der Linde H, Lans H, de Lange D, Hirst MC, Oostra BA (2001): Instability of a (CGG)$_n$ repeat in the *Fmr1* promoter. Hum Mol Genet 10:1693-1699

Burman RW, Yates PA, Green LD, Jacky PB, Turker MS, Popovich BW (1999): Hypomethylation of an expanded FMR1 allele is not associated with a global DNA methylation defect. Am J Hum Genet 65:1375-1386

Cameron EE, Bachman KE, Myohanen S, Herman JG, Baylin SB (1999): Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. Nature Genet 21:103-107

Chang JG, Hsieh-Li HM, Jong YJ, Wang NM, Tsai CH, Li H (2001): Treatment of spinal muscular atrophy by sodium butyrate. Proc Natl Acad Sci U S A 98:9808-9813

Chiurazzi P, Kozak L, Neri G (1994): Unstable triplets and their mutational mechanism: Size reduction of the CGG repeat vs. germline mosaicism in the Fragile X syndrome. Am J Med Genet 51:517-521

Chiurazzi P, Macpherson J, Sherman S, Neri G (1996a): Editorial Comment: Significance of linkage disequilibrium between the fragile X locus and its flanking markers. Am J Med Genet 64:203-208

Chiurazzi P, Genuardi M, Kozak L, Giovannucci-Uzielli ML, Bussani C, Dagna-Bricarelli F, Grasso M, Perroni L, Sebastio G, Sperandeo MP, Oostra BA, Neri G (1996b): Fragile X founder chromosomes in Italy: a few initial events and possible explanation for their heterogeneity. Am J Med Genet 64: 209-215

Chiurazzi P, Destro-Bisol G, Genuardi M. Oostra BA, Spedini G, Neri G (1996c): Extended gene diversity at the FMR1 locus and neighbouring CA repeats in a sub-Saharan population. Am J Med Genet 64:216-219

Chiurazzi P, Pomponi MG, Willemsen R, Oostra BA, Neri G (1998): *In vitro* reactivation of the *FMR1* gene involved in fragile X syndrome. Hum Mol Genet 7:109-113

Chiurazzi P, Pomponi MG, Sharrock A, Macpherson J, Lormeau S, Morel ML, Rousseau F (1999a): DNA panel for interlaboratory standardization of haplotype studies on the fragile X syndrome and proposal for a new allele nomenclature. Am J Med Genet 83:347-349

Chiurazzi P, Pomponi MG, Pietrobono R, Bakker CE, Neri G, Oostra BA (1999b): Synergistic effect of histone hyperacetylation and DNA demethylation in the reactivation of the *FMR1* gene. Hum Mol Genet 8:2317-2323

Crawford DC, Schwartz CE, Meadows KL, Newman JL, Taft LF, Gunter C, Brown WT, Carpenter NJ, Howard-Peebles PN, Monaghan KG, Nolin SL, Reiss AL, Feldman GL, Rohlfs EM, Warren ST, Sherman SL (2000): Survey of the fragile X syndrome CGG repeat and the Short-Tandem-Repeat and Single-Nucleotide-Polymorphism haplotypes in an African American population. Am J Hum Genet 66:480-493

Cummings CJ, Zoghbi HY (2000): Fourteen and counting: unraveling trinucleotide repeat diseases. Hum Mol Genet 9:909-916

Darlow JM, Leach DRF (1998): Secondary Structures in d(CGG) and d(CCG) Repeat Tracts. J Mol Biol 275:3-16

de Graaff E, Willemsen R, Zhong N, de Die-Smulders CE, Brown WT, Freling G, Oostra B (1995): Instability of the CGG repeat and expression of the FMR1 protein in a male fragile X patient with a lung tumor. Am J Hum Genet 57:609-618

de Vries BB, van den Ouweland AM, Mohkamsing S, Duivenvoorden HJ, Mol E, Gelsema K, van Rijn M, Halley DJ, Sandkuijl LA, Oostra BA, Tibben A, Niermeijer MF (1997): Screening and diagnosis for the fragile X syndrome among the mentally retarded: an epidemiological and psychological survey. Collaborative Fragile X Study Group. Am J Hum Genet 61:660-667

Dode C, Durr A, Pecheux C, Mouret JF, Belal S, Bachner L, Agid Y, Kaplan JC, Brice A, Feingold J (1993): Huntington's disease in French families: CAG repeat expansion and linkage disequilibrium analysis. C R Acad Sci III 316:1374-1380

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 instability in the FMR1 gene. Nat Genet 8:88-94

Eichler EE, Macpherson JN, Murray A, Jacobs PA, Chakravarti A, Nelson DL (1996): Haplotype and interspersion analysis of the FMR1 CGG repeat identifies two different mutational pathways for the origin of the fragile X syndrome. Hum Mol Genet 5:319-330

Feng Y, Zhang F, Lokey LK, Chastain JL, Lakkis L, Eberhart D, Warren ST (1995): Translational suppression by trinucleotide repeat expansion at FMR1. Science 268:731-734

Fu YH, Kuhl DP, Pizzuti A, Pieretti M, Sutcliffe JS, Richards S, Verkerk AJ, Holden JJ, 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

Gibbons A (1995): Out of Africa- at last? Science 267:1272-1273

Goldman A, Krause A, Ramsay M, Jenkins T (1996): Founder effect and prevalence of myotonic dystrophy in South Africans: molecular studies. Am J Hum Genet 59:445-452

Haataja R, Vaisanen ML, Li M, Ryynanen M, Leisti J (1994): The fragile X syndrome in Finland: demonstration of a founder effect by analysis of microsatellite haplotypes. Hum Genet 94:479-483

Hansen RS, Canfield TK, Lamb MM, Gartler SM, Laird CD (1993): Association of fragile X syndrome with delayed replication of the FMR1 gene. Cell 73:1403-1409

Hinton VJ, Brown WT, Wisniewski K, Rudelli RD (1991): Analysis of neocortex in three males with the fragile X syndrome. Am J Med Genet 41:289-294

Hirst MC (1995): FMR1 triplet arrays: paying the price for perfection. J Med Genet 32:761-763

Irwin SA, Patel B, Idupulapati M, Harris JB, Crisostomo RA, Larsen BP, Kooy F, Willems PJ, Cras P, Kozlowski PB, Swain RA, Weiler IJ, Greenough WT (2001): Abnormal dendritic spine characteristics in the temporal and visual cortices of patients with fragile-X syndrome: a quantitative examination. Am J Med Genet 98:161-167

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

Kenneson A, Zhang F, Hagedorn K, Warren ST (2001): Reduced FMRP and increased *FMR1* transcription is proportionally associated with CGG repeat number in intermediate-length and premutation carriers. Hum Mol Genet 10:1449-1454

Knight SJL, Flannery AV, Hirst MC, Campbell L, Christodoulou Z, Phelps SR, Pointon J, Middleton-Price HR, Barnicoat A, Pembrey ME, Holland J, Oostra BA, Bobrow M, Davies KE (1993): Trinucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation. Cell 74:127-134

Kumari D, Usdin K (2001): Interaction of the transcription factors USF1, USF2, and α-Pal/Nrf-1 with the FMR1 promoter. Implications for Fragile X mental retardation syndrome. J Biol Chem 276:4357-4364

Labuda M, Labuda D, Miranda C, Poirier J, Soong BW, Barucha NE, Pandolfo M (2000): Unique origin and specific ethnic distribution of the Friedreich ataxia GAA expansion. Neurology 54:2322-2324

Lauterborn JC, Lynch G, Vanderklish P, Arai A, Gall CM (2000): Positive modulation of AMPA receptors increases neurotrophin expression by hippocampal and cortical neurons. J Neurosci 20:8-21

Lyko F, Ramsahoye BH, Jaenisch R (2000): DNA methylation in *Drosophila Melanogaster*. Nature 408:538-540

Macpherson JN, Curtis G, Crolla JA, Dennis N, Migeon B, Grewal PK, Hirst MC, Davies KE, Jacobs PA (1995): Unusual (CGG)n expansion and recombination in a family with fragile X and DiGeorge syndrome. J Med Genet 32:236-239

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

Mathews DJ, Kashuk C, Brightwell G, Eichler EE, Chakravarti A (2001): Sequence variation within the fragile X locus. Genome Res 11: 1382-1391

Morris A, Morton NE, Collins A, Lawrence S, Macpherson JN (1995): Evolutionary dynamics of the FMR1 locus. Ann Hum Genet 59:283-289

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

Neri G, Chiurazzi P (2000): Fragile X syndrome screening: A current opinion. Community Genet 3:38-40

Nolin SL, Houck GE Jr, Gargano AD, Blumstein H, Dobkin CS, Brown WT (1999): FMR1 CGG-repeat instability in single sperm and lymphocytes of fragile-X premutation males. Am J Hum Genet 65:680-688.

Parrish JE, Oostra BA, Verkerk AJ, Richards CS, Reynolds J, Spikes AS, Shaffer LG, Nelson DL (1994): Isolation of a GCC repeat showing expansion in FRAXF, a fragile site distal to FRAXA and FRAXE. Nat Genet 8:229-235

Razin A (1998): CpG methylation, chromatin structure and gene silencing-a three-way connection. EMBO J 17:4905-4908

Razin A, Shemer R (1995): DNA methylation in early development. Hum Mol Genet 4:1751-1755

Reyniers E, Vits L, De Boulle K, Van Roy B, van Velzen D, de Graaff E, Verkerk AJMH, Jorens HZ, Darby JK, Oostra BA, Willems PJ (1993): The full mutation in the FMR1 gene of male fragile X patients is absent in their sperm. Nat Genet 4:143-146

Richard GF, Paques F (2000): Mini- and microsatellite expansions: the recombination connection. EMBO Rep 1:122-126.

Richards RI, Holman K, Friendendend 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

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

Saleem Q, Choudhry S, Mukerji M, Bashyam L, Padma MV, Chakravarthy A, Maheshwari MC, Jain S, Brahmachari SK (2000): Molecular analysis of autosomal dominant hereditary ataxias in the Indian population: high frequency of SCA2 and evidence for a common founder mutation. Hum Genet 106:179-187

Schwarze SR, Ho A, Vocero-Akbani A, Dowdy SF (1999): In vivo protein transduction: delivery of a biologically active protein into the mouse. Science 285:1569-1572

Schwemmle S (1999): In vivo footprinting analysis of the FMR1 gene: proposals concerning gene regulation in high-functioning males. Am J Med Genet 84:266-267.

Sinden RS (2001): Origins of instability. Nature 411:757-758

Tassone F, Hagerman RJ, Taylor AK, Gane LW, Godfrey TE, Hagerman PJ (2000a): Elevated levels of FMR1 mRNA in carrier males: a new mechanism of involvement in the fragile-X syndrome. Am J Hum Genet 66:6-15.

Tassone F, Hagerman RJ, Loesch DZ, Lachiewicz A, Taylor AK, Hagerman PJ (2000b): Fragile X males with unmethylated, full mutation trinucleotide repeat expansions have elevated levels of FMR1 messenger RNA. Am J Med Genet 94:232-236.

Tassone F, Hagerman RJ, Taylor AK, Hagerman PJ (2001): A majority of fragile X males with methylated full mutation alleles have significant levels of *FMR1* messenger RNA. J Med Genet 38:453-456

Toledano-Alhadef HL, Basel-Vanagaite L, Magal N, Davidov B, Ehrlich S, Drasinover V, Taub E, Halpern GJ, Ginott N, Shohat M (2001): Fragile-X carrier screening and the prevalence of premutation and full-mutation carriers in Israel. Am J Hum Genet 69: 351-360

Torrioli MG, Vernacotola S, Mariotti P, Bianchi E, Calvani M, De Gaetano A, Chiurazzi P, Neri G (1999): Double-blind, placebo-controlled study of L-acetylcarnitine for the treatment of hyperactive behavior in fragile X syndrome. Am J Med Genet 87: 366–368

Turner G, Webb T, Wake S, Robinson H (1996): Prevalence of fragile X syndrome. Am J Med Genet 64:196-197

Verkerk AJMH, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DPA, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang F, Eussen BE, van Ommen GJB, Blonden LAJ, Riggins GJ, Chastain JL, Kunst CB, Galjaard H, Caskey CT, Nelson DL DL, Oostra BA, Warren ST (1991): Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting lenght variation in fragile X syndrome. Cell 65:905-914.

Walsh CP, Bestor TH (1999): Cytosine methylation and mammalian development. Genes Develop 13:26-34

Weiler IJ, Irwin SA, Klintsova AY, Spencer CM, Brazelton AD, Miyashiro K, Comery TA, Patel B, Eberwine J, Greenough WT (1997): Fragile X mental retardation protein is translated near synapses in response to neurotransmitter activation. Proc Natl Acad Sci USA 94:5395-5400

White PJ, Borts RH, Hirst MC (1999): Stability of the human fragile X (CGG)(n) triplet repeat array in Saccharomyces cerevisiae deficient in aspects of DNA metabolism. Mol Cell Biol 19:5675-5684.

Wohrle D, Schwemmle S, Steinbach P (1996):DNA methylation and triplet repeat stability: new proposals addressing actual questions on the CGG repeat of fragile X syndrome. Am J Med Genet 64:266-267

Wohrle D, Salat U, Glaser D, Mucke J, Meisel-Stosick M, Schindler D, Vogel W, Steinbach P (1998): Unusual mutations in high functioning fragile X males: apparent instability of expanded unmethylated CGG repeats. J Med Genet 35:103-111

Woehrle D, Salat U, Hameister H, Vogel W, Steinbach P (2001): Demethylation, reactivation and destabilization of human fragile X full-mutation alleles in mouse embryocarcinoma cells. Am J Hum Genet 69: 504-515

Wolffe A, Jones PL, Wade PA (1999): DNA demethylation. Proc Natl Acad Sci USA 96:5894-5896

.

Summary

The fragile X syndrome represents the most common inherited cause of mental retardation worldwide. Fragile X belongs to a large group of more than 200 mental retardation conditions caused by mutations in X-linked genes (XLMR), that have a collective frequency of up to 1 in 1000 males. Fragile X syndrome is also unique because it was the first genetic condition caused – in the overwhelming majority of cases – by the expansion of an unstable CGG repeat, becoming the prototype of a growing list of inherited disorders due to the instability of trinucleotide repeats.

Ten years after the cloning of the *FMR1* gene involved in fragile X syndrome, we still don't know all the molecular players that allow the destabilization of the CGG repeat located close to the gene's CpG island. However, the finding of a founder effect in fragile X syndrome indicates that only few of the unstable repeats eventually reached the pathological range, while many of the potentially unstable CGG arrays are hidden in the upper limit of the normal range thus escaping detection. It also points to the multistep nature of the fragile X mutational history: a fragile X 'full' mutation is not created abruptly at a certain point of time – as it is the case of all traditional types of mutations – but requires passage through many generations before manifesting itself, thus allowing the constitution of 'pools' of at-risk alleles in the normal population. The long 'evolution' of fragile X mutations – normal alleles that are gradually transformed into full mutations – bears some resemblance to tumorigenesis. The first line of research illustrated in this Thesis focuses on this early phase of the natural history of fragile X: namely, founder effects and haplotypes at the FMR1 locus.

The second line of research was aimed at understanding what happens *after* the fragile X mutation has reached a pathological size. What we showed with our 'reactivation' experiments is that the size of the CGG expansion *per se* does not cause the silencing of the *FMR1* gene: it's the methylation that is added to the expansion that leads to the transcriptional block. Also a few individuals with normal intellect (or slight impairment) and an unmethylated full mutation have been observed in some fragile X families, confirming that the presence or absence of tiny chemical groups modifying the DNA can determine the activity of the *FMR1* gene. Thus, by studying the 'reactivation' of fragile X full mutations, we try to learn more about their 'inactivation' and – hopefully – about possible ways of preventing or 'reverting' their inactivation in fragile X children.

Samenvatting

Het fragiele X syndroom is wereldwijd de meest voorkomende erfelijke vorm van mentale retardatie. Het behoort tot een groep van meer dan 200 vormen van mentale retardatie veroorzaakt door mutaties in X gebonden genen (XLMR), hetgeen bij 1 op 1000 mannen voorkomt. Het fragiele X syndroom is uniek, omdat het de eerste erfelijke ziekte was veroorzaakt door - in de meeste gevallen- een verlenging van een instabiele CGG repeat, waardoor het prototype is geworden van een groeiende lijst van erfelijke ziekten veroorzaakt door instabiliteit van trinucleotide repeats. Tien jaar na het kloneren van het *FMR1* gen betrokken bij het fragiele X syndroom, weten we nog steeds niet alle moleculaire spelers die van belang zijn bij de destabilisatie van de CGG repeat dichtbij het CpG island van het gen. Het vinden van een founder effect in het fragiele X syndroom impliceert dat slechts enkele van de instabiele repeats uiteindelijk de pathologische lengte bereiken, terwijl vele van de potentieel instabiele CGG repeats zich in het bovenste gebied van de normale allelen bevinden, waardoor ze niet gedetecteerd worden. Verder wijst het naar een meerstaps proces van de fragiele X mutatie geschiedenis; een fragiele X 'volledige' mutatie ontstaat niet abrupt op een bepaald tijdstip, -wat wel geldt voor alle traditionele type mutaties- maar vereist passage door vele generaties voordat deze zich manifesteert. Ten gevolge hiervan treedt er constitutie op van 'pools' van risico allelen in de normale populatie. De lange 'evolutie' van fragiele X mutaties - normale allelen die langzaam stapsgewijs transformeren naar volledige mutaties- vertoont overeenkomsten met tumorgenese. De eerste lijn van onderzoek in dit proefschrift richt zich op de vroege fase van de natuurlijke herkomst van het fragiele X syndroom; namelijk, founder effects en haplotypes van het FMR1 locus.

De tweede lijn van onderzoek was er op gericht te begrijpen wat er gebeurt wanneer de fragiele X mutatie de pathologische grootte heeft bereikt. De reactivatie experimenten tonen aan dat de lengte van de CGG repeat op zich niet bepaalt of er inactivatie van *FMR1* optreedt; de methylering die wordt toegevoegd aan de repeat verlenging is bepalend voor het blokkeren van de transcriptie. De beschrijving van enkele individuen met een normaal (of licht verlaagd) IQ en een ongemethyleerde volledige mutatie bevestigt dat de aan- of afwezigheid van kleine chemische groepen die het DNA modificeren, de activiteit van het *FMR1* gen kunnen bepalen. Op deze manier, door het bestuderen van de 'reactivatie' van fragiele X volledige mutaties, proberen we meer te weten te komen over de inactivatie en hopelijk over mogelijkheden tot preventie of omkeerbaarheid van deze inactivatie in fragiele X kinderen.

Curriculum Vitae

Pietro Chiurazzi was born on August 25, 1967 in Rome, Italy

1986 - 1992   School of Medicine, graduated on October 20, 1992 at the Medical Faculty of the Catholic University (Rome) with the experimental thesis: "Molecular genetic study and diagnosis of the Fragile X syndrome"

1992 - 1996   Fellow in Medical Genetics at the Institute of Medical Genetics of the Catholic University (Rome), specialized in Medical Genetics on November 26, 1996 with the experimental thesis: "Fragile X founder chromosomes in Italy"

1996 - 1998   Research associate at the Institute of Medical Genetics of the Catholic University (Rome); two periods at the Department of Clinical University of the Erasmus University (Rotterdam) for a total of 4 months

1998 - 2000   Italian researcher abroad at the Department of Clinical Genetics, Erasmus University (Rotterdam) with a two-year grant from Fondazione Telethon on a project entitled: "Mechanisms and possible prevention of hypermethylation of full mutations causing the fragile X syndrome: perspective of *FMR1* gene reactivation"

1999 - now   Research scientist in Medical Genetics at the Department of Pediatrics, University Hospital of Messina (Italy) since July 26, 1999

Publications

Neri G, Chiurazzi P, Arena F, Lubs HA, Glass IA (1992): XLMR Genes: Update 1992. Am J Med Genet 43:373-382

Kozak L, Chiurazzi P, Genuardi M, Neri G (1993): Mapping of a gene for non-specific X-linked mental retardation: evidence for linkage to chromosomal region Xp21.1-Xp22.3. J Med Genet 30:866-869

Genuardi M, Chiurazzi P, Capelli A, Neri G (1993): X-linked VACTERL with hydrocephalus: the VACTERL-H syndrome. BDOAS 29(1):235-241

Chiurazzi P, de Graaff E, Ng J, Verkerk AJMH, Wolfson S, Fisch GS, Kozak L, Neri G, Oostra BA (1994): No apparent involvement of the FMR-1 gene in five patients with phenotypic manifestations of the Fragile X syndrome. Am J Med Genet 51:309-314

Chiurazzi P, Kozak L, Neri G (1994): Unstable triplets and their mutational mechanism: Size reduction of the CGG repeat vs. germline mosaicism in the Fragile X syndrome. Am J Med Genet 51:517-521

Neri G, Chiurazzi P, Arena JF, Lubs HA (1994): XLMR genes. Update 1994. Am J Med Genet 51:542-549

Gurrieri F, Genuardi M, Chiurazzi P, Gillessen-Kaesbach G, Neri G (1994): Exclusion of linkage between autosomal dominant split hand-split foot and markers from chromosome 7q. Further evidence for genetic heterogeneity of this anomaly. Am J Hum Genet 55:853-855

Neri G, Chiurazzi P, Pomponi MG (1994): The molecular diagnosis of Fragile X and X-linked mental retardation syndromes. Int Pediat 9:205-208

Genuardi M, Bardoni B, Floridia G, Chiurazzi P, Scarano G, Zollino M, Garcea N, Martini-Neri ME, Neri G (1995): Dicentric chromosome Y associated with Leydig cell agenesis and sex reversal. Clinical Genetics 47:38-41

Lubs HA, Chiurazzi P, Arena JF, Schwartz C, Tranebjaerg L, Neri G (1996): XLMR Genes: Update 1996. Am J Med Genet 64:147-157

Chiurazzi P, Macpherson J, Sherman S, Neri G (1996): Editorial Comment: Significance of linkage disequilibrium between the fragile X locus and its flanking markers. Am J Med Genet 64:203-208

Chiurazzi P, Genuardi M, Kozak L, Giovannucci-Uzielli ML, Bussani C, Dagna-Bricarelli F, Grasso M, Perroni L, Sebastio G, Sperandeo MP, Oostra BA, Neri G (1996): Fragile X founder chromosomes in Italy: a few initial events and possible explanation for their heterogeneity. Am J Med Genet 64:209-215

Chiurazzi P, Destro-Bisol G, Genuardi M, Oostra BA, Spedini G, Neri G (1996): Extended gene diversity at the FMR1 locus and neighbouring CA repeats in a sub-saharan population. Am J Med Genet 64:216-219

Lapenta V, Chiurazzi P, van der Spek P, Pizzuti A, Hanaoka F, Brahe C (1997): *SMT3A*, a human homologue of the *S. cerevisiae SMT3* gene, maps to chromosome 21qter and defines a novel gene family. Genomics 40:362-366

Chiurazzi P, Pomponi MG, Willemsen R, Oostra BA, Neri G (1998): *In vitro* reactivation of the *FMR1* gene involved in fragile X syndrome. Hum Mol Genet 7:109-113

Holden JJ, Percy M, Allingham-Hawkins D, Brown WT, Chiurazzi P, Fisch G, Gane L, Gunter C, Hagerman R, Jenkins EC, Kooy RF, Lubs HA, Murray A, Neri G, Schwartz C, Tranebjaerg L, Villard L, Willems PJ (1999): Eighth International workshop on the fragile X syndrome and X-linked mental retardation, August 16-22, 1997. Am J Med Genet 83:221-236

Lubs H, Chiurazzi P, Arena J, Schwartz C, Tranebjaerg L, Neri G (1999): XLMR genes: update 1998. Am J Med Genet 83:237-247

Chiurazzi P, Pomponi MG, Sharrock A, Macpherson J, Lormeau S, Morel ML, Rousseau F (1999): DNA panel for interlaboratory standardization of haplotype studies on the fragile X syndrome and proposal for a new allele nomenclature. Am J Med Genet 83:347-349

Grasso M, Faravelli F, Nigro CL, Chiurazzi P, Sperandeo MP, Argusti A, Pomponi MG, Lecora M, Sebastio GF, Perroni L, Andria G, Neri G, Bricarelli FD (1999): Mosaicism for the full mutation and a microdeletion involving the CGG repeat and flanking sequences in the FMR1 gene in eight fragile X patients. Am J Med Genet 85:311-316

Neri G, Chiurazzi P (1999): X-linked mental retardation. Adv Genet 41:55-94

Chiurazzi P, Pomponi MG, Pietrobono R, Bakker CE, Neri G, Oostra BA (1999): Synergistic effect of histone hyperacetylation and DNA demethylation in the reactivation of the *FMR1* gene. Hum Mol Genet 8:2317-2323

Torrioli MG, Vernacotola S, Mariotti P, Bianchi E, Calvani M, De Gaetano A, Chiurazzi P, Neri G (1999): Double-blind, placebo-controlled study of L-acetylcarnitine for the treatment of hyperactive behavior in fragile X syndrome. Am J Med Genet 87: 366–368

Hamel BC, Chiurazzi P, Lubs HA (2000): Syndromic XLMR genes (MRXS): update 2000. Am J Med Genet 94: 361-363

Neri G, Chiurazzi P (2000): Fragile X syndrome screening: A current opinion. Community Genet 3:38-40

Chiurazzi P, Oostra BA (2000): Genetics of mental retardation. Curr Op Pediat 12: 529-535

Neri G, Chiurazzi P (2000): X-linked mental retardation. Am J Med Genet 97:173
Introduction as Guest Editors of the Seminars in Medical Genetics issue

Chiurazzi P, Hamel BC, Neri G (2001): XLMR genes: update 2000. Eur J Hum Genet 9:71-81

DANKWOORD

All you need is love. *The Beatles (among others...)*

In het begin was 'JOE HOU'... I even received a T-shirt with that printed on sorrounded by many kisses (Coleta's) and the signatures of all you friends. September 1994 was my first time in Rotterdam and in The Netherlands: what a great time! I was in the same laboratory where 3 years earlier the *FMR1* gene had been cloned. It all started as a short visit to update my expertise in the Fragile X field and it became just the first of many times: twice two months in 1996 and finally two years from 1998 until 2000.

Beste Ben, zonder jou was dit helemaal onmogelijk: van harte bedankt! You gave me a book on Rotterdam the first time and suggested to come again to check the "everchanging" skyline: well, I did that! Since the very start you believed in me: I hope that after the "promotie" we can start thinking about science again... is there space for me in the toilet to do some experiments? En nooit meer te laat ! Wij moeten ook naar Hotel New York voor een lekker dinertje met Hennie....

And now "Grazie" Professor Neri (Giovanni): thank you for taking me to the 1991 Workshop in Strasbourg even before graduation. Thank you for sending me to Rotterdam and helping in any possible way. "Grazie" for your continuous support and friendship: you have been (and are) my true teacher, providing guidance and inspiration. Grazie anche ad Enrica per essere qui oggi!

Piet "Rotterdam" has been helped by many in the Department of Clinical Genetics, but one special person must be thanked now: Professor Hans Galjaard. Duizend mal bedankt, Professor! I was not only your guest in Westzeedijk for more than a time, but you always found time to invite me out for dinner or to a concert: that made me feel very special. You have a gift to make everyone feel at ease and I will never forget your generosity and the drinks prepared by your lovely wife at your place looking at the garden.

En nu, my two original paranymphen Cathy Bakker (my real and only Dutch teacher !) and Carola Bontekoe (you came with flowers at Schiphol !): you both gave me much scientifically (if I can clone constructs it's because of you, Cathy!) and not only.... I will never forget missing my plane to Amsterdam when Carola married Andre: I arrived with the next flight, just in time to dance with the bride (eh, Caroltje?!) at de Pelgrim, where my party should take place. Unfortunately Cathy had to stop her activity a bit in advance of her term for delivery and had to delegate her role to another paranymph, but she knows that for me it will just be as if I have three paranymphs: I'll keep the fingers crossed for you, lieve Cathy ... ik weet dat je zal jouw best doen !!!

Esther de Graaff is back in town and she has been my angel since 1994: lieve E*, you have always been there for me, you answered all my questions and also organized a "borrel" in 1996 to celebrate my specialty in Genetics (and I still have the bear with the papillon!)..... and finally you accepted without hesitating to be my paranymph in place of Cathy: Bedankt E* !

bedankt ook aan Cilesta (hemelkindje!)... and finally thanks to Peter (van Vuuren) en Jeroen who always helped me with the computers !

I will now take the lift down to the 7th floor and greet Melle (magic lady of orders!), Tom de Vries en Co. (how many slides?... and chats over wine and Mediterranean cousine, eh?). Just before getting out I want to say thanks to all the "portieren" (we always had a chat and a smile, expecially when I came in late at night or during the weekend: niets is onmogelijk, Paul!), Gabriela Bos (the smile of ABN AMRO) en de medewerkers van de Mensa (we met every day at luch time!). Ik ben nu op Dr. Molewaterplein bij de fietsstalling.... Who's there ? Simona (la ginecologa) and Fabiola (la cardiologa, soon PhD!), two italian friends in Rotterdam like me: let's go biking!

In the evening I went back home: once very far (Mw. Vermeulen and her dogs!) but since 1998 very close... Gerard (en Nelly), without you my Dutch experience would have been very boring: no talks on philosophy and sociology until 2 am, no special soup on Saturday evening, no rides to the swimming pool,   no lunch at the "Groene Passage",   no chats on kultuurverschillende, no walks with your dog Bunkie (=), less bier at the Wester Paviljoen, no good music, no trips to Papendrecht en Dordrecht: van harte "Bedankt" grote broer Gerard ... ik weet dat je bent altijd daar.

Another very special "brother" I have in Holland (when he is not in Belgium or somewhere else)... Peter "Joker" van der Spek. Crazy, smily, quick, friendly, helping, always joking, caring, Peter escapes all classifications... maybe it's because of his mother (Hallo en altijd bedankt voor je gastvriendschap, Mw. V.d. Spek!), anyhow 'Joker' is a special friend who always motivated me: BEDANKT Peter (also for not forgetting my birthday this year!!).. It all started with SMT3 by chance, where will we go now ??

On Sunday morning I had an appointment: kwart voor 11 in het Paradijskerk! Climbing the stairs and ringing the church bells ("suonare i campanelli") was a privilege I will never forget. Wouter Blaquiere, onze organist en koor dirigent, merci!   How music can speak to the heart..... Bedankt Albert (en Petra !) voor jullie vriendschap, Zondag was echt een vrolijke dag met jullie! Kees, thanks for the many invitations to your beautiful house near the Park: your wife always offered me nice food and you made me feel at home when I was in Rotterdam: bedankt! Wannecke, many thanks for your hospitality and dinner: Daniela still remembers making the bath to your little Pablo that night... last but not least, Padre Eric, onze pastor: van harte bedankt!

The experimental part of this PhD was partly done at the Institute of Medical Genetics, Catholic University, Rome, so let me thank the coworkers who put so much passion in this research and always encouraged me: the "FRAXA girls" Maria Grazia (the boss!), Roberta (i primers per il bisulfito!) and Elisabetta (e Carlo??) plus Ilaria Zito (who got her PhD this year too). Many colleagues supported me with their comments, thoughts and friendship: Christina ("Pietro, avrei un cDNA da farti vedere...."), Maurizio (Advance Australia fair !), Manuela (Manuuu !), Oana (Moana sei unica !), Eugenio (Good luck in Utah, Gege'), Fiorella 'little flower', Danilo, Accursio, Donatella e Teresa, Francesca e Daniela, Jolanda, Lucia, Luciana e Angela (dulcis in fundo). Everyone of you made my life in the Lab in Roma unique: GRAZIE. Special thanks to Paola Tanci and Libor Kozak (Ahoi!): with you I first approached molecular biology and learned not to get angry.

Finally I want to thank Padre Angelo Serra and Renato Bova: you are my inspiration and remind every day that science is all about trying (provare e riprovare) and organizing (ordine !).

The end of knowledge is wisdom; the end of wisdom, humility. *Peggy Claude-Pierre*

Many friends and colleagues in Italy and around the world collaborated and helped me: Franca Dagna-Bricarelli e Marina Grasso (Genova), Maria Pia Sperandeo e il Prof. Sebastio (Napoli), Maria Luisa Giovannucci (Firenze), Giovanni Destro-Bisol (Roma), Stephanie Sherman (Atlanta) and James Macpherson (Salisbury), Francois Rousseau (Montreal), Herb Lubs (Miami) and Ben Hamel (Nijmegen) and many many others that I had the pleasure to meet also in Frascati last september for the tenth International Workshop on Fragile X and XLMR: GRAZIE, THANKS, BEDANKT !

Special thanks to Katie Clapp from the FRAXA Foundation for her personal encouragement and invitation to Banbury last March (inspiring meeting under the snow!) and to all the american friends of the National Fragile X Foundation, with a special mention to Stephen and Deborah Le Cover, who organized the unforgettable Fragile X Conference in Los Angeles (July 2000) where I again met special people that we 'researchers' keep calling "fragile"....

Truly I say to you, whoever does not receive the kingdom of God
like a little child, will not come into it at all. *Mark 10, 15*

Let me now thank all the children with fragile X that I had the privilege to meet, their gifted parents and relatives, the friends and colleagues that gather each year for the exciting meeting of the Italian Association of Fragile X syndrome... Luigi, Valerio, Filippo, Gilberto e Paolo, Andrea, Mauro, Giacomo... I have learned much, much more from you than from any book. We are all special, there is no such thing as 'normality'... I wish that everyone be touched by your kindness and realize what a great treasure is hidden in your soul. I hope that - looking at you - we all remember to "receive the kingdom of God" like little children.

This thesis was actually written in Sicily, in the Department of Pediatrics at the University of Messina (after 2 years of cold Holland I needed a warm island!). Special colleagues and friends have provided me the calm and support needed to concentrate: let me thank Prof. Ignazio Barberi (il 'capo' siciliano!), Luisa Gitto (Dr. Melatonin), Francesco & Nilla (e la loro casa ospitale..), Aldo (la nuotata a Licata), Rosy & Tanino (Chez Toi e biciclette), Tanino V. ("il caffe, Petruzzo!"), Pippo Vita, Mariangela (il latte di mandorla!), Adolfa Crifo', Graziella (che insalate e che panini!) e Bruno (che pasta alla carrettera indimenticabile!!). Messina somehow reminds me of Rotterdam: it has a harbour (though smaller), is often windy (though warmer!) and I ride a bicycle also there!

Truly I say to you, if a seed of grain does not go into the earth and come to an end,
it is still a seed and no more; but through its death it gives much fruit. *John 12,24*

Life isn't for work, it's your work that should serve your life and make it fuller and richer... these pages would be incomplete if I didn't thank and acknowledge the love, the encouragement, the smiles, the fun, the emotions that I shared with my friends and family. If I am Pietro, I owe it also to all of them (including my cats Tigre and Tino !)...........

Grazie Rosy e Gigi (Mum & Dad): you gave me love since I was born (and before). You gave me a happy family and continue to love me now. If I smile today (= it's only because you smiled at me. I love you and hope you will always be proud of me. All my loving memory to my grandparents Giuseppe, Pasquale, Elena and Enza, I know they stand by me somehow every day... thank you to my uncles and aunts, to my cousins and nephews, far and close I never forgot you even in the "sounds of silence"... will I be able to travel more and come and see you ?!?

Grazie ai miei amici di sempre: Luca, Francesco, Massimo ed i vostri cari, grazie dal profondo del cuore ! Thanks to all my friends of the summer, Lavinio may not be the best beach of the world (e.g. compare it with the Great Barrier Reef – Australia) but it was great fun for all of us to meet for the holidays: I still treasure in my heart the precious memory of my birthday parties in the garden under the pine trees...

Liebe Familie Wameling: Vielen Dank fuer immer da sein ! Be und Edel, (kartellbruder) Sandor und Geza: es war fantastisch in Osnabrueck immer "zu Hause" zu sein.... Danke! Jetzt muss ich auch zurueck zum Tonalepass.... Ein grossen Kuss auch der ganze Familie Gerbaulet: herzlichen Willkommen Julius Gerbaulet (ich gratuliere mich Stephan & Ute !)

Grazie Don Roberto, grazie per il gruppo Tobia, ma soprattutto per quel tuo primo saluto quando sapevi il mio nome prima che io ti conoscessi... thanks for teaching us to smile and to touch the heart of the others, of our friends !

Love is strong as death, passion enduring as the grave. *Song of Songs 8,6*

Dulcis in fundo: sweet Daniela. We trusted each other right from the start, because our love is a precious gift that can only be unconditional. Our strenght is not hiding anything: absolute sincerity and attention to each other... honesty is rare, but it's the only way to live truly. You brought (sun)flowers in my life (and Winnie!), you give me courage to go ahead, you make the rainbow shine in my heart... you are my love and I love you: Stand by me ! (=

Gratefully yours,

Pietro   =)

Smile ☺ and the world will smile with you: happiness is contagious.

1 SMILE 4 YOU!

JOKER

Stellingen behorende bij het proefschrift

Origin and reactivation of the fragile X gene

I

A "dynamic" mutation is a multistep process taking place over many generations, involving both expansions and contractions of repeated DNA sequences.
This thesis

II

The fragile X mutations associated with the most frequent founder chromosomes originated thousands of years ago.
This thesis

III

The initial events leading to the instability of a wild-type *FMR1* allele are rare compared to those determining the final transition from premutation to full mutation.
This thesis

IV

Repeat instability may occur to a greater degree in non-dividing cells.
Sinden (2001) Nature 411:757-758
Kovtun and McMurray (2001) Nat Genet 27:407-411

V

A long expanded CGG repeat can be seen as a 'parasitic' element that becomes neutralized by cytosine methylation.
This thesis

VI

CpG methylation is the most important factor determining transcriptional repression of the *FMR1* gene.
This thesis

VII

Epigenetic phenomena orchestrate the DNA genetic music.
Pennisi (2001) Science 293:1064-1067

VIII

Histone hyperacetylating drug sodium butyrate has been successfully used
to induce expression of the fetal globin gene in patients with sickle cell disease.
Atweh et al. (1999) Blood 93:1790-1797

IX

The notion of permanence of our bodies is largely an illusion . We are a mosaic
of organs and tissues, some of which are continuously deconstructing and renewing.
*Jean Claude Ameisen, La sculpture du vivant: Le suicide cellulaire
ou la mort creatrice. 1999 Ed. du Seuil*

X

New neurons are probably produced even in the adult brain
and activation of neuronal stem cells might restore brain function.
Kempermann et al. (2000) Prog Brain Res 127:35-48

XI

The fruit of silence is prayer, the fruit of prayer is faith, the fruit of faith is love,
the fruit of love is service, the fruit of service is peace.
Mother Theresa

Pietro Chiurazzi
Rotterdam, 24 october 2001