The molecular basis of the fragile X syndrome

expansion of a trinucleotide repeat,
a new mutational mechanism

Annemieke Verkerk

The molecular basis of the fragile X syndrome

expansion of a trinucleotide repeat, a new mutational mechanism

De moleculaire basis van het fragiele X syndroom

verlenging van een trinucleotide repeat,
een nieuw mutatie mechanisme

PROEFSCHRIFT
Ter verkrijging van de graad van doctor
aan de Erasmus universiteit Rotterdam
op gezag van de rector magnificus
Prof. Dr. P.W.C. Akkermans M. Lit.
en volgens besluit van het College van Dekanen.

De openbare verdediging zal plaatsvinden op
woensdag 23 februari 1994 om 13.45 uur

door
Johanna Maria Henriëtta Verkerk
Geboren te Tiel

PROMOTIECOMMISIE

Promotor:	Prof. Dr. H. Galjaard
Co-Promotor:	Dr. B.A. Oostra
Overige leden:	Prof. Dr. M.F. Niermeijer Prof. Dr. J.A. Grootegoed Prof. Dr. G.J.B. van Ommen

Omslag: Jan Maarten van den Boom
Bert Eussen

Het onderzoek beschreven in dit proefschrift werd mogelijk gemaakt door financiële steun van de Stichting Klinische Genetica Rotterdam.

Print: Offsetdrukkerij Ridderprint B.V., te Ridderkerk

"On the gates of hell and a most unusual gene."

"And then as always, one stops to recollect with total astonishment and great reverence the massive burden of pain carried so patiently by the mothers, fathers, sibs, grandparents and the many others involved so closely on a daily basis with apparent failure, defect, handicap, disability, and disappointment in the many thousands of Martin-Bell syndrome families throughout the world."

John M. Opitz, Am. J. Med. Genet. 23, 1-10, 1986.

CONTENTS

PREFACE

The aim of the research described in this thesis was to isolate and characterize the gene involved in fragile X syndrome. In *Chapter 1* available techniques are described, used for the localization, identification and cloning of genes. By means of "positional cloning" a growing number of disease genes has been identified and isolated in the past years. Positional cloning is used when no information is available about the protein or function of the responsible gene. The fragile X syndrome is an X linked heritable disease that affects about 1 in 1200 males (*Chapter 2*). A cytogenetically visible "fragile site" at the end of the long arm of the X chromosome (Xq27-28) in cultured cells of affected males gave a first indication where to search on the X chromosome for a gene involved in the fragile X syndrome. In *Chapter 3* an overview is given of the different ways used by our and other research groups to search for the fragile X gene. The search started with the isolation of markers in the Xq27-28 region. Using markers closely linked to the fragile X locus, a yeast artificial chromosome was isolated that spanned the fragile X site (*Publications 7.1 and 7.2*) and contained the fragile X gene, designated *FMR1* (*Publication 7.1*). The characteristics of the *FMR1* gene are described in *Chapter 3*. In *Chapter 4* examples are given of DNA diagnosis of the fragile X syndrome before and after the isolation of the *FMR1* gene.

The fragile X syndrome was the first disease where an amplification of a trinucleotide repeat appeared to be responsible for the disease phenotype. This finding was soon followed by discoveries of other disorders where amplification of a trinucleotide repeat was responsible for causing the disease (*Chapter 5*). In *Chapter 6* possible mechanisms of repeat expansion and timing of expansion are discussed.

Chapter 7 contains the various articles of the experimental work which form the basis of this thesis. In *Publication 7.1* the isolation and identification of the *FMR1* gene (Fragile X Mental Retardation) is reported. In the *FMR1* gene, an unusual CGG repeat was found which has increased in size in the *FMR1* gene of fragile X carriers and patients. This discovery, that amplification of a trinucleotide repeat is the molecular genetic basis of the fragile X syndrome, revealed a new mechanism of transgenerational inheritance. The data in *Publication 7.2*, obtained by means of fluorescence *in situ* hybridization, show that the yeast artificial chromosome containing the *FMR1* gene spans the fragile X site. Examples of DNA diagnosis of the syndrome in fragile X families, using an intragenic probe, are given in *Publication 7.3*. Alternative splicing of the *FMR1* gene, resulting in 12 possible

mRNA transcripts, is described in *Publication 7.4*. In *Publication 7.5* one patient is described with a point mutation in the *FMR1* gene, resulting in the fragile X phenotype.

CHAPTER 1

THE SEARCH FOR DISEASE GENES

GENESANDDISEASESGENESANDDISEASESGENESANDDISEASESGENESANDDISEASESGENESANDDISEASESGENESANDDISEASESGENESANDDISEASESGENESANDDISEASESGENESANDDISEASESGENESANDDISEASESGENESANDDISEASES

Introduction

The human genome has been estimated to contain between 50,000 and 100,000 genes (1,2). Mutations in these genes may cause disease. To date more than 3000 conditions in man have been attributed with good certainty to a single gene mutation (1). Genes for more then 700 disorders have been assigned to specific regions of the different chromosomes (2). An increasing number of disease genes are being cloned and analyzed using recombinant DNA technology (2,3).

In all instances, knowledge of the gene structure provides the opportunity to identify the mutation(s) involved in different clinical and/or biochemical variants of a genetic disease. Also transgenic animals can be produced which offer new perspectives for studying the pathogenesis of a genetic disease and to test therapeutic strategies.

In the case of genetic disease, where the responsible protein defect is not yet known, the identification of a disease gene and its mutants provides new possibilities for early diagnosis of patients, carrier detection and prenatal diagnosis. In general two strategies can be used to identify human disease genes: *functional cloning* and *positional cloning* (3,4,5).

Functional cloning

In *functional cloning* the disease gene is isolated based on knowledge about the basic biochemical defect. In this case there is information available about the protein and/or function of the responsible gene (6,7,8). Two methods are used to isolate the responsible gene. The first is to raise antibodies against the protein of interest. These antibodies can then be used to screen cDNA expression libraries. The second method is to purify the normal protein product and determine a partial amino acid sequence. From this sequence possible DNA sequences can be deduced. Degenerate oligonucleotides can then be made which can be used to screen cDNA libraries (8). In *functional cloning*, the cloning of the gene is followed by mapping to a chromosomal region.

Positional cloning

For more than 90% of the known genetic diseases the responsible protein defect has not yet been elucidated. Here *positional cloning* is of use: the gene is first mapped to a specific chromosomal region, using large families in which a disease cosegregates with a genetic marker. Subsequently, the gene can be cloned before any protein function is known (9,10).

To establish potential regions of interest, chromosomal rearrangements such as

translocations or deletions in patients, can be of tremendous help in identifying the location of a disease gene. For instance in the search for the genes involved in Duchenne muscular dystrophy (DMD) (11,12), retinoblastoma (Rb) (13,14), Wilms' tumour (WT) (15), and von Hippel-Lindau disease (VHL) (16), the presence of translocations (DMD, VHL) and deletions (DMD, Rb, WT) in patients gave away the chromosomal area where the gene eventually was found. However, such rearrangements are rare, and most often the gene has to be mapped to a specific chromosomal region with linkage analysis. Huntington's disease (17), cystic fibrosis (CF) (18), and myotonic dystrophy (19), are three examples of disorders where locating the disease gene was dependent on the use of genetic linkage mapping. Genetic linkage mapping is based on the cosegregation of a disease gene of unknown location with a genetic marker of a known chromosomal localization, using large families (10). The closer two loci are to each other on a chromosome, the more likely it is that they segregate together during meiosis and the less likely it is that a recombination occurs between them. With linkage analysis, families in which a certain disease segregates are studied with multiple polymorphic markers until linkage of the disease is found with one or more of these markers (9,10). Many disease genes have been mapped with this genetic linkage approach. In first instance, a gene is roughly mapped to a specific chromosomal region, between distant markers. Subsequently, the region in which the defective gene must reside can be narrowed down to 1 - 5 megabases (Mb), using polymorphic markers between these distant markers, by scoring recombination events between these markers and the disease gene. This is then the candidate region in which a specific gene has to be identified. The genetic distance between a gene and a marker is expressed in centimorgans (cM). One cM corresponds to a recombination frequency of 1 in 100 meioses, and an average physical distance of about 1000 kilobases (kb). This means that an average of 100 meioses has to be studied to find one recombination between the gene and a marker within a region of about 1000 kb. This is usually the maximum resolution to narrow down the candidate region with genetic mapping. The region can then be further characterized by physical mapping.

Physical mapping

Physical maps can be generated with the aid of pulsed field gel electrophoresis (PFGE), using restriction enzymes which cut the DNA infrequently (20,21). Actual physical lengths of DNA fragments and distances between probes up to several thousands of kilobases can be established (22). In addition the use of somatic cell hybrids containing human chromosomes against a rodent background (*Chapter 3*)

and fluorescence *in situ* hybridization (FISH) can help to establish the order of probes (23). The FISH technique is based on labelling of DNA fragments with reporter molecules (such as biotin or dioxigenin), that, after hybridization, can be detected using fluorescent affinity reagents. In this way the chromosomal localization of DNA sequences can be directly determined (24). The order of two sequences labelled in different colours can be established relative to each other along metaphase chromosomes if they are separated by > 1 Mb (25,26). If interphase nuclei are used, the relative order of sequences separated by as little as 50 kb can be determined (23,27). With high resolution FISH different coloured DNA fragments can be ordered on highly extended DNA loops (halo's), which are arranged around the nucleus, with a resolution of 10-200 kb (28).

Many techniques and strategies have been developed to search for genes in a candidate region (29). Large contigs of genomic DNA fragments from the candidate region can be assembled by using yeast artificial chromosomes (YACs) (30,31). YACs can contain DNA inserts of up to 1000 kb or more, and are a way to obtain large continuous stretches of a desired DNA region (32,33). Complete physical maps of the human chromosome 21 and the human Y chromosome have been established by YAC contig building (34,35), and more than half of the X chromosome has now been covered by YAC contigs (36). The use of YACs has been instrumental in the cloning of the genes involved in for instance cystic fibrosis (37) and the fragile X syndrome (*Chapter 3 and Publication 7.1*). To obtain more manageable sized DNA fragments, YACs can be subcloned into cosmids (38) which can contain DNA fragments with an average size of 40 kb.

Isolation of coding sequences

The next step is to determine whether the YACs or cosmids contain gene sequences and subsequently to identify the matching gene(s). Genomic DNA consists of coding exons, but also of non coding sequences: introns, promoter regions and untranslated regions of genes, repeat sequences such as Alu or Line, etc. Exons contain coding sequences and these sequences are often conserved between species. To determine if a cosmid or YAC clone contains conserved sequences, these clones can be used to probe genomic DNA of different species (zoo blot). To identify whether sequences are transcribed, clones can be hybridized to Northern blots.

Most vertebrate genes are associated with the presence of a CpG island (promoter region) (39), and thus identification of CpG islands can give an indication of the gene density of a certain region and lead to the isolation of a gene (40,41,42).

Exon trapping is an alternative method to identify exon sequences in genomic

DNA (43). This method is based on the presence of splice sites within the DNA. A genomic fragment is cloned into a "splicing plasmid" and transfected into COS cells. When the inserted sequence contains an exon, this exon is contained within the mature transcript and can be amplified by means of flanking primer sequences.

To obtain expressed sequences, cDNA libraries can be screened with whole cosmids (44; *Publication 7.1*), YACs (45), or pools of microdissected DNA fragments (46). This can be done directly without knowledge of possible present gene sequences, or with the certainty that gene sequences must be present by using information obtained from previous tests (signals on zoo blots, Northern analysis, presence of CpG islands, or exon trapping). In order to prove that a candidate gene is a disease causing gene, this gene has to be screened for abnormalities in affected individuals.
Positional cloning has been proven successful for the isolation of many disease genes. In Table 1 a number of diseases is listed for which the gene has been cloned using this strategy.

Candidate gene approach
A number of genes has been cloned using the "candidate gene approach". With this approach the gene is mapped to a position on a chromosome where genes are already known to be located. One of these genes might be selected as a candidate disease gene on the basis of a possible known function which could be implicated in the pathogenesis of the disease. For example the gene for fibrillin, which is an element of elastic tissue, mapped to the same region on chromosome 15q as Marfan syndrome, a heritable disorder of connective tissue (59,60,61). In the fibrillin gene, indeed point mutations were found in several Marfan patients. X-linked spinal and bulbar muscular atrophy, or Kennedy's disease, mapped to the same region on Xq11-12 as the androgen receptor gene (84,96). Symptoms in patients suggested a defect in the androgen receptor and a CAG repeat in exon 1 was found to be amplified in patients. For Waardenburg's syndrome mutations were found in the PAX-3 gene, which was already mapped to chromosome 2q35-37 (73,74). Waardenburg's syndrome is characterized by pigmentary disturbances and deafness. One of the reasons to focus on the chromosome 2q region was that "Splotch" mice have Waardenburg like symptoms. In this mouse a deletion was found in the PAX-3 gene on the proximal part of chromosome 1 (97). This chromosome region contains regions that are homologous to the long arm of the human chromosome 2.
While the emphasis to isolate genes may shift more and more to the candidate

Table 1. Human diseases for which the gene has been cloned using positional cloning.

Disease	Chromosomal position	Reference
Autosomal:		
Aniridia	11p13	47
Charcot Marie Tooth type 1A	17p11.2	48
Charcot Marie Tooth type 1B	1q22-23	49,50
Cystic fibrosis	7q31.2	51-53
Familial polyposis coli	5q21-22	54,55
Familial breast cancer	17q21.3	56
Greig syndrome	7p13	57
Huntington's disease	4q16.3	58
Marfan syndrome	15q15-21	59-61
Myotonic dystrophy	19q13.2-13.3	62-67
Neurofibromatosis type 1	17q11.2	68,69
Neurofibromatosis type 2	22q	70
Retinoblastoma	13q14.1-14.2	71
Von Hippel Lindau disease	3p25-26	72
Waardenburg's syndrome type 1	2q35-37	73,74
Wilms tumour	11p13	75
X-linked:		
Adrenoleukodystrophy	Xq28	76
Choroideremia	Xq21.2	77
Chronic granulomatous disease	Xp21.1	78
Duchenne muscular dystrophy	Xp21.2	79
Fragile X syndrome	Xq27.3	80
Hunter syndrome	Xq28.1	81
Kallmann syndrome	Xp22.3	82,83
Kennedy's disease	Xq11-12	84
Lowe syndrome	Xq26.1	85
Menkes disease	Xq12-13	86,87
Nephrogenic diabetes insipidus	Xq28	88
Norrie disease	Xp11.4	89,90
X-linked agammaglobulinaemia	Xq21.3-22	91
X-linked glycerol kinase deficiency	Xp21.2-21.3	92,93
X-linked hydrocephalus	Xq28	94,95

gene approach, as genes are being isolated at a rapid pace in the framework of the Human Genome Project (98), the initial mapping of disease genes will continue to occur via one or more of the different modes of genetic linkage mapping.

Implications for diagnosis and treatment

In many cases the cloning of a (disease) gene and knowledge of gene defects has led to improved diagnostics at the molecular level, and has opened up possibilities for disease prevention as well as for treatment. Defects at the DNA level can be readily detected in for instance DMD (99), CF (100) and fragile X syndrome (*Chapter 4 and Publication 7.3*). Families in which these diseases occur can be offered carrier detection and prenatal diagnosis.

A better understanding of molecular defects could lead to early diagnosis and a timely intervention in the form of surgery or of a diet. For a disease such as familial adenomatous polyposis, where it is now possible to follow persons at risk on basis of the presence of a gene defect in the APC gene, early removal of the entire colon prevents the development of colonic carcinoma (54,55,101).

Approaches to treating genetic diseases include administration of proteins or drugs and somatic gene therapy. For example patients with the adult type of Gaucher disease, which is caused by a deficiency of the enzyme glucocerebrosidase, are since a few years treated by the administration of modified enzyme that allows recognition and uptake by macrophages (102,103). A lot of research is being done in the field of somatic gene therapy. Somatic gene therapy is the introduction of normal genes into somatic cells to correct defective genes. This is only possible for recessive autosomal and X-linked diseases. The first clinical trial of somatic gene therapy involved introduction of a normal adenosine deaminase gene into mature T cells in adenosine deaminase deficient patients (104). Much effort has been put in the research on gene therapy for cystic fibrosis (105). Experiments with CF mice, in which the ion transport defect is corrected by introducing the normal CF transmembrane conductance regulator gene coupled to liposomes into lung epithelium cells, look quit promising. Hopefully this therapy can be applied to human patients in the future (105).

Novel mechanisms of inheritance

By the isolation and study of genes, novel mechanisms of inheritance have been discovered. Genes are identified that are subject to imprinting, including the genes involved in the Angelman and Prader-Willi syndrome and hereditary paragangliomas (106-109). Children might inherit identical genes from their fathers and mothers, but these genes may have different effects. The Prader-Willi

syndrome occurs when the paternal copy of a gene on chromosome 15q11-13 is missing (107). A functional paternal copy of this gene is required for normal development. The maternal copy of this gene seems to be imprinted and not expressed. In Angelman syndrome a similar but opposite situation is found (106). The maternal copy of a gene on chromosome 15q11-13 is missing, but is needed for normal development, while the paternal copy is not expressed.
A new type of mutation has been found in several genes where an unstable trinucleotide repeat is responsible for causing the disease (*Chapter 5 and Publication 7.1*). Accurate diagnosis of these diseases is now possible by determining the repeat lengths belonging to the relevant genes in healthy individuals, normal carriers and patients.

The X chromosome
The X chromosome has been estimated to code for more than 400 disease genes (1). As X-linked segregation patterns are easy to recognize by the predominance of affected males in subsequent generations, the X-linked disease genes involved in colour blindness, haemophilia, and DMD were among the first genes that were mapped to a chromosome (36). To date more than 150 separate genes have been assigned to the X chromosome and more than 20 have been cloned (1,2,36). Table 1 shows a number of genes associated with X-linked diseases, cloned in the last few years.
Mental retardation of 11 males within one pedigree was a first indication for an X-linked pattern of inheritance in the Martin-Bell or fragile X syndrome (110). The discovery of a cytogenetically visible abnormality in these patients at the end of the long arm of the X chromosome in combination with macro-orchidism separated this group of patients from other X-linked mental retardation syndromes (111,112). This so called "fragile site" was the starting point in the search for the fragile X gene, which is the subject of this thesis.

REFERENCES

1. McKusick V.A. Mendelian inheritance in man. (10th ed.) The Johns Hopkins University press, Baltimore and London, 1992.
2. McKusick V.A. and Amberger J.S. The morbid anatomy of the human genome: chromosomal location of mutations causing disease. J. Med. Genet. 30, 1-26, 1993.
3. Collins F.S. Positional cloning: let's not call it reverse anymore. Nature Genet. 1, 3-6, 1992.
4. Orkin S.H. Reverse genetics and human disease. Cell 47, 845-850, 1986.

5. Ballabio A. The rise and fall of positional cloning? Nature Genet. 3, 277-279, 1993.
6. Orkin S.H. and Kazazian H.H. The mutation and polymorphism of the human ß-globin gene and its surrounding DNA. Ann. Rev. Genet. 18, 131-171, 1984.
7. Brown M.S. and Goldstein J.L. A receptor mediated pathway for cholesterol homeostasis. Science 323, 34-47, 1986.
8. Lee C.C., Wu X., Gibbs R.A., Cook R.G., Muzny D.M. and Caskey C.T. Generation of cDNA probes directed by amino acid sequence: cloning of urate oxidase. Science 239, 1288-1291, 1988.
9. Botstein D., White R.L., Skolnick M. and Davis R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphism. Am. J. Hum. Genet. 32, 314-331, 1980.
10. Ott J. Analysis of human genetic linkage: revised edition. The Johns Hopkins University Press, Baltimore and London, 1991.
11. Worton G.R., Duff C., Sylvester J.E., Schmickel R.D. and Willard H.F. Duchenne muscular dystophy involving translocation of the DMD gene next to ribosomal RNA genes. Science 224, 1447-1449, 1984.
12. Francke U., Ochs H.D., de Martinville B., Giacalone J., Lindgren V., Distèche, Pagon R.A., Hofker M.H., van Ommen G.J.B., Pearson P.L. and Wedgwood R.J. Minor Xp21 chromosome deletion in a male associated with expression of Duchenne muscular dystrophy, chronic granulomatous disease, retinitis pigmentosa and McLeod syndrome. Am. J. Hum. Genet. 37, 250-267, 1985.
13. Francke U. Retinoblastoma and chromosome 13. Cytogenet. Cell Genet. 16, 131-134, 1976.
14. Yunis J.J. and Ramsay N. Retinoblastoma and subband deletion of chromosome 13. Am. J. Dis. Child. 132, 161-163, 1978.
15. Fearon E.R., Vogelstein B. and Feinberg A.P. Somatic deletion and duplication of genes on chromosome 11 in Wilms' tumours. Nature 309, 176-178, 1984.
16. Yoshida M.A., Ohyashiki K., Ochi H., Gibas Z., Prout G.R., Pontes E.J., Huben R. and Sandberg A.A. Rearrangement of chromosome 3 in renal cell carcinoma. Canc. Genet. Cytogenet. 19, 351-354, 1986.
17. Gusella J.F., Wexler N.S., Conneally P.M., Naylor S.L., Anderson M.A., Tanzi R.E., Watkins P.C., Ottina K., Wallace M.R., Sakaguchi A.Y., Young A.B., Shoulson I., Bonilla E. and Martin J.B. A polymorphic DNA marker genetically linked to Huntington's disease. Nature 306, 234-238, 1983.
18. Tsui L-C., Buchwald M., Barker D., Knowlton R., Schumm J.W., Eiberg H., Mohr J., Kennedy D., Plavsic N., Zsiga M., Markiewicz D., Akots G., Brown V., Helms C., Gravius T., Parker C., Rediker K. and Donis-Keller H. Cystic fibrosis locus defined by a genetically linked polymorphic DNA marker. Science 230, 1054-1057, 1985.
19. Renwick J.H. and Bolling D.R. An analysis procedure illustrated on a triple linkage of use for prenatal diagnosis of myotonic dystrophy. J. Med. Genet. 8, 399-416, 1971.
20. Schwarz D.C. and Cantor C.R. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. Cell 37, 67-75, 1984.
21. Carle G.F. and Olson M.V. Separation of chromosomal DNA molecules from yeast by orthogonal-field-alternation gel electrophoresis. Nucleic Acids Res. 12, 5647-5664,

1984.

22. Van Ommen G.J.B., Verkerk J.M.H., Hofker M.H., Monaco A.P., Kunkel L.M., Ray P., Worton R., Wieringa B., Bakker E. and Pearson P.L. A physical map of 4 million bp around the Duchenne muscular dystrophy gene on the human X chromosome. Cell 47, 499-504, 1986.
23. Trask B.J. Fluorescence *in situ* hybridization: applications in cytogenetics and gene mapping. Trends Genet. 7, 149-154, 1991.
24. Troelstra C., Landsvater R.M., Wiegant J., van der Ploeg M., Viel G., Buys C.H. and Hoeijmakers J.H. Localization of the nucleotide excision repair gene ERCC6 to human chromosome 10q11-q12. Genomics 12, 745-749, 1992.
25. Kluck P.M., Wiegant J., Raap A.K., Vrolijk H., Tanke J.J., Willemze R. and Landegent J. The order of human hematopoietic growth factor and receptor genes on the long arm of chromosome 5, as determined by fluorescence *in situ* hybridization. Ann. Hematol. 66, 15-20, 1993.
26. Lebo R.V., Lynch E.D., Wiegant J., Moore K., Trounstine M. and van der Ploeg M. Multicolor fluorescence *in situ* hybridization and pulsed field electrophoresis dissect CMT1B gene region. Hum. Genet. 88, 13-20, 1991.
27. Brandriff B., Gordon L. and Trask B. A new system for high resolution DNA sequence mapping in interphase pronuclei. Genomics 10, 75-82, 1991.
28. Wiegant J., Kalle W., Mullenders L., Brookes S., Hoovers J.M.N., Dauwerse J.G., van Ommen G.J.B. and Raap A.K. High-resolution *in situ* hybridization using DNA halo preparations. Hum. Mol. Genet. 1, 587-591, 1992.
29. Hochgeschwender U. Toward a transcriptional map of the human genome. Trends Genet. 8, 41-44, 1992.
30. Burke D.T., Carle G.F. and Olson M.V. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science 236, 806-812, 1987.
31. Schlessinger D. Yeast artificial chromosomes: tools for mapping and analysis of complex genomes. Trends Genet. 6, 248-258, 1990.
32. Anand R., Villasante A. and Tyler-Smith C. Construction of yeast artificial chromosome libraries with large inserts using fractionation by pulsed-field gel electrophoresis. Nucleic Acids Res. 17, 3425-3433, 1989.
33. Larin Z., Monaco A.P. and Lehrach H. Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. Proc. Natl. Acad. Sci. USA 88, 4123-4127, 1991.
34. Chumakow I., Rigault P., Guillou S., Ougen P., Billaut A., Guasconi G., Gervy P., LeGall I., Soularue P., Grinas L., Bougueleret L., Bellanné-Chantelot C., Lacroix B., Barillot E., Gesnouin P., Pook S., Vaysseix G., Frelat G., Schmitz A., Sambucy J.L., Bosch A., Estivill X., Weissenbach J., Vignal A., Riethman H., Cox D., Patterson D., Gardiner K., Hattori M., Sakaki Y., Ichikawa H., Ohki M., Le Paslier D., Heilig R., Antonarakis S. and Cohen D. Continuum of overlapping clones spanning the entire human chromosome 21q. Nature 359, 380-387, 1992.
35. Foote S., Vollrath D., Hilton A. and Page D.C. The human Y chromosome: overlapping DNA clones spanning the euchromatic region. Science 258, 60-66, 1992.
36. Schlessinger D., Mandel J.L., Monaco A.P., Nelson D.L. and Willard H.F. Report of

the fourth international workshop on human X chromosome mapping 1993.
37. Green E.D. and Olson M.V. Chromosomal region of the cystic fibrosis gene in yeast artificial chromosomes: a model for human genome mapping. Science 250, 94-98, 1990.
38. Evans G.A., Lewis K. and Rothenberg B.E. High efficiency vectors for cosmid microcloning and genomic analysis. Gene 79, 9-20, 1989.
39. Bird A.P. CpG islands as gene markers in the vertebrate nucleus. Trends Genet. 3, 342-347, 1987.
40. Barlow D.P. and Lehrach H. Genetics by gel electrophoresis: the impact of pulsed field gel electrophoresis on mammalian genetics. Trends Genet. 3, 167-171, 1987.
41. Meastrini E., Rivella S., Tribioli C., Purtilo D., Rocchi M., Archidiacono N. and Toniolo D. Probes for CpG islands on the distal long arm of the human X chromosome are clustered in Xq24 and Xq28. Genomics 8, 664-670, 1990.
42. Estivill X., Farrall M., Scambler P.J., Bell G.M., Hawley K.M.F., Lench N.J., Bates G.P., Kruyer H.C., Frederick P.A., Stanier P., Watson E.K., Williamson R. and Wainwright B.J. A candidate for the cystic fibrosis locus isolated by selection for methylation-free islands. Nature 326, 840-845, 1987.
43. Buckler A.J. Chang D.D., Graw S.L., Brook J.D., Haber D.A., Sharp P.A. and Housman D.E. Exon amplification: a strategy to isolate mammalian genes based on RNA splicing. Proc. Natl. Acad. Sci. USA 88, 4005-4009, 1991.
44. Korn B., Sedlacek Z., Manca A., Kioschis P., Konecki D., Lehrach H. and Poustka A. A strategy for the selection of transcribed sequences in the Xq28 region. Hum. Mol. Genet. 1, 235-242, 1992.
45. Elvin P., Slynn G., Black D., Graham A., Butler R., Riley J., Anand R. and Markham A.F. Isolation of cDNA clones using yeast artificial chromosome probes. Nucleic Acids Res. 18, 3913-3917, 1990.
46. Djabali M., Nguyen C., Biunno I., Oostra B.A., Mattei M.G., Ikeda J.E. and Jordan B.R. Laser microdissection of the fragile X region: identification of cosmid clones and of conserved sequences in this region. Genomics 10, 1053-1060, 1991.
47. Ton C.C., Hirvonen H., Miwa H., Weil M.M., Monaghan P., Jordan T., van Heyningen C., Hastie N.D., Meijers-Heijboer H., Drechsler M., Royer-Pokora B., Collins F., Swaroop A., Strong L.C. and Saunders G.F. Positional cloning and characterization of a paired box- and homeobox-containing gene from the aniridia region. Cell 67, 1059-1074, 1991.
48. Valentijn L.J., Baas F., Wolterman R.A., Hoogendijk J.E., van den Bosch N.H.A., Zorn I., Gabreëls-Festen A.A.W.M., de Visser M. and Bolhuis P.A. Identical point mutation of PMP-22 in trembling-J mouse and Charcot-Marie-Tooth disease type 1A. Nature Genet. 2, 288-291, 1992.
49. Hayasaka K., Himoro M., Sato W., Takada G., Uyemura K., Shimizu N., Bird T.D., Conneally P.M. and Chance P.F. Charcot-Marie-Tooth neuropathy type 1B is associated with mutations of the myelin P_0gene. Nature Genet. 5, 31-34, 1993.
50. Kulkens T., Bolhuis P.A., Wolterman R.A., Kemp S., te Nijenhuis S., Valentijn L.J., Hensels G.W., Jennekens F.G.I., de Visser M., Hoogendijk J.E. and Baas F. Deletion of the serine 34 codon from the major peripheral myelin protein P_0gene in Charcot-

Marie-Tooth disease type 1B. Nature Genet. 5, 35-39, 1993.
51. Rommens J.M., Iannuzzi M.C., Kerem B.S., Drumm M.L., Melmer G., Dean M., Rozmahel R., Cole J.L., Kennedy D., Hidaka N., Zsiga M., Buchwald M., Riordan J.R., Tsui L.C. and Collins F.S. Indentification of the cystic fibrosis gene: chromosome walking and jumping. Science 245, 1059-1065, 1989.
52. Riordan J.R., Rommens J.M., Kerem B.S., Alon N., Rozmahel R., Grzelczak Z., Zielenski J., Lok S., Plavsic N., Chou J.L., Drumm M.L., Iannuzzi M.C., Collins F.S. and Tsui L.C. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science 245, 1066-1073, 1989.
53. Kerem B.S., Rommens J.M., Buchanan J.A., Markiewicz D., Cox T.K., Chakravarti A., Buchwald M. and Tsui L.C. Idenfification of the cystic fibrosis gene: genetic analysis. Science 245, 1073-1080, 1989.
54. Kinzler K.W., Nilbert M.C., Su L.K., Vogelstein B., Bryan T.M., Levy D.B., Smith K.J., Preisinger A.C., Hedge P., McKechnie D., Finniear R., Markham A., Groffen J., Boguski M.S., Altschul S.F., Horii A., Ando H., Miyoshi Y., Miki Y., Nishisho I. and Nakamura Y. Identification of the FAP locus genes from chromosome 5q21. Science 253, 661-665, 1991.
55. Groden J., Thliveris A., Samowitz W.,Carlson M., Gelbert L., Albertsen H., Joslyn G., Stevens J., Spirio L., Robertson M., Sargeant L., Krapcho K., Wolff E., Burt R., Hughes J.P., Warrington J., McPherson J., Wasmuth J., Le Paslier D., Abderrahim H., Cohen D., Leppert M. and White R. Identification and characterization of the familial adematous polyposis coli gene. Cell 66, 589-600, 1991.
56. Emi M., Katagiri T., Harada Y., Saito H., Inazawa J., Ito I., Kasumi F. and Nakamura Y. A novel metalloprotease/disintegrin-like gene at 17q21.3 is somatically rearranged in two primary breast cancers. Nature Genet. 5, 151-157, 1993.
57. Vortkamp A., Gessler M. and Grzeschik K.H. GL13 zinc-finger gene interrupted by translocations in Greig syndrome families. Nature 352, 539-540, 1991.
58. The Huntington's disease collaborative research group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell 72, 971-983, 1993.
59. Lee B., Godfrey M., Vitale E., Hori H., Mattei M.G., Sarfarazi M., Tsipouras P., Ramirez F. and Hollister D.W. Linkage of Marfan syndrome and a phenotypically related disorder to two different fibrillin genes. Nature 352, 330-334, 1991.
60. Masien C.L., Corson G.M., Maddox B.K., Glanville R.W. and Sakai L.Y. Partial sequence of a candidate gene for the Marfan syndrome. Nature 352, 334-352, 1991.
61. Dietz H.C., Cutting G.R., Pyeritz R.E., Maslen C.L., Sakai L.Y., Corson G.M., Puffenberger E.G., Hamosh A., Nanthakumar E., Curristin S.M., Stetten G., Meyers D.A. and Francomano C.A. Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene. Nature 352, 337-339, 1991.
62. Harley H.G., Brook J.D., Rundle S.A., Crow S., Reardon W., Buckler A.J., Harper P.S., Housman D.E. and Shaw D.J. Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. Nature 355, 545-546, 1992.
63. Buxton J., Shelbourne P., Davies J., Jones C., Van Tongeren T., Aslanidis C., de Jong P., Jansen G., Anvret M., Riley B., Williamson R. and Johnson K. Detection of an

unstable fragment of DNA specific to individuals with myotonic dystrophy. Nature 355, 547-548, 1992.

64. Aslanidis C., Jansen G., Amemiya C., Shutler G., Mahadevan M., Tsilfidis C., Chen C., Alleman J., Wormskamp N.G.M., Vooijs M., Buxton J., Johnson K., Smeets H.J.M., Lennon G.G., Carrano A.V., Korneluk R.G., Wieringa B. and de Jong P. Cloning of the essential myotonic dystrophy region and mapping of the putative defect. Nature 355, 548-551, 1992.
65. Brook J.D., McCurrach M.E., Harley H.G., Buckler A.J., Church D., Aburatani H., Hunter K., Stanton V.P., Thirion J.P., Hudson T., Sohn R., Zemelman B., Snell R.G., Rundle S.A., Crow S., Davies J., Shelbourne P., Buxton J., Jones C., Juvonen V., Johnson K., Harper P.S., Shaw D.J. and Housman D.E. Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. Cell 68, 799-808, 1992.
66. Mahadevan M., Tsilfidis C., Sabourin L., Shutler G., Amemiya C., Jansen G., Neville C., Narang M., Barcelo J., O'Hoy K., Leblond S., Earle-Macdonald J., de Jong P.J., Wieringa B. and Korneluk R.G. Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. Science 255, 1253-1255, 1992.
67. Fu Y.H., Pizzuti A., Fenwick R.G., King J., Rajnarayan S., Dunne P.W., Dubel J., Nasser G.A., Ashizawa T., de Jong P., Wieringa B., Korneluk R., Perryman M.B., Epstein H.F. and Caskey C.T. An unstable triplet repeat in a gene related to myotonic dystrophy. Science 255, 1256-1258, 1992.
68. Wallace M.R., Marchuk D.A., Andersen L.B., Letcher R., Odeh H.M., Saulino A.M., Fountain J.W., Brereton A., Nicholson J., Mitchell A.L., Brownstein B.H. and Collins F.S. Type 1 neurofibromatosis gene: identification of a large transcript disrupted in three NF1 patients. Science 249, 181-186, 1990.
69. Cawthon R.M., Weiss R., Xu G., Viskochil D., Culver M., Stevens J., Robertson M., Dunn D., Gesteland R., O'Connell P. and White R. A major segment of the neurofibromatosis type 1 gene: cDNA sequence, genomic structure, and point mutations. Cell 62, 193-201, 1990.
70. Rouleau G.A., Merel P., Lutchman M., Sanson M., Zucman J., Marineau C., Hoang-Xuan K., Demczuk S., Desmaze C., Plougastel B., Pulst S.M., Lenoir G., Bijlsma E., Fashold R., Dumanski J., de Jong P., Parry D., Eldrige R., Aurias A., Delattre O. and Thomas G. Alteration in a new gene encoding a putative membrane-organizing protein causes neuro-fibromatosis type 2. Nature 363, 515-521, 1993.
71. Friend S.A., Bernards R., Rogelj S., Weinberg R.A., Rapaport J.M., Albert D.M. and Dryja T.P. A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. Nature 323, 643-646, 1986.
72. Latif F., Tory K., Gnarra J., Yao M., Duh F.M., Orcutt M.L., Stackhouse T., Kuzmin I., Modi W., Geil L., Schmidt L., Zhou F.W., Li H., Wei M.H., Chen F., Glenn G., Choyke P., Walther M.M., Weng Y.K., Duan D.S.R., Dean M., Glavac D., Richards F.M., Crossey P.A., Ferguson-Smith M.A., Le Paslier D., Chumakov I., Cohen D., Chinault A.C., Maher E.R., Linehan W.M., Zbar B. and Lerman M.I. Indentification of the von Hippel-Lindau disease tumor suppressor gene. Science 260, 1317-1320, 1993.
73. Tassabehji M., Read A.P., Newton V.E., Harris R., Balling R., Gruss P. and Strachan

T. Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene. Nature 355, 635-636, 1992.

74. Baldwin C.T., Hoth C.F., Amos J.A., da-Silva E.O. and Milunsky A. An exonic mutation in the HuP2 paired domain gene causes Waardenburg's syndrome. Science 355, 637-638, 1992.
75. Rose E.A., Glaser T., Jones C., Smith C.L., Lewis W.H., Call K.M., Minden M., Champagne E., Bonetta L., Yeger H. and Housman D.E. Complete physical map of the WAGR region of 11p13 localizes a candidate Wilm's tumor gene. Cell 60, 495-508, 1990.
76. Mosser J., Douar A.M., Sarde C.O., Kioschis P., Feil R., Moser H., Poustka A.M., Mandel J.L. and Aubourg P. Putative X-linked adrenoleukodystrophy gene shares unexpected homology with ABC transporters. Nature 361, 726-730, 1993.
77. Cremers F.P.M., van de Pol D.J.R., van Kerkhoff L.P.M., Wieringa B. and Ropers H.H. Cloning of a gene that is rearranged in patients with choroideraemia. Nature 347, 674-677, 1990.
78. Royer-Pokora B., Kunkel L.M., Monaco A.P., Goff S.C., Newburger P.E., Baehner R.L., Cole F.S., Curnutte J.T. and Orkin S.H. Cloning the gene for an inherited human disorder -chronic granulomatous disease- on the basis of its chromosomal location. Nature 322, 32-38, 1986.
79. Monaco A.P., Neve R.L., Colletti-Feener C., Bertelson C.J., Kurnit D.M. and Kunkel L.M. Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. Nature 323, 646-650, 1986.
80. Verkerk A.J.M.H., Pieretti M., Sutcliffe J.S., Fu Y.H., Kuhl D.P.A., Pizutti 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. and Warren S.T. 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, 1991.
81. Flomen R.H., Green E.P., Green P.M., Bentley D. and Giannelli F. Determination of the organisation of coding sequences within the iduronate sulphate sulphatase (IDS) gene. Hum. Mol. Genet. 2, 5-10, 1993.
82. Franco B., Guioli S., Pragliola A., Incerti B., Bardoni B., Tonlorenzi R., Carrozzo R., Maestrini E., Pieretti M., Taillon-Miller P., Brown C.J., Willard H., Lawrence C., Persico M.G., Camerino G. and Ballabio A. A gene deleted in Kallmann's syndrome shares homology with neural cell adhesion and axonal path-finding molecules. Nature 353, 529-563, 1991.
83. Legouis R., Hardelin J.P., Levilliers J., Claverie J.M., Compain S., Wunderle V., Millasseau P., Le Paslier D., Cohen D., Caterina D., Bougueleret L., Delemarre-van de Waal H., Lutfalla G., Weissenbach J. and Petit C. The candidate gene for the X-linked Kallmann syndrome encodes a protein related to adhesion molecules. Cell 67, 423-435, 1991.
84. La Spada A.R., Wilson E.M., Lubahn D.B., Harding A.E. and Fischbeck K.H. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. Nature 352, 77-79, 1991.

85. Attree O., Olivos I.M., Okabe I., Bailey L.C., Nelson D., Lewis R.A., McInnes R.R. and Nussbaus R.L. The Lowe's oculocerebrorenal syndrome gene encodes a protein highly homologous to inositol polyphosphate-5-phosphatase. Nature 358, 239-242, 1992.
86. Vulpe C., Levinson B., Whitney G., Packman S. and Gitschier J. Isolation of a candidate gene for Menkes disease and evidence that it encodes a copper-transporting ATPase. Nature Genet. 3, 7-13, 1993.
87. Chelly J., Tümer Z., Tonnesen T., Petterson A., Ishikawa-Brush Y., Tommerup N., Horn N. and Monaco A.P. Isolation of a candidate gene for Menkes disease that encodes a potential heavy metal binding protein. Nature Genet. 3, 14-19, 1993.
88. Rosenthal W., Seibold A., Antaramian A., Lonergan M., Arthus M.F., Hendy G.N., Birnbaumer M. and Dichet D.G. Molecular identification of the gene responsible for congenital nephrogenic diabetes insipidus. Nature 359, 233-235, 1992.
89. Berger W., Meindl A., van de Pol T.J.R., Cremers F.P.M., Ropers H.H., Döerner C., Monaco A., Bergen A.A.B., Lebo R., Warburg M., Zergollern L., Lorenz B., Gal A., Bleeker-Wagemakers E.M. and Meifinger T. Isolation of a candidate gene for Norrie disease by positional cloning. Nature Genet. 1, 199-203, 1992.
90. Chen Z.Y., Hendriks R.W., Jobling M.A., Powell J.F., Breakefield X.O., Sims K.B. and Craig I.W. Isolation and characterization of a candidate gene for Norrie disease. Nature Genet. 1, 204-208, 1992.
91. Vetrie D., Vorechovsky I., Sideras P., Holland J., Davies A., Flinter F., Hammarström L., Kinnon C., Levinsky R., Bobrow M., Smith C.I.E. and Bentley D.R. The gene involved in X-linked agammaglobulinaemia is a member of the src family of protein-tyrosine kinases. Nature 361, 226-233, 1993.
92. Walker P., Muscatelli F. and Monaco A.P. Isolation of the human Xp21 glycerol kinase gene by positional cloning. Hum. Mol. Genet. 3, 107-114, 1993.
93. Sargent C.A., Affara N.A., Bentley E., Pelmear A., Bailey D.M.D., Davey P., Dow D., Leversha M., Aplin H., Besley G.T.N. and Ferguson-Smith M.A. Cloning of the X-linked glycerol kinase deficiency gene and its identification by sequence comparison to the Bacillus subtilis homologue. Hum. Mol. Genet. 2, 97-106, 1993.
94. Rosenthal A., Jouet M. and Kenwrick A. Aberrant splicing of neural cell adhesion molecule L1 mRNA in a family with X-linked hydrocephalus. Nature Genet. 2, 107-112, 1992.
95. Van Camp G., Vits L., Coucke P., Lyonnet S., Schrander-Stumpel C., Darby J., Holden J., Munnich A. and Willems P.J. A duplication in the L1CAM gene associated with X-linked hydrocephalus. Nature Genet. 4, 421-425, 1993.
96. Brown C.J., Goss S.J., Lubahn D.B., Wilson E.E., French F.S. and Willard H.F. Androgen receptor locus on the human X chromosome: regional localization to Xq11-12 and description of a DNA polymorphism. Am. J. Hum. Genet. 44, 264-269, 1989.
97. Epstein D.J., Vekemans M. and Gros P. Splotch (Sp^{2H}), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of Pax-3. Cell 67, 767-774, 1991.
98. Olson M.V. The human genome project. Review. Proc. Natl. Acad. Sci. USA 90, 4338-4344, 1993.

99. Chamberlain J.S. Multiplex PCR for the diagnosis of Duchenne muscular dystrophy. In PCR protocols: a guide to methods and applications. Innis M. et al. (Eds). Academic press, Orlando, pp 272-281, 1990.
100. Ferrie R.M., Schwarz M.J., Robertson N.H., Vaudin S., Super M., Malone G. and Little S. Development, multiplexing and application of ARMS test for common mutations in the CFTR gene. Am. J. Hum. Genet. 51, 251-262, 1992.
101. Nishisho I., Nakamura Y., Miyoshi Y., Miki Y., Ando H., Horii A., Koyama K., Utsunomiya J., Baba S., Hedge P., Markham A., Krush A.J., Petersen G., Hamilton S.R., Nilbert M.C., Levy D.B., Bryan T.M., Preisinger A.C., Smith K.J., Su L.K., Kinzler K.W., Vogelstein B. Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. Science 253, 665-669, 1991.
102. Barton N.W., Brady R.O., Dambrosia J.M., Di Bisceglie A.M., Doppelt S.H., Hill S.C., Mankin H.J., Murray G.J., Parker R.I., Argoff C.E., Grewal R.P., Yu K.T. and collaborators. Replacement therapy for inherited enzyme deficiency-macrophage-targeted glucocerebrosidase for Gaucher's disease. New Engl. J. Med. 324, 1464-1470, 1991.
103. Sato Y. and Beutler E. Binding, internalization and degradation of mannose-terminated glucocerebrosidase by macrophages. Clin. Invest. 91, 1909-1917, 1993.
104. Blaese R.M. and Anderson W.F. The ADA human gene therapy clinical protocol. Hum. Gene Ther. 1, 327-329, 1990.
105. Hyde S.C., Gill D.R., Higgins C.F., Trezise A.E.O., MacVinish L.J., Cuthbert A.W., Ratcliff R., Evans M.J. and Colledge W.H. Correction of the ion transport defect in cystic fibrosis transgenic mice by gene therapy. Nature 362, 250-255, 1993.
106. Williams C.A., Zori R.T., Stone J.W., Gray B.A., Cantu E.S. and Ostrer H. Maternal origin of 15q11-13 deletions in Angelman syndrome suggests a role for genomic imprinting. Am. J. Med. Genet. 35, 350-353, 1990.
107. Nicholls R.D., Knoll J.H.M., Butler M.G., Karam S. and Lalande M. Genetic imprinting suggested by maternal heterodisomy in non-deletion Prader-Willi syndrome. Nature 342, 281-285, 1989.
108. Meijers-Heijboer E.J., Sandkuijl L.A., Brunner H.G., Smeets H.J.M., Hoogeboom A.J.M., Deelen W.H., van Hemel J.O., Nelen M.R., Smeets D.F.C.M., Niermeijer M.F. and Halley D.J.J. Linkage analysis with chromosome 15q11-13 markers shows genomic imprinting in familial Angelman syndrome. J. Med. Genet. 29, 853-857, 1992.
109. Heutink P., van der Mey A.G.L., Sandkuijl L.A., van Gils A.P.G., Bardoel A., Breedveld G.J., van Vliet M., van Ommen G.J.B., Cornelisse C.J., Oostra B.A., Weber J.L. and Devilee P. A gene subject to genomic imprinting and responsible for hereditary paragangliomas maps to chromosome 11q23-qter. Hum. Mol. Genet. 1, 7-10, 1992.
110. Martin J.P. and Bell J. A pedigree of mental defect showing sex linkage. J. Neurol. Psych. 6, 154-157, 1943.
111. Lubs H.A. A marker X chromosome. Am. J. Hum. Genet. 34, 552-557, 1969.
112. Glass I.A. X linked mental retardation. Review article. J. Med. Genet 28, 361-371, 1991.

CHAPTER 2

DESCRIPTION OF THE FRAGILE X SYNDROME

The fragile X or Martin-Bell syndrome is an X-linked heritable disease. After trisomy 21 or Down syndrome, the fragile X syndrome is the most frequent inherited cause of mental retardation. The syndrome is associated with a (rare) fragile site (referred to as FRAXA) which is cytogenetically visible at the end of the long arm of the X chromosome, in band Xq27.3. For an X-linked disease this syndrome shows an unusual pattern of inheritance, because part of the carrier females are affected and because of the occurrence of normal (i.e. unaffected) male carriers. These males have been termed "normal transmitting males", since they transmit the fragile X mutation to their daughters and are themselves, just as their daughters, clinically and cytogenetically normal. The prevalence of individuals phenotypically affected with the fragile X syndrome is estimated to be 1 in 1250 males and 1 in 2500 females (1,2,3). These numbers are based on the incidence of mentally retarded individuals with cytogenetic expression. Including the normal transmitting males, the total number of males estimated to carry the fragile X mutation is about 1 in 1000 (3). For females this number is estimated to be 1 in about 400 - 500 (4,5).

2.1 Some historical aspects

In 1943 the fragile X syndrome was described by Martin and Bell, who reported the first pedigree with clear familial X-linked mental retardation (6). In 1969 Lubs noted a secondary constriction, referred to as a fragile site, at the end of the long arm of the X chromosome (Xq27.3) in cultured cells from patients with X-linked mental retardation (7). The existence of a fragile X site in cells of fragile X patients, distinguished these patients from other patients with X-linked mental retardation syndromes. Another important observation was the macro-orchidism (large testes) in a high proportion of male fragile X patients (8,9).
Although Lubs discovered a fragile site at Xq27.3 in cells of fragile X patients, this observation could not be reproduced for many years. In 1977 Sutherland discovered that the finding of a fragile site depends on the type of medium in which lymphocytes were cultured (10,11). Culture medium deficient in folic acid was necessary to induce fragile chromosome sites. Following this discovery, a series of families with apparently X-linked non specific mental retardation were tested using folate-deficient medium, and the fragile X chromosome was identified in a high proportion (12). From this time on the presence of a fragile site was used as a diagnostic criterium.
A reexamination of the family described by Martin and Bell indeed demonstrated a

fragile site at Xq27.3 in cells of male patients (13). The term Martin-Bell syndrome has since been used as a synonym for the fragile X syndrome.

As a result of a growing interest and a rapidly expanding research into the fragile X syndrome, since 1983 international workshops on this theme are held every 2 years, and the proceedings have been published in special issues of the American Journal of Medical Genetics (14).

2.2 Clinical aspects

The main characteristics of male fragile X patients are moderate to severe mental retardation (15,16), macro-orchidism, and a long face with a prominent forehead and jaw and large everted ears (17). These features represent the classical Martin-Bell phenotype (18). Figure 1 shows a male patient with typical facial characteristics. Macro-orchidism is found in 65-70% of all affected adult fragile X males and usually develops after puberty (19). In approximately 10% of the patients mental retardation is considered to be the only symptom (17). Fragile X patients are found to have a normal life span (20). There is a large variability in the clinical expression of the syndrome in patients, and besides mental retardation the fragile X mutation causes a spectrum of additional problems. Normal transmitting males and their daughters always have a normal phenotype.

Figure 1. Fragile X patient with typical facial features. (Courtesy of L.B.A. de Vries.)

Physical features

A long and narrow face is often less notable in prepubertal boys, but prominent ears are common in fragile X adults as well as in children (17,21). In males, a slight overgrowth of the head is observed as compared to normal controls (21-23). A high arched palate has been found in 48% of fragile X males often associated with slight dental abnormalities (17). Ophthalmologic problems occur in more than 50% of fragile X patients and include strabismus and short-sightedness (24,25).

The skin of fragile X patients is often smooth and soft and shows abnormal elastin fibres (12,17). This feature, in addition to a high palate, hyperextensible finger joints, pectus excavatum and flat feet, suggested abnormalities of connective tissue. In connective tissue disorders like Marfan syndrome, the majority of patients suffer from severe cardiovascular complications (26). To assess the suggestion that connective tissue is involved in fragile X syndrome, patients were cardiologically investigated. In some patients dilatation of the aortic root and insufficient functioning of the mitral heart valve were found (27). In Table 1 physical features are listed that are identified in affected fragile X males.

Table 1. Physical features of fragile X males.

Feature	% of patients with feature
long ears	66
prominent ears	63
high arched palate	48
prominent jaw	28
long face	74
pectus excavatum	43
hyperextensible fingers	64
hand calluses	45
double-jointed thumbs	41
single palmar crease	35
flat feet	65
macro-orchidism	74
scoliosis	20
strabismus	33

Adapted from Hagerman R.J. 1991 (17).

In approximately 17% of fragile X males, epileptic seizures have been reported (17, 28-30). In view of the mental retardation found in fragile X patients it is of interest whether abnormalities can be detected in the brain of these patients. Only 4 post-mortem brain studies have been done in fragile X males (31-34). In one 62 year old fragile X male, abnormalities of dendritic spines were observed in different layers of the cortex. In addition there was a reduction in the length of the synapses as compared to controls (31). The significance of these findings remains to be established in a larger number of cases. Another patient showed mildly dilated ventricles (32), which are also found during life by neuroimaging on CT scan in approximately 38% of the patients (31). In a third patient brain changes were typical for ALS (amyotrophic lateral sclerosis) which had developed in this patient at a relatively young age (33).

Behavioral aspects

Behavioral abnormalities include hyperactivity, hand-flapping, handbiting, shyness and avoidance of eye contact, and perseverative speech (17,19).

A possible association between fragile X syndrome and autism has been studied by several groups, because many fragile X patients show autistic like features such as gaze avoidance and hand-flapping. This suggested that autistic behaviour may be common in fragile X males. Fish pooled data from previously published reports and evaluated these data from an epidemiological perspective (35). He could not establish a statistically significant correlation between fragile X syndrome and autism.

Fragile X females

About 30% of all females with the fragile X mutation show some degree of mental impairment, ranging from learning disabilities to mental retardation (36). Females with fragile X syndrome may have behavioral and emotional problems (37). Facial and other characteristics are usually less pronounced and less frequent than in males (17,19,36,37).

Subphenotypes of fragile X syndrome

Besides the Martin-Bell phenotype other less frequent phenotypes have been observed:

1. Fragile X males with a Prader-Willi like phenotype (38,39). These patients show extreme obesity and some have a short stature. All these patients have short broad hands and feet and resemble the classical Prader-Willi syndrome patients.

2. Fragile X males with a Sotos like phenotype (40). These patients show overgrowth in the first years of life, especially of the head.

Detailed reviews of clinical aspects in fragile X syndrome patients have been published by Hagerman (17) and Frijns (19).

2.3 Cytogenetic expression of the fragile X site

For a number of years confirmation of the clinical diagnosis was carried out by cytogenetic detection of the fragile site at Xq27.3 in cultured lymphocytes or fibroblasts (7,11,41). The appearance of the fragile site, microscopically visible as a gap or discontinuity in chromatids and chromosomes (see Figure 2) is induced by culturing cells in folic acid deficient medium before chromosome spreads are made (11,41,42). In mentally retarded males the fragile site is usually seen in 2-60% of the cells. The site can be detected in about 50% of obligate carrier females (43). Normal transmitting males and their daughters do not express the fragile site cytogenetically. Therefore, the cytogenetic test is not reliable in identifying carriers of the fragile X mutation.
Over one hundred different fragile sites have been described dispersed over different chromosomes (44), of which only two are now associated with a disease phenotype, the FRAXA site associated with the fragile X syndrome (7) and the X-linked mental retardation syndrome associated with the FRAXE fragile site (for

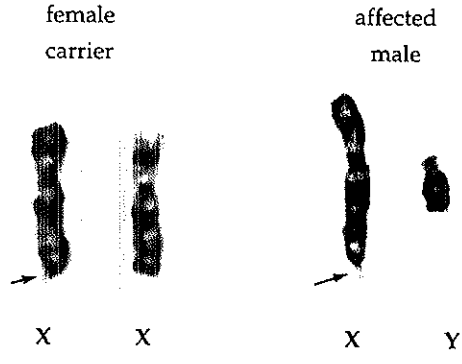


Figure 2. Sex chromosomes of a female carrier and an affected male. The fragile site at Xq27.3 is clearly visible and indicated with an arrow.
(Chromosome preparations kindly provided by Dr. J.O. Van Hemel).

details see *Chapter 3, paragraph 3.6*), which is located about 600 kilobases distal from FRAXA (45).

Fragile sites can be classified into two major groups: rare and common (15). Both groups can be further classified according to the conditions of tissue culture under which they are expressed. The exact mechanism underlying the fragile site expression is unknown. The FRAXA site is defined as a folate-sensitive site, since its cytogenetic expression can be induced by media deficient in folic acid and thymidine, with excess thymidine or containing the dihydrofolate reductase inhibitor methotrexate (46). This influences the synthesis of pyrimidines necessary for DNA synthesis and results in low levels of thymidine or cytidine (15,46). Therefore it was proposed that a folate sensitive fragile site results from a naturally occurring amplificated sequence of which the replication is hindered when either dCTP or dTTP is limited (47).

2.4 Unusual pattern of inheritance

In general, genetic diseases segregate throughout a family following the rules of mendelian inheritance. For X-linked recessive diseases this means that females who carry a defective gene on one of their X chromosomes usually are unaffected because of the presence of a normal X chromosome. Their daughters have a risk of 50% of being unaffected carriers and their sons have a risk of 50% of being affected. This is depicted in Figure 3A. Fragile X syndrome deviates from this pattern on the following points:

1. approximately 20% of males who carry the fragile X mutation are phenotypically normal. They will pass their X chromosome with the non-penetrant mutation on to all their daughters. As these apparently non-penetrant fathers pass on the mutation and are themselves clinically unaffected, they are called normal transmitting males (NTMs) (48,49). Their daughters, who are also normal, are however at risk of having affected children. A transmitting male was previously detected only if one of his direct descendants had affected offspring.

   By performing extensive segregation analysis in fragile X syndrome families, Sherman calculated the risks for individuals in a pedigree to be mentally retarded (48,49). She determined that mothers and daughters of intellectually normal transmitting males are rarely if ever mentally impaired. In addition it was noted that the siblings of transmitting males are much less likely to be retarded than the siblings of mentally impaired males. Mothers and daughters

of NTMs are similar in phenotype, both are obligate carriers and intellectually normal, but the effect of the gene in their offspring is apparently different: the penetrance of mental impairment is higher in offspring of intellectually normal daughters of transmitting males than in offspring of intellectually normal mothers of transmitting males. The risk for separate individuals to be mentally retarded seems to be dependent upon the position of an individual in the pedigree, a phenomenon that has become known as the Sherman paradox (48,49). This risk is expressed in percentages in Figure 3B.

Normal transmitting males have a 100% chance of having normal daughters, though carriers. Sons of these normal daughters have a 38% chance of being affected, and daughters have a 16% chance of being affected. Sisters and brothers of normal transmitting males have a low probability of clinically expressing the disease, 5% and 9% respectively. If a mother is mentally impaired, her sons have a 50% chance of being affected, and her daughters

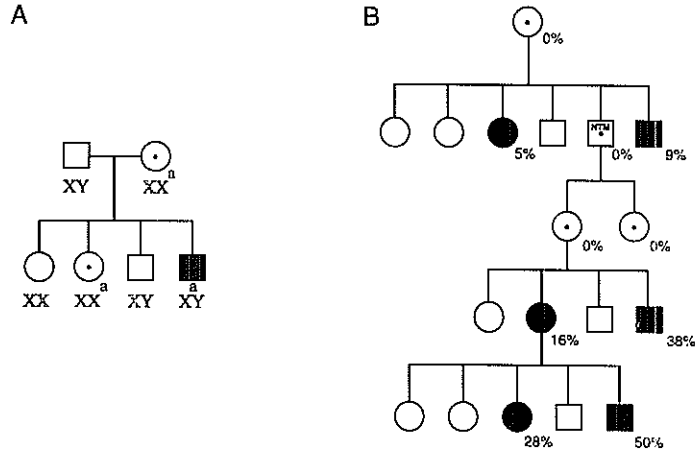


Figure 3. A: Inheritance pattern for an X-linked recessive disease.
Daughters of a female carrier have a 50% chance of being a normal carrier. Sons of a female carrier have a 50% chance of being affected.
X^a = X chromosome with a defective gene.

B: Fragile X pedigree illustrating the Sherman paradox. Mothers and daughters of normal transmitting males have different chances of having affected offspring.

○ = normal female
⊙ = normal female carrier
● = affected female
□ = normal male
NTM = normal male carrier
■ = affected male

have a 28% chance of being affected.
The Sherman paradox can now be explained at the molecular level by the size of the CGG repeat that is found in the fragile X gene (4). This is described in *Chapter 3*.

2. About 35% of the carrier females are mentally retarded, although they are often less severely affected than males (16,17). Sons of these affected females have the usual risk of 50% of being affected.
   Before the fragile X gene was identified, female carriers were defined as women who showed cytogenetic expression of the fragile site at Xq27.3, or who, based on pedigree data, were obligate carriers. After identification of the fragile X gene, female carriers could be divided into two groups (*Chapter 3*), and only those females with a "full mutation" are at risk of being mentally retarded.

3. The disease phenotype is only found in offspring after transmission of the mutation by a female. Daughters of normal male carriers are always normal. The daughters of the few male patients that have reproduced are also normal (50,51,52) (*Chapter 6*).

Several theories were proposed to explain this non-mendelian inheritance pattern. These theories have been the subject of many discussions in the past but after the elucidation of the molecular defect none of these theories has proven to be completely correct.

A: Premutation hypothesis by Pembrey.
= X chromosome carrying a premutation
= X chromosome carrying a recombination enhancing sequence
= X chromosome carrying definitive Martin-Bell mutation
= clinically normal female
= normal transmitting male
= mentally retarded
Adapted from Pembrey et al., 1985 (53).

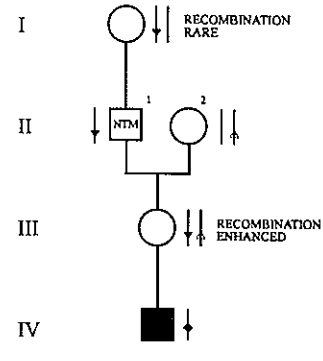
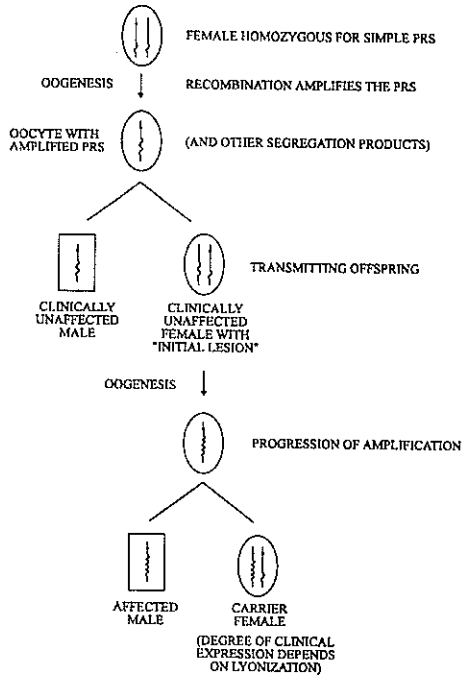
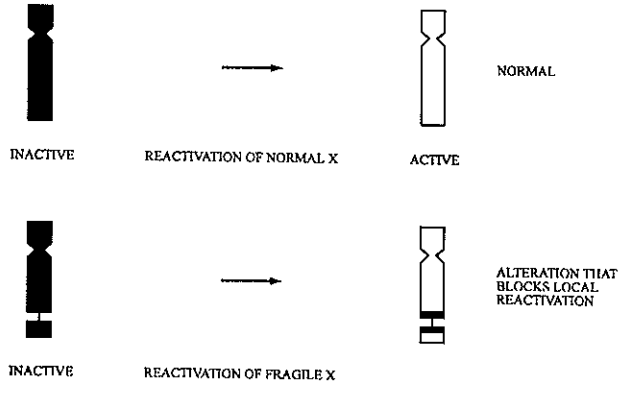


Figure 4. Theories that tried to explain the non-mendelian inheritance pattern in fragile X syndrome.



B: Amplification theory by Nussbaum. Adapted from Nussbaum et al., 1986 (54).
PRS = pyrimidine rich sequence



C: Incomplete reactivation theory by Laird. Adapted from Laird, 1987 (56).
A normal X chromosome is completely reactivated, but a mutation on the fragile X chromosome blocks reactivation in Xq27.3.

Pembrey suggested a premutation hypothesis (53) which is summarized in Figure 4A: a normal transmitting male would have a harmless "premutation" on his X chromosome that causes no phenotypic effect. This "premutation" is inherited by his daughter(s). At the same time this daughter inherits an X chromosome from her mother that has an unidentified recombination enhancing sequence. Recombination then occurs between these X chromosomes during meiosis in the oocytes of the daughter which causes a mutation resulting in the Martin-Bell phenotype in her children. This model predicts that recombination should be found systematically. However, genetic analysis with flanking polymorphic DNA markers demonstrated that this is clearly not the case.

Nussbaum proposed that unaffected male and female carriers might have an "initial lesion" on their X chromosome (54). This initial lesion would be a small amplification of a pyrimidine-rich sequence (46) that has occurred during non-homologous crossing over in meiosis during oogenesis in the mother (see Figure 4B). Offspring of these mothers would be normal transmitting individuals. Further amplification can then occur during pairing of the X chromosomes in oogenesis in the unaffected female carriers. This extensive amplification then causes the Martin-Bell syndrome in offspring of these females. In unaffected male carriers this amplification cannot occur, as they have only one X chromosome, and therefore they will have normal daughters. According to this model again recombinations would have to be found, which is not the case.

X chromosome inactivation in female mammals is part of the process of dosage compensation, which ensures that most cells have only one active X chromosome (lyonization, termed after Lyon (55)). Which chromosome remains active is determined by chance (55). During oogenesis both X chromosomes become active, which means that the inactive X has been reactivated. Laird postulated that the fragile X chromosome carries a mutation which prevents reactivation of this inactive X chromosome in the Xq27.3 region, which results in a local "imprinted block" on the fragile X chromosome (see Figure 4C). Through this imprinting one or more genes at Xq27.3 might not be transcribed (56), resulting in mental retardation in some of the offspring. Individuals who are normal carriers would have inherited the mutation without imprinting of the fragile X region and would therefore not be mentally retarded. Mothers and daughters of NTMs both are normal carriers, and according to this model both should have the same chance of affected offspring. This is however not the case in fragile X syndrome. In reality the chance of affected offspring for normal female carriers is closely related to the size of the CGG repeat in the fragile X gene (4,5; *Chapter 3, paragraph 3.4*).

REFERENCES

1. Webb T.P., Bundey S.E., Thake A.I. and Todd J. Population incidence and segregation ratios in the Martin-Bell syndrome. Am. J. Med. Genet. 23, 573-580, 1986.
2. Gustavson K.H., Blomquist H. and Holmgren G. Prevalence of the fragile-X syndrome in mentally retarded boys in a Swedish county. Am. J. Med. Genet. 23, 581-587, 1986.
3. Kähkönen M., Alitalo T., Airaksinen E., Matilainen R., Launiala K., Autio S. and Leisti J. Prevalence of the fragile X syndrome in four birth cohorts of children of school age. Hum. Genet. 77, 85-87, 1987.
4. 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. and Caskey C.T. Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. Cell 67, 1047-1058, 1991.
5. Heitz D., Devys D., Imbert G., Kretz C. and Mandel J.L. 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, 1992.
6. Martin J.P. and Bell J. A pedigree of mental defect showing sex-linkage. J. Neurol. Psych. 6, 154-157, 1943.
7. Lubs H.A. A marker X chromosome. Am. J. Hum. Genet. 34, 552-557, 1969.
8. Escalante J.A., Grunspun H. and Frota-Pessoa O. Severe sex-linked mental retardation. J. Génét. Hum. 19, 137-140, 1971.
9. Turner G., Eastman C., Casey J., McLeay A., Procopis P. and Turner B. X-linked mental retardation associated with macro-orchidism. J. Med. Genet. 12, 367-371, 1975.
10. Sutherland G.R. Marker X chromosomes and mental retardation. New Eng. J. Med. 296, 1415, 1977.
11. Sutherland G.R. Fragile sites on human chromosomes: demonstration of their dependence on the type of tissue culture medium. Science 197, 265-266, 1977.
12. Turner G., Daniel A. and Frost M. X-linked mental retardation, macro-orchidism, and the Xq27 fragile site. J. Pediatr. 96, 837-841, 1980.
13. Richards B.W., Sylvester P.E. and Brooker C. Fragile X-linked mental retardation: the Martin-Bell syndrome. J. Ment. Defic. Res. 25, 253-256, 1981.
14. Special issue on X-linked mental retardation. Am. J. Med. Genet. 17, 1-394, 1984; 23, 1-738, 1986; 30, 1-702, 1988; 38, 153-516, 1991; 43, 1-510, 1992.
15. Sutherland G.R. and Hecht F. Fragile sites on human chromosomes. Oxford University press, pp. 1-94, 1985.
16. Fish G.S., Arinami T., Froster-Iskenius U., Fryns J.P., Curfs L.M., Borghgraef M., Howard-Peebles P.N., Schwartz C.E., Simensen R.J. and Shapiro L.R. Relationship between age and IQ among fragile X males: a multicenter study. Am. J. Med. Genet. 38, 481-487, 1991.
17. Hagerman R.J. Physical and behavioral phenotype. In: Fragile X syndrome: diagnosis, treatment and research. Hagerman R.J. and Silverman A.C. (Eds.). The Johns Hopkins University Press, Baltimore and London, pp. 1-68, 1991.
18. Sutherland G.R. and Ashford P.L.C. X-linked mental retardation with macroorchidism and the fragile site at Xq27 or 28. Hum. Genet. 48, 117-120, 1979.

19. Fryns J.P. Clinical studies. In: The fragile X syndrome. Davies K.E. (Ed.) Oxford University Press, Oxford, pp. 1-39, 1989.
20. Partington M.W., Robinson H., Laing S. and Turner G. Mortality in the fragile X syndrome: preliminary data. Am. J. Med. Genet. 43, 120-123, 1992.
21. Partington M.W. The fragile X syndrome II: preliminary data on growth and development in males. Am. J. Med. Genet. 17, 175-194, 1984.
22. Butler M.G., Allen G.A., Singh D.N., Carpenter N.J. and Hall B.D. Preliminary communication: photoanthropometric analysis of individuals with the fragile X syndrome. Am. J. Med. Genet. 30, 165-168, 1988.
23. Butler M.G., Brunschwig A., Miller L.K. and Hagerman R.J. Standards for selected anthropometric measurements in males with the fragile X syndrome. Pediatr. 89, 1059-1062, 1992.
24. Schinzel A. and Largo R.H. The fragile X syndrome (Martin-Bell syndrome): clinical and cytogenetic findings in 16 prepubertal boys and in 4 of their 5 families. Helv. Paediatr. Acta 40, 133-152, 1985.
25. Storm R.L., De Benito R. and Ferretti C. Ophthalmologic findings in the fragile X syndrome. Arch. Ophthalmol. 105, 1099-1102, 1987.
26. Peyeritz R.E. In: Principle and practice of medical genetics, 2nd ed., pp. 1047-1063. Emery E.H. and Rimoin D.L. (eds.), Churchill Livingstone, New York, 1990.
27. Loehr J.P., Synhorst D.P., Wolfe R.R. and Hagerman R.J. Aortic root dilatation and mitral valve prolapse in the fragile X syndrome. Am. J. Med. Genet. 23, 189-194, 1986.
28. Hecht F. Seizure disorders in the fragile X chromosome syndrome. Am. J. Med. Genet. 38, 509, 1991.
29. Bergonzi P., Ferri R., Elia M. and Musumeci S.A. Epileptic seizures and fra(X) syndrome. Am. J. Med. Genet. 38, 510, 1991.
30. Musumeci S.A., Ferri R., Elia M., Colognola R.M., Bergonzi P. and Tassinari C.A. Epilepsy and fragile X syndrome: a follow-up study. Am. J. Med. Genet. 38, 511-513, 1991.
31. Wisniewski K.E., Segan S.M., Miezejeski C.M., Sersen E.A. and Rudelli R.D. The fra(X) syndrome: neurological, electrophysiological and neuropathological abnormalities. Am. J. Med. Genet. 38, 476-480, 1991.
32. Dunn H.G., Renpenning H., Gerrard J.W., Miller J.R., Tabata T. and Federofff S. Mental retardation as a sex-linked defect. Am. J. Ment. Defic. 67, 827-848, 1963.
33. Desai H.B., Donat J., Shokeir M.H.K. and Munoz D.G. Amyotrophic lateral sclerosis in a patient with fragile X syndrome. Neurol. 40, 378-380, 1990.
34. Fryns J.P., Moerman P., d'Espallier L. and Van den Berghe H. Suggestively increased incidence of infant death in children of fra(X) positive mothers. Am. J. Med. Genet. 30, 73-75, 1988.
35. Fisch G.S. Is autism associated with the fragile X syndrome? Am. J. Med. Genet. 43, 47-55, 1992.
36. Mazzocco M.M.M., Hagerman R.J. and Pennington B.F. Problem solving limitations among cytogenetically expressing fragile X women. Am. J. Med. Genet. 43, 78-86, 1992.
37. Lachiewicz A.M. Abnormal behaviors of young girls with fragile X syndrome. Am. J.

Med. Genet. 43, 72-77, 1992.

38. Fryns J.P., Haspeslagh M., Dereymaeker A.M., Volcke P. and Van den Berghe H. A peculiar subphenotype in the fra(X) syndrome: extreme obesity-short stature-stubby hands and feet-diffuse hyperpigmentation. Further evidence of disturbed hypothalamic function in the fra(X) syndrome? Clin. Genet. 32, 388-392, 1987.
39. De Vries B.B.A., Fryns J.P., Butler M.G., Wesby-van Swaay E., Oostra B.A., Halley D.J.J. and Niermeijer M.F. Clinical and molecular studies in fragile X patients with a Prader-Willi like phenotype. J. Med. Genet. 30, 761-766, 1993.
40. Beemer F.A., Veenema H. and Pater J.M. Cerebral gigantism (Sotos syndrome) in two patients with Fra(X) chromosomes. Am. J. Med. Genet 23, 221-226, 1986.
41. Opitz J.M. and Sutherland G.R. Conference report: International workshop on the fragile X and X linked mental retardation. Am. J. Med. Genet. 17, 5-94, 1984.
42. Jacky P.B., Ahuja Y.R., Anyane-Yeboa K., Breg W.R., Carpenter N.J., Froster-Iskenius U.G., Fryns J.P., Glover T.W., Gustavson K.H., Hoegerman S.F., Holmgren G., Howard-Peebles P.N., Jenkins E.C., Krawczun M.S., Neri G., Pettigrew A., Schaap T., Schonberg S.A., Shapiro L.R., Spinner N., Steinbach P., Vianna-Morgante A.M., Watson M.S. and Wilmot. Guidelines for the preparation and analysis of the fragile X chromosome in lymphocytes. Am. J. Med. Genet. 38, 400-403, 1991.
43. Nielsen K.B., Tommerup N., Poulsen H., Jacobsen P., Beck B. and Mikkelsen M. Carrier detection and X-inactivation studies in the fragile X syndrome. Hum. Genet. 64, 240-245, 1983.
44. Sutherland G.R. and Ledbetter D.H. Report of the committee on cytogenetic markers. Cytogenet. Cell Genet. 51, 452-458, 1989.
45. Knight S.J.L., Flannery A.V., Hirst M.C., Campbell L., Christodoulou Z., Phelps S.R., Pointon J., Middleton-Price H.R., Barnicoat A., Pembrey M.E., Holland J., Oostra B.A., Bobrow M. and Davies K.E. Trinucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation. Cell 74, 127-134, 1993.
46. Sutherland G.R. and Baker E. Effects of nucleotides on expression of the folate sensitive fragile sites. Am. J. Med. Genet. 23, 409-417, 1986.
47. Sutherland G.R., Baker E. and Fratini A. Excess thymidine induces folate sensitive fragile site. Am. J. of Med. Genet. 22, 433-443, 1985.
48. Sherman S.L., Morton N.E., Jacobs P.A. and Turner G. The marker (X) chromosome: a cytogenetic and genetic analysis. Ann. Hum. Genet. 48, 21-37, 1984.
49. Sherman S.L., Jacobs P.A., Morton N.E., Froster-Iskenius U., Howard-Peebles P.N., Nielsen K.B., Partington N.W., Sutherland G.R., Turner G. and Watson M. Further segregation of the fragile X syndrome with special reference to transmitting males. Hum. Genet. 289-299, 1985.
50. Moric-Petrovic S. and Laca Z. A father and daughter with fragile X chromosome. J. Med. Genet. 20, 476-478, 1983.
51. Mulley J.C., Yu S., Gedeon A.K., Donnelly A., Turner G., Loesch D., Chapman C.J., Gardner R.J.M., Richards R.I. and Sutherland G.R. Experience with direct molecular diagnosis of fragile X. J. Med. Genet. 29, 368-374, 1992.
52. Willems P.J., Van Roy B., De Boulle K., Vits L., Reyniers E., Beck O., Dumon J.E., Verkerk A. and Oostra B. Segregation of the fragile X mutation from an affected male

to his normal daughter. Hum. Mol. Genet. 1, 511-515, 1992.
53. Pembrey M.E., Winter R.M. and Davies K.E. A premutation that generates a defect at crossing over explains the inheritance of fragile X mental retardation. Am. J. Med. Genet. 21, 709-717, 1985.
54. Nussbaum R.L., Airhart S.D. and Ledbetter D.H. Recombination and amplification of pyrimidine-rich sequences may be responsible for initiation and progression of the Xq27 fragile site: an hypothesis. Am. J. Med. Genet. 23, 715-721, 1986.
55. Lyon M.F. Gene action in the X chromosome of the mouse. Nature 2, 372-373, 1961.
56. Laird C.D. Proposed mechanism of inheritance and expression of the human fragile-X syndrome of mental retardation. Genetics 117, 587-599, 1987.

CHAPTER 3

THE SEARCH FOR THE FRAGILE X GENE

The close association between the fragile X syndrome in males and the presence of the fragile site at Xq27.3 indicated that at least one gene involved in this disease must be located at, or very near to, the fragile site. Accordingly, the efforts of many laboratories have been aimed at obtaining DNA probes and fragments as close as possible to this fragile site, with the ultimate goal of cloning the gene involved in the fragile X syndrome. The search for the fragile X gene started with the finding of genetic markers segregating with the disease, followed by the isolation of physically closely linked markers that were used as a starting point for the isolation of the fragile X mutation. In this chapter an overview is given of the many strategies and approaches, used by other research groups and our own in the search for the fragile X gene.

3.1 Linkage analysis and the isolation of new markers

Genetic linkage

The frequent observation of a fragile site at Xq27 in mentally retarded males with macro-orchidism suggested a localization of the gene involved in the fragile X syndrome at the end of the long arm of the X chromosome. This hypothesis turned out to be correct, because in 1983 genetic linkage was reported between the glucose 6-phosphate dehydrogenase (G6PD)-colour blindness cluster at Xq28 and the fragile X locus in several fragile X pedigrees (1). In 1984 close genetic linkage was established in a large fragile X family between factor IX at Xq27 and the fragile X locus (2). No recombinants were observed in 15 informative meioses in this family which included an apparently non-penetrant grandfather. This placed the locus for the fragile X gene between factor IX at Xq27 and G6PD at Xq28. Other groups also reported families with tight linkage to factor IX, however, other pedigrees were found to show significant crossover frequencies (3,4,5). Brown suggested genetic heterogeneity as a possible explanation for this discrepancy (4,5). This could not be confirmed by others (6) and the apparent heterogeneity may reflect uncertainties due to limited statistics and mistyping of individuals in families.

Isolation of new markers

Pinning down a (disease) gene requires the isolation of DNA fragments or probes from the region of interest (see *Chapter 1*). Aiming at the fragile X gene, new probes were isolated continuously in the Xq27-q28 region. X chromosome specific probes were isolated from a genomic library of flow sorted human X chromosomes

(7,8,9), from libraries enriched for X chromosome sequences (10), from microdissection of the fragile X region (11,12), and from rodent/human hybrid cell lines containing the whole human X chromosome or parts of the human X chromosome against a mouse or hamster background (13,14). New hybrid cell lines were obtained containing discrete parts of the distal long arm of the X chromosome, to position markers more precisely to the Xq27-28 region and to establish the genetic location of the fragile X gene (10,15-20).
The first probes localized within a region of 5 cM from the fragile X locus on the distal site, U6.2 and VK21, were described by Dahl et al. (21) and Suthers et al. (22) in 1989, and on the proximal site the probes RN1 and VK23 were described by Oostra et al. (17) and Suthers et al. (23), in 1990 and 1991, respectively. These flanking markers have been very useful for carrier testing and prenatal diagnosis in a number of families (22) (*Chapter 4*). Later, other markers were isolated that mapped between these probes (10,18,24-26; *Publication 7.1*). The physical location, order and distances between probes were established using linkage analysis, somatic cell hybrids (9,10,13,16,18,19,22,27-30), *in situ* hybridization (31), and pulsed field gel electrophoresis (32,33).

3.2 Physical mapping

Construction and use of somatic cell hybrids
Somatic cell hybrids consist of hamster or mouse cells that contain a single (or a limited number of) human chromosome(s) or part of a human chromosome. These hybrids have played an important role in localizing and ordering probes on the human genome (34-38). Hybrids containing defined fragments of human chromosomes can provide information about the order of probes along the chromosome and can again be a source for cloning of human sequences from specific chromosomal regions (10,13-20). The usefulness of such cell hybrids was easily recognized and therefore many research groups started to isolate somatic rodent-human hybrid cell lines containing the whole or different parts of the human X chromosome (22,27,29,30,39,40). In Figure 1 the X chromosome content of a number of such hybrids is depicted. The rough location of a probe on a chromosome is determined by the presence or absence of this marker in a certain hybrid cell line. By hybridizing DNA probes to Southern blots of DNA made from various hybrids, the probes can be placed in groups on different fragments of the human chromosome and the order of the different grouped probes can be established. Additional information about the order of probes and the genetic

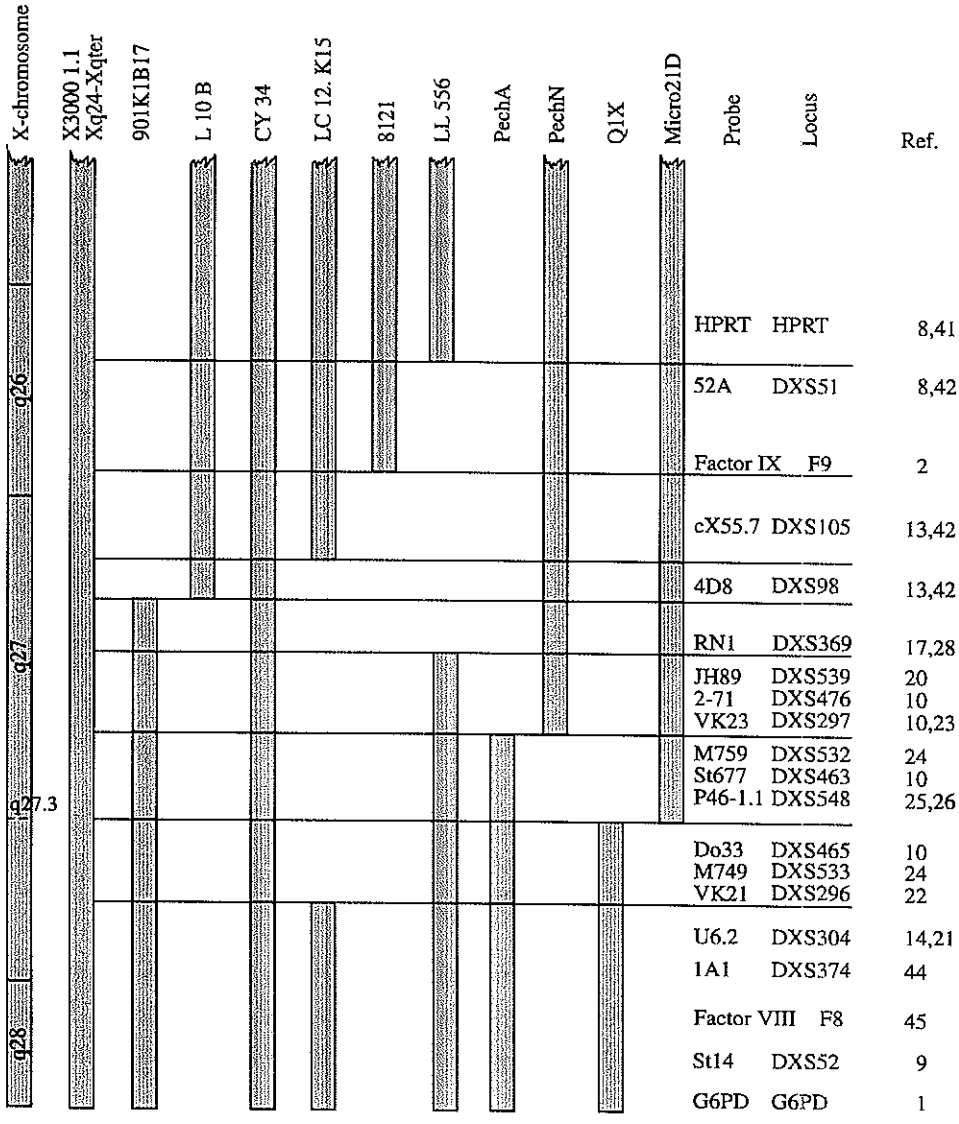


Figure 1. Localization of markers in the Xq26-q28 region utilizing somatic cell hybrids containing different regions of the X chromosome. Only the distal part of the Xq content of each hybrid is shown. Distances are not drawn to scale. Refs. for the cell lines: 22,28-30,40,46-51.

distances between the DNA markers can be obtained from linkage mapping. A number of probes were mapped in this way in the Xq27-28 region (see Figure 1). Subsequently, pulsed field gel electrophoresis can be used to establish a linear order of the probes within one group and to obtain a map of higher resolution with physical distances between probes in terms of kilobases. Two cell lines, termed Micro21D and Q1X, have been essential in the cloning of the fragile X gene (26; *Publication 7.1*). These human/hamster cell lines, each containing a reciprocal part of a human fragile X chromosome translocated to a hamster chromosome were developed by Warren et al. and were shown to break in the middle of the fragile X site (49-51). Micro21D contains the Xpter-Xq27.3 part of the human X chromosome, and Q1X contains the Xq27.3-Xqter part of the human X chromosome (Figure 1). It was hypothesized that by isolating the translocation breakpoints, which contain both human and hamster DNA sequences, it should be possible to obtain part of the fragile site and the fragile X gene. A cosmid library was constructed from the translocation hybrids, and filters were probed with total human genomic DNA and hamster genomic DNA, respectively. Although several hamster/human positive cosmids were isolated in this way (50,51), this method never led to the isolation of clones containing the fragile X region, which might be caused by the sequence composition of the fragile X mutation (26; *Publication 7.1*). However, these cell hybrids have been extremely useful for the mapping of X chromosomal markers (Figure 1), and have played a key role in the eventual identification of the fragile X gene, (26; *Publication 7.1*).

3.3 Long range mapping and CpG islands

A detailed physical map of a chromosomal area and distances in terms of kilobases between markers can be obtained by using pulsed field gel electrophoresis (PFGE). PFGE is based on the separation of long DNA fragments after digestion with infrequently cleaving restriction enzymes (52,53). Many of these enzymes (e.g. NotI, SacII, BssHII and EagI) have one or more CpG dinucleotides in their recognition site and as these sequences are under-represented in the genome (54), large DNA fragments can be generated. Often, these CpG sequences are methylated, in that case methylation occurs on cytosine.

Also clusters of CpG sequences, called CpG islands, are found (52,55,56). In general these islands are characterized by lack of methylation, except on the inactivated mammalian X chromosome (57,58) or in tissues where certain genes are inactive. CpG islands are usually found at the 5' end of mammalian genes and are

associated with the promoter regions of genes (55). Thus the presence of a cluster of "rare cutter restriction enzyme" sites is a strong indication of the presence of a promoter and an associated gene (58), and consequently the finding of CpG islands can lead to the identification of genes in the mammalian genome (19,26, 59-62; *Publication 7.1).*

Pulsed field map of Xq27-q28 and distribution of CpG islands

In the course of time several small pulsed field maps were established from the Xq27-28 region using only a small number of probes (32,44,56,63). A long-range physical map of the entire Xq27.2-qter region using PFGE analysis was developed by Poustka et al. (33). A total map, extending from the telomere to and beyond the fragile X site, of about 12 Megabases (Mb) was established. It was estimated that the fragile site is located approximately 9 Mb from the telomere.

Maestrini et al. identified CpG islands on the distal long arm of the X chromosome (19) and showed that there is a high density of CpG islands in the Xq24 and Xq27-28 regions (19,60). It was concluded that these regions are presumably rich in genes.

A more detailed physical map narrowing down on the fragile X region itself was obtained using probes mapping close to the breakpoints of the cell lines Q1X and Micro21D, which were thought to have breakpoints in the middle of the fragile site. The order of these probes was resolved by using hybrid cell lines and PFGE analysis and they were localized within a distance of 2-2,5 Mb that should contain the fragile site and thus the fragile X mutation (26,61,62; *Publication 7.1*). In

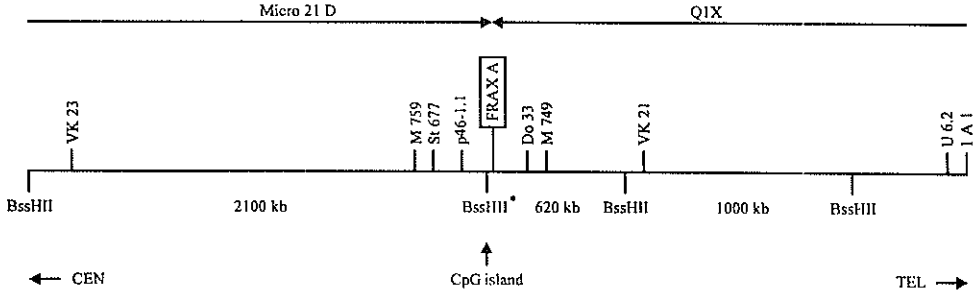


Figure 2. Physical map of the fragile X region constructed with tightly linked markers. Results from refs. 26, 61, 62, 64-66 and *Publication 7.1* are combined. The order of the probes is resolved using hybrid cell lines and PFGE analysis.

Figure 2 a physical map of the fragile X region is depicted. BssHII restriction sites and the length of the BssHII fragments are indicated.

Unexpected findings during long-range mapping:
methylation of DNA in mentally retarded fragile X males

The first difference noted between normal males and fragile X patients was methylation of a BssHII site in DNA from fragile X patients. In DNA from normal males and normal transmitting males a 620 kb BssHII fragment could be detected with the probes Do33 (61) or M749 (62) (Figure 2). In contrast, in DNA from the majority of unrelated fragile X patients this fragment was absent or present to a lesser extent. Instead of a normal 620 kb band, a large BssHII band appeared of about 2,7 Mb (2100 + 620 kb) (64) (Figure 2). As DNA from some male patients showed a less intense normal fragment, and the probe VK21 gave normal patterns with BssHII and other restriction enzymes, it was not likely that large deletions or insertions were present in the DNA of affected males (61,62,64). It seemed more likely that the absence or decreased intensity of the normal band was due to methylation of the BssHII site, indicated with an asterisk in Figure 2. The enzyme BssHII, which recognizes the sequence G/CGCGC, is sensitive to methylation and no digestion will occur in the case of methylation of this site in the DNA.
Further pulsed field gel analysis demonstrated that the BssHII site, which is methylated in patients, is part of a CpG island, indicating the presence of a possible neighbouring gene (26,61,62,65; *Publication 7.1*). It was suggested that the methylation of this CpG island might have an influence on the mRNA expression of this neighbouring gene. This assumption turned out to be true.

YACs spanning the Fragile X site

A major step forward was the development of yeast artificial chromosome cloning or YAC systems, which made it possible to clone and isolate DNA fragments from 100 kb to over 1 Mb (67,68). By isolating overlapping clones, a large stretch of human DNA from a specific chromosomal region can be obtained. Therefore, to bridge and clone the 2-3 Mb physically mapped fragile X area the YAC cloning technique was the obvious method of choice. YAC libraries became available from normal total human DNA (67,68), some enriched for X chromosomes (69; unpublished results), as well as from fragile X chromosomes (70).
By using markers tightly linked to the fragile site, shown in Figures 1 and 2, several YACs were isolated with insert sizes varying from 200 to 950 kb (26,64-66,71,72; *Publications 7.1 and 7.2*). Several of these YACs contained the translocation breakpoints of the cell lines Q1X and Micro21D generated by Warren et al. (26,51,

65,71; *Publication 7.1*). Fluorescence *in situ* hybridization of these YACs was performed on metaphase chromosomes from fragile X patients, which displayed the fragile X site. Fluorescent signals were scored on the proximal and on the distal site of the "fragile gap" (65,66,71,72; *Publication 7.2*) indicating that the YAC clones indeed spanned the fragile site. In *Publication 7.2*, YAC 209G4 is shown to cross the fragile site (72). In Figure 1A, page 147, an X chromosome (DAPI staining) from a male fragile X patient is shown, with a clearly visible fragile X site. In Figure 1B, page 147, YAC 209G4 is shown to span the fragile site of this X chromosome. This YAC was isolated from the YAC library constructed by Albertsen et al. (68) and was used by us to isolate the fragile X gene as described in *Publication 7.1* (26).

From YACs to the finding of variable fragments in fragile X individuals

PFGE mapping showed that the CpG island, which was found to be methylated in fragile X patients, was contained in the YACs spanning the fragile site. This CpG island contains the recognition sites for several rare cutting enzymes including BssHII, EagI and SacII, which are clustered within a very small region. The YACs were subcloned and genomic sequences around the CpG island were used as probes on Southern blots of DNA from patients and normal individuals. The abnormal methylation pattern was confirmed in affected males (73,74), while no methylation of the CpG island was found in normal males and normal transmitting males. In normal females about 50% methylation of the island was detected (as is expected with regard to X chromosome inactivation), but in affected females an excess of methylated fragments was seen. In addition to the methylation abnormalities, a fragment on the telomeric side of the CpG island was found to be increased in length in individuals carrying the fragile X mutation. This increase in length was variable in different fragile X individuals and was restricted to a small DNA fragment that also contained the CpG island (26,73,74; *Publication 7.1*). This DNA fragment turned out to be part of the fragile X gene, which we termed *FMR1*.

3.4 The fragile X mental retardation gene, *FMR1*

The isolation and identification of the fragile X mental retardation gene *FMR1* is described in detail in *Publication 7.1* (26). In this paragraph the characteristics of the gene will be discussed.

The *FMR1* gene has a genomic size of approximately 40 kb and consists of 17

exons and 16 introns, and is transcribed in a proximal to distal orientation (75). In Figure 3 the intron/exon distribution of *FMR1* is given. The length of the mRNA is about 4,4 kb, with an open reading frame of about 1,9 kb. *FMR1* mRNA is particularly abundant in brain and testis, concordant with the pathology found in fragile X patients. In the 5′ untranslated region of the *FMR1* gene an unusual CGG trinucleotide repeat is present in the first exon. In the original *FMR1* cDNA clone the CGG repeat is interspersed with 2 AGG triplets, but as most triplets consist of CGG sequence, the repeat is referred to as CGG repeat. The ATG start codon for translation is found 69 bases after the CGG repeat, indicating that the repeat is not translated into protein (76,77). The previously identified CpG island mentioned above (61,62,73,74) is located 250 bp proximal of the repeat. This CpG island as well as the CGG repeat itself are methylated in DNA of fragile X patients (78-80). No methylation of the island or the repeat is found in DNA of normal males and normal transmitting males. Methylation of a CpG island is usually correlated with lack of gene expression. Indeed, in fragile X males, in whom methylation of the CpG island of the *FMR1* was found, no mRNA is transcribed from the *FMR1* gene (81,82). In some patients with partial methylation of the CpG island a lowered RNA expression was demonstrated (81). These patients seemed not less severely affected (83).

The *FMR1* gene was shown to have been conserved in evolution and is found in many animal species, even in species as divergent as yeast and *Caenorhabditis elegans* (26,76,84,85; *Publication 7.1*). There is a significant homology between the mouse and human gene sequences, with a nucleotide identity of 95% within the coding region (76). At the protein level even a 97% amino acid identity is found. The order of genes found in the human Xq27 region (*F9-FMR1-IDS-GABRA3*) has also been conserved in the mouse (86).

The CGG repeat, which is confined to the 5′ untranslated region of the *FMR1* gene and which is not translated into protein, has also been evolutionary conserved in many species (76,85,87). This might point to an important regulatory role, at the DNA and/or RNA level. Richards et al. demonstrated that double and single stranded CGG repeats and other simple repeat sequences are binding sites for specific nuclear proteins (80).

Alternative splicing and the fragile X protein, FMRP

In general, splicing involves removing introns from a transcript and forming a bond between the ends of exons, resulting in one species of mRNA molecules and subsequently in one species of protein. The *FMR1* gene, however, is subject to alternative splicing, as is described in detail in *Publication 7.4* (88). Due to

alternative splicing of the precursor mRNA at 3 different locations in the gene, the *FMR1* gene can give rise to as many as 12 partly different mRNA molecules and thus to 12 possible proteins, which differ in various segments. In these splicing events, involving exons 12, 15 and 17 (Figure 3), the open reading frame is maintained. The longest possible transcript codes for a protein of 631 amino acids, the smallest for 568 amino acids. In addition, at the 3' end of the gene alternative use of different polyadenylation signals is found, resulting in a longer or shorter 3' untranslated region. The alternative splicing in *FMR1* does not seem to be tissue specific, for no differences were detected between the various tissues analyzed. For more details the reader is referred to *Publication 7.4.* In addition to the splicing described in *Publication 7.4,* Ashley et al. (76) also identified alternative splicing involving exon 14 in both human and mouse. In the mouse this was further analyzed and it was found that exclusion of exon 14 causes a frameshift in the *Fmr1* (mouse) open reading frame. Exon 13 can be joined to several splice acceptor sites in exon 15. Multiple protein isoforms are predicted with new amino acid sequences at the C terminus (Figure 3).

Knowledge of the sequence of the *FMR1* gene gave us the opportunity to synthesize synthetic peptides (15-30 amino acids) which were used to raise antibodies against the *FMR1* protein (*FMRP*) in rabbit. These antibodies recognize

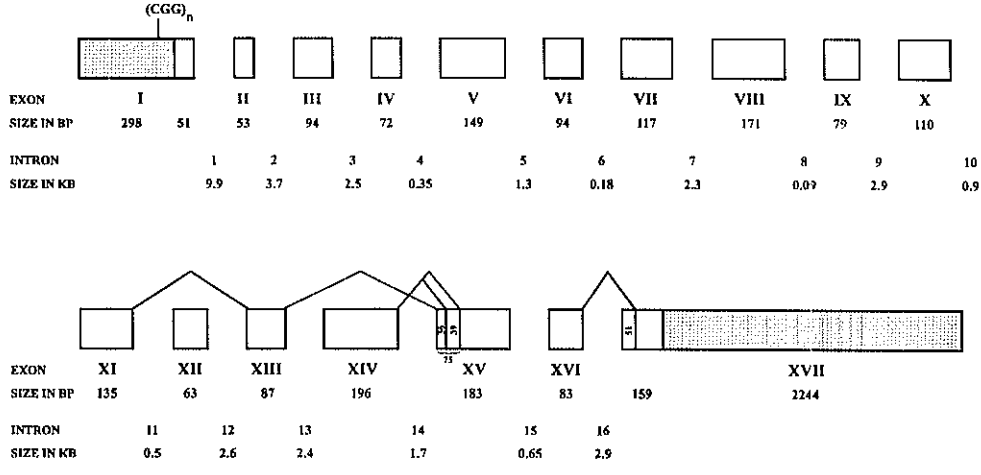


Figure 3. Intron/exon distribution in the *FMR1* gene. Exons are depicted as boxes and numbered in roman numerals. Sizes of introns and exons are indicated. Shaded portions at the 3' and 5' ends signify untranslated regions of the gene. Alternative splicing is depicted. Adapted from Eichler et al., 1993 (75,88; *Publication 7.4*).

more than one protein product in white blood cells of healthy individuals, confirming the alternative splicing in *FMR1* (77,88,89; *Publication 7.4*). In patients, not expressing *FMR1* mRNA, no *FMRPs* could be detected. Using immunofluorescence, the intracellular localization of different splice variants of *FMRP* was investigated in COS cells (monkey kidney cells) and was found to be predominantly cytoplasmic (77,89), in spite of the presence of a putative nuclear translocation signal in the C-terminal part of the gene (26; *Publication 7.1*). Unexpectedly, Devys et al. found a predominantly nuclear localization when COS cells were transfected with expression vector containing only the N-terminal half of the *FMR1* coding sequence (89).

A lot is still to be learned about the function of the normal *FMR1* protein and the relation between the absence of *FMRP* and the pathology in fragile X syndrome patients. Initially, no homology was found between the predicted *FMRP* sequence and any other known protein in database analyses (26; *Publication 7.1*). Recently, two domains have been identified resembling a motif thought to be involved in RNA binding (84,90). *FMRP* indeed was found to have RNA binding activity (84,90; *paragraph 3.5*).

mRNA expression

Various levels of *FMR1* mRNA expression have been found in different organs (26,91-93; *Publication 7.1*), with high expression in human adult testis and brain (91). Studies of mRNA expression by *in situ* hybridization to human tissues of 8 and 9 weeks old fetuses showed high expression in the brain, especially in the cerebellum and hippocampus, spinal cord, ganglia, neural retina and cartilaginous structures, but not in testis (92). Further study of human brain tissue of 25 week old fetuses showed a high level of *FMR1* mRNA in differentiating neurons in numerous cerebral structures (92). In the mouse high levels of mRNA are found during early embryogenesis (91). High expression of *Fmr1* was found in murine spermatogonia, which suggests a role of *FMR1* in sperm development (94).

Basis of the fragile X mutation:
variation in the number of CGG repeats and different classes of mutation

With the discovery in 1991 of a trinucleotide repeat in the *FMR1* gene (26; *Publication 7.1*) a new type of mutational mechanism has been revealed. Transgenerational amplification of this repeat appeared to be the molecular genetic basis of the fragile X syndrome.

PCR and sequence analysis showed that the number of CGG repeats in the *FMR1*

gene is found to be polymorphic in the human, mouse and pig (85,87). The repeat in the published *FMR1* gene contained two AGG sequences (26; *Publication 7.1*). Not all *FMR1* alleles studied contain AGG sequences, and nothing is known about any distribution of AGG triplets in the *FMR1* CGG repeat in different individuals.

In normal individuals the number of repeats was found to vary from 6-54, with an average of 29 (87). In fragile X syndrome carriers and patients there is a considerable increase in size of the region containing the CGG repeat. (26,73,74,87,95; *Publication 7.1*). No differences were observed in sequences immediately flanking the CGG repeats between fragile X patients, carriers and normal individuals (87), suggesting that the increase in DNA fragment size is indeed restricted to the CGG repeat itself and that the *FMR1* gene might have a primary role in the pathogenesis of the fragile X syndrome. Analysis of CGG repeat PCR products on sequencing gels showed spacing of bands at 3 bp intervals (87) and sequencing within the repeat region confirmed amplification of the CGG repeat itself (79).

Two classes of DNA abnormalities can be distinguished in fragile X syndrome.

1. In normal transmitting males and their normal carrier daughters the number of CGG repeats in the *FMR1* gene varies from 43-200. This is termed a ***premutation,*** as this number of repeats exceeds the normal range and is not directly associated with the disease phenotype (73,74,87,95). Premutation alleles are normally transcribed (81).
2. In male patients more than 200 repeats are found in the *FMR1* gene, with some expansions exceeding 2000 repeats (73,83,87,96,97; *Publication 7.3).* This number of repeats is directly associated with the disease phenotype and is called a ***full mutation.*** In case of a full mutation, the promoter region c.q. CpG island of the *FMR1* gene and the CGG repeat are methylated (78-80). This results in repression of *FMR1* transcription (81) and absence of the *FMR1* protein (77), which leads to the fragile X phenotype.

Female carriers have either a premutation or a full mutation. Females carrying a premutation are always phenotypically normal. About 50-70% of females carrying a full mutation on one of their X chromosomes are mildly to moderately affected (96). No correlation in lymphocytes has been demonstrated between skewing of X inactivation and mental status in affected females. This might be different in the brain.

In Figure 4 the molecular basis of the fragile X syndrome is shown schematically.

The length of the CGG repeat is correlated with the percentage of cells in which cytogenetic expression is found (83). Normal transmitting males and females with a premutation never show the fragile site, whereas cells from practically all male patients with a full mutation show cytogenetic expression. Only in a number of the

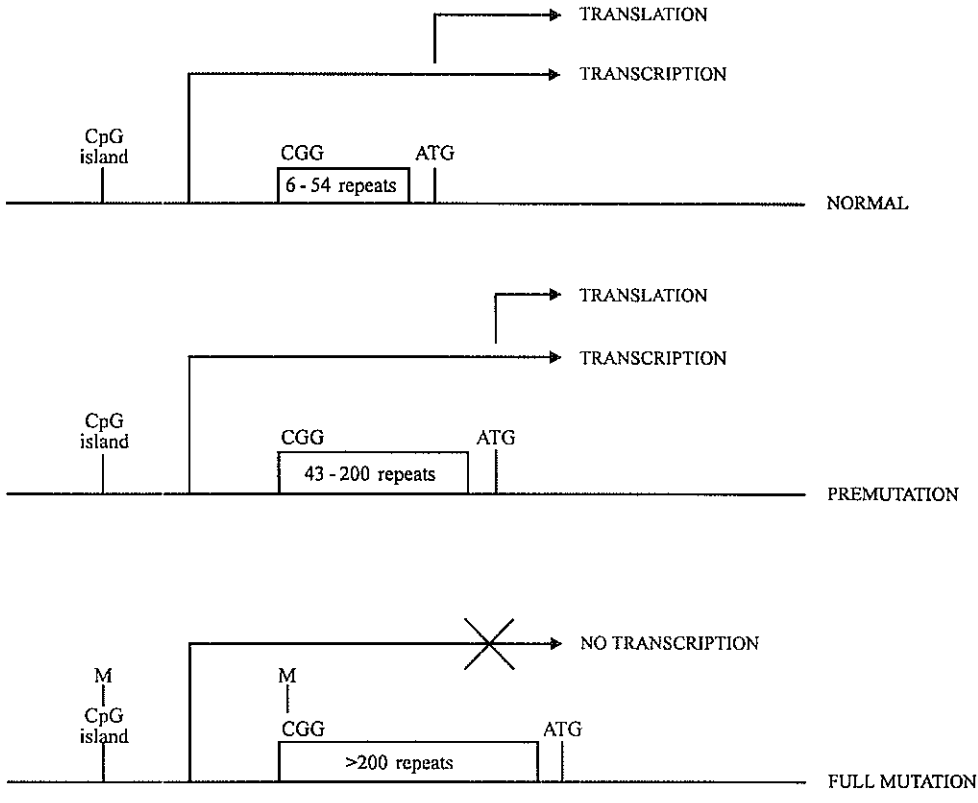


Figure 4. Schematic representation of the *FMR1* gene and the molecular basis of the fragile X syndrome.

Premutation alleles are found in normal transmitting males and in normal female carriers. No methylation of the CpG island is found, resulting in normal transcription of the gene and a normal phenotype.

Full mutations are found in 100% of affected males. Of the females carrying a full mutation about 50% - 75% are mildly to moderately mentally retarded (96). In full mutations the CpG island of the *FMR1* gene is methylated (M) and transcription is absent, resulting in the disease phenotype.

Table 1. Phenotype, genotype, mRNA and cytogenetic expression of the *FMR1* gene in different individuals.

Allele	Phenotype	Number of repeats	MR	mRNA expression	Cytogenetic expression %
normal	normal	6-54	-	+	-
premutation	normal transmitting males	43-200	-	+	-
	female premutation carriers	43-200[a]	-	+	-
full mutation	male fragile X patients	> 200	+	-	100%
	female full mutation carriers	> 200[a]	+[b]	-[c]	> 50%

MR = mental retardation

[a] in addition these females also have a normal allele.

[b] in about 50-70% of females carrying a full mutation mild to moderate mental retardation is found.

[c] in females *FMR1* mRNA is only transcribed from the normal active X.

females with a full mutation, the fragile site can be detected.

Table 1 gives an overview of the differences between normal, premutation and full mutation alleles.

Instability and mosaicism of the CGG repeat

Previously unexplained findings, such as the Sherman paradox, and the transmission of an X-linked disease by apparently healthy males, could be clarified by the elucidation of the molecular defect in the fragile X syndrome.

Normal alleles containing 6-54 repeats are stable. This means that there is no change in repeat size when the gene is passed on to the next generation. Premutation alleles containing 43-200 repeats are unstable and may increase in size with transmission to the next generation (73,74,87,95-97; *Publication 7.3*). Occasionally a decrease in size is observed (73,87,96,98). There is an overlap in the number of CGG repeats between normal and premutation alleles. These can be distinguished by studying the (in)stability of the alleles in a pedigree. Premutation

alleles transmitted by normal transmitting males always remain in the premutation range in their daughters. Therefore daughters of normal transmitting males are always phenotypically normal (*Chapter 6*). Premutation alleles transmitted by a female carrier may change to a larger premutation or a full mutation in the offspring. Only premutation alleles transmitted by females have a chance of increasing to a full mutation in their offspring. It is still unclear why the expansion of CGG repeats is limited to female germline transmission.
The risk of a premutation in a female to expand to a full mutation in the next generation is correlated with the size of the premutation allele of the mother. A low repeat number (50-65 repeats) has a low risk, and a high repeat number (70-80 repeats) has a high risk of expanding to a full mutation (73,74,87,99). When a premutation contains more than circa 90 repeats, the chance that it becomes a full mutation in the next generation is almost 100% (87). These risks correspond to the risks of mental retardation calculated by Sherman for individuals in a fragile X pedigree (87).

Due to mitotic instability of the CGG repeat, length mosaicism is frequently observed in individuals carrying a fragile X mutation when somatic DNA is studied.
Two kinds of mosaicism are observed in fragile X individuals:

1. Extensive somatic instability is observed in individuals carrying a full mutation (73,96,97; *Publication 7.3*). Mitotic divisions may result in a variable number of repeats in different cells of the same individual. This is visualized in a Southern blot analysis as a "smear" instead of a distinct band. Several such individuals are described in *Publication 7.3*, page 157. For example in Figure 1A, in patients 3 and 6; and in Figure 1B, in patient 4 and in a normal female carrier 5.
2. In 25% (81) of male patients, in addition to a full mutation a premutation is found in part of their cells (73,97; *Publication 7.3*). Despite transcription of the (unmethylated) premutation alleles these males are mentally retarded. It is assumed that the absence of the *FMR1* protein in relevant cell types is responsible for the disease phenotype. In female carriers this kind of mosaicism is also found. An example of a male patient mosaic for a premutation and a full mutation is shown in *Publication 7.3*, page 158, Figure 2A, individual 3.

Evidence accumulates that the somatic mosaicism as seen in adults is established during early embryogenesis, as identical patterns of mosaicism were found in different tissues of a 13 week old fetus, in adult males and in monozygotic twins (100,101). If mitotic instability would occur in every mitotic division during life,

proliferation of a single cell would probably result in variable random patterns in different tissues. In vitro studies of proliferation of a single cell with one particular size demonstrated that the repeat length was stably maintained in clonal cell lines after as many as 25 doublings (100). These data together point in the direction that the mosaicism is established early in development.

3.5 Fragile X patients without CGG repeat amplification in the *FMR1* gene

Although an abnormally increased repeat length of the *FMR1* CGG repeat is the hall mark of most fragile X patients, there are a few exceptions. Several patients have been identified with the typical clinical phenotype of the fragile X syndrome, but without CGG repeat expansion or cytogenetic expression of the fragile site. In one such patient a *de novo* deletion of approximately 250 kb was found that included the CpG island and at least five exons of the *FMR1* gene (102). In another patient a deletion of more than 2 Mb of chromosomal DNA, including the entire *FMR1* gene, was found (103). The possible involvement of other genes in the deletion, contributing to the disease phenotype could not be excluded.
One patient has recently been described (104; *Publication 7.5*) with the fragile X phenotype, without CGG repeat expansion and with normal transcription of the *FMR1* gene. In the *FMR1* gene of this patient a single *de novo* point mutation was found, resulting in an aberrant protein causing the fragile X phenotype. The search for and identification of this mutation by means of reverse transcriptase-PCR and sequencing is described in detail in *Publication 7.5*.
Recently, RNA binding proteins have been identified in several organisms that contain specific RNA binding domains, called RGG boxes and KH domains (84). Sequences for an RGG box and two KH domains have now also been identified in the *FMR1* gene (84,90). The point mutation in the *FMR1* gene of the patient described in *Publication 7.5* is located in the middle of the second KH domain. However, the aberrant *FMR1* protein from this patient, did not lose its RNA binding activity (Verheij, unpublished results).
The observation of a point mutation in the *FMR1* gene of this patient supports the hypothesis that the fragile X syndrome is a single gene disorder and that loss of function of or a mutation in the *FMR1* gene leads to the clinical phenotype of the fragile X syndrome.
Analysis at the molecular level of other patients with the fragile X phenotype but without CGG repeat amplification and without cytogenetic expression must reveal whether there might exist a variety of additional mutations in the *FMR1* gene. A

small number of such patients have now been analyzed at the RNA level, and no mutations have been found in the *FMR1* gene of these males (105). One explanation for these findings could be that these patients closely resemble fragile X patients, but in reality are suffering from another syndrome. A second possibility is that normal *FMRP* binds to specific RNA(s), which themselves could be mutated and lead to the fragile X phenotype.

3.6 Other fragile sites on the X chromosome

The majority of individuals with a fragile site at Xq27.3 and mental retardation have been confirmed to have the fragile X syndrome, by identifying an increased CGG repeat in the *FMR1* gene. However, several families have been described in which individuals exhibiting a fragile site at Xq27 do not have a repeat

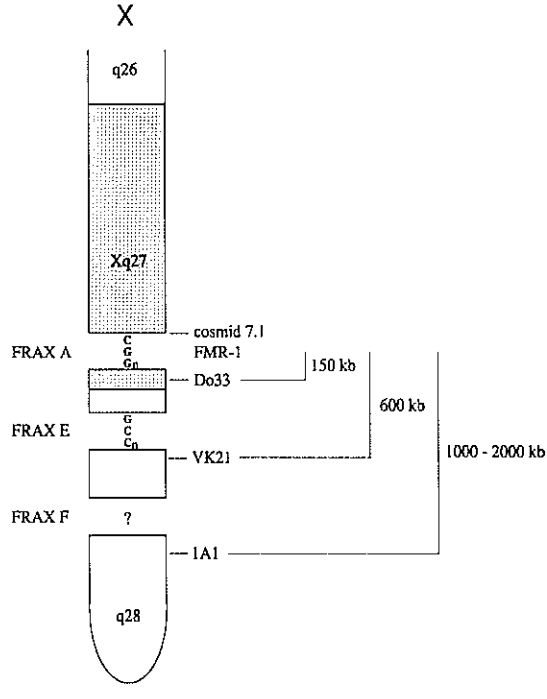


Figure 5. Different fragile sites are indicated in the Xq27-28 region, FRAXA containing a CGG repeat, FRAXE containing a GCC repeat, and FRAXF of which the molecular composition is unknown. These sites can be distinguished by fluorescence *in situ* hybridization with different probes.

amplification in the *FMR1* gene, even though they show cytogenetic expression of the fragile site in 5 to 75% of the cells examined (106-110). In most cases, after the finding of a normal CGG repeat in the *FMR1* gene, the affected individuals were clinically reexamined, and were assessed to have a non-fragile X phenotype involving mild mental retardation. In these patients a neighbouring fragile site was found in Xq27.3.

Refined cytogenetic and DNA based methods have allowed differentiation of two other different fragile sites, called FRAXE (111-113) and FRAXF (114), in different families that were initially cytogenetically indistinguishable from FRAXA (111,114). These sites can be distinguished by fluorescence *in situ* hybridization using different DNA probes or cosmids from the Xq27-q28 region on fragile site induced metaphase chromosomes. In Figure 5 this is schematically depicted. The FRAXA site is located between cosmid 7.1 (a subclone from YAC 209G4) (72; *Publication 7.2*) and Do33 (65). The FRAXE fragile site is located between Do33 and VK21 and is located 600 kb distal to FRAXA (111,112). FRAXF is located between VK21 and 1A1 and is located more than 1000 kb distal from FRAXA (114).

The FRAXE site is associated with mild mental retardation in patients and an extended GCC repeat (>200 repeats) (113). In normal individuals, 6-25 copies of the GCC sequence are found with an average of 15. Essentially the FRAXA and FRAXE repeat are similar in sequence. The CpG island associated with FRAXE is located 600 kb distal to the *FMR1* associated CpG island and is methylated in FRAXE patients. This suggests that methylation may play a role in the inactivation of a postulated gene associated with FRAXE resulting in mental impairment. FRAXE patients have a normal methylation pattern across the FRAXA region (113). The molecular nature of FRAXF is still unknown. Cloning this region of fragility will determine whether the FRAXF region also contains a trinucleotide repeat amplification.

REFERENCES

1. Filippi G., Rinaldi A., Archidiacono N., Rocchi M., Balazs I and Siniscalco M. Brief report: linkage between G6PD and Fragile-X syndrome. Am. J. Med. Genet. 15, 113-119, 1983.
2. Camerino G., Mattei M.G., Mattei J.F., Jaye M. and Mandel J.L. Close linkage of fragile X-mental retardation syndrome to haemophilia B and transmission through a normal male. Nature 306, 701-704, 1983.
3. Warren S.T., Glover T.W., Davidson R.L. and Jagadeeswaran P. Linkage and recombination between fragile X-linked mental retardation and the factor IX gene.

Hum. Genet. 69, 44-46, 1985.
4. Brown W.T., Gross A.C., Chan C.B. and Jenkins E.C. Genetic linkage heterogeneity in the fragile X syndrome. Hum. Genet. 71, 11-18, 1985.
5. Brown W.T., Gross A., Chan C., Jenkins E.C., Mandel J.L., Oberlé I., Arveiler B., Novelli G., Thibodeau S., Hagerman R., Summers K., Turner G., White B.N., Mulligan L., Forster-Gibson C., Holden J.J.A., Zoll B., Krawczak M., Goonewardena P., Gustavson K.H., Pettersson U., Holmgren G., Schwartz C., Howard-Peebles P.N., Murphy P., Breg W.R., Veenema H. and Carpenter N.J. Multilocus analysis of the fragile X syndrome. Hum. Genet. 78, 201-205, 1988.
6. Suthers G.K., Mulley J.C., Voelckel M.A., Dahl N., Vaisanen M.L., Steinbach P., Glass I.A., Schwartz C.E., van Oost B.A., Thibodeau S.N., Haites N.E., Oostra B.A., Schinzel A., Carballo M., Morris C.P., Hopwood J.J. and Sutherland G.F. Linkage homogeneity near the fragile X locus in normal and fragile X families. Genomics 10, 576-582, 1991.
7. Davies K.E., Young, B.D., Elles R.G., Hill M.E. and Williamson R. Cloning of a representative genomic library of the human X chromosome after sorting by flow cytometry. Nature 293, 374-376, 1981.
8. Drayna D., Davies K., Hartley D., Mandel J.L., Camerino G., Williamson R. and White R. Genetic mapping of the human X chromosome by using restriction fragment length polymorphisms. Proc. Natl. Acad. Sci. USA 81, 2836-2839, 1984.
9. Oberlé I., Drayna D., Camerino G., White R. and Mandel J.L. The telomeric region of the human X chromosome long arm: presence of a highly polymorphic DNA marker and analysis of recombination frequency. Proc. Natl. Acad. Sci. USA 82, 2824-2828, 1985.
10. Rousseau F., Vincent A., Rivella S., Heitz D., Triboli C., Maestrini E., Warren S.T., Suthers G.K., Goodfellow P., Mandel J.L., Toniolo D. and Oberlé I. Four chromosomal breakpoints and four new probes mark out a 10 cM region encompassing the Fragile-X locus (FRAXA). Am. J. Hum. Genet. 48, 108-116, 1991.
11. MacKinnon R.N., Hirst M.C., Bell M.V., Watson J.E.V., Claussen U., Ludecke H.J., Senger G., Horsthemke B. and Davies K.E. Microdissection of the Fragile X region. Am. J. Hum Genet. 47, 181-187, 1990.
12. Djabali M., Nguyen C., Biunno I., Oostra B.A., Mattei M.G., Ikeda J.E. and Jordan B.R. Laser microdissection of the Fragile X region: identification of cosmid clones and of conserved sequences in this region. Genomics 10, 1053-1060, 1991.
13. Hofker M.H., Bergen A.A.B., Skraastad M.I., Carpenter N.J., Veenema H., Connor J.M., Bakker E., van Ommen G.J.B. and Pearson P.L. Efficient isolation of X-chromosome-specific single-copy probes from a cosmid library of a human X/hamster hybrid-cell line: mapping of new probes close to the locus for X-linked mental retardation. Am. J. Hum. Genet 40, 312-328, 1987.
14. Dahl N., Hammarstrom-Heeroma K., Goonewardena P., Wadelius C., Gustavson K.H., Holmgren G., van Ommen G.J.B. and Pettersson U. Isolation of a DNA probe of potential use for diagnosis of the fragile-X syndrome. Hum. Genet. 82, 216-218, 1989.
15. Poustka A., Rackwitz H.R., Frischauf A.M., Hohn B. and Lehrach H. Selective isolation of cosmid clones by homologous recombination in Escherichia coli. Proc. Natl. Acad. Sci. USA 81, 4129-4133, 1984.

16. Hyland V.J., Fernandez K.E.W., Callen D.F., MacKinnon R.N., Baker E., Friend K. and Sutherland G.R. Assignment of anonymous DNA probes to specific intervals of human chromosomes 16 and X. Hum. Genet. 83, 61-66, 1989.
17. Oostra B.A., Hupkes P.E., Perdon L.F., van Bennekom C.A., Bakker E., Halley D.J.J., Schmidt M., du Sart D., Smits A., Wieringa B. and van Oost B.A. New polymorphic DNA marker close to the fragile site FRAXA. Genomics 6, 129-132, 1990.
18. Hulsebos T.J.M., Oostra B.A., Broersen S., Smits A., van Oost B.A. and Westerveld A. New distal marker closely linked to the fragile X locus. Hum. Genet. 87, 369-372, 1991.
19. Maestrini E., Rivella S., Tribioli C., Purtilo D., Rocchi M., Archidiacono N. and Toniolo D. Probes for CpG islands on the distal long arm of the human X chromosome are clustered in Xq24 and Xq28. Genomics 8, 664-670, 1990.
20. Dreesen J.C.F.M., van den Hurk J.A.J.M., Smits A.P.T., van den Ouweland A.M.W., Markslag P.W.B. and van Oost B.A. DXS539, a polymorphic DNA marker proximal of the fragile-X gene. Hum. Genet. 91, 80-82, 1993.
21. Dahl N., Goonewardena P., Malmgren H., Gustavson K.H., Holmgren G., Seemanova E., Anneren G., Flood A. and Pettersson U. Linkage analysis of families with Fragile-X mental retardation, using a novel RFLP marker (DXS304). Am. J. Hum. Genet. 45, 304-309, 1989.
22. Suthers G.K., Callen D.F., Hyland V.J., Kozman H.M., Baker E., Eyre H., Harper P.S., Roberts S.H., Hors-Cayla M.C., Davies K.E., Bell M.V. and Sutherland G.R. A new DNA marker tightly linked to the Fragile X locus (FRAXA). Science 246, 1298-1299, 1989.
23. Suthers G.K., Mulley J.C., Voelckel M.A., Dahl N., Vaisanen M.L., Steinbach P., Glass I.A., Schwartz C.E., van Oost B.A., Thibodeau S.N., Haites N.E., Oostra B.A., Gine R., Carballo M., Morris C.P., Hopwood J.J. and Sutherland G.R. Genetic mapping of new DNA probes at Xq27 defines a strategy for DNA studies in the Fragile X syndrome. Am. J. Hum. Genet. 48, 460-467, 1991.
24. Hirst H.C., Roche A., Flint T.J., Mackinnon R.N., Bassett J.H.D., Nakahori Y., Watson J.E., Bell M.V., Patterson M.N., Boyd Y., Thomas N.S.T., Knight S.J.L., Warren S.T., Hors-Cayla M., Schmidt M. and Davies K.E. Linear order of new established markers around the fragile site at Xq27.3. Genomics 10, 243-249, 1991.
25. Riggins G.J., Sherman S.L., Oostra B.A., Sutcliffe J.S., Feitell D., Nelson D.L., van Oost B.A., Smits A.P.T., Ramos F.J., Pfendner E., Kuhl D.P.A., Caskey C.T. and Warren S.T. Characterization of a highly polymorphic dinucleotide repeat 150 kb proximal to the fragile X site. Am. J. Med. Genet. 43, 237-243, 1992.
26. 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. and Warren S.T. 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, 1991.
27. Wieacker P., Davies K.E., Cooke H.J., Pearson P.L., Williamson R., Bhattacharya S., Zimmer J. and Ropers H.H. Toward a complete linkage map of the human X

chromosome: regional assignment of 16 cloned single-copy DNA sequences employing a panel of somatic cell hybrids. Am. J. Hum. Genet. 36, 265-276, 1984.

28. Oostra B.A., Majoor-Krakauer D.F., van Hemel J.O., Bakker E., Callen D.F., Schmidt M. and van Oost B.A. Mapping of a new RFLP marker RN1 (DXS369) close to the fragile site FRAXA on Xq27-q28. Am. J. Med. Genet. 38, 332-335, 1991.
29. Suthers G.K., Hyland V.J., Callen D.F., Oberlé I., Rocchi M., Thomas N.S., Morris C.P., Schwartz C.E., Schmidt M., Ropers H.H., Baker E., Oostra B.A., Dahl N., Wilson P.J., Hopwood J.J. and Sutherland G.R. Physical mapping of new DNA probes near the fragile X mutation (FRAXA) by using a panel of cell lines. Am. J. Hum. Genet. 47: 187-195, 1990.
30. Oberlé I., Vincent A., Abbadi N., Rousseau F., Hupkes P.E., Hors-Cayla M.C., Gilgenkrantz S., Oostra B.A. and Mandel J.L. New polymorphism and a new chromosome breakpoint establish the physical and genetic mapping of DXS369 in the DXS98-FRAXA interval. Am. J. Med. Genet. 38, 336-342, 1991.
31. Suthers G.K., Hyland V.J., Callen D.F., Baker E., Eyre H., Roberts S.H., Hors-Cayla M.C., Davies K.E., Bell M.V. and Sutherland G.R. Physical and genetic mapping of new DNA probes at Xq26-qter. Cytogenet. Cell Genet. 51, 1087, 1989.
32. Patterson M., Kenwrick S., Thibodeau S., Faulk K., Mattei M.G., Mattei J.F. and Davies K.E. Mapping of DNA markers close to the fragile site on the human X chromosome at Xq27.3. Nucleic Acids Res. 15, 2639-2651, 1987.
33. Poustka A., Dietrich A., Langenstein G., Toniolo D., Warren S.T. and Lehrach H. Physical map of human Xq27-qter: localizing the region of the fragile X mutation. Proc. Natl. Acad. Sci. USA 88, 8302-8306, 1991.
34. Willard H.F., Goss S.J., Holmes M.T. and Munroe D.L. Regional localization of the phosphoglycerate kinase gene and pseudogene on the human X chromosome and assignment of a related DNA sequence to chromosome 19. Hum. Genet. 71, 138-143, 1985.
35. Benham F., Hart K., Crolla J., Bobrow M., Francavilla M. and Goodfellow P.N. A method for generating hybrids containing nonselected fragments of human chromosomes. Genomics 4, 509-517, 1989.
36. Cremers F.P.M., van de Pol T.J.R., Wieringa B., Hofker M.H., Pearson P.L., Pfeiffer R.A., Mikkelsen M., Tabor A. and Ropers H..H. Molecular analysis of male-viable deletions and duplications allows ordering of 52 DNA probes on proximal Xq. Am. J. Hum. Genet. 43, 452-461, 1988.
37. Nelson D.L., Ballabio A., Victoria M.F., Pieretti M., Bies R.D., Gibbs R.A., Maley J.A., Chinault A.G., Webster T.D. and Caskey C.T. Alu-primed polymerase chain reaction for regional assignment of 110 yeast artificial chromosome clones from the human X chromosome: identification of clones associated with a disease locus. Proc. Natl. Acad. Sci. USA 88, 6157-6161, 1991.
38. Burmeister M., Kim S., Price E.R., De Lange T., Tantravahi U., Myers R.M. and Cox D.R. A map of the distal region of the long arm of human chromosome 21 constructed by radiation hybrid mapping and pulsed-field gel electrophoresis. Genomics 9, 19-30, 1991.
39. Nussbaum R.L, Airhart S.D. and Ledbetter D.H. A rodent-human hybrid containing

Xq24-qter translocated to a hamster chromosome expresses the Xq27 folate-sensitive fragile site. Am. J. Med. Genet. 23, 457-466, 1986.

40. Schonk D., Boerwinkel-Driessen M., van Dalen I., Oerlemans F., Smeets B., Schepens J., Hulsebos T., Cockburn D., Boyd Y., Davis M., Rettig W., Shaw D., Roses A., Ropers H. and Wieringa B. Definition of subchromosomal intervals around the myotonic dystrophy gene region at 19q. Genomics 4, 384-396, 1989.
41. Pai G.S., Sprenkle J.A., Do T.T., Mareni C.E. and Migeon B.R. Localization of loci for hypoxanthine phosphoribosyltransferase and glucose-6-phosphate dehydrogenase and biochemical evidence of nonrandom X chromosome expression from studies of a human X-autosome translocation. Proc. Natl. Acad. Sci. USA 77, 2810-2813, 1980.
42. Davies K.E., Mattei M.G., Mattei J.F., Veenema H., McGlade S., Harper K., Tommerup N., Nielsen K.B., Mikkelsen M., Beighton P., Drayna D., White R. and Pembrey M.E. Linkage studies of X-linked mental retardation: high frequency of recombination in the telomeric region of the human X chromosome. Hum. Genet. 70, 249-255, 1985.
43. Vincent A., Oberlé I., Feil R.F., Arveiler J.L. and Mandel J.L. Physical and genetic mapping of the Xq27-q28 region. Cytogenet. Cell Genet. 51, 1099, 1989.
44. Patterson M., Bell M.V., Bloomfield J., Flint T., Dorkins H., Thibodeau S.N., Schaid D., Bren G., Schwartz C.E., Wieringa B., Ropers H.H., Callen D.F., Sutherland G., Froster-Iskenius U., Vissing H. and Davies K.E. Genetic and physical mapping of a novel region close to the fragile X site on the human X chromosome. Genomics 4, 570-578, 1989.
45. Gitschier J., Drayna D., Tuddenham E.G.D., White R.L. and Lawn R.M. Genetic mapping and diagnosis of haemophilia A achieved through a BclI polymorphism in the factor VIII gene. Nature 314, 738-740, 1985.
46. Suthers G.K., Hyland V.J., Callen D.F., Baker E., Eyre H., Roberts S.H., Hors-Cayla M.C., Davies K.E., Bell M.V. and Sutherland G.R. Physical and genetic mapping of new DNA probes at Xq26-qter. Cytogenet. Cell Genet. 51, 1087, 1989.
47. Schmidt M., Certoma A., Du Sart D., Kalitsis P., Leversha M., Fowler K., Sheffield L., Jack I. and Danks D.M. Unusual X chromosomal inactivation in a mentally retarded girl with an interstitial deletion Xq27. Implications for the fragile X syndrome Hum. Genet. 84, 347-352, 1990.
48. Ledbetter S.A., Schwartz C.E., Davies K.E. and Ledbetter D.H. New somatic cell hybrids for physical mapping in distal Xq and the fragile X region. Am. J. Med. Genet. 38, 418-420, 1991.
49. Warren S.T., Zhang F., Licameli G.R. and Peters J.F. The Fragile X site in somatic cell hybrids: an approach for molecular cloning of fragile sites. Science, 237, 420-423, 1987.
50. Warren S.T., Zhang F., Sutcliffe J.S. and Peters J.F. Strategy for molecular cloning of the Fragile X site DNA. Am. J. Med. Genet. 30, 613-623, 1988.
51. Warren S.T., Knight S.J.L., Peters J.F., Stayton C.L., Consalez G.G. and Zhang F. Isolation of the human chromosomal band Xq28 within somatic cell hybrids by Fragile X site breakage. Proc. Natl. Acad. Sci. USA 87, 3856-3860, 1990.
52. Brown W.R.A. and Bird A.P. Long-range restriction site mapping of mammalian genomic DNA. Nature 322, 477-481, 1986.
53. Van Ommen G.J.B. and Verkerk J.M.H. Restriction analysis of chromosomal DNA in

a size range up to two million basepairs by pulsed field gradient electrophoresis. In human genetic diseases: A practical approach (KE Davies ed) pp. 113-133. IRL Press, Oxford, 1986.

54. Bird A.P., DNA methylation and the frequency of CpG in animal DNA. Nucleic Acids Res. 8, 1499-1504, 1980.
55. Bird A.P. CpG-rich islands and the function of DNA methylation. Nature 321, 209-213, 1986.
56. Kenwrick S. and Gitschier J. A contiguous, 3-Mb physical map of Xq28 extending from the colorblindness locus to DXS15. Am. J. Hum. Genet. 45, 873-882, 1989.
57. Bird A.P., CpG islands as gene markers in the vertebrate nucleus. Trends Genet. 3, 342-347, 1987.
58. Pfeifer G.P., Tanguay R.L., Steigerwald D.D. and Riggs A.D. In vivo footprint and methylation analysis by PCR-aided genomic sequencing: comparison of active and inactive X-chromosomal DNA at the CpG island and promoter of human PGK-1. Genes Dev. 4, 1277-1287, 1990.
59. Lindsay S. and Bird A.P. Use of restriction enzymes to detect potential gene sequences in mammalian DNA. Nature 327, 336-338, 1987.
60. Tribioli C., Tamanini F., Patrosso C., Milanesi L., Villa A., Pergolizzi R., Meastrini E., Rivella S., Bione S., Mancini M., Vezzoni P. and Toniolo D. Methylation and sequence analysis around EagI sites: identification of 28 new CpG islands in Xq24-Xq28. Nucleic Acids Res. 20, 727-733, 1992.
61. Vincent A., Heitz D., Petit C., Kretz C., Oberlé I. and Mandel J.L. Abnormal pattern detected in fragile X patients by pulsed-field gel electrophoresis. Nature 349, 624-626, 1991.
62. Bell M.V., Hirst M.C., Nakahori Y., MacKinnon R.N., Roche A., Flint T.J., Jacobs P.A., Tommerup N., Tranebjaerg L., Froster-Iskenius U., Kerr B., Turner G., Lindenbaum R.H., Winter R., Pembrey M., Thibodeau S. and Davies K.E. Physical mapping across the fragile X: hypermethylation and clinical expression of the fragile X syndrome. Cell 64, 861-866, 1991.
63. Patterson M., Schwartz C., Bell M., Sauer S., Hofker M., Trask B., van den Engh G. and Davies K.E. Physical mapping studies on the human X chromosome in the region Xq27-Xqter. Genomics 1, 297-306, 1987.
64. Dietrich A., Kioschis P., Monaco A.P., Gross B., Korn B., Williams S.V., Sheer D., Heitz D., Oberlé I., Toniolo D., Warren S.T., Lehrach H. and Poustka A. Molecular cloning and analysis of the fragile X region in man. Nucleic Acids Res. 19, 2567-2572, 1991.
65. Heitz D., Rousseau F., Devys D., Saccone S., Abderrahim H., Le Paslier D., Cohen D., Vincent A., Toniolo D., Della Valle G., Johnson S., Schlessinger D., Oberlé I. and Mandel J.L. Isolation of sequences that span the fragile X and identification of a fragile X-related CpG island. Science 251, 1236-1239, 1991.
66. Hirst M.C., Rack K., Nakahori Y., Roche A., Bell M.V., Flynn G., Christadoulou Z., MacKinnon R.N., Francis M., Littler A.J., Anand R., Poustka A.M., Lehrach H., Schlessinger D., D'Urso M., Buckle V.J. and Davies K.E. A YAC contig across the fragile X site defines the region of fragility. Nucleic Acids Res. 19, 3283-3288, 1991.

67. Burke D.T., Carle G.F. and Olson M.V. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science 236, 806-812, 1987.
68. Albertsen H.M, Abderrahim H., Cann H.M., Dausset J., Le Paslier D. and Cohen D. Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents. Proc. Natl. Acad. Sci. USA 87, 4256-4260, 1990.
69. Anand R., Riley J.H., Smith J.C. and Markham A.F. A 3.5 genome equivalent multi access YAC library: construction, characterization, screening and storage. Nucleic Acids Res. 18, 1951-1956, 1990.
70. Abidi F.E., Wada M., Little R.D. and Schlessinger D. Yeast artificial chromosomes containing human Xq24-Xq28 DNA: library construction and representation of probe sequences. Genomics 7, 363-376, 1990.
71. Kremer E.J., Yu S., Pritchard M., Nagaraja R., Heitz D., Lynch M., Baker E., Hyland V.J., Little R.D., Wada M., Toniolo D., Vincent A., Rousseau F., Schlessinger D., Sutherland G.R. and Richards R.I. Isolation of a human DNA sequence which spans the fragile X. Am. J. Hum. Genet. 49, 656-661, 1991.
72. Verkerk A.J.M.H., Eussen B.H.J., Van Hemel J.O. and Oostra B.A. Limited size of the fragile X site shown by fluorescence in situ hybridization. Am. J. Med. Genet. 43, 187-191, 1992.
73. Oberlé I., Rousseau F., Heitz D., Kretz C., Devys D., Hanauer A., Boue J., Bertheas M.F. and Mandel J.L. Instability of a 550 base pair DNA segment and abnormal methylation in Fragile X syndrome. Science, 252, 1097-1102, 1991.
74. Yu S., Pritchard M., Kremer E., Lynch M., Nancarrow J., Baker E., Holman K., Mulley J.C., Warren S.T., Schlessinger D., Sutherland G.R. and Richards R.I. Fragile X genotype characterized by an unstable region of DNA. Science 252, 1179-1181, 1991.
75. Eichler E.E., Richards S., Gibbs R.A. and Nelson D.L. Fine structure of the human *FMR-1* gene. Hum. Mol. Genet. 2, 1147-1153, 1993.
76. Ashely C.T., Sutcliffe J.S., Kunst C.B., Leiner H.A., Eichler E.E., Nelson D.L. and Warren S.T. Human and murine *FMR-1* alternative splicing and translational initiation downstream of the CGG repeat. Nature Genet. 4, 244-251, 1993.
77. Verheij C., Bakker C.E., de Graaff E., Keulemans J., Willemsen R., Verkerk A.J.M.H., Galjaard H., Reuser A.J.J., Hoogeveen A.T. and Oostra B.A. Characterization and localization of the *FMR-1* gene product associated with fragile X syndrome. Nature 363, 722-724, 1993.
78. Hansen R.S., Gartler S.M., Scott C.R., Chen S.H. and Laird C.D. Methylation analysis of CGG sites in the CpG island of the human FMR-1 gene. Hum. Mol. Genet. 1, 571-578, 1992.
79. Hornstra I.K., Nelson D.L., Warren S.T. and Yang T.P. High resolution methylation analysis of the *FMR-1* trinucleotide repeat region in fragile X syndrome. Hum. Mol. Genet. 2, 1659-1665, 1993.
80. Richards R.I., Holman K., Yu S. and Sutherland G.R. Fragile X syndrome unstable element, p(CGG)n, and other simple tandem repeat sequences are binding sites for specific nuclear proteins. Hum. Mol. Genet. 2, 1429-1435, 1993.
81. Pieretti M., Zhang F., Fu Y.H., Warren S.T., Oostra B.A., Caskey C.T. and Nelson D.L.

Absence of expression of the *FMR-1* gene in Fragile X syndrome. Cell 66, 817-822, 1991.

82. Sutcliffe J.S., Nelson D.L., Zhang F., Pieretti M., Caskey C.T., Saxe D. and Warren S.T. DNA methylation represses *FMR-1* transcription in fragile X syndrome. Hum. Mol. Genet. 1, 397-400, 1992.
83. De Vries B.B.A., Wiegers A.M., de Graaff E., Verkerk A.J.M.H., Van Hemel J.O., Halley D.J.J., Fryns J.P., Curfs L.M.G., Niermeijer M.F. and Oostra B.A. Mental status and fragile X expression in relation to *FMR-1* gene mutation. Eur. J. Hum. Genet. 1, 72-79, 1993.
84. Siomi H., Siomi M.C., Nussbaum R.L. and Dreyfuss G. The protein product of the fragile X gene, *FMR1*, has characteristics of an RNA-binding protein. Cell 74, 291-298, 1993.
85. Deelen W., Bakker C., Halley D.J.J. and Oostra B.A. Conservation of CGG region in *FMR-1* gene in mammals. Am. J. Med. Genet., in press.
86. Faust C.J., Verkerk A.J.M.H., Wilson P.J., Morris C.P., Hopwood J.J., Oostra B.A. and Herman G.E. Genetic mapping on the mouse X chromosome of human cDNA clones for the fragile X and Hunter syndromes. Genomics 12, 814-817, 1992.
87. Fu Y-H., Kuhl D.P.A., Pizutti 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., and Caskey C.T. Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. Cell 67, 1047-1058, 1991.
88. Verkerk A.J.M.H., de Graaff E., De Boulle K., Eichler E.E., Konecki D.S., Reyniers E., Manca A., Poustka A., Willems P.J., Nelson D.L. and Oostra B.A. Alternative splicing in the fragile X gene *FMR1*. Hum. Mol. Genet. 2, 399-404, 1993.
89. Devys D., Lutz Y., Rouyer N., Bellocq J.P. and Mandel J.L. 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, 1993.
90. Ashley C.T., Wilkinson K.D., Reines D. and Warren S.T. *FMR1* protein: conserved RNP family domains and selective RNA binding. Science 262, 563-566, 1993.
91. Hinds H.L., Ashley C.T., Sutcliffe J.S., Nelson D.L., Warren S.T., Housman D.E. and Schalling M. Tissue specific expression of *FMR-1* provides evidence for a functional role in fragile X syndrome. Nature Genet. 3, 36-43, 1993.
92. Abitbol M., Menini C., Delezoide A.L., Thyner T., Vekemans M. and Mallet J. Nucleus basalis magnocellularis and hippocampus are the major sites of *FMR-1* expression in the human fetal brain. Nature Genet. 4, 147-153, 1993.
93. Hanzlik A.J., Osemlak-Hanzlik M.M., Hauser M.A. and Kurnit D.M. A recombination-based assay demonstrates that the fragile X sequence is transcribed widely during development. Nature Genet. 3, 44-47, 1993.
94. Bächner D., Steinbach P., Wöhrle D., Just W., Vogel W., Hameister H., Manca A. and Poustka A. Enhanced *Fmr1* expression in testis. Nature Genet. 4, 115-116, 1993.
95. Kremer E.J., Pritchard M., Lynch M., Yu S., Holman K., Baker E., Warren S.T., Schlessinger D., Sutherland G.R. and Richards R.I. Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. Science 252, 1711-1714, 1991.
96. Rousseau F., Heitz D., Biancalana V., Blumenfeld S., Kretz C., Boué J., Tommerup N.,

van der Hagen C., DeLozier-Blanchet C., Croquette M.F., Gilgenkrantz S., Jalbert P., Voelckel M.A., Oberlé I. and Mandel J.L. Direct diagnosis by DNA analysis of the fragile X syndrome of mental retardation. New. Engl. J. Med. 325, 1673-1681, 1991.

97. Verkerk A.J.M.H., de Vries B.B.A., Niermeijer M.F., Fu Y.H., Nelson D.L., Warren S.T., Majoor-Krakauer D.F., Halley D.J.J. and Oostra B.A. Intragenic probe used for diagnostics in fragile X families. Am. J. Med. Genet. 43, 192-196, 1992.
98. Snow K., Doud L., Hagerman R., Hull C., Hirst M.C., Davies K.E. and Thibodeau S.L. Analysis of mutations at the Fragile X locus using the DNA probe Ox1.9. Am. J. Med. Genet 43, 244-254, 1992.
99. Heitz D., Devys D., Imbert G., Kretz C. and Mandel J.L. Inheritance of the fragile X syndrome: size of the fragile premutation is a major determinant of the transition to full mutation. J. Med. Genet. 29, 794-801, 1992.
100. Wöhrle D., Hennig I., Vogel W. and Steinbach P. Mitotic stability of fragile X mutations of differentiated cells indicates early post-conceptional trinucleotide repeat expansion. Nature Genet. 4, 140-142, 1993.
101. Devys D., Biancalana V., Rousseau F., Boué J., Mandel J.L. and Oberlé I. 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, 1992.
102. Wöhrle D., Kotzot D., Hirst M.C., Manca A., Korn B., Schmidt A., Barbi G., Rott H.D., Poustka A., Davies K.E. and Steinbach P. A microdeletion of less than 250 kb, including the proximal part of the *FMR-1* gene and the fragile-X site, in a male with the clinical phenotype of fragile X syndrome. Am. J. Hum. Genet. 51, 299-306, 1992.
103. Gedeon A.K., Baker E., Robinson H., Partington M.W., Gross B., Manca A., Korn B., Poustka A., Yu S., Sutherland G.R. and Mulley J.C. Fragile X syndrome without CCG amplification has an *FMR1* deletion. Nature Genet. 1, 341-344, 1992.
104. De Boulle K., Verkerk A.J.M.H., Reyniers E, Vits L., Hendrickx J., Van Roy B., Van Den Bos F., de Graaff E., Oostra B.A. and Willems P.J. A point mutation in the *FMR-1* gene associated with mental retardation and the fragile X phenotype. Nature Genet. 3, 31-35, 1993.
105. Chiurazzi P., de Graaff E., Ng J., Verkerk A.J.M.H., Wolfson S., Fisch G.S., Kozak L., Neri G. and Oostra B.A. No apparent involvement of the *FMR-1* gene in five patients with phenotypic manifestations of the fragile X syndrome. Am. J. Med. Genet., in press.
106. Voelckel M.A., Philip N., Piquet C., Pellissier M.C., Oberlé I., Birg F., Mattei M.G. and Mattei J.F. Study of a family with a fragile site of the X chromosome at Xq27-28 without mental retardation. Hum. Genet. 81, 353-357, 1989.
107. Nakahori Y., Knight S.J.L., Holland J., Schwartz C., Rochhe A., Tarleton J., Wong S., Flint T.J., Froster-Iskenius U., Bentley D., Davies K.E. and Hirst M.C. Molecular heterogeneity of the fragile X syndrome. Nucleic Acids Res. 19, 4355-4359, 1991.
108. Romain D.R. and Chapman C.J. Fragile site Xq27.3 in a family without mental retardation. Clin. Genet. 41, 33-35, 1992.
109. Dennis N.R., Curtis G., Macpherson J.N. and Jacobs P.A. Two families with Xq27.3 fragility, no detectable insert in the *FMR-1* gene, mild mental impairment and

absence of the Martin-Bell phenotype. Am. J. Med. Genet. 43, 232-236, 1992.

110. Oberlé I., Boué J., Croquette M.F., Voelckel M.A., Mattei M.G. and Mandel J.L. Three families with high expression of a fragile site at Xq27.3, lack of anomalies at the *FMR-1* CpG island, and no clear phenotypic association. Am. J. Med. Genet. 43, 224-231, 1992.
111. Sutherland G.S. and Baker E. Characterization of a new rare fragile site easily confused with the fragile X. Hum. Mol. Genet. 1, 111-113, 1992.
112. Flynn G.A., Hirst M.C., Knight S.J.L., Macpherson J.N., Barber J.C.K., Flannery A.V., Davies K.E. and Buckle V.J. Identification of the FRAXE fragile site in two families ascertained for X linked mental retardation. J. Med. Genet. 30, 97-100, 1993.
113. Knight S.J.L., Flannery A.V., Hirst M.C., Campbell L., Christodoulou Z., Phelps S.R., Pointon J., Middleton-Price H.R., Barnicoat A., Pembrey M.E., Holland J., Oostra B.A., Bobrow M. and Davies K.E. Trinucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation. Cell 74, 127-134, 1993.
114. Hirst M.C., Barnicoat A, Flynn G., Wang Q., Daker M., Buckle V.J., Davies K.E. and Bobrow M. The identification of a third fragile site, FRAXF, in Xq27-q28 distal to both FRAXA and FRAXE. Hum. Mol. Genet. 2, 197-200, 1993.

CHAPTER

4

DNA DIAGNOSTICS OF THE FRAGILE X SYNDROME PAST AND PRESENT

Natural variations in DNA sequence exist between individuals and are passed on to their offspring. Some of these variations can be detected by the use of restriction enzymes and appear as variations in fragment lengths, known as restriction fragment length polymorphisms or RFLPs (1). Probes can recognize two (RFLP) or more (Variable Number of Tandem Repeats or VNTRs) different fragment lengths or alleles (2). In addition di-, tri- and tetranucleotide repeats are present in the genome. Repeats, which are abundant in the human genome and can be polymorphic in length, can be analyzed by means of the polymerase chain reaction (3). Segregation of alleles at polymorphic loci can be studied in families and can be used to follow the segregation of a disease gene.
The longer the distance of a marker is to the disease gene, the more chance there will be that a recombination occurs between the disease locus and this marker. Polymorphic markers which are genetically closely linked to a disease locus can be used for carrier testing and prenatal diagnosis of a disease when mutations in the (disease) gene are not known. For a reliable diagnosis the markers have to be informative in the families to be investigated. For X-linked diseases this means that female carriers have to be heterozygous for markers on both sides of the disease gene locus, so that the different haplotypes can be followed in their offspring. When flanking markers are available, only a small chance on a double recombination event remains, resulting in an erroneous diagnosis. The major limitation of linkage analysis with distant flanking markers is the substantial risk of a single recombination event, in which case no conclusion can be drawn.

4.1 Diagnosis in families with fragile X syndrome using polymorphic markers

For a number of years, diagnosis of the fragile X syndrome has been based on the detection of the fragile site at Xq27.3 (4,5,6). However, with this method many carrier females could not be detected, as approximately 50% of all obligate carrier females do not show cytogenetic expression of the fragile site (7). These females include all carriers of a premutation and a number of carriers of a full mutation. Normal male carriers were also missed using this method, because they never show cytogenetic expression. The availability of polymorphic markers on both sides of the fragile X locus opened up new possibilities for the diagnosis of the syndrome.

Since markers very close to the fragile X locus were not available for many years (*Chapter 3, paragraph 3.1*), clinical geneticists have been reluctant to use risk

analysis based on linked polymorphic DNA markers in prenatal diagnosis, and the method has been largely restricted to carrier detection.
In Figure 1 an example of linkage analysis is given in a fragile X family using the polymorphic markers RN1 (8,9), JH89 (10), and II-10 (11). The order of these markers is as follows: centromere - RN1 - JH89 - FRAXA - II-10 - telomere (8-11). The markers RN1 and II-10 are estimated to have a distance to the fragile X locus of approximately 5 cM (8,9,11). The distance from JH89 to the fragile X locus is estimated to be 4 cM (10).

Individual II-8 is a mentally retarded fragile X male with cytogenetic expression of the fragile site at Xq27.3 in 22% of his cells. The X chromosome carrying the fragile X mutation in the affected male II-8 has a specific haplotype determined by the alleles of the different markers used. The two other individuals with cytogenetic expression in this pedigree, II-4 and III-2, have the same haplotype as II-8.
The female carrier I-2 is informative for the markers proximal and distal to the fragile X locus, therefore it is possible to asses the carrier status of individuals in

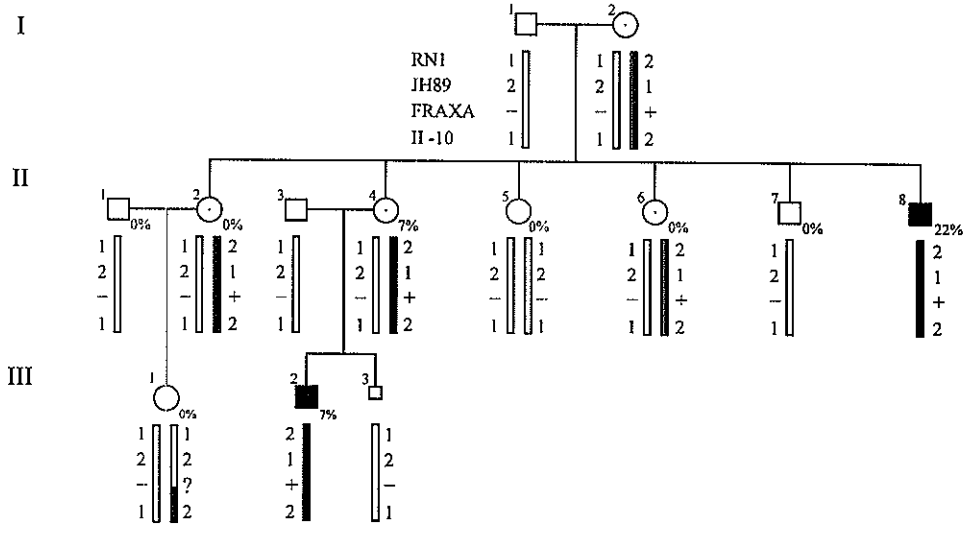


Figure 1. An example of carrier detection in a fragile X family, using closely linked markers. The markers are indicated on the left side of individual I-1. For the majority of individuals the percentage of cytogenetic expression is indicated.

FRAXA - = normal allele FRAXA + = allele with fragile X mutation

○ = normal female □ = normal male

⊙ = carrier female ■ = affected male

generation II. Both sisters II-2 and II-6 show no cytogenetic expression of the fragile site, but based on haplotype analysis both are almost certainly (> 99% risk) carriers. Sister II-5 has inherited the normal X chromosome from her mother, reducing her risk of being a carrier to less than 1%.
Sister II-4, who already has one affected son, wanted a prenatal diagnosis. She is also informative for markers on both sides of the fragile X locus. Her male fetus, III-3, has inherited the normal X chromosome from his mother and has a > 99% chance of being normal. Daughter III-1 has inherited a recombined chromosome from her mother II-2. Since it is unknown whether this single recombination has occurred proximal or distal from the disease locus, no conclusion about her carrier status could be made.

This example illustrates the possibilities and limitations of linkage analysis. In a high number of cases a clear conclusion about carrier status can be drawn. However, with the flanking markers used (with estimated distances of 4 cM proximal and 5 cM distal of the fragile X locus) a high chance (9%) of a single recombination exists. If a recombination is found, no conclusion can be drawn.

4.2 DNA diagnosis of the fragile X syndrome by determining the size of the CGG repeat in the *FMR1* gene

After the identification of the gene defect, postnatal and prenatal diagnosis of the fragile X syndrome can now be performed with a high degree of accuracy by detecting the CGG repeat amplification at the DNA level using Southern blot analysis. In addition, a PCR test can be performed to determine the exact number of CGG repeats in normal and premutation alleles. Due to the inefficiency of detection of the high number of CGG repeats in full mutations by the PCR method used (12), it is difficult to visualize these alleles by means of PCR. An example of such a PCR is given in Figure 5 in this chapter.

Southern blot analysis
DNA fragments from individuals are visualized using Southern blot analysis and hybridization with a radioactively labelled specific probe. For DNA diagnosis of the fragile X syndrome a specific probe is used to demonstrate the presence of a genomic DNA fragment containing the CGG repeat. In various laboratories different probes have been isolated, that can be used for diagnostic purposes (13-19, *Publication 7.3*), these probes are indicated in Figure 2 as a, b, c and d.

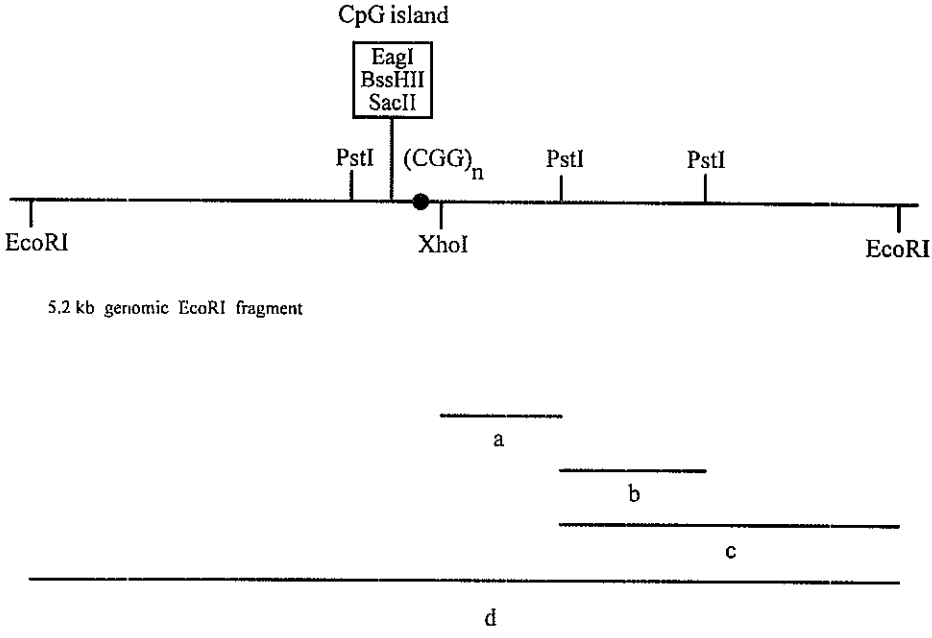


Figure 2. Representation of DNA probes used in fragile X diagnosis.
a=pfxa3=Ox0.55=pPX6
b=StB12.3=pfxa7=pP2
c=Ox1.9
d=pE5.1=pfxa1
Refs. 13-19, *Publication 7.3.*

A schematic representation of the different EcoRI fragments found in normal individuals and in fragile X carriers and patients is given in Figure 3. In normal individuals only a 5.2 kb EcoRI fragment is observed (lanes 1 and 2). In normal transmitting males premutations are found, and these are visible as somewhat larger bands up to ± 5.7 kb (lane 3). In male patients, full mutations are seen as a broad band or smear, due to the presence of different repeat lengths above 200 repeats in different cells (lane 5). In fragile X carrier females, whether they have a premutation (lane 4) or a full mutation (lane 6), always a normal 5.2 kb band is visible, representing their normal X chromosome.
Lanes 7 and 8 represent a fragile X male and female, both mosaic for a premutation and a full mutation.

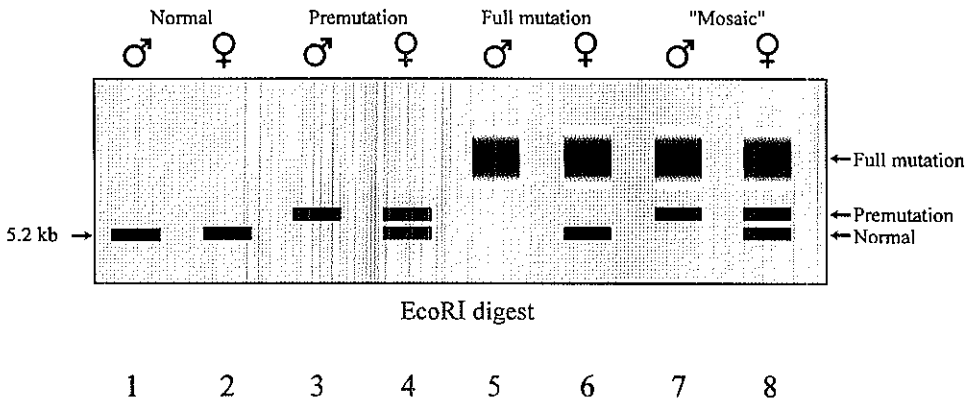


Figure 3. Schematic representation of Southern blot analysis to detect abnormal DNA fragments in fragile X syndrome. CGG repeat containing EcoRI fragments can be detected by for example probe pP2, see Figure 2 (fragment b).

1 = normal male
2 = normal female
3 = normal transmitting male
4 = normal carrier female
5 = mentally retarded male
6 = female carrying a full mutation*
7 = mentally retarded male, mosaic for a premutation and a full mutation
8 = female, mosaic for a premutation and a full mutation*

*50-75% of the women with a full mutation are mildly to moderately mentally retarded (19).

In Figure 4 an example is given of DNA analysis by means of Southern blot analysis in a fragile X family. In this example the probe pP2 (b in Figure 2) was used in combination with an EcoRI restriction digest of the DNA .
In normal individuals, the probe pP2 identifies a 5.2 kb EcoRI fragment (lanes 4, 8, 9, 11 and 12). This fragment contains the CGG repeat and the *FMR1* gene associated CpG island, which is methylated in fragile X patients (13,19; *Publication 7.3*). Altered sizes of this fragment, usually increased lengths, are detected in female carriers, affected individuals and transmitting males.
Female carriers can have either a premutation (lanes 1 and 7) or a full mutation (lane 10). In male fragile X patients a full mutation is found (lane 5).

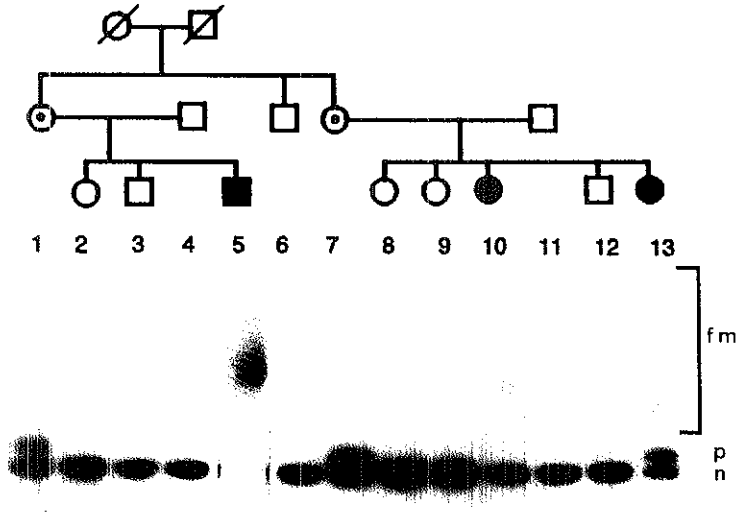


Figure 4. Southern blot analysis of a fragile X family.
○ = normal female □ = normal male
⊙ = carrier female ■ = affected male
n = normal, p = premutation, fm = full mutation.
(Southern blot analysis kindly provided by Dr. D.J.J. Halley and Dr. A.M.W. van den Ouweland.)

The premutation of the female carrier in lane 1 has increased to a full mutation in her son (lane 5). Individual 13 is a female patient who, besides a normal allele, has a mosaic pattern consisting of a premutation and a full mutation.

In *Publication 7.3* the results of DNA analysis are described for several fragile X families using the probe pE5.1 (fragment d in Figure 2) on EcoRI restriction digests. An example of the principle of gradual change of the insert size during transgenerational passage is depicted in Figure 1B, page 157.

Polymerase chain reaction amplification of the CGG repeat
By means of the polymerase chain reaction (PCR) and separation of the CGG repeat PCR products on polyacrylamide gels, it is possible to determine the exact CGG repeat numbers of normal and premutation alleles (12). Due to the high

number of CGG repeats in full mutations, it is difficult to determine repeat numbers above 200 with this method (12).

PCR analysis of the CGG repeat is useful for determining the carrier status of females. The exact number of repeats in normal and premutation alleles can be determined with this method, whereas with Southern blotting, alleles having only small size differences can not always be distinguished. PCR can be used to exclude carrier status of females by demonstrating normal alleles. It can also be used to study the stability or instability of alleles in successive generation.
In Figure 5 an example is given of a PCR analysis in a family with fragile X syndrome, including a prenatal diagnosis. The number of CGG repeats in the *FMR1* gene of the different alleles are indicated.
The PCR products of individuals II-1, III-1, III-2 and III-4 are too large to be detected using this type of PCR analysis. Full mutations in these individuals were confirmed by Southern blot analysis by the presence of a smear. Individual II-2 is a normal transmitting male who has inherited the premutation allele (increased in length) from his mother. Individual III-5 is a male fetus who inherited the normal maternal allele (27 repeats). Prenatal diagnosis by PCR was confirmed by Southern blot analysis by demonstrating the presence of an unchanged normal maternal allele in this male fetus.

Pergolizzi et al. described a different method to visualize full mutations, which combines PCR with Southern blot analysis (20,21). After PCR amplification of the CGG repeats, the PCR products are separated on an agarose gel, followed by Southern blotting. The resulting filters are then hybridized with a radioactively labelled CGG repeat oligonucleotide probe. Due to the presence of a large number of copies of CGG repeats in full mutation alleles, this gives an intense signal. However, this method has some disadvantages. Because this method is limited in the separation of the DNA, it is not possible to determine the exact number of normal and premutation alleles in this way. In addition, the amplification of large alleles will be less efficient in the presence of a normal allele (21), for instance in a female carrying a full mutation or in a male who is mosaic for a premutation and full mutation allele. In these cases a large allele might be missed and a reliable diagnosis requires an additional Southern blot analysis. At the same time Southern blot analysis can provide information about deletions in the repeat containing region or about the methylation status of the CpG island of the *FMR-1* gene, which is not obtained by PCR.
Guidelines for the diagnosis of fragile X syndrome have been formulated,

indicating the best strategy to diagnose repeat amplification in the *FMR1* gene (18).

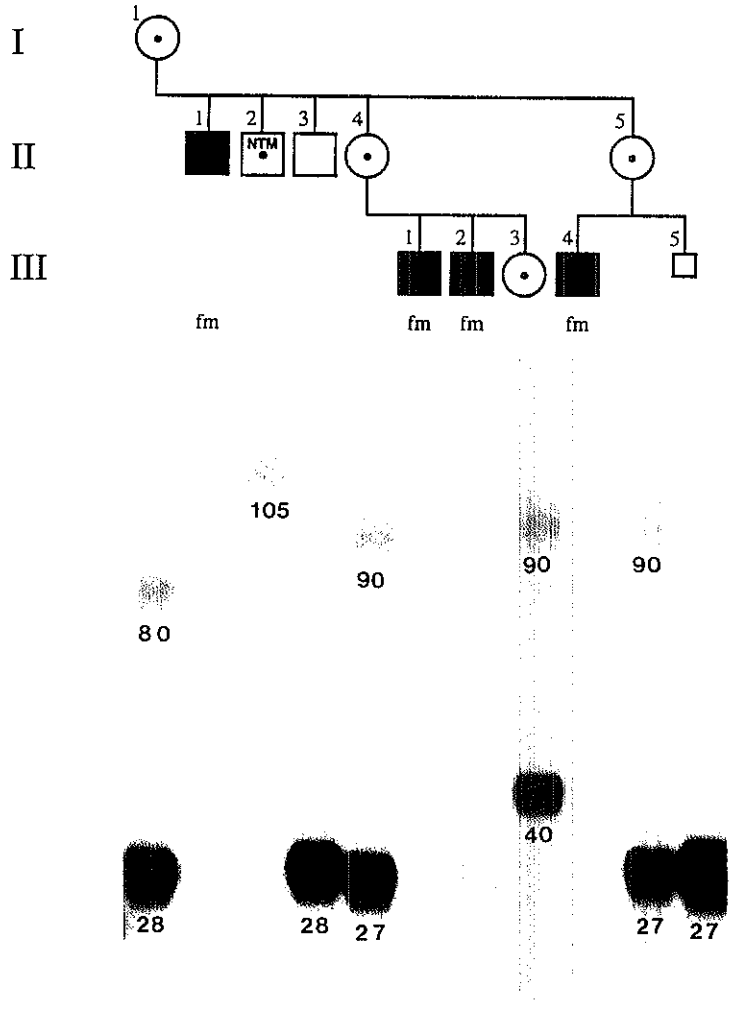


Figure 5. PCR analysis of the CGG repeat in the *FMR1* gene in a pedigree with fragile X syndrome. The number of CGG repeats is indicated. fm = full mutation.

⊙ = female carrier □ = normal male

NTM = normal transmitting male ■ = affected male

(PCR analysis kindly provided by Dr. D.J.J. Halley and Dr. A.M.W. van den Ouweland.)

Conclusions

The availability of analysis of the *FMR1* gene at the molecular level has changed the way in which fragile X diagnosis is performed during the last years.
Cytogenetic testing of the fragile site has been replaced by DNA analysis. DNA testing is more reliable in detection of normal transmitting males and female carriers of a premutation, who never show a fragile site, and of female carriers with a full mutation, in whom a fragile site can not always be detected. Southern blot analysis remains the standard method for detection of full mutations. PCR based methods can be used for characterizing the sizes of premutation and normal alleles (18).

DNA of new cases of mentally retarded individuals, whom are suspected of having the fragile X syndrome, will be tested for the presence of an amplified CGG repeat in the *FMR1* gene by using DNA analysis. Cytogenetic testing of chromosomes of these individuals for chromosomal abnormalities is still important, because such abnormalities can also be the cause of mental retardation. In unrelated mentally retarded individuals, chromosomal abnormalities are seen roughly at the same frequency (~4%) as fragile X syndrome (22).

In the future it might be possible to offer women, who consider having children, the possibility of determining their carrier status for the fragile X syndrome. The prevalence of fragile X premutations was studied in X chromosomes of more than 10,000 women (23). It was found that 1/400 women in the population is carrier of a repeat in the premutation range of 60-200 repeats (23). Taking also into account the women with small premutations, containing between 40 and 60 repeats, the frequency of female carriers is higher and might even be 1/200.
To determine repeat numbers in these cases, PCR is likely to be the method of choice. However, an overlap exists between repeat numbers in normal and premutation alleles. Alleles containing a CGG repeat number below 40 are considered to be normal. The smallest unstable allele found contained 43 CCG repeats (Nelson, pers. comm.). Alleles containing more than 60 repeats will virtually all be premutation alleles. It will be difficult to give a judgement about alleles in the "grey area" containing 40-60 CGG repeats. To determine whether these are normal or premutation alleles, it will have to be examined if such an allele is stably inherited in the family of the individual concerned.

The examples in this chapter illustrate that cloning and characterization of the *FMR1* gene and the identification of the gene defect in fragile X patients has

greatly improved the way by which the syndrome can be diagnosed. Increases in size of the CGG repeat and the number of repeats in the *FMR1* gene of individuals, can be determined by means of a simple Southern blot analysis and PCR test, and can be correlated to the clinical phenotype. On the other hand, additional questions, especially concerning expansion-proneness of "grey area" repeats, remain to be assessed in a molecular epidemiologic approach.

REFERENCES

1. Kan Y.W. and Dozy A.M. Polymorphisms of DNA sequence adjacent to human ß-globin structural gene: relation to sickle cell mutation. Proc. Natl. Acad. Sci. USA 75, 5631-5635, 1978.
2. Nakamura Y., Leppert M., O'Connell P., Wolff R., Holm T., Culver M., Martin C., Fujimoto E., Hoff M., Kumlin E. and White R. Variable number of tandem repeat (VNTR) markers for human gene mapping. Science 235, 1616-1622, 1987.
3. Weber J.L. and May P.E. Abundant class of human DNA polymerase chain reaction. Am. J. Hum. Genet. 44, 388-396, 1989.
4. Lubs H.A. A marker X chromosome. Am. J. Hum. Genet. 34, 552-557, 1969.
5. Sutherland G.R. Marker X chromosomes and mental retardation. New Eng. J. Med. 296, 1415, 1977.
6. Opitz J.M. and Sutherland G.R. Conference report: International workshop on the fragile X and X linked mental retardation. Am. J. Med. Genet. 17, 5-94, 1984.
7. Nielsen K.B., Tommerup N., Poulsen H., Jacobsen P., Beck B. and Mikkelsen M. Carrier detection and X-inactivation studies in the fragile X syndrome. Hum. Genet. 64, 240-245, 1983.
8. Oostra B.A., Hupkes P.E., Perdon L.F., van Bennekom C.A., Bakker E., Halley D.J.J., Schmidt M., du Sart D., Smits A., Wieringa B. and van Oost B.A. New polymorphic DNA marker close to the fragile site FRAXA. Genomics 6, 129-132, 1990.
9. Wieacker P., Davies K.E., Cooke H.J., Pearson P.L., Williamson R., Bhattacharya S., Zimmer J. and Ropers H.H. Toward a complete linkage map of the human X chromosome: regional assignment of 16 cloned single-copy DNA sequences employing a panel of somatic cell hybrids. Am. J. Hum. Genet. 36, 265-276, 1984.
10. Dreesen J.C.F.M., van den Hurk J.A.J.M., Smits A.P.T., van den Ouweland A.M.W., Markslag P.W.B. and van Oost B.A. DXS539, a polymorphic DNA marker proximal of the fragile-X gene. Hum. Genet. 91, 80-82, 1993.
11. Hulsebos T.J.M., Oostra B.A., Broersen S., Smits A., van Oost B.A. and Westerveld A. New distal marker closely linked to the fragile X locus. Hum. Genet. 87, 369-372, 1991.
12. Fu Y-H., Kuhl D.P.A., Pizutti 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., and Caskey C.T. Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. Cell 67, 1047-1058, 1991.

13. Oberlé I., Rousseau F., Heitz D., Kretz C., Devys D., Hanauer A., Boué J., Bertheas M.F. and Mandel J.L. Instability of a 550 base pair DNA segment and abnormal methylation in Fragile X syndrome. Science, 252, 1097-1102, 1991.
14. Yu S., Pritchard M., Kremer E., Lynch M., Nancarrow J., Baker E., Holman K., Mulley J.C., Warren S.T., Schlessinger D., Sutherland G.R. and Richards R.I. Fragile X genotype characterized by an unstable region of DNA. Science 252, 1179-1181, 1991.
15. Kremer E.J., Pritchard M., Lynch M., Yu S., Holman K., Baker E., Warren S.T., Schlessinger D., Sutherland G.R. and Richards R.I. Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. Science 252, 1711-1714, 1991.
16. Nakahori Y., Knight S.J.L., Holland J., Schwartz C., Roche A., Tarleton J., Wong S., Flint T.J., Froster-Iskenius U., Bentley D., Davies K.E. and Hirst M.C. Molecular heterogeneity of the fragile X syndrome. Nucleic Acids Res. 19, 4355-4359, 1991.
17. Rousseau F., Heitz D., Biancalana V., Oberlé I. and Mandel J.L. On some technical aspects of direct DNA diagnosis of the Fragile X syndrome. Am. J. Med. Genet. 43, 197-207, 1992.
18. Oostra B.A., Jacky P.B., Brown W.T. and Rousseau F. Guidelines for the diagnosis of fragile X syndrome. J. Med. Genet. 30, 410-413, 1993.
19. Rousseau F., Heitz D., Biancalana V., Blumenfeld S., Kretz C., Boué J., Tommerup N., van der Hagen C., DeLozier-Blanchet C., Croquette M.F., Gilgenkrantz S., Jalbert P., Voelckel M.A., Oberlé I. and Mandel J.L. Direct diagnosis by DNA analysis of the fragile X syndrome of mental retardation. New. Engl. J. Med. 325, 1673-1681, 1991.
20. Pergolizzi R.G., Erster S.H., Goonewaardena P. and Brown W.T. Detection of full fragile X mutation. Lancet 339, 271-272, 1992.
21. Erster S.H., Brown W.T., Goonewardena P., Dobkin C.S., Jenkins E.C. and Pergolizzi R.G. Polymerase chain reaction analysis of fragile X mutations. Hum. Genet. 90, 55-61, 1992.
22. Van den Ouweland A.M.W., de Vries B.B.A., Bakker P.L.G., Deelen W.H., de Graaff E., van Hemel J.O., Oostra B.A., Niermeijer M.F. and Halley D.J.J. 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., in press.
23. Rousseau F., Réhel R., Rouillard P., DeGrandpré P., Morgan K. and Khandjian E.W. Mutational prevalence of fragile X premutations in 10,624 females from the general population by Southern blotting. Am. J. Hum. Genet. 53, Abstract 3, 1993.

CHAPTER 5

TRINUCLEOTIDE REPEATS

CAGCAGCAGCAGCAGCAGCAGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCAGCAGCAGCAGCAGCAGCAG

5.1 Disease genes containing trinucleotide repeats

The discovery in 1991 that amplification of a trinucleotide repeat is the molecular genetic basis of the fragile X syndrome revealed a new mechanism of transgenerational inheritance.
This finding was soon followed by discoveries of other diseases in which triplet amplifications at the DNA level are responsible for the disease phenotype: Kennedy's disease or spinal and bulbar muscular atrophy (SBMA) (1991) (1), myotonic dystrophy (DM=dystrophia myotonica) (1992) (2,3,4), Huntington's disease (HD) (1993) (5), spinocerebellar ataxia type 1 (SCAI) (1993) (6), and the mental retardation syndrome associated with the FRAXE site (7; *Chapter 3, paragraph 3.6*).
The genes associated with these disorders all contain a polymorphic GC-rich trinucleotide repeat which is unstable and is expanded in affected individuals. Only two kinds of trinucleotide repeat sequences are found in these genes, CGG/CCG and CAG/CTG repeats. For example, in the fragile X syndrome and the mental retardation syndrome associated with the FRAXE site, the composition of the repeat DNA is identical, though opposite sequences are coding.
In the case of fragile X syndrome, DM, SBMA, HD and SCAI, the repeat is present within the mature transcript. In the *FMR1* gene and the *DM* gene the repeat is present in the untranslated region. The location of the repeat in FRAXE is not yet known.

Kennedy's disease
Kennedy's disease or spinal and bulbar muscular atrophy (SBMA) is an X-linked recessive disease that has been described in about 50 families. Affected males show an adult onset of progressive muscle weakness, atrophy and may have gynaecomastia and reduced fertility, the latter symptoms suggesting a defect in androgen receptor function.
PCR amplification and sequencing of the eight exons of the androgen receptor on chromosome Xq11-12 showed variation in the first exon of the gene between patients and published control sequences (1). A polymorphic CAG repeat is present in this exon of the androgen receptor which falls within the coding region of the AR gene, and in patients the CAG repeat is increased in length. The repeat is highly polymorphic in length both in normal individuals (12-30 repeats) (1,8,9,10) and in SBMA patients (40-62 repeats) (1,8). Transmission of the expanded CAG allele from one generation to the next is associated with instability, both expansions and contractions have been observed (8,11). No somatic mosaicism has

been found in SBMA, like in fragile X syndrome and myotonic dystrophy.

Myotonic dystrophy

Myotonic dystrophy or DM is an autosomal dominant neuromuscular disease with an estimated incidence of 1 in 8000 (12). Patients exhibit progressive muscle weakness and wasting, myotonia and cataracts. Age of onset and severity of the disease show extreme variation both within and between families. Patients can be divided into 3 groups: 1) minimally affected or late onset, 2) the classical form with onset in early adult life and adolescence (both termed non congenital DM), and 3) the most severe congenital form of myotonic dystrophy also abbreviated as CMD. CMD is only seen in the offspring of mothers who themselves have DM and is often associated with neonatal death. The congenital form of DM is exclusively maternally transmitted.

The DM gene is located to chromosome 19q13.2-13.3 (13), and an unstable DNA region was identified in this region (2,3,4). Isolation and identification of the DM gene showed the presence of a CTG repeat in the 3' untranslated region of the gene (14,15,16). The CTG repeat is polymorphic in the normal population (5-30 repeats) and is increased in length in DM patients (50 - >2000 repeats) (14,15,16). A protein was predicted with high similarity to numerous protein kinases (14,16,17).

Huntingon's disease

Huntington's disease or HD is a progressive neurodegenerative disorder. It is inherited in an autosomal dominant fashion and affects about 1 in 10.000 individuals (18). Onset of the disease generally starts in the fourth to fifth decade of life (19) and gradually worsens over a course of 10-20 years until death occurs (20). Symptoms start with characteristic uncontrolled movements and include progressive mental disturbances; there is loss of specific neurons in different parts of the brain.

A polymorphic CAG repeat is found in the 5' region of the Huntington's gene on chromosome 4p16.3, which is most likely translated into the protein (5). In Huntington's patients the CAG repeat is increased in length and is unstable, as in the DNA of affected children of an affected individual different numbers of repeats are found. The function of the HD protein is unknown; no homology to any other known protein is found (5).

Spinocerebellar ataxia type 1

Spinocerebellar ataxia type 1 or SCAI is an autosomal dominant progressive neurodegenerative disease (6). The SCAI gene is located on chromosome 6p22-23.

Table 1. Data concerning the trinucleotide repeat containing genes involved in fragile X syndrome, Kennedy's disease, myotonic dystrophy, Huntington's disease and spinocerebellar ataxia type 1.

Gene	FMR1	AR	DM	HD	SCA1
Disease involved	Fragile X syndrome	Kennedy's disease	Myotonic dystrophy	Huntington's disease	Spino-cerebellar ataxia type 1
Chromosomal location	Xq27.3	Xq11-12	19q13.3	4p16.3	6p22-23
Repeat	CGG	CAG	CTG	CAG	CAG
Location in the gene	5'UTR	5'coding	3'UTR	5'coding	coding
Number of repeats:					
in normal individuals	6-53	12-30	5-30	10-34	19-36
in patients	200->2000	40-62	50- >2000	30-128	43-81
Gene characteristics:					
genomic size	~40 kb	>90 kb	~14 kb	~210 kb	nk
mRNA transcript	~4 kb	10,6 kb	~3,4 kb	~10-11 kb	10 kb
open reading frame	1,9 kb*	2,7 kb	1,9 kb*	9,4 kb	nk
protein	631 aa*	910 aa	629 aa*	3144 aa	nk

AR = Androgen receptor
DM = Dystrophia myotonica
SCA1 = Spinocerebellar ataxia type 1
nk = not known
* = due to alternative splicing several ORF's and differently sized proteins are possible. The longest possible protein is indicated.

Refs.: 1-6,8,9,11,14-17,21-28 and references mentioned in *Chapters 3, 6 and 8.*

In patients, an extended CAG repeat has been identified in this region (6,28). The SCAI gene itself has not been isolated yet.

In Table 1 an overview is given of genetic diseases based on trinucleotide amplification within the genes involved.

5.2 Amplification and anticipation

Four of the above mentioned disorders, Kennedy's disease, myotonic dystrophy, Huntington's disease and spinocerebellar ataxia type 1, show the phenomenon of anticipation.
Anticipation is the term given to the apparent occurrence of increasing severity of symptoms of an inherited disorder (through successive generations) with progressively earlier age of onset of the disease in successive generations. In more severely affected individuals, higher copy numbers of repeats are found. With increasing repeat lengths, the age of onset of the diseases decreases, a phenomenon clearly visible in myotonic dystrophy and Huntington's disease. The anticipation can be explained by the molecular findings: the number of repeats increases with the transmission of the mutation from generation to generation (6,8,25-29).

The "anticipation" as seen in fragile X syndrome is an all or nothing phenomenon and does not fit the definition. In carriers of a premutation, normal transcription and translation take place, resulting in a normal phenotype. When the amplification passes the border of about 200 copies of the CGG repeat, methylation of the promoter region of the *FMR1* gene and the repeat itself occurs. This prevents transcription, and the resulting protein deficiency leads to the disease phenotype in all males. In these males, the severity of the disease is not dependent on the number of repeats found in their *FMR1* gene (30). In DM, HD and SCAI, the anticipation is evident, whereas in SBMA it is present but not as evident as in the other three diseases. It might be speculated that genes involving other disorders that exhibit anticipation are also candidates to contain an unstable trinucleotide repeat.

Sex specificity of transmission
In fragile X syndrome, affected children are exclusively born, after passage of the mutation through a female. Offspring of normal transmitting males and male

patients are always normal. Sperm of affected males was found to contain only mutations in the premutation range (31), which remain a premutation in their daughters. Why only maternal premutations can expand to full mutations is not yet understood (*see Chapter 6*).
Congenital DM is inherited exclusively from mothers (32,33). In CDM, very large expansions of the CTG repeat (1000-2000 repeats) are found. It could be speculated that these large repeat sizes could not be transmitted by males. No imprinting is involved in the expression of the *DM* gene (34,35). In contrast, the juvenile onset forms of HD and SCAI are inherited from fathers (25-28,36). In these cases a larger number of repeats are found.
The basis of the sex specific transmissions of especially large mutations is unknown.

Size and location of trinucleotide repeats within disease genes
Amplifications on a small scale are found when the repeat is located within the coding region of the gene and thus translated into protein, as is the case in the *HD*, *SCAI*, and androgen receptor genes. Amplifications of a much larger scale are found when the repeat is located within the untranslated region, as in the *FMR1* (5' UTR) and *DM* (3' UTR) genes. A possible explanation for these differences might be that large amplifications in the coding region of the genes involved may be lethal in early embryogenesis.

Repeats and methylation
There are 10 possible trinucleotide repeat motifs, of which only two can be methylated: CGG/CCG and ACG/CGT. In Fragile X syndrome expansion above 200 repeats is accompanied by methylation of the *FMR1* promoter region, resulting in transcriptional silencing of the gene and absence of a protein product. In the mental retardation syndrome associated with the FRAXE fragile site, methylation of the associated CpG island occurs as well, suggesting the same kind of mechanism as in fragile X syndrome (7). It will be of interest to determine whether methylation would play a role in any other disease associated with a CGG/CCG or ACG/CGT unstable repeat.
The increase of the CTG repeat in the myotonic dystrophy gene has not been associated with methylation (35). CAG/CTG repeats have been found in four of the above mentioned diseases (Table 1, page 93).

Gain of function
In Kennedy's disease, Huntington's disease and spinocerebellar ataxia type 1, the

CAG repeat is located within the open reading frame of the gene. This results in a glutamine stretch in the respective proteins. It is speculated that the protein with an above normal increased repeat gains a new, erroneous function. Proteins containing glutamine stretches can be substrates for the enzyme transglutaminase. For instance the protein involucrin contains a number of glutamine stretches. This protein serves as a substrate for the enzyme keratinocyte transglutaminase in the eye, that cross-links involucrin to other proteins to form the insoluble envelope of the corneocyte (37). As a result of an increased stretch of glutamines in the proteins encoded by the *SBMA*, *HD* and *SCAI* genes, these proteins might now serve as substrates for transglutaminases and are cross linked to other proteins, forming aggregates (38). The brain contains transglutaminases as well, also bound to synaptosomal membranes (39,40). Aggregates formed in the brain might be difficult to degrade, in contrast to more efficient degradation in other cell types, and might be lethal to neurons (37).
In the late-onset disorders SBMA, HD, and SCA1, death of neuronal cells is found, but there is variability in the affected region in the brain.

5.3 The search for trinucleotide repeats in other human genes

The finding of trinucleotide repeats in disease genes such as fragile X syndrome, Kennedy's disease, myotonic dystrophy and Huntington's disease has initiated the search for other genes containing this kind of repeats. This may lead to the discovery of other expanding genes causing (neurological) diseases.
A search in GenBank revealed numerous trinucleotide repeat containing cDNAs, of which CGG, CAG and CCT were most frequently found (41,42). This indicates that trinucleotide repeats in genes are quite common in the human genome.
Screening of cDNA libraries with CGG and CAG repeat oligonucleotides has identified a number of independent cDNAs containing these sequences, most of which were isolated from a fetal brain cDNA library (41,42,43). In about one third of the investigated genes the repeat was not polymorphic, one third revealed a two allele polymorphism, and one third revealed length variation with greater then 2 alleles (42).
The triplet repeats that were found in the 5' untranslated region of the different genes were more polymorphic than those where the repeat appeared to be translated (42). This might be anticipated since there is probably a higher selective pressure against variation within coding regions.
Genes involved in other disorders that exhibit anticipation might also be a

candidate to contain a trinucleotide repeat.

REFERENCES

1. La Spada A.R., Wilson E.M., Lubahn D.B., Harding A.E. and Fischbeck K.H. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. Nature 352, 77-79, 1991.
2. Harley H.G., Brook J.D., Rundle S.A., Crow S., Reardon W., Buckler A.J., Harper P.S., Housman D.E. and Shaw D.J. Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. Nature 355, 545-546, 1992.
3. Buxton J., Shelbourne P., Davies J., Jones C., Van Tongeren T., Aslanidis C., de Jong P., Jansen G., Anvret M., Riley B., Williamson R. and Johnson K. Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. Nature 355, 547-548, 1992.
4. Aslanidis C., Jansen G., Amemiya C., Shutler G., Mahadevan M., Tsilfidis C., Chen C., Alleman J., Wormskamp N.G.M., Vooijs M., Buxton J., Johnson K., Smeets H.J.M., Lennon G.G., Carrano A.V., Korneluk R.G., Wieringa B. and de Jong P. Cloning of the essential myotonic dystrophy region and mapping of the putative defect. Nature 355, 548-551, 1992.
5. The Huntington's disease collaborative research group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell 72, 971-983, 1993.
6. Orr H.T., Chung M., Banfi S., Kwiatkowski T.J., Servadio A., Beaudet A., McCall A.E., Duvick L.A., Ranum L.P.W. and Zoghbi H.Y. Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. Nature Genet. 4, 221-226, 1993.
7. Knight S.J.L., Flannery A.V., Hirst M.C., Campbell L., Christodoulou Z., Phelps S.R., Pointon J., Middleton-Price H.R., Barnicoat A., Pembrey M.E., Holland J., Oostra B.A., Bobrow M. and Davies K.E. Trinucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation. Cell 74, 127-134, 1993.
8. La Spada A.R., Roling D.B., Harding A.E., Warner C.L., Spiegel R., Hausmanowa-Petrusewicz I., Yee W-C. and Fischbeck K.H. Meiotic stability and genotype-phenotype correlation of the trinucleotide repeat in X-linked spinal and bulbar muscular atrophy. Nature Genet. 2, 301-304, 1992.
9. Edwards A., Civitello A., Hammond H.A. and Caskey C.T. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am. J. Hum. Genet. 49, 746-756, 1991.
10. Edwards A., Hammond H.A., Jin L., Caskey C.T. and Chakraborty R. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics 12, 241-25, 1992.
11. Biancalana V., Serville F., Pommier J., Julien J., Hanauer A. and Mandel J.L. Moderate instability of the trinucleotide repeat in spino bulbar muscular atrophy. Hum. Mol. Genet. 1, 255-258, 1992.

12. Harper P.S. Myotonic dystrophy 2nd edn (Saunders, London and Philadelphia) 1989.
13. Brunner H.G., Kornelk R.G., Coerwinkel-Driessen M., MacKenzie A., Smeets H., Lambermon H.M.M., van Oost B.A., Wieringa B. and Ropers H.H. Myotonic dystrophy is closely linked to the gene for muscle-type creatine kinase (CKMM). Hum. Genet. 81, 308-310, 1989.
14. Brook J.D., McCurrach M.E., Harley H.G., Buckler A.J., Church D., Aburatani H., Hunter K., Stanton V.P., Thirion J-P., Hudson T., Sohn R., Zemelman B., Snell R.G., Rundle S.A., Crow S., Davies J., Shelbourne P., Buxton J., Jones C., Juvonen V., Johnson K., Harper P.S., Shaw D.J. and Housman D.E. Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. Cell 68, 799-808, 1992.
15. Mahadevan M., Tsilfidis C., Saborin L., Shutler G., Amemiya C., Jansen G., Neville C., Narang M., Barcelo J., O'Hoy K., Leblond S., Earle-Macdonald J., de Jong P.J., Wieringa B. and Korneluk R.G. Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. Science 255, 1253-1255, 1992.
16. Fu Y-H., Pizzuti A., Fenwick R.G. Jr., King J., Rajnarayan S., Dunne P.W., Dubel J., Nasser G.A., Ashizawa T., de Jong P., Wieringa B., Korneluk R., Perryman M.B., Epstein H.F. and Caskey C.T. An unstable triplet repeat in a gene related to myotonic muscular dystrophy. Science 255, 1256-1258, 1992.
17. Jansen G., Mahadevan M., Amemiya C., Wormskamp N., Segers B., Hendriks W., O'Hoy K., Baird S., Sabouring L., Lennon G., Jap P.L., Iles D., Coerwinkel M., Hofker M., Carrano A.V., de Jong P.J., Korneluk R.G. and Wieringa B. Characterization of the myotonic dystrophy region predicts multiple protein isoform-encoding mRNAs. Nature Genet. 1, 261-266, 1992.
18. Harper P.S., Morris M.J., Quarrell O., Shaw D.J., Tyler A. and Youngman S. Huntington disease. (Philadelphia) WB Saunders Company Ltd, London, 1991.
19. Di Maio L., Squitieri F., Napolitano G., Campanella G., Trofatter J.A. and Conneally P.M. Onset symptoms in 510 patients with Huntington's disease. J. Med. Genet. 30, 289-292, 1993.
20. Conneally P.M. Huntington disease: genetics and epidemiology. Am. J. Hum. Genet. 36, 506-526, 1984.
21. Eichler E.E., Richards S., Gibbs R.A. and Nelson D.L. Fine structure of the human FMR1 gene. Hum. Mol. Genet. 2, 1147-1153, 1993.
22. Ashley C.T., Sutcliffe J.S., Kunst C.B., Leiner H.A., Eichler E.E., Nelson D.L. and Warren S.T. Human and murine FMR1: alternative splicing and translational initiation downstream of the CGG repeat. Nature Genet. 4, 244-251, 1993.
23. Faber P.W., van Rooij H.C.J., van der Korput H.A.G.m., Baarends W.M., Brinkmann A.O., Grootegoed J.A. and Trapman J. Characterization of the human androgen receptor transcription unit. J. Biol. Chem. 266, 10743-10749,1991.
24. Fu, Y.H., Pizzuti A., Fenwick R.G., King J., Rajnarayan S., Dunne P.W., Dubel J., Nasser G.A., Ashizawa T., de Jong P., Wieringa B., Korneluk R., Perryman M.B., Epstein H.F. and Caskey C.T. An unstable triplet repeat in a gene related to myotonic muscular dystrophy. Science 255, 1256-1258, 1992.
25. Duyao M., Ambrose C., Myers R., Novelletto A., Persichetti F., Frontali M., Folstein S.,

Ross C., Franz M., Abbott M., Gray J., Conneally P., Young A., Penney J., Hollingsworth Z., Shoulson I., Lazzarini A., Falek A., Koroshetz W., Sax D., Bird E., Vonsattel J., Bonilla E., Alvir J., Conde J.B., Cha J.H., Dure L., Gomez F., Ramos M., Sanchez-Ramos J., Snodgrass S., de Young M., Wexler N., Moscowitz C., Penchaszadeh G., MacFarlane H., Anderson M., Jenkins B., Srinidhi J., Barnes G., Gusella J. and MacDonald M. Trinucleotide repeat length instability and age of onset in Huntington's disease. Nature Genet. 4, 387-392, 1993.

26. Snell R.G., MacMillan J.C., Cheadle J.P., Fenton I., Lazarou L.P., Davies P., MacDonald M.E., Gusella J.F., Harper P.S. and Shaw J. Relationship between trinucleotide repeat expansion and phenotypic variation in Huntington's disease. Nature Genet. 4, 393-397, 1993.
27. Chung M.Y., Ranum L.P.W., Duvick L.A., Servadio A., Zoghbi H.Y. and Orr H.T. Evidence for a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type I. Nature Genet. 5, 254-258, 1993.
28. Andrew S.E., Goldberg Y.P., Kremer B., Telenius H., Theilmann J., Adam S., Starr E., Squitieri F., Lin B., Kalchman M.A., Graham R.K. and Hayden M.R. The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. Nature Genet. 4, 398-403, 1993.
29. Igarashi S., Tanno Y., Onodera O., Yamazaki M., Sato S., Ishikawa A., Miyatani N., Nagashima M., Ishikawa Y., Sahashi K., Ibi T., Miyatake T. and Tsuji S. Strong correlation between the number of CAG repeats in androgen receptor genes and the clinical onset of features of spinal and bulbar muscular atrophy. Neurology 42, 2300-2302, 1992.
30. De Vries B.B.A., Wiegers A.M., de Graaff E., Verkerk A.J.M.H., Van Hemel J.O., Halley D.J.J., Fryns J.P., Curfs L.M.G., Niermeijer M.F. and Oostra B.A. Mental status and fragile X expression in relation to *FMR-1* gene mutation. Eur. J. Hum. Genet. 1, 72-79, 1993.
31. Reyniers E., Vits L., De Boulle K., Van Roy B., Van Velzen D., de Graaff E., Verkerk A.J.M.H., Jorens H.Z.J., Darby J.K., Oostra B. and Willems P.J. The full mutation in the *FMR1* gene of male fragile X patients is absent in their sperm. Nature Genet. 4, 143-146, 1993.
32. Hageman A.T.M., Gabreëls F.J.M., Liem K.D., Renkawek K. and Boon J.M. Congenital myotonic dystrophy; a report in thirteen cases and a review of the literature. J. Neurol. Sci. 115, 95-101, 1993.
33. Mulley J.C., Staples A., Donnelly A., Gedeon A.K., Hecht B.K., Nicholson G.A., Haan E.A. and Sutherland G.R. Explanation for exclusive maternal origin for congenital form of myotonic dystrophy. Lancet 341, 236-237, 1993.
34. Shaw D.J., Chaudhary S., Rundle S.A., Crow S., Brook J.D., Harper P.S. and Harley H.G. A study of DNA methylation in myotonic dystrophy. J. Med. Genet. 30, 189-192, 1993.
35. Jansen G., Bartolomei M., Kalscheuer. Merkx G., Wormskamp N., Mariman E., Smeets D., Ropers H.H. and Wieringa B. No imprinting involved in the expression of DM-kinase mRNAs in mouse and human tissues. Hum. Mol. Genet. 2, 1221-1227, 1993.
36. Chung M.Y., Ranum L.P.W., Duvick L.A., Servadio A., Zoghbi H.Y. and Orr H.T.

Evidence for a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type 1. Nature Genet. 5, 254-258, 1993.

37. Rice R.H. and Green H. Presence in human epidermal cells of a soluble protein precursor of the cross-linked envelope: activation of the cross-linking by calcium ions. Cell 18, 681-694, 1979.
38. Green H. Human genetic diseases due to codon reiteration: relationship to an evolutionary mechanism. Cell 74, 955-956, 1993.
39. Greenberg C.S., Birckbichler P.J. and Rice R.H. Transglutaminases: multifunctional cross-linking enzymes that stabilize tissues. Faseb J. 5, 3071-3077, 1991.
40. Facchiano F., Benfenati F., Valtorta F. and Luini A. Covalent modification of Synapsin I by a Tetanus toxin-activated transglutaminase. J. Biol. Chem. 268, 4588-4591, 1993.
41. Caskey C.T., Pizzuti A., Fu Y-H., Fenwick R.G. Jr. and Nelson D.L. Triplet repeat mutations in human disease. Science 256, 784-789, 1992.
42. Riggins G.J., Lokey L.K., Chastin J.L., Leiner H.A., Sherman S.L., Wilkinson K.D. and Warren S.T. Human genes containing polymorphic trinucleotide repeats. Nature Genet. 2, 186-191, 1993.
43. Li S.H., McInnis M.G., Margolis R.L., Antonarakis S.E. and Ross C.A. Novel triplet repeat containing genes in human brain: cloning, expression, and length polymorphisms. Genomics 16, 572-579, 1993.

CHAPTER 6

DISCUSSION

The elucidation of the DNA sequence of the *FMR1* gene not only answered many questions, but has raised many new ones as well. The amplification of a CGG repeat in fragile X patients was an unexpected finding and represents a new and complicated genetic mutational mechanism that is not yet understood. The discovery that trinucleotide repeats can change in size during passage from one generation to the next and in addition can cause diseases is a fundamentally new concept in genetics.
There are still unanswered questions: how and when normal alleles expand to premutations and premutations to full mutations; how full mutations induce methylation, silence *FMR1* transcription and induce cytogenetic fragile X expression.
In this chapter possible mechanisms of repeat expansion and the moment that repeat expansion might occur will be discussed.

6.1 Founder effect

Fragile X mutations are lost constantly from the population, as affected males in general do not reproduce and because mentally retarded fragile X females also tend not to reproduce (1). According to Haldane's theory, this loss of mutations should be compensated by a gain of new mutations (2). Likewise, it was expected that the high population frequency of the fragile X syndrome was due to a high rate of new mutations (1,3,4). The isolation and identification of the *FMR1* gene has revealed an amplification of the CGG repeat in the 5′ end of the gene in the majority of fragile X patients (*Chapter 3*). If mutations would occur random on either X chromosome, then these mutations would not be correlated with a specific haplotype (2). However, to date no new mutations involving the CGG repeat length have been detected in any of the fragile X families investigated. Detailed molecular analysis revealed that all mothers of fragile X males are fragile X carriers (5-7), which means that no fragile X males are the result of a new mutation in the oocyte of the mother. Genealogical analysis of fragile X families showed that the fragile X mutation could be traced back to common ancestors up to 6 generations ago (7-9). It is suggested that initial mutations can be transmitted by numerous generations prior to increasing to a high-risk premutation (10,11). In addition, haplotype analysis with polymorphic markers flanking the CGG repeat revealed distinct differences in haplotype frequency between normal and fragile X chromosomes (10-13). A limited number of haplotypes are found in the majority of all fragile X chromosomes examined, suggesting that a small number of founder

chromosomes account for the present day fragile X mutations (7,10,12-14). In the normal population one of the haplotypes frequently found on fragile X chromosomes, is associated with a much higher than average number of CGG repeats (in the normal range), suggesting that this higher CGG copy number is more likely to expand (to a premutation).
If this model is correct, then a large pool of mutant (premutation) alleles should be present in the population. Preliminary data suggests that 1/400 females do have an allele in the premutation range (60-200 repeats) (15). They are potential carriers of an allele that might be at risk in generations to come.

6.2 Timing of repeat expansion

There is an overlap in the number of CGG repeats present in normal and premutation alleles (16; *Chapter 3, paragraph 3.4*). A normal allele (6-54 repeats) can be defined by the fact that it is stably transmitted from one generation to the next. An allele is called a premutation (43-200 repeats) and thus unstable, when it increases or decreases in length during passage to the next generation.
Determination of the number of CGG repeats in the *FMR1* gene in normal individuals revealed a 54 repeat allele which was clearly unstable during transmission through several generations in a clinically normal family (14). This allele likely represents a fragile X premutation that has yet to undergo expansion to a full mutation causing the fragile X phenotype. In a fragile X family, a female was identified with a premutation of 43 repeats that was unstable (Nelson, pers. comm.).

The risk of expansion from a premutation in the parent to a full mutation in the offspring is dependent on two factors: the size of the premutation (16,17) and the parental origin. A fundamental question is when and how the expansion takes place. Is the expansion already present in the germ cells of the parent, or does it occur at an early embryological stage? Some evidence has now been gathered that the expansion most likely occurs in the early embryo.

Transmission by a female
When a premutation is transmitted by a female, the chance of a full mutation in the offspring is correlated with the size of the premutation in the mother (16,17). The larger the premutation is, the higher the risk of expansion in the next generation will be. Premutations with an allele size above 90 repeats have an

almost 100% risk of becoming a full mutation in the offspring (16; *Chapter 3*). Females with a full mutation never have children with a premutation. They do sometimes have children, who are mosaic for a premutation and a full mutation.

Transmission by a male

When a premutation is transmitted by a normal transmitting male, expansion to the full mutation never occurs, irrespective of the size of the repeat. Increases however do occur, but they stay in the premutation range in the offspring. Mentally retarded male fragile X patients usually do not reproduce because of their mental handicap, but the few male patients who have reproduced, have phenotypically normal daughters (18-22). One such a case has been studied extensively at the molecular level (21). In this case the father was mosaic for a full mutation and a premutation. His daughter carried a premutation. A premutation in the daughter could be explained by the fact that the germline of the father is mosaic as well, and that he passed the premutation on to his daughter. Another possibility would be that a full mutation has been transmitted by the father, which has decreased to a premutation in the daughter. A third possibility is that germ cells of fragile X patients never carry a full mutation but only a premutation. To address this question, the *FMR1* mutation was examined in sperm cells of several male fragile X patients (23). The sperm cells of these males only contained a premutation, despite the fact that full mutations were present in somatic cells of each of the males.

The size of the mutation in oocytes is not known. Moreover, it is unknown whether in mature oocytes of fragile X females, like in sperm cells of fragile X males, only premutations are present. The expansion of a premutation to a full mutation might occur after fertilization in the early embryo.

The fact that expansion to a full mutation only occurs after female transmission is not yet explained, but could be due to a parental imprinting mechanism in the *FMR1* region: a distinction is made between paternally and maternally derived premutations, and for some unknown reason only maternally derived alleles can expand to a full mutation in the offspring. This expansion is then a somatic event.

Another possibility to explain expansion to a full mutation only after female transmission is that sperm cells only contain premutations and that oocytes contain either a premutation or a full mutation.

Models

In this paragraph two possible models are discussed that could explain the absence of the full mutation in male gametes:

1) gametes with a premutation have a selective advantage over gametes with a full mutation,
2) the germline only contains cells with a premutation.

Model 1.

Both premutations and full mutations could be present in spermatogonia and/or primordial germ cells. Through selection of spermatogonia for *FMRP* production only cells containing a premutation may develop to mature sperm cells. Bächner et al. described a high expression of *FMR1* mRNA in proliferating stages of spermatogonia but not at later stages of spermatogenesis in normal 4 and 6 week old mice (24). The presence of *FMR1* protein in human spermatogonia, but not in more mature germ cells was confirmed by Devys et al. using antibodies directed against *FMR1* (25). This could imply that *FMR1* expression is necessary for germ cell proliferation, and that only spermatogonia with a premutation allele, which allows for mRNA synthesis, may enter spermatogenesis.

Two arguments can be used against this model of selective pressure for *FMR1* expression in spermatogonia. First, transgenic male mice have been generated with a "knocked out" *Fmr1* gene. These mice have a knocked out *Fmr1* gene also in their germ cells. Breeding experiments have demonstrated, that these mice can produce normal numbers of offspring (Bakker, pers. comm.). This indicates that *FMR1* expression is not essential for spermatogenesis.
Secondly, in a fragile X family a deletion was found of the promoter region of the *FMR1* gene in four sisters and four of their sons (Meijer, pers. comm.). In this family the deletion is likely to have been transmitted by the grandfather. All his daughters inherited the deletion, and haplotype analysis showed that the four affected grandsons inherited the same allele from their grandfather. This deletion results in lack of transcription of the *FMR1* gene (de Graaff, pers. comm.). This again suggests that *FMR1* is not a prerequisite for spermatogenesis.

Model 2.

An alternative model predicts that all the cells in the germ line contain a premutation (23). It is clear that male fragile X patients with somatically the full mutation, only have the premutation in their sperm cells (23), and according to this model, only a premutation would be present in the mature oocytes of fragile X carriers. Model 2 involves that the expansion to a full mutation is post-conceptional and occurs exclusively in somatic cells, after separation of the germ-line cells (which keep the premutation), Figure 1. An argument in favour of

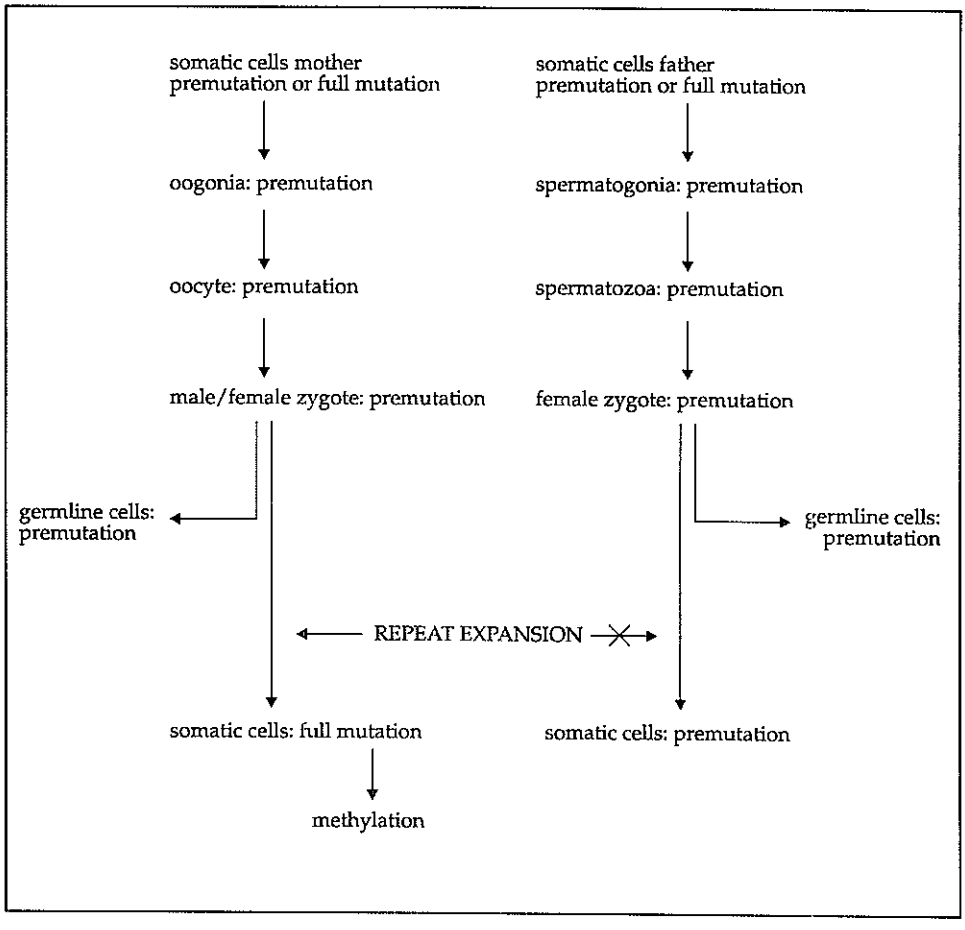


Figure 1. Model for the expansion of the CGG repeat in the *FMR1* gene in the early embryo.

this model is that the obvious somatic variation of CGG repeat length which is observed in cells of patients with a full mutation also has to be a result of ongoing mitotic instability of the expanded repeat, as only a single *FMR1* allele is passed on via an oocyte. Wöhrle et al. demonstrated that the pattern of repeat lengths in somatic cells is already established in the 13th week of fetal development (26). All tissues revealed near identical patterns of repeat lengths.

According to this model, also premutation alleles have to be found in oocytes of a

female with a somatic full mutation. It has to be assumed that the size of this premutation is in the high-risk range as females with a full mutation never gave birth to a child with a premutation. Females with a full mutation do sometimes have children who are mosaic for a premutation and a full mutation. Resolution of this model awaits the direct analysis of oocytes in fragile X carrier females.

6.3 Possible mechanisms of expansion

Premutation alleles often change in size with transmission to individuals in the next generation. In most cases the allele increases in size, but reduction in repeat size is also observed in fragile X syndrome (16,20,27). Premutations can be passed on through many generations before a full mutation and subsequently a fragile X phenotype occurs in a pedigree (7,8). The mechanism through which the repeat expansion takes place is still unknown. Studies on micro- and minisatellites have demonstrated high mutation frequencies in these sequences (28,29). Genes have been localized that control the stability of repeat sequences, which are mutated in high frequency in tumour cells (30,31). Several different models have been proposed to explain the instability of trinucleotide repeat sequences. Most of these models could explain why expansions take place, but at the same time these models introduce the occurrence of regressions, which would have to take place as frequently as expansions. Expansions in repeat length, however, are found much more often than regressions.

Model 1. Amplification through genetic recombination in meiosis
Expansion of the repeat is assumed to occur during meiosis. Hot spots of meiotic recombination co-localize with simple sequence repeats (32). The minimal length for an efficient recombination is about 150-200 base pairs (33). This length corresponds to the repeat size above which instability is observed (16). This model cannot explain why in most cases an increase in the repeat size is observed. Although this model could account for small changes in repeat length, recombinations at the site of the repeat have not been detected in fragile X carriers. If recombination would occur, it would have been strictly limited to the CGG repeat. Recombination also cannot explain the expansion in male cells, where only one X chromosome is present.

Model 2. Repeat amplification through recombination during mitosis
In this model expansion is assumed in mitosis. Before germ cells enter meiosis,

they go through a number of mitotic divisions. In mitosis homologous chromosomes do not pair, but an exchange between sister chromatids can occur. Unequal sister chromatid exchange would eventually result in gametes carrying alleles with different numbers of repeats. This model, in contrast to model 1, at least explains that premutation alleles from normal transmitting males can change in size, because sister chromatid exchange can occur in a single X chromosome. This model would also explain expansion during mitosis in somatic cells as seen for instance in male fragile X patients who have different repeat lengths in different cells. But according to this model, increases should occur as often as decreases, and large increases cannot be explained.

Model 3. Amplification through errors during replication

A third possibility is that mistakes occur during replication of C/G rich areas. This might happen during mitotic divisions in primordial germ cells which eventually produce sperm and oocytes, or during mitosis in the early embryo. The rate of DNA synthesis of C/G rich areas might be slow as compared to other regions of the chromosome. This might lead to incomplete single strands. Out of register pairing between the repeats on the original and the newly produced strand, and reinitiation of replication might then lead to longer alleles, a process called "slipped mispairing". A single slippage could not account for the massive amplification observed from a premutation size to a full mutation size. To explain these large expansions, repeated or multiple slippage events would have to occur. This can either happen in one replication event or in subsequent replication events. DNA can have alternative forms that differ distinctly from the classical Watson-Crick double helix (34). The DNA structure could be involved in the occurrence of mutations. It is known that Z-DNA regions can be formed in C/G or A/T rich areas, and regions of Z-DNA are prone to slipped mispairing and recombination (34,35). Regions of triplet repeats might form unusual and complex structures with alternative hydrogen bonding, which could hinder the formation of a normal replication fork. Replication could be slowed down by the presence of such a complex structure, or by a protein tightly bound to the repeat region. This replication block is indicated as a shaded box in Figure 2. The polymerase would be unable to proceed, and slippage might occur in either the leading or the lagging strand, resulting in the generation of extra repeats (34). That Xq27.3 region is a late replicating area, has been shown by several groups (36,37). In male fragile X patients, who have large repeats in their *FMR1* gene, band Xq27 is even later replicating than in normal controls.

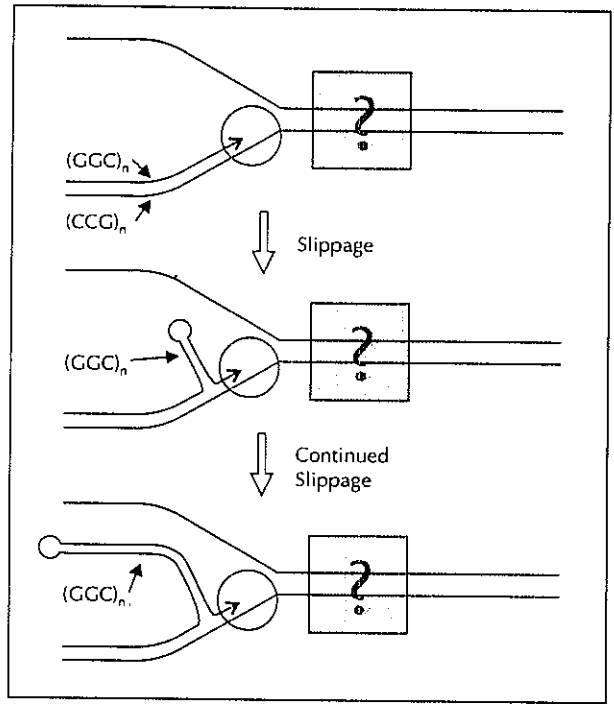


Figure 2. Hypothetical mechanism for expansion of triplet repeats. Adapted from Sinden and Wells, 1992 (34).

Formation of complex secondary structures of long repeats upon denaturation and annealing might explain why thermophilic DNA polymerase used for polymerase chain reactions cannot easily replicate these long repeats (34).

This model, however, does not solve questions such as: why are these kind of replication errors restricted to a short period in development, as in the 13th week of development the pattern of repeat lengths seems already established (26,35). One possibility could be that, once methylation has occurred, no changes in repeat length take place anymore. Another question is, if amplification is restricted to somatic cells, why does it not occur in germ cells.

Model 4. Gap repair

Loop formation of the DNA might occur, when the DNA repeat has reached a critical length (38). To remove the loop in one strand, for example the loop present in the "upper" strand in Figure 3, the opposite strand could be cleaved with an endonuclease. As a consequence, the upper strand could stretch and the lower strand can be filled in. In this way strand extension might occur.

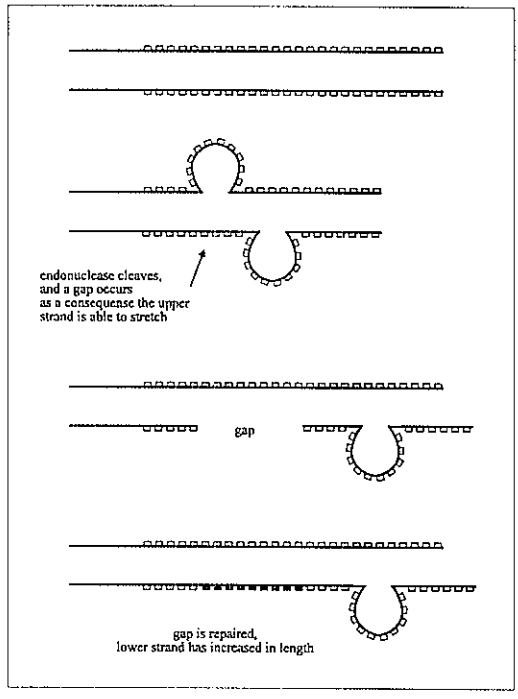


Figure 3. Hypothetical model for repeat expansion due to gap repair.
represents a repeat unit.
Adapted from Jessberger and Berg, 1991 (38).

6.4 Relation between methylation and repeat size

Methylation of the *FMR1* CpG island has been examined in 10-11 week old tissues and chorionic villi of male fetuses with a full mutation (39,40). In the different fetal tissues, almost complete methylation of the *FMR1* associated CpG island was found, and no or a very low level of *FMR1* mRNA was present in these tissues. In general, methylation is also found in the chorionic villi, but in several cases no (or only partial methylation of the CpG island was found, with *FMR1* mRNA present (5,39,40). This indicates that in absence of the methylation, transcription of the gene with a full mutation is possible (40). Another argument in favour of the occurrence of transcription of an unmethylated expanded repeat is the identification of a few normal males who have unmethylated "full mutations" with repeat numbers exceeding 200 and with normal *FMR1* protein.
Chorionic villi and fetal tissues are derived from the same embryonal tissue and

are separated late in the 2nd week of gestation. In 10-11 week old fetal tissues, complete methylation of the *FMR1* CpG island is found, whereas in villi from the same time period sometimes no or partial methylation is found (5,39,40). This means that:

1) methylation of the expanded repeat in the fetus probably occurs after the second week (after separation of the villi) and is completed around or after the eleventh week of gestation (26),
2) extension of the repeat is preceded by methylation.

It is unclear at what time of pregnancy methylation of the *FMR1* promoter in extra embryonal tissue of the chorion is completed. Therefore, the methylation status of villi should not be used in prenatal diagnostic testing.

It can be speculated that *FMR1* is variably expressed during development of an affected fetus due to variability in the timing of onset of methylation. This might explain some of the variability seen in the fragile X phenotype.

6.5 Repeats and cytogenetic expression

Fragile X syndrome and FRAXE mental retardation are both associated with a cytogenetically visible fragile site (41,42). Other diseases caused by repeat expansions are not associated with fragile sites. The repeats in fragile X syndrome (CGG) and in FRAXE mental retardation (GCC) both contain CpG sequences which can be methylated on cytosine. Methylation occurs when these repeats expand beyond 200 repeats. Therefore it is reasonable to assume that the existence of fragility in these two syndromes, is correlated with the nature of the repeat (solely composed of C and G nucleotides) and/or with the methylation of this kind of repeat. The percentage of cells that show cytogenetic expression of the fragile site is determined by the length of the repeat (43). The mechanism of the origin of the cytogenetic fragile site has yet to be elucidated.

6.6 Abnormal cyclic AMP production in platelets from fragile X patients

Fragile X patients have problems with short-term memory and processing of new information as well as with habituation to sensory stimuli (44). This prompted Berry-Kravis et al. to study cyclic adenosine monophosphate (cAMP) metabolism in platelets from fragile X patients (45,46), because the cAMP cascade is a second

messenger system that is thought to be required for short-term neuronal retention of information (47). Interruption of this cascade appears to produce a deficiency in the registration of perceptions (cognitive deficiency) (47). Berry-Kravis et al. demonstrated that fragile X patients have a diminished production of cAMP in platelets as compared to control groups. It has to be kept in mind though that the situation in blood platelets may not be representative for the situation in the brain. The *FMR1* gene does not appear to code for a protein involved in cyclic AMP metabolism but could possibly regulate other genes which might have an influence on the cAMP cascade.

6.7 Concluding remarks

The isolation and characterization of the *FMR1* gene has revealed a trinucleotide repeat in the 5′ region of the *FMR1* gene. This repeat is increased in length in the *FMR1* gene of fragile X carriers and patients. The repeat length in individuals can be determined using Southern blotting and PCR, and it can be used in DNA diagnostics. However, many questions still are unanswered. The mechanism of repeat expansion, the timing of repeat expansion and the actual function of the *FMR1* protein are still unclear. In the mouse *Fmr1* gene, which has a high homology to the human *FMR1* gene (48), a polymorphic CGG repeat is also present. Using the mouse as a model, the study of *Fmr1* in this animal might give an indication of the function of *FMR1* in humans. Studying transgenic mice in which the *Fmr1* gene has been knocked out or in which an enlarged CGG repeat in the *Fmr1* gene is introduced, might give some answers related to function, and timing and mechanism of repeat expansion.
The study of transgenic animal models may also provide more insight in the pathogenesis of mental retardation and other congenital abnormalities. Once the pathogenesis is understood, it might become possible to test therapeutic strategies on these animals.

REFERENCES

1. Sherman S.L., Jacobs P.A., Morton N.E., Froster-Iskenius U., Howard-Peebles P.N., Nielsen K.B., Partington M.W., Sutherland G.R., Turner G. and Watson M. Further segregation analysis of the fragile X syndrome with special reference to transmitting males. Hum. Genet. 69, 289-299, 1985.

2. Haldane J.B.S. The rate of spontaneous mutation of a human gene. J. Genet. 31, 317-326, 1935.
3. Brown T.W. The fragile X: progress toward solving the puzzle. Am. J. Hum. Genet. 47, 175-180, 1990.
4. Nussbaum R.L. and Ledbetter D.M. The fragile X syndrome. In: The metabolic basis of inherited disease, 6th edition. New York. N.Y. MacCraw-Hill. pp. 327-341, 1989.
5. Rousseau F., Vincent A., Rivella S., Heitz D., Tribioli C., Maestrini E., Warren S.T., Suthers G.K., Goodfellow P., Mandel J.L., Toniolo D and Oberlé I. Direct diagnosis by DNA analysis of the fragile X syndrome of mental retardation. New Engl. J. Med. 325, 1673-1681, 1991.
6. Yu S., Mulley J., Loesch D., Turner G., Donnelly A., Gedeon A., Hillen D., Kremer E., Lynch M., Pritchard M., Sutherland G.R. and Richards R.I. Fragile X genotype characterized by an unstable region of DNA. Am. J. Hum. Genet. 50, 968-980, 1992.
7. Smits A., Smeets D., Hamel B. and van Oost B. High prevalence of the Fra(X) syndrome cannot be explained by a high mutation rate. Am. J. Med. Genet. 43, 345-352, 1992.
8. Smits A.P.T., Dreesen J.C.F.M., Post J.G., Smeets D.F.C.M., de Die-Smulders C., Spaans-van der Bijl T., Govaerts L.C.P., Warren S.T., Oostra B.A. and van Oost B.A. The fragile X syndrome: no evidence for any recent mutation. J. Med. Genet. 30, 94-96, 1993.
9. Drugge U., Holmgren G., Blomquist H.K., Dahl N., Gustavson K.H. and Malmgren H. Study of individuals possibly affected with the fragile X syndrome in a large Swedish family in the 18th to 20th centuries. Am. J. Med. Genet. 43, 353-354, 1992.
10. Chakravarti A. Fragile X founder effect? Nature Genet. 1, 237-238, 1992.
11. Morton N.E. and Macpherson J.N. Population genetics of the fragile X syndrome: multiallelic model for the *FMR-1* locus. Proc. Natl. Acad. Sci. USA. 89, 4215-4217, 1992.
12. Richards R.I., Holman K., Friend K., Kremer E., Hillen D., Staples A., Brown W.T., Goonewardena P., Tarleton J., Schwartz C. and Sutherland G.R. Evidence of founder chromosomes in fragile X syndrome. Nature Genet. 1, 257-260, 1992.
13. Oudet C., Mornet E., Serre J.L., Thomas F., Lentes-Zengerling S., Kretz C., Deluchat C., Tejada I., Boué J., Boué A. and Mandel J.L. 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, 1993.
14. Buyle S., Reyniers E., Vits L., de Boulle K., Handig I., Wuyts F.L.E., Deelen W., Halley D.J.J., Oostra B.A. and Willems P.J. Founder effect in a Belgian-Dutch fragile X population. Hum. Genet. 92, 269-272, 1993.
15. Rousseau F., Réhel R., Rouillard P., DeGrandpré P., Morgan K. and Khandjian E.W. Mutational prevalence of fragile X premutations in 10.624 females from the general population by southern blotting. Am. J. Hum. Genet. 53, abstract 3, 1993.
16. 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. and Caskey C.T. Variation of the CGG repeat at the fragile X site results in genetic

instability: resolution of the Sherman paradox. Cell 67, 1047-1058, 1991.
17. Heitz D., Devys D., Imbert G., Kretz C. and Mandel J.L. Inheritance of the fragile X syndrome: size of the fragile premutation is a major determinant of the transition to the full mutation. J. Med. Genet. 29, 794-801, 1992.
18. Van Roy B.C., de Smedt M.C., Raes R.A., Dumon J.E. and Leroy J.G. Fragile X trait in a large kindred: transmission also through normal males. J. Med. Genet. 20, 286-289, 1983.
19. Moric-Petrovic S. and Laca Z. A father and daughter with fragile X chromosome. J. Med. Genet. 20, 476-478, 1983.
20. Mulley J.C., Yu S., Gedeon A,K., Donnelly A., Turner G., Loesch D., Chapman C.J., Gardner R.J.M., Richards R.I. and Sutherland G.R. Experience with direct molecular diagnosis of fragile X. J. Med. Genet. 29, 368-374, 1992.
21. Willems P.J., Van Roy B., De Boulle K., Vits L., Reyniers E., Beck O., Dumon J.E., Verkerk A. and Oostra B. Segregation of the fragile X mutation from an affected male to his normal daughter. Hum. Mol. Genet. 1, 511-515, 1992.
22. Hori T., Yamauchi M., Seki N., Tsuji S. and Ikuko K. Heritable unstable DNA sequences and hypermethylation associated with fragile X syndrome in Japanese families. Clin. Genet. 43, 34-38, 1993.
23. Reyniers E., Vits L., De Boulle K., Van Roy B., Van Velzen D., de Graaff E., Verkerk A.J.M.H., Jorens H.Z.J., Darby J.K., Oostra B. and Willems P.J. The full mutation in the *FMR-1* gene of male fragile X patients is absent in their sperm. Nature Genet. 4, 143-146, 1993.
24. Bächner D., Steinbach P., Wöhrle D., Just W., Vogel W., Hameister H., Manca A. and Poustka A. Enhanced *FMR-1* expression in testis. Nature Genet. 4, 115-116, 1993.
25. Devys D., Lutz Y., Rouyer N., Bellocq J.P. and Mandel J.L. 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, 1993.
26. Wöhrle D., Hennig I., Volgel W. and Steinbach P. Mitotic stability of fragile X mutations in differentiated cells indicates early post-conceptional trinucleotide repeat expansion. Nature Genet. 4, 140-142, 1993.
27. Oberlé I., Rousseau F., Heitz D., Kretz C., Devys D., Hanauer A., Boué J., Bertheas M.F. and Mandel J.F. Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. Science 252, 1097-1102, 1991.
28. Gibbs M., Collick A., Kelly R.G. and Jeffreys A.J. A tetranucleotide repeat mouse minisatellite displaying substantial somatic instability during early preimplantation development. Genomics 17, 121-128, 1993
29. Weber J.L. and Wong C. Mutation of human short tandem repeats. Hum. Mol. Genet. 2, 1123-1128, 1993.
30. Aaltonen L.A., Peltomäki P., Leach F.S., Sistonen P., Pylkkänen L., Mecklin J.P., Härvinen H., Powell S.M., Jen J., Hamilton S.R., Petersen G.M., Kinzler K.W., Vogelstein B. and de la Chapelle. Clues to the pathogenesis of familial colorectal cancer. Science 260, 812-815, 1993.
31. Thibodeau S.N. and Schaid G.B. Microsatellite instability in cancer of the proximal colon. Science 260, 816-819, 1993.

32. Steinmetz M., Stephan D. and Lindahl K. Gene organization and recombinational hotspots in the murine major histocompatibility complex. Cell 44, 895-904,1986.
33. Rubnitz J. and Subramani S. Rapid assay for extrachromosomal homologous recombination in monkey cells. Mol. Cell Biol. 4, 2253-2258, 1984.
34. Sinden R.R. and Wells R.D. DNA structure, mutations, and human genetic disease. Current Opinion in Biotechn. 3, 612-622, 1992.
35. Nelson D.L. and Warren S.T. Trinucleotide repeat instability: when and where? Nature Genet. 4, 107-108, 1993.
36. Webb T. Delayed replication of Xq27 in individuals with the fragile X syndrome. Am. J. Med. Genet. 43, 1057-1062, 1992.
37. Hansen R.S., Canfield T.K., Lamb M.M., Gartler S.M. and Laird C.D. Association of fragile X syndrome with delayed replication of the *FMR-1* gene. Cell 73, 1403-1409, 1993.
38. Jessberger R. and Berg P. Repair of deletions and double-strand gaps by homologous recombination in a mammalian in vitro system. Mol. Cell. Biol. 11, 445-457, 1991.
39. Sutherland G.R., Gedeon A., Kornman L., Donnelly A., Byard R.W., Mulley J.C., Kremer E., Lynch M., Pritchard M., Yu S. and Richards R.I. Prenatal diagnosis of fragile X syndrome by direct detection of the unstable DNA sequence. New Engl. J. Med. 325, 1720-1722, 1991.
40. Sutcliffe J.S., Nelson D.L., Zhang F., Pieretti M., Caskey C.T., Saxe D. and Warren S.T. DNA methylation represses *FMR-1* transcription in fragile X syndrome. Hum. Mol. Genet. 6, 397-400, 1992.
41. Lubs H.A. A marker X chromosome. Am. J. Hum. Genet. 34, 552-557, 1969.
42. Knight S.J.L., Flannery A.V., Hirst M.C., Campbell L., Christodoulou Z., Phelps S.R., Pointon J., Middleton-Price H.R., Barnicoat A., Pembrey M.E., Holland J., Oostra B.A., Bobrow M. and Davies K.E. Trinucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation. Cell 74, 127-134, 1993.
43. De Vries B.B.A., Wiegers A.M., de Graaff E., Verkerk A.J.M.H., Van Hemel J.O., Halley D.J.J., Fryns J.P., Curfs L.M.G., Niermeijer M.F. and Oostra B.A. Mental status and fragile X expression in relation to *FMR-1* gene mutation. Eur. J. Hum. Genet. 1, 72-79, 1993.
44. Freund L.S. and Reiss A.L. Cognitive profiles associated with the fra(X) syndrome in males and females. Am. J. Med. Genet. 38, 542-547, 1991.
45. Berry-Kravis E. and Huttenlocher P.R. Cyclic AMP metabolism in fragile X syndrome. Ann. Neurol. 31, 22-26, 1992.
46. Berry-Kravis E. and Sklena P. Demonstration of abnormal cyclic AMP production in platelets from patients with fragile X syndrome. Am. J. Med. Genet. 45, 81-87, 1993.
47. Kandel E.R. and Schwartz J.H. Molecular biology of learning: modulation of transmitter release. Science 218, 433-443, 1982.
48. Ashley C.T., Sutcliffe J.S., Kunst C.B., Leiner H.A., Eichler E.E., Nelson D.L. and Warren S.T. Human and murine *FMR-1*: alternative splicing and translational initiation downstream of the CGG-repeat. Nature Genet. 4, 244-251, 1993.

CHAPTER 7

PUBLICATIONS OF THE EXPERIMENTAL WORK

PUBLICATION 7.1

IDENTIFICATION OF A GENE (*FMR-1*) CONTAINING A CGG REPEAT COINCIDENT WITH A BREAKPOINT CLUSTER REGION EXHIBITING LENGTH VARIATION IN FRAGILE X SYNDROME

Annemieke J.M.H. Verkerk*, Maura Pieretti¶,
James S. Sutcliffe‡, Ying-Hui Fu¶, Derek P. A. Kuhl¶,
Antonio Pizzuti¶, Orly Reiner¶, Stephen Richards¶, Maureen F. Victoria¶,
Fuping Zhang‡, Bert E. Eussen*, Gert-Jan B. v. Ommen§, Lau A.J. Blonden§,
Gregory J. Riggins‡, Jane L. Chastain‡, Catherine B. Kunst‡, Hans Galjaard*,
C.Thomas Caskey¶, David L. Nelson¶, Ben A. Oostra#, and Stephen T. Warren‡.

* Department of Clinical Genetics, Erasmus University Rotterdam, The Netherlands.

¶ Institute for Molecular Genetics, Human Genome Center and Howard Hughes Medical Institute, Baylor College of Medicine Houston, Texas 77030.

‡ Departments of Biochemistry and Pediatrics, Emory University School of Medicine Atlanta, Georgia 30322.

§ Department of Human Genetics, Sylvius Laboratory Leiden, The Netherlands.

\# Department of Cell Biology, Erasmus University Rotterdam, The Netherlands.

Cell 65, 905-914, 1991.

SUMMARY

Fragile X syndrome is the most frequent form of inherited mental retardation and is associated with a fragile site at Xq27.3. We identified human YAC clones that span fragile X site-induced translocation breakpoints coincident with the fragile X site. A gene (*FMR-1*) was identified within a four cosmid contig of YAC DNA that expresses a 4.8 kb message in human brain. Within a 7.4 kb EcoRI genomic fragment, containing *FMR-1* exonic sequences distal to a CpG island previously shown to be hypermethylated in fragile X patients, is a fragile X site-induced breakpoint cluster region that exhibits length variation in fragile X chromosomes. This fragment contains a lengthy CGG repeat that is 250 bp distal of the CpG island and maps within a *FMR-1* exon. Localization of the brain expressed *FMR-1* gene to this EcoRI fragment suggests the involvement of this gene in the phenotypic expression of the fragile X syndrome.

INTRODUCTION

Fragile X syndrome is the most frequently encountered form of inherited mental retardation in humans, with a prevalence estimated to be 1/1250 males (Gustavson et al., 1986; Webb et al., 1986). Fragile X syndrome segregates as an X-linked dominant disorder with reduced penetrance, since either sex, when carrying the fragile X mutation, may exhibit mental deficiency. Sherman et al. (1984; 1985) have shown that approximately 30% of carrier females are penetrant and that 20% of males carrying the fragile X chromosome are phenotypically normal but may transmit the disorder and have fully penetrant grandsons. In addition to the mental retardation, which is variable in severity, penetrant males exhibit additional phenotypic involvement including macroorchidism and distinctive facies (Nussbaum and Ledbetter, 1990). Since fully penetrant males rarely reproduce, it has been suggested that the frequency of new fragile X mutations may be as high as 1/3000 germ cells to maintain the population frequency (Brown, 1990).

Fragile X syndrome, as implied by the name, is associated with a fragile site, expressed as an isochromatid gap in the metaphase chromosome, at map position Xq27.3 (Krawczun et al., 1985). The fragile X site is induced by cell culture conditions which perturb deoxypyrimidine pools and is rarely observed in greater than 50% of the metaphase spreads (Sutherland and Hecht, 1985). Neither the molecular nature of the fragile X site nor its relationship to the gene(s) responsible for the clinical expression of the syndrome is understood. However, based upon

genetic linkage studies as well as *in situ* hybridizations, the fragile X site and its associated gene(s) are tightly linked, if not coincident (Oostra et al., 1990; Rousseau et al., 1991; Hirst et al., 1991).

To elucidate the fragile X site at the molecular level, somatic cell hybrids were isolated that contained translocations between rodent chromosomes and the human fragile X chromosome, retaining either human Xpter→q27.3 or human Xq27.3→qter, referred to as proximal or distal translocations relative to the fragile X site (Warren et al., 1987, 1988, 1990). Since the high frequency and specificity of the chromosome breakage was not observed in normal X hybrids and since the translocation breakpoints map within the same interval defined by polymorphic loci which flank the fragile X locus (Warren et al., 1990; Rousseau et al, 1991; Hirst et al., 1991), these breakpoints are likely to coincide with the fragile X site.

A yeast artificial chromosome (YAC) has been isolated which spans some of these translocation breakpoints and includes polymorphic loci which flank the fragile X locus (Heitz et al., 1991). Within this region, a fragile X-related CpG island was identified which is aberrantly hypermethylated in patients and most carriers of the fragile X syndrome (Vincent et al., 1991; Bell et al., 1991). Although the significance of this CpG-island hypermethylation remains unclear, these data do imply the presence of a gene, perhaps inactivated by methylation, within a genomic region which includes the fragile X-associated hybrid breakpoints. We report below the isolation and characterization of a brain-expressed gene, designated *FMR-1*, that is associated with this CpG-island; we show that the majority of the hybrid translocation breakpoints cluster within a 7.4 kb fragment, containing both the CpG island and sequences at the 5′ end of *FMR-1*, which exhibits length variation in patients, suggesting the association of the fragile X site with *FMR-1*.

RESULTS

Isolation of YACs Spanning the Fragile X Translocation Breakpoints

Through regional mapping of YAC clones containing DNA inserts derived from the distal human Xq (Nelson et al., 1991), an 80 kb YAC (RS46) was found to map within Xq27.3, proximal to the fragile X-associated hybrid breakpoints. A 4.0 kb subclone (p46-1.1) of RS46 identified a normal 600 kb SalI fragment on pulsed field gel electrophoresis (PFGE), that was altered in size in 6 of 8 proximal translocation hybrids (Figure 1). Previous PFGE analyses of these hybrids, with more distant X-linked probes, had shown identical band sizes and therefore similar methylation patterns as might be expected since the hybrids were all derived from the same

parental fragile X somatic cell hybrid (Y75-1B-M1). These data therefore suggest that in 75% of the proximal translocation hybrids, the human breakpoint is within the 600 kb SalI fragment observed in the parental, intact fragile X hybrid. In the translocation hybrids, the distal human SalI site is lost and replaced by heterologous translocations containing different rodent SalI sites.

Since YAC RS46 does not hybridize to the DNA of the distal translocation hybrids and therefore does not cross these translocation breakpoints, additional YACs were sought of this region. A YAC library developed at the Human Polymorphism Study Center (CEPH; Albertson et al., 1990) was screened using RS46 specific oligonucleotide primers 1625 and 1626 (Nelson et al., 1991). A YAC of 475 kb (209G4) was identified which completely overlaps YAC RS46 and includes sequences distal to the proximal translocation breakpoints which are present in 13 of 14 distal translocation breakpoints (data not shown). YAC 209G4 encompasses

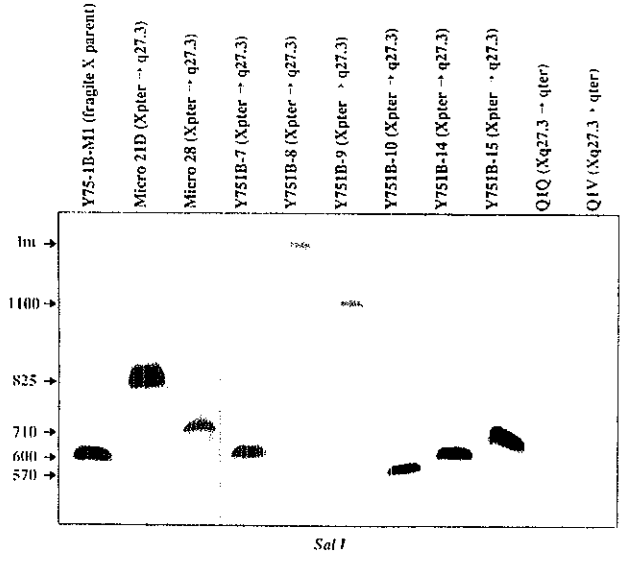


Figure 1. Southern blot analysis of pulsed field gel-resolved SalI digested DNA of proximal translocation hybrids probed with p46-1.1.

Y75-1B-M1 is a somatic cell hybrid containing the intact fragile X chromosome from which all other hybrids were derived. Lanes 2-9 are proximal translocation hybrids containing centric human Xpter→q27.3 translocated to different rodent chromosome arms. Q1Q and Q1V are distal translocation hybrids containing human Xq27.3→qter translocated to different centric rodent chromosomes. The distal translocation hybrids have lost the human sequence detected by p46-1.1. Hybrids Y751B-7 and Y751B-14 show the same 600 kb SalI fragment as the parental hybrid; however all other proximal translocation hybrids show variant bands indicating that probe p46-1.1 detects sequence within 600 kb of these translocation breakpoints.

86% (19/22) of both the proximal and distal translocation breakpoints and thus identifies a fragile X-associated breakpoint cluster region. *In situ* hybridization using YAC 209G4 showed localization to the expressed fragile X site with signal on both flanking boundaries of the isochromatid gap of the fragile site as well within the gap itself, suggesting the presence of uncondensed DNA within the fragile site and indicating that YAC 209G4 includes this region (data not shown).

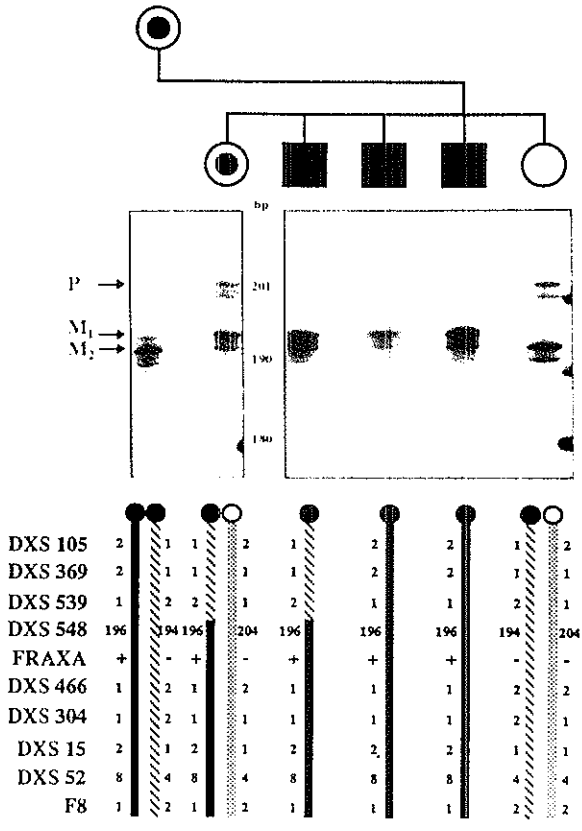


Figure 2. PCR analysis of DXS548 alleles in a fragile X family with key recombinant individuals.

The carrier mother shows two DXS548 alleles at 196 and 194 bp (M_1 and M_2, respectively). The paternal 204 allele of the father (not shown) is seen in the carrier daughter (II-1) who also inherited the maternal 196 bp allele. All three affected males inherited the 196 bp maternal allele (compare with the 194 allele of the normal daughter II-5). The carrier daughter (II-1) and affected son (II-2) are both recombinants between proximal markers DXS105, DXS 369 and DXS539. However, these individuals are non-recombinant with DXS548, placing this locus distal to the crossovers and closer to the fragile X locus.

The close proximity of these YACs to the fragile X locus was independently supported by genetic linkage studies between a polymorphism identified in YAC RS46 and the fragile X locus. DXS548 is a dinucleotide repeat that reveals nine alleles of variable length that are informative in >80% of fragile X families (Riggins et al., 1992). In highly selected families previously shown to have crossovers with tightly linked flanking markers, DXS548 cosegregated, without recombination, with the fragile X locus (lod of 6.95 at $\Theta = 0$). As shown in Figure 2, a carrier daughter and affected son are recombinant between the fragile X locus (*FRAXA*) and proximal markers DXS539 (probe JH89) and DXS369 (probe RN1) which map approximately 5 cM proximal to *FRAXA* with lod scores >40 (Oostra et al., 1990 and B. Oostra, unpublished data). Therefore, DXS548 positions YACs RS46 and 209G4 near the mutation responsible for the clinical phenotype of the fragile X syndrome.

A physical map of YAC 209G4 and of the corresponding genomic region was developed and is shown in Figure 3. A CpG-island containing five infrequently

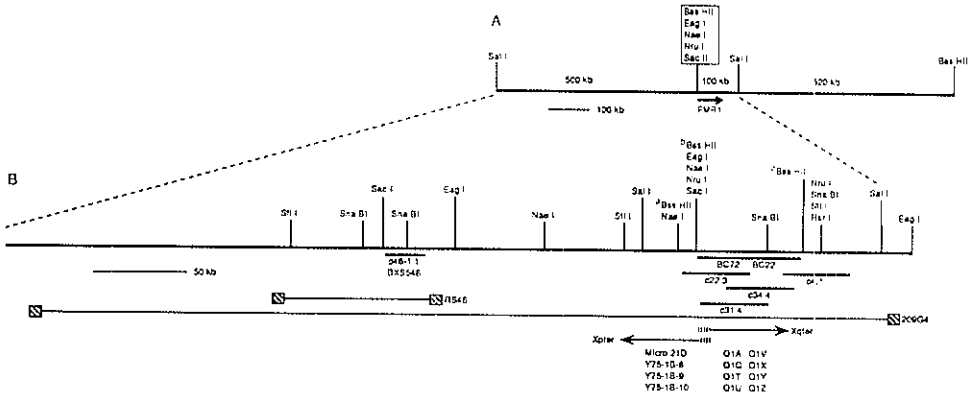


Figure 3. Physical map of the fragile X region of genomic and YAC DNA.

(A) Physical map of the fragile X chromosome in the vicinity of the fragile X locus. Shown are the SalI sites which give rise to the 600 kb fragment seen in hybrid Y75-1B-M1 probed with p46-1.1 and the normal 620 kb BssHII fragment observed in normal X chromosomes. The sites within the box are those previously shown to be methylated on the fragile X chromosome. The position and orientation of *FMR-1* are shown.

(B) Higher resolution physical map derived from both YAC inserts and genomic DNA. Probe p46-1.1 and the DXS548 loci are shown, as are the positions of cDNAs and cosmids. YACs RS46 and 209G4 are shown below in alignment with the map (cross-hatched boxes indicate YAC vector sequences). The positions of the translocation breakpoints are shown as well as the orientation of the map relative to the X chromosome telomeres.

cleaving restriction endonuclease sites was identified 150 kb distal to DXS548. This CpG island is that identified by a previous independent isolation of YAC 209G4 from the CEPH library (Heitz et al., 1991). This CpG island appears hypermethylated on the fragile X chromosome. Vincent et al. (1991) and Bell et al. (1991) have both reported the absence of a normal 620 kb BssHII fragment (Figure 3A) in patients and most carriers of the fragile X syndrome. The absence of the fragment appears to be due to the methylation (and therefore resistance to cleavage) of the BssHII site (b in Figure 3B), leading to a very large band that fails to resolve on PFG electrophoresis. Since CpG islands often are found 5′ to mammalian genes (Lindsay and Bird, 1987) and since methylation of such islands may influence expression of associated genes (Wolf et al., 1984; Yen et al., 1984), it is possible a gene may reside near this fragile X-related CpG island and its expression (or lack of) may be responsible for at least a portion of the fragile X phenotype.

Cosmid Contig Surrounding the Fragile X-Related CpG Island and Breakpoint Cluster Region

To characterize the region surrounding this CpG island, a cosmid library was constructed from the yeast clone harboring YAC 209G4 and cosmids containing human DNA were identified by hybridization to human-specific repetitive elements. A four cosmid contig was identified which spans the fragile X-related CpG island (Figure 3B) from BssHII site a (cosmid 22.3) through BssHII site c (cosmid 4.1).

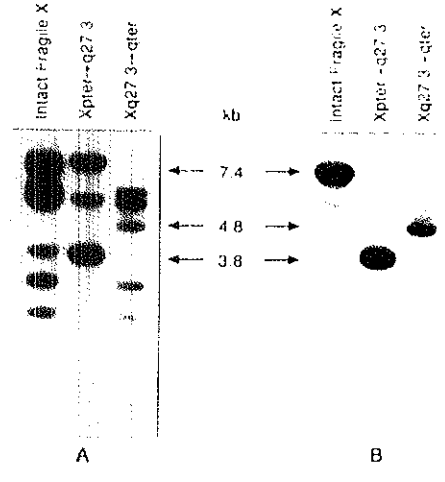


Figure 4. Southern blot analysis of fragile X-associated translocation breakpoints.

DNA of hybrid Y75-1B-M1, containing the intact fragile X chromosome, proximal hybrid micro21D and distal hybrid Q1X was cleaved with EcoRI and 10 µg of each sample loaded onto the gel.

(A) Southern blot hybridized with cosmid 22.3

and

(B) the same filter hybridized with pE5.1.

Cosmid 22.3 was found to include the breakpoints of 11 of 16 tested translocation hybrids (4/5 proximal translocations and 7/11 distal translocations; all 16 breakpoints map within YAC 209G4). As shown in Figure 4A, nine bands (including doublet bands at 5.6 and 5.5 kb), surveying approximately 44 kb of genomic DNA, are observed on Southern analysis of EcoRI digested DNA of the intact fragile X hybrid (Y75-1B-M1) following hybridization with radiolabelled and preannealled cosmid 22.3. Of these nine bands, three are present in the proximal hybrid micro21D (with a novel 3.8 kb junctional fragment) and five are present in the distal hybrid Q1X (with a novel 4.8 kb junctional fragment). The 7.4 kb band of the intact X hybrid Y75-1B-M1 is absent in both translocation hybrids indicating that both breakpoints fall within this interval. The other nine hybrids all exhibited patterns similar to either micro21D or Q1X, with distinct junctional fragments allowing identification of a fragile X-associated breakpoint cluster region (FXBCR) with this 7.4 kb fragment.

The 7.4 kb EcoRI fragment observed above on the fragile X chromosome was not observed in restriction digests of the overlapping cosmids 22.3 and 31.4. However, comparison of the cosmid restriction maps with the genomic EcoRI fragments detected by cosmid 22.3 shows the replacement of the 7.4 kb fragment of the fragile X chromosome with a 5.1 kb fragment of normal DNA which, similar to the 7.4 kb fragment, includes the BssHII site of the fragile X-related CpG island (Figure 5A). The normal 5.1 kb fragment was subcloned to further analyze this difference.

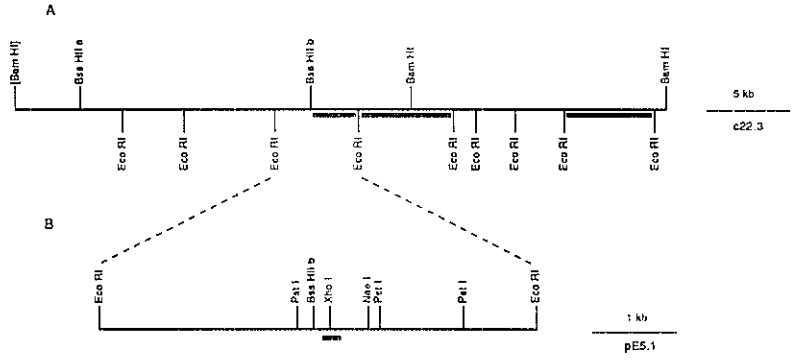


Figure 5. Restriction map of cosmid 22.3 and pE5.1.

(A) Cosmid 22.3 showing BssHII sites a and b as well as EcoRI and BamHI sites. BamHI site in brackets was destroyed during cloning. Solid lines below the map show fragments that hybridize to cDNAs BC72 and BC22.

(B) Map of the cloned 5.1 kb EcoRI fragment of cosmid 22.3 (pE5.1). Solid line below map shows the position of *FMR-1* exonic sequence, which contains the XhoI site.

As shown in Figure 4B, the 5.1 kb fragment (pE5.1; Figure 5B) hybridizes specifically to the 7.4 kb EcoRI fragment of the fragile X chromosome and clearly shows the junctional fragments in micro21D and Q1X. Thus a fragment length difference exists between the normal DNA used to construct YAC 209G4 and the fragile X chromosome of hybrid Y75-1B-M1, and this fragment identifies the FXBCR.

Fragile X Breakpoint Cluster Region Rearranged in Fragile X Patients

Figure 6 shows the results of Southern hybridization of EcoRI digested DNA from two normal and seven unrelated fragile X individuals using pE5.1 as probe. The normal samples (two of five normal samples are shown) exhibit the expected 5.1 kb fragment while all seven fragile X patient DNAs exhibited larger EcoRI fragments with variable increases in size, including the 7.4 kb fragment observed above from hybrid Y75-1B-M1. These data suggest an insertion or amplification event within the normal 5.1 kb fragment that is specific for the fragile X chromosome and is coincident with the fragile X-associated breakpoint cluster region and the fragile X-related CpG island.

Figure 6. Length Variation of EcoRI fragments from normal and fragile X human chromosomes detected with probe pE5.1.

Lanes 1, 6 and 7 show hybridization of the normal 5.1 kb EcoRI fragment in placental DNA (lane 1) and cloned into a cosmid (22.3) or YAC vector (209G4) and seeded into hamster DNA at single-copy level. Somatic cell hybrids containing portions of fragile X chromosomes in hamster backgrounds show bands of altered size from the normal 5.1 kb fragment. Lane 2 contains the hybrid X3000-11.1 (Nussbaum, et al., 1986a). Lane 3 contains DNA from micro28D, a proximal hybrid with a breakpoint distal to the fragile site; lane 4 contains DNA from micro21D, a proximal hybrid with the same chromosome as micro28D, however with a breakpoint detected by pE5.1 (see Figure 5). Lane 5 contains hamster DNA. Lanes 8-12 contain DNA from 5 unrelated fragile X patients' lymphoblastoid lines. These are GM03200, GM04025c, GM06897, GM06912, GM07294. Bands altered from the normal 5.1 kb are seen in each fragile X sample.

Identification and Characterization of *FMR-1*

To search for transcripts associated with the fragile X region, the cosmid subclones of YAC 209G4 were used as hybridization probes to screen a cDNA library derived from normal human fetal brain RNA. Cosmid 4.1, containing BssHII site c (Figure 3B), identified cDNA clone BC22. Restriction digestion and sequence analysis revealed an insert in BC22 of 2832 bp, with an open reading frame at one end extending 1033 bp to a stop codon (Figure 7). Since the reading frame remains open at the 5′ end of the clone, BC22 was used to identify related cDNAs from the same library. Several overlapping clones were isolated, one of which, BC72, was characterized in greater detail. This clone extended the cDNA sequence another 933 bp in the 5′ direction, and overlapped BC22 for approximately 2000 bp toward the 3′ end. Sequence analysis demonstrated that the same reading frame remained open through the 5′ end of BC72, indicating that the 5′ end of the mRNA has not yet been reached, and allowing prediction of a portion (657 amino acids) of the encoded protein. It remains unclear if the entire 3′ portion also was isolated since no poly(A) tract was found at the end of BC22, however a putative polyadenylation addition signal is observed in position 3741 following numerous in frame stop codons (Figure 7). A consensus sequence for nuclear translocation signal KKXK, often followed by a proline, has been described (Roberts, 1989) and is underlined in Figure 7B.

Northern hybridization using the BC22 insert as probe detects an mRNA of approximately 4.8 kb in human brain and placenta (Figure 8), indicating that the 3.8 kb of cDNA obtained does not contain the entire mRNA of this gene. The probe failed to detect signal in human liver, fetal lung and fetal kidney but did detect message in lymphocytes (data not shown).

Figure 9 shows hybridization of BC22 to DNA samples from a number of different organisms. Hybridization signals were observed with all organisms with the exception of Drosophila melanogaster. Since this blot was washed under very stringent conditions (final wash in 0.2x SSC at 65°C for 5 min), cross-hybridization may be observed in Drosophila under less stringent conditions. However, the high stringency of the final wash does indicate the highly conserved nature of this sequence particularly in Caenorhabditis elegans.

A repeated DNA sequence is found close to the 5′ end of BC72 with 28 CGG triplets interspersed with two AGG triplets. In the predicted open reading frame, this repeat would generate a protein domain composed of 30 contiguous arginine residues. Homology searches with the predicted protein sequence identify significant overlaps with a number of arginine-rich proteins, although none contain a polyarginine stretch of equivalent length. The remainder of the protein shows no

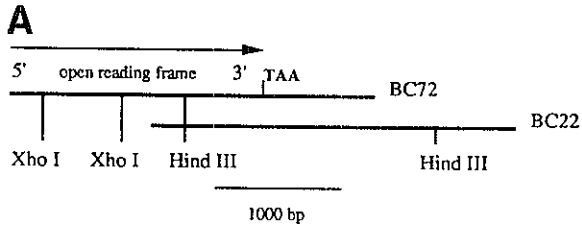


Figure 7. Map and sequence of *FMR-1* cDNA clones.

(A) Map of *FMR-1* cDNA clones.

(B) DNA sequence of *FMR-1* cDNA clones. The 3766 nucleotides of DNA sequence generated from cDNA clones BC72 and BC22 defining the *FMR-1* gene are shown with translation of the 657 amino acid open reading frame. Nucleotides 1-1027 derive from BC72 and nucleotides 934-3766 are from BC22. The CGG repeat encoding 30 contiguous arg residues begins with base 37 and extends to base 127.

significant homology in protein database searches. However, searches against DNA sequence databases identify several related sequences, the strongest of which is with the human androgenreceptor (AR). This is an X-linked gene (mapping to Xq12) with an identical, though smaller, CGG repeat in the first exon which encodes a polyglycine stretch (Tilley et al., 1989).

Location of *FMR-1* Gene Relative to the Fragile X-related CpG island and FXBCR

BC22 demonstrates hybridization to the 70 kb fragment of YAC 209G4 between BssHII sites b and c as well as to cosmids 4.1, 34.4, 31.4 and 22.3 (see Figure 3), indicating exons spanning over 80 kb of DNA. The proximal/distal orientation of the transcript was determined by hybridizing end fragments of BC22 to the cosmid contig. Since the 3′ end of BC22 detected cosmid 4.1 and the 5′ end detected cosmid 22.3, the transcriptional orientation was distal from BssHII site b toward the Xq telomere. This suggests the potential involvement of the fragile X-related CpG island in the regulation of this gene. A 1 kb 5′ fragment of BC72 (to the HindIII site at position 1026 of Figure 7) was used to study the location of the exons encoding this portion of the mRNA in the cosmid and YAC clones. In cosmid 22.3, this probe identifies three EcoRI fragments (see Figure 5A) distal to the BssHII site b. One of the fragments contains the BssHII site (b) as well as the breakpoint cluster region and exhibits length variation in fragile X patients.

B

```
1: GAC GGA GGC GCC CGT GCC AGG GGG CGT GCG GCA GCG CGG CGG CGG CGG CGG CGG CGG CGG CGG AGG CGG CCG CGG CGG CGG CGG CGG
1: asp gly gly ala arg ala arg gly arg ala ala ala arg arg arg arg arg arg arg arg arg arg arg arg arg arg arg arg arg arg

91: CGG CGG AGG CGG CGG CGG CGG CGG CGG CGG CGG CGG CTG GGC CTC GAG CGC CCG CAG CCC ACC TCT CGG GGG CGG GCT CCC GGC GGT AGC
31: arg arg arg arg arg arg arg arg arg arg arg arg leu gly leu glu arg pro gln pro thr ser arg gly arg ala pro gly ala ser

181: AGG GCT GAA GAG AAG ATG GAG GAG CTG GTG GTG GAA GTG CGG GGC TCC AAT GGC GCT TTC TAC AAG GCA TTT GTA AAG GAT GTT CAT GAA
61: arg ala glu glu lys met glu glu leu val val glu val arg gly ser asn gly ala phe tyr lys ala phe val lys asp val his glu

271: GAT TCA ATA ACA GTT GCA TTT GAA AAC AAC TGG CAG CCT GAT AGG CAG ATT CCA TTT CAT GAT GTC AGA TTC CCA CCT CCT GTA GGT TAT
91: asp ser ile thr val ala phe glu asn asn trp gln pro asp arg gln ile pro phe his asp val arg phe pro pro pro val gly tyr

361: AAT AAA GAT ATA AAT GAA AGT GAT GAA GTT GAG GTG TAT TCC AGA GCA AAT GAA AAA GAG CCT TGC TGT TGG TGG TTA GCT AAA GTG AGG
121: asn lys asp ile asn glu ser asp glu val glu val tyr ser arg ala asn glu lys glu pro cys cys trp trp leu ala lys val arg

451: ATG ATA AAG GGT GAG TTT TAT GTG ATA GAA TAT GCA GCA TGT GAT GCA ACT TAC AAT GAA ATT GTC ACA ATT GAA CGT CTA AGA TCT GTT
151: met ile lys gly glu phe tyr val ile glu tyr ala ala cys asp ala thr tyr asn glu ile val thr ile glu arg leu arg ser val

541: AAT CCC AAC AAA CCT GCC ACA AAA GAT ACT TTC CAT AAG ATC AAG CTG GAT GTG CCA GAA GAC TTA CGG CAA ATG TGT GCC AAA GAG GCG
181: asn pro asn lys pro ala thr lys asp thr phe his lys ile lys leu asp val pro glu asp leu arg gln met cys ala lys glu ala

631: GCA CAT AAG GAT TTT AAA AAG GCA GTT GGT GCC TTT TCT GTA ACT TAT GAT CCA GAA AAT TAT CAG CTT GTC ATT TTG TCC ATC AAT GAA
211: ala his lys asp phe lys lys ala val gly ala phe ser val thr tyr asp pro glu asn tyr gln leu val ile leu ser ile asn glu

721: GTC ACC TCA AAG CGA GCA CAT ATG CTG ATT GAC ATG CAC TTT CGG AGT CTG CGC ACT AAG TTG TCT CTG ATA ATG AGA AAT GAA GAA GCT
241: val thr ser lys arg ala his met leu ile asp met his phe arg ser leu arg thr lys leu ser leu ile met arg asn glu glu ala

811: AGT AAG CAG CTG GAG AGT TCA AGG CAG CTT GCC TCG AGA TTT CAT GAA CAG TTT ATC GTA AGA GAA GAT CTG ATG GGT CTA GCT ATT GGT
271: ser lys gln leu glu ser ser arg gln leu ala ser arg phe his glu gln phe ile val arg glu asp leu met gly leu ala ile gly

901: ACT CAT GGT GCT AAT ATT CAG CAA GCT AGA AAA GTA CCT GGG GTC ACT GCT ATT GAT CTA GAT GAA GAT ACC TGC ACA TTT CAT ATT TAT
301: thr his gly ala asn ile gln gln ala arg lys val pro gly val thr ala ile asp leu asp glu asp thr cys thr phe his ile tyr

991: GGA GAG GAT CAG GAT GCA GTG AAA AAA GCT AGA AGC TTT CTC GAA TTT GCT GAA GAT GTA ATA CAA GTT CCA AGG AAC TTA GTA GTA ATA
331: gly glu asp gln asp ala val lys lys ala arg ser phe leu glu phe ala glu asp val ile gln val pro arg asn leu val val ile

1081: GGA AAA AAT GGA AAG CTG ATT CAG GAG ATT GTG GAC AAG TCA GGA GTT GTG AGG GTG AGG ATT GAG GCT GAA AAT GAG AAA AAT GTT CCA
361: gly lys asn gly lys leu ile gln glu ile val asp lys ser gly val val arg val arg ile glu ala glu asn glu lys asn val pro

1171: CAA GAA GAG GAA ATT ATG CCA CCA AAT TCC CTT CCT TCC AAT AAT TCA AGG GTT GGA CCT AAT GCC CCA GAA GAA AAA AAA CAT TTA GAT
391: gln glu glu glu ile met pro pro asn ser leu pro ser asn asn ser arg val gly pro asn ala pro glu glu lys lys his leu asp

1261: ATA AAG GAA AAC AGC ACC CAT TTT TCT CAA CCT AAC AGT ACA AAA GTC CAG AGG GGT ATG GTA CCA TTT GTT TTT GTG GGA ACA AAG GAC
421: ile lys glu asn ser thr his phe ser gln pro asn ser thr lys val gln arg gly met val pro phe val phe val gly thr lys asp

1351: AGC ATC GCT AAT GCC ACT GTT CTT TTG GAT TAT CAC CTG AAC TAT TTA AAG GAA GTA GAC CAG TTG CGT TTG GAG AGA TTA CAA ATT GAT
451: ser ile ala asn ala thr val leu leu asp tyr his leu asn tyr leu lys glu val asp gln leu arg leu glu arg leu gln ile asp

1441: GAG CAG TTG CGA CAG ATT GGA GCT AGT TCT AGA CCA CCA CCA AAT CGT ACA GAT AAG GAA AAA AGC TAT GTG ACT GAT GAT GGT CAA GGA
481: glu gln leu arg gln ile gly ala ser ser arg pro pro pro asn arg thr asp lys glu lys ser tyr val thr asp asp gly gln gly

1531: ATG GGT CGA GGT AGT AGA CCT TAC AGA AAT AGG GGG CAC GGC AGA CGC GGT CCT GGA TAT ACT TCA GGA ACT AAT TCT GAA GCA TCA AAT
511: met gly arg gly ser arg pro tyr arg asn arg gly his gly arg arg gly pro gly tyr thr ser gly thr asn ser glu ala ser asn

1621: GCT TCT GAA ACA GAA TCT GAC CAC AGA GAC GAA CTC AGT GAT TGG TCA TTA GCT CCA ACA GAG GAA GAG AGG GAG AGC TTC CTG CGC AGA
541: ala ser glu thr glu ser asp his arg asp glu leu ser asp trp ser leu ala pro thr glu glu glu arg glu ser phe leu arg arg

1711: GGA GAC GGA CGG CGG CGT GGA GGG GGA GGA AGA GGA CAA GGA GGA AGA GGA CGT GGA GGA GGC TTC AAA GGA AAC GAC GAT CAC TCC CGA
571: gly asp gly arg arg arg gly gly gly gly arg gly gln gly gly arg gly arg gly gly gly phe lys gly asn asp asp his ser arg

1801: ACA GAT AAT CGT CCA CGT AAT CCA AGA GAG GCT AAA GGA AGA ACA ACA GAT GGA TCC CTT CAG AAT ACC TCC AGT GAA GGT AGT CGG CTG
601: thr asp asn arg pro arg asn pro arg glu ala lys gly arg thr thr asp gly ser leu gln asn thr ser ser glu gly ser arg leu

1891: CGC ACG GGT AAA GAT CGT AAC CAG AAG AAA GAG AAG CCA GAC AGC GTG GAT GGT CAG CAA CCA CTC GTG AAT GGA GTA CCC TAA
631: arg thr gly lys asp arg asn gln lys lys glu lys pro asp ser val asp gly gln gln pro leu val asn gly val pro ***

1975: ACT GCA TAA TTC TGA AGT TAT ATT TCC TAT ACC ATT TCC GTA ATT CTT ATT CCA TAT TAG AAA ACT TTG TTA GGC CAA AGA CAA ATA GTA
2065: GGC AAG ATG GCA CAG GGC ATG AAA TGA ACA CAA ATT ATG CTA AGA ATT TTT TAT TTT TTG GTA TTG GCC ATA AGC AAC AAT TTT CAG ATT
2155: TGC ACA AAA AGA TAC CTT AAA ATT TGA AAC ATT GCT TTT AAA ACT ACT TAG CAC TTC AGG GCA GAT TTT AGT TTT ATT TTC TAA AGT ACT
2245: GAG CAG TGA TAT TCT TTG TTA ATT TGG ACC ATT TTC CTG CAT TGG GTG ATC ATT CAC CAG TAC ATT CTC AGT TTT TCT TAA TAT ATA GCA
2335: TTT ATG GTA ATC ATA TTA GAC TTC TGT TTT CAA TCT CGT ATA GAA GTC TTC ATG AAA TGC TAT GTC ATT TCA TGT CCT GTG TCA GTT TAT
2425: GTT TTG GTC CAC TTT TCC AGT ATT TTA GTG GAC CCT GAA ATG TGT GTG ATG TGA CAT TTG TCA TTT TCA TTA GCA AAA AAA GTT GTA TGA
2515: TCT GTG CCT TTT TTA TAT CTT GGC AGG TAG GAA TAT TAT ATT TGG ATG CAG AGT TCA GGG AAG ATA AGT TGG AAA CAC TAA ATG TTA AAG
2605: ATG TAG CAA ACC CTG TCA AAC ATT AGT ACT TTA TAG AAG AAT GCA TGC TTT CCA TAT TTT TTT CCT TAC ATA AAC ATC AGG TTA GGC AGT
2695: ATA AAG AAT AGG ACT TGT TTT TGT TTT TGT TTT GTT GCA CTG AAG TTT GAT AAA TAG TGT TAT TGA GAG AGA TGT GTA ATT TTT CTG TAT
2785: AGA CAG GAG AAG AAA GAA CTA TCT TCA TCT GAG AGA GGC TAA AAT GTT TTC AGC TAG GAA CAA ATC TTC CTG GTC GAA AGT TAG TAG GAT
2875: ATG CCT GCT CTT TGG CCT GAT GAC CAA TTT TAA CTT AGA GCT TTT TTT TTT AAT TTT GTC TGC CCC AAG TTT TGT GAA ATT TTT CAT ATT
2965: TTA ATT TCA AGC TTA TTT TGG AGA GAT AGG AAG GTC ATT TCC ATG TAT GCA TAA TAA TCC TGC AAA GTA CAG GTA CTT TGT CTA AGA AAC
3055: ATT GGA AGC AGG TTA AAT GTT TTG TAA ACT TTG AAA TAT ATG GTC TAA TGT TTA AGC AGA ATT GGA AAA GAC TAA GAT CGG TTA ACA AAT
3145: AAC AAC TTT TTC TTT TTT TCT TTT GTT TTT TGA AGT GTT GGG GTT TGG TTT TGT TTT TTG AGT CTT TTT TTT TTA AGT GAA ATT TAT
3235: TGA GGA AAA ATA TGT GAA GGA CCT TCA CTC TAA GAT GTT ATA TTT TTC TTA AAA AGT AAC TCC TAG TAG GGG TAC CAC TGA ATC TGT ACA
3325: GAG CCG TAA AAA CTG AAG TTC TGC CTC TGA TGT ATT TTG TGA GTT TGT TTC TTT GAA TTT TCA TTT TAC AGT TAC TTT TCC TTG CAT ACA
3415: AAC AAG CAT ATA AAA TGG CAA CAA ACT GCA CAT GAT TTC ACA AAT ATT AAA AAG TCT TTT AAA AAG TAT TGC CAA ACA TTA ATG TTG ATT
3505: TCT AGT TAT TTA TTC TGG GAA TGT ATA GTA TTT GAA AAC AGA AAT TGG TAC CTT GCA CAC ATC ATC TGT AAG CTG TTT GGT TTT AAA ATA
3595: CTG TAG ATA ATT AAC CAA GGT AGA ATG ACC TTG TAA TGT AAC TGC TCT TGG GCA ATA TTC TCT GTA CAT ATT AGC GAC AAC AGA TTG GAT
3685: TTT ATG TTG ACA TTT GTT TGG TTA TAG TGC AAT ATA TTT TGT ATG CAA GCA GTT TCA ATA AAG TTT GAT CTT CCT CTG CTA CCG
```

Figure 8. Northern blot analysis of poly(A) RNA hybridized with cDNA BC22.

Five µg of poly(A) selected RNA from human brain (lane 1) and normal placenta (lane 2) were electrophoresed, blotted onto a GeneScreen Plus filter and hybridized with radiolabelled BC22 insert. A single hybridizing species of approximately 4.8 kb is observed in each lane.

Restriction mapping and direct sequencing of the 5.1 kb EcoRI fragment using a primer derived from BC72 sequence (position 223 to 246) demonstrated an exon immediately distal to the BssHII site b. This exon contains an XhoI site (position 137 in *FMR-1* cDNA sequence) that is found 310 nucleotides from the BssHII site in genomic DNA (see Figure 5B), allowing the 5′ end of BC72 to be positioned 220 bp from the BssHII site. This exon, within a 1 kb PstI fragment (Figure 5B), also contains the block of CGG repeats which are seen in the sequence analysis of the genomic DNA as well. This PstI fragment has been found to contain two translocation breakpoints (micro21D and Q1X) distal to the BssHII site (b) as well as to be the region of length variation observed in fragile X chromosomes (Yu et al., 1991). Therefore the CGG repeat is found within the fragile X-related CpG island, constituting a portion of this CpG-rich region, and is in the immediate proximity of the breakpoint cluster region and the region of length variation.

DISCUSSION

We report above the YAC cloning of 475 kb of human DNA corresponding to the region involved in the fragile X syndrome and show the presence of a gene, designated *FMR-1*, adjacent to the fragile X-related CpG island. It is further demonstrated that within a 7.4 kb EcoRI fragile X genomic fragment, which contains the CpG island and *FMR-1* exonic sequences, is a breakpoint cluster region where the majority of fragile X-associated translocation breakpoints position. This same EcoRI fragment, which is 5.1 kb in normal X chromosomes, undergoes a variable increase in size in the fragile X chromosome. The fragile X variation and breakpoint clustering may involve a highly unusual CGG repeat found within the coding region of *FMR-1* and also present within this same EcoRI fragment, 220 bp distal to the BssHII site, demonstrating fragile X-related

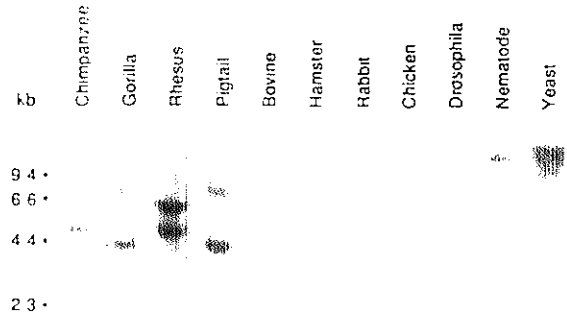


Figure 9. Zoo blot analysis of DNA isolated from several species hybridized with cDNA BC22. Ten µg of DNA from each species was cleaved with EcoRI and electrophoresed and blotted onto nylon membrane. Hybridization was carried out with labelled cDNA overnight using standard conditions and washed to a final stringency of 0.2x SSC for 5 min at 65°C.

methylation.

The data presented above suggest the identification of the region containing the fragile X site. A breakpoint cluster region is demonstrated where the majority of fragile X-associated translocation breakpoints fall. These translocations were identified by the segregation of the marker loci *HPRT* and *G6PD* which flank the fragile X site (Warren et al., 1987, 1990). Since the translocation breakpoints are flanked by the same polymorphic loci which genetically flank the fragile X syndrome locus (Rousseau et al, 1991; Hirst et al., 1991), the breakpoints were believed to cluster at the fragile X site. We show above that these breakpoints indeed cluster within a 7.4 kb EcoRI fragment. Since almost 70% (11/16) of tested translocation breakpoints fall within 7.4 kb, an interval of less than 0.03% of the estimated 30 Mb which separate the marker loci used for selection, the clustering is quite remarkable and is compelling evidence for the presence of the fragile X site within this fragment. This is further supported by the observation that this EcoRI fragment is markedly smaller (5.1 kb) in normal chromosomes and when the normal fragment (pE5.1) is used as probe on Southern blot analysis of fragile X DNA, significantly increased size variation is observed. This suggests that an insertion and/or amplification event within the normal 5.1 kb fragment occurred in the fragile X chromosome. Finally, *in situ* hybridization of cloned material from this region maps precisely within the isochromatid gap characteristic of the fragile X site.

Within this 5.1 kb fragment, the fragile X related CpG island is found (Heitz et al., 1991) which contains a BssHII site hypermethylated in fragile X chromosomes (Vincent et al., 1991; Bell et al., 1991). The correlation of this CpG island and its methylation status also support the involvement of the 5.1 kb fragment, reported above, in the fragile X syndrome. Indeed, linkage analysis using a highly polymorphic dinucleotide repeat (DXS548) located within 150 kb (Figure 3) of the CpG island demonstrates tight linkage with the fragile X locus (*FRAXA*) without recombination. Such linkage is of significance as it represents evidence independent of the physical mapping studies and is correlated to the clinical expression of the fragile X syndrome, indicating the close proximity of the fragile X site and a gene involved in phenotype.

Such a gene is described above, designated *FMR-1* (Fragile X Mental Retardation - 1). There is compelling evidence that *FMR-1* is important in the phenotypic consequences of inheriting the fragile X locus: first, the 5' association of *FMR-1* with the CpG island shown to be selectively associated with hypermethylation in the fragile X chromosome; second, the presence of a *FMR-1* exon within the 5.1 kb fragment, which is increased in size in the fragile X chromosome and contains the fragile X-associated breakpoint cluster region; third the presence of a highly unusual trinucleotide repeat CGG, which is reiterated 30 times in the normal mRNA of *FMR-1* and the placement of this repeat 250 bp distal to the BssHII site of the CpG island within the 5.1 kb fragment; and fourth, the expression of *FMR-1* in human brain, as would be expected for an inherited disorder whose major phenotype is mental retardation. Although the finding of *FMR-1* does not preclude the additional involvement of other genes in the fragile X phenotype, *FMR-1* is a likely candidate, particularly since there are no other readily apparent CpG islands in the immediate vicinity (Figure 3 and Heitz et al., 1991).

Further study of *FMR-1* is required to show a definitive involvement with the fragile X syndrome including currently ongoing expression studies in normal and affected individuals. However, the data presented above clearly indicate this gene to be of considerable interest independent of this involvement. The CGG repeat within the coding region is quite striking since the reading frame of the message would suggest a protein containing 30 contiguous arginine residues. If this repeat is not processed off, *FMR-1* would encode an exceptionally basic protein. No significant protein matches were found on database searches, except for the polyarginine stretch found in histones and protamines of different organisms, although the arginine stretches in these proteins are shorter and interspersed with other amino acids (Lee et al., 1987; Martin-Ponthieu et al., 1991). Considering the fact that DNA binding proteins such as histones and protamines contain arginine

stretches of variable length and the fact that a nuclear translocation signal is found, it might well be considered that *FMR-1* has a DNA binding function in the cell nucleus. *FMR-1* appears to have been strongly conserved through evolution with stringent hybridization through yeast. Hybridization to such organisms should allow the isolation and sequencing of the homologous sequences which may reveal conserved protein domains within *FMR-1*, indicating regions of functional importance. Also, such homologies, particularly with C. elegans, may allow experimental approaches in other organisms to unravel the function of *FMR-1*.

It is tempting to speculate that the unusual *FMR-1* CGG repeat, found within the same 5.1 kb fragment associated with the breakpoint cluster region and fragile X patient length variation, is a component of the fragile site itself. It has long been speculated that the fragile X site is a repeat of variable length (Nussbaum et al., 1986b; Ledbetter et al., 1986; Warren et al., 1987). This possibility awaits further analysis of the repeat region in normal and fragile X chromosomes as well as the junctional sequences of the translocation chromosomes. If the fragile site is not the CGG repeat, the data presented above strongly suggest that the responsible sequence is adjacent to the repeat within the implicated EcoRI fragment. However, the high CpG density of this repeat may influence local methylation status and if expanded may result in the abnormal methylation levels observed in fragile X chromosomes.

The finding of larger fragment lengths in fragile X chromosomes relative to normal chromosomes does indicate a molecular diagnostic approach which should be superior to either linkage analysis or cytogenetic examination for the fragile X site. In particular, Southern blot hybridization with pE5.1 of DNA digested with both EcoRI and BssHII should document both the methylation status of the BssHII site and the size of the resulting fragment(s).

In summary, we report the presence of a brain-expressed gene (*FMR-1*) containing a CGG repeat that is coincident with a fragile X breakpoint cluster region, which itself exhibits length variation in fragile X chromosomes. The breakpoint cluster region and the interval of length variation are likely to contain the sequence responsible for the cytologic expression of the fragile X site. *FMR-1* exonic sequence, particularly the CGG repeat, also map to this same interval and therefore are likely related in some manner. Although the exact nature of the fragile X site, the mechanisms of the mutation and phenotypic consequences remain to be elucidated, these data provide the elements needed to now explore the molecular biology, biochemistry and cell biology of this unique and puzzling genetic disease.

EXPERIMENTAL PROCEDURES

Pulsed Field Gel Electrophoresis

Southern blot analysis of genomic DNA or YAC DNA resolved by pulsed field gel electrophoresis (PFGE) was performed essentially as described (Smith et al., 1988). Trypsinized and washed mammalian cells were suspended in molten agarose (final concentration 0.5% wt/vol; Baker) prepared in SE buffer (75 mM NaCl, 25 mM EDTA, pH 8.0) at a final concentration of 1.5 x 10^7cells/ml. Chromosomal DNAs were isolated from YAC clones as described (Anand, et al., 1989). Yeast cells from a 10 ml saturated culture were harvested, rinsed once in 50 mM EDTA, pH 8.0 and recovered in 0.5 ml SBE-zymolase (1 M sorbitol, 25 mM EDTA pH 8.0, 14 mM 2-mercaptoethanol, 1 mg/ml zymolase [ICN]). 0.5 ml 1% Seaplaque agarose (FMC Corp.) in SBE (without zymolase) was added and the suspension transferred to plug molds. Spheroplast generation (for yeast cells) was for 5 hours to overnight in SBE-zymolase. Cell lysis (mammalian or yeast cells) was for 2 days in ESP (0.5 M EDTA, pH 9.5, 1% N-lauroylsarcosine, 1 mg/ml proteinase K) at 50°C. Restriction endonuclease digestion was performed using the manufacturer's recommended buffers and conditions with a 50 µl plug slice in 250 µl of buffer containing 50 units of enzyme. For double digests, the plugs were rinsed and equilibrated, following digestion with the first enzyme, with the second buffer several times prior to digestion with the second enzyme. PFGE was carried out on a Bio-Rad contour-clamped homogeneous electric field (CHEF) DRII apparatus through 1% agarose (BRL) at 200 V and 14° C in 0.5x TBE buffer (45 mM Tris-borate, 1 mM EDTA). For resolution of fragments of ~200-1200 kb, switch time was 60 sec for 17 hrs followed by 90 sec for 10 hrs; for resolution of fragments ~10-500 kb, the switch times were ramped from 5 sec to 50 sec over 27 hrs. Southern blotting and hybridization were carried out as described (Southern, 1975; Feinberg and Vogelstein, 1984) with the exception that acid depurination in 0.25 M HCl was allowed to proceed 20 min for pulsed-field gels. Radiolabeled probes were synthesized by random priming from 50 ng gel purified fragments except when intact cosmids were used which were nick translated (Boehringer Mannheim kit; following manufacturer's recommendations). For genomic probes containing repetitive elements, repeat suppression was accomplished by preassociation with 1-3 mg of sonicated human placental DNA in 100-300 µl of 5x SSC (1x SSC is 150 mM NaCl, 15 mM NaCitrate, pH 7.0) for 3-10 min at 65°C prior to the addition to the filter. Washing was carried out to a final stringency wash of 0.2x SSC for 15 min at 65°C prior to autoradiography. S. cerevisiae strain YNN295 chromosomes (BioRad), concatamers of phage lambda (BioRad) or high molecular weight markers (BRL) were used as size standards.

Cosmid Library Construction of YAC 209G4

Agarose plugs (0.5%, SeaPlaque FMC) containing 5-10 µg of yeast DNA were prepared as described (Van Ommen and Verkerk, 1986). Blocks (100 µl) of DNA were equilibrated on

ice in 0.5 ml of MboI digestion buffer, containing 0.1 mg/ml bovine serum albumin (BSA, MB grade; Boehringer Mannheim). After 2-3 hrs, the buffer was replaced by 150 µl of fresh buffer to which MboI was added (0.0001-0.0007 units). Following overnight incubation on ice, digestion was carried out for 40 min at 37°C. The agarose blocks were melted, the DNA dephosphorylated with 1 unit calf intestinal alkaline phosphatase (Boehringer Mannheim), and treated with 2.5 units of agarase (Calbiochem). The solution was extracted twice with phenol-chloroform, once with chloroform, and the DNA precipitated with ethanol and dissolved in 10 mM Tris, 0.1 mM EDTA (pH 7.4) at a concentration of 500 ng/µl. DNA (250 ng) was ligated to 500 ng of BstBI (dephosphorylated) and BamHI digested vector (p2CpG; Douwerse et al., 1989). Ligation and packaging was carried out according to standard procedures. Cosmids containing human inserts were selected by hybridizing with human specific Alu-repeat probe.

YAC and Cosmid Subcloning

YACs were subcloned following isolation of the intact chromosome by preparative pulsed-field gel electrophoresis and EcoRI digestion of the DNA in molten agarose (Seaplaque; FMC). Fragments were phenol/chloroform extracted, ethanol precipitated, recovered, and ligated into EcoRI cut, dephosphorylated, lambda ZAP II arms according to manufacturer's recommendations (Stratagene). Cosmids were subcloned following an alkaline lysis isolation and EcoRI digestion. Fragments were phenol-chloroform extracted and ethanol precipitated prior to ligation into lambda ZAP II arms as with YAC fragments. In the case of both cosmids and YACs, 75 ng EcoRI fragments were ligated to 1 ug vector arms. Selected phage were converted into pBluescript II SK$^-$ clones following in vivo excision of plasmid with insert according to manufacturer's guidelines.

Polymerase Chain Reaction Analysis of DXS548 Alleles

Analysis of DXS548 was carried out as described (Riggins et al., 1992). Amplification was carried out on 0.2-0.5 µg of genomic DNA in a 10 µl total reaction containing 0.25 mM dNTPs, 40 ng of primers 5′ AGAGCTTCACTATGCAATGGAATC and 5′ GTACATTAGAGTCACCTGTGGTGC, and 0.25 units of Taq polymerase in a buffer of 10 mM Tris-HCl, 50 mM KCl, 12 mM MgCl and 0.01% gelatin. Twenty three cycles of PCR was carried out in the following fashion: 3 cycles of 1 min each at 97°C, 62°C annealing and 72°C extension followed by 20 additional cycles with the annealing temperature lowered to 55°C. The reaction volume was then increased to 50 µl with the same reaction components and concentrations except that one primer was 5′ end-labelled with [^{32}P]-ATP as described (Sambrook et al., 1989). PCR was continued for 10 cycles of 1 min each at 95°C denaturation, 62°C annealing and 72°C extension. PCR products were analyzed by electrophoresis of 2 µl of reaction through a 40 cm 6% polyacrylamide denaturing sequencing gel for approximately 2.25 hrs. The gel was dried without fixing and exposed to X-ray film overnight at room temperature.

cDNA Library Screening

A human fetal brain lambda gt11 cDNA library (Clonetech, Palo Alto, CA) of 1.3 x 10^6 independent clones with insert lengths of 0.7-4.0 kb was used. The library was plated on 15 cm plates at a density of 50,000 pfu per dish using strain LE392. Filter lifts were prepared according to standard techniques and the library screened with cosmid DNA hexanucleotide labelled with [^{32}P]-dATP and [^{32}P]-dCTP. The labelled DNA was first prehybridized with 100 µg of total sheared human genomic DNA and 100 µg cosmid vector DNA in 5x SSC at 65°C for 2 hrs. Following hybridization for 16 hrs, the filters were washed to a stringency of 0.1x SSC. The filters were exposed to Fuji film with intensifing screens for 2 days at -80°C.

Northern Blot Analysis

Total RNA was extracted using guanidinium isothiocyanate followed by centrifugation through cesium chloride as described (Sambrook et al., 1990). Poly(A)$^+$ RNA was selected by passage through oligo(dT) cellulose (Aviv and Leder, 1972). Human brain, liver, and fetal poly(A) RNA was purchased from Clontech Laboratories (Palo Alto, CA).

Five µg of poly(A) containing RNA or 25 µg of total RNA were precipitated and dissolved in 20 µl of 50% (vol/vol) formaldehyde and 1x MEN (20 mM MOPS, pH 6.8, 5 mM sodium acetate, 1 mM EDTA) and incubated for 10 min at 60°C; 5 µl of dye marker (50% sucrose, 0.5% bromophenol-blue) was added and the samples were loaded on a formaldehyde-agarose gel. Electrophoresis was carried out for 3 hrs. at 100 V and the gel then soaked for 30 min in 20x SSC and blotted onto a nitrocellulose or nylon (GeneScreen Plus, Dupont) overnight in 10x SSC (Thomas, 1980). The RNA was fixed to the membranes by baking under vacuum for 2 hrs at 80°C. The membranes were prehybridized in 50% formamide, 5x Denhart's, 50 mM sodium phosphate, pH 6.8, 10% dextran sulfate and 100 µg of denatured salmon sperm DNA at 42°C for 2-4 hrs. Hybridization with the probe was for 16-20 hrs at 42°C in the above buffer. Filters were washed with 3x SSC, 0.1% SDS at 50°C and then the SSC concentration was lowered according to the level of background, with a final wash in 0.1x SSC, 0.1% SDS.

DNA Sequencing

Sequence of BC22 was obtained via a shotgun strategy as described in Bankier et al. (1987) using dideoxynucleotide termination. The sequencing reactions were analyzed on an automated DNA sequencer (ABI 373) using fluorescently labelled oligonucleotide primers. Sequence information derived from 25 random M13 clones was assembled using the Sequence Assembly manager software developed by the Molecular Biology Information Resource (MBIR) of the Department of Cell Biology, Baylor College of Medicine.

An EcoRI to HindIII fragment at the 5′ end of BC72 was subcloned into Bluescript II KS$^-$ and sequenced with vector and internal primers. Sequence at the extreme 5′ end is incomplete due to the difficulty of extension by polymerases through the repeated CGG sequence.

Acknowledgements

We thank K. E. Bakker, R. M. v.d. Helm, J. F. Peters, and C. L. Stayton for assistance and helpful discussion; S. Ledbetter and D. Ledbetter for their continuous support and D. Le Paslier of CEPH (Paris) for the generous gift of the YAC library/clone. G. J. R. is a predoctoral fellow of the March of Dimes Birth Defects Foundation (18-90-57). O. R. is a postdoctoral fellow of the Human Frontier Science Program. This research was supported by the Foundation of Clinical Genetics to A. J. M. H. V.; a grant of the MGCZWN to B. A. O.; a DOE grant (FG05-88ER60692) to D. L. N. and C. T. C.; NIH Genome Center Grant and the Howard Hughes Medical Institute to C. T. C.; and NIH grants (HD20521 and HG00038) to S. T. W.

The costs of publication of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 USC Section 1734 solely to indicate this fact.

Received May 9, 1991; revised may 16, 1991.

REFERENCES

Albertsen, H. M., Abderrahim, H., Cann, H. M., Dausset, J., Le Paslier, D., and Cohen, D. (1990). Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents. Proc. Natl. Acad. Sci. USA *87*, 4256-4260.

Anand, R., Villasante, A., and Tyler-Smith, C. (1989). Construction of yeast artificial chromosome libraries with large inserts using fractionation by pulsed-field gel electrophoresis. Nucl. Acids Res. *17*, 3425-3433.

Aviv, H., and Leder, P. (1972). Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci. USA *69*, 1408-1412.

Bankier, A.T., Weston, K. M., and Barrel, B. G. (1987). Random cloning and sequencing by the M13 dideoxynucleotide termination method. Methods in Enzymology *155*, 51-93.

Bell, M. V., Hirst, M. C., Nakahori, Y., MacKinnon, R. N., Roche, A., Flint, T. J., Jacobs, P. A., Tommerup, N., Tranebjaerg, L., Froster-Iskenius, U., Kerr, B., Turner, G., Lindenbaum, R. H., Winter, R., Pembry, M., Thibodeau, S., and Davies, K. E. (1991). Physical mapping across the fragile X: hypermethylation and clinical expression of the fragile X syndrome. Cell *64*, 861-866.

Brown, W. T. (1990). The fragile X: progress towards solving the puzzle. Am. J. Hum. Genet. *47*, 175-180.

Dauwerse, H. G., van Ommen, G. B., Breuning, M. H., and Pearson, P. L. (1989). Cosmid vector pCpG and plasmid vector pKNUNI, and their uses for cloning DNA sequences

adjacent to sites for rare cutting restriction endonucleases. Nucl. Acids Res. *17*, 3603.

Feinberg, A. P., and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction fragments to high specific activity. Anal. Biochem. *132*, 6-13.

Gustavson, K.-H., Blomquist, H., and Holmgren, G. (1986). Prevalence of fragile-X syndrome in mentally retarded children in a Swedish county. Am. J. Med. Genet. *23*, 581-588.

Heitz, D., Rousseau, F., Devys, D., Saccone, S., Abderrahim, H., Le Paslier, D., Cohen, D., Vincent, A., Toniolo, D., Della Valle, G., Johnson, S., Schlessinger, D., Oberle, I., and Mandel, J.-L. (1991). Isolation of sequences that span the fragile X and identification of a fragile X-related CpG island. Science *251*, 1236-1239.

Hirst, M. C., Roche, A., Flint, T. J., MacKinnon, R. N., Bassett, J. H. D., Nakahori, Y., Watson, J. E. V., Bell, M. V., Patterson, M. N., Boyd, Y., Thomas, N. S. T., Knight, S. J. L., Warren, S. T., Hors-Cayla, M., Schmidt, M., and Davies, K. E. (1991). Linear order of new and established markers around the fragile site at Xq27.3. Genomics *10*, 243-249.

Krawczun, M. S., Jenkins, E. C., and Brown, W. T. (1985). Analysis of the fragile-X chromosome: localization and detection of the fragile site in high resolution preparations. Hum. Genet. *69*, 209-211.

Ledbetter, D. H., Ledbetter, S. A., and Nussbaum, R. L. (1986). Implications of fragile X expression in normal males for the nature of the mutation. Nature *324*, 161-163.

Lee, C-H, Hoyer-Fender, S., and Engel, W. (1987). The nucleotide sequence of a human protamine 1 cDNA. Nucl. Acids Res. *15*, 7639.

Lindsay, S., and Bird, A. P. (1987). Use of restriction enzymes to detect potential gene sequences in mammalian DNA. Nature *327*, 336-338.

Martin- Ponthieu, A., Wouters-Tyrou, D., Belaiche, D., Sautiere, P., Schindler, P., and Van Dorsselaer, A. (1991). Cuttlefish sperm protamines. Eur. J. Biochem. *195*, 611-619.

Nelson, D. L., Ballabio, A., Victoria, M. F., Pieretti, M., Bies, R. D., Gibbs, R. A., Maley, J. A., Chinault, A. C., Webster, T. D., and Caskey, C. T. (1991). Alu PCR for regional assignment of 110 yeast artificial chromosome clones from the human X chromosome: identification of clones associated with a disease locus. Proc. Natl. Acad. Sci. USA *88*, 6157-6161.

Nussbaum, R. L., Airhart, S. D., and Ledbetter, D. H. (1986). A rodent-human hybrid containing Xq24-qter translocated to a hamster chromosome expresses the Xq27 folate-sensitive fragile site. Am. J. Med. Genet. *23*, 457-466.

Nussbaum, R. L., Airhart, S. D., and Ledbetter, D. H. (1986). Recombination and amplification of pyrimidine-rich sequences may be responsible for initiation and progression of the Xq27 fragile site. Am. J. Med. Genet. *23*, 715-721.

Nussbaum, R. L., and Ledbetter, D. H. (1990). The fragile X syndrome. In The Metabloic Basis of Inherited Disease, Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds. (New York: McGraw Hill), pp.327-341.

Oostra, B. A., Hupkes, P. E., Perdon, L. F., van Bennekom, C. A., Baker, E., Halley, D. J. J., Schmidt, M., Du Sart, D., Smits, A., Wieringa, B., and van Oost, B. A. (1990). New polymorphic DNA marker close to the fragile site FRAXA. Genomics. *6*, 129-132.

Riggins, G.J., Sherman, S.L., Oostra, B.A., Sutcliffe, J.S., Feitell, D., Nelson, D.L., vaan Oost, B.A., Smits, A.P.T., Ramos, F.J., Pfendner, E., Kuhl, D.P.A., Caskey, C.T., and Warren, S.T. (1992). Characterization of a highly polymorphic dinucleotide repeat 150 Kb proximal to the fragile X site. Am. J. Med. Genet. *43*, 237-243, 1992.

Roberts, B. (1989). Nuclear location signal-mediated protein transport. Biochem. Biophy. Acta *1008*, 263-280.

Rousseau, F., Vincent, A., Rivella, S., Heitz, D., Tribioli, C., Maestrini, E., Warren, S. T., Suthers, G. K., Goodfellow, P., Mandel, J.-L., Toniolo, D., and Oberle, I. (1991). Four chromosomal breakpoints and four new probes mark out a 10 cM region encompassing the FRAXA locus. Am. J. Hum. Genet. *48*, 108-116.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor: Cold Spring Harbor Press), pp. 7.19-7.22, 10.59-10.66.

Sherman, S. L., Morton, N. E., Jacobs, P. A., and Turner, G. (1984). The marker (X) syndrome: a cytogenetic and genetic analysis. Ann. Hum. Genet. *48*, 21-37.

Sherman, S. L., Jacobs, P. A., Morton, N. E., Froster-Iskenius, U., Howard-Peebles, P. N., Nielsen, K. B., Partington, N. W., Sutherland, G. R., Turner, G., and Watson, M. (1985). Further segregation analysis of the fragile X syndrome with special reference to transmitting males. Hum. Genet. *69*, 3289-3299.

Smith, C. L., Klco, S. R., and Cantor, C. R. (1988). Pulsed-field gel electrophoresis and the technology of large DNA molecules. In Genome Analysis: A Practical Approach, K. E. Davies, ed. (Oxford: IRL Press), pp. 41-72.

Sutherland, G. R., and Hecht, F. (1985). Fragile Sites on Human Chromosomes. (New York: Oxford University Press), p. 53.

Tilley, W. D., Marcelli, M., Wilson, J. D., and McPhaul, M. J. (1990). Characterization and expression of a cDNA encoding the human androgen receptor. Proc. Natl. Acad. Sci. USA *86*, 327-331.

Thomas, P. S. (1980). Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA *77*, 5201-5205.

Van Ommen, G. J. B., and Verkerk, J. M. H. (1986). Restriction analysis of chromosomal DNA in a size range up to two million base pairs by pulsed-field gradient electrophoresis. In Human Genetic Diseases: A Practical Approach, K. E. Davies, ed. (Oxford: IRL Press), pp. 113-133.

Vincent, A., Heitz, D., Petit, C., Kretz, C., Oberle, I., and Mandel, J.-L. (1991) Abnormal pattern detected in fragile X patients by pulsed-field gel electrophoresis. Nature *349*, 624-626.

Warren, S. T., Zhang, F., Licamelli, G. R., and Peters, J. F. (1987) The fragile X site in

somatic cell hybrids: an approach for molecular cloning of fragile sites. Science 237, 420-423.

Warren, S. T., Zhang, F., Sutcliffe, J. S., and Peters, J. F. (1988) Strategy for the molecular cloning of the fragile X site DNA. Am. J. Med. Genet. *30*, 613-623.

Warren, S. T., Knight, S. J. L., Peters, J. F., Stayton, C. L., Consalez, G. G., and Zhang, F. (1990) Isolation of the human chromosomal band Xq28 within somatic cell hybrids by fragile X site breakage. Proc. Natl. Acad. Sci. USA *87*, 3856-3860.

Webb, T. P., Bundey, S. E., Thake, A. I., and Todd, J. (1986). Population incidence and segregation ratios in the Martin-Bell syndrome. Am. J. Med. Genet. 23, 573-580.

Wolf, S. F., Jolly, D. J., Lunnen, K. D., Axelman, J., and Migeon, B. R. (1984). Methylation of the hypoxanthine phosphoribosyltransferase locus on the human inactive X chromosome: implications for X chromosome inactivation. Proc. Natl. Acad. Sci. USA *81*, 2806-2810.

Yen, P. H., Patel, P., Chinault, A. C., Mohandas, T., and Shapiro, L. J. (1984). Differential methylation of hypoxanthine phosphoribosyltransferase genes on active and inactive human X chromosomes. Proc. Natl. Acad. Sci. USA *81*, 1759-1763.

Yu, S., Kremer, E., Pritchard, M., Lynch, M., Nancarrow, J., Baker, E., Holman, K., Mulley, J. C., Warren, S. T., Schlessinger, D., Sutherland, G. R., and Richards, R. I. (1991) The fragile X genotype is characterized by an unstable region of DNA. Science 252, 1179-1181.

PUBLICATION 7.2

THE LIMITED SIZE OF THE FRAGILE X SITE SHOWN BY FLUORESCENCE IN SITU HYBRIDIZATION

Annemieke J.M.H. Verkerk, Bert H.J. Eussen, Jan O. Van Hemel, Ben A. Oostra.

Department of Clinical Genetics, University Hospital Dijkzigt
(A.J.M.H.V., B.H.J.E., J.O.V.H.), Department of Cell Biology (B.A.O.),
Erasmus University, Rotterdam, The Netherlands.

Am. J. Med. Genet. 43, 187-191, 1992.

ABSTRACT

Cosmids, isolated from a 475 kb YAC that spans the fragile X region, and the YAC itself, were used for fluorescence *in situ* hybridization (FISH) on metaphase chromosomes from fragile X patients. Cosmid 22.3, containing most of the hybrid translocation breakpoints, shows *in situ* hybridization signals distal and proximal from the fragile X site. We propose that the size of the fragile site is limited to 20 kb.

KEY WORDS: Fragile X syndrome, fluorescence in situ hybridization, fragile site, X-linked mental retardation.

INTRODUCTION

The fragile X [fra(X)] syndrome is the most frequent form of inherited mental retardation and is associated with a fragile site, expressed as a isochromatid gap in the metaphase chromosome at map position Xq27.3 [Krawczun et al., 1985]. The fragile site is induced under specific cell culture conditions [Sutherland and Hecht, 1985] and is thought to be a local decondensation of the DNA. It was demonstrated that the translocations involving the fra(X) chromosome and rodent chromosomes were observed after induction of the fragile site. [Warren et al., 1987; 1990]. There was strong evidence that the human breakpoint of these translocations was non random within the fra(X) sequence. The isolation of fra(X) site DNA sequences was guided by the identification of the human/rodent breakpoints, genetic mapping and long range PFGE mapping that allowed the discovery of methylation anomalies. From yeast artificial chromosome (YAC) libraries several groups have isolated YACs that span the fragile X site [Heitz et al., 1991; Dietrich et al., 1991; Verkerk et al., 1991], resulting in the identification of the fragile site at molecular level. We have isolated a YAC (209G4) of 475 kb that spans the fra(X) region and contained the breakpoints of the hybrid cell lines [Verkerk et al., 1991]. The FMR-1 gene was identified within a four cosmid contig isolated from the YAC. A 5.2 kb EcoRI fragment was isolated from cosmid 22.3 that contains most of the hybrid translocation breakpoints, containing both the CpG island and sequences at the 5′ end of FMR-1, which exhibits length variation in patients, suggesting the association of the fra(X) site with FMR-1.

YAC 209G4 and several cosmids were used for fluorescence *in situ* hybridization on fra(X) induced chromosomes from different patients.

RESULTS AND DISCUSSION

Cytogenetically the fra(X) site can be induced in cultured cells by folate deprivation and is visible as a gap in the chromosome at Xq27.3, see Figure 1A, C and E. YAC 209G4 was used for FISH and is shown to be spanning the fragile site (Fig. 1B), with signals both proximal and distal from the gap. From this we conclude that the size of the fragile site must lie within 475 kb. From the YAC a cosmid library was constructed and a cosmid contig was isolated that spans the fragile X-related CpG island (Fig. 2) from BssHII site a (cosmid 22.3) through BssHII site c (cosmid 4.1). Cosmid 7.1 , which maps on the proximal SfiI fragment from YAC 209G4 (Fig. 2) lies minimal 190 kb proximal from the hybrid translocation breakpoints. After FISH, this cosmid was found to show signal proximal from the fragile X site (Fig. 1D). The opposite was found for cosmid 4.1, which is located about 40 kb distal from the hybrid translocation breakpoints and contains the 3' end of the FMR-1 gene [Verkerk et al., 1991], Fig. 2. This cosmid was found to always hybridize to the distal side of the fra(X) site (data not shown). Cosmid 22.3 (40 kb) contains the translocation breakpoints as well as BssHII site b (Fig. 2) that is hypermethylated in fra(X) chromosomes [Vincent et al., 1991; Bell et al., 1991] and a 5' exon of the FMR-1 gene including an unusual repetitive CGG sequence. *In situ* hybridization of cosmid 22.3 shows signals proximal and distal from the fragile site in 40% of the chromosomes (Fig. 1F). In total 40 fra(X) chromosomes were analyzed with cosmid 22.3. A single spot proximal to the fragile as well as in the fragile site was shown in 23% of the chromosomes. In 37% of the cases hybridization was found distal as well as in the fragile site. From these results we conclude that the minimal fragile site is localized in cosmid 22.3 and is less then 40 kb. Not in all cases hybridization was seen proximal or distal.

One explanation for this could be that the fragile site can sometimes extend in one or the other direction; another explanation is an incomplete *in situ* hybridization signal. If the *in situ* hybridization signal on either side of the gap is accounted for by 5-10 kb of DNA, then the gap can be as small as 20 kb. This is in agreement with the size of 10-100 kb as proposed by Heitz et al. [1991]. It is striking that overall after hybridization of cosmid 22.3 three signals are seen. The separation between two spots is too large to account only for the unfolding of the DNA. An

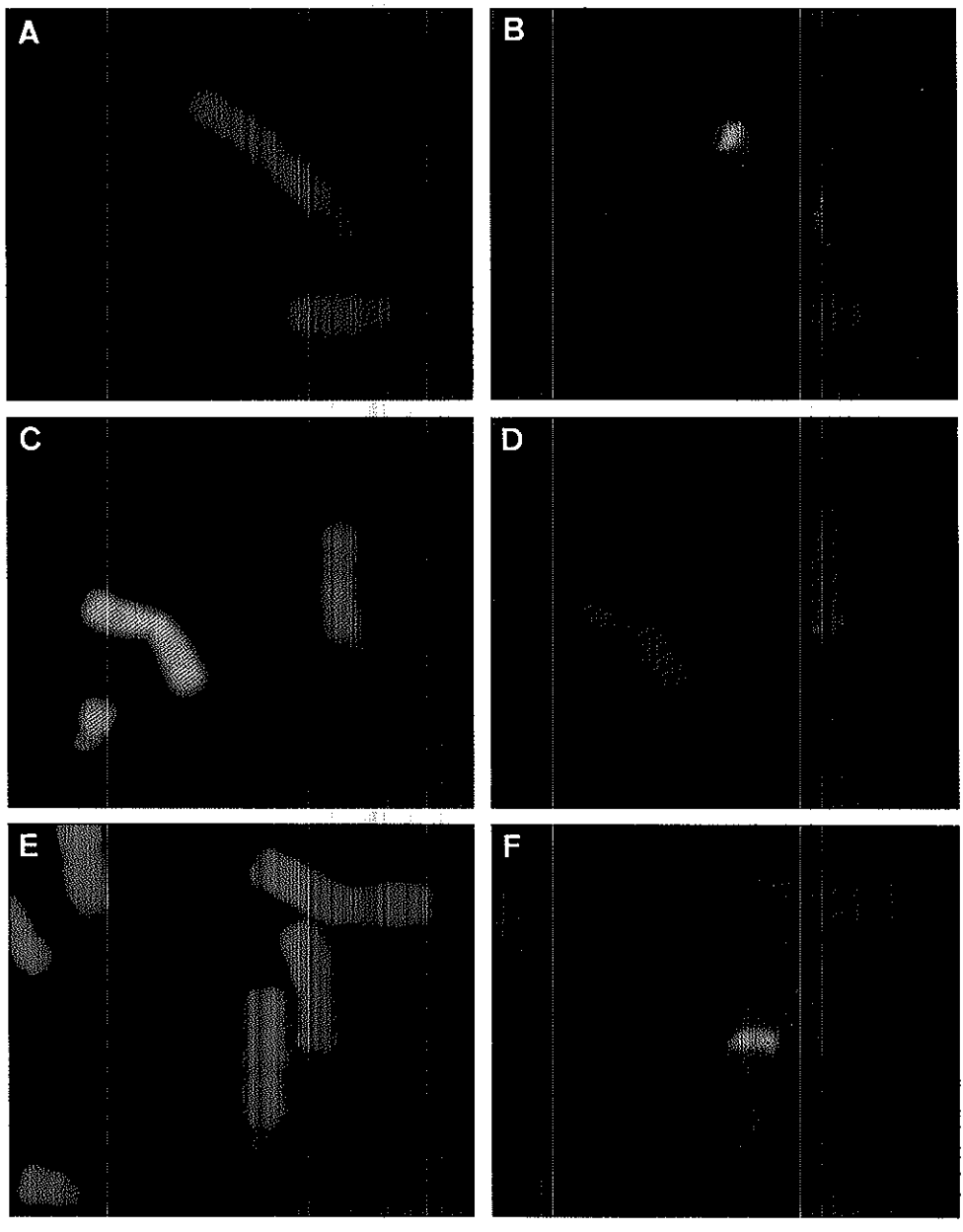


Figure 1. Fluorescence *in situ* hybridization.

In situ hybridization of YAC 209G4 (B), cosmid 7.1 (D) and cosmid 22.3 (F) to chromosome preparations of fra(X) patients. A centromere probe (pBamX5) specific for the X chromosome was also hybridized to confirm the identity of the chromosome examined. To show the fragile site the chromosomes were counterstained with DAPI (A,C,E).

explanation could be that in one of the chromatids a break has occurred, while the other chromatid is still intact. The region from cosmid 22.3 to 4.1 contains the FMR-1 gene [Verkerk et al., 1991]. It was found that the 3' end of the gene was located on cosmid 4.1, distal from the fragile X region. The 5' end of FMR-1 is located in the middle of cosmid 22.3, 250 bp away from the CpG island that is methylated in fragile X patients. This region is located in the gap (Fig. 1F). From this we conclude that the CpG island is localized in the middle of the gap from the fragile site itself. The differential methylation in this region together with the insertion in the first exon of the FMR-1 gene in fra(X) patients [Verkerk et al., 1991; 1992; Fu et al., 1991; Oberlé et al., 1991; Kremer at al., 1991] can be the cause of the chromosomal fragility. This is emphasized by the observation that most of the translocation breakpoints lie within this region. The first step in this mechanism would be a small increase in size in the gene. Normal transmitting males have a small insertion, do not show any methylation of the CpG island and cytogenetically no fra(X) site is seen. In patients however when there is a large insertion in the gene, the gene is methylated [Oberlé et al, 1991]. Partial

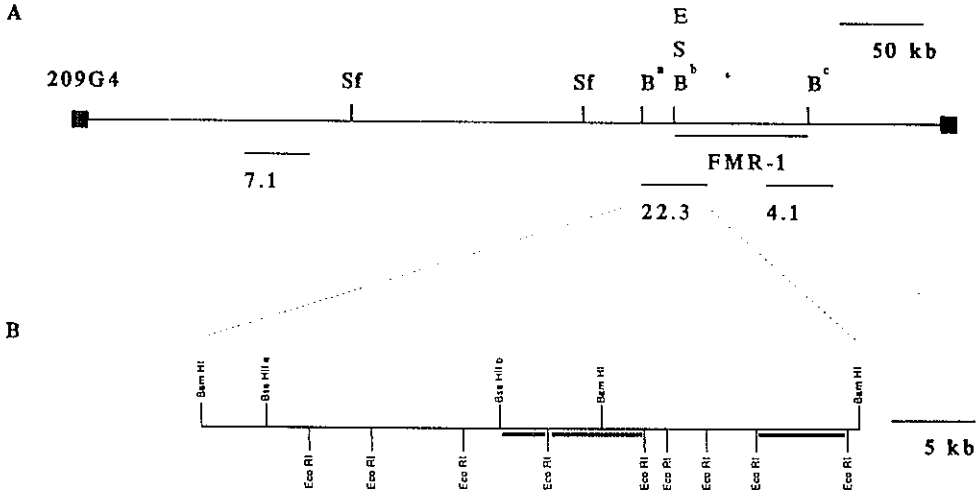


Figure 2. Physical maps.

A. Physical map of YAC 209G4 with restriction sites: B=BssHII; S=SacII; Sf=SfiI; E=EagI. The position of cosmids and the FMR-1 gene are shown. Closed boxes indicate YAC vector sequences.

B. Physical map of cosmid 22.3. Solid lines below the map show fragments containing exon sequences. The fragment between BssHII site b and the proximal EcoRI site contains the 5' exon of FMR-1, as well as fra(X) associated translocation breakpoints [Verkerk et al., 1991].

methylation almost always correspond to the presence of both large and small insertion within the cell population. Both the nature of the insertion as well as the methylation together could be the cause of the fragility.

MATERIALS AND METHODS

Fluorescence In Situ Hybridization

In situ hybridizations of total yeast DNA and cosmids were performed according to Wada et al. [1990] and Kievits et al. [1990], respectively. Fra(X) expression was induced by methotrexate (10 µg/ml)) (Lederle), which was supplemented for the last 24 hours to phytohaemagglutinin stimulated lymphocyte cultures (medium 199, Gibco, 10% FCS) of a fra(X) patient. After fixation the suspensions were spotted on slides, and air-dried overnight. The slides were washed with PBS, RNase (100 µg/ml)/2 *SSC incubated for 1 h at 37°C, digested for 10 min with pepsin (Serva) (10 mg/100 ml)/0.01 N HCL at 37°C, post-fixed in 1% formaldehyde (Merck)/ PBS/50 mM $MgCl_2$ and dehydrated in cold ethanol.

The biotinylated total yeast and cosmid DNA were preannealled for 1-4 h in presence of 50 times excess of sonicated total human genomic DNA to reduce the background for repetitive sequences. The overnight hybridization of total yeast (150 ng) and cosmid (40 ng) DNA was performed in 10 µl 50% formamide, 2x SSC, 10% Dextran sulfate solution under a 18*18 mm coverslip sealed with rubber. In the experiments 2 ng/µl pBamX5, a human repetitive sequences hybridizing mainly to pericentromeric region of the X chromosome [Willard et al., 1983], was separately denaturated and added to the hybridization solution.

After amplification of the signal by two layers of avidin-FITC (Vector), and one layer of biotinylated goat anti-avidin (Vector) the samples were washed with PBS and mounted in propidium iodide/DAPI (0.03 µg/ml/0.6 µg/ml) antifade (2% DABCO/glycerol) medium (Sigma).

Microscopic analysis was performed with a Leitz Aristoplan microscope equiped for FITC (K3 block) and DAPI (A block) detection. The results were photographed on Kodak Ektachrome 400 (Slides) daylight film.

DNA Probes

YAC 209G4 is a 475 kb YAC isolated from the CEPH YAC library [Albertsen et al., 1990; Verkerk et al., 1991]. Cosmids were isolated from a cosmid library prepared from YAC 209G4 [Verkerk et al., 1991].

ACKNOWLEDGMENTS

We thank T. Kievits and J.G. Dauwerse for help in establishing the FISH techniques; K.E. Bakker-van Kempen and R.M. v.d. Helm for expert technical assistence. We thank Mrs. J. van Deursen for her secretarial assistence. B.A.O. was supported by a grant from MGC.
Received for publication October 15, 1991; revision received Januari 2, 1992.

REFERENCES

Albertsen HM, Abderrahim H, Cann HM, Dauset J, Le Paslier D, Cohen D (1990): Construction and characterization of a yeast artificial chromosome library containing seven haploid genome equivalents. Proc. Natl. Acad. Sci. USA 87:4256-4260.

Bell MV, Hirst MC, Nakahori Y, MacKinnon RN, Roche A, Flint TJ, Tommerup N, Tranebjaerg L, Froster-Iskenius U, Kerr B, Turner G, Lindebaum D, Winter R, Pembrey M, Thibodeau S, Davies KE (1991): Physical mapping across the fragile X: Hypermethyl-ation and clinical expression of the fragile X syndrome. Cell 64:861-866.

Dietrich A, Kioschis P, Monaco AP, Gross B, Korn B, Williams SV, Sheer D, Heitz D, Oberlé I, Toniolo D, Warren ST, Lehrach H, Poustka A (1991): Molecular cloning and analysis of the fragile X region in man. Nucleic Acids Res 19:2567-2572.

Fu Y-H, Kuhl DPA, Pizzutti 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.

Heitz D, Rousseau F, Devys D, Saccone S, Abderrahim H, Le Paslier D, Cohen D, Vincent A, Toniolo D, Della Valle G, Johson S, Schlessinger D, Oberlé I, Mandel JL (1991): Isolation of sequences that span the fragile X and identification of a fragile X-related CpG island. Science 251:1236-1239.

Kievits T, Dauwerse JG, Wiegant J, Devilee P, Breuning MH, Cornelisse CJ, van Ommen G-JB, Pearson PL (1990): Rapid subchromosomal localization of cosmids. Cytogenet Cell Genet 53:134-136.

Krawczun MS, Jenkins EC, Brown WT (1985): Analysis of the fragile-X chromosome: Localisation and detection of the fragile site in high resolution preparations. Hum Genet 69:209-211.

Kremer EJ, Pritchard M, Lynch M, Yu S, Holman K, Baker E, Warren ST, Schlessinger D, Sutherland GR, Richards RI (1991): Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. Science 252:1711-1714.

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

Sutherland GR, Hecht F (1985): "Fragile Sites on Human Chromosomes". New York: Oxford University Press, pp 53.

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

Verkerk AJHM, de Vries BBA, Niermeijer MF, Fu Y-H, Nelson DL, Warren ST, Majoor-Krakauer DF, Halley DJJ, Oostra BA (1992): An intragenic probe used for diagnostics in fragile X families. Am. J. Med. Genet. 43:192-196.

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

Wada M, Little RD, Abidi F, Porta G, Labella T, Cooper T, Valle GD, D'Urso M, Schlessinger D (1990): Human Xq24-Xq28: Approaches to mapping with yeast artificial chromosomes. Am J Hum Genet 46:95-106.

Warren ST, Zhang F, Licamelli GR, Peters JF (1987): The fragile X site in somatic cell hybrids: an approach for molecular cloning of fragile sites. Science 237:420-423.

Warren ST, Knight SJL, Peters JF, Stayton CL, Consalez GG, Zhang F (1990): Isolation of the human chromosome band Xq28 with somatic cell hybrids by fragile X breakage. Proc Natl Acad Sci USA 87:3856-3860.

Willard HF, Smith KD, Sutherland J (1983): Isolation and characterization of a major tandem repeat family from the human X chromosome. Nucleic Acids Res 11:2017-2033.

PUBLICATION 7.3

AN INTRAGENIC PROBE USED FOR DIAGNOSTICS IN FRAGILE X FAMILIES

Annemieke J.M.H. Verkerk, Bert B.A. de Vries, Martinus F. Niermeijer, Ying-Hui Fu, David L. Nelson, Stephen T. Warren, Danielle F. Majoor-Krakauer, Dicky J.J. Halley, Ben A. Oostra.

Department of Clinical Genetics, University Hospital Dijkzigt (A.J.M.H.V., B.B.A.dV, M.F.N., D.F.M-K., D.J.J.H.), Department of Cell Biology (B.A.O), Erasmus University, Rotterdam, The Netherlands; Institute for Molecular Genetics (Y-H.F., D.L.N), Baylor College of Medicine, Houston, Texas; Howard Hughes Medical Institute and Departments of Biochemistry and Pediatrics (S.T.W.), Emory University School of Medicine, Atlanta, Georgia.

Am. J. Med. Genet. 43, 192-196, 1992.

ABSTRACT

The intragenic (FMR-1) probe pE5.1 was used for DNA analysis in fragile X families. With this probe fragments of altered size can be detected in female carriers, affected individuals and transmitting males. The length of the altered fragments was found to vary from one generation to another as well as between sibs. This instability of the DNA detected by pE5.1 was also seen in peripheral blood within single individuals. These phenomena are illustrated by 4 exemplary families segregating the fragile X syndrome. We demonstrate the diagnostic contribution of intragenic analysis to carrier detection as well as the identification of normal transmitting males carrying premutations. One of the families illustrates the passage of a premutation to a male through 2 generations.

KEY WORDS: Fragile X syndrome, carrier detection, repeat sequence, FMR-1, DNA diagnostic, X-linked mental retardation.

INTRODUCTION

We have recently described the isolation of the FMR-1 gene at the fragile X [fra(X)] locus, which may be directly involved in the manifestation of the fra(X) syndrome [Verkerk et al., 1991]. A 5.2 kb EcoRI fragment was isolated, containing the majority of breakpoints in hybrid cell lines at the fra(X) site [Warren et al., 1990], a CpG island that is preferentially methylated in fra(X) patients [Vincent et al., 1991; Bell et al., 1991], and an exon of the FMR-1 gene [Verkerk et al., 1991] including an unusual repetitive CGG sequence. The CGG repeat exhibits polymorphic length variation in the normal population. The number of CGG repeats ranged from 6-46, with an average of 29, in more than 200 individuals [Fu et al., 1991]. In fra(X) patients the 5.2 kb EcoRI fragment is considerably increased in size and shows somatic length variation.No difference was found in sequences flanking the CGG repeat region in carriers and patients, placing the source of the length variation within the CGG repeat itself [Kremer et al., 1991, Fu et al., 1991].

The diagnosis of the fra(X) syndrome has been based on cytogenetic testing of the presence of a fragile site at Xq27.3. Cytogenetics has serious limitations in carrier detection because 50% of obligate carriers fail to express the fragile site. In addition, the fragile site normally is not observed in normal transmitting males [Sherman et al., 1985].

The possibilities for carrier detection and prenatal diagnosis have been extended by DNA analysis, using closely linked polymorphic markers flanking the fra(X) locus [Suthers et al., 1991]. However, a widespread application of polymorphic markers was hampered by limited informativeness. Using the 5.2 kb EcoRI fragment as a probe in diagnosis is a major improvement [Oberlé et al.,1991; Kremer et al., 1991; Pieretti et al., 1991]. Carriers and patients now can be detected by showing the presence of fragments of abnormal size.

In this paper we describe the analysis of four fra(X) families using the 5.2 kb EcoRI fragment (pE5.1) as a probe. We detected female carriers and distinguished between normal, affected and transmitting males. The diagnostic potential of intragenic DNA analysis is discussed.

RESULTS AND DISCUSSION

A family with 2 affected brothers, A3 and A6 is shown in Figure 1A. They showed cytogenetic fra(X) expression in 4% and 14% of their cells, respectively. Abnormal fragments appeared in EcoRI digests of both patients. Instead of discrete bands, smears of DNA fragments were observed, indicating somatic mosaicism: different cells may have fragments with different increased sizes in the CGG repeat sequence. A band of normal size was observed in their sister A5 as well as a band with an increased fragment length, although its size is smaller than is seen in her brothers. Besides a normal band, a band that is increased in size, but that is smaller than in A5, was found in the mother A1.

In this family it is clear that the fragment increases in size from the mother A1 to her children A5,3,6. Daughter A5 inherits the fra(X) allele from her mother and is therefore a carrier of the syndrome. Daughter A2 has inherited the normal allele and thus is not a carrier.

This phenomenon of a fragment increasing in size is often observed in fra(X) families. It is not clear how and when this takes place. The most likely explanation would be that the initial increase in size is produced during oogenesis. The germ cell of the mother carries one abnormal fragment, but to account for the mosaicism in the patients it has to be assumed that during embryogenesis the CGG repeat "grows" in length. Clearly there is a great variability in the size in the CGG repeat that is passed on from the mother to her different children.

Figure 1B shows an example of a family where the mutation changes in size through 3 generations. None of the 3 women in this pedigree showed any cytogenetic expression of the fragile site. Marker analysis had not been conclusive to establish the carrier status of B5. However, a clearcut result was obtained by Southern blot analysis using probe pE5.1. The grandmother B1 shows a fragment with a very small increase in the CGG repeat that barely resolves from the normal allele on the Southern blot. By PCR analysis using primers flanking the CGG repeat [Fu et al., 1991] it was confirmed that B1 had an increase in size of 108 bp [Nelson, unpublished results]. In her daughter B3 the fragment became larger, continuing its growth in the granddaughter B5 and her patient brother B4, who had 20% of cytogenetic fra(X) expression.

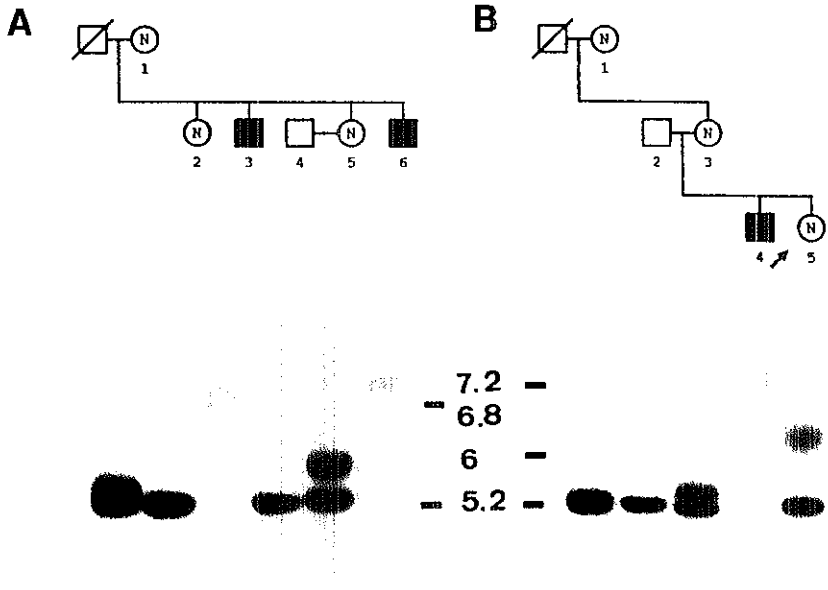


Figure 1. Pedigrees of fra(X) families.

The following symbols are used: Unaffected ○ female and □ male;
Normal Ⓝ female and 🄽 male, without cytogenetic expression of the fragile X site;
■ Mentally retarded fragile X male, with cytogenetic expression of the fragile X site;
◑ Female with cytogenetic expression of the fragile X site.

In the pedigree in Figure 2A the individual A3 is a fra(X) patient. He has low (4%) cytogenetic fra(X) expression and is mentally retarded. On a Southern blot analysis he shows a band of near normal size and in addition a smear with a population of larger fragments. This type of mosaicism, first observed by Oberlé et al. [1991], is different from the one observed in the patients A3 and A6 in family 1A. Other members in this pedigree were tested. The woman A6 shows a normal allele as well as one that is increased in size. This latter allele is passed on to her son A7. This fragment which is characterized by a small increase in size, is associated with lack of clinical and cytogenetic expression, which marks A7 as a normal transmitting male. A7 is related to A3 through their grandfathers, which strongly suggests that they too were normal transmitting males. This was supported by pE5.1 probing of DNA from their spouses, who were shown to have normal 5.2 kb alleles (A5, data not shown).

Transmitting males, who have no cytogenetic or clinical expression can be detected by using probe pE5.1. They show fragments slightly increased in size. When this

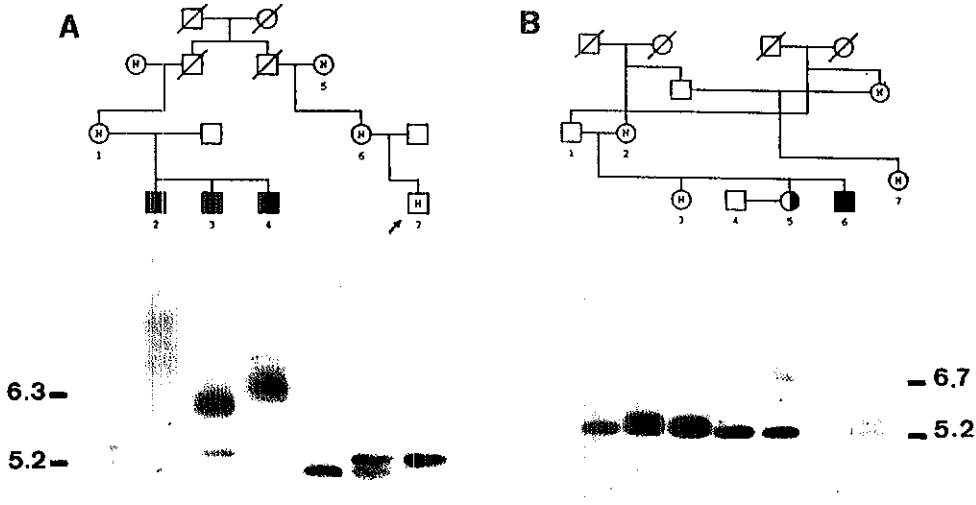


Figure 2. Pedigrees of Fragile X families. For explanation of the symbols, see legend Figure 1.

fragment is passed on to their daughters, often a change is seen in the size, although it stays in the premutation range, with fragment sizes increased roughly between 150 and 300 bp [Oberlé et al., 1991; Fu et al., 1991]. Patients in the next generation show much larger fragment lengths. Apparently, the increase in size in the CGG repeat has to pass a certain critical size to result in the fra(X) phenotype. When this limit is reached is not clear yet, but it is thought to be around a size of 500 bp [Oberlé et al., 1991].

In Figure 2B, a pedigree is shown which illustrates that passage of the fra(X) allele from one generation to another is not always accompanied by an increase of the fragment length. In the woman B2, besides a normal allele, an allele with an increased size is observed. Two of her children, B5 and B6 have cytogenetic fra(X) expression and both show smears composed of different enlarged fragments. In the daughter B3 a clearly visible increased fragment is seen that is, however, smaller than in her mother. In this case we find a reduction in the fragment size. In the woman B7 also an allele of increased size is found, indicating her as a carrier.

Our analysis shows that the detection of abnormal fragments using pE5.1 is a powerful tool in fra(X) diagnosis. Abnormal patterns can be seen in normal transmitting males, female carriers and patients.

In addition, if there is doubt after Southern analysis about the presence of a fragment with a very small increase in size, amplification by PCR of the CGG repeat should give a clearcut answer [Fu et al., 1991]. It should be pointed out, that no amplification by PCR is found if the increase in the CGG repeat is too large. Using both techniques reliable genotyping is possible resulting in very much improved genetic counseling in families affected by fra(X) syndrome.

MATERIALS AND METHODS

The DNA probe used was pE5.1 [Verkerk et al., 1991], which was later determined to be 5222 bp long [Fu et al., 1991].

Total genomic DNA was isolated from leukocytes from individuals from different Fra(X) families as described [Miller et al., 1988]. DNA (8 µg) was digested to completion with the restriction endonuclease EcoRI, separated by gel electrophoresis and subjected to Southern blot analysis. The probe was labeled by the random oligonucleotide priming method

[Feinberg and Vogelstein, 1983] and was competed for 2 hours with 120 µg of total human DNA in 5 x SSC at 65°C. After prehybridization and hybridization the filters were washed to 0.1 x SSC at 65°C prior to autoradiography.

ACKNOWLEDGMENTS

We thank M.N. van der Est and W.H. Deelen for excellent technical assistance and Dr. M.L. Giovannucci Uzielli from Florence, Italy, for kindly providing family material. We thank Mrs. J. van Deursen for her secretarial assistence in the preparation of the manuscript. S.T.W. is an investigator of the Howard Hughes Medical Institute.

Received for publication October 15, 1991; revision received January 2, 1992.

REFERENCES

Bell MV, Hirst MC, Nakahori Y, MacKinnon RN, Roche A, Flint TJ, Tommerup N, Tranebjaerg L, Froster-Iskenius U, Kerr B, Turner G, Lindebaum D, Winter R, Pembrey M, Thibodeau S, Davies KE (1991): Physical mapping across the fragile X: Hypermethyl-ation and clinical expression of the fragile X syndrome. Cell 64:861-866.

Feinberg AP, Vogelstein B (1983): A technique for radiolabeling DNA restriction endo-nuclease fragments to high specific activity. Analyt Biochem 132: 6-13.

Fu Y-H, Kuhl DPA, Pizzutti A, Pieretti M, Sutcliffe JS, Richards S, Verkerk AJMH, Holden JJA, Fenwick RG, Warren ST, Oostra BA, Nelson DL, Caskey CT (1991): Fragile X site: A polymorphic and highly mutable CGG repeat in the FMR-1 gene. Cell 67, 1047-1058.

Kremer EJ, Pritchard M, Lynch M, Yu S, Holman K, Baker E, Warren ST, Schlessinger D, Sutherland GR, Richards RI (1991): Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. Science 252:1711-1714.

Miller SA, Dykes DD, Polesky HF (1988): A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16:1214.

Nelson DL (1991): Unpublished results.

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

Pieretti M, Zhang F, Fu Y-H, 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.

Sherman SL, Jacobs PA, Morton NE, Froster-Iskenius U, Howard-Peebles PN, Brøndum-Nielsen K, 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.

Suthers GK, Mulley JC, Voelckel MA, Dahl N, Vaisanen ML, Steinbach P, Glass IA, Schwartz CE, van Oost BA, Thibodeau SN, Haites NE, Oostra BA, Giné R, Carballo M, Morris CP, Hopwood JJ, Sutherland GR (1991): Genetic mapping of new DNA probes at Xq27 defines a strategy for DNA studies in the fragile X syndrome. Am J Hum Genet 48:460-467.

Verkerk AJMH, Pieretti M, Sutcliffe JS, Fu Y, 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, Oostra BA, Warren ST (1991): Identification of a gene (FMR-1) containing a CGG repeat coincident with a fragile X breakpoint cluster region exhibiting length variation in fragile X syndrome. Cell 65:905-914.

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

Warren ST, Knight SJL, Peters JF, Stayton CL, Consalez GG, Zhang F (1990): Isolation of the human chromosome band Xq28 with somatic cell hybrids by fragile X breakage. Proc Natl Acad Sci USA 87:3856-3860.

PUBLICATION 7.4

ALTERNATIVE SPLICING IN THE FRAGILE X GENE *FMR1*

Annemieke J.M.H. Verkerk, Esther de Graaff, Kristel De Boulle*, Evan E. Eichler#, David S. Konecki@, Edwin Reyniers*, Antonella Manca@, Annemarie Poustka@, Patrick J. Willems*, David L. Nelson#, Ben A. Oostra.

Department of Clinical Genetics, Erasmus University, Dr Molewaterplein 50, 3153 DR Rotterdam, The Netherlands.

* Department of Medical Genetics, University of Antwerp-UIA, 2610 Antwerp, Belgium.

\# Institute for Molecular Genetics, Baylor College of Medicine, Houston, TX 77030, USA.

@ Deutsches Krebsforschungszentrum, Heidelberg, Germany.

Hum. Mol. Genet. 2, 399-404, 1993.

ABSTRACT

The *FMR1* gene, associated with fragile X syndrome, has recently been cloned and the sequence of partial cDNA clones is known. We have determined additional cDNA sequences both at the 5′ and 3′ end. We have characterized the expressed gene by means of RT-PCR in various tissues and have found that alternative splicing takes place in the *FMR1* gene, which does not seem to be tissue specific. When the different alternative splicing events are combined, 12 distinct mRNA products could result from *FMR1* expression in each tested tissue. In all these transcripts the open reading frame is maintained until the same stop codon. At the 3′ end alternative use of polyadenylation signals is found. The alternative splicing allows functional diversity of the FMR-1 gene. Whether all the possible proteins will be synthesized and whether they will be functionally active has to be determined.

INTRODUCTION

Fragile X syndrome is an X-linked heritable disease and is the most common cause of inherited mental retardation, with an estimated prevalence of 1 in 1250 males (1, 2). Two important clinical features, mental retardation and macroorchidism, indicate that brain and testis are affected tissues. The recent isolation of the fragile X gene *FMR1* (3) revealed a CGG repeat at the 5′ end of the gene. Variation in the length of this repeat comprises the majority of mutations leading to fragile X syndrome. Methylation of a CpG island (4, 5, 6) that is located 250 bp upstream of the CGG repeat correlates with loss of expression of the *FMR1* gene in patients (7). In the normal population the CGG repeat is polymorphic and varies in length from 6 - 50 repeats (8). In phenotypically normal transmitting males the extended repeat number ranges from 52 up to 200 (premutation). In most fragile X patients the CGG repeat is significantly increased in length to more than 200 repeats (full mutation) (6, 8, 9, 10). In addition to the full mutation, a premutation is found in about 20% of male patients (7). In most of the cells in these mosaic patients a full mutation is found and in these cells there is no mRNA or protein production. The premutation alleles in the other cells are transcribed, resulting in protein. Overall, there is an apparent insufficiency of protein production in the appropriate tissues resulting in the abnormal phenotype.

A few patients lacking expansion of the CGG repeat have been described with the fragile X phenotype. In one patient the entire *FMR1* gene was deleted, as well as

approximately 2,5 megabases of flanking sequence (11). In a second patient at least five exons of the *FMR1* gene were missing (12). Recently, transcription of the *FMR1* gene has been described in a patient with the fragile X phenotype without CGG expansion or cytogenetic expression. In this patient the phenotype was the result of a single point mutation (13). An A → T mutation was found in the open reading frame (ORF) resulting in an Ile → Asn substitution in the *FMR1* gene of this patient. In order to facilitate screening for *FMR1* mutations in patients with the fragile X phenotype without CGG expansion, RT-PCR studies of transcription of the normal *FMR1* gene in different tissues were performed. This revealed several alternatively spliced products of the gene that could give rise to as many as 12 different mRNAs and to 12 possible proteins.

RESULTS

Transcription of the *FMR1* gene

Transcription of the *FMR1* gene was studied by RT-PCR. Total RNA was isolated from HEPG2 cells and control EBV transformed human lymphocytes. After first strand cDNA synthesis, PCR was performed with 4 partially overlapping primer sets, K1-K2, K3-K4, K5-K6 and K7-K8 (for position of the primers see Material and Methods and fig. 3). PCR reactions with primer sets K1-K2 and K3-K4 each showed one discrete band on ethidium bromide stained agarose gels. The size of the PCR products was in agreement with the size expected from the cDNA sequence (3) (Fig. 1, lanes 1 and 2). However, PCR with primer set K5-K6 showed 2 bands (Fig. 1, lanes 3), whereas 4 bands were visible with primer set K7-K8 (Fig. 1, lanes 4), indicating the possibility of several alternative splice sites in the *FMR1* gene.

To test whether the alternative splicing would be tissue specific, RT-PCR products of various tissues were analyzed, including testis and brain as these two tissues demonstrate the most interesting clinical features (mental retardation and macroorchidism) of the fragile X syndrome. Using primer set K5-K6 2 bands were seen in all tissues analyzed: adult human testis and EBV transformed lymphocytes, fetal pancreas, liver and brain and adult mouse brain (Fig. 2a). Sequencing both bands revealed an additional 63 basepairs between positions 1126 - 1188 in the upper RT-PCR product (Fig. 2a: upperband and Fig. 3: B). The originally published human cDNA did not contain this 63 bp sequence and corresponds to the lower band. In mouse this 63 bp sequence is found as well (Warren, Nelson, Bakker, unpublished results). The ratios of the upper to the lower band were about 1:4 in

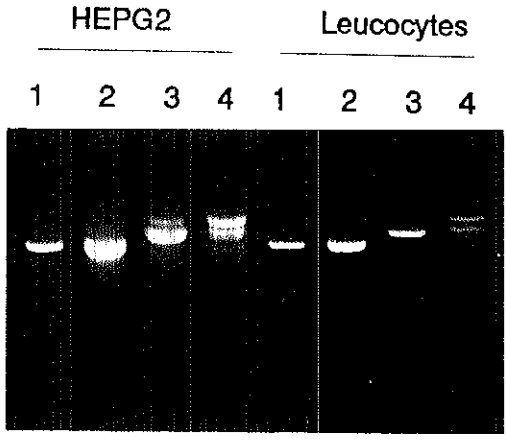


Figure 1. RT-PCR products of HEPG2 cells and EBV transformed lymphocytes using 4 overlapping primer sets covering the ORF of the *FMR1* gene, were separated on 3% ethidium bromide stained agarose gels. Lane 1 to 4 show RT-PCR products with primer sets K1-K2, K3-K4, K5-K6 and K7-K8, respectively.

all tissues analyzed (Fig. 2a).

Sequencing the K5-K6 PCR products also revealed 6 basepairs that were not present in the published *FMR1* cDNA (3). These 6 bases, GCAAAG, are positioned between base 881 - 886 (Fig. 3: A) and were consistently present in all mRNA products and cDNA clones analyzed. The fact that all sequences analyzed contain the 6 bases makes it likely that the absence of this sequence in the original BC22 cDNA clone could have been due to a splicing artefact. Analysis of these tissues with primer set K7-K8 showed 4 bands in all tissues tested, except in testis where a fifth band was found (Fig. 2b). The upper band (band 1) was missing in mouse brain. Sequencing band 1 disclosed 51 additional bases compared to the published cDNA sequence, positioned between bases 1738 - 1788 (Fig. 3: D). These 51 basepairs were also found in band 3. Sequence data also showed that band 3 and 4 (Fig. 2b) were missing 75 bases positioned from base 1472 - 1546 (Fig. 3: C), as compared to the published cDNA sequence. Direct sequencing of the RT-PCR product in band 2 from fetal brain and EBV transformed lymphocytes showed overlapping sequences, indicating more than one fragment in this band. We presumed that as in testis material, this band consisted of 2 different PCR products (band 2 → a and b) that were not separated due to relatively high amounts of these 2 products with a minor size difference. To investigate this further primer A3 (see Fig. 3 and Material and Methods) was used in combination with K7. RT-PCR with

this primer set excluded the 51 bases. Only the area around the 75 basepairs was amplified and this gave 3 well separated bands in testis, fetal brain (Fig. 4), EBV transformed lymphocytes and mouse (data not shown). Sequencing the upper product identified a sequence of 242 bases in complete alignment with the published cDNA sequence. The lower band of 167 bp was missing the above

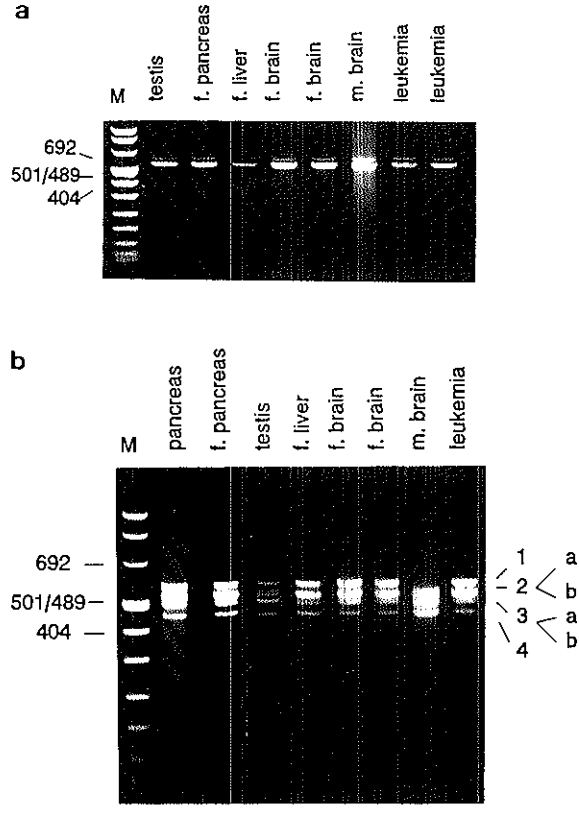


Figure 2. RT-PCR products from different tissues, using primer set K5-K6 (**a**) and primerset K7-K8 (**b**); f=fetal tissue, m=mouse tissue. Fetal brain was from a 16 and a 17 week old fetus, respectively. In Figure 2b the bands contain: band 1: +75 bp, +51 bp = C and D in Fig. 3; band 2a: +39 bp, +51 bp = C (last 39 bp) and D in Fig. 3; band 2b: +75 bp = C in Fig. 3, -51 bp; band 3a: -75 bp, +51 = D in Fig. 3; band 3b: +39 bp = C (last 39 bp) in Fig. 3, -51 bp; band 4: -75 bp, -51 bp. The original published BC22 cDNA clone corresponds to band 2b.

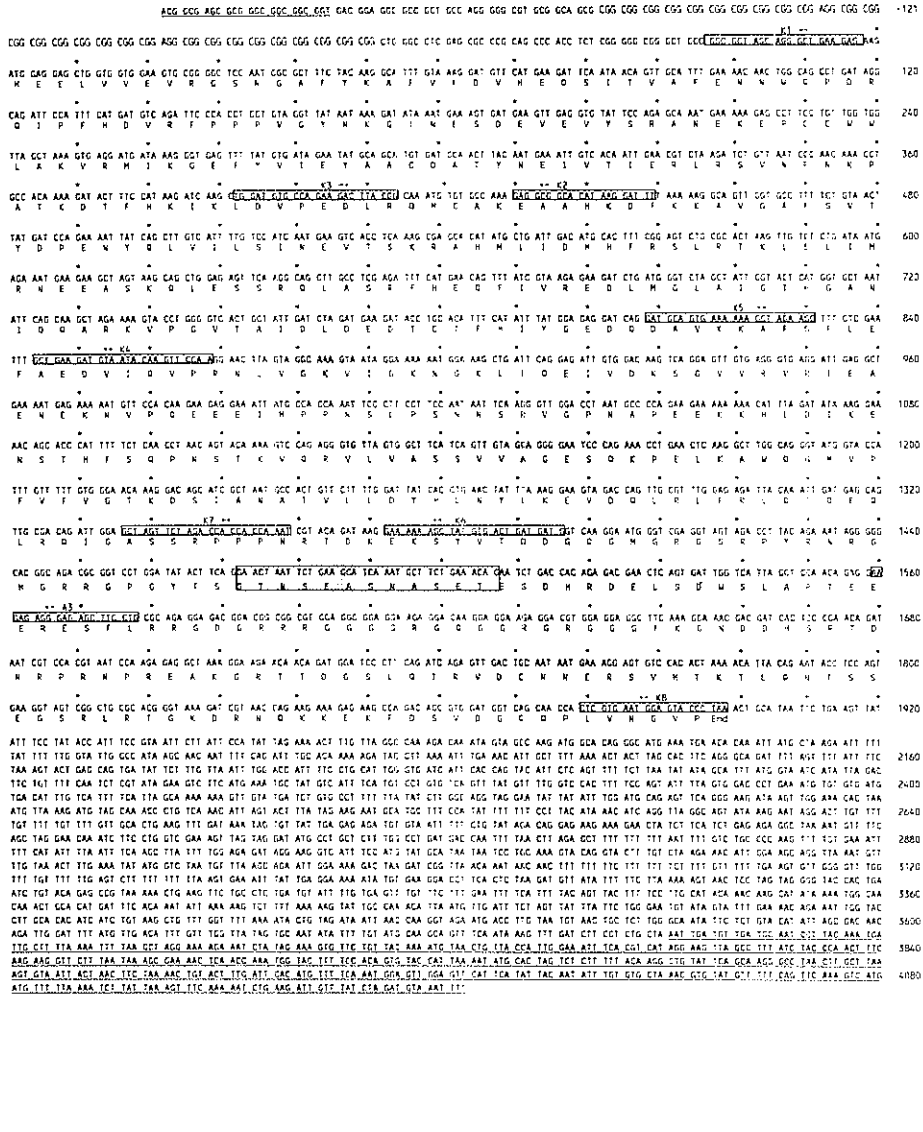


Figure 3. *FMR1* cDNA sequence. Additional (exon) sequences are integrated in the sequence of the original BC22/BC72 cDNA clones (3). Sequences that undergo alternative splicing are indicated in grey and are indicated A, B, C and D in the left margin. Additional sequences (underlined) have been added at the 5′ and 3′ end of the *FMR1* cDNA. The location of the primers is given and the sequences are contained in boxes. The sequence has been renumbered as compared to the prior published sequence (3) and the numbering starts at the first ATG 3′ from the CGG repeat.

mentioned 75 bp. This is consistent with data found with primer set K7-K8. However, the middle band contained only part (39 bases present) of these 75 bases, with 36 bp missing from the 5' end, indicating another splice site situated within the 75 basepairs (Fig. 3: C). Using primers K7 and K8, a sixth band could be visualised after extended running time (Fig. 2b: band 3b). According to the size of this product, this band contains the 39 bases without the 51 bases. When using primers that select for the presence or absence of the 51 basepairs, in mouse we only found bands indicating the absence of this sequence. In human tissues there is a 1:1 ratio for the presence or absence of splice site D (data not shown). In tabel 1 the sequences that are present in the alternatively spliced regions C and D found with primers K7 and K8 as well as their relative abundance are summerized.
In figure 3 the above mentioned differential sequences are visualised as grey areas

Table 1. Sequences present in the alternatively spliced regions C and D found with primers K7/K8 and the relative abundance of the different transcripts.

Product	Bases present at C	Bases present at D	Total bases present	Relative abundance		
				Human somatic	Human testis	Mouse brain
1	+75	+51	126	+++	+	0
2a	+39	+51	90	++	+	0
2b	+75	0	75	+++	+	++++
3a	0	+51	51	++	+	0
3b	+39	0	39	+	+/-	++
4	0	0	0	+	+	++

The relative abundance of the different transcripts was determined by comparing the ratios of the different PCR products obtained with different primersets in the alternatively spliced regions C and D.

in the original cDNA. The individual differential sequences B,C and D were foundto be present in different cDNA clones. In all cases the open reading frame is maintained until the original stop codon.

Additional sequence from fetal brain and liver FMR1 cDNA clones

New *FMR1* clones were isolated from human fetal brain and liver cDNA libraries to extend the sequence of the original published *FMR1* cDNA at the 5′ end. Several clones that were longer at the 5′ end, contained the same 24 basepairs at the 5′ end preceding the original sequence (underlined in Figure 3 at the 5′ end) which is in agreement with the genomic sequence published by Fu et al (8). In front of this (underlined) shared sequence, all the clones show a different sequences.

The original BC72/22 clone was isolated from fetal brain and contains a polyadenylation signal and the first A of a poly A tail. Several clones isolated from a liver cDNA library show the same 3′ end, but a number extended beyond the published sequence at the 3′ end (underlined in Fig. 3 at the 3′ end). The original 3′ sequence and the extended sequence are adjacent in the genomic sequence.

Intron-exon boundaries

To determine the nature of the alternative splicing, the relevant splice boundaries were analyzed. Sequences of the intron-exon boundaries of the different splice sites were determined by comparing genomic and cDNA sequences (Figure 5). The sequence of 63 bases (B in Figure 5) is a complete exon that has intron sequences

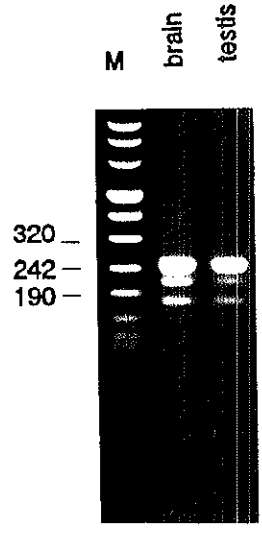


Figure 4. RT-PCR products from fetal brain and testis, using primers K7 and A3. For EBV transformed lymphocytes and mouse brain the same bands were found in the same ratios (data not shown).

on either side in genomic DNA. Intron sequences are also found adjacent to the other 5′ splice sites C and D and to site A (Figure 5). However, at the 3′ splice boundaries of these sites, splicing occurs inside exon sequences (Figure 5, C, D and A). Splice junction sites in higher eukaryotes are characterized by a short but well conserved consensus sequence. By comparing the actual splice sites in a number of known gene sequences, Shapiro and Senapathy (14) have found that splice sites can be identified with a great degree of accuracy in different species. The most frequently occurring consensus sequence found in primate mRNAs at the 5′ junction site (CAG/gta) is present in all above described alternative splice sites in *FMR1*, except in the first 5′ junction site of splice site B (Fig. 5) where AGG/gta is found. All introns preceding A, B, C and D start with gt, which is found at 5′ junctions of all introns studied in different organisms, including primates, plants and invertebrae. For splice site D the 5′ intron sequence was not determined. All 3′ ends of the introns of the *FMR1* splice sites of A, B, C and D conform to the ag consensus sequence (tccag/G) (14).

It is striking that 3 of the 3′ acceptor sites (Figure 5 and 6, C and D) are in the

	5′ splice site	3′ splice site
A	TTAGTAGgtaagtgcaga	atttcttagGCAAAG ttagcaaagTAATAG
B	CCAGAGGgtaagaattac TTGGCAGgtaggaaaaca	actgcttagGTGTTA ctttttcagGGTATG
C	ACTTCAGgtacanctaag	ccncgatagGAACTA ctgaaacagAATCTG ggtcattagCTCCAA
D	CCTTCAG.....	ttgttttagATCAGA acattacagAATACC
consensus	AGgt	agG

Figure 5. Intron/exon boundaries of the alternative splice sites. A, B, C and D give the intron/exon sequences for the 6, 63, 75, and 51 basepairs, respectively. Capitals indicate exon sequences, lower case letters (not underlined) indicate intron sequences. Underlined lower case letters indicate exon sequences that are (alternatively) spliced out of the mRNA.

middle of exon sequences. The AG consensus sequence is found to directly precede the 3' acceptor site, however here it is found in exon sequences. The 6 bases at site A in Figure 6 are present in all tested cDNA clones, with the exception of the originally described BC22 (3). Therefore the absence of this short sequence in BC22 is probably a splicing artefact. The splicing in BC22 has occurred at an AG in exon sequence that is downstream from the ag of the intron. In this case the original splice site has been bypassed, which is very unusual: when 2 ag sequences are within 10 bases of each other, splicing normally occurs at the upstream ag (15).
Splicing at the alternative sites (Fig. 5, B, C and D) occurs at different ratios, but no differences were detected between the various tissues analyzed.

DISCUSSION

Our goal was to pursue qualitative transcription information in various normal tissues and to examine if *FMR1* would be differentially expressed in various tissues, especially with regard to those tissues, brain and testis, that are affected in fragile X patients. Northern blot analysis indicated high expression of *FMR1* in brain, testis, placenta and lymphocytes, whereas in other organs a lower level of expression is seen (3). As RT-PCR is a very sensitive and rapid method to study mRNA transcripts and is also the method of choice to detect rare mRNAs or mRNAs in small amount of tissues or cells, we used RT-PCR to study normal

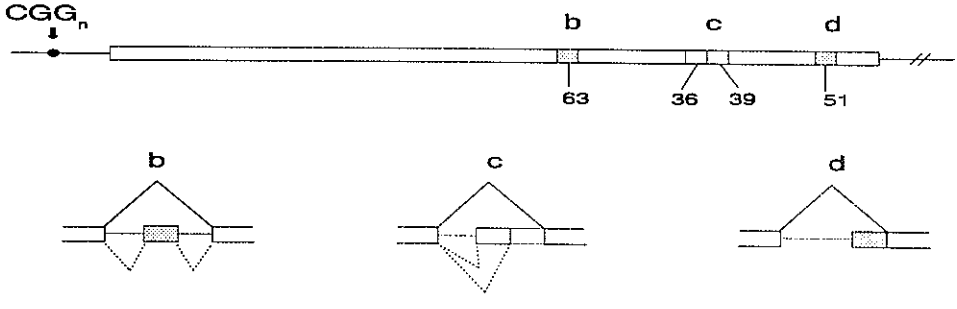


Figure 6. Schematic representation of the different splice sites B, C and D (not drawn to scale). In the upper part of the figure the *FMR1* cDNA is represented containing the different regions B, C and D that undergo splicing. In the lower part of the figure the different splice sites are depicted.

transcription of *FMR1* in different tissues.

We focused on the open reading frame of *FMR1* as alternative transcripts might encode isoproteins responsible for functional diversity of *FMR1*. No alternative splicing was found in the first half of the *FMR1* gene. The CGG repeat is on the same (first) exon as the start ATG and cannot be spliced out without removing the start ATG (8). Alternative splicing occurs in the *FMR1* gene but does not appear to be tissue specific as similar ratios of transcripts were found in all tissues analyzed. In these tissues, using primers K7 and K8, PCR products 1 and 2 are considerably more abundant than 3 and 4. In testis we find slightly different ratios of transcripts using primers K7 and K8. Although RT-PCR is not a quantitative method, in one reaction the same primers are used resulting in different alternative spliced mRNA products in different ratios.

The mouse *FMR1* sequence has a high homology to the human sequence. Using the primers K5/K6 the same alternative splicing is found in the mouse and human *FMR1* gene with the upperband containing the 63 basepairs in both species. But, when primers K7/K8 are used, a difference is observed. In adult mouse brain using these primers, the largest transcript is missing. In mouse brain no transcripts are found containing the 51 basepairs. In human tissues there is a 1:1 ratio for the presence or absence of splice site D (data not shown). Taking into account the relative abuncance of the different transcripts in human somatic tissues (Table 1), it seems that the majority of *FMR1* RNA products use the upstream 3' splice site at C (no bases excluded). The upstream and downstream 3' splice sites at D are used equally efficient in human. In mouse only the downstram 3' splice site at D is used. The product ratios at splice site B in favour of the absence of B is not influenced by the size of the preceding intron (which is larger than 500 bp), as there is a minimum size of about 80 base pairs for efficient splicing (15).

The likelihood that all the transcripts found are exon sequences and do not represent incompletely spliced mRNA products, is strengthened by the maintenance of an open reading frame and the apparent prevalence of different splice products in cDNA libraries. Two types of alternative splicing are found in *FMR1*. The 63 basepair sequence is a complete exon located between two introns and spliced into or out of the mRNA in a conventional way. Alternative splicing at sites A, C and D involves use of alternative acceptors. 3' splicing at sites A, C and D occurs inside exon sequences with AG conserved and preceding all 3' acceptor sites. As the absence of the 75 bases and the absence of 36 bases are not independent events, 3 combinations of different transcripts are possible for C (Fig. 6). Combining these 3 possible transcripts with the splice sites B and D, 12 distinct mRNA products could result from *FMR1* expression, each of which could be

translated into protein. Western blot analysis using antibodies directed to a portion of the *FMR1* protein has shown the presence of 4 discrete proteins of different mobilities in lymphoblastoid cell lines. These proteins are absent in cell lines of fragile X patients (C. Verhey et al, submitted). These 4 products could represent more than 4 proteins with overlapping lengths (the largest possible product being 631 amino acids and the smallest being 568 amino acids). Expression of different cDNA constructs in COS cells showed distinct single proteins with different lengths excluding extensive post-translational modification events. Whether all these proteins will have a different function has to be determined.

The finding of 12 possible different mRNAs in *FMR1* does not necessarily indicate that 12 isoforms of the protein are produced. However, the presence of at least 4 proteins has been confirmed by using *FMR1* specific antibodies.

New *FMR1* cDNA clones have been isolated from human fetal brain and liver cDNA libraries. The cDNA clones that were extended at the 5′ end as compared to the sequence of the original published *FMR1* cDNA showed a different sequence in front of the underlined shared sequence (Figure 3). A possible explanation for this phenomenon could be the presence of unstable secondary structures through which the reverse transcriptase is not able to proceed correctly. At the 3′ end the original BC72/22 cDNA clone contains a polyadenylation signal and the first A of a poly A tail (followed by vector sequence CGG (3)). This was confirmed by sequencing cDNA clones that did have longer poly A tails. Several liver cDNA clones are extended at the 3′ end beyond the originally published sequence. This suggests alternative use of different polyadenylation signals, although no new polyadenylation signal was found at the 3′ end of these liver clones, which are probably incomplete. Mouse cDNA clones that extend even further than the human clones at the 3′ end are found as well (Warren, pers. comm.). These mouse clones do contain a polyadenylation signal and are at the 3′ end very homologous to the human sequence. Combining all the sequence data from the different cDNA clones, the longest *FMR1* cDNA sequence consists of 4362 basepairs.

Alternative splicing is found in many genes and can introduce functional diversity into the products of a single gene. In most cases this gives rise to protein isoforms sharing extensive regions of identity and varying only in specific domains, thus allowing for the fine regulation of protein function. The functional significance of alternative splicing is described for several genes. 1) The function of differential splice products can be tissue specific, e.g.: transcripts derived from the *CALC1* gene are differentially processed in a tissue specific manner to include exon 4 in the thyroidal C-cells (encoding the precursor to the hormone calcitonin) and to exclude this exon in neuronal cells (encoding the precursor of the neuropeptide

CGRP) (16,17). 2) Differential splicing products can have different functions in separate developmental stages, e.g.: mRNA transcribed from the transformer (*tra*) gene of Drosophila melanogaster is subjected to sex-specific alternative splicing during embryogenesis (18,19). 3) Alternative splicing can create an altered coding frame, for instance in myotonic dystrophy. The 3' ends of the ORFs of brain and heart transcripts of the DM-kinase gene differ in length and codon usage (20). Consequently, different hydrophobic C termini are predicted for the putative kinases. Differential splicing could have a regulatory function by creating differential anchoring of the kinase to different cellular structures. 4) By means of alternative splicing domains with suggested functions (glycosylation or phosphorylation sites etc.) can be added to transcripts. Also in Wilms' tumor alternative splicing in the *WT1* gene results in protein isoforms with differing binding affinities to the EGR-1 consensus sequence (21,22). For the *FMR1* protein no homology is found to any other known protein, so no functions can be assigned to different domains. Which function can be attributed to the alternative splicing in the *FMR1* gene will be subject to further research. At least splicing in this gene does not seem to be tissue specific. However, *FMR1* like DM shows alternative splicing confined to the 3' half of the gene.

In conclusion, we have detected 12 possible *FMR1* transcripts that could result in 12 possible isoproteins. The next challenge will be to decipher the functions of these proteins and to determine whether the different isoforms will have different functions in different tissues.

MATERIAL AND METHODS

FMR1 liver cDNA isolation and characterization

A human liver cDNA library utilizing the vector lambda gt11 was constructed as described by Konecki et al. (23). Approximately 3 million recombinants were screened by filter hybridization using ^{32}P-labelled BglI/BamHI fragment (base 430 - 1730) from the FMR-1 isolated from a fetal human brain library (Stratagene). Hybridization-positive recombinants were plaque-purified by three rounds of purification. Phage DNA inserts were isolated, purified and subcloned into pBlueScript II-SK (-) vector for DNA sequence analysis. Sequencing was performed by the double-stranded dideoxy chain termination technique using the Pharmacia T7 sequencing kit.

RNA isolation

RNA was isolated according to Auffray and Rougeon (24). The LiCl method was used (procedure C) with several modifications. After overnight incubation at 0^0C in 3M

LiCl/6M urea the samples were spun down at 25K at 4^0C for 20 min. Proteinase K treatment (10 µg/ml) was performed (30 min at 37^0C) prior to phenol/chloroform extraction.

First strand cDNA synthesis

5 µg of total RNA was reversed transcribed as described by Pieretti et al (7) with minor modifications. Instead of precipitating the cDNA, 2 µl of the reverse transcribed reaction was directly used for PCR.

PCR

PCR reactions were done on 2 µl of cDNA solution with different primer sets. Amplification was carried out in a total volume of 50 µl containing 10 mM Tris-HCl (pH 8.3), 50 mM KCL, 1.5 mM $MgCl_2$, 0,25 mM of each dNTP and 0,5 units of Taq polymerase. The reaction was heated to 94^0C for 5 min, followed by 28 cycles of DNA denaturation (1 min at 94^0), annealing (1 min at 55^0C) and extension (1 min and 45 sec. at 72^0C). Final extension was at 72^0C for 10 min. 5-10 µl of PCR product was analyzed on a 1 or 3% agarose gel (1.5% regular and 1.5% nusieve agarose) depending on band sizes, stained with ethidium bromide.

The following primers were used:

Primer K1: 5' GGCGCTAGCAGGGCTGAAGAGA 3'

Primer K2: 5' AAAATCCTTATGTGCCGCCTCTTT 3'

K1 and K2 were derived from positions -24 - -3 and 430 - 449 (Fig. 3)

Primer K3: 5' TTGAACTTGTATTACATCTTCAGC 3'

Primer K4: 5' TTGGAACTTGTATTACATCTTCAGC 3'

K3 and K4 were derived from positions 392-414 and 844-868.

Primer K5: 5' GATGCAGTCAAAAAAGCTAGAAGC 3'

Primer K6: 5' CATCATCAGTCACATAGCTTTTTTC 3'

K5 and K6 were derived from positions 808-831 and 1372-1396.

Primer K7: 5' GCTAGTTCTAGACCACCACCAAAT 3'

Primer K8: 5' CTTAGGGTACTCCATTCAGGAG 3'

K7 and K8 were derived from positions 1336-1359 and 1878-1898.

Primer A3: 5' CAGGAAGCTCTCCCTCTCTT 3' is derived from position 1559-1578.

Sequencing

To sequence the PCR products bands were excised from the agarose gel and equilibrated in 1 ml of 300 mM NaAc pH 6.5 + 1 mM EDTA for 15 min at room temperature. The agarose containing the DNA was then cut in little pieces and transferred to 0.5 ml eppendorf tube that was punctured at the bottom and on top of some glasswool. The tube with agarose was kept at -80^0C for 30 min. and spun down in a 1.5 ml tube for 10 min at room temperature at 13000 rpm. To the supernatant 1/100 vol 1M $MgCl_2$ + 10% HAc was added and the DNA was precipitated with 2 volumes of ethanol at -80^0C for 30 min. The

DNA was spun down, washed with 70% ethanol and dissolved in 50µl of Tris-HCl (10 mM)/EDTA (0,1mM). Direct sequencing of approximately 30 ng of RT-PCR-product was performed using the BRL cycle sequencing kit.

ACKNOWLEDGEMENTS

We thank Jo Höppener, Annelies de Klein, Marianne Kroos and Marcel Koken for providing mRNA of different tissues, and Richard A. Gibbs and the DNA sequence core of the BCM Human Genome Center for sequence data, and M.F. Victoria and A. McAughan for technical assistance. This research was supported by the Foundation of Clinical Genetics to A.J.M.H. Verkerk en E. de Graaff, a grant of the MGCZWN to B.A. Oostra, and NIH grant NICHHDIROIHD29256 and NIH Human Genome Center grant P30-HG00210 to D.L.Nelson.

REFERENCES

1. Gustavson, K.H., Blomquist, H. and Holmgren G. (1988) Am. J. Med. Genet., 23, 58-588.
2. Webb, T.P., Bundey, SE., Thake, A.I. and Todd, J. (1986) Am. J. Med. Genet. 23, 573-580.
3. Verkerk, A.J.M.H., Pieretti, M., Sutcliffe, J.S., Fu, Y-H., Kuhl, D.P.A., Pizzuti, A., Reiner, O., Richards, S., Fuping Zhang, M.F.V., 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. and Waren, S.T. (1991) Cell, 65, 905-914.
4. Vincent, A., Heitz, D., Petit, C., Kretz, C., Oberle, I. and Mandel, J.L. (1991) Nature 349, 624-626.
5. Bell, M.V., Hirst, M.C., Nakahori, Y., Mackinnon, R.N., Roche, A., Flint, T.J., Jacobs, P.A., Tommerup, N., Tranebjaerg, L., Froster-Iskenius, U., Kerr, B., Turner, G., Lindenbaum, R.H., Winter, R., Pembrey, M., Thibodeau, S. and Davies, K.E. (1991) Cell 64, 861-866.
6. Oberle, I., Rousseau, F., Heitz, C., Kretz, C., Devys, D., Hanauer, A., Boue, J., Bertheas, M.F. and Mandel J.L. (1991) Science, 252, 1097-1102.
7. Pieretti, M., Zhang, F., Fu, Y-H., Warren, S.T., Oostra, B.A., Caskey, C.T. and Nelson, D.L. (1991) Cell, 66, 1-20.
8. 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, Jr., R.G., Warren, S.T., Oostra, B.A., Nelson, D.L. and Caskey, C.T. (1991) Cell, 67, 1047-1058.
9. Yu, S., Pritchard, M., Kremer, E., Lynch, M., Nancarrow, J., Baker, E., Holman, K., Mulley, I.C., Warren, S.T., Schlessinger, D., Sutherland, G.R. and Richards, R.I. (1991) Science, 252, 1179-1181.
10. Kremer, E., Pritchard, M., Lynch, M., Yu, S., Holman, K., Baker, E., Warren, S.T.,

Schlessinger, D., Sutherland, G.R., Richards, R.I. (1991) Science 252, 1711-1714.
11. Gedeon, A.K., Baker, E., Robinson, H., Partington, M.W., Gross, B., Manca, A., Korn, B., Poustka, A., Yu, S., Sutherland, G.R. and Mulley, J.C. (1991) Nature Genetics 1, 341-344.
12. Wöhrle, D., Kotzot, D., Hirst, M.C., Manca, A., Korn, B., Schmidt, A., Barbi, G., Rott, H.D., Poustka, A., Davies, K.E. and Steinbach, P. (1992) Am. J. Hum. Genet. 51, 299-306.
13. De Boulle, K., Verkerk, A.J.M.H., Reyniers, E., Vits, L., Hendrickx, J., Van Roy, B., van den Bos, F., de Graaff, E., Oostra, B.A. and Willems, P.J. (1993) Nature Genetics 3, 31-35, 1993.
14. Shapiro, M.B. and Senapathy, P. (1987) Nucleic Acids Res. 15, 7155-7174.
15. Wieringa, B., Hofer, E. and Weissman, C. (1984) Cell 37, 915-925.
16. Amara, S.G., Jonas, V., Rosenfeld, M.G., Ong, E.S. and Evans R.M. (1982) Nature 298, 240-244.
17. Cote, G.J., Stolow, D.T., Peleg, S., Berget, S.M. and Gagel, R.F. (1992) Nucleic Acids Res. 20, 2361-2366.
18. Baker, B.S. (1989) Nature 340, 521-524.
19. O'Neil, M.T. and Belote, J.M. (1992) Genetics 131, 113-128.
20. Jansen, G., Mahadevan, M., Amemiya, C., Wormskamp, N., Segers, B., Hendriks, W., O'Hoy, K., Baird, S., Sabourin, L., Lennon, G., Jap, P.L., Iles, D., Coerwinkel, M., Hofker, M., Carrano, A.V., de Jong, P.J., Korneluk, R.G. and Wieringa, B. (1992) Nature Genetics 1, 261-266).
21. Brenner, B., Wildhardt, G., Schneider, S. and Royer-Pokora, B. (1992) Oncogene, 7, 1431-1433.
22. Bickmore, W.A., Oghene, K., Little, M.H., Seawright, A., van Heyningen, V. and Hastie, N.D. (1992) Science 257, 235-237.
23. Konecki, D.S., Wang, Y., Trefz, F.K., Lichter-Konecki, U. and Woo, S.L.C. (1992) Biochemistry 31, 8363-8368.
24. Auffray, C. and Rougeon, F. (1980) Eur. J. Biochem. 107, 303-314.

PUBLICATION 7.5

A POINT MUTATION IN THE *FMR-1* GENE ASSOCIATED WITH MENTAL RETARDATION AND THE FRAGILE X PHENOTYPE

Kristel De Boulle[1], Annemieke J.M.H. Verkerk[2], Edwin Reyniers[1], Lieve Vits[1], Jan Hendrickx[1], Bernadette Van Roy[1], Feikje Van Den Bos[3], Esther de Graaff[2], Ben A. Oostra[4] & Patrick J. Willems[1].

[1] Department of Medical Genetics, University of Antwerp-UIA, Universiteitsplein 1, 2610 Antwerp, Belgium.

[2] Department of Clinical Genetics and [4]Cell Biology, Erasmus University, Dr Molewaterplein 50, 3153 DR Rotterdam, The Netherlands.

[3] Institute of Mentally Retarded Het Bouwhuis, Welnaweg 100, 7542 AK Enschede, The Netherlands.

Nature Genet. 3, 31-35, 1993.

SUMMARY

The vast majority of patients with fragile X syndrome show a folate-sensitive fragile site at Xq27.3 (FRAXA) at the cytogenetic level, and both amplification of the $(CGG)_n$ repeat and hypermethylation of the CpG island in the 5′ fragile X gene (*FMR-1*) at the molecular level. We have studied the *FMR-1* gene of a patient with the fragile X phenotype but without cytogenetic expression of FRAXA, a $(CGG)_n$ repeat of normal length, and an unmethylated CpG island. We find a single point mutation in *FMR-1* resulting in an $Ile^{367}Asn$ substitution. This *de novo* mutation is absent in the patient's family and in 130 control X chromosomes, suggesting that the mutation causes the clinical abnormalities. Our results suggest that mutations in *FMR-1* are directly responsible for fragile X syndrome, irrespective of possible secondary effects caused by FRAXA.

INTRODUCTION

The recent isolation of the fragile X gene (*FMR-1*)[1] has led to a search for mutations in *FMR-1* responsible for the clinical and cytogenetic expression of the fragile X syndrome. The vast majority of fragile X patients have a significant increase in the $(CGG)_n$ repeat length[1-5]. This $(CGG)_n$ repeat is located between the CpG island that is hypermethylated in fragile X patients[6-9], and the start codon of the open reading frame of *FMR-1* (S. Warren, personal communication). Males with repeat inserts greater than 600 basepairs (bp) (full mutations) almost invariably show clinical and cytogenetic expression of the disease[5]. This suggests that expansion of the $(CGG)_n$ repeat to the full mutation length induces the fragile site (FRAXA). Further evidence that the fragile site is located in the $(CGG)_n$ repeat is provided by: (i) the absence of cytogenetic expression of FRAXA in two patients with clinical stigmata of fragile X syndrome and a molecular deletion of the whole *FMR-1* gene[10], or part of *FMR-1* including the $(CGG)_n$ repeat[11]; (ii) the physical proximity of *FMR-1* and FRAXA as evidenced by fluorescent *in situ* hybridization of *FMR-1* to fragile X chromosomes[12]; and (iii) the localization of fragile X site-induced translocation breakpoints in the $(CGG)_n$ repeat region[2] (S. Warren, personal communication).

The role of *FMR-1* in the fragile X syndrome is evidenced by the presence of an unstable microsatellite in the *FMR-1* gene, similar to that identified in myotonic dystrophy[13-15] and spinal and bulbar muscular atrophy[16], the finding of hypermethylation of the CpG island located 5′ of the open reading frame of *FMR-1* (refs 6-9), which probably represents a regulatory mechanism, and the observation of markedly reduced *FMR-1* mRNA levels in male fragile X patients[17]. However, the mechanism of the mutation remains unclear. Although *FMR-1* is not expressed in

the majority of male fragile X patients[17], it is not yet clear whether or not deficient activity of *FMR-1* protein is directly responsible for the clinical abnormalities of the fragile X syndrome. It is possible that the abnormalities of the chromatin structure induced by the fragile site, or hypermethylation of the Xq27.3 region down regulates genes in the proximity of *FMR-1*, which might cause the clinical abnormalities of the fragile X syndrome. In this respect it is interesting to note that the activity of iduronate sulphatase, encoded by *IDS* located approximately 1 centiMorgan (cM) distal to the *FMR-1* gene, is reduced in affected fragile X males[18]. The fact that the two patients with a partial or complete deletion of *FMR-1* (refs 10,11) exhibit fragile X syndrome in the absence of the fragile site, is only weak evidence for a direct role of *FMR-1*, as the patients' deletions presumably encompass adjacent genes.

We therefore searched for an intragenic mutation in *FMR-1* of a patient with clinical stigmata of the fragile X syndrome, but without cytogenetic expression of FRAXA. This work reports a single point mutation in the open reading frame of *FMR-1* of this patient, which was not found in 130 control X chromosomes, suggesting that mutations in *FMR-1* can be responsible directly for fragile X syndrome.

RESULTS

Clinical description and cytogenetic analysis

The patient was born after an uneventful pregnancy and vaginal delivery at 37 weeks gestation. Psychomotor delay was present from an early age on, and developing milestones were delayed with sitting at 20 months, standing without support at 4 years and walking alone at 5 years. Because of this developmental delay and also because of hepatomegaly he was admitted to hospital. X-linked liver glycogenosis due to phosphorylase kinase deficiency with growth retardation and hepatomegaly was diagnosed at the age of 3 years. The patient belongs to an extended pedigree affected with this disorder[19]. The major symptoms of this glycogenosis, located at Xp22 (ref. 19), are hepatomegaly and growth retardation during infancy with gradual disappearance of the symptoms thereafter[20]. Both symptoms had been present in the index case until the age of approximately 10 years, but liver size and height are normal at the age of 26. Mental retardation, neurological and phenotypic abnormalities such as present in the index case, are not present in the family, also not in 29 relatives affected with X-linked liver glycogenosis.

As well as his liver glycogenosis the patient developed severe mental retardation with an IQ below 20. He is institutionalized and has only primitive contact with his surroundings. There has never been any verbal communication except for some

a *b* *c*

Figure 1. *a*, Appearance of the patient's face with coarse acromegaloid features, large forehead, asymmetric long face, large ears, thick lips, mandibular prognathism; *b*, impressive macroorchidism; *c*, hypotrophic lower extremities.

sounds. Focal seizures are successfully treated with antiepileptic medication. Examination at the age of 27 years showed a severely retarded male with a peculiar phenotype. The face was asymmetric and elongated with prominent supraorbital ridges, sunken eyes, mandibular prognathism and coarse features giving an acromegalic impression (Fig. 1). The ears were large (8 cm), but normally implanted. Other phenotypic abnormalities include a thick lower lip, brachydactyly type E of the fourth ray of the foot, and short toe nails. The thorax and upper extremities were very muscular in contrast to the lower extremities which were hypotrophic. Impressive macroorchidism was present with both testicles exceeding 100 ml volume (Fig 1). Height was 1.85 m (90th centile), weight was 65 kg (50th centile) and OFC was 58 cm (between 50th and 98th centile). The liver was palpable at the costal margin. Neurological examination showed spastic paraparesis of the lower extremities with increased deep tendon reflexes. Flat feet, and X-deformity of the knees together with paraparesis made normal gait impossible.

A total of 100 mitoses of a folate-deprived culture without trimethoprim and 30 mitoses of a culture with trimethoprim were examined in our laboratory. In none of them was the fragile X site at Xq27.3 (FRAXA) observed. The analysis was repeated in two different laboratories at two separate occasions. A total of 280 mitoses was investigated in folate-deprived culture conditions and none showed FRAXA. Routine cytogenetic analysis in three laboratories revealed a normal male karyotype 46,XY, with no abnormality in the Xp22 or Xq27.3 region.

$(CGG)_n$ repeat analysis

Southern blot analysis of an EcoRI digest of DNA of the patient (P) with the probe pP2, showed the normal 5.2 kb EcoRI fragment without amplification of the $(CGG)_n$ repeat (Fig. 2a). A double digest with both EcoRI and the methylation-sensitive restriction enzyme EagI was performed, allowing detection of both the amplification of the $(CGG)_n$ repeat and the methylation status of the CpG island preceeding the $(CGG)_n$ repeat. The patient (P), his mother (M), his brother (B) and the male control (C), have a normal 2.8 kb unmethylated DNA fragment. The male with a premutation (X) has an unmethylated DNA band of 3-3.5 kb. The male fragile X patient with a full mutation (F) shows a smear in the full mutation range without a normal 2.8 kb fragment due to complete methylation of the CpG island (Fig. 2b). Polymerase chain reaction (PCR) amplification of the $(CGG)_n$ repeat region of our patient indicated a repeat number of 25, which is in the normal range of 5 to 54 repeats[4].

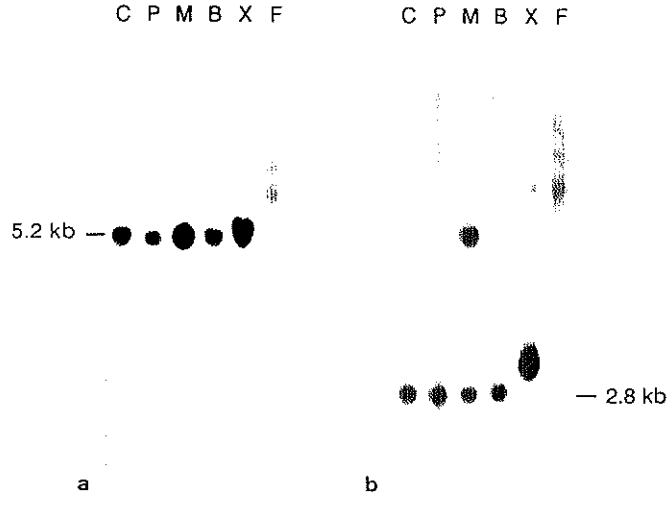


Figure 2. Southern blot of EcoRI (a) and EcoRI/EagI (b) digested genomic DNA of a male control (C), the patient (P), the patient's mother (M), the patient's brother (B), a male with a premutation (X) and a male fragile X patient with a full mutation (F).

On EcoRI digests C,P,M and B show a normal 5.2 kb fragment, whereas X shows a 5.5 kb fragment due to a premutation in the $(CGG)_n$ repeat, and F shows a smear due to a large insertion in the $(CGG)_n$ repeat. The double digests with EcoRI and EagI show both the $(CGG)_n$ amplification and the methylation status of the CpG island. C,P and B have a normal unmethylated fragment of 2.8 kb. M has both a normal unmethylated (active) 2.8 kb fragment, and a normal methylated (inactivated by the lyonisation) 5.2 kb fragment. X shows an unmethylated premutation band between 3 and 3.5 kb. F has a methylated smear in the full mutation range.

mRNA quantification

Because of absence of expression of *FMR-1* in the majority of male fragile X patients[17], a PCR reaction was carried out on first strand cDNA to check the level of *FMR-1* mRNA in our patient. First strand cDNA was synthesised with reverse transcriptase from RNA isolated from an Epstein-Barr virus transformed cell line of the patient. The resulting cDNA was amplified with a set of primers derived from the *FMR-1* sequence in combination with a pair of primers derived from the hypoxanthine-guanine phosphoribosyltransferase (*HPRT*) sequence as an internal control[17]. The patient shows bands of the same size and intensity as a normal control (data not shown). There is also no difference in intensity between the *FMR-1* and the *HPRT* band which suggests that the patient has a normal level of *FMR-1* mRNA.

Sequence analysis

Using first strand cDNA, four overlapping *FMR-1* cDNA fragments (Fig. 3) were amplified by PCR. The four generated reverse transcriptase PCR fragments (RT-PCR I-IV) of the patient showed a normal length and intensity as compared to RT-PCR products of a control. The four RT-PCR products were subcloned and sequenced. Comparison of the complete open reading frame of *FMR-1* revealed only one alteration, a T to A conversion (Fig. 4) at position 1100 in the *FMR-1* sequence[1]. This substitution converts an Ile codon at amino acid position 367 to an Asn codon. To exclude a PCR or sequencing artefact, the RT-PCR III reaction was repeated and the

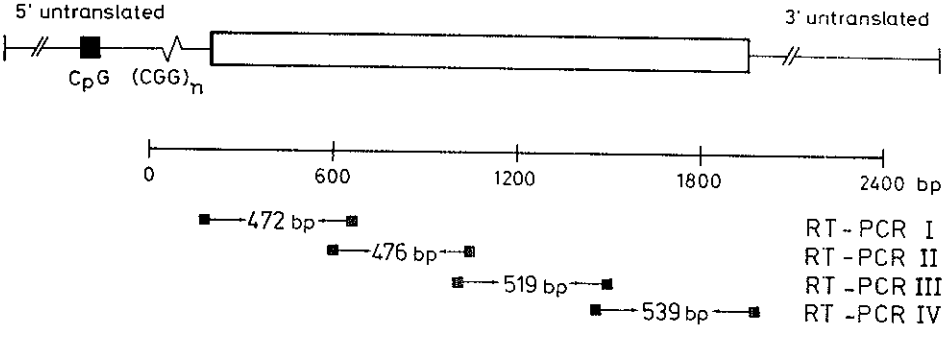


Figure 3. Reverse PCR amplification of cDNA from the *FMR-1* gene.
Schematic representation of the cDNA structure of *FMR-1* and the design of the four reverse transcriptase PCR reactions (RT-PCR I-IV). The boxes represent the respective primer sets. The indented area indicates the $(CGG)_n$ repeat and the hatched area represents the CpG island hypermethylated in the majority of fragile X patients. Numbering of base pairs and amino acids is according to the published sequence[1]. Recent findings indicate that the actual startcodon is located at position 66 of the published[1] sequence (S. Warren, personal communication).

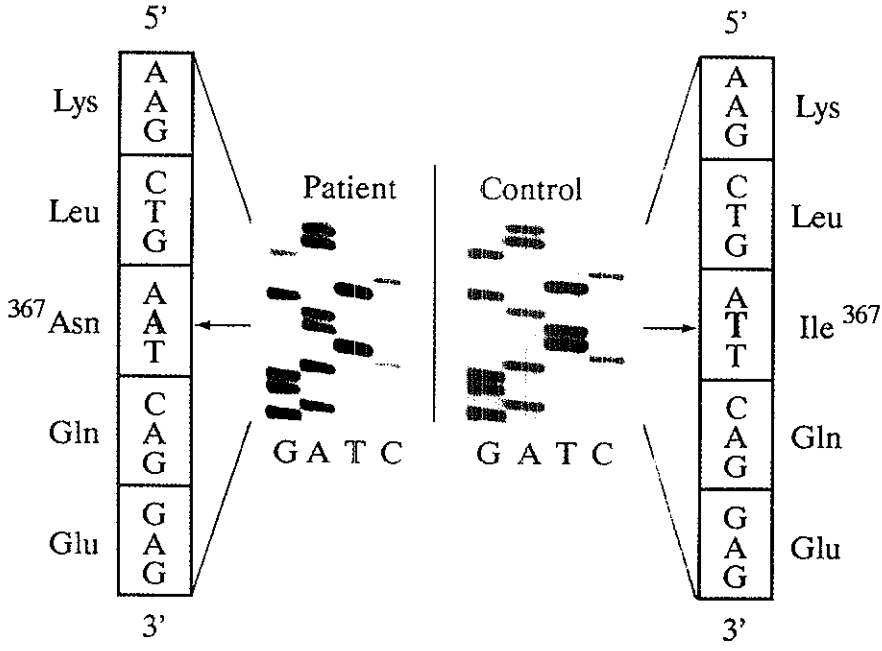


Figure 4. Autoradiograph of a sequencing gel of the middle portion of the RT-PCR III product of the *FMR-1* cDNA. The figure shows the single T to A substitution in the cDNA of the affected patient. This point mutation results in an Ile367Asn substitution (arrowhead).

product resequenced both after subcloning and by direct sequencing. This confirmed the presence of the T to A point mutation.

Mutation analysis

To determine whether this point mutation is genetically associated with mental retardation in this family, pedigree analysis was performed. Using primers flanking the T to A point mutation and genomic DNA of the patient and his family as a template, PCR products were subjected to slot blot analysis with wild type and mutant allele specific oligonucleotides (ASO) used as hybridization probes. Figure 5 shows that the T to A mutation is present in the patient, but not in the mother, brother and nephews who have normal intellectual abilities and no stigmata of fragile X syndrome. Therefore, this point mutation represents a new mutation in the patient.

To assess the possible presence of the T to A mutation in the normal population we also investigated genomic DNA from 130 control X chromosomes by slot blot analysis using allele-specific oligonucleotides (ASO). The mutant allele, however, was not present in the normal population.

DISCUSSION

Although *FMR-1* has been isolated and the most prevalent fragile X mutation identified, little is known about the role of *FMR-1* as the predicted protein sequence has given little hint of its function. It remains unclear if the amplification of the *FMR-1* $(CGG)_n$ repeat of fragile X patients is responsible for their clinical abnormalities solely by impairing the function of *FMR-1*. As the amplification of the $(CGG)_n$ repeat induces a fragile site FRAXA with gross alterations of the chromatine structure of the Xq27.3 region, whereas the hypermethylation of the CpG island 5' of the open reading frame of *FMR-1* might down regulate adjacent genes, the mutation in *FMR-1* might affect the expression of other genes in Xq27.3. Therefore, it is not impossible that the fragile X syndrome is a contiguous gene syndrome in which more than one gene is affected. In this study, we asked the question whether intragenic mutations in *FMR-1* might not be directly responsible for the clinical abnormalities of the fragile X syndrome.

We therefore analysed the *FMR-1* gene of a patient with clinical evidence of fragile X, but without cytogenetic expression of FRAXA, amplification of the $(CGG)_n$ repeat or hypermethylation of the CpG island. We specifically choose to study this particular patient as he not only has facial stigmata compatible with the fragile X syndrome, but also macroorchidism. Whereas the facial stigmata of the

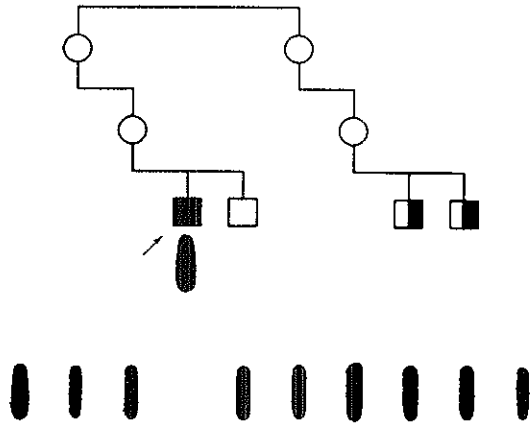


Figure 5. Pedigree of the family. The arrowhead represents the index patient.
◧ indicates mental retardation, ◨ indicates liver glycogenosis. The corresponding slot blot analysis is shown below: the upper row of DNA is hybridised to the mutant ASO recognizing the T to A point mutation, whereas the lower row is hybridised to the wild type ASO. The DNA of the family is flanked at both sides by DNA of a control. The figure illustrates that the point mutation is only present in the index patient and not in his relatives.

fragile X syndrome are more or less aspecific, impressive macroorchidism (such as present in this patient) suggests fragile X syndrome. Our patient is also affected with X-linked liver glycogenosis (*XLG*) due to phosphorylase kinase deficiency. *XLG* is a mild hepatomegalic glycogen storage disease causing hepatomegaly, growth retardation, hypercholesterolaemia and liver transaminase elevations during infancy and adolescence, but few symptoms thereafter[20]. The *XLG* gene has been mapped to Xp22 by linkage analysis[19], whereas the gene encoding the liver α-subunit of phosphorylase kinase (*PHKA2*) was physically mapped to the same chromosomal region[21,22]. Therefore, *XLG* is probably caused by mutations in *PHKA2* (ref. 23). The mental retardation and macroorchidism of our patient can not be attributed to this disorder as the clinical spectrum of *XLG* does not include the latter symptoms[20]. Furthermore, mental retardation was absent in all other relatives affected with *XLG*. This suggested that not *XLG* but a different condition, possibly the fragile X syndrome, was responsible for the clinical abnormalities of this patient.

The significance of our patient's *FMR-1* mutation is suggested by the fact that his $(CGG)_n$ repeat number (25) lies well within the normal range (5-54). The mutation arose *de novo*, and is not seen in 130 normal X chromosome. As little is known of the overall function of *FMR-1* and the genomic organisation of *FMR-1*, we cannot speculate on the functional effects of the Ile367Asn mutation in our patient. It is surprising, however, that a mutant protein which differs by only one amino acid, apparently produces more severe clinical abnormalities than the classical fragile X mutation where hardly any *FMR-1* mRNA is present, or the two deletion patients reported recently[10,11]. Neurological impairment with spastic paraparesis, muscle atrophy, hypotrophy of the lower extremities, and phenotypic abnormalites such as brachydactyly type E are uncommon in fragile X syndrome, and mental retardation and macroorchidism of our patient was particularly severe.
A protein suicidal effect of the mutant *FMR-1*, such as occurs in certain collagen disorders, is a possibility. A regulatory function of *FMR-1* might be more impaired by a mutation than by a reduction in the amount of normal protein. In any case, the $Ile^{367}Asn$ substitution must alter significantly the function of an important domain of *FMR-1*. In mouse *Fmr-1*, the Ile at this position is conserved (B.A.O., unpublished results). Secondary structure algorithms[24] indicate an alteration in the ß sheet structure of *FMR-1* by the $Ile^{367}Asn$ substitution. However, the predictive value of secondary structure algorithms is very limited. A formal proof that the T to A point mutation is responsible for the clinical abnormalities of our patient, would require site-directed mutagenesis introducing the mutation *in vitro* and/or *in vivo* transgenic constructions. However, this remains impossible as long as there are no biological parameters to assess the activity of *FMR-1*. As our *FMR-1* mutation does not lead to amplification of the $(CGG)_n$ repeat and/or methylation of the

adjacent CpG island, it probably plays a direct role in the development of the mental retardation and neurological abnormalities, which suggests an important physiological function of *FMR-1* in neurological tissue.
Our findings suggests that the fragile X syndrome is a single gene disorder and not a contiguous gene syndrome. The presence of an intragenic *FMR-1* mutation different from the classical amplification of the $(CGG)_n$ repeat also illustrates the necessity to scan the *FMR-1* gene in fragile site-negative patients with mental retardation, and especially in patients with mental retardation and macroorchidism. As it becomes clear that the clinical and cytogenetic definition of the fragile X syndrome has been biased towards fragile site-positive patients, continued molecular analysis of the *FMR-1* gene will probably lead to a change from an indefinite homogeneity to a definite heterogeneity in the fragile X syndrome.

Methodology

Subjects. The patient belongs to an extended Dutch pedigree with more than 20 individuals affected with X-linked liver glycogenosis due to deficiency of phosphorylase kinase activity. The pedigree has been reported before[19] and was used to localize the gene responsible for X-linked liver glycogenosis by linkage analysis. In this report the *FMR-1* gene was studied in part of the pedigree (Fig 5). Genomic DNA was extracted from peripheral lymphocytes, whereas cDNA was prepared from RNA obtained from EB-transformed lymphoblastoid cell lines. We used genomic DNA from peripheral lymphocytes from 65 female Caucasians as controls.

Cytogenetic analysis. Fragile X was analysed in 96 hours whole blood cultures in low folate medium (M199) with 2% fetal calf serum (FCS), as previously described[25], and in the same medium supplemented with 20 mg/ml trimethoprim[26]. Slides were stained with 5% Giemsa and examined for fragile sites. After destaining the slides were GTG-banded to identify the chromosome involved. The cytogenetic analysis was repeated in two different genetic laboratories at two separated occasions.

Southern blots and PCR amplification. Genomic DNA of the patient was prepared from peripheral blood lymphocytes according to standard procedures. Approximately 6 µg of DNA was digested with EcoRI and EcoRI/EagI for Southern blot analysis with pP2 used as a probe. pP2 is a 1 kb PstI fragment derived from pE5.1 which identifies the $(CGG)_n$ repeat and the preceeding CpG island[27]. Amplification of genomic DNA of the region around the $(CGG)_n$ repeat was carried out according to published procedures[4], with the modification that the concentration of $MgCl_2$ was 2.5mM.

RNA extraction. Total cellular RNA was extracted from an EBV transformed lymphoblastoid cell line. A pellet of lymphoblastoid cells was mixed in 10 ml icecold LiCl-Urea (3 M LiCl, 6 M Urea) for 2 min. After overnight incubation on ice the suspension was pelleted for 20

min at 25,000 rpm at 0°C. The pellet was dissolved in 1 ml of 0.01 M Tris-HCl, 0.01 M EDTA, 1 % SDS, 10 µg/ml prot K and incubated at 37°C for 30 min. The solution was extracted once with phenol-chloroform and once with chloroform. RNA was precipitated with ethanol and dissolved in 300 µl H_2O.

cDNA synthesis. Ten µg of total cellular RNA obtained from lymphoblastoid cells was incubated at 42°C for 1 h in a total volume of 40 µl containing 2 µg of random hexamer primers $d(N)_6$, 3 µg oligo(dT)primer $d(T)_{15}$, 1.2 mM of each dNTP, 60 mM Tris-HCl (pH 8.3), 60 mM $MgCl_2$, 4 mM dithiothreitol, 25 U of RNAase inhibitor and 5 µl of Superscript RNAase H-RT (200 u/µl). After first-strand synthesis the cDNA was collected by ethanol precipitation and resuspended in 20 µl H_2O.

RNA quantification. PCR was carried out on a DNA thermal cycler (Perkin-Elmer, Cetus) as described[17] using 2 µl first strand cDNA as a template. After initial denaturation at 94°C for 5 min, the PCR consisted of 28 cycles of 1 min at 94°C, 1 min at 55°C and 1 min 45 sec at 72°C. Final extension was at 72 °C for 10 min. The *FMR-1* primers used were 5' TAGCTAACCACCAACAGCAAGGC 3' and 5'AACTGGCAGCCTGATAGGCAGATTC 3', derived from positions 297-320 and 420-442 of the published *FMR-1* sequence[1]. PCR products were visualised on a 3 % agarose gel stained with ethidiumbromide.

PCR amplification and sequencing. Based on the cDNA sequence of the normal *FMR-1* gene[1], we designed 4 pairs of primers to amplify the entire protein-coding region of *FMR-1* (Fig. 3). The primers used were K1 (5' GGAATTCCGGCGCTAGCAGGGCTGAAGAGA 3') and K2 (5' GGAATTCCAAAATCCTTATGTGCCGCCTC TTT 3') derived from positions 172-192 and 622-645 of the *FMR-1* sequence[1]; K3 (5'GGAATTCCTGGATGTGCCAGAAGACT TACGG 3') and K4 (5' GGAATTCCTTGG AACTTGTATTACATCTTCAGC 3') from positions 586-609 and 1039-1063; K5 (5' GGAATTCCGATGCAGTCAAAAAAGCTAGAAGC 3') and K6 (5' GGAATTCCC ATCATCAGTCACATAGCTTTTTTC 3') from positions 1003-1027 and 1498-1521; K7 (5' GGAATTCCGCTAGTTCTAGACCACCACCAAAT 3') and K8 (5' GGAATT CCGGAAATATAACTTCAGAATTATGCAGT 3') from positions 1462-1485 and 1975-2001 of the *FMR-1* sequence (Fig. 3). PCR was carried out as described before[17]. After initial denaturation at 94°C for 7 min, 40 cycles of PCR amplification were performed. Each cycle consisted of: 94°C for 1 min; 55°C for 1 min and 72°C for 1 min 45s. Final extension was at 72°C for 10 min. The PCR products were subcloned into Bluescript SK- and sequenced with vector primers using the dideoxynucleotide termination method (Sequenase version 2.0, USB). Additionally, direct sequencing of the RT-PCR III product,using the BRL cycle sequencing kit, was performed to confirm the point mutation.

Slot blot detection of the point mutation. Amplification of 500 ng genomic DNA was carried out in a total volume of 50 µl containing 1 x Cetus Taq DNA polymerase buffer, 100 pmol of each primer, 125 mM of each dNTP and 2.5 U Taq Pol. After an initial 4 min

at 94°C, the PCR consisted of 32 cycles of 1 min at 94°C, 2 min at 55°C, and 2 min at 72°C. Final extension was at 72°C for 10 min. The primers used were 5' CTGCTTTGAGGTATGTGTTTTTA 3' (intron sequence in intron at position 1073 of the *FMR-1* sequence[1]) and 5' AATATTCGCACTGTAACATACCT 3' (starting from position 1177 of the *FMR-1* sequence[1]). PCR products were denaturated in 0.4 M NaOH and 25 mM EDTA for 10 min and applied directly on Hybond N^+ membranes using a slot blot apparatus. The filters were air dried, prehybridized for 1 h at 42°C in 5 x SSPE, 1% SDS and 50 µg/ml salmon sperm DNA and subsequently hybridized for O/N in the same solution to [^{32}P] labelled oligonucleotides. The filters were washed in 3 x SSC for 15 min. The following allele-specific oligonucleotides (ASO) were used: GGAAAGCTGATTCAGG AGAT (wild type) and GGAAAGCTGAATCAGGAGAT (mutant).

Acknowledgements

We thank G.P.A. Smit and I. Stolte for clinical information, A. Hoogeveen for secundary structure algorithm analysis, R. Bernaerts and K. Denecker for secretarial assistance and M. D'Hollander and O. Beck for technical assistance. We are indebted to D.L. Nelson for release of sequence data prior to publication. The continuous encouragement of J. Fernandes in the study of this family is greatly acknowledged. Part of this study was supported by a grant of the Belgian National Foundation for Scientific Research (NFWO) to P.J.W., a grant of the Foundation of Clinical Genetics to A.J.M.H.V. and a grant of the MGCZWN to B.A.O.

REFERENCES

1. Verkerk, A.J.M.H. *et al.* 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 (1991).
2. Kremer, E. *et al.* Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence $p(CCG)_n$. *Science* **252**, 1711-1714 (1991).
3. Yu, S. *et al.* Fragile X genotype characterized by an unstable region of DNA. *Science* **252**, 1179-1181 (1991).
4. Fu, Y.H. *et al.* Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* **67**, 1047-1058 (1991).
5. Rousseau, F. *et al.* Direct diagnosis by DNA analysis of the fragile X syndrome of mental retardation. *N. Engl. J. Med.* **325**, 1673-1681 (1991).
6. Bell, M.V. *et al.* Physical mapping across the fragile X: Hypermethylation and clinical expression of the fragile X syndrome. *Cell* **64**, 861-866 (1991).
7. Heitz, D. *et al.* Isolation of sequences that span the fragile X and identification of a fragile X-related CpG Island. *Science* **251**, 1236-1239 (1991).
8. Oberlé, I. *et al.* Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science* **252**, 1097-1102 (1991).

9. Vincent, A. *et al.* Abnormal pattern detected in fragile X patients by pulsed-field gel electrophoresis. *Nature* **349**, 624-626 (1991).
10. Gedeon, A.K. *et al.* Fragile X syndrome without CGG amplification has an *FMR-1* deletion. *Nature genetics* **1**, 341-344 (1992).
11. Wöhrle, D. *et al.* A microdeletion of less than 250 kb, including the proximal part of the *FMR-1* gene and the fragile X site, in a male with the clinical phenotype of the fragile X syndrome. *Am. J. Hum. Genet.* **51**, 299-306 (1992).
12. Verkerk, A.J.M.H., Eussen, B.H.J., Van Hemel, J.O. & Oostra, B.A. Limited size of the fragile X site shown by fluorescence in situ hybridization. *Am. J. Med. Genet.* **43**, 187-191 (1992).
13. Brook, J.D. *et al.* Molecular basis of myotonic dystrophy: Expansion of trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* **68**, 799-808 (1992).
14. Fu, Y.-H. *et al.* An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* **255**, 1256-1258 (1992).
15. Mahadevan, M. *et al.* Myotonic dystrophy mutation: An unstable CTG repeat in the 3' untranslated region of the gene. *Science* **255**, 1253-1255 (1992).
16. La Spada, A.R. *et al.* Androgen receptor gene mutation in X-linked spinal and bulbar muscular atrophy. *Nature* **352**, 77-79 (1991).
17. Pieretti, M. *et al.* Absence of expression of the *FMR-1* gene in fragile X syndrome. *Cell* **66**, 817-822 (1991).
18. Clarke, A. *et al.* Fragile X mental retardation and the *iduronate sulphatase* locus: Testing Laird's model of Fra(X) inheritance. *Am. J. Med. Genet.* **43**, 299-306 (1992).
19. Willems, P.J. *et al.* Mapping of the gene for X-linked liver glycogenosis due to phosphorylase kinase deficiency to human chromosome region Xp22. *Genomics* **9**, 565-569 (1991).
20. Willems, P.J., Gerver, W.J.M., Berger, R. & Fernandes, J. The natural history of liver glycogenosis due to phosphorylase kinase deficiency: A longitudinal study of 41 patients. *Eur. J. Pediatr.* **149**, 268-271 (1990).
21. Wauters, J.G. *et al.* Regional mapping of a liver α-subunit gene of phosphorylase kinase to the distal region of chromosome Xp. *Cytogen. Cell Genet.* **60**, 194-196 (1992).
22. Davidson, J.J. *et al.* cDNA cloning of a liver isoform of the phosphorylase kinase α-subunit and mapping of the gene to Xp22.2-p22.1, the region of human X-linked liver glycogenosis. *Proc. Natl. Acad. Sci. USA* **89**, 2096-2100 (1992).
23. Chou, P.Y. & Fasman, G.D. Prediction of the secondary structure of protein from their amino acid sequence. *Adv. Enzym.* **47**, 45-147 (1978).
24. Van Roy, B.C., De Smedt, M.C., Raes, R.A., Dumon, J.E. & Leroy, J.G. Fragile X trait in a large kindred: transmission also through normal males. *J. Med. Genet.* **20**, 286-289 (1983).
25. McLean, D.A. & Faed, M.J.W. Improved fragile site detection with trimethoprim. *Hum. Genet.* **85**, 241-243 (1990).
26. Oostra,B.A. & Verkerk, A.J.M.H. The fragile X syndrome: Isolation of the *FMR-1* gene and characterization of the fragile X mutation. *Chromosoma* **101**, 381-387 (1992).

SUMMARY

The objective of the research described in this thesis was to isolate and characterize the gene involved in the fragile X syndrome. For this purpose we used the approach of positional cloning, a technique by which a growing number of genes has been isolated in the past years (*Chapter 1*).

The fragile X syndrome, also called the Martin-Bell syndrome, is one of the most common genetic diseases, with mental handicap and enlarged testes in males as its main characteristics. About 1 in 1250 males is affected with this disease, and although the syndrome shows an X-linked pattern of inheritance, normal male carriers do exist, and about 50-70% of the females carrying the "full mutation" gene defect are mildly to moderately affected. *Chapter 2* gives some historical aspects about the "discovery" of the syndrome. In that chapter also an overview is given of the physical and behavioral aspects found in fragile X patients. In addition, the unusual pattern of inheritance of the fragile X syndrome is discussed.

The prevalence of the syndrome, the severity of the disease in male patients, and the urge to understand the defects at the molecular and biochemical level started the search for the gene involved in fragile X syndrome. In *Chapter 3* an overview is given of the different approaches used by us and other research groups in the hunt for this gene. The finding of a "fragile site" at the end of the long arm of the X chromosome in cells of male patients indicated that the fragile X gene was to be found in the Xq27-28 region of the X chromosome. This fragile site was thought to be associated with the gene involved in fragile X syndrome. With the aid of genetic and physical mapping, the region of interest was eventually restricted to 2-2,5 megabases. We used probes from this region, that mapped close to the fragile site, to isolate yeast artificial chromosomes (YACs) containing human X chromosome material from the Xq27.3 region (*Chapter 7; Publication 7.1*). We isolated one YAC of 475 kb (YAC 209G4), which was shown by fluorescence *in situ* hybridization (FISH) to cross the fragile site (*Chapter 7; Publication 7.2*). This YAC contained a CpG island that was found to be preferentially methylated on the fragile X chromosome of male patients and many female carriers of the syndrome. Since CpG islands often are found 5' to mammalian genes, it was possible that a gene might reside near this fragile site related CpG island. We constructed a cosmid library from the yeast clone harbouring YAC 209G4, and human DNA containing cosmids were used to screen a cDNA library derived from normal human fetal brain RNA. This resulted in the isolation of a gene, which we have named *FMR1* (Fragile X Mental Retardation 1) (*Publication 7.1*). The *FMR1* gene

was found to contain a CGG repeat in its 5' untranslated region. This repeat was located 250 basepairs distal from the CpG island which is hypermethylated in the DNA of fragile X patients. One of the cosmids (cosmid 22.3) used to isolate the gene, contained this CpG island and the CGG repeat. With FISH it was found that this cosmid showed signals proximal and distal from the fragile site (*Publication 7.2*), suggesting that the *FMR1* gene was closely associated with the fragile site. It was demonstrated that the CGG repeat is involved in the origin of the fragile X syndrome. In the *FMR1* gene of normal individuals, an average number of 29 CGG repeats is found, varying from 6-54 repeats. In fragile X carriers and patients this repeat is increased in length. In most of the phenotypically normal carriers, this increase is small and called a premutation, as this small increase is not associated with the disease phenotype. In fragile X syndrome patients the repeat has increased to over 200 copies. Due to this large number of repeats, methylation of the promoter region of the *FMR1* gene occurs, resulting in lack of transcription and absence of *FMR1* protein. This large increase in repeat size is termed a full mutation, as this number of repeats is associated with the disease phenotype. In fragile X families, the CGG repeat is unstable, which means that the number of repeats may increase from one generation to the next, eventually resulting in the fragile X phenotype in a number of individuals. This discovery, that amplification of a trinucleotide repeat is the molecular basis of the fragile X syndrome, revealed a new mechanism of transgenerational inheritance and has been a fundamentally new concept in genetics. This finding was soon followed by the identification of other diseases, described in *Chapter 5*, in which triplet amplifications are responsible for the disease phenotype. The genes associated with these disorders all contain a polymorphic GC rich trinucleotide repeat, that is unstable and expanded in affected individuals.

The *FMR1* gene is highly expressed in human brain and testes, tissues involved in the pathogenesis of the syndrome, and encodes a mRNA of about 4,4 kb. The gene has been conserved during evolution and can be found in organisms as divergent as mouse, yeast and *Caenorhabditis elegans*. Several domains have been identified in the predicted *FMR1* protein sequence that were recently found in RNA binding proteins, suggesting an RNA binding activity for the *FMR1* protein.

We found that the *FMR1* gene is subject to alternative splicing (*Publication 7.4*). This could give rise to as many as twelve different *FMR1* mRNA transcripts and hence to twelve possible isoproteins. Indeed different isoproteins are found in lymphocytes of normal individuals. The splicing described in *Publication 7.4* does not seem to be tissue specific, because no differences were detected between various tissues analyzed.

Cloning and characterization of the *FMR1* gene and the identification of the gene defect in fragile X patients, has improved the way in which the syndrome can be diagnosed. Now the number of CGG repeats in the *FMR1* gene of individuals can be determined by means of Southern blot analysis and a PCR test and can be correlated with the clinical phenotype. In *Chapter 4* an overview is given of the means of fragile X diagnosis before and after the identification of the *FMR1* gene. In *Publication 7.3*, several examples are given of DNA diagnosis in fragile X families by means of Southern blot analysis.

In the majority of fragile X patients an increased number of CGG repeats is found, with the absence of *FMR1* protein. It was however unclear, whether nearby genes were also involved in the etiology of the syndrome. We have identified one patient with a *de novo* point mutation in the *FMR1* gene (*Publication 7.5*). This point mutation results in the fragile X phenotype in this patient, strongly suggesting that the fragile X syndrome is a single gene disorder.

Questions as to when and how repeat expansion occurs are still unanswered. Several models of mechanisms through which repeat expansion might occur, and the timing of expansion are discussed in *Chapter 6*. However, none of the proposed mechanisms can completely explain the observed phenomena. Evidence accumulates that premutations are present in germ cells of carriers and patients, and that large expansions occur in a distinct time period in the early embryo. Mice in which the *Fmr1* gene has been knocked out, or mice with an enlarged CGG repeat in the *Fmr1* gene might serve as a model to study these issues and might give answers related to the function of the *FMR1* protein.

SAMENVATTING

Het doel van het onderzoek, beschreven in dit proefschrift, was het gen te isoleren en te karakteriseren dat betrokken is bij het ontstaan van het fragiele X syndroom. Om dit doel te bereiken hebben we gebruik gemaakt van "positional cloning", een techniek waarmee in de afgelopen jaren een flink aantal genen is gekloneerd (*Hoofdstuk 1*).

Het fragiele X syndroom, ook wel Martin-Bell syndroom genoemd, is een van de meest voorkomende erfelijke aandoeningen, met als voornaamste kenmerken geestelijke handicap en vergrote testes in mannelijke patiënten. Bij ongeveer 1 op de 1250 mannen komt deze ziekte voor, en hoewel het syndroom een X-gebonden overervingspatroon vertoont, zijn er ook normale mannelijke dragers, en is 50-70% van de vrouwelijke draagsters met een "volledige mutatie" op een van haar X chromosomen mild tot matig geestelijk gehandicapt. In *Hoofdstuk 2* worden een aantal historische aspecten over de "ontdekking" van het syndroom beschreven. In dit hoofdstuk wordt ook een overzicht gegeven van fysieke en gedragskenmerken van fragiele X patiënten. Tevens wordt het ongewone overervingspatroon van het syndroom besproken.

De prevalentie van het syndroom, de ernst van de ziekte in mannelijke patiënten en het belang om een defect op moleculair en biochemisch niveau te onderzoeken, startte de speurtocht naar het gen betrokken bij het fragiele X syndroom. In *Hoofdstuk 3* wordt een overzicht gegeven van de verschillende manieren van aanpak die zijn gebruikt door ons en andere onderzoeksgroepen in de jacht op dit gen.

De aanwezigheid van een "fragiele plaats" aan het uiteinde van de lange arm van het X chromosoom in cellen van mannelijke patiënten gaf aan, dat het gen gezocht moest worden in het Xq27-28 gebied van het X chromosoom. De aanname was, dat deze fragiele plaats geassocieerd was met het gen betrokken bij het fragiele X syndroom. Met behulp van genetische en fysische kartering werd het gebied waar het gen zich zou kunnen bevinden beperkt tot 2-2,5 megabasen. Probes uit deze regio, die zich vlak bij de fragiele plaats bevonden, werden door ons gebruikt om artificiële gistchromosomen (YACs) te isoleren die grote fragmenten bevatten afkomstig van het humane X chromosoom uit het Xq27-28 gebied (*Publikatie 7.1*). Door onze groep werd een YAC geïsoleerd van 475 kilobasen. Met behulp van de fluorescerende *in situ* hybridisatie techniek (FISH) werd aangetoond, dat deze YAC zich bevond ter plekke van de fragiele plaats (*Publikatie 7.2*). Deze YAC bevatte het CpG eiland dat preferentieel gemethyleerd was op de fragiele X chromosomen van

mannelijke patiënten en een deel van de vrouwelijke draagsters van het syndroom. Omdat CpG eilanden vaak worden gevonden aan het 5' einde van genen, bestond de mogelijkheid dat dit CpG eiland ook geassocieerd was met een zich in de buurt bevindend gen. Om te onderzoeken of op YAC 209G4 een gen gelegen was, werd hiervan een cosmidebank geconstrueerd. Cosmides met een humaan fragment werden gebruikt om een complementaire DNA bank, gemaakt van normaal foetaal humaan hersen mRNA, te screenen. Dit resulteerde in de isolatie van een gen, dat we *FMR1* (Fragiele X Mentale Retardatie 1) hebben genoemd (*Publikatie 7.1*). In het 5' onvertaalde gebied van het *FMR1* gen werd een CGG repeat aangetroffen. Deze repeat lag 250 baseparen distaal van het CpG eiland, dat gehypermethyleerd was in het DNA van fragiele X patiënten. Een van de cosmiden (cosmide 22.3), gebruikt om het gen te isoleren, bevatte dit CpG eiland en de CGG repeat. Met de FISH techniek werd aangetoond, dat dit cosmide signalen vertoonde aan beide zijden van de fragiele plaats (*Publikatie 7.2*), wat suggereerde dat het *FMR1* gen nauw geassocieerd moest zijn met de fragiele plaats.

Er is aangetoond, dat de CGG repeat is betrokken bij het ontstaan van het fragiele X syndroom. In het *FMR1* gen van normale individuen, wordt een gemiddeld aantal van 29 CGG repeats gevonden, variërend van 6-54 repeats. In fragiele X draagsters en dragers en in patiënten is deze repeat verlengd. In alle fenotypisch normale mannelijke dragers en een groot aantal van de fenotypisch normale draagsters, is deze verlenging klein. Deze kleine verlenging wordt een premutatie genoemd, omdat deze niet wordt geassocieerd met het fenotypische ziektebeeld. In fragiele X patiënten is de repeat verlengd tot meer dan 200 kopieën. Door dit grote aantal repeats vind methylering plaats van het promoter gebied van het gen, wat resulteert in afwezigheid van transcriptie en afwezigheid van *FMR1* eiwit. Deze grote verlenging wordt een volledige mutatie genoemd, omdat een repeat aantal boven de 200 kopieën wel wordt geassocieerd met het fenotypische ziektebeeld. In fragiele X families is de CGG repeat instabiel. Dit wil zeggen, dat het aantal repeats toe kan nemen van generatie op generatie, uiteindelijk resulterend in het fragiele X fenotype in een aantal nakomelingen. Deze ontdekking, dat verlenging van een trinucleotide repeat de moleculaire basis is van het fragiele X syndroom, heeft een nieuw mechanisme van overerving onthuld en is een fundamenteel nieuw concept in de genetica. Deze vondst werd spoedig gevolgd door de identificatie van andere ziekten, beschreven in *Hoofdstuk 5*, waarbij verlenging van een basen triplet verantwoordelijk is voor het ontstaan van een ziektebeeld. De genen, geassocieerd met deze ziekten, bevatten allemaal een polymorfe, CG rijke, trinucleotide repeat die instabiel is en verlengd is in DNA van patiënten.

Het *FMR1* gen komt hoog tot expressie in humane hersenen en testes, organen

betrokken bij de pathogenese van het syndroom, en codeert voor een mRNA van ongeveer 4,4 kilobasen. Het gen is geconserveerd gedurende de evolutie en wordt gevonden in organismen als de muis, gist en *Caenorhabditis elegans*. In de voorspelde *FMR1* aminozuur volgorde zijn een aantal domeinen geïdentificeerd, die kort geleden zijn gevonden in een aantal RNA bindende eiwitten. Dit suggereert dat ook het *FMR1* eiwit een RNA bindende functie kan hebben.
Ook hebben we aangetoond dat het *FMR1* gen onderhevig is aan alternatieve splicing (*Publikatie 7.4*). Dit zou aanleiding kunnen geven tot minimaal 12 verschillende *FMR1* mRNA transcripten en dus tot 12 mogelijke isovormen van het eiwit. Er zijn inderdaad een aantal verschillende isovormen van het *FMR1* eiwit gevonden in witte bloedcellen van normale individuen. De alternatieve splicing, beschreven in *Publikatie 7.4*, lijkt niet weefsel specifiek te zijn, omdat geen verschillen werden waargenomen tussen de verschillende onderzochte weefsels.
Het kloneren en karakteriseren van het *FMR1* gen en de identificatie van het gendefect in fragiele X patiënten, heeft de mogelijkheden voor diagnose van het syndroom aanzienlijk verbeterd. Nu kunnen verlenging van de repeat en het aantal CGG repeats in het *FMR1* gen van individuen bepaald worden met behulp van Southern blot analyse en een PCR test. Het aantal repeats kan gecorreleerd worden met het klinische beeld. In *Hoofdstuk 4* wordt een overzicht gegeven van de manieren waarop diagnostiek werd en wordt bedreven voor en na de isolatie van het *FMR1* gen. In *Publikatie 7.3* worden nog een aantal voorbeelden gegeven van DNA diagnostiek in fragiele X families met behulp van Southern blot analyse.
In bijna alle fragiele X patiënten wordt een verhoogd aantal CGG repeats gevonden in het *FMR1* gen, en afwezigheid van *FMR1* eiwit. Het was echter onduidelijk of ook andere genen in de buurt van *FMR1* betrokken waren bij het ontstaan van de ziekte. We hebben een patiënt geïdentificeerd met een *de novo* puntmutatie in het *FMR1* gen (*Publikatie 7.5*). Deze puntmutatie resulteert in het fragiele X fenotype in deze patiënt, wat suggereert dat het fragiele X syndroom door een enkel gen veroorzaakt wordt.
Vragen wat betreft wanneer en hoe repeat verlenging optreedt, zijn nog onbeantwoord. Verschillende modellen van mechanismen waardoor repeat verlenging zou kunnen optreden, of wanneer repeat verlenging plaats vindt, worden besproken in *Hoofdstuk 6*. Geen van de voorgestelde mechanismen kan echter alle waargenomen verschijnselen verklaren. Er zijn nu aanwijzingen, dat er zich slechts premutaties bevinden in geslachtscellen van draagsters, dragers en patiënten, en dat grote verlengingen van de repeat plaats vinden in somatische cellen in een bepaalde periode tijdens de vroege embryonale ontwikkeling.
Muizen, waarbij het *Fmr1* gen onwerkzaam is gemaakt, of muizen waarbij een

verlengde repeat in het *Fmr1* gen wordt gebracht, kunnen als model fungeren om deze zaken te bestuderen. Tevens zou bestudering van deze muizen antwoord kunnen geven op vragen betreffende de functie van het *FMR1* eiwit.

LIST OF PUBLICATIONS

Verkerk A.J.M.H, Pieretti M., Sutcliffe J.S., Fu Y., 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. and Warren S.T. Identification of a gene (*FMR-1*) containing a CGG repeat coincident with a fragile X breakpoint cluster region exhibiting length variation in fragile X syndrome. Cell 65, 905-914, 1991.

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. and Caskey C.T. Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. Cell 67, 1047-1058, 1991.

Verkerk A.J.M.H., Eussen B.H.J., van Hemel J.O. and Oostra B.A. The limited size of the fragile X site shown by fluorescence in situ hybridization. Am. J. Med. Genet. 43, 187-191, 1992.

Verkerk A.J.M.H., de Vries B.B.A., Niermeijer M.F., Fu Y.H., Nelson D.L., Warren S.T., Majoor-Krakauer D.F., Halley D.J.J. and Oostra B.A. An intragenic probe used for diagnostics in fragile X families. Am. J. Med. Genet. 43, 192-196, 1992.

Faust C.J., Verkerk A.J.H.M., Wilson P.J., Morris C.P., Hopwood J.J., Oostra B.A., and Herman G.E. Genetic mapping on the mouse X chromosome of human cDNA clones for the fragile X and Hunter syndromes. Genomics 12, 814-817, 1992.

Oostra B.A. and Verkerk A.J.M.H. The fragile X syndrome: isolation of the *FMR-1* gene and characterization of the fragile X mutation. Chromosoma 101, 381-387, 1992.

Willems P.J., Van Roy B., De Boulle K., Vits L., Reyniers E., Beck O., Dumon J.E., Verkerk A. and Oostra B.A. Segregation of the fragile X mutation from an affected male to his normal daughter. Hum. Mol. Genet. 1, 511-516, 1992.

De Vries L.B.A., Verkerk J.M.H., Niermeijer M.F, Oostra B.A. en Halley D.J.J. Het fragiele X syndroom: basaal effect, diagnostiek en erfelijkheidsadvies. Ned. Tijdschr. Geneeskd. 136, 1247-1251, 1992.

De Boulle K., Verkerk A.J.M.H., Reyniers E., Vits L., Hendrickx J., Van Roy B., Van den Bos F., De Graaff E., Oostra B.A. and Willems P.J. A point mutation in the *FMR-1* gene associated with mental retardation and the fragile X phenotype. Nature Genet. 3, 31-35, 1993.

De Vries B.B.A., Wiegers A.M., De Graaff E., Verkerk A.J.M.H., Van Hemel J.O., Halley D.J.J., Fryns J.P., Curfs L.M.G., Niermeijer M.F. and Oostra B.A. Mental status and fragile X expression in relation to *FMR-1* gene mutation. Eur. J. Hum. Genet. 1, 72-79, 1993.

Verkerk A.J.M.H., de Graaff E., De Boulle K., Eichler E.E., Konecki D.S., Reyniers E., Manca A., Poustka A., Willems P.J., Nelson D.L. and Oostra B.A. Alternative splicing in the fragile X gene *FMR-1*. Hum. Mol. Genet. 2, 399-404, 1993.

Reyniers E., Vits L., De Boulle K., Van Roy B., De Graaff E., Verkerk A.J.M.H., Darby J.K., Oostra B.A. and Willems P.J. The full mutation in the *FMR-1* gene of fragile X patients is absent in their sperm. Nature Genet. 4, 143-146, 1993.

Oostra B.A., Willems P.J. and Verkerk A.J.M.H. Fragile X syndrome: a growing gene. In: Genome analysis, volume 6. Davies K.E. and Tilghman S.M. (eds.). Cold Spring Harbor Laboratory Press, 1993.

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

Chiurazzi P., de Graaff E., Ng J., Verkerk A.J.M.H., Wolfson S., Fisch G.S., Kozak L., Neri G. and Oostra B.A. No apparent involvement of the *FMR1* gene in five patients with phenotypic manifestations of the fragile X syndrome. Am. J. Med. Genet., in press.

Verkerk A.J.M.H. and Oostra B.A. Fragile X linked mental retardation. In: Encyclopedia of Molecular Biology and Biotechnology. Meyers R.A. and Solomon S. (eds.), VCH publishers Inc., New York. In press.

Verkerk A.J.M.H. and Oostra B.A. Fragile X linked mental retardation syndrome. In: Encyclopedia of Molecular Biology: fundamentals and applications, volume 1 and 2. Meyers R.A. and Solomon S. (eds.), VCH publishers Inc., New York. In press.

CURRICULUM VITAE

28 september 1957	:	geboren te Tiel
juni 1973	:	eindexamen MAVO aan het Katholieke MAVO Centrum te 's-Gravenhage
juni 1975	:	eindexamen HAVO aan de R.K. Pedagogische Academie "Mariahoeve" te 's-Gravenhage
juni 1977	:	eindexamen VWO aan het Aloysiuscollege te 's-Gravenhage
sept. 1977 - aug. 1981	:	studie Slavistiek aan de Rijksuniversiteit Leiden
september 1981	:	aanvang studie biologie aan de Rijksuniversiteit Leiden
november 1984	:	kandidaatsexamen biologie (B1)
sept. 1984 - maart 1985	:	kandidaatsrichting medische biologie (B6)
april 1985 - aug. 1986	:	doctoraalhoofdvak Anthropogenetica, Prof. P.L. Pearson, Sylvius laboratoria, Leiden begeleider: Dr. G.J.B. van Ommen
febr. 1987 - dec. 1987	:	doctoraalhoofdvak afdeling cellulaire biochemie, Dr. H.L. Ploegh, Nederlands Kanker Instituut, Amsterdam begeleider: Drs. J.J. Neefjes
23 februari 1988	:	doctoraalexamen biologie
juli 1988 - sept. 1988	:	opzetten van het YAC kloneringssysteem bij de vakgroep Anthropogenetica, Sylvius laboratoria, Leiden, Dr. G.J.B. van Ommen en Dr. J. den Dunnen
1 okt. 1988 - 31 sept. 1993	:	promotieonderzoek, afdeling Klinische Genetica, Erasmus universiteit, Rotterdam bij Prof. Dr. H. Galjaard en Dr. B.A. Oostra

TOT SLOT

Vijf jaar onderzoek aan het fragiele X syndroom zijn voorbij gevlogen. Vijf jaar die in alle opzichten de moeite waard zijn geweest.

Ik herinner mij nog de eerste refereeravond van onze afdeling, waarbij ik professor Galjaard probeerde te overtuigen van het feit dat wij dat gen wel konden kloneren. Ik geloof dat ik toen nog niet zo goed besefte dat er een geduchte concurrentie strijd gaande was tussen de verschillende groepen in Europa, Amerika en Australië. Maar al met al (een beetje werk, een beetje geluk, en een goede samenwerking met de juiste mensen op het juiste moment) is het ons toch gelukt! Geachte professor Galjaard, ik denk dat we tevreden mogen zijn. Zonder uw (financiële) steun zou dit alles niet mogelijk zijn geweest. Hiervoor mijn hartelijke dank. Dat extra jaartje had ik nog net even nodig en gaf me de mogelijkheid om er een afgerond geheel van te maken.

BEN! Een gen kloneren, dat "zag ik wel zitten". Maar zonder jouw bezielende leiding en persoonlijke betrokkenheid zou dat niet gelukt zijn en zou ook dit boekje er nooit zijn gekomen. Ik denk dan ook dat dit "ons" boekje is geworden en dat we er best een beetje trots op mogen zijn. Je enthousiasme en energie waren onuitputtelijk (het heeft me altijd verbaasd dat je nooit moe was te krijgen), voor alles was een oplossing te vinden en je regelde alles wat los en vast zat. En toen het nodig was heb je ook jezelf nog in de strijd (of beter gezegd op het lab) geworpen (wat zo af en toe wel wat consequenties had voor de dagelijkse orde). Het waren soms chaotische tijden, waar ik altijd met veel plezier aan terug zal denken.

Professor Niermeijer, professor Grootegoed en professor van Ommen wil ik bedanken voor het kritisch lezen van het manuscript. Vooral professor Niermeijer bedank ik voor zijn bemoedigende woorden en altijd positieve manier van benaderen.

Beste Gert Jan. Ik ben je altijd zeer erkentelijk geweest voor het feit dat je mij, na mijn afstuderen, achter de typemachine hebt weggehaald (waar ik terecht was gekomen omdat ik met een AIO plaats geen genoegen wilde nemen). Er was nog ergens een "potje" waar voor drie maanden een "normaal" salaris in zat en ik stortte mij op het "gist kloneren". Na mijn aanstelling in Rotterdam ben ik nog een paar maanden blijven hangen op de afdeling Anthropogenetica in Leiden. Het waren goede tijden en ik heb nog nooit zo gelachen als toen met ons gist klonerings clubje: Marijke, Kurt (yeast cloning is fun!) en Martin (wanneer gaat het nu gebeuren?).

Rakesh Anand, thank you for, so patiently, introducing me into the wonders of

yeast cloning.
Johan en Petra, jullie wil ik bedanken voor het maken van alle gistblokjes en filters van de YAC's die uit Parijs binnen kwamen en waar, naar later bleek, onze voltreffer tussen zat. Door de hectische toestanden ontbreken jullie namen in de acknowledgements van hét artikel. Lau, het is onbegrijpelijk dat jouw manier van "stempelen" (op de grond en even het volle gewicht erop) niet heeft geleid tot de groei van allerlei ongewenste organismen.
I would like to thank our colleagues from Houston and Atlanta, David Nelson and Steve Warren and co-workers for an unique, pleasant and fruitful collaboration. The fast action and effort of many people, the daily communication by telephone and the almost daily exchange of material, have helped to make this project a big succes.
In de loop der jaren hebben een aantal studenten hun steentje bijgedragen aan het onderzoek: Shirley, Yvonne, Goedele (lunchen? Hè, néééé, ander keertje.) en Jessica (wat een teleurstelling, geen mutatie! Maar wel een publicatie, toch een hele prestatie!).
André (altijd positief, logica voorop), Arnold (niet te beschrijven), Rob (heb je wel in de goede vriezer gekeken? Ja, ik kan het echt niet vinden!), Joke K. en Coleta worden bedankt voor een prettige en waardevolle samenwerking. Elke vrijdagochtend zal ik nog even terugdenken aan de fraX werkbesprekingen (zijn we niet wat afgedwaald?). Ook de samenwerking met onze Belgische vrienden, Patrick (jij hebt zo je eigen wijze om dingen te formuleren en illustreren) en medewerkers was altijd uiterst plezierig.
Pauline, twee jaar lang hebben we goed samengewerkt en waren we op gang. Het is jammer dat je weg ging. Goede analisten zijn zeldzaam. Cathy en Robert, bedankt voor alle knipjes en hybjes. Cathy en Arjenne, bedankt voor de gezelligheid, vooral gedurende de lunchpauzes (doen we nog eens een koninginnedagje?). Toen jullie kwamen werd het eindelijk mode om boven te lunchen en zat ik niet meer in m'n eentje te pruimen.
Peter (Heut) en Bart (Brutus? Hoe kwam je daar nou aan?). Gedrieën zijn we gestart als promovendi op de 24e en hebben we het bijna vijf jaar lang met elkaar uitgehouden op dat kleine kamertje. Hoe is het mogelijk! (Kan het even stil zijn? Ik kan me hier niet concentreren. Wat is het koud hier zeg. Goedemiddag, mag ik van U een nummer in .....?) Mijn bijnaam kwam wat later (Kees???), nee, geen familie van!
De mensen van de afdeling diagnostiek, onder leiding van Dicky en Ans, bedank ik voor het produceren van alle mooie "plaatjes". Wout en Mieke, vooral jullie Southerns en PCR plaatjes hebben hun weg gevonden in gepubliceerde artikelen.
Jan Van Hemel, díe fraX chromosomen blijven het doen!

Bert de V., bedankt voor het bezoeken en "prikken" van talloze fragiele X patienten en hun familieleden. Tevens dank aan alle families die hun medewerking hebben verleend en zonder wie een groot deel van dit onderzoek niet had kunnen plaatsvinden. In het bijzonder worden de heer en mevrouw Meiners, Maarten, Norbert, Jeroen, Veronique en Marloes bedankt voor hun medewerking en gastvrijheid.
Rein (klachtenbalie; heb je nog een kopieerkaartje?), Mieke en Madeleine bedankt voor het doen van alle bestellingen. Piet (reparatiebalie; gaat toch goed met die oplaadbare batterijen?) bedankt voor al je kunstwerkjes (die giststempels worden een heuse curiositeit, ik zal ze goed bewaren). Peter van V. (soms wat moeilijk te bereiken) bedankt voor het oplossen van de computer en printer paniek. Op het secretariaat zorgden Jolanda en Jacqueline altijd voor een gevulde kast en was er vaak wel even tijd voor een gezellig praatje. Jopie, Joke B. en Elly worden bedankt voor het altijd maar weer produceren van schoon glaswerk en het autoclaveren van liters medium. Tom (driepoot), Ruud en Mirko, het viel niet altijd mee om aan mijn verzoeken te voldoen (dat kraagje moet er op; hier iets donkerder, daar iets lichter graag; ik heb het wel snel nodig), maar alles is altijd goed gelukt.
Bert E. en Esther, twee pilaren om op te steunen. Bert, jou heb ik leren kennen als een betrouwbaar collega, iemand waar je van op aan kunt. Er kon altijd nog wel een YACje bij (is het X of is het 19?) en je mooie in situ plaatjes mogen er zijn. Dat jij mijn paranimf zou zijn was vanaf het begin duidelijk!
Esther, je sequence talenten zijn alom bekend en werden (en worden nog steeds) zeer gewaardeerd. Je staat altijd klaar voor iedereen, hetgeen onder andere heeft geresulteerd in een groot aantal co-auteurschappen. Nog even doorzetten (PCR, PCR) en je bent ook klaar (jij taarten, ik soepen en salades?). Verder worden alle mensen van de 24e, 22e, 10e, 8e, 7e en 5e verdieping en de Westzeedijk, die hier niet met name zijn genoemd, bedankt voor hun belangstelling, samenwerking en gezelligheid.
Jan Maarten (heb je nog vierkante ogen?), jij wordt bedankt voor de vele (nachtelijke) uurtjes die je hebt gestoken in het maken van de omslag en van talloze (onbegrijpelijke) figuren. Het eindeloze geduld dat je daarbij wist op te brengen is benijdenswaardig. Je hulp heb ik zeer op prijs gesteld. En als je dacht dat je klaar was, dan waren er de volgende dag (na dag, na dag) nog wel wat veranderingen. Tar (onverstoorbaar tot de laatste pixel), enorm bedankt voor de adviezen en technische begeleiding bij het drukklaar maken van de omslag (dat zo'n simpel plaatje nog zo veel problemen kan geven).
Mijn familie, vrienden en kennissen wil ik bedanken voor hun interesse en enthousiasme. Met mijn uitleg, "ik werk aan een erfelijke ziekte, maar verder is het wat moeilijk om uit te leggen", nam (gelukkig) niet iedereen genoegen.

Stellingen behorende bij het proefschrift

The molecular basis of the fragile X syndrome

expansion of a trinucleotide repeat, a new mutational mechanism

Annemieke Verkerk

Rotterdam, 23 februari 1994

1. The anticipation of a triplet amplification in genes involved in diseases with the phenomenon of anticipation is strongly amplified.

La Spada et al. Nature Genet. 2, 301-304, 1992.
Harley et al. Nature 355, 545-546, 1992.
Tsilfidis et al. Nature Genet. 1, 192-195, 1992.
Snell et al. Nature Genet. 4, 393-397, 1993.
Andrew et al. Nature Genet. 4, 398-403, 1993.
Orr et al. Nature Genet. 4, 221-226, 1993.
Pulst et al. Nature Genet. 5, 8-10, 1993.
Koide et al. Nature Genet. 6, 9-13, 1994.
Nagafuchi et al. Nature Genet. 6, 14-18, 1994.

2. De tegenstrijdige resultaten gerapporteerd door Fu *et al.* en Sabouri *et al.* (respectievelijk verlaagde en verhoogde expressie van Mt-Pk mRNA in patiënten met myotone dystrofie) laten zien dat het kwantificeren van mRNA nog veel te wensen over laat.

Fu et al. Science 260, 235-238, 1993.
Sabouri et al. Nature Genet. 4, 233-238, 1993.

3. Bij het opsporen van vermoedelijk grote mutaties, zoals gevonden in het Factor VIII gen bij ernstig aangedane hemofilie patiënten, (over)wint een techniek als pulsed field gel electrophorese de RACE.

Lakich et al. Nature Genet. 5, 236-241, 1993.
Naylor et al. Hum. Mol. Genet. 2, 1773-1778, 1993.

4. Het verschil in interpretatie door Hinds *et al.* en Bächner *et al.* van resultaten, verkregen met *FMR1* mRNA *in situ* hybridizatie, geeft aan dat ook voor moleculair biologen een gedegen morfologische kennis onontbeerlijk is.

Hinds et al. Nature Genet. 3, 36-43, 1993.
Bächner et al. Nature Genet. 4, 115-116, 1993.

5. "The case of the cataractous kangaroo" is een goed voorbeeld van de betrekkelijkheid van de term "single gene disorder" waar het gaat om complexe organismen in een complexe wereld.

Stephens et al. Med. J. Aust. 2, 910-911, 1974.

6. Streptomycine geïnduceerde perceptieve slechthorendheid kan mede berusten op een erfelijke predispositie vastgelegd in het mitochondriële DNA.

Hu et al. J. Med. Genet. 28, 79-83, 1991.

Prezant et al. Nature Genet. 4, 289-294, 1993.

7. Dragerschap voor *Pseudomonas cepacia* leidt tot een bacteriologische barrière tussen patiënten met cystische fibrose en hun familieleden.

Govan et al. Lancet 342, 15-19, 1993.

8. De kans op een epileptische aanval bij het spelen van videospelletjes wordt aanzienlijk verminderd door er enige afstand van te nemen.

9. The Fragile X Mental Retardation gene was isolated by a Fast Moving Research group from the Faculty of Medicine in Rotterdam.

10. Het ophangen van een rookgordijn is geen oplossing om één ruimte te verdelen in een rokers en een niet-rokers gedeelte.

11. Als iedere automobilist de beschikking zou hebben over een boordkanon, dan zouden nog weinig weggebruikers op hun werk arriveren.

12. Bewuste kinderloosheid reflecteert niet per definitie een antipathie tegen kinderen.