CGG Repeat Instability and FXR Proteins

C.J.M. Ubink-Bontekoe


Gedrukt bij Offsetdrukkerij Ridderprint B.V., Ridderkerk

CGG Repeat Instability and FXR Proteins

CGG repeat instabiliteit en FXR eiwitten

Proefschrift

ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de Rector Magnificus
Prof. dr. ir. J.H. van Bemmel
en volgens besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op
woensdag 13 juni 2001 om 13.45 uur
door

Carola Jacoba Maria Ubink-Bontekoe
geboren te Poortugaal

Promotiecommissie

Promotor: Prof. dr. B.A. Oostra

Overige leden: Prof. dr. J.A. Grootegoed
Prof. dr. R. Kanaar
Prof. dr. C.I. de Zeeuw

Omslag: F.R. Ubink

The studies described in this thesis were performed in the Department of Clinical Genetics at the Erasmus University Rotterdam.
This project was financially supported by a grant of the Dutch Science Foundation (N.W.O). through GB-MW (Medical Sciences).

Contents

Chapter 1

The fragile X syndrome

1.1 The fragile X genotype/phenotype

The name "fragile X syndrome" refers to the association found between a fragile site at the long arm of the X chromosome, Xq27.3 (FRAXA), and a form of X-linked mental retardation (Lubs 1969). Previously, several other names were used to describe this syndrome, like the Martin-Bell syndrome and the Gillian Turner-type X-linked mental deficiency syndrome (Martin *et al.* 1943; Turner *et al.* 1975; Turner *et al.* 1980). The prevalence of the syndrome is estimated to be 1 in 4000 males and 1 in 6000 females based on molecular studies (Turner *et al.* 1996; De Vries *et al.* 1997). The fragile site at Xq27.3 was found in patients' cells when cultured in folic acid depleted medium. The expression of this fragile X was extensively studied to understand the causative mechanism of the fragile X syndrome, but without success. In 1991, the gene responsible for the fragile X syndrome, the fragile X mental retardation gene, *FMR1*, was cloned (Oberlé *et al.* 1991; Verkerk *et al.* 1991; Yu *et al.* 1991). The gene defect causing the inactivation of the gene, and thus the absence of the *FMR1* protein, FMRP, was an expanded CGG repeat present in the 5' UTR of *FMR1*. In patients, this CGG repeat varies between 200 and more than 1000 CGG triplets. The fragile X syndrome was the first discovered inherited disease caused by the expansion of a trinucleotide repeat. Nowadays, a whole class of so-called trinucleotide repeat diseases is known.

The mental retardation is the main characteristic of the fragile X syndrome. All patients show mental impairment, although the mental retardation varies from moderate to severe in male patients. Because the fragile X syndrome is an X-linked disease, it is unexpected that in 35 to 53% of the females with a full mutation (hereinafter: full mutation females) a phenotype is observed (Hagerman *et al.* 1992). Besides the mental retardation other physical and behavioural characteristics are described. Physical characteristics include macroorchidism in males and facial features, such as a long face and large protruding ears. These characteristics develop or are most prominent in adult patients. Connective tissue abnormalities, such as hyperextensive joints, hand calluses and flat feet, are also observed (reviewed in thesis of de Vries, 1997) (De Vries 1997).

The most frequent behavioural problems in fragile X patients are hyperactivity, decreased attention span, speech and language problems, and a number of autistiform features including handflapping, poor eye contact and gaze avoidance.

Although the typical fragile X phenotype has many more characteristics than the mental retardation itself, many clinicians have difficulties recognising the phenotype. Reasons are the relatively low prevalence of the syndrome, and the difficulty to recognise a number of the facial characteristics and the macroorchidism during childhood. A number of

behavioural problems like attention deficit and hyperactivity are most pronounced at a young age during childhood. At this age most patients are referred to clinicians, but these symptoms as such are not specific for the fragile X syndrome. However, the combination of features increases the likelihood of the fragile X diagnosis by more than tenfold (De Vries *et al.* 1999).

1.2 *FMR1*

In 1991 the gene involved in the fragile X syndrome was cloned, the fragile X mental retardation gene, *FMR1*. *FMR1* spans approximately 38 kb of genomic sequence. The longest mRNA consists of approximately 4.0 kb and the coding region is approximately 1.9 kb. Alternative splicing is observed for exons 12, 14, 15 and 17 (Ashley *et al.* 1993b; Verkerk *et al.* 1993). This generates different mRNAs coding for protein isoforms. By Western blotting using antibodies against FMRP several proteins with a molecular mass around 70-80 kDa are detected. The detected bands most likely represent different isoforms. The 5' and 3' UTRs are, respectively, 0.2 kb and 1.8 kb. In the 5' UTR a CGG trinucleotide repeat is located, 69 bp upstream of the start ATG. The promoter region needed for proper FMRP expression *in vivo* is thought to reside within a 2.8 kb fragment including the 5' UTR of *FMR1* and the CGG repeat. Deletion mapping identified a minimal promoter region allowing expression in transfection experiments. This minimal promoter region was defined as the 466 bp *PstI*/*Xho*I fragment containing the CpG island and the CGG repeat (Hwu *et al.* 1993).

FMR1 is highly conserved among different species. Nucleotide identity between human and mouse is 95 % in the open reading frame, amino acid identity is 97 % (Ashley *et al.* 1993b). Also in chicken, *Xenopus leavis* and monkey, significant homologies are found (Eichler *et al.* 1995; Siomi *et al.* 1995; Price *et al.* 1996). Previously, it was suggested that homologues were also present in *C. elegans* and yeast (Verkerk *et al.* 1991). For both species the total genome is sequenced by now, and no homologues could be identified. In *Drosophila,* one *FMR1*/*FXR* homologue was identified (Wan *et al.* 2000). The CGG repeat located in the 5' UTR of *FMR1* is also conserved between species, although the length of the CGG repeat is smaller compared to the CGG repeat in humans (Deelen *et al.* 1994; Eichler *et al.* 1995). The observed homology between species of the 5' UTR and the 3' UTR is surprising and might indicate that they have regulatory functions. For the homology found between human and mouse intron 1, the presence of regulatory elements was also suggested (Kumari *et al.* 2000).

With the cloning of *FMR1* a new mutational mechanism was discovered, the expansion of trinucleotide repeats. In humans, the alleles are divided into three groups: the normal, the premutation and the full mutation alleles. Normal alleles range between 5 and 50 repeat units. The most common alleles have 29 or 30 triplets (Fu *et al.* 1991). These alleles are stably transmitted to offspring. The allele in the parent has the same size in the children and there is normal FMRP expression.

Premutation alleles range between 50 and 200 triplets. These repeat sizes can change upon transmission. Both expansions and contractions occur, depending among others on the length of the premutation. An increasing length of the premutation increases the chance of expansions. Premutations do show instability upon transmission to next generations, but they do not block FMRP expression completely (Devys *et al.* 1993; Feng *et al.* 1995a; Feng *et al.* 1995b; Tassone *et al.* 2000c). Therefore, premutations do not result in the fragile X phenotype.

Upon maternal transmission the CGG repeat can expand above the threshold of 200 CGG triplets, called full mutations (Kremer *et al.* 1991; Oberlé *et al.* 1991). Full mutations coincide with methylation of the CGG repeat and the surrounding CpG island, including the promoter region (Hansen *et al.* 1992; Hornstra *et al.* 1993). This methylation blocks the transcription of *FMR1*, and thus results in the absence of FMRP (Pieretti *et al.* 1991; Verheij *et al.* 1993). Due to the absence of FMRP, these full mutations are causing the fragile X phenotype. The observation that full mutation males with an unmethylated promoter are not affected provided evidence that the repeat expansion alone is not sufficient to cause the fragile X phenotype (McConkie-Rosell *et al.* 1993; Hagerman *et al.* 1994; Smeets *et al.* 1995; de Vries *et al.* 1996; Wohrle *et al.* 1996; Wohrle *et al.* 1998; Taylor *et al.* 1999). Methylation of the promoter region plays an important role in FMRP expression. Chiurazzi *et al* have demonstrated that demethylation of the *FMR1* promoter re-activates *FMR1* transcription and FMRP expression (Chiurazzi *et al.* 1998; Chiurazzi *et al.* 1999; Coffee *et al.* 1999).

A limited number of patients have been described without a CGG repeat expansion. Most of these patients showed the fragile X phenotype because of a deletion of the promoter region, resulting in the absence of FMRP (De Graaff *et al.* 1995). These deletions range from a few hundred base pairs including the CGG repeat and flanking regulatory elements up to a few megabases including FMR1 completely (reviewed in thesis of de Graaff, 1996) (De Graaff *et al.* 1995). Only one patient is described with a mutated protein (De Boulle *et al.* 1993). This severely affected patient contains an Ile304Asn mutation (see Chapter 3.2).

1.3 Mode of inheritance

The fragile X syndrome does not inherit as a regular X-linked trait. A number of observations in fragile X families do not fit the classic Mendelian mode of inheritance. Once the fragile site was found to be associated with transmission of the mutant X chromosome, a lot of effort has been put in trying to explain the inheritance of this fragile site, but with little success. It was not before the cloning of *FMR1* that the mode of inheritance was understood.

In normal X-linked disorders all males carrying the mutation are affected. In families with the fragile X syndrome males were found that transmitted the mutated X chromosome although they were not affected. These males were called "normal transmitting males". These males do not express the fragile site. Once *FMR1* was cloned, these males were recognised as premutation carriers. Expansion of the premutation allele could give rise to full mutation alleles in subsequent generations when transmitted via the female germ line.

Another peculiar observation is the phenomenon known as the Sherman (Sherman *et al.* 1985; Fu *et al.* 1991). The risk of expressing the fragile X phenotype depends on the position in the pedigree and increases in subsequent generations. Nowadays, this is explained by the increasing length of the premutation allele. When the size of the premutation increases over generations, the chance of an expansion to a full mutation increases (Fisch *et al.* 1995). This shows that the cloning of *FMR1* gave us more insight into the mode of inheritance of the CGG repeat and thus the fragile X syndrome. However, other questions are still not answered. In both males and females, premutations behave unstable, and for both males and females expansions and contractions are observed. It is not understood why expansions to full mutation only occur upon female transmission. Males never transmit a full mutation (Oberlé *et al.* 1991). The reason for this is still under debate.

1.4 Fragile X diagnostics

The identification of the mutational mechanism responsible for the fragile X syndrome has provided reliable methods for diagnostics. The length of the CGG repeat can be determined by Southern blot analysis or PCR (Oostra *et al.* 1993). Normal alleles will give a fragment of 5.2 kb. Longer fragments indicate the presence of a pre- or full mutation. By using an additional methylation sensitive enzyme, like *Eag*I or *Bss*HII, the length and the methylation status can be determined (Fu *et al.* 1991). This is applied to distinguish large premutations from small full mutations. With PCR analysis the exact length of normal and premutation carriers can be determined. This can be informative to identify carrier females and to identify small changes in repeat length in successive generations. Using standard PCR

protocols it is impossible to identify full mutations. However, a recently developed PCR-based method (Fragile X Size Polymorphism assay, Perkin Elmer Biosystems) claims to identify full mutations.

To perform prenatal testing, DNA is isolated from chorionic villi or amniotic fluid cells. By Southern blot analysis the repeat length can be determined. Methylation analysis of chorionic villi samples revealed a discrepancy between the embryonic and the extra-embryonic tissues in a number of cases. Some reports describe no or less methylation in the chorionic villi (Hirst *et al.* 1991; Rousseau *et al.* 1991; Sutherland *et al.* 1991; Devys *et al.* 1992; Sutcliffe *et al.* 1992). Other reports do find methylation in the sampled chorionic villi (Suzumori *et al.* 1993). Since in the fetal tissues the methylation is complete, this indicates that the methylation pattern found in chorionic villi is not always reflecting the methylation pattern of the fetal tissues. Therefore, the methylation status of the chorionic villi cannot give additional information in prenatal diagnosis. Only the repeat length has to be considered. The methylation status for female fetuses is even more complex, because of additional methylation due to X-inactivation. In humans, not much is known about X-inactivation in chorionic villi. This is discussed in more detail in Chapter 5.3 of this thesis.

There is also a diagnostic test available based on the absence of FMRP in fragile X patients. By using an antibody test, FMRP can be detected in lymphocytes in blood smears of control individuals (Willemsen *et al.* 1995; Willemsen *et al.* 1997b). In the blood smears of male fragile X patients no FMRP can be detected in the lymphocytes. In full mutation females approximately 50 % of the lymphocytes do not express FMRP due to X-inactivation of the normal X chromosome.

This antibody test has also been used for prenatal diagnosis, although the number of samples is still limited (Willemsen *et al.* 1997a). The test was performed on chorionic villi, as well as amniotic fluid cells. FMRP could be detected in cytotrophoblasts of control individuals, but not in chorionic villi samples of full mutation male fetuses. The gestational age of the chorionic villi has to be taken into account. In samples with a gestational age below 12.5 weeks FMRP might still be present. Most likely, FMRP expression is not completely inactivated at that time. This might correlate to the hypomethylation of chorionic villi seen at 11 weeks of gestational age. The prenatal protein test has also been tested on chorionic villi of full mutation female fetuses. The results are discussed in Chapter 5.3 of this thesis.

Quite a high percentage of fragile X males shows a mosaic DNA pattern. Besides the methylated full mutation, a percentage of their blood cells contain an unmethylated full- or premutation. These unmethylated full mutation or premutation cells express FMRP. It has to

be taken into account that the percentage of cells expressing FMRP in blood might be different from the percentage FMRP-expressing cells in brain. The same applies to full mutation females. Approximately 50 to 75 % of these females show some mental impairment. Assuming a random X-inactivation, 50 % of the cells will express FMRP. This percentage is probably not enough to allow normal cognitive functioning. It has been hypothesized that more than 60 % of the cells need to express FMRP for a normal cognitive functioning. A clear correlation between IQ and FMRP expression in blood could not be found. Recently, the antibody test was also performed on hairbulbs (Willemsen *et al.* 1999). The percentage of hairbulbs expressing FMRP in full mutation females revealed a correlation with the IQ. Probably, the hairbulbs reflect the situation in brain better than the FMRP pattern in blood. In the future this test might be used as a predictive test for the development of full mutation females (Willemsen *et al.* 2000).

Chapter 2

CGG trinucleotide repeat instability

2.1 Mechanism of CGG repeat expansion

The fragile X syndrome was the first identified trinucleotide repeat disease. The CGG repeat in the 5' UTR of *FMR1* proved to be highly polymorphic. The length of the CGG repeat correlates with the extent of amplification in the next generation. There is overlap between the high end of normal alleles and the low end of premutation alleles. The observation of AGG interruptions led to the suggestion that these interruptions might provide stability to the repeat, and that instability might result from loss of AGG interruptions (Kunst *et al.* 1994). Analysis of the AGG interruptions showed that in the most common normal alleles AGG interruptions are found downstream of 9 or 10 CGG triplets (Eichler *et al.* 1996). In longer repeats relatively less AGG interruptions are found. Loss of these AGG interruptions always occurs at the 3' end of the repeat tract. The longest pure CGG tract found within a pre- or full mutation is always located at the most 3' end. It is believed that a pure CGG tract at the 3' end of more than 38 repeats is enough to cause instability (Eichler *et al.* 1994). The reason for this polar variation might involve the mechanisms underlying the repeat instability.

Besides the AGG interruptions, other determinants involved in fragile X repeat instability were observed. Most of these determinants were identified in studies with yeast and *E. coli*. These determinants include the direction of replication, the genetic background (including repair systems), transcription, and growth conditions (Shimizu *et al.* 1996; Wells 1996; White *et al.* 1999). A number of haplotypes are identified that coincides with premutations. This founder effect might indicate the presence of cis-acting elements, reflecting the mutational mechanism or pathway. Studies in a large unaffected and fragile X African-American population suggest that these haplotypes probably reflect the mutational history (Crawford *et al.* 2000a; Crawford *et al.* 2000c), and are not necessarily linked to cis-acting elements. Different polymorphisms in the region of the repeat have been studied, but none of these could be identified as a cis-acting element (Gunter *et al.* 1998; Crawford *et al.* 2000b).

The exact mechanism causing repeat instability is not understood. For all short-tandem repeats small changes over generations occur. These changes involve one or two repeat units and result in what is called a "polymorphism". These variabilities are observed within the normal alleles. Premutations show more instability compared to other polymorphisms. Contractions and expansions involving ten to twenty triplets occur quite frequently. Once a certain threshold is passed, the repeats behave very unstable preferentially resulting in large expansions. This expansion-biased instability is sometimes referred to as "dynamic

mutations" (Richards *et al.* 1992). These dynamic mutations most likely result from other mechanisms than the polymorphic changes.

Repeat instability is thought to be instigated by the formation of d(CGG) secondary structures. Several structures of the CGG repeat, like hairpins and tetrahelical structures have been suggested. The stability of the different structures seems to differ between both templates. Pausing of DNA polymerases at both hairpin and tetrahelical structures was observed *in vitro* as well as *in vivo* (Kang *et al.* 1995; Usdin *et al.* 1995; Samadashwily *et al.* 1997). This stalling of DNA polymerase might increase the probability of DNA polymerase slippage. According to the "slippage model" this can result in expansion of the CGG repeat. The slippage model assumes the formation of hairpins within the CGG repeat. These hairpins can form when during replication DNA polymerase slippage occurs. When such a hairpin forms on the Okazaki fragments of the lagging strand reannealing might be at a different position within the repeat. This leaves a gap that will be repaired resulting in a slightly longer repeat than the original template. Alternatively, a hairpin on the template of the leading strand will result in a shorter repeat since the DNA polymerase will skip a part of the CGG repeat. This slippage model can explain relatively small changes that occur within the premutation group. The formation of hairpins or tetrahelical structures (dimerization of hairpins) was found to be changed by AGG interruptions (Pearson *et al.* 1998; Weisman-Shomer *et al.* 2000). This was shown in experiments with CGG repeat containing plasmids in *E. coli* and yeast. An increasing amount of evidence obtained from studies in yeast and *E. coli* favour the slippage model, although other models, like unequal meiotic sister chromatid exchange and/or gene conversion are difficult to rule out. The involvement of repair mechanisms is studied in different systems. Studies are performed in yeast, *E. coli* and to a lesser extent in mice. Especially proteins involved in mismatch repair and FEN1, a protein involved in long-patch base excision repair, are thought to play a role in repeat instability (Freudenreich *et al.* 1998; Parniewski *et al.* 2000).

2.2 Timing of CGG repeat expansion

The timing of the repeat expansion and the difference observed for male and female transmission of the repeat, is an intriguing feature of the mode of inheritance of the fragile X syndrome. How and when expansions from premutation to full mutation occur is not known. The repeat length can vary from one generation to the next. This means that instability has to occur somewhere during gametogenesis in the parent and/ or early embryogenesis of the child. Two models have been postulated, based on the presence of either a full mutation or a premutation in the oocyte.

The first model assumes a full mutation to be present in the oocyte. Expansion has to occur somewhere before or during female meiosis. After fertilisation, the zygote will develop into an embryo with a full mutation present in all somatic and germ cells. However, in some cells regression can occur. These regression events are thought to occur randomly, resulting in cells with a premutation. This explains the high percentage of full mutation males that are mosaic for a premutation.

To explain why full mutations are only transmitted via the female germ line, a kind of selection against full mutations is proposed for the male germ line. This idea is supported by the observation that in sperm of full mutation males only premutations were found (Reyniers *et al.* 1993). Immunohistochemistry of testes of full mutation fetuses showed that during development only a few primordial germ cells produce FMRP. This suggests that a regression event has occurred in those cells. During development the number of primordial germ cells expressing FMRP in full mutation male fetuses increases (personal communication, Willemsen). There seems to be a kind of selection against full mutations or in favour of premutations. This selection can be based on either the presence of FMRP or the absence of the expanded CGG repeat. The latter is the most favoured hypothesis, as the *FMR1* knockout mice lacking FMRP completely are fertile. Recently, it was shown that culturing fibroblasts from a male who presented heterogeneous unmethylated expansions in the pre- and full mutation range led to accumulation of shorter alleles, thereby increasing the FMRP expression (Salat *et al.* 2000). The same might account for the primordial germ cells with a premutation. If they proliferate faster, this might result in overgrowing the primordial germ cells with a full mutation. This hypothesis is supported by the observed delay in replication in cells containing an expanded CGG repeat (Hansen *et al.* 1993; Hansen *et al.* 1997; Samadashwily *et al.* 1997).

Difficult to explain by the first model is that the premutation in mosaic individuals and the premutation present in sperm of full mutation males is mostly present as one distinct band, indicating the presence of one size of premutation. This suggests that these premutation cells are the result from one regression event, or that full mutations contracted to the same length in different cells.

The second model assumes the presence of a premutation in the oocyte, which expands to a full mutation during early embryogenesis. Expansion is thought to occur only in somatic cells after separation of the germ line (Devys *et al.* 1992; Wöhrle *et al.* 1993). This would result in premutation alleles in both the male and the female germ line. This model easily explains that in mosaics and in sperm of full mutation males only one premutation size is present. However, in oocytes from full mutation female fetuses, full mutations were already

present (Malter *et al.* 1997). This observation makes model two less likely. However, this does not give direct information about the transition from a premutation to a full mutation. Therefore, oocytes of a premutation female should be studied. For obvious reasons, it is difficult to obtain these oocytes. A simulation study showed that if transition to full mutation is postzygotic, one should expect a much higher proportion of mosaics in offspring of mothers with small premutations (Moutou *et al.* 1997). However, a transition at an early morula stage (before day 3) cannot be excluded.

Because of the presence of premutations in sperm of full mutation males and the presence of full mutations in oocytes of full mutation females, model one is the most favoured. To study the exact timing of repeat expansion an animal model will be required. The availability of an animal model to study the CGG repeat instability will allow us to study the repeat length at different time points during development.

2.3 Other trinucleotide repeat diseases

Disorder	Repeat	Location	Normal	Affected	Sex-bias
Kennedy's disease	CAG	Coding	9-36	38-62	Paternal
Huntington's disease	CAG	Coding	6-35	36-121	Paternal
SCA1	CAG	Coding	6-35	40-81	Paternal
SCA2	CAG	Coding	14-32	33-77	Paternal
SCA3	CAG	Coding	12-40	67-82	
SCA6	CAG	Coding	4-17	20-30	Paternal
SCA7	CAG	Coding	7-17	38-130	Paternal
SCA8	CTG	Coding mRNA		110-250	
SCA12	CAG	5'UTR			
DRPLA	CAG	Coding	3-36	49-88	Paternal
Myotonic dystrophy	CTG	3'UTR	5-30	>50	Maternal
FRAXE mental retardation	GCC	5'UTR	6-25	<130	
Fragile X syndrome	CGG	5'UTR	6-50	>200	Maternal
Friedreich's ataxia	GAA	Intron	7-22	>200	

Table 1. Summary of features associated with trinucleotide repeat diseases. Normal and affected depict the sizes of the repeat found in normal and affected individuals respectively. Sex-bias refers to the sex of the transmitting parent in which the repeat is most often expanded.

In the last ten years more than a dozen trinucleotide repeat diseases have been identified (reviewed in Baldi *et al.* 1999; Cummings *et al.* 2000). Different trinucleotide repeats are involved in these diseases, and they are located in different regions of the responsible genes. In Table 1 the known trinucleotide repeat diseases are listed together with the

involved trinucleotide repeat, the localisation of the repeat and the different classes of alleles.

Despite these differences, several features are shared amongst the trinucleotide repeat diseases. All trinucleotide repeat diseases show both somatic and germ line instability; expansions are more frequent than contractions upon transmissions. Secondly, anticipation is observed. The earlier age of onset and the increasing severity of phenotype correlate with a larger repeat length. Friedreich's ataxia and fragile X syndrome are the only trinucleotide repeat diseases that do not show anticipation. For fragile X syndrome it is observed that an increasing repeat length gives a higher chance of expansion to a full mutation allele. This observation is known as the Sherman paradox, but sometimes it is referred to as anticipation. For many of these disorders paternal transmission has a higher risk of expansion, the reason for this is not known.

The trinucleotide repeat disorders can be divided into two groups, based on the localisation of the repeat in the coding or the non-coding region of the involved gene. All coding trinucleotide repeat diseases involve a CAG repeat, which codes for a polyglutamine stretch. These disorders are characterised by progressive dysfunction beginning mostly in mid-life and they result in severe neurodegeneration. Animal models and tissue culture systems demonstrated that a direct effect of the expanded glutamine stretches is at least partially responsible for this pathogenesis (Lin *et al.* 1999; Zoghbi *et al.* 1999). The expansions, observed in these diseases, are smaller in size and variation than those observed in non-coding trinucleotide repeat diseases. This raises the question whether the same mechanisms play a role in these disorders. It is possible that other mechanisms are involved in these relatively small changes.

Most of the non-coding trinucleotide repeat disorders are caused by (partial) loss of function of the involved gene. The trinucleotide repeat can be located in the 5' UTR, as is seen in the fragile X syndrome, fragile XE mental retardation (Knight *et al.* 1993; Chakrabarti *et al.* 1996) and SCA12 (Holmes *et al.* 1999). For myotonic dystrophy the CTG trinucleotide repeat is located in the 3' UTR of protein kinase gene, *DMPK* (Brook *et al.* 1992; Fu *et al.* 1992; Mahadevan *et al.* 1992). However, the repeat might also influence flanking genes, like *DMAHP* and *DMWD*. It is not known how the repeat expansion influences the different genes, and how the different expression levels of these genes cause the disease phenotype (Groenen *et al.* 1998). For Friedreich's ataxia the GAA trinucleotide repeat is located in an intron of *X25*, the frataxin gene (Campuzano *et al.* 1996). Expansion of this repeat reduces the expression levels of frataxin, suggesting that Friedreich's ataxia results from a partial loss of frataxin function.

More trinucleotide repeats are present in the human genome. In total 5 folate-sensitive fragile sites were cloned. Only FRAXA and FRAXE were found to be involved in human diseases as described before. Three other fragile sites, FRAXF, FRA11B and FRA16A were also cloned. All these sites showed an expanded CGG/CCG repeat. This indicates that the formation of the folic-acid sensitive fragile site is induced by specific properties of the CGG repeat. Most likely these specific properties involve the DNA structures formed by these repeats. The cytogenetic expression of the fragile site in two unmethylated full mutation males showed that the expression of the fragile site was not dependent of the methylation of the repeat. Besides these folic-acid sensitive sites also other fragile sites are described. Most likely all these fragile sites represent regions in the genome that under certain conditions are capable of forming extreme DNA structures that appear as fragile sites.

2.4 Mouse models for trinucleotide repeat instability

How and when repeat instability occurs, is one of the most intriguing questions for all trinucleotide repeat diseases. To study the mechanisms in and timing of repeat instability, an animal model is required. Only by using an animal model it will be possible to study repeat instability during the development of the germ line and all developmental stages of early embryogenesis. Several transgenic mouse models were generated to study the behaviour of the CGG repeat involved in the fragile X syndrome. The construct we used contained the *FMR1* promoter region (*Eco*RI-*Nhe*I fragment) fused to a *LacZ* reporter gene (Hergersberg *et al.* 1995). The CGG repeat present in this promoter was an $(CGG)_{81}$ allele, wherein two interruptions were present. Three independent lines were generated and in total more than 300 of both maternal and paternal transmissions were studied. No major instabilities were found (Bontekoe *et al.* 1997). Lavedan, using a $(CGG)_{88}$ (also including two interruptions), studied different lines, but also found no instabilities (Lavedan *et al.* 1997). An explanation for the observed stability could be the presence of interruptions within the CGG repeat. However, a mouse model containing a pure CGG repeat also showed no major instabilities (Lavedan *et al.* 1998).

Several explanations were put forward to explain the observed stability of the CGG repeats in the different mouse models. Most of these explanations involve the genomic localisation of the CGG repeat. The expanded CGG repeat is not present on the X chromosome in the endogenous *FMR1*. Flanking cis-acting elements required for repeat instability might be missing on the transgenes. Therefore, the best animal model would be to replace the $(CGG)_8$ allele present in the endogenous *Fmr1* by a human expanded allele. Cloning and characterisation of the mouse *Fmr1* promoter region showed that the human and the mouse

promoter regions share the regulatory elements that were previously identified by footprinting analysis (Drouin *et al.* 1997; Schwemmle *et al.* 1997; Schwemmle 1999). The CGG repeat present in the mouse promoter was only 8-12 triplets. This repeat was, allowing minimal changes to the flanking sequence, replaced by a human *Nhe*I-*Xho*I fragment containing a $(CGG)_{98}$ allele using a homologous recombination technique. For this CGG trinucleotide repeat moderate instability is found mainly upon male transmission. The exact data are discussed in Chapter 5.2 of this thesis.

Repeat	Reference	Instability
CAG	Kennedy et al. 2000	Yes
CAG	Kovtun et al. 2000	Yes
CAG	Lorenzetti et al. 2000	Yes
CTG	Seznec et al. 2000	Yes
CTG	Fortune et al. 2000	Yes
CAG	Sato et al. 1999	Yes
CAG	Bingham et al. 1995	No
CAG	Goldberg et al. 1996	No
CAG	Burright et al. 1995	No
CTG	Gourdon et al. 1997	Yes
CGG	Bontekoe et al. 1997	No
CTG	Monckton et al. 1997	Yes
CGG	Lavedan et al. 1997	No
CGG	Lavedan et al. 1998	No
CAG	Kaytor et al. 1997	Yes
CAG	Wheeler et al. 1999	Yes
CAG	Mangiarini et al. 1997	Yes
CAG	LaSpada et al. 1998	Yes
CAG	Shelbourne et al. 1999	Yes
CAG	Manley et al. 1999	Yes
CTG	Lia et al. 1998	Yes

Table 2. Mouse models generated to study trinucleotide repeat instability are listed, together with the involved trinucleotide and information whether instability was observed. The rate and percentage of instability was found to be different for the different models.

Mouse models containing repeats involved in other trinucleotide repeat diseases are also described. Most of these mouse models are mice transgenic for an expanded repeat, but also mouse models were described where the endogenous mouse trinucleotide repeat is replaced by an expanded allele (Table2). For most models instability is observed in both somatic tissues and the germ line. These mouse models showed a number of variables

influencing the behaviour of the repeat. Most models showed differences between maternal and paternal transmission as observed in the human situation. In mice it was described that also the gender of the progeny could influence instability (Kovtun *et al.* 2000). Overall, the intergenerational changes observed in mice are to a lesser extent, compared to the human situation. In mice expansion biased instability was observed, but the "big jumps" or "dynamic mutations" were not observed for these mouse models. To study the mechanisms involved in trinucleotide repeat instability, and especially the mechanisms involved in these "big jumps", it might be worthwhile to cross trinucleotide repeat mice with mice deficient for proteins involved in DNA repair or replication since these proteins proved to influence the instability in yeast and bacteria (Jaworski *et al.* 1995; Schumacher *et al.* 1998; Jakupciak *et al.* 1999; Jakupciak *et al.* 2000; Parniewski *et al.* 2000).

In a transgenic Scal (CAG) mouse model, increased (CAG) trinucleotide repeat instability through the germ line was found with advanced maternal age. This increased instability with advanced ageing is also found in somatic tissue for many trinucleotide repeat diseases. For a number of the described mouse models increased instability was found in certain somatic tissues (Fortune *et al.* 2000; Kennedy *et al.* 2000). It appeared that these somatic instabilities showed larger expansions than observed for germ line transmission. The reason for this is not known. Increased somatic instability was observed in among others striatum and kidney. This shows that these "big jumps" were not linked to cell turnover, and indicates that trinucleotide repeat instability can occur by mechanisms that are not replication biased. These mouse models might give more insight into the mechanisms involved in repeat expansion and the timing of repeat expansion.

Chapter 3

FMRP

3.1 FMRP expression

The cloning of *FMR1* revealed that the fragile X phenotype is caused by the absence of FMRP, the fragile X mental retardation protein. This raised new questions about the normal expression pattern of FMRP and the function of FMRP. Although the precise function of FMRP has not been elucidated yet, many characteristics of FMRP have been identified. FMRP is a protein of 70-80 kDa. The different bands observed by Western blotting represent different splice variants of the protein. Alternative splicing is described, allowing theoretically 24 transcripts, but it is not known whether all these transcripts are present and translated into different isoforms (Ashley *et al.* 1993b; Verkerk *et al.* 1993). FMRP has predominantly a cytoplasmatic localisation, but *in vitro* experiments showed that certain isoforms are localised in the nucleus (Eberhart *et al.* 1996; Fridell *et al.* 1996; Sittler *et al.* 1996; Willemsen *et al.* 1996).

FMRP is expressed in almost all tissues tested; exceptions are heart, aorta and skeletal muscle. The level of FMRP is relatively high in the organs affected in the fragile X phenotype, brain and testes. In brain, all neurons show FMRP expression. FMRP is concentrated in the cell body and the proximal dendrites, but no FMRP could be detected in the axons (Feng *et al.* 1997b; Tamanini *et al.* 1997). FMRP is also found in synapses. It has been suggested that FMRP is important in the normal maturation of synaptic connections (Weiler *et al.* 1997; Weiler *et al.* 1999). Non-neuronal cells show less or no FMRP expression (Devys *et al.* 1993).

In testes, primordial germ cells in a human fetus and spermatogonia in adults show high FMRP expression (Tamanini *et al.* 1997). This relatively high expression of FMRP in spermatogonia could suggest that FMRP is necessary for spermatogenesis. This could explain why full mutations are not transmitted through the male germ line. However, the knockout mice completely lacking FMRP are fertile, indicating that FMRP is not necessary for spermatogenesis (Bakker *et al.* 1994). Also, a male with a deletion of *FMR1*, and thus lacking FMRP in his spermatogonia, transmitted this deletion to his daughters (Meijer *et al.* 1994).

3.2 RNA-binding capacities and association to ribosomes

FMRP contains two KH domains and multiple RGG boxes. These domains were previously found to be present in a class of proteins with RNA binding capacity (Mattaj 1993; Burd *et al.* 1994). KH domains are found either as single or multiple copies in RNA binding proteins. The only common property of these proteins is that they are found in close association with RNA. KH domains are capable of binding ssRNA, mRNA, and rRNA *in vivo*. RGG boxes are

found in nuclear and nucleolar RNA-binding proteins (Kiledjian *et al.* 1992). The RGG box found in FMRP shows striking similarity to the RGG box found in the nucleolar protein fibrillin (Siomi *et al.* 1993).

These domains appeared to be functional, since *in vitro* studies showed that FMRP can bind to homopolymeric RNA (Siomi *et al.* 1993). Furthermore, FMRP showed a selectivity for a fraction of mRNAs expressed in brain (Ashley *et al.* 1993a; Brown *et al.* 1998). Another indication that the KH domain was functional was given by a severely affected fragile X patient which showed a point mutation in the second KH domain (De Boulle *et al.* 1993). This mutation changes an isoleucine into an asparagine at position 304. This was the only mutation that is found in FMRP. This mutated protein showed impaired mRNA binding capacity (Siomi *et al.* 1994; Verheij *et al.* 1995). Structural analysis showed a different folding of this KH domain (Musco *et al.* 1996; Musco *et al.* 1997). Together this shows that RNA binding plays a role in the function of FMRP, although the exact RNA targets are not known.

Cellular fractionation studies and biochemical analysis showed that FMRP is associated with ribosomes. Cellular FMRP cosediments with polyribosomes on sucrose gradients (Khandjian *et al.* 1996). In the cytoplasm FMRP is found in close proximity of free ribosomes, membrane-bound ribosomes and polyribosomes (Willemsen *et al.* 1996; Feng *et al.* 1997b). These associations of FMRP to ribosomes are thought to occur via RNA, since RNase treatment removes FMRP from the polyribosomes. Dissociation of the polyribosomes by EDTA treatment followed by further characterisation of the EDTA-dissociated polyribosomes led to the conclusion that FMRP is a component of the messenger ribonucleoprotein particle (mRNP) in actively translating ribosomes (Corbin *et al.* 1997; Feng *et al.* 1997a).

3.3 Shuttling of FMRP between nucleus and cytoplasm

FMRP is predominantly a cytoplasmatic protein, but some studies also report an additional nuclear staining. Expression studies showed that FMRP lacking the domain encoded by exon 14, was localised in the nucleus with exception of the nucleolus (Willemsen *et al.* 1996). In exon 14 a nuclear export signal (NES) was identified (Fridell *et al.* 1996). This NES is also described in Rev and PKI, and it exists of ten amino acid residues of which four are hydrophobic (Fischer *et al.* 1995; Wen *et al.* 1995). A peptide of this NES can direct the nuclear export of a microinjected protein in a temperature-dependent way. Mutant peptides do not have this capacity (Eberhart *et al.* 1996). When the NES in FMRP is mutated, FMRP is directed to the nucleus, not to the nucleolus (Fridell *et al.* 1996). This NES is identified in

an increasing number of apparently functionally unrelated proteins. Some of these proteins are capable of RNA binding like FMRP, other proteins are not.

Rev was one of the first proteins in which this NES was identified. Certain RNAs are recognised by Rev, which functions as an adaptor molecule. Rev binds these RNAs and interacts with a nuclear export factor called exportin 1/CRM1. Once this complex is formed in the nucleus it can penetrate the nuclear envelope as a nuclear pore complex. In the cytoplasm this complex dissociates. Leptomycin B, an antibiotic agent, acts on exportin1. Leptomycin B blocks this nuclear export process in the nucleus. It was found that also the shuttling of FMRP can be blocked by leptomycin (Tamanini *et al.* 1999).

In addition to the NES also a nuclear localisation signal (NLS) was identified in FMRP. Through expression of truncation and fusion constructs the localisation of the NLS in FMRP could be studied. The NLS was identified in the amino terminus of the protein. This NLS did not show similarities to other described NLS. Different studies located the NLS between amino acid 115 and 154 (Eberhart *et al.* 1996; Bardoni *et al.* 1997). The identification of the NES and the NLS suggests that FMRP can shuttle between the nucleus and the cytoplasm. This, together with the RNA-binding and association with ribosomes led to the hypothesis that FMRP might be involved in mRNA transport.

FMRP, as part of the messenger ribonucleoprotein particle, might bind a certain subset of mRNAs to direct them through the nuclear envelope. In the cytoplasm most of the nuclear proteins are exchanged for cytoplasmatic proteins in the RNP particle. However, some RNA-binding proteins in the nucleus and the cytoplasm are the same. Because FMRP is found in both the cytoplasm and the nucleus, it is not known where FMRP becomes associated with this complex.

The FMRP-containing mRNP complex contains several additional proteins and mRNA. Some of these proteins have been identified, such as FXR1P, FXR2P and nucleolin (Ceman *et al.* 1999). At least three other unrecognised proteins are present in this mRNP complex. Using the yeast two-hybrid system, another protein interacting with FMRP was identified. This nuclear FMRP interacting protein (NUFIP) was localised in neurons, like FMRP. An interaction of NUFIP with FXR1P and/ or FXR2P could not be demonstrated (Bardoni *et al.* 1999). Among the mRNAs associated with this complex *FMR1* mRNA was identified. This was the first *in vivo* evidence that FMRP associates with its own mRNA.

How FMRP enters the nucleus after translation in the cytoplasm is not known. The import of the protein is an active type of transport, but an import receptor has not been identified. It might be that already in the cytoplasm complexes are formed and that these preassembled complexes are transported to the nucleus.

3.4 FXR family

Two homologues of *FMR1* are described: *FXR1* and *FXR2*. These three genes together form a small family of fragile X related proteins. First *FXR1* was identified by cross hybridisation experiments (Siomi *et al.* 1995). A yeast two-hybrid screening using FMRP as a bait, identified FXR2P as well as FXR1P (Zhang *et al.* 1995). *FXR1* and *FXR2* are autosomal genes and are localised at chromosome 3q28 and 17p13.1, respectively. The three proteins show a high structure similarity, especially for the domains identified in FMRP. Since the yeast two-hybrid system is used to identify proteins that can interact with each other, an interaction between the three FXR proteins was suggested.

FXR1P shows two major isoforms on Western blot of 70 and 78 kDa; FXR2P has a mass of 95 kDa. In skeletal muscle and heart, two additional isoforms of FXR1P (81 and 84 kDa) were identified (Khandjian *et al.* 1998; Kirkpatrick *et al.* 1999). FMRP is highly homologous to FXR1P and FXR2P. Especially, the amino-terminal and central regions are very homologous. In these regions the identified functional domains are located. The carboxy-terminus is very divergent, and shows only 6% similarity. Recently, in the carboxy terminal region of some isoforms of FXR1P and in FXR2P a nucleolar-targeting signal was identified (Tamanini *et al.* 2000). This domain is not present in FMRP. The expression of FXR1P and FXR2P shows similarities, as well as differences with the tissue distribution of FMRP. Like FMRP, FXR1P and FXR2P are expressed in neurons and testis, the organs affected in the fragile X syndrome. FXR1P is also highly expressed in muscle.

In mice the expression of the different proteins is extensively studied by immunohistochemistry and biochemistry (Bakker *et al.* 2000). In neurons FMRP, FXR1P and FXR2P could be identified in the cytoplasm. FXR1P was also found in the nucleolus in some neurons. Immunoelectronmicroscopy showed that the majority of the three proteins was found in the cytoplasm in association with ribosomes. A minority was found in the nucleus. This tissue distribution of the FXR proteins suggests a nucleocytoplasmatic shuttling for these proteins. This colocalisation is consistent with similar cellular functions of these proteins in those neurons.

In testis a differential expression is observed. In human fetal testes FMRP and FXR1P expression is high in all primordial germ cells. To a lesser extent, FXR1P is localised in the non-spermatogenic cells (Khandjian *et al.* 1998) (Tamanini *et al.* 1997). FXR2P is highly expressed in the interstitial cells. In human adult testes FMRP is expressed in the spermatogonia close to the basal membrane of the tubuli seminiferi, together with FXR1P. However, FXR1P is predominantly expressed in the more advanced spermatogenic cells

more towards the lumen of the tubuli seminiferi. FXR2P gave a less intense labelling in all cells of the tubuli seminiferi. In testes of a fragile X fetus the labelling of FXR1P and FXR2P was as in testes of a normal fetus. FMRP was present only in a few spermatogonia. These spermatogonia most likely have a premutation allele due to contraction of the full mutation allele. These results indicate that, due to the different expression, the FXR proteins may have independent functions. However, in the spermatogonia coexpression of the three proteins was detected.

3.5 The *Fmr1* and *Fxr2* knockout mouse models

The generation of a knockout mouse model for the fragile X syndrome facilitated the study of the function of FMRP, as well as the pathogenesis of the fragile X syndrome. The *Fmr1* knockout mouse has characteristics in common with fragile X patients (Bakker *et al.* 1994). The mental retardation as seen in the fragile X patients, was tested in knockout mice by the Morris water maze test. The results of these tests showed abnormalities in cognitive function. The learning in these knockout mice is delayed. Furthermore, it was found that knockout mice are susceptible to epileptic seizures after audiogenic stimulation (Musumeci *et al.* 2000). Previously, it was described that in fragile X patients an increased incidence of epilepsy was found (Kluger *et al.* 1996; Musumeci *et al.* 1999). This might indicate that absence of FMRP increases the cortical excitability.

In fragile X patients some non-specific abnormalities, like brain-atrophy, were described. Furthermore, the dendritic spines of pyrimidal neurons are abnormal (Rudelli *et al.* 1985; Hinton *et al.* 1991; Wisniewski *et al.* 1991). In knockout mice the dendritic spines are longer and thinner, and they have a higher density along apical dendrites compared to normal mice. These characteristics are also seen after sensory deprivation. Therefore, these observations suggest a deficit in the normal selection or "pruning" of synaptic contacts during early synaptogenesis
(Comery *et al.* 1997).

The macroorchidism as observed in male fragile X patients was also found in male *Fmr1* knockout mice, albeit to a lesser extent. The macroorchidism in these mice develops through time. This corresponds with the observation in fragile X patients where the macroorchidism occurs after puberty. The macroorchidism observed in the knockout mice is caused by increased Sertoli cell proliferation during testis development (Slegtenhorst-Eegdeman *et al.* 1998).

The function of FXR1P and FXR2P might resemble the function of FMRP, because of the homology between the three proteins and their partly overlapping expression. It was also

suggested that FXR1P and/ or FXR2P might complement the function of FMRP to a certain level. The levels of FXR1P and FXR2P appear to be normal in fragile X patients. The generation of knockout mice for Fxr1P and Fxr2P would facilitate the study of the functions of FXR1P and FXR2P. Furthermore, these knockout models might reveal a phenotype due to the absence of FXR1P or FXR2P, since in humans no disease is known to involve one of these two proteins. To study the interactions between these proteins, and to study in more detail whether there is some kind of redundancy, the generation of double-knockout mice might be informative. In Chapter 5.4 the generation of a knockout model for Fxr2p is described. The results about the observed phenotype are only preliminary. A knockout mouse is also generated for Fxr1p. These knockout mice die a few hours after birth (Siomi, personal communication). Normally, Fxr1p is highly expressed in muscle. It was suggested that *Fxr1* knockout mice die because of problems with the respiratory tract.

Chapter 4

Discussion

Since the cloning of *FMR1* the research concerning the fragile X syndrome has focused on two fields of research. For studies concerning the pathogenesis of the fragile X syndrome, understanding the function of FMRP is very important. Many characteristics of FMRP have been revealed. Although the precise function is not yet known, the available data led to speculation about the function of FMRP, as well as of the fragile X related proteins, FXR1P and FXR2P.

With the cloning of *FMR1* a new mutational mechanism involved in human diseases was discovered: the expansion of trinucleotide repeats. A new field of research developed trying to answer questions about the CGG repeat instability observed in fragile X syndrome and other trinucleotide repeat expansions involved in other human diseases. Despite all these studies, the exact timing of and the mechanisms involved in repeat expansions are still not known. Furthermore, the role of methylation in CGG repeat instability and in the regulation of FMRP expression is studied.

4.1 CGG repeat instability

The highly polymorphic CGG repeat present in the 5' UTR of *FMR1* can expand to more than 200 CGG repeats. These full mutations can result in methylation of the promoter region thereby blocking the transcription of *FMR1*, giving rise to the fragile X syndrome. The causative mechanism(s) and the timing of repeat instability are not known. A lot of studies have been performed to determine the behaviour of different trinucleotide repeats in bacteria, yeast and cell culture systems. The results of these studies are influenced by the length of the trinucleotide repeat, the purity of the repeat (the number of interruptions) and the sequences flanking the trinucleotide repeat. All studies described that instability increases with the length of the trinucleotide repeat. The opposite is true for the number of interruptions; more interruptions in a trinucleotide repeat decreases the instabilities observed. The influence of flanking sequences is difficult to determine. However, in transgenic mouse models the influence of flanking sequences seems to be limited. The observed instability in mice does not dramatically increase when additional human sequences are added. Furthermore, transgenic models using different founders give in most cases comparable instability rates.

The direction of replication through the CGG repeat, studied in bacteria and yeast, is important in determining the rate of instability. When the CGG triplet constitutes the template of the lagging strand, the CGG repeat behaves less stable (Shimizu *et al.* 1996; Hirst *et al.* 1998). In this orientation relatively more contractions are observed (Balakumaran *et al.* 2000), whereas in the other orientation the CGG repeats have a higher chance to

expand. Still, the majority of instabilities observed in yeast and bacteria are contractions, in contrast to the human situation where the premutations in fragile X families predominantly expand upon transmission. The exact distribution and transmission of premutations in the general population is not known. Therefore a certain bias for known fragile X families cannot be excluded. Theoretically, it is possible that the likelihood of premutation expansion is higher in individuals from known fragile X families (Geva *et al.* 2000).

Other factors influencing the instability of trinucleotide repeats concern the genetic background of the host. These factors include proteins involved in DNA repair, recombination and/ or replication events. The influence of such factors is still not understood and it is not known if and how the different systems interact with each other. Mismatch repair (MMR) was suggested to play a role in trinucleotide repeat instability. More instability was found in hereditary non-polyposis colorectal cancer patients than in controls, especially when one of the mismatch repair proteins, hMLH1, was mutated. However, the normal length of the FMR1 CGG repeat in human cell lines in which hMLH1 or hMSH2 were mutated did not show instability. By introduction of normal and premutation size uninterrupted CGG repeats in yeast, a low frequency of orientation-dependent large expansions was observed. The frequency of these expansions was not affected by deletion of several DNA mismatch repair genes (White *et al.* 1999). Also other studies about the involvement of mismatch-repair in CGG trinucleotide repeat expansion led to different conclusions.

Conflicting results were also detected for CTG repeat instability. Mutations in mismatch repair proteins stabilised the behaviour of an $(CTG)_{175}$ (containing two interruptions) (Jaworski *et al.* 1995), but for a much shorter and uninterrupted $(CTG)_{64}$ tract MMR mutations led to deletions (Schumacher *et al.* 1998). The influence of MMR may depend on the length and the purity of the repeat. Repeats with interruptions will lead to mismatches when slippage occurs. These mismatches trigger MMR, resulting in deletions. When, due to mutations in MMR or the absence of interruptions the MMR system is not activated, deletions will not occur. A difference is made between small deletions, which are thought to arise for all sizes of CTG repeats, and larger deletions, which are more frequently observed in longer CGG tracts. These larger deletions more readily occur in *E. coli* when an active MMR system is present. Interruptions were found to increase the number of large deletions (Parniewski *et al.* 2000).

For expanded CGG repeats replication was found to be delayed. This might be the result of stalling of DNA polymerase at trinucleotide repeats (Iyer *et al.* 2000) or of pausing of DNA synthesis as was found in trinucleotide repeats *in vitro* (Kang *et al.* 1995). Stalling of DNA

polymerase can result in slippage, which allows the formation of alternative stable DNA structures. The formation of hairpins, tetraplex structures and slipped-strand structures has been described. These alternative stable DNA structures may mediate errors during DNA replication, repair and/or recombination of the trinucleotide repeat, leading to expansion. AGG interruptions were found to decrease the ability to form alternative DNA structures, which can explain the stabilising effect of AGG interruptions.

Flap endonuclease 1 (FEN1) is involved in long-patch base excision repair and is suggested to play a role in trinucleotide repeat expansion. In yeast, loss of flap endonuclease activity increases instability of trinucleotide repeats (Freudenreich *et al.* 1998; Schweitzer *et al.* 1998). However, this instability is not restricted to trinucleotide repeats, but is observed throughout the whole genome. This in contrast to the site-restricted trinucleotide repeat expansion observed in human diseases. Secondary structures inhibit FEN1 activity at trinucleotide repeats (Spiro *et al.* 1999). For normal and premutation size CGG repeats, a 10-fold elevated frequency of expansion in FEN1 mutant yeast strain was found (White *et al.* 1999). These results are similar to those described for CTG repeats. This suggests that FEN1 plays a role in the processing of branched DNA structures in trinucleotide repeat expansion.

Most likely, the different factors, that were found to influence the behaviour of trinucleotide repeats, form a complex network of interactions *in vivo*. A precise balance between replication, repair and recombination is necessary. This balance might in some critical situations cause the dynamic mutations observed in trinucleotide repeat expansions. It is not known whether the same mechanisms are involved in CTG and CGG repeat expansions. Presumably, the mechanisms involved are the same, but the combination of the involved factors together, like length and interruptions of the repeat, repair, formed DNA structures and cell cycle conditions, determine the consequences for repeat instability. The complexity of these mechanisms and the involvement of many factors makes it difficult to study the influence of each factor separately. Furthermore, it is important to distinguish instability from expansions. Instability includes both contractions and expansions. The influence of some factors might be different for expansions and/or contractions, which makes it difficult to compare studies. The role of expansions is important in causing trinucleotide repeat diseases in humans. In models like *E. coli* and yeast, contractions are much more frequent. For this reason results obtained from yeast and *E. coli* models are difficult to extrapolate to the human situation.

To study trinucleotide repeat expansions in mice, mainly two types of mouse models are used: transgenic mouse models and so-called knock-in mouse models. These mouse

models would enable researchers to study the mechanisms involved in trinucleotide repeat expansions in more detail, especially the timing of the repeat expansions, as well as the differences due to the parental origin of the repeat. Most mouse models show instabilities, both in somatic tissues and the germ line. A difference in instability due to the parental origin of the expanded repeat is also observed in mice. In most cases the observed percentage instability as well as the size changes in mice are less compared to the human situation. The strong preference for expansions as observed in humans seems to be absent in mice. It was hypothesized that the small changes in repeat length occur in mice, but no "big jumps" or "dynamic mutations" were observed upon transmission. For some somatic tissues (striatum, kidney) large expansions occurred. The tissues where these dynamic mutations occurred do not show a high cell turnover. This shows that dynamic mutations can occur without replication. The mechanisms mentioned to be involved in repeat instability are present in both mouse and human. However, the complex interactions between trinucleotide repeat, DNA secondary structure, replication, repair and recombination might be different in mice. It is also possible that in mice the threshold for dynamic mutations is higher. Crossing the trinucleotide repeat mice with mice deficient for certain aspects of repair, recombination and/or replication might give more insight in the balance determining repeat instability.

Different fragile X CGG repeat mouse models have been described (Bontekoe *et al.* 1997; Lavedan *et al.* 1997; Lavedan *et al.* 1998). Only two models, including the $(CGG)_{98}$ knock-in mouse model described in this thesis, show instability upon transmission (D. Nelson, personal communication). The observed instability is less than in humans both in frequency as in size changes. For the $(CGG)_{98}$ knock-in model 16 instabilities were found; 14 expansions and 2 contractions. The contractions and 11 expansions were found upon paternal transmission. Only one expansion was observed upon maternal transmission; for 2 other expansions the parental origin of the repeat could not be determined. The timing of these small changes has not been determined yet. In humans expansions in the premutation range are observed upon both paternal and maternal transmission. For unknown reasons, expansions to full mutations only arise upon maternal transmission. Based on the results obtained for other trinucleotide repeat mouse models, it may be informative to search for instability in somatic tissues of both males and females.

Besides CGG repeat instability, our knock-in $(CGG)_{98}$ mouse model can be used to study the inactivation of FMRP expression upon expansion. FMRP expression is present in these mice suggesting that the promoter region is not methylated. To get more insight into the methylation and inactivation of FMRP expression, both the expansion of the repeat and the

FMRP expression in these mice should be determined. Recently, it was described that premutation carriers have an elevated *FMR1* mRNA level, whereas the expression of FMRP was normal in lymphocytes; for premutation > 100 CGG the number of FMRP-expressing lymphocytes might even be reduced (Tassone *et al.* 2000a; Tassone *et al.* 2000b; Tassone *et al.* 2000c). Further experiments for these $(CGG)_{98}$ knock-in mice can be performed to study the mRNA levels in premutations, the diminished transcription of FMRP mRNA containing expanded CGG repeats and the mechanistic switch when *FMR1* is silenced.

4.2 FMRP and its function

To study the fragile X phenotype, knowledge about the expression and function of FMRP is essential. It is known that FMRP is highly expressed in brain and testis, the organs known to be involved in fragile X syndrome. Furthermore, FMRP is expressed in fast-dividing cells. The expression in brain is mainly in differentiated neurons. These findings in first instance seem to contradict, because differentiated neurons are non-dividing cells. However, a similarity between fast-dividing cells and differentiated neurons is the high turnover in translation that is required. Dividing cells need a high transcription/ translation because of their proliferation. Neurons need the ability to respond very fast to stimuli immediately resulting in protein synthesis depending on the required response. This resemblance might shed light on the function of FMRP, based on the tissue distribution.

The characteristics found in FMRP, namely RNA-binding, polyribosome association and shuttling between the nucleus and the cytoplasm suggested that FMRP might have a function in the transcription/ translational machinery. FMRP might bind specific mRNAs to form functional mRNP complexes. The mRNP complexes will be presented to the translation machinery to allow translation of these mRNAs. Instead of promoting translation of these specific mRNAs it is also possible that binding of these mRNAs to FMRP masks these mRNAs, thereby repressing the translation. When the protein is needed a certain stimulus can direct the mRNA to be presented to the translation machinery.

The characteristics described of FMRP suggest a very general function of FMRP. If FMRP would play a general role in the transcription/translation machinery this is difficult to combine with the relatively mild phenotype seen in the fragile X patients. Studies are performed to describe the exact intracellular localisation of the FXR proteins and to unravel the transport routes they use or facilitate. Putting together the characteristics of FMRP together with those of FXR1P and FXR2P, it seems likely that the FXR proteins have a function in the transport of certain mRNAs. It is thought that FMRP has a function in the fine regulation of the transport of a number of mRNAs. FMRP might facilitate the transport of

certain mRNAs. In that way FMRP enables the cell to respond very fast and precise to certain stimuli. In the absence of FMRP, responses to certain stimuli will be slower or less efficient. This fits with the observed phenotype in fragile X patients. The absence of FMRP does not cause major physiological difficulties but the precise regulation of more downstream targets in this transport pathway will be disturbed.

Since the phenotype of *Fmr1* knockout mice resembles the phenotype observed in fragile X patients, it is possible to study the pathology of the fragile X phenotype. By unravelling the function of FMRP it might become clear how the absence of FMRP can cause the characteristics of the fragile X phenotype. For instance, it might be possible to study whether the observed macroorchidism is a direct effect of the absence of FMRP in the testis. Another explanation might be that the absence of FMRP in the brain might indirectly cause the macroorchidism. To address this question, transgenic mice expressing FMRP solely in the brain or in the testes are generated. Furthermore, it is not known whether the absence of FMRP causes developmental abnormalities. FMRP expression is observed during development, but it is not known whether the absence of FMRP has an irreversible effect during development. In other words, it is not known whether there is a possibility that addition of FMRP might restore mental capacities. By using conditional knockouts or transgenic mice expressing FMRP postnatally it will be possible to address these questions. By creating transgenic mice expressing certain isoforms of FMRP or expressing FMRP in certain tissues a number of questions can be addressed. Recently, different reports describe the attempt to create a total "rescue" mouse. Therefore transgenic mice expressing either *FMR1* cDNA or the entire *FMR1* locus from a human cosmid or YAC were used. None of these rescue lines showed a complete rescue of the *Fmr1* knockout mice phenotype. Only for the rescue mice containing the *FMR1*-YAC, rescue of the macroorchidism was observed.

The behavioural characterisation of these rescue mice did not show a rescue of the fragile X like phenotype, but they did show differences between the knockout, rescue, control and overexpression mice. It seemed that the amount of FMRP is closely correlated to the observed behaviour. Not only the absence of FMRP influences the behaviour of mice. Changes in behaviour were also observed in the overexpression mice. The rescue mice showed behaviour comparable to the behaviour of the overexpression mice.

These results have significant implications for substitution or gene therapy for the fragile X syndrome. These results suggest that the amount of FMRP required in the cell is strictly regulated. Decreased as well as increased levels of FMRP have implications for behaviour. Previous observations support the idea that overexpression of FMRP has consequences for

cells. Difficulties were encountered producing stable cell lines overexpressing FMRP (Peier, personal communication) and also by the generation of transgenics overexpressing FMRP. Only by decreasing the amount of DNA used for microinjection transgenic progeny was obtained. This most likely is the result of a lower copynumber, and thus less overexpression, in these cells.

Biochemical studies try to unravel the localisation of the FXR proteins and the transport routes they use or facilitate. Putting together the characteristics of FMRP known sofar it seems likely that FMRP has a function in the transport of certain mRNAs. For FXR1P and FXR2P analogous functions are suggested. The *Fxr2* knockout mice phenotype resembles the phenotype observed in the *Fmr1* knockout mouse model. *Fxr2* knockout mice show learning deficits, but sofar there is no evidence of macroorchidism. The identification of mRNA targets for the FXR proteins will enable us to study the exact consequences of decreased and/ or increased levels of the FXR proteins on the expression of these more downstream in the pathway mRNA targets. Furthermore, the generation of double knockouts for *Fxr2/ Fmr1* might shed more light on the *in vivo* interaction of the FXR proteins. Conditional knockouts as well as transgenics with a regulated expression of the FXR proteins may be useful towards the understanding of the biological function of the FXR proteins.

References

Ashley, C., Jr., K. D. Wilkinson, et al. (1993a). "FMR1 protein: conserved RNP family domains and selective RNA binding." Science 262(5133): 563-568.

Ashley, C. T., J. S. Sutcliffe, et al. (1993b). "Human and murine FMR-1: alternative splicing and translational initiation downstream of the CGG-repeat." Nature Genet 4(3): 244-251.

Bakker, C. E., Y. de Diego Otero, et al. (2000). "Immunocytochemical and biochemical characterization of FMRP, FXR1P, and FXR2P in the mouse." Exp Cell Res 258(1): 162-70.

Bakker, C. E., C. Verheij, et al. (1994). "Fmr1 knockout mice: A model to study fragile X mental retardation." Cell 78: 23-33.

Balakumaran, B. S., C. H. Freudenreich, et al. (2000). "CGG/CCG repeats exhibit orientation-dependent instability and orientation-independent fragility in Saccharomyces cerevisiae." Hum Mol Genet 9(1): 93-100.

Baldi, P., S. Brunak, et al. (1999). "Structural basis for triplet repeat disorders: a computational analysis [In Process Citation]." Bioinformatics 15(11): 918-29.

Bardoni, B., A. Schenck, et al. (1999). "A novel RNA-binding nuclear protein that interacts with the fragile X mental retardation (FMR1) protein." Hum Mol Genet 8(13): 2557-2566.

Bardoni, B., A. Sittler, et al. (1997). "Analysis of domains affecting intracellular localization of the FMRP protein." Neurobiol Dis 4(5): 329-36.

Bingham, P. M., M. O. Scott, et al. (1995). "Stability of an expanded trinucleotide repeat in the androgen receptor gene in transgenic mice." Nature Genet 9: 191-196.

Bontekoe, C. J. M., E. de Graaff, et al. (1997). "FMR1 premutation allele is stable in mice." Eur J Hum Genet 5: 293-298.

Brook, J. D., M. E. McCurrach, et al. (1992). "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.

Brown, V., K. Small, et al. (1998). "Purified recombinant Fmrp exhibits selective RNA binding as an intrinsic property of the fragile X mental retardation protein." J Biol Chem 273(25): 15521-15527.

Burd, C. G. and G. Dreyfuss (1994). "Conserved structures and diversity of functions of RNA-binding proteins." Science 265(5172): 615-21.

Burright, E. N., H. B. Clark, et al. (1995). "SCA1 transgenic mice: A model for neurodegeneration caused by an expanded CAG trinucleotide repeat." Cell 82: 937-948.

Campuzano, V., L. Montermini, et al. (1996). "Friedreich's ataxia: Autosomal recessive disease caused by an intronic GAA triplet repeat expansion." Science 271(5254): 1423-1427.

Ceman, S., V. Brown, et al. (1999). "Isolation of an FMRP-Associated Messenger Ribonucleoprotein Particle and Identification of Nucleolin and the Fragile X-Related Proteins as Components of the Complex." Mol Cell Biol 19(12): 7925-7932.

Chakrabarti, L., S. J. L. Knight, et al. (1996). "A candidate gene for mild mental handicap at the FRAXE fragile site." Hum Mol Genet 5(2): 275-282.

Chiurazzi, P., M. G. Pomponi, et al. (1999). "Synergistic effect of histone hyperacetylation and DNA demethylation in the reactivation of the FMR1 gene." Hum Mol Genet 8(12): 2317-2323.

Chiurazzi, P., M. G. Pomponi, et al. (1998). "In vitro reactivation of the FMR1 gene involved in fragile X syndrome." Hum Mol Genet 7(1): 109-113.

Coffee, B., F. Zhang, et al. (1999). "Acetylated histones are associated with FMR1 in normal but not fragile X-syndrome cells." Nat Genet 22(1): 98-101.

Comery, T. A., J. B. Harris, et al. (1997). "Abnormal dendritic spines in fragile X knockout mice: Maturation and pruning deficits." Proc Natl Acad Sci USA 94(10): 5401-5404.

Corbin, F., M. Bouillon, et al. (1997). "The fragile X mental retardation protein is associated with poly(A)(+) mRNA in actively translating polyribosomes." Hum Mol Genet 6(9): 1465-1472.

Crawford, D. C., C. E. Schwartz, et al. (2000a). "Survey of the fragile X syndrome CGG repeat and the short-tandem-repeat and single-nucleotide-polymorphism haplotypes in an African American population." Am J Hum Genet 66(2): 480-93.

Crawford, D. C., B. Wilson, et al. (2000b). "Factors involved in the initial mutation of the fragile X CGG repeat as determined by sperm small pool PCR." Hum Mol Genet 9(19): 2909-2918.

Crawford, D. C., F. Zhang, et al. (2000c). "Fragile X CGG repeat structures among African-Americans: identification of a novel factor responsible for repeat instability." Hum Mol Genet 9(12): 1759-1769.

Cummings, C. J. and H. Y. Zoghbi (2000). "Fourteen and still counting: unraveling trinucleotide repeat diseases." Human Molecular Genetics 9(6): 909-916[Baldi, 1999 #1523].

De Boulle, K., A. J. Verkerk, et al. (1993). "A point mutation in the FMR-1 gene associated with fragile X mental retardation." Nature Genet 3(1): 31-35.

De Graaff, E., P. Rouillard, et al. (1995). "Hotspot for deletions in the CGG repeat region of FMR1 in fragile X patients." Hum Mol Genet 4: 45-49.

de Vries, B. B., S. Mohkamsing, et al. (1999). "Screening for the fragile X syndrome among the mentally retarded: a clinical study. The Collaborative Fragile X Study Group." J Med Genet 36(6): 467-70.

De Vries, B. B., A. M. van den Ouweland, et al. (1997). "Screening and diagnosis for the fragile X syndrome among the mentally retarded: an epidemiological and psychological survey. Collaborative Fragile X Study Group." Am J Hum Genet 61(3): 660-667.

De Vries, B. B. A. (1997). "The fragile X syndrome. Clinical, genetic and large scale diagnostic studies among mentally retarded individuals." Thesis.

De Vries, B. B. A., C. A. M. Jansen, et al. (1996). "Variable FMR1 gene methylation leads to variable phenotype in 3 males from one fragile X family." J Med Genet 33: 1007-1010.

Deelen, W., C. Bakker, et al. (1994). "Conservation of CGG region in FMR1 gene in mammals." Am J Med Genet 51(4): 513-516.

Devys, D., V. Biancalana, et al. (1992). "Analysis of full fragile X mutations in fetal tissues and monozygotic twins indicate that abnormal methylation and somatic heterogeneity are established early in development." Am J Med Genet 43(1-2): 208-216.

Devys, D., Y. Lutz, et al. (1993). "The FMR-1 protein is cytoplasmic, most abundant in neurons and appears normal in carriers of a fragile X premutation." Nature Genet 4(4): 335-340.

Drouin, R., M. Angers, et al. (1997). "Structural and functional characterization of the human FMR1 promoter reveals similarities with the hnRNP-A2 promoter region." Hum Mol Genet 6(12): 2051-2060.

Eberhart, D. E., H. E. Malter, et al. (1996). "The fragile X mental retardation protein is a ribosonucleoprotein containing both nuclear localization and nuclear export signals." Hum Mol Genet 5: 1083-1091.

Eichler, E. E., J. Holden, et al. (1994). "Length of uninterrupted CGG repeats determines instability in the FMR1 gene." Nature Genet 8(1): 88-94.

Eichler, E. E., C. B. Kunst, et al. (1995). "Evolution of the cryptic FMR1 CGG repeat." Nature Genet 11(3): 301-308.

Eichler, E. E., J. N. Macpherson, et al. (1996). "Haplotype and interspersion analysis of the FMR1 CGG repeat identifies two different mutational pathways for the origin of the fragile X syndrome." Hum Mol Genet 5(3): 319-330.

Feng, Y., D. Absher, et al. (1997a). "FMRP associates with polyribosomes as an mRNP, and the I304N mutation of severe fragile X syndrome abolishes this association." Mol Cell 1(1): 109-118.

Feng, Y., C. A. Gutekunst, et al. (1997b). "Fragile X mental retardation protein: Nucleocytoplasmic shuttling and association with somatodendritic ribosomes." J Neurosci 17(5): 1539-1547.

Feng, Y., C Lakkis, et al. (1995a). "Quantitative comparison of FMR1 gene expression in normal and premutation alleles." Am J Hum Genet 56: 106-113.

Feng, Y., F P. Zhang, et al. (1995b). "Translational suppression by trinucleotide repeat expansion at FMR1." Science 268(5211): 731-734.

Fisch, G. S. K. Snow, et al. (1995). "The fragile X premutation in carriers and its effect on mutation size in offspring." Am J Hum Genet 56(5): 1147-1155.

Fischer, U., J. Huber, et al. (1995). "The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs." Cell 82: 475-483.

Fortune, M T., C. Vassilopous, et al. (2000). "Dramatic, expansion-biased, age-dependent, tissue-specific somatic mosaicism in a transgenic mouse model of triplet repeat instability." Human Molecular Genetics 9(3): 439-445.

Freudenreich, C. H., S. M. Kantrow, et al. (1998). "Expansion and length-dependent fragility of CTG repeats in yeast." Science 279(5352): 853-6.

Fridell, R. A., R. E. Benson, et al. (1996). "A nuclear role for the fragile X mental retardation protein." EMBO J 15(19): 5408-5414.

Fu, Y. H., D. P. Kuhl, et al. (1991). "Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox." Cell 67(6): 1047-1058.

Fu, Y. H., A. Pizzuti, et al. (1992). "An unstable triplet repeat in a gene related to myotonic muscular dystrophy." Science 255: 1256-1258.

Geva, E., Y Yaron, et al. (2000). "The risk of fragile X premutation expansion is lower in carriers detected by general prenatal screening than in carriers from known fragile X families." Genet Test 4(3): 289-92.

Goldberg, Y P., M. A. Kalchman, et al. (1996). "Absence of disease phenotype and intergenerational stability of the CAG repeat in transgenic mice expressing the human Huntington disease transcript." Hum Mol Genet 5(2): 177-185.

Gourdon, G., F. Radvanyi, et al. (1997). "Moderate intergenerational and somatic instability of a 55-CTG repeat in transgenic mice." Nature Genet 15: 190-192.

Groenen, P. and B. Wieringa (1998). "Expanding complexity in myotonic dystrophy." Bioessays 20(11): 901-12.

Gunter, C., W. Paradee, et al. (1998). "Re-examination of factors associated with expansion of CGG repeats using a single nucleotide polymorphism in FMR1." Hum Mol Genet 7(12): 1935-46.

Hagerman, R. J., C. E. Hull, et al. (1994). "High functioning fragile X males: Demonstration of an unmethylated fully expanded FMR-1 mutation associated with protein expression." Amer J Med Genet 51(4): 298-308.

Hagerman, R. J., C. Jackson, et al. (1992). "Girls with fragile X syndrome: physical and neurocognitive status and outcome." Pediatrics 89(3): 395-400.

Hansen, R. S., T. K. Canfield, et al. (1997). "A variable domain of delayed replication in FRAXA fragile X chromosomes: X inactivation-like spread of late replication." Proc Natl Acad Sci USA 94(9): 4587-4592.

Hansen, R. S., T. K. Canfield, et al. (1993). "Association of fragile X syndrome with delayed replication of the FMR1 gene." Cell 73(7): 1403-1409.

Hansen, R. S., S. M. Gartler, et al. (1992). "Methylation analysis of CGG sites in the CpG island of the human FMR1 gene." Hum Mol Genet 1(8): 571-578.

Hergersberg, M., K. Matsuo, et al. (1995). "Tissue-specific expression of a FMR1/beta-galactosidase fusion gene in transgenic mice." Hum Mol Genet 4(3): 359-366.

Hinton, V. J., W. T. Brown, et al. (1991). "Analysis of neocortex in three males with the fragile X syndrome." Am J Med Genet 41(3): 289-294.

Hirst, M. C., S. J. L. Knight, et al. (1991). "Prenatal diagnosis of fragile X syndrome." Lancet 338(8772): 956-957.

Hirst, M. C. and P. J. White (1998). "Cloned human FMR1 trinucleotide repeats exhibit a length- and orientation-dependent instability suggestive of in vivo lagging strand secondary structure." Nucleic Acids Res 26(10): 2353-2358.

Holmes, S. E., E. E. O'Hearn, et al. (1999). "Expansion of a novel CAG trinucleotide repeat in the 5' region of PPP2R2B is associated with SCA12." Nature Genetics 23: 391-392.

Hornstra, I. K., D. L. Nelson, et al. (1993). "High resolution methylation analysis of the FMR1 gene trinucleotide repeat region in fragile X syndrome." Hum Mol Genet 2(10): 1659-1665.

Hwu, W.-L., Y.-M. Lee, et al. (1993). "In vitro DNA methylation inhibits FMR1 promoter." Biochem Biophys Res Comm 193: 324-329.

Iyer, R. R., A. Pluciennik, et al. (2000). "DNA polymerase III proofreading mutants enhance the expansion and deletion of triplet repeat sequences in Escherichia coli." J. Biol. Chem. 275(3): 2174-84.

Jakupciak, J. P. and R. D. Wells (1999). "Genetic instabilities in (CTG.CAG) repeats occur by recombination." J. Biol. Chem. 274(33): 23468-79.

Jakupciak, J. P. and R. D. Wells (2000). "Gene Conversion (Recombination) Mediates Expansions of CTG{middle dot}CAG Repeats.*." J Biol Chem.

Jaworski, A., W. A. Rosche, et al. (1995). "Mismatch repair in Escherichia coli enhances instability of (CTG)n triplet repeats from human hereditary diseases." Proc Natl Acad Sci USA 92(24): 11019-23.

Kang, S. M., K. Ohshima, et al. (1995). "Pausing of DNA synthesis in vitro at specific loci in CTG and CGG triplet repeats from human hereditary disease genes." J Biol Chem 270(45): 27014-27021.

Kaytor, M. D., E. N. Burright, et al. (1997). "Increased trinucleotide repeat instability with advanced maternal age." Hum Mol Genet 6(12): 2135-2139.

Kennedy, L. and P. F. Shelbourne (2000). "Dramatic mutation instability in HD mouse striatum: does polyglutamine load contribute to cell-specific vulnerability in Huntington's disease?" Human Molecular Genetics 9(17): 2539-44.

Khandjian, E., F. Corbin, et al. (1996). "The fragile X mental retardation protein is associated with ribosomes." Nature Genet 12: 91-93.

Khandjian, E. W., B. Bardoni, et al. (1998). "Novel isoforms of the fragile X related protein FXR1P are expressed during myogenesis." Hum Mol Genet 7(13): 2121-2128.

Kiledjian, M. and G. Dreyfuss (1992). "Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box." Embo J 11(7): 2655-64.

Kirkpatrick, L. L., K. A. McIlwain, et al. (1999). "Alternative splicing in the murine and human FXR1 genes." Genomics 59(2): 193-202.

Kluger, G., I. Bohm, et al. (1996). "Epilepsy and fragile X gene mutations." Pediatr Neurol 15(4): 358-60.

Knight, S. J., A. V. Flannery, et al. (1993). "Trinucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation." Cell 74(1): 127-134.

Kovtun, I. V., T. M. Therneau, et al. (2000). "Gender of the embryo contributes to CAG instability in transgenic mice containing a Huntington's disease gene." Hum Mol Genet 9(18): 2767-75.

Kremer, E. J., M. Pritchard, et al. (1991). "Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n." Science 252(5013): 1711-1714.

Kumari, D. and K. Usdin (2000). "Interactions of the transcription factors USF1, USF2 and alpha-Pal/Nrf-1 at the FMR1 promoter: implications for Fragile X Mental Retardation syndrome." J Biol Chem 31: 31.

Kunst, C. B. and S. T. Warren (1994). "Cryptic and polar variation of the fragile X repeat could result in predisposing normal alleles." Cell 77(6): 853-861.

La Spada, A. R., K. R. Peterson, et al. (1998). "Androgen receptor YAC transgenic mice carrying CAG 45 alleles show trinucleotide repeat instability." Hum Mol Genet 7(6): 959-67.

Lavedan, C., E. Grabczyk, et al. (1998). "Long uninterrupted CGG repeats within the first exon of the human FMR1 gene are not intrinsically unstable in transgenic mice." Genomics 50(2): 229-240.

Lavedan, C. N., L. Garrett, et al. (1997). "Trinucleotide repeats (CGG)22TGG(CGG)43TGG(CGG)21 from the fragile X gene remain stable in transgenic mice." Hum Genet 100: 407-414.

Lia, A. S., H. Seznec, et al. (1998). "Somatic instability of the CTG repeat in mice transgenic for the myotonic dystrophy region is age dependent but not correlated to the relative intertissue transcription levels and proliferative capacities." Human Molecular Genetics 7(8): 1285-91.

Lin, X., C. J. Cummings, et al. (1999). "Expanding our understanding of polyglutamine diseases through mouse models." Neuron 24: 499-502.

Lorenzetti, D., K. Watase, et al. (2000). "Repeat instability and motor incoordination in mice with a targeted expanded CAG repeat in the Sca1 locus." Human Molecular Genetics 9(5): 779-85.

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

Mahadevan, M., C. Tsilfidis, et al. (1992). "Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene." Science 255: 1253-1255.

Malter, H. E., J. C. Iber, et al. (1997). "Characterization of the full fragile X syndrome mutation in fetal gametes." Nature Genet 15: 165-169.

Mangiarini, L., K. Sathasivam, et al. (1997). "Instability of highly expanded CAG repeats in mice transgenic for the Huntington's disease mutation." Nature Genet 15: 197-200.

Manley, K., T. L. Shirley, et al. (1999). "Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice." Nature Genetics 23(4): 471-3.

Martin, J. P. and J. Bell (1943). "A pedigree of mental defect showing sex-linkage." J Neurol Psych 6: 154-157.

Mattaj, I. W. (1993). "RNA recognition: a family matter?." Cell 73(5): 837-40.

McConkie-Rosell, A., A. Lachiewicz, et al. (1993). "Evidence that methylation of the FMR1 locus is responsible for variant phenotypic expression of the fragile X syndrome." Am J Hum Genet 53: 800-809.

Meijer, H., E. De Graaff, et al. (1994). "A deletion of 1.6 kb proximal to the CGG repeat of the FMR1 gene causes the clinical phenotype of the Fragile X syndrome." Hum Mol Genet 3(4): 615-620.

Monckton, D. G., M. I. Coolbaugh, et al. (1997). "Hypermutable myotonic dystrophy CTG repeats in transgenic mice." Nature Genet 15: 193-196.

Moutou, C., M. C. Vincent, et al. (1997). "Transition from premutation to full mutation in fragile X syndrome is likely to be prezygotic." Hum Mol Genet 6(7): 971-979.

Musco, G., A. Kharrat, et al. (1997). "The solution structure of the first KH domain of FMR1, the protein responsible for the fragile X syndrome." Nature Structural Biology 4(9): 712-716.

Musco, G., G. Stier, et al. (1996). "Three-dimensional structure and stability of the KH domain: molecular insights into the fragile X syndrome." Cell 85: 237-245.

Musumeci, S. A., P. Bosco, et al. (2000). "Audiogenic seizures susceptibility in transgenic mice with fragile X syndrome." Epilepsia 41(1): 19-23.

Musumeci, S. A., R. J. Hagerman, et al. (1999). "Epilepsy and EEG findings in males with fragile X syndrome." Epilepsia 40(8): 1092-9.

Oberlé, I., F. Rousseau, et al. (1991). "Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome." Science 252(5010): 1097-1102.

Oostra, B. A., P. B. Jacky, et al. (1993). "Guidelines for the diagnosis of fragile X syndrome." J Med Genet 30(5): 410-413.

Parniewski, P., A. Jaworski, et al. (2000). "Length of CTG.CAG repeats determines the influence of mismatch repair on genetic instability." Journal of Molecular Biology 299(4): 865-74.

Pearson, C. E., E. E. Eichler, et al. (1998). "Interruptions in the triplet repeats of SCA1 and FRAXA reduce the propensity and complexity of slipped strand DNA (S-DNA) formation." Biochemistry 37(8): 2701-8.

Pieretti, M., F. P. Zhang, et al. (1991). "Absence of expression of the FMR-1 gene in fragile X syndrome." Cell 66(4): 817-822.

Price, D. K., F. Zhang, et al. (1996). "The chicken FMR1 gene is highly conserved with a CCT 5'-untranslated repeat and encodes an RNA-binding protein." Genomics 31: 3-12.

Reyniers, E., L. Vits, et al. (1993). "The full mutation in the FMR-1 gene of male fragile X patients is absent in their sperm." Nature Genet 4(2): 143-146.

Richards, R. I. and G. R. Sutherland (1992). "Heritable unstable DNA sequences." Nature Genet 1(1): 7-9.

Rousseau, F., D. Heitz, et al. (1991). "Direct diagnosis by DNA analysis of the fragile X syndrome of mental retardation." N Engl J Med 325(24): 1673-1681.

Rudelli, R. D., W. T. Brown, et al. (1985). "Adult fragile X syndrome. Clinico-neuropathologic findings." Acta Neuropathol 67: 289-295.

Salat, U., B. Bardoni, et al. (2000). "Increase of FRMP expression, raised levels of FMR1 mRNA, and clonal selection in proliferating cells with unmethylated fragile X repeat expansions: a clue to the sex bias in the transmission of full mutations?" J Med Genet 37(11): 842-850.

Samadashwily, G. M., G. Raca, et al. (1997). "Trinucleotide repeats affect DNA replication in vivo." Nat Genet 17(3): 298-304.

Sato, T., M. Oyake, et al. (1999). "Transgenic mice harboring a full-length human mutant DRPLA gene exhibit age-dependent intergenerational and somatic instabilities of CAG repeats comparable with those in DRPLA patients." Hum Mol Genet 8(1): 99-106.

Schumacher, S., R. P. P. Fuchs, et al. (1998). "Expansion of CTG repeats from human disease genesis dependent upon replication mechanisms in Escherichia coli: The effect of long patch mismatch repair revisited." J Mol Biol 279(5): 1101-1110.

Schweitzer, J. K. and D. M. Livingston (1998). "Expansions of CAG repeat tracts are frequent in a yeast mutant defective in Okazaki fragment maturation." Hum Mol Genet 7(1): 69-74.

Schwemmle, S. (1999). "In vivo footprinting analysis of the FMR1 gene: proposals concerning gene regulation in high-functioning males." Am J Med Genet 84(3): 266-7.

Schwemmle, S., E. de Graaff, et al. (1997). "Characterization of FMR1 promoter elements by in vivo-footprinting analysis." Am J Hum Genet 60(6): 1354-62.

Seznec, H., A. S. Lia-Baldini, et al. (2000). "Transgenic mice carrying large human genomic sequences with expanded CTG repeat mimic closely the CM CTG repeat intergenerational and somatic instability." Human Molecular Genetics 9(8): 1185-94.

Shelbourne, P. F., N. Killeen, et al. (1999). "A Huntington'sdisease CAG expansion at the Hdh locus is unstable and associated with behavioural abnormalities in mice." Hum Mol Genet 8(5): 763-74.

Sherman, S. L., P. A. Jacobs, et al. (1985). "Further segregation of the fragile X syndrome with special reference to transmitting males." Hum Genet 69: 289-299.

Shimizu, M., R. Gellibolian, et al. (1996). "Cloning and characterization, and properties of plasmids containing CGG triplet repeats from the FMR-1 gene." J Mol Biol 258: 614-626.

Siomi, H., M. Choi, et al. (1994). "Essential role for KH domains in RNA binding: impaired RNA binding by a mutation in the KH domain of FMR1 that causes fragile X syndrome." Cell 77(1): 33-39.

Siomi, H., M. C. Siomi, et al. (1993). "The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein." Cell 74(2): 291-298.

Siomi, M. C., H. Siomi, et al. (1995). "FXR1, an autosomal homolog of the fragile X mental retardation gene." EMBO J 14(11): 2401-2408.

Sittler, A., D. Devys, et al. (1996). "Alternative splicing of exon 14 determines nuclear or cytoplasmic localisation of FMR1 protein isoforms." Hum Mol Genet 5(1): 95-102.

Slegtenhorst-Eegdeman, K. E., H. J. G. van de Kant, et al. (1998). "Macro-orchidism in FMR1 knockout mice is caused by increased Sertoli cell proliferation during testis development." Endocrinology 139: 156-162.

Smeets, H., A. Smits, et al. (1995). "Normal phenotype in two brothers with a full FMR1 mutation." Hum Mol Genet 4(11): 2103-2108.

Spiro, C., R. Pelletier, et al. (1999). "Inhibition of FEN-1 processing by DNA Secundary structure at trinucleotide repeats." Molecular Cell 4: 1079-85.

Sutcliffe, J. S., D. L. Nelson, et al. (1992). "DNA methylation represses FMR-1 transcription in fragile X syndrome." Hum Mol Genet 1(6): 397-400.

Sutherland, G. R., A. Gedeon, et al. (1991). "Prenatal diagnosis of fragile X syndrome by direct detection of the unstable DNA sequence." N Engl J Med 325(24): 1720-1722.

Suzumori, K., M. Yamauchi, et al. (1993). "Prenatal diagnosis of a hypermethylated full fragile X mutation in chorionic villi of a male fetus." J Med Genet 30(9): 785-7.

Tamanini, F., C. Bontekoe, et al. (1999). "Different targets for the fragile X-related proteins revealed by their distinct nuclear localizations." Hum Mol Genet 8(5): 863-869.

Tamanini, F., L. L. Kirkpatrick, et al. (2000). "The fragile X-related proteins FXR1P and FXR2P contain a functional nucleolar-targeting signal equivalent to the HIV-1 regulatory proteins [In Process Citation]." Hum Mol Genet 9(10): 1487-93.

Tamanini, F., R. Willemsen, et al. (1997). "Differential expression of FMR1, FXR1 and FXR2 proteins in human brain and testis." Hum Mol Genet 6: 1315-1322.

Tassone, F., R. J. Hagerman, et al. (2000a). "Fragile X males with unmethylated, full mutation trinucleotide repeat expansions have elevated levels of FMR1 messenger RNA." Am J Med Genet 94(3): 232-6.

Tassone, F., R. J. Hagerman, et al. (2000b). "Elevated levels of FMR1 mRNA in carrier males: A new mechanism of involvement in the Fragile-X syndrome." Am J Hum Genet 66(1): 6-15.

Tassone, F., R. J. Hagerman, et al. (2000c). "Clinical involvement and protein expression in individuals with the FMR1 premutation." Am J Med Genet 91(2): 144-152.

Taylor, A. K., F. Tassone, et al. (1999). "Tissue heterogeneity of the FMR1 mutation in a high-functioning male with fragile X syndrome." Am J Med Genet 84(3): 233-9.

Turner, G., A. Daniel, et al. (1980). "X-linked mental retardation, macro-orchidism, and the Xq27 fragile site." J Pediatr 96: 837-841.

Turner, G., C. Eastman, et al. (1975). "X-linked mental retardation associated with macro-orchidism." J Med Genet 12(4): 367-71.

Turner, G., T. Webb, et al. (1996). "Prevalence of fragile X syndrome." Am J Med Genet 64(1): 196-197.

Usdin, K. and K. J. Woodford (1995). "CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis in vitro." Nucleic Acids Res 23(20): 4202-4209.

Verheij, C., C. E. Bakker, et al. (1993). "Characterization and localization of the FMR-1 gene product associated with fragile X syndrome." Nature 363(6431): 722-724.

Verheij, C., E. De Graaff, et al. (1995). "Characterization of FMR1 proteins isolated from different tissues." Hum Mol Genet 4(5): 895-901.

Verkerk, A. J., E. De Graaff, et al. (1993). "Alternative splicing in the fragile X gene FMR1." Hum Mol Genet 2(4): 399-404.

Verkerk, A. J., M. Pieretti, et al. (1991). "Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome." Cell 65(5): 905-914.

Wan, L., T. C. Dockendorff, et al. (2000). "Characterization of dFMR1, a Drosophila melanogaster Homolog of the Fragile X Mental Retardation Protein." Molecular and Cellular Biology 20(22): 8536-8547.

Weiler, I. J. and W. T. Greenough (1999). "Synaptic synthesis of the Fragile X protein: possible involvement in synapse maturation and elimination." Am J Med Genet 83(4): 248-52.

Weiler, I. J., S. A. Irwin, et al. (1997). "Fragile X mental retardation protein is translated near synapses in response to neurotransmitter activation." Proc Natl Acad Sci USA 94: 5395-5400.

Weisman-Shomer, P., E. Cohen, et al. (2000). "Interruption of the fragile X syndrome expanded sequence d(CGG)(n) by interspersed d(AGG) trinucleotides diminishes the formation and stability of d(CGG)(n) tetrahelical structures." Nucleic Acids Res 28(7): 1535-41.

Wells, R. D. (1996). "Molecular basis of genetic instability of triplet repeats." J Biol Chem 271(6): 2875-2878.

Wen, W., J. L. Meinkoth, et al. (1995). "Identification of a signal for rapid export of proteins from the nucleus." Cell 82: 463-473.

Wheeler, V. C., W. Auerbach, et al. (1999). "Length-dependent gametic CAG repeat instability in the Huntington's disease knock-in mouse." Hum Mol Genet 8(1): 115-122.

White, P. J., R. H. Borts, et al. (1999). "Stability of the human fragile X (CGG)(n) triplet repeat array in saccharomyces cerevisiae deficient in aspects of DNA metabolism [In Process Citation]." Mol Cell Biol 19(8): 5675-5684.

Willemsen, R., B. Anar, et al. (1999). "Noninvasive test for Fragile X syndrome, using hair root analysis." Am J Hum Genet 65(1): 98-103.

Willemsen, R., C. Bontekoe, et al. (1996). "Association of FMRP with ribosomal precursor particles in the nucleolus." Biochem Biophys Res Comm 225: 27-33.

Willemsen, R., F. Los, et al. (1997a). "Rapid antibody test for prenatal diagnosis of fragile X syndrome on amniotic fluid cells: A new appraisal." J Med Genet 34(3): 250-251.

Willemsen, R., S. Mohkamsing, et al. (1995). "Rapid antibody test for fragile X syndrome." The Lancet 345: 1147-1148.

Willemsen, R., R. Olmer, et al. (2000). "Twin sisters : monozygotic with the fragile X mutation, but with a different phenotype." J Med Genet 37: 603-604.

Willemsen, R., A. Smits, et al. (1997b). "Rapid antibody test for diagnosing fragile X syndrome: A validation of the technique." Hum Genet 99(3): 308-311.

Wisniewski, K. E., S. M. Segan, et al. (1991). "The fra(X) syndrome: neurological, electrophysiological and neuropathological abnormalities." Am J Med Genet 38: 476-480.

Wöhrle, D., I. Hennig, et al. (1993). "Mitotic stability of fragile X mutations in differentiated cells indicates early post-conceptional trinucleotide repeat expansion." Nature Genet 4(2): 140-142.

Wohrle, D., U. Salat, et al. (1998). "Unusual mutations in high functioning fragile X males: apparent instability of expanded unmethylated CGG repeats." J Med Genet 35(2): 103-11.

Wohrle, D., S. Schwemmle, et al. (1996). "DNA methylation and triplet repeat stability: New proposals addressing actual questions on the CGG repeat of fragile X syndrome - Letter to the editor." Am J Med Genet 64(2): 266-267.

Yu, S., M. Pritchard, et al. (1991). "Fragile X genotype characterized by an unstable region of DNA." Science 252(5010): 1179-1181.

Zhang, Y., J. P. Oconnor, et al. (1995). "The fragile X mental retardation syndrome protein interacts with novel homologs FXR1 and FXR2." EMBO J 14(21): 5358-5366.

Zoghbi, H. Y. and H. T. Orr (1999). "Polyglutamine diseases: protein cleavage and aggregation." Curr. Opin. neurobiol. 9: 566-570.

Chapter 5

Experimental work

Chapter 5.1

FMR1 Premutation Allele $(CGG)_{81}$ Is Stable in Mice

Carola J M Bontekoe, Esther de Graaff, Ingeborg M Nieuwenhuizen,
Rob Willemsen, Ben A Oostra

MGC Department of Clinical Genetics, Erasmus University, Rotterdam, The Netherlands

Eur. J. Hum. Genet. (1997) 5: 293-298

Abstract

Fragile X syndrome is caused by an expansion of the CGG repeat present in the 5' UTR of the *FMR1* gene. A lot has been elucidated about the genetics of the disease, but not much is known about the mechanisms involved in repeat instability. Transgenic animals with a premutation allele [$(CGG)_{11}AGG(CGG)_{60}CAG(CGG)_8$] in the human *FMR1* promoter were generated to study the inheritance of this repeat in mice. Three independent lines B6, B7 and B29, in total 263 transgenic animals, were tested for repeat instability. In all meiosis and mitosis tested the repeat inherited stably. This suggests that other factors might be important in repeat (in)stability.

Introduction

Since 1991 more then 12 diseases have been identified that are caused by amplifications of trinucleotide repeats. Although we have learned a lot about the genetics of the "trinucleotide repeat diseases" not much is known about the mechanism of repeat expansion. The fragile X syndrome, one of these trinucleotide disease diseases, is an X-linked disorder affecting 1 in 4000 males and 1 in 6000 females (Turner *et al.* 1996). The main characteristics of the syndrome are mental retardation and macroorchidism (Hagerman 1996). The trinucleotide repeat causing the disease is a CGG repeat present in the 5' untranslated region (UTR) of the Fragile X Mental Retardation 1 gene *(FMR1)* (Verkerk *et al.* 1991). Expansions of this repeat above 200 triplets, full mutations, cause methylation of the CpG island including the promoter region and the repeat itself (Vincent *et al.* 1991; Hansen *et al.* 1992). This methylation blocks the transcription and therefore results in the lack of the Fragile X Mental Retardation Protein (FMRP) (Sutcliffe *et al.* 1992; Devys *et al.* 1993).

In non-affected individuals the highly polymorphic CGG repeat varies between 6 and 200 triplets with 29-32 triplets the most common. These alleles are divided into two classes. Normal alleles, repeats between 6 and 54 triplets, inhered stably, whereas so-called premutation alleles of 43 to 200 repeat units, show instability upon transmission (Fu *et al.* 1991; Oberlé *et al.* 1991). These premutations do not cause the disease, but they are prone to expand to full mutations in next generations.

Instable inheritance of the CGG repeat is observed via paternal and maternal transmission, but full mutations only arise upon female transmission. Inheritance via the paternal line

never gives rise to full mutation alleles. Because of this parent of origin effect it has been suggested that expansions occur during gametogenesis (Malter *et al.* 1997). However, it is also possible that expansions occur during early embryonic development. For obvious reasons data about oocytes and early embryogenesis are limited. Recently, female fetuses showing a full mutation in somatic tissues were described whereas full mutations were already present in the oocytes (Malter *et al.* 1997). Still the exact timing of the repeat expansion is not known, but it shows that the female germline is not protected for full mutation expansion.

In sperm of patients only premutations were found (Reyniers *et al.* 1993), suggesting contraction of full mutations in the immature testis. A hypothesis is that during spermatogenesis a kind of selection mechanism prevent germline precursors with a full mutation to arise or to develop into a mature spermatocyte (Malter *et al.* 1997). It has been suggested that FMRP is necessary for gametogenesis although the knock-out mice for *FMR1,* lacking Fmrp completely, are fertile (Bakker *et al.* 1994). Previously, transmission of a deletion causing fragile X syndrome through a family could be traced back to the deceased grandfather. This grandfather transmitted the deletion to three daughters indicating that the deletion was already present in his sperm cells (Meijer *et al.* 1994). These findings make it less likely that FMRP is necessary to produce spermatocytes.

Moreover, somatic instability does also occur. Most patients show different lengths of full mutations often present as a smear on Southern blot analysis (Rousseau *et al.* 1991). This shows that repeat length changes still occur in the early embryo. It has been estimated that in 40 % of the fragile X patients also a premutation is present (Nolin *et al.* 1994). These mosaics show always one premutation allele and a smear for the full mutation. From previous studies it is clear that this mosaic pattern is not caused by permanent mitotic instability. It is thought that mitotic instability can arise during a fixed window in early embryogenesis (Wöhrle *et al.* 1992; Wöhrle *et al.* 1993; Wohrle *et al.* 1996).

At this moment it is not known which mechanisms are involved in expansions, mitotic or meiotic. To study these mechanisms an animal model is required. Only in an animal model it will be possible to study gametogenesis and early embryogenesis at specific time points. To study the behavior of a premutation allele upon next generations we generated transgenic mice with a premutation allele, $(CGG)_{81}$ in the *FMR1* promoter. The *FMR1* promoter was fused to the reporter gene *LacZ*. A similar construct with a $(CGG)_{16}$ repeat showed expression similarly as the complete *FMR1* gene (Hergersberg *et al.* 1995).

Assuming that the mouse model resembles the human situation regarding repeat instability, expansion of the repeat from a premutation to a full mutation is expected to be followed by methylation of the promoter region of *FMR1* (Hansen *et al.* 1992; Hornstra *et al.* 1993). It is therefore expected that if the repeat in these transgenic mice expands to a full mutation expression of LacZ will be absent due to methylation of the *FMR1* promoter. Using this fusion gene as a transgene the instability of the repeat upon transmission was studied.

Material and Methods

Construct with the expanded CGG repeat

The construct of the *FMR1* promoter fused to the *LacZ* reporter (*FMR1/lacZ*) was kindly provided by Hergersberg (Hergersberg *et al.* 1995). The *NruI/XhoI* fragment containing a repeat of 16 CGGs in this construct was replaced by a *NruI/XhoI* fragment containing an premutation allele, present in pRN2 (Shimizu *et al.* 1996). The repeat present in this fragment was amplified from a premutation carrier $(CGG)_{81}$, $[(CGG)_{11}AGG(CGG)_{60}CAG(CGG)_8]$. This repeat contains a pure $(CGG)_{60}$ tract. This length is sufficient to cause instability in humans. Downstream of the pure $(CGG)_{60}$ repeat there is a CAG interruption and a $(CGG)_8$.

Transgenic mice

The transgene was isolated from agarose gel after a *NotI* digest. For microinjection a transgene concentration of 5 ng/µl was used. The transgene was injected in the pronuclei of fertilized oocytes of FVB mice. After microinjection the oocytes were transferred to a BCBA foster.

Transgenic animals were identified with PCR performed at tail DNA. PCR conditions were 20" 94°C, 20" 65°C and 45" 72°C and primers E (5' TGGGCCTCGAGCGCCCGCAGCCCA CCTCTC 3') and AGß (5' GGGATGTGCTGCAAGGCGATTTAAG 3') were used. The product was visualized on a 2 % agarose gel.

Repeat stability

To determine the repeat length in the transgenic animals radioactive PCR was performed with primers C (5' GCTCAGCTCCGTTTCGGTTTCACTTCCGGT 3') and ACß (5' GTACCCGGGGATCCTCTAGCGCCGGGA 3'). PCR conditions were as described by

Deelen (Deelen *et al.* 1994). The PCR products were run on a polyacrylamide gel. Small changes in repeat length could be detected in this way.
To determine whether there were large expansions of the repeat Southern blot analysis was performed for a number of transgenic animals.

Results

Transgenic animals with enlarged CGG repeat

Unstable CGG repeats in humans of a premutation size have a length between 43 and 200 repeat units. To introduce a repeat length in this range into mouse, we have replaced the normal CGG repeat in the *FMR1* promoter by a repeat of 81 repeat units. The *FMR1* promoter was located upstream a *LacZ* reporter gene (fig1).

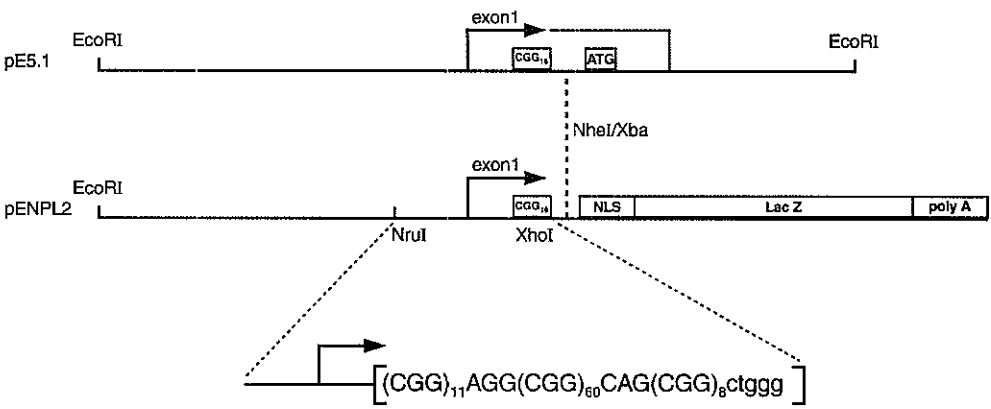

Figure 1. Schematic representation of pENPL2 containing the *FMR1* promoter region with $(CGG)_{81}$ repeat upstream of the *LacZ* reporter gene.

Transgenic mice were obtained by microinjection of the construct pENPL2 into fertilized eggs. Tail DNA from pups was tested for the presence of the transgene and the number of CGG repeats. Three independent transgenic lines B6, B7 and B29 were obtained. The copy number of the transgene was one for B7 and B29 and two for B6. This was determined by Southern blot analysis (data not shown). The repeat length in the founder mice was exactly 81 repeat units as in the original DNA construct used for microinjection. In order to study the

stability of the repeat upon germline transmission the transgenic founders were crossed with wildtype FVB mice. Transgenics of the first generation were crossed with transgenic litter mates if possible. When there were no transgenic litter mates available transgenic animals were crossed with wildtype FVBs. Line B6 and B7 were crossed till the seventh generation, line B29 was crossed till the sixth generation. For each line the transgene is situated on one of the autosomes as was concluded from the ratio of male and female transgenic progeny.

In total 263 transgenic animals were identified as shown in table 1. Because of the parent of origin effect for the repeat instability seen in fragile X syndrome the sex of the transmitting parent is depicted. In total 25 times the repeat was transmitted by a male and 25 times by a female. When transgenic litter mates were crossed, we scored them for both male and female transmission. In this way 101 transgenic males and 98 transgenic females were generated.

	Transgenic lines			
	B6	B7	B29	total
Male transmission	9	8	8	25
Female transmission	8	10	7	25
Transgenic male offspring	31	42	28	101
Transgenic female offspring	36	34	28	98
F1 males*	11	9	15	35
F1 females*	10	9	10	29
Transgenic offspring total	88	94	81	263

* Offspring of a cross between a male/female homozygous transgenic mice and a control mouse.

Table 1 Transgenic offspring of transgenic mouse strains

In addition the most likely homozygous transgenic mice from the last generation were crossed with wildtype mice. For each line one female and one male mouse was crossed with a wildtype animal. The next generation showed 64 heterozygous mice, 35 males and 29 females.

Repeat stability

For all transgenic animals radioactive PCR was performed to score for possible (small) differences in repeat length. All transgenic animals showed exactly the same size of PCR product, indicating that the repeat inherited stably (Fig 2.). Southern blot analysis was performed on a number of samples to detect possible larger expansions especially when the radioactive PCR failed or gave a weak signal. Also by Southern blot analysis no different repeat length could be identified, indicating that the $(CGG)_{81}$ repeat inherited stably in all mitosis and meiosis tested. LacZ expression in testis was tested on the three different strains. LacZ expression was found in testis of the B7 line. Strain B29 failed to show expression due to a deletion in the *LacZ* gene in the transgenic mice. Strain B6 did not show LacZ expression for unknown reasons. No methylation of the *FMR1* promoter region was detected in this transgenic strain (data not shown).

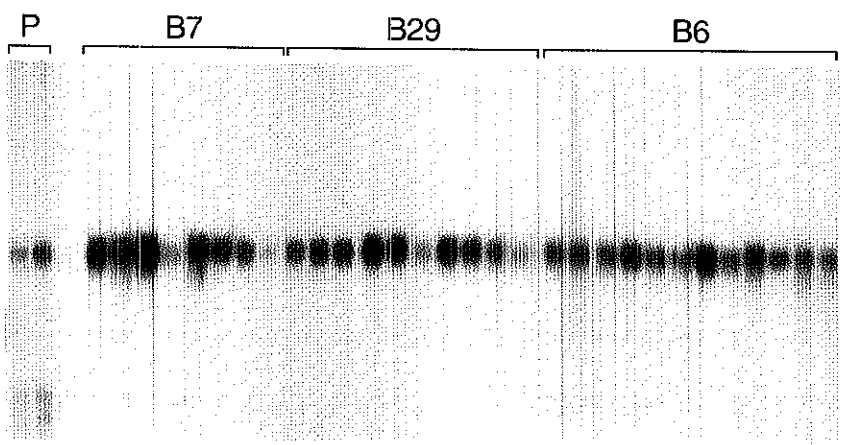

Figure 2. The size of the (CGG) in the tree independent transgenic lines was determined by PCR analysis. P shows the size of the (CGG) repeat in the original plasmid: $(CGG)_{81}$.

Discussion

In this study we describe the inheritance of a CGG repeat of 81 repeat units in the human *FMR1* promoter when present as a transgene. Expanded repeats of this size show instable inheritance in humans upon transmission. To our knowledge instability of the CGG repeat in mice has not been observed. We used a repeat length of 81 repeat units. The longest pure CGG tract within this repeat is 60 CGGs. This number has shown to be sufficient to cause instability in humans (Eichler *et al.* 1995a). The transgenic mice for the CGG repeat showed stable inheritance in all mitosis and meiosis studied, regardless of the sex of the transmitting parent. There might be a number of explanations for this.

In the murine *FMR1* CGG repeats of 9 to 12 repeat units have been identified in different mice strains. The size of the repeat necessary to cause instability in mice is not known. It is not known whether CGG repeat instability in mice exists. It might be a human specific phenomenon. We assumed that underlying mechanisms causing instability would be present in humans and mice. Transgenic mice containing an expanded repeat in the androgen receptor cDNA, the Huntington cDNA or the SCA1 cDNA showed stable inheritance through many meiosis (Bingham *et al.* 1995; Burright *et al.* 1995; Goldberg *et al.* 1996), but recently, studies for other trinucleotide diseases such as Huntington and myotonic dystrophy have shown that small changes in repeat length can occur (Gourdon *et al.* 1997; Mangiarini *et al.* 1997; Monckton *et al.* 1997).

Expansions of the CGG repeat in humans are only found at the 3'end of the repeat (Eichler *et al.* 1995a; Zhong *et al.* 1995). The observation of AGG triplets interrupting the CGG tract has led to the suggestion that these interruptions provide stability to the repeat (Kunst *et al.* 1994; Eichler *et al.* 1995a). Likewise, instability might be the result from loss of AGG interruptions or from the growth of the repeat at the 3' end. It has been proposed that repeats with an uninterrupted tract of 34 triplets are becoming unstable (Eichler *et al.* 1995a). The CGG repeat in our construct contains a pure CGG tract of 60 repeat units and therefore this length should be sufficient to cause instability. Downstream of the pure $(CGG)_{60}$ tract a CAG triplet is found followed by a tract of 8 CGG repeats. This CAG interruption, although present in some mammals (Eichler *et al.* 1995b), has never been detected in humans. It is therefore likely that this CAG triplet originated during the cloning procedures, either in the *E.coli* bacteria or as a PCR artefact. It is possible that this CAG interruption, like AGG interruptions in humans has a stabilizing effect on the repeat.

The site of integration of the transgene might determine the stability of the repeat. Sequences in the interrupted chromosomal region might influence the behavior of the transgene. The *FMR1* gene is normally present at the X-chromosome, while the transgene was found to be present at one of the autosomes. Since we do not know which mechanisms are involved in CGG expansions the X chromosome might play an important role. Possibly mechanisms involved in X-inactivation might play a role in the stability of the repeat. Furthermore, flanking sequences of the *FMR1* promoter might be important in generating instability. The region we used as promoter is sufficient to allow normal gene expression as described by Hergersberg for the fusion gene (Hergersberg *et al.* 1995), but the behavior of the repeat upon replication might be influenced by additional flanking sequences. A way to circumvent these influences would be to generate a mouse model with an expanded CGG repeat in the endogenous *FMR1* gene. However, this is for technical reasons extremely difficult. Homologous recombination by itself is not difficult (Bakker *et al.* 1994), but it is hard to generate a large amount of a long CGG repeat by PCR; a repeat up to 120 units can be amplified with the technique used, but the product can then be only visualized by radioactive PCR. To achieve homologous recombination first the repeat has to be cloned in a fragment of DNA containing the promoter and then has be propegated in E.coli. The repeat in E.coli is very unstable and it is almost impossible to isolate enough DNA with a long repeat necessary for homologous recombination.

The variation of the CGG repeat is polar and it has been found that the repeat is unstable at the 3' end (Eichler *et al.* 1995a). This might be influenced by the direction of replication of the repeat sequence. The direction of replication has shown to be important for inheritance of CGG repeats in *E.coli*. When the CGG strand is in the leading strand it tends to behave more stable. The repeat is highly instable, showing deletions and expansions when the CCG strand is in the leading strand (Shimizu *et al.* 1996; Wells 1996). This might also play a role in mice and humans as well. Although we describe three independent transgenic strains, the stable inheritance can still be a consequence of the site of integration and the direction of replication in that region. It has been suggested that slippage events during replication are causing the instability of the CGG repeat. Misalignment of the repeat during replication might lead to continued polymerization or skipping of template sequences. If this is followed by an incorrect repair repeat expansion can occur. Crossing of transgenic animals carrying CGG repeats and mice deficient for mismatch repair systems might reveal pathways that play an essential role in generating repeat instability.

In this study we were not able to generate a mouse model to study the instable behavior of the CGG repeat present in the *FMR1* promoter. It is known that repeat expansions occur somewhere between gametogenesis and early embryogenesis. To study this an animal model is required. Only then it will be possible to look at specific time points during development.

Furthermore, the differences in repeat stability upon paternal and maternal transmission and the importance of the methylation of the repeat remain to be elucidated. These differences can only be studied when a lot of material at specific points of development is available. The only way to obtain this material is the presence of an animal model which will enable us to study the behavior of the repeat and possibly to identify factors involved in the mechanism of CGG repeat expansion.

Acknowledgments

We like to thank Dr Hergersberg for the generous gift of plasmid *FMR1/lacZ*, and C.Bakker and S.Mohkamsing for technical assistance. This study was supported by the Foundation of Clinical Genetics (to E.d.G), by a BIOMED2 grant (to B.A.O) and by the Netherlands Organisation for Scientific Research (NWO) (to C.B).

References

Bakker, C. E., C. Verheij, et al. (1994). "Fmr1 knockout mice: A model to study fragile X mental retardation." Cell **78**: 23-33.

Bingham, P. M., M. O. Scott, et al. (1995). "Stability of an expanded trinucleotide repeat in the androgen receptor gene in transgenic mice." Nature Genet **9**: 191-196.

Burright, E. N., H. B. Clark, et al. (1995). "*SCA1* transgenic mice: A model for neurodegeneration caused by an expanded CAG trinucleotide repeat." Cell **82**: 937-948.

Deelen, W., C. Bakker, et al. (1994). "Conservation of CGG region in FMR1 gene in mammals." Am J Med Genet **51**(4): 513-516.

Devys, D., Y. Lutz, et al. (1993). "The FMR-1 protein is cytoplasmic, most abundant in neurons and appears normal in carriers of a fragile X premutation." Nature Genet **4**(4): 335-340.

Eichler, E. E., H. A. Hammond, et al. (1995a). "Population survey of the human FMR1 CGG repeat substructure suggests biased polarity for the loss of AGG interruptions." Hum Mol Genet **4**(12): 2199-2208.

Eichler, E. E., C. B. Kunst, et al. (1995b). "Evolution of the cryptic FMR1 CGG repeat." Nature Genet **11**(3): 301-308.

Fu, Y. H., D. P. Kuhl, et al. (1991). "Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox." Cell **67**(6): 1047-1058.

Goldberg, Y. P., M. A. Kalchman, et al. (1996). "Absence of disease phenotype and intergenerational stability of the CAG repeat in transgenic mice expressing the human Huntington disease transcript." Hum Mol Genet **5**(2): 177-185.

Gourdon, G., F. Radvanyi, et al. (1997). "Moderate intergenerational and somatic instability of a 55-CTG repeat in transgenic mice." Nature Genet **15**: 190-192.

Hagerman, R. J. (1996). Physical and behavioral phenotype. Fragile X syndrome: diagnosis, treatment and research. R. J. Hagerman and A. C. Silverman. Baltimore and London, The John Hopkins University Press: 3-87.

Hansen, R. S., S. M. Gartler, et al. (1992). "Methylation analysis of CGG sites in the CpG island of the human FMR1 gene." Hum Mol Genet **1**(8): 571-578.

Hergersberg, M., K. Matsuo, et al. (1995). "Tissue-specific expression of a FMR1/beta-galactosidase fusion gene in transgenic mice." Hum Mol Genet **4**(3): 359-366.

Hornstra, I. K., D. L. Nelson, et al. (1993). "High resolution methylation analysis of the FMR1 gene trinucleotide repeat region in fragile X syndrome." Hum Mol Genet **2**(10): 1659-1665.

Kunst, C. B. and S. T. Warren (1994). "Cryptic and polar variation of the fragile X repeat could result in predisposing normal alleles." Cell **77**(6): 853-861.

Malter, H. E., J. C. Iber, et al. (1997). "Characterization of the full fragile X syndrome mutation in fetal gametes." Nature Genet **15**: 165-169.

Mangiarini, L., K. Sathasivam, et al. (1997). "Instability of highly expanded CAG repeats in mice transgenic for the Huntington's disease mutation." Nature Genet **15**: 197-200.

Meijer, H., E. De Graaff, et al. (1994). "A deletion of 1.6 kb proximal to the CGG repeat of the FMR1 gene causes the cli7nical phenotype of the Fragile X syndrome." Hum Mol Genet **3**(4): 615-620.

Monckton, D. G., M. I. Coolbaugh, et al. (1997). "Hypermutable myotonic dystrophy CTG repeats in transgenic mice." Nature Genet **15**: 193-196.

Nolin, S. L., A. Glicksman, et al. (1994). "Mosaicism in fragile X affected males." Am J Med Genet **51**(4): 509-512.

Oberlé, I., F. Rousseau, et al. (1991). "Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome." Science **252**(5010): 1097-1102.

Reyniers, E., L. Vits, et al. (1993). "The full mutation in the FMR-1 gene of male fragile X patients is absent in their sperm." Nature Genet **4**(2): 143-146.

Rousseau, F., D. Heitz, et al. (1991). "Direct diagnosis by DNA analysis of the fragile X syndrome of mental retardation." N Engl J Med **325**(24): 1673-1681.

Shimizu, M., R. Gellibolian, et al. (1996). "Cloning and characterization, and properties of plasmids containing CGG triplet repeats from the FMR-1 gene." J Mol Biol **258**: 614-626.

Sutcliffe, J. S., D. L. Nelson, et al. (1992). "DNA methylation represses FMR-1 transcription in fragile X syndrome." Hum Mol Genet **1**(6): 397-400.

Turner, G., T. Webb, et al. (1996). "The prevalence of the fragile X syndrome." Am J Med Genet **64**: 196-197.

Verkerk, A. J., M. Pieretti, et al. (1991). "Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome." Cell **65**(5): 905-914.

Vincent, A., D. Heitz, et al. (1991). "Abnormal pattern detected in fragile-X patients by pulsed-field gel electrophoresis." Nature **349**(6310): 624-626.

Wells, R. D. (1996). "Molecular basis of genetic instability of triplet repeats." J Biol Chem **271**(6): 2875-2878.

Wöhrle, D., I. Hennig, et al. (1993). "Mitotic stability of fragile X mutations in differentiated cells indicates early post-conceptional trinucleotide repeat expansion." Nature Genet **4**(2): 140-142.

Wöhrle, D., M. C. Hirst, et al. (1992). "Genotype mosaicism in fragile X fetal tissues." Hum Genet **89**(1): 114-6.

Wohrle, D., S. Schwemmle, et al. (1996). "DNA methylation and triplet repeat stability: New proposals addressing actual questions on the CGG repeat of fragile X syndrome - Letter to the editor." Am J Med Genet **64**(2): 266-267.

Zhong, N., W. H. Yang, et al. (1995). "Fragile X gene instability: Anchoring AGGs and linked microsatellites." Am J Hum Genet **57**(2): 351-361.

Chapter 5.2

Instability of an $(CGG)_{98}$ Repeat in the *Fmr1* Promoter

C.J.M. Bontekoe, C.E. Bakker, I.M. Nieuwenhuizen, H. van der Linde, H. Lans,
D. de Lange, M.C.Hirst[1], B.A. Oostra.

CBG Department of Clinical Genetics, Erasmus University Rotterdam, The Netherlands.
[1]Institute of Molecular Medicine, University of Oxford, United Kingdom.

submitted

Abstract

Fragile X syndrome is one of the fourteen known trinucleotide repeat diseases. Expansions occurring in the CGG repeat present in the 5'UTR of the *FMR1* gene can cause the fragile X phenotype, characterised by mental retardation. The mechanisms involved in trinucleotide repeat expansions are poorly understood. Transgenic mouse models containing an expanded CGG repeat present in a transgene failed to show instability as observed in humans. As cis-acting factors and the endogenous genomic localisation of the expanded CGG repeat are thought to play a role, we here report the generation of a knock-in expanded $(CGG)_{98}$ repeat present in the mouse *Fmr1* gene. In these mice the murine $(CGG)_8$ repeat was exchanged by a human $(CGG)_{98}$ repeat. This model shows moderate instability upon transmission, and will enable studies on the timing and the mechanism of repeat expansion in mice.

Introduction

The fragile X syndrome is one of the 14 identified trinucleotide repeat diseases. Although a lot has been elucidated about the genetics of the trinucleotide repeat diseases not much is known about the mechanism(s) causing the repeat instability. The highly polymorphic CGG trinucleotide repeat located in the 5' untranslated region (UTR) of the fragile X mental retardation gene (*FMR1*) is associated with the disease phenotype when the allele contains more than 200 triplets (Oberlé *et al.* 1991; Verkerk *et al.* 1991; Yu *et al.* 1991). A repeat of more than 200 triplets coincides with methylation of the CpG island including the promoter region and the CGG repeat. This methylation blocks the transcription and thus translation into the fragile X mental retardation protein (FMRP) (Pieretti *et al.* 1991; Verheij *et al.* 1993). This absence of FMRP results in the fragile X phenotype. The main characteristics of the fragile X syndrome are mental retardation and macroorchidism (Hagerman 1996) . In adult males, mental retardation ranges from profound to borderline with an average IQ in the moderate range. Macroorchidism is a common finding in post-pubescent affected males. With an incidence of 1:4000 males and 1:6000 females this X-linked disorder is the most common form of inherited mental retardation (Turner *et al.* 1996; De Vries *et al.* 1997).
The CGG repeat alleles can be divided into three groups: normal, premutation and full mutation alleles. The normal alleles range between 5 and 50 triplets. These alleles behave

stable upon transmission to next generations. Premutations are alleles of 50 to 200 triplets (Fu *et al.* 1991; Rousseau *et al.* 1991). These repeats are expanded and they can behave unstable upon transmission to next generations. Both expansions and contractions can occur. Since they allow FMRP expression (Devys *et al.* 1993; Verheij *et al.* 1993) they do not cause the fragile X phenotype, but they are prone to expand to full mutations of more than 200 triplets in the next generation. Only full mutations result in the absence of FMRP and thus the fragile X phenotype (Devys *et al.* 1993; Verheij *et al.* 1993). Full mutations only arise upon transmission through the female germ line. Males never transmit a full mutation to their daughters.

The exact timing of the repeat expansion is still under debate. Repeat expansions have to occur somewhere during gametogenesis or early embryonic development (Malter *et al.* 1997; Moutou *et al.* 1997). The most accepted model assumes a full mutation to be present already in the oocyte. In this way all cells in the embryo will have a full mutation. One or a few events of mitotic regression to a premutation can explain the mosaic pattern quite often found in patients. For the male germline, where it is found that patients with a full mutation only show a premutation in their sperm, some kind of selection mechanism has to be assumed (Reyniers *et al.* 1993). This mechanism protects the male germ line against transmitting full mutations. The basis of such a selection mechanism is not known. The observation that in oocytes of full mutation female fetuses a full mutation was found confirms this model, although it can not be excluded that the expansion from a premutation to a full mutation occurs during early embryogenesis (Malter *et al.* 1997; Moutou *et al.* 1997). The repeat length(s) present in oocytes of premutation females is not known.

Examining the mode of inheritance of premutations in fragile X families showed that the risk of expansion to a full mutation increases with the repeat size of the premutation. Small premutations behave unstable, both expansions and contractions are observed. Premutations of 90 or more triplets almost always expand to a full mutation in the next generation. The risk of expansion to a full mutation increases with the length of the CGG repeat (Bat *et al.* 1997). This variation in risk accounts for the Sherman paradox (Sherman *et al.* 1985; Fu *et al.* 1991). More detailed insight in the repeat length and its behaviour upon transmission was gained by sequencing a large number of normal and premutation alleles. It was found that most normal alleles were interspersed by AGG triplets (Kunst *et al.* 1994). The most common alleles of 29 to 32 triplets are interspersed by 2 AGG interruptions. These AGG interruptions are normally found downstream tracts of 9 or 10

CGG triplets. In premutations, less AGG interruptions are present compared to normal alleles. The variation in repeat length appears polar, instability occurs always at the 3' end of the repeat, in the region where no or less AGG interruptions are present (Kunst *et al.* 1994; Eichler *et al.* 1995). The most 3' pure CGG tract is the most important element of the CGG repeat in determining the instability of the repeat. It is thought that a pure CGG tract of more than 34-38 CGG triplets is enough to cause instability (Eichler *et al.* 1994). Most premutation alleles contained only one or two AGG repeats, which means that the 3' pure CGG tract is more than 35 CGG triplets. Since the longest pure CGG tract is always found at the 3' end of the CGG repeat, the region where instability occurs in fragile X families, this might give some insight in the mechanism of instability.

To study the timing and the mechanism of the CGG repeat expansion observed in the fragile X syndrome, it is for obvious reasons necessary to have an animal model. In that way it would be possible to study the behaviour of the repeat through the female and male germline and during (early) embryonic development. We previously reported that a [$(CGG)_{11}$ $AGG(CGG)_{60}CAG(CGG)_8$] repeat was transmitted without detectable changes through several generations of transgenic mice (Bontekoe *et al.* 1997). Similar results were obtained in mice with a [$(CGG)_{22}TGG(CGG)_{43}TGG(CGG)_{21}$] tract and with tracts having a pure 3' CGG tract up to 97 triplets (Lavedan *et al.* 1997; Lavedan *et al.* 1998). Several hypotheses were put forward trying to explain the stability of these repeats in mice. Several hypotheses involved the structure of the repeat tract. The interruptions present in the two first transgenes were mentioned as stabilisers, but the last report using a pure CGG tract at the 3' end showed that other factors might be involved as well. Other possible explanations concerned the importance of the genomic localisation of the CGG repeat.

FMR1 is highly conserved among vertebrates. The murine homologue *Fmr1* is 97% identical in amino acid sequence and exhibits an expression pattern similar to that observed in humans (Ashley *et al.* 1993; Bakker *et al.* 2000) .The homology extends to the repeat region and the promoter region. We generated a mouse model in which the endogenous mouse CGG repeat was replaced by a human CGG repeat of 98 CGG triplets. This was done using a homologous recombination strategy using an DNA construct where the mouse endogenous $(CGG)_8$ repeat was exchanged for an $(CGG)_{98}$ repeat of human origin in the mouse promoter. With the generation of such a "knock-in" mouse model we have studied the behaviour of the premutation allele in the *Fmr1* gene.

Material and Methods

Construction of pCB66 and ES cell electroporation

The construct pCB66 contained an 7 kb *Hin*dIII fragment of the mouse *Fmr1* promoter region inserted in an pBR322 derived vector (pBR322 minus *Bam*HI 375-*Nar*I 1205). In this fragment a *neo* cassette flanked by *lox*P sites was cloned in the *Bam*HI site present in intron. The mouse $(CGG)_6$ repeat was exchanged for a $(CGG)_{98}$ repeat of human origin. The human derived $(CGG)_{98}$ repeat was isolated as an *Sfo*I/*Xho*I fragment from a cloned expanded CGG repeat. To clone the expanded human CGG repeat in the murine *Fmr1* promoter minimal changes were made to the flanking sequence. These changes involved a point mutation to abolish an *Nar*I site in intron 1 and the substitution of 4 bp (TCGA) to abolish an *Xho*I site in intron 1. The *Xho*I site flanking the CGG repeat in the human situation was not present in the mouse promoter. To facilitate the cloning of the expanded $(CGG)_{98}$ repeat an *Xho*I site was generated by an A→G transition based on the homology between the human and the mouse promoter. The integrity of the cloned fragment was determined by sequencing and restriction enzyme digestion. Cloning was performed using standard procedures.

For electroporation of E14 ES-cells to allow homologous recombination plasmid pCB66 was linearised by an *Xba*I digest. After purification linearised plasmid DNA was used to transfect ES-cells. Electroporation was performed with 10^7 ES cells in 400 µl PBS using a Progenetor II Gene Pulser (1200 µF and 117V during 10ms). Using double selection with G418 (200 µg/ml) and Fiau (2 µM) the cells were cultured to allow colony forming. Colonies were picked and cultured separately. From each clone one batch was frozen and one batch was used for DNA isolation.

DNA analysis

ES cell clones were grown to confluency in a 24-wells plate. The medium was removed and 300 µl (10mM Tris-(HCl), 400mM NaCl, 2mM EDTA pH7.3-7.4) and 30 µl 10mg/ml Prot K was added to lyse the cells during overnight incubation at 55°C. 150 µl 6M NaCl was added and the suspension was centrifuged. To the supernatant two volumes of 96% ethanol were added to precipitate the DNA. DNA was dissolved in 50 microliter H_20. For radioactive PCR 1 µg DNA was used.

Radioactive PCR was performed to determine the repeat length in the ES-clones. Primers C (5' GCTCAGCTCCGTTTCGGTTTCACTTCCGGT 3') and F (AGCCCCGCACTTCCACCAC CAGCTCCTCCA 3') were used. PCR conditions were as described by Deelen (Deelen *et al.* 1994). PCR products were run at a 6 % denaturing polyacrylamide gel.

Generation of knockout mice

ES clone 651 was used for injection into C57/BL6J blastocysts. These blastocysts were transferred to pseudopregnant female mice. Three chimearas were generated and crossed with wt FVB and C57/Bl6J females. Female offspring of these chimearas was tested for the presence of the expanded $(CGG)_{98}$ repeat. Mice containing the expanded CGG repeat were crossed with wt mice as well as with $(CGG)_{98}$ littermates. Repeat instability upon transmission to offspring was tested by radioactive PCR.

The *neo* cassette inserted into the *Bam*HI was flanked by *lox*P sites. Because the presence of the *neo* cassette might disturb the natural environment of the CGG repeat, knock-in mice were crossed with mice expressing Cre recombinase. In this way the changes made to the *Fmr1* gene were kept minimal. These $(CGG)_{98}$/*neo*- mice were also crossed with wt mice, and with $CGG)_{98}$/*neo*- littermates.

Fragile X size polymorphism assay

Radioactive PCR as described above is informative to determine whether instability occurs, but it does not give the exact length changes observed. Also small changes, +/-1 CGG triplet might be missed. The fragile X size polymorphism assay (Perkin Elmer Biosystems) allows us to determine the exact length of the CGG repeat. This test was used to determine the exact size changes. PCR conditions were as described by the Perkin Elmer Biosystems. PCR samples were analysed at an ABI377 sequencer (PE Biosystems).

Results

Construct *FMR1* promoter region

The mouse homologue of *FMR1* was isolated from an E14 ES cell phage library (kindly provided by D. Meijer). An *neo* cassette flanked by *lox*P sites was cloned in the *Bam*HI site. The endogenous $(CGG)_8$ present in the mouse promoter was replaced by an $(CGG)_{98}$ of human origin. To allow cloning of the human CGG repeat minimal changes were made to

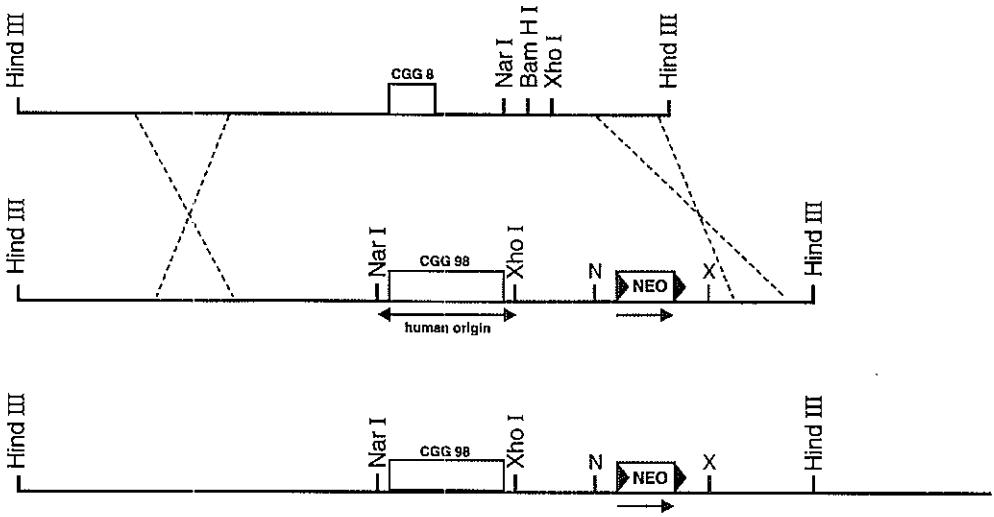

Figure 1. The targeting construct pCB 66. The mouse endogenous *Fmr1* $(CGG)_8$ is exchanged for an $(CGG)_{98}$ repeat of human origin. X= removed *Xho*I site, N= removed *Nar*I site.

the mouse promoter region (see Figure 1). Cloning of the CGG repeat was the last cloning step, because often deletions were found by propagation of a plasmid containing an CGG repeat.

Cloning of the mouse promoter region revealed that the mouse promoter region cloned into a vector and propagated in bacteria was prone to deletions too. Even without an expanded CGG repeat deletions were often observed after digestion, religation and transformation into bacteria (different strains were used). The frequency of these deletions is dependent of the restriction enzyme used or the localisation of the restriction sites in the construct, and the ligation buffer used (data not shown). Since for most DNA constructs digestion, religation and transformation is a very straightforward experiment this could be an indication that cis-acting factors important for instability to occur might be present in the mouse promoter region. Comparison of the promoter sequences of *FMR1* and *Fmr1* showed that all the identified regulatory elements were conserved. Whether or not all these elements are functional, is not known.

Generation of the mice

Homologous recombinants were recognised by the absence of the endogenous mouse $(CGG)_8$ allele, and the presence of the expanded $(CGG)_{98}$ allele. Of the 1200 screened ES-cell clones 13 clones were identified as possible homologous recombinants. The length of the observed expanded CGG repeat differed between the clones. These different lengths most likely represent deletions of the repeat and/or flanking sequences in the plasmid DNA. Clone 651 showed by PCR exactly the same length as the cloned $(CGG)_{98}$ repeat in the plasmid. Therefore this clone was selected for blastocyst injection. Three chimearas with the same repeat length were identified. These animals were crossed into FVB background and the repeat length was determined in the next generation. The first heterozygous females were crossed with a WT male or a male mouse expressing Cre-recombinase. Cre-recombinase allows deletion of the *neo* cassette present between the *lox*P sites. This recombination minimilizes the changes to the flanking sequence. Both $(CGG)_{98}$ mice with and without *neo* were crossed and checked for repeat instability.

In total 155 $(CGG)_{98}$ transmissions were studied. $(CGG)_{98}$/*neo* was transmitted 34 times, 15 male transmissions, and 19 female transmissions. $(CGG)_{98}$/*neo*- was transmitted 121 times, 80 maternal transmissions and 41 paternal transmissions. Offspring homozygous for the expanded repeat were included for paternal (13) and maternal (13) transmission. In total 15 instabilities were found using radioactive PCR (figure 2). Confirmation and exact sizing of the repeat was carried out using the CGG expansion method of Perkin-Elmer. An example is shown in figure 3. Two regressions were found and 13 expansions. Both regressions (-11) were observed after paternal transmission. Six expansion and two regression events were found after paternal transmission of the repeat. Five expansions were observed via the female germline. Two instabilities were observed in homozygous $(CGG)_{98}$ female offspring. The origin of the instability can be either paternal or maternal. These numbers are summarised in table 1. The shortest repeat is now 87 CGGs and after passages through several generations the longest repeat is 108 CGGs.

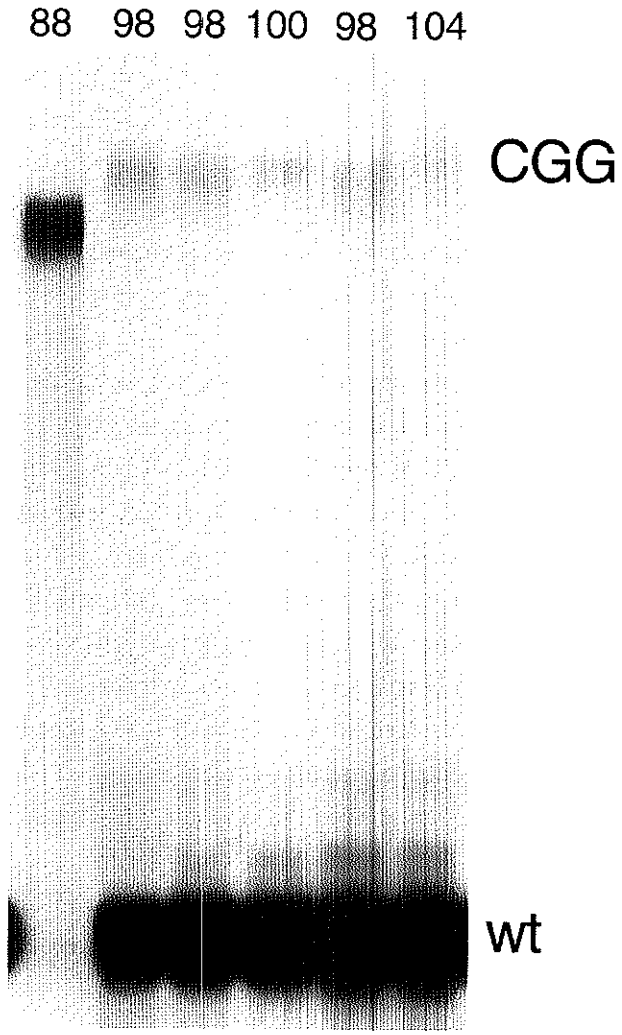

Figure 2. CGG repeat instabilities detected by radioactive PCR. On top of the lanes are the sizes as determined with the Fragile X polymorphism assay.

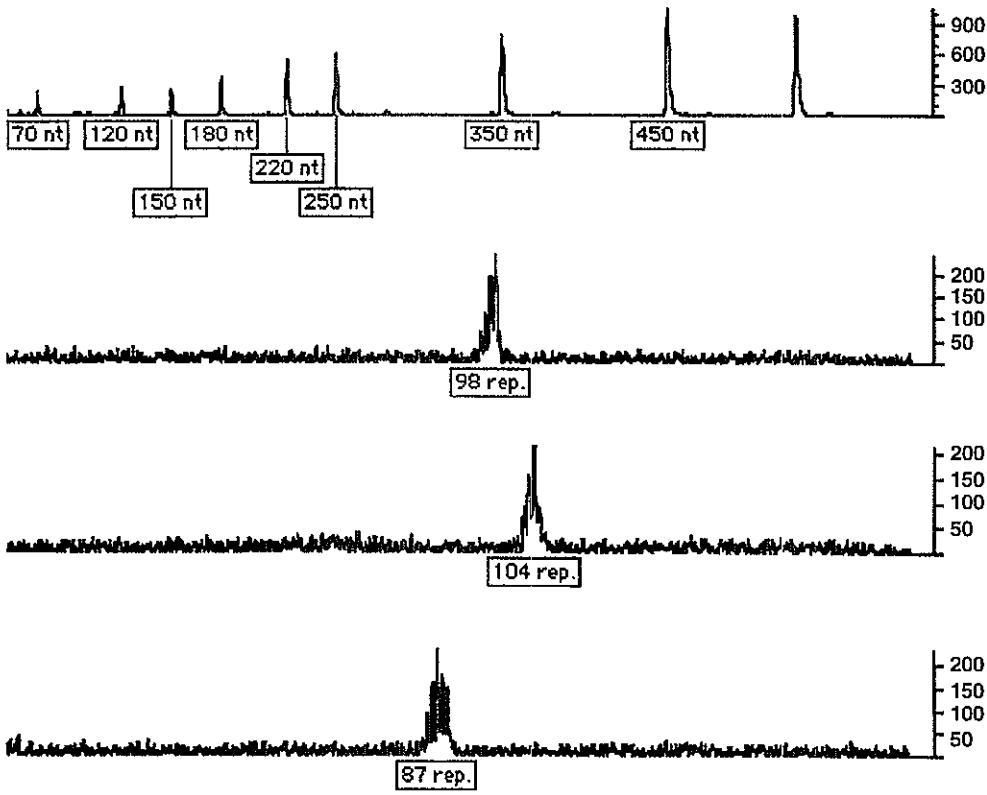

Figure 3. The Fragile X polymorphism assay is used to determine the exact repeat length of the observed instabilities.

	Female/male transmissions	**Instabilities**
$(CGG)_{98}$/*neo*	Female 19	0
$(CGG)_{98}$/*neo*	Male 15	4 (+2, +6, +4, +4)
$(CGG)_{98}$/*neo*-	Female 67	5 (+2, +2, +1, +1,+1)
$(CGG)_{98}$/*neo*-	Male 28	4 (-11, -11, +5,+3)
$(CGG)_{98}$/*neo*-	Female/Male 13/13	2(+2, +3)

Table 1. Total number of male and female transmissions for both $(CGG)_{98}$/*neo* and $(CGG)_{98}$/*neo*- are listed. Most instabilities are found upon male transmission. Only for a limited number the size of the instabilities is determined. Length changes are depicted between bracklets.

Discussion

The mechanism causing the repeat instability as observed in the fragile X syndrome is still not known. It has been postulated that during replication, slippage occurs in the repeat. Due to secondary structures, like hairpin formation or single-stranded breaks of the repeat, both expansions and contractions can be explained (Gacy *et al.* 1995; Nadel *et al.* 1995). Also the formation of tetraplex structures has been reported, but the existence of either hairpin or tetraplex structures remains to be established *in vivo* (Usdin 1998). It was shown that the instability of CGG repeats was orientation-dependent (Hirst *et al.* 1998). From *in vitro* studies as well as studies in bacteria and yeast it was obvious that the stability of the repeat was dependent on, like in fragile X families, the length of the repeat and the number of AGG interruptions (Shimizu *et al.* 1996; Weisman-Shomer *et al.* 2000). Furthermore, for yeast and bacteria it was observed that the host cell genotypes, the orientation of the repeat and the position of the cloned repeat in the vector influenced the mutation rate.

All transgenic mouse models containing an autosomal transgene with an expanded CGG repeat failed to show instabilities of the CGG repeat (Bontekoe *et al.* 1997; Lavedan *et al.* 1997; Lavedan *et al.* 1998). In the human situation the repeat size can increase over generations. Once it passes a certain threshold the repeat becomes very unstable, and can expand up to a few thousand CGG triplets. This phenomenon has given rise to the term "dynamic mutation" (Richards *et al.* 1992). Although for some trinucleotide repeat mouse models small expansions or contractions are observed, dynamic mutations are never observed. The absence of important cis-acting factors and the random integration of the transgene on one of the autosomes instead of the X chromosome were suggested as possible explanations for the relative stability of the CGG repeats in these mouse models (Bontekoe *et al.* 1997). To circumvent these influences we generated a mouse model with an expanded CGG repeat in the endogenous *Fmr1* promoter by using a homologous recombination technique. In this way the behaviour of the expanded CGG repeat can be studied in the endogenous genomic localisation.

In total 155 transmissions of the expanded $(CGG)_{98}$ allele present in the endogenous *Fmr1* promoter were studied. In total 15 instabilities were observed 2 regression events and 13 expansions. This means that instability observed is at least 10% as only differences clearly seen by radioactive PCR have been counted. Both regressions occurred by paternal transmission as well as six of the 13 expansions. Two expansions could be of either

paternal or maternal origin. Given the total number of maternal and paternal transmission these results are surprising. Although the numbers are still limited, there appears to be a tendency of higher instability upon paternal transmission. In humans, these small changes as observed in these mice are observed upon both male and female transmission.

For dynamic mutations in humans a difference is observed depending on the parental origin. Dynamic mutations are only observed upon female transmission. However, in this $(CGG)_{98}$ mouse model no dynamic mutations were observed. In contrast to the human situation for the fragile X syndrome, most other trinucleotide repeat disorders in human show instability preferentially upon male transmission. For one transgenic mouse model containing an expanded CAG repeat it was found that also the gender of the offspring determined the rate of instability (Kovtun *et al.* 2000). Whether the gender of the offspring in this CGG repeat mouse model plays a role could not be determined.

The human and mouse promoter regions are very homologous. Four footprints have been identified in the human *FMR1* promoter. They reflect DNA-protein interactions. These footprints correspond to consensus binding sites of transcription factors. In cells from fragile X patients these footprints were absent, indicating that they reflect functional regulatory elements (Schwemmle *et al.* 1997). These detected regulatory elements are also present in the mouse promoter region. However, it is not known whether these sequences in the promoter region might be influencing the behaviour of the repeat instability. Therefore, the overall homology between the mouse and the human promoter does not necessarily mean that sequences involved in repeat instability are also conserved between mouse and human. Theoretically, there is still a possibility that the mouse promoter region does not contain the cis-acting elements involved in (large) repeat instability in the human *FMR1* gene.

Cloning of the *Fmr1* promoter region showed that this region was prone to deletions in bacteria. These deletions occurred after digestion of the DNA, religation and transformation into bacteria. Most, although not all, deletions occurred in the region directly down and upstream of the $(CGG)_8$ repeat. These results indicate that the *Fmr1* promoter region itself is unstable in bacteria. The reason for this instability might be the high GC percentage in this region flanking the CGG repeat. Although this is no direct evidence, a DNA region which proved to be highly unstable in bacteria is less likely to have a stabilising effect in the mouse genome. Therefore, we hypothesise that other factors than cis-acting factors alone may play a role in determining repeat instability.

The variation found in the fragile X alleles appeared to be polar, instability occurs always at the 3' end. This might be influenced by the direction of replication (Shimizu *et al.* 1996; Hirst *et al.* 1998). The direction of replication proved to be important for determining the repeat instability in *E. coli* and yeast. The origin of replication in the DNA constructs determines whether the 5'-CGG-3' or the 5'-CCG-3' strand is the leading or the lagging strand. The importance of the location of the origin of replication can be explained by the involvement of the "slippage during replication" model. The Okazaki fragments formed on the lagging strand may form hairpin loops during replication. Reannealing in a different position on the template may leave a gap that is later filled in resulting in an increased repeat length. Hairpin loops formed in the leading strand will result in a decreased repeat length. This mechanism can explain the minor expansions and deletions observed within the premutation range. The dynamic mutations can be explained by two single strand breaks within the pure CGG tract (consistent with the size of an Okazaki fragment). Since the pure CGG strand between the two breaks is not anchored at either side the strand is free to slide during replication. Repair of this "not anchored" strand can introduce many more triplets at once. The longer the pure repeat the more likely that two single strand breaks will occur simultaneously. For the mouse and human *FMR1* loci the direction of replication is not known. The direction of replication in the human situation might favour the occurrence of expansions, the situation might be different for the mouse *Fmr1* locus.

Flap endonuclease 1 (FEN1), a protein involved in long-patch base excision repair, is thought to play a role in trinucleotide repeat instability (Tischkoff *et al.* 1997). In yeast loss of flap endonuclease activity (rad27) increases instabilities throughout the whole genome including trinucleotide repeats (Freudenreich *et al.* 1998; Schweitzer *et al.* 1998; Spiro *et al.* 1999). For normal and premutation size CGG repeats an 10-fold elevated frequency of expansion in rad27 mutant yeast strain was found (White *et al.* 1999). This suggests that FEN1 can play a role in CGG trinucleotide repeat instability.

The endogenous *Fmr1* 5' UTR contains an CGG repeat of 8-12 triplets. In our ES cells the naturally $(CGG)_8$ repeat is exchanged for an $(CGG)_{98}$ repeat. In humans repeat instability occurs when the number of CGG triplets is more than 50. The possible threshold for repeat instability in mice is not known. For mice containing an expanded Huntington CAG repeat it was found that the observed rate of instability for mice was less than observed in humans (Shelbourne *et al.* 1999). Extrapolation of the data obtained from the Huntington expanded

trinucleotide mice suggests that if there is a threshold for instability in mice the threshold might be higher in mouse than in human.

The mechanisms causing repeat instability are not known. Results of this CGG repeat expanded mouse model, together with the studies of other trinucleotide repeats suggest that mouse might not be a good model to study repeat instability. The unknown mechanism involved in repeat instability in humans might be absent in mouse, or the environment leading to the instability in humans might be absent in mice. The mechanisms, which are thought to play a role in repeat instability, involve mechanisms that are normally functional in replication, meiosis, recombination or repair. During these processes duplication of the DNA occurs. Therefore the secondary structures like hairpins or tetraplex structures must be removed. When the DNA is duplicated, proof-reading and repair occurs. These processes try to keep the changes, mistakes in the DNA to a minimum. On the other hand recombination occurs during replication and meiosis to allow variation for the species. Although these processes are known to occur in humans as well as in mice the balance between these processes might give rise to repeat instability in humans, but not in mice.

The expected involvement of repair mechanisms in repeat instability can be studied by crossing mice with expanded CGG repeats with mice deficient for different repair mechanisms. In yeast and *E. coli* it was shown that the rate of instability changed in certain repair deficient strains (Jakupciak *et al.* 1999; White *et al.* 1999; Iyer *et al.* 2000). Crossings between trinucleotide repeat mice and mice with certain repair deficiencies might give more insight in repeat instability in mice as well as in humans. Furthermore, it will be worthwhile to study the timing of the small repeat instabilities that occur. This might give more insight in the difference observed between male and female transmission.

Eventually, this $(CGG)_{98}$ knock-in mouse model may be used to study inactivation of FMRP expression. In our $(CGG)_{98}$ mice FMRP expression was present, suggesting that the promoter region was not inactivated (data not shown). As small expansions are observed, it can be possible that upon expansion over several generations a larger expanded allele can be generated in which the expression of FMRP will be inactivated. Then both the expansion of the repeat as well as the FMRP expression can be studied in next generations. Recently, it was described that premutation alleles resulted in elevated mRNA levels (Tassone *et al.* 2000). The FMRP expression in lymphocytes with premutations of more than 100 CGG triplets the level of FMRP expression was reduced. This mouse model can also be used to

study the mRNA levels in premutation alleles, the translation of these mRNAs, and maybe the mechanistic switch, which occurs when *Fmr1* is silenced.

Acknowledgements

We like to thank T. de Wit for the generous gift of the mouse transgenic for Cre recombinase, D. Meijer for the mouse ES cell phage library and R. van Haperen for blastocyst injections. This study was supported by the Netherlands Organisation for Scientific Research (N.W.O.) and the National Institute of Health (NIH 5 RO1 HD38038). We like to thank Perkin Elmer for their gift of their fragile X size polymorphism assay.

References

Ashley, C. T., J. S. Sutcliffe, et al. (1993). "Human and murine FMR-1: alternative splicing and translational initiation downstream of the CGG-repeat." Nature Genet 4(3): 244-251.

Bakker, C. E., Y. de Diego Otero, et al. (2000). "Immunocytochemical and biochemical characterization of FMRP, FXR1P, and FXR2P in the mouse." Exp Cell Res 258(1): 162-70.

Bat, O., M. Kimmel, et al. (1997). "Computer simulation of expansions of DNA triplet repeats in the fragile X syndrome and Huntington's disease." Journal of Theoretical Biology 188(1): 53-67.

Bontekoe, C. J. M., E. de Graaff, et al. (1997). "FMR1 premutation allele is stable in mice." Eur J Hum Genet 5: 293-298.

De Vries, B. B., A. M. van den Ouweland, et al. (1997). "Screening and diagnosis for the fragile X syndrome among the mentally retarded: an epidemiological and psychological survey. Collaborative Fragile X Study Group." Am J Hum Genet 61(3): 660-667.

Deelen, W., C. Bakker, et al. (1994). "Conservation of CGG region in FMR1 gene in mammals." Am J Med Genet 51(4): 513-516.

Devys, D., Y. Lutz, et al. (1993). "The FMR-1 protein is cytoplasmic, most abundant in neurons and appears normal in carriers of a fragile X premutation." Nature Genet 4(4): 335-340.

Eichler, E. E., H. A. Hammond, et al. (1995). "Population survey of the human FMR1 CGG repeat substructure suggests biased polarity for the loss of AGG interruption." Hum Mol Genet 4(12): 2199-2208.

Eichler, E. E., J. Holden, et al. (1994). "Length of uninterrupted CGG repeats determines instability in the FMR1 gene." Nature Genet 8(1): 88-94.

Freudenreich, C. H., S. M. Kantrow, et al. (1998). "Expansion and length-dependent fragility of CTG repeats in yeast." Science 279(5352): 853-6.

Fu, Y. H., D. P. Kuhl, et al. (1991). "Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox." Cell 67(6): 1047-1058.

Gacy, A. M., G. Goellner, et al. (1995). "Trinucleotide repeats that expand in human disease form hairpin structures in vitro." Cell 81: 533-540.

Hagerman, R. J. (1996). Physical and behavioral phenotype. Fragile X syndrome: diagnosis, treatment and research. R. J. Hagerman and A. C. Silverman. Baltimore and London, The John Hopkins University Press: 3-87.

Hirst, M. C. and P. J. White (1998). "Cloned human FMR1 trinucleotide repeats exhibit a length- and orientation-dependent instability suggestive of in vivo lagging strand secondary structure." Nucleic Acids Res 26(10): 2353-2358.

Iyer, R. R., A. Pluciennik, et al. (2000). "DNA polymerase III proofreading mutants enhance the expansion and deletion of triplet repeat sequences in Escherichia coli." J. Biol. Chem. 275(3): 2174-84.

Jakupciak, J. P. and R. D. Wells (1999). "Genetic instabilities in (CTG.CAG) repeats occur by recombination." J. Biol. Chem. 274(33): 23468-79.

Kovtun, I. V., T. M. Therneau, et al. (2000). "Gender of the embryo contributes to CAG instability in transgenic mice containing a Huntington's disease gene." Hum Mol Genet 9(18): 2767-75.

Kunst, C. B. and S. T. Warren (1994). "Cryptic and polar variation of the fragile X repeat could result in predisposing normal alleles." Cell 77(6): 853-861.

Lavedan, C., E. Grabczyk, et al. (1998). "Long uninterrupted CGG repeats within the first exon of the human FMR1 gene are not intrinsically unstable in transgenic mice." Genomics 50(2): 229-240.

Lavedan, C. N., L. Garrett, et al. (1997). "Trinucleotide repeats (CGG)22TGG(CGG)43TGG(CGG)21 from the fragile X gene remain stable in transgenic mice." Hum Genet 100: 407-414.

Malter, H. E., J. C. Iber, et al. (1997). "Characterization of the full fragile X syndrome mutation in fetal gametes." Nature Genet 15: 165-169.

Moutou, C., M. C. Vincent, et al. (1997). "Transition from premutation to full mutation in fragile X syndrome is likely to be prezygotic." Hum Mol Genet 6(7): 971-979.

Nadel, Y., P. Weismanshomer, et al. (1995). "The fragile X syndrome single strand d(CGG)(n) nucleotide repeats readily fold back to form unimolecular hairpin structures." J Biol Chem 270(48): 28970-28977.

Oberlé, I., F. Rousseau, et al. (1991). "Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome." Science 252(5010): 1097-1102.

Pieretti, M., F. P. Zhang, et al. (1991). "Absence of expression of the FMR-1 gene in fragile X syndrome." Cell 66(4): 817-822.

Reyniers, E., L. Vits, et al. (1993). "The full mutation in the FMR-1 gene of male fragile X patients is absent in their sperm." Nature Genet 4(2): 143-146.

Richards, R. I. and G. R. Sutherland (1992). "Dynamic mutations: a new class of mutations causing human disease." Cell 70(5): 709-712.

Rousseau, F., D. Heitz, et al. (1991). "Direct diagnosis by DNA analysis of the fragile X syndrome of mental retardation." N Engl J Med 325(24): 1673-1681.

Schweitzer, J. K. and D. M. Livingston (1998). "Expansions of CAG repeat tracts are frequent in a yeast mutant defective in Okazaki fragment maturation." Hum Mol Genet 7(1): 69-74.

Schwemmle, S., E. de Graaff, et al. (1997). "Characterization of FMR1 promoter elements by in vivo-footprinting analysis." Am J Hum Genet 60(6): 1354-62.

Shelbourne, P. F., N. Killeen, et al. (1999). "A Huntington'sdisease CAG expansion at the Hdh locus is unstable and associated with behavioural abnormalities in mice." Hum Mol Genet 8(5): 763-74.

Sherman, S. L., P. A. Jacobs, et al. (1985). "Further segregation of the fragile X syndrome with special reference to transmitting males." Hum Genet 69: 289-299.

Shimizu, M., R. Gellibolian, et al. (1996). "Cloning and characterization, and properties of plasmids containing CGG triplet repeats from the FMR-1 gene." J Mol Biol 258: 614-626.

Spiro, C., R. Pelletier, et al. (1999). "Inhibition of FEN-1 processing by DNA Secundary structure at trinucleotide repeats." Molecular Cell 4: 1079-85.

Tassone, F., R. J. Hagerman, et al. (2000). "Elevated levels of FMR1 mRNA in carrier males: A new mechanism of involvement in the Fragile-X syndrome." Am J Hum Genet 66(1): 6-15.

Tischkoff, D. X., N. Filosi, et al. (1997). "A novel mutation avoidance mechanism dependent on S. cerevisiae RAD27 is distinct from DNA mismatch repair." Cell 88: 5027-31.

Turner, G., T. Webb, et al. (1996). "The prevalence of the fragile X syndrome." Am J Med Genet 64: 196-197.

Usdin, K. (1998). "NGG-triplet repeats form similar intrastrand structures: implications for the triplet expansion diseases." Nucleic Acids Res 26: 4078-85.

Verheij, C., C. E. Bakker, et al. (1993). "Characterization and localization of the FMR-1 gene product associated with fragile X syndrome." Nature 363(6431): 722-724.

Verkerk, A. J., M. Pieretti, et al. (1991). "Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome." Cell 65(5): 905-914.

Weisman-Shomer, P., E. Cohen, et al. (2000). "Interruption of the fragile X syndrome expanded sequence d(CGG)(n) by interspersed d(AGG) trinucleotides diminishes the formation and stability of d(CGG)(n) tetrahelical structures." Nucleic Acids Res 28(7): 1535-41.

White, P. J., R. H. Borts, et al. (1999). "Stability of the human fragile X (CGG)(n) triplet repeat array in saccharomyces cerevisiae deficient in aspects of DNA metabolism [In Process Citation]." Mol Cell Biol 19(8): 5675-5684.

Yu, S., M. Pritchard, et al. (1991). "Fragile X genotype characterized by an unstable region of DNA." Science 252(5010): 1179-1181.

Chapter 5.3

Inactivation of *FMR1* Full Mutations in Chorionic Villi

Carola J.M. Bontekoe, Rob Willemsen, Ben A. Oostra

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

submitted

Abstract

Fragile X syndrome is caused by expansion of the CGG repeat in the 5' UTR of the fragile X mental retardation gene (*FMR1*). This expansion leads to methylation of the *FMR1* promoter region thereby blocking FMR1 protein (FMRP) expression. Prenatal diagnosis can be performed on chorionic villi samples (CVS) by Southern blot analysis. Alternatively, for male fetuses an immunohistochemical method on CVS has been introduced and the results at different stages of gestational ages is shown. In the present study we have used both methods for CVS in two cases of full mutation female fetuses. FMRP expression in CVS was found to be absent completely in a number of villi, whereas other villi showed normal FMRP expression. Unlabelled villi can only be present if inactivation of the full mutation has occurred together with X-inactivation of the normal X allele. FMRP positive villi can be explained by an active normal X allele. The presence of both positive and negative villi indicates that X-inactivation in human CVS is a random process. No villi were found with a mixture of both FMRP expressing and non-FMRP expressing cells. This suggests that X-inactivation occurs very early in development, before the villi start to proliferate. As a result the X-inactivation in villi is clonal. Furthermore, it was found that both the normal allele and the full mutation allele were hypomethylated in CVS. Also in control villi and placental tissue from female fetuses the *FMR1* promoter is hypomethylated on the inactive X chromosome.

Introduction

Fragile X syndrome is the most common form of inherited mental retardation affecting one in 4000 males (Turner *et al.* 1996). Mental retardation and developmental delay are the most significant clinical features of fragile X syndrome (Hagerman 1996). This X-linked disorder is caused by the absence of the fragile X mental retardation protein (FMRP) for review see (Kooy *et al.* 2000). The gene defect causing the absence of FMRP is usually an expansion of the trinucleotide $(CGG)_n$ repeat present in the 5' UTR of the fragile X mental retardation gene 1 (*FMR1*) (Oberlé *et al.* 1991; Verkerk *et al.* 1991; Yu *et al.* 1991). The trinucleotide repeat is highly polymorphic and alleles can be divided into three groups. The first group contains the alleles ranging between 5 and 50 repeat units. Repeats of this size remain stable upon transmission. The two other groups, called pre- and full mutations, behave unstable upon transmission to the next generation. Both contractions and expansions are observed, the latter being the most prominent (Rousseau *et al.* 1991).

Alleles between 50 and 200 CGGs are called premutations. Although premutations are expanded and behave unstable, they do not block the expression of FMRP (Devys *et al.* 1993; Verheij *et al.* 1993). Premutations can expand upon both male and female transmission. Expansion to a full mutation, > 200 CGGs, occurs only upon female transmission. Only females with a premutation can transmit a full mutation to their offspring. Full mutations usually coincide with methylation of the promoter region of *FMR1*, thereby blocking the transcription of *FMR1*. This results in the absence of FMRP, which causes the fragile X disease phenotype (Pieretti *et al.* 1991; Verheij *et al.* 1993).

The methylation status can be determined by Southern blot analysis using methylation sensitive restriction enzymes (Oostra *et al.* 1993). The methylation status of the promoter region is always correlated with the expression of FMRP. Why and how this methylation occurs is not known. It has been hypothesized that not the repeat expansion itself, but the methylation is the most important factor causing the disease phenotype (McConkie-Rosell *et al.* 1993). The methylation of the promoter region, including the CGG repeat is responsible for the blocking of the transcription, resulting in the absence of FMRP. The observation that full mutation males who do not show methylation but do express FMRP, supports the idea that an expanded repeat itself is not enough to block the expression completely (McConkie-Rosell *et al.* 1993; Hagerman *et al.* 1994; Smeets *et al.* 1995; De Vries *et al.* 1996; Wohrle *et al.* 1996; Wohrle *et al.* 1998; Taylor *et al.* 1999).

In affected females the situation is more complex, because their cells contain two X chromosomes. Dosage compensation in somatic cells of normal females is necessary for the expression of equal amounts of X-linked genes compared to males and this is achieved by inactivation of one of the two X chromosomes. The process of X-inactivation occurs shortly after blastocyst implantation during embryonic development (Tan *et al.* 1993). Once the X-inactivation is established, it is maintained during further cell proliferation and differentiation of the embryo. In humans the choice, which X chromosome is inactivated, is a random process. For females with one normal and one full mutation allele, this implicates that in 50% of the cells the normal X chromosome will be inactivated and in 50% of the cells the X chromosome carrying the full mutation allele will be inactivated. If the full mutation allele is inactivated by X-inactivation, the normal allele will produce normal amounts of FMRP. In theory, this will be the case in 50% of the cells. In the other 50% of the cells, the normal allele will be subject to X-inactivation. As a consequence, the full mutation allele will be at the active X chromosome. However, since the full mutation is inactivated too, this

allele will not be active, which results in two inactive *FMR1* loci in these cells. One allele is inactivated because of the full mutation (If= inactivation of the *FMR1* full mutation), and the normal allele is inactivated because of X-inactivation (Ix= X-inactivation of the *FMR1* allele). Thus, in these cells there is no FMRP expression. These two mechanisms of inactivation complicate the disease phenotype in females. Depending on the X-inactivation a certain percentage of cells are normally expressing FMRP, whereas other cells lack FMRP expression.

Approximately 65% of the full mutation females show some form of mental impaired. No correlation could be determined between the methylation status in females and the mental status as determined by DNA analysis in blood cells. Diagnostics for the fragile X syndrome is routinely performed by DNA analysis (Oostra *et al.* 1993). DNA isolated from blood is used for Southern blotting to determine the size of the CGG repeat. Together with the size of the repeat, the methylation status of the *Eag*I site present in the promoter region can be determined. The methylation status of the *Eag*I site might be extra informative in a number of cases. First, in females it can provide information about methylation that coincides with the X-inactivation pattern (Mx= methylation due to X-inactivation), and indirectly about the percentage of blood cells that express normal amounts of FMRP. Second, since only the full mutations coincide with the methylation (Mf= methylation due to the full mutation) of the *Eag*I site the double digest technique might help to distinguish small full mutations from large premutations.

In case of prenatal diagnosis, DNA isolated from chorionic villi can be used. In the literature a number of studies have been described in which prenatal diagnosis is performed on chorionic villi samples (CVS) (Sutherland *et al.* 1991; Devys *et al.* 1992; Suzumori *et al.* 1993; Iida *et al.* 1994; Grasso *et al.* 1996). In most described male cases a methylation-sensitive restriction enzyme has been used. In this way both the repeat size and the methylation status of CVS were determined. An alternative prenatal diagnostic method on chorionic villi is based on the absence of FMRP in cytothrophoblasts from full mutation male fetuses (Willemsen *et al.* 1996; Willemsen *et al.* 1997). In the present study we have used this method to study FMRP expression in extra embryonic tissue of female full mutation fetuses. Again, like for blood cells, a distinction has to be made between two independent causes of inactivation. First, normal X-inactivation has to occur as dosage compensation (Ix). Secondy, the full mutation in female fetuses will be inactivated (If). The present study

on CVS from female fetuses with a full mutation should give more insight into the pattern and timing of X inactivation in human extra embryonic tissues.

Material and Methods

Tissue processing and immunohistochemistry

Chorionic villi obtained from pregnancies at risk for the fragile X syndrome were tested for FMRP expression next to the standard DNA analysis. Biopsy material was sent to our laboratory from different Clinical Genetic Centers in The Netherlands. In these centers the chorionic villi biopsies were taken at different times of gestational age, varying from 10.5 weeks - 12.5 weeks. More than 10 male fetuses with a full mutation has been studied. Two cases of full mutation female fetuses are described in more detail below.

Chorionic villi biopsy from both female fetuses showed a full mutation allele by Southern blot analysis. The parents decided to terminate the pregnancy. After informed consent by the parents, fetal tissues, including chorionic villi and somatic tissue, were collected and analyzed. Both DNA analysis and immunohistochemical detection of FMRP was performed on chorionic villi (whole mount and sections) and on brain sections of case 1. Briefly, chorionic villi and fetal tissues were either embedded in Tissue-Tek (Miles Inc., USA) and immediately frozen in liquid nitrogen or chorionic villi were whole mount fixed for 10 min. in 3% paraformaldehyde and subsequently permeabilized in 100% methanol for 20 min. Immunohistochemistry, using monoclonal antibodies against FMRP (Devys *et al.* 1993) followed by an indirect immunoperoxidase technique, was performed on both cryostat sections (8 µm) and whole mount chorionic villi (Willemsen *et al.* 1996). Sections were counterstained with haematoxylin, dehydrated and mounted with Entellan. Finally, slides were examined with a Zeiss Axioskop microscope.

DNA analysis

DNA was extracted from CVS and embryonic tissues according to standard procedures (Miller *et al.* 1988). Approximately 7 µg DNA was digested with *Hin*dIII and the methylation sensitive enzyme *Eag*I. After electrophoresis through a 0.7 % agarose gel Southern blotting was performed (Sambrook *et al.* 1989). Hybridization was performed using ^{32}P-dATP random-primed labeled pP2 as a probe (Rousseau *et al.* 1991; Oostra *et al.* 1993).

For the full term control placenta and the pooled control chorionic villi sample, Southern blotting was performed using the same method. For the pooled Chorionic villi samples different methylation sensitive enzymes (*Ava*I, *Bss*HII, *Eag*I, *Nru*I, *Sac*II, *Sfo*I and *Xho*I) in combination with *Hin*dIII were used to digest the DNA.

Results

Immunohistochemistry

For full mutation male fetuses the FMRP expression pattern was studied in cryostat sections of chorionic villi at different stages of gestational age. Figure 1 illustrates the FMRP expression pattern in chorionic villi from full mutation male fetuses at 10.5, 11.5 and 12.5 weeks of gestational age. At 10.5 weeks the FMRP expression was similar to the expression pattern observed in control male fetuses (data not shown), showing a strong labeling in cytotrophoblast cells. An intermediate pattern is observed at 11.5 weeks, showing a weak labeling in most cytotrophoblast cells and only a strong labeling in some cells. In contrast, at 12.5 weeks of gestational age all the cytotrophoblast cells are totally devoid of FMRP.

Chorionic villi from the female fetuses obtained from the first biopsies were not available for FMRP expression studies, because the quantity was too limited and only allowed DNA analysis. DNA analysis showed a full mutation (data not shown) and the parents decided to terminate the pregnancies. Fetal tissues, including chorionic villi and brain, obtained from the termination at 13.0 weeks of gestational age (case 1) and 13.5 weeks of gestational age (case 2) were used for immunohistochemistry. For case one two strategies were used to study the FMRP expression in these tissues. First, immediately after termination the chorionic villi were fixed as whole mount in order to investigate the overall expression of FMRP in the different villi. This method (only performed for case 1) enables us to compare the expression pattern between one villus and the other villus. The results are illustrated in Figures 2A and 2B. With this method we observed chorionic villi that were absolutely devoid of FMRP (Fig. 2A) and villi that were positively labeled for FMRP (Fig. 2B). This pattern can be explained by assuming that in the FMRP negative villi (Fig 2A) the normal allele is subjected to X-inactivation while in the FMRP positive villi the normal allele is active.

Second, for both female fetuses cryostat sections from sampled (different) chorionic villi and brain tissue were cut and immunoincubated for FMRP expression. Both fetuses showed the

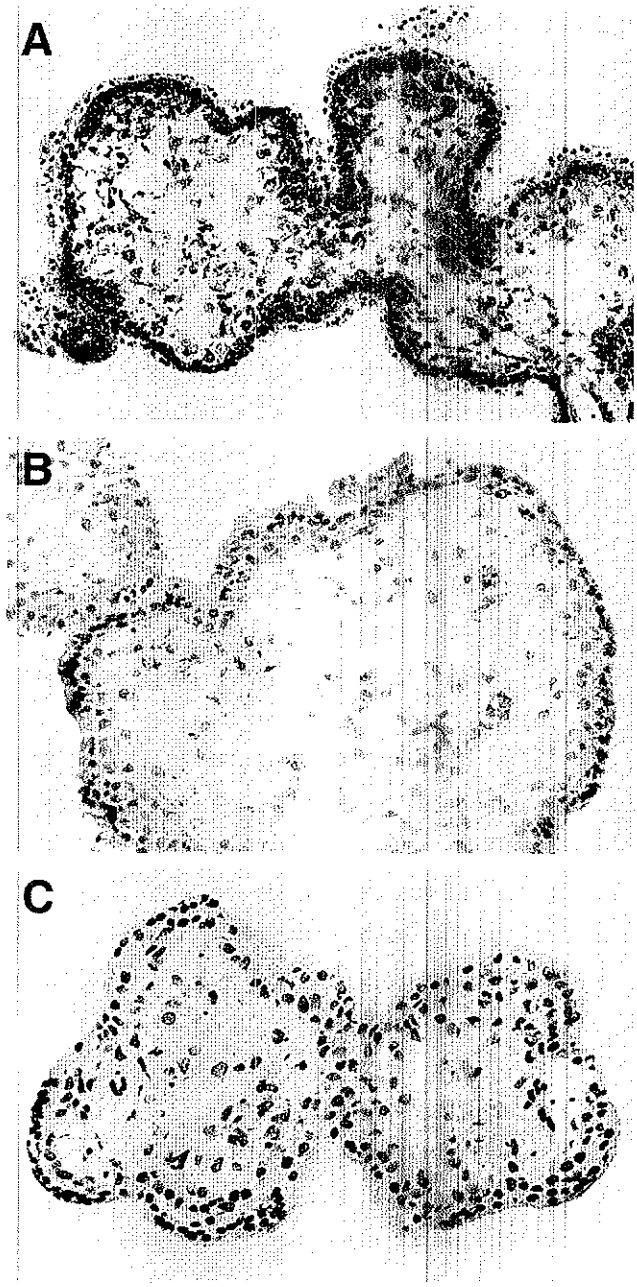

Figure 1. Expression pattern of FMRP in chorionic villi of full mutation males fetuses at 10.5 (A), 11.5 (B), and 12.5 (C) weeks of gestational age. A: FMRP expression is present in cytotrophoblast cells. B: An intermediate pattern of FMRP expression is observed. C: All cytotrophoblast cells are completely devoid of FMRP.

same labeling pattern. For villi, the labeling pattern is shown in Figure 2C illustrating high FMRP expression in cytotrophoblast cells, however, not all villi showed a positive labeling. Apparently, in the plane of a section positive and negative labeled villi are present. Striking is the fact that within one cross section of a villus all the cytotrophoblast cells are either positively labeled or not labeled. In brain tissue a similar labeling pattern is observed (Fig. 2D). Some neurons are totally devoid of labeling, whereas other neurons are strongly labeled. Interestingly, the positive or negative labeled neurons were always located in small groups together suggesting a clonal origin of nearby neurons.

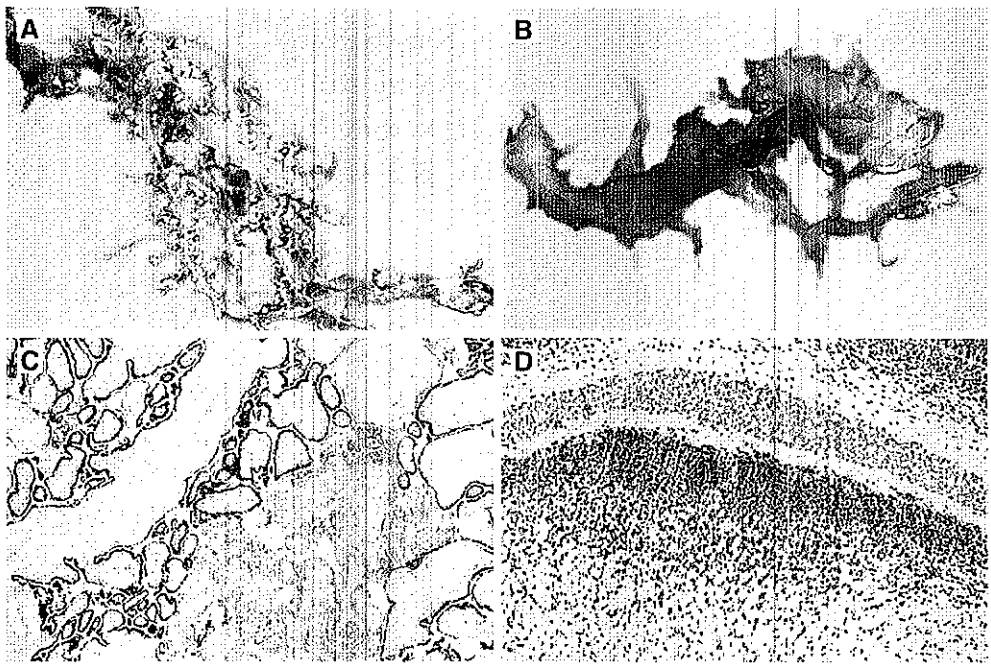

Figure 2. Chorionic villi from full mutation female fetuses were either completely negative (A) or completely positive (B) for FMRP. (C) The labeling pattern for FMRP of cryostat sections of the chorionic villi showed expression of FMRP in cytotrophoblast cells in a number of villi. Other villi were completely devoid for FMRP. (D) In brain tissue of the full mutation female fetus FMRP-positive and FMRP-negative neurons were always located in small groups together.

DNA analysis

Of the two female fetuses DNA analysis after Southern blotting was performed on both embryonic and extra embryonic tissues. The results are shown in figure 3.

In the first female fetus an expanded allele and a normal allele are present (figure 3 lane 4 to 7). The methylation pattern is different in the different tissues. Of this fetus two samples of chorionic villi and two samples of somatic tissue (brain and leg) were digested with *Hind*III and *Eag*I. In the somatic tissues mainly three types of bands are present. The bands of 2.8 kb and 5.2 kb are representing the normal allele. The 2.8 kb band is the unmethylated band, and the 5.2 kb is the methylated band. This methylation reflects the X-inactivation (Mx). Approximately 50% of the normal allele is represented by the 2.8 kb band and 50% is represented by the 5.2 kb band, illustrating random X-inactivation in these tissues. The bands presenting the expanded allele are mainly seen as a smear of more than 5.2 kb. This indicates a methylated allele similar to what is seen in affected males. This methylation occurs due to the expansion of the CGG repeat (Mf) and not due to the X-inactivation.

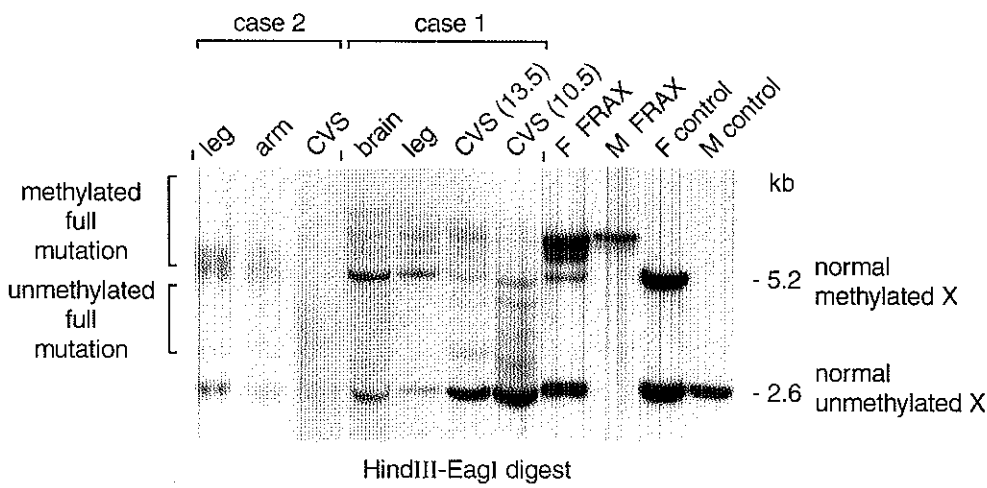

Figure 3. Southern blot analysis on DNA from chorionic villi and somatic tissues of two full mutation female fetuses. DNA was digested with *Eag*I and *Hind*III, pP2 was used as probe. In somatic tissues, the full mutation allele was almost completely methylated. Of the normal allele approximately 50 % was methylated. In chorionic villi, both alleles were hypomethylated.

A different pattern is seen for CVS. The normal allele is presented as bands of 2.8 kb and 5.2 kb, as observed in the somatic tissues. However, in CVS the 2.8 kb fragment is more intense than the 5.2 kb fragment. This indicates that the methylation of the normal allele is seen in less than 50% of the cells. The less abundant 5.2 kb fragment represents the minority of the normal allele, which is methylated (Mx). The full mutation is seen as a smear rather than a clear band, thus quantitation on the basis of the intensity of the bands is more difficult. However, it is clear that most of the expanded repeat allele is unmethylated. These results show that up to 13.5 weeks of gestational age the methylation in CVS is less, compared to the methylation observed in fetal somatic tissues of the same age (see leg DNA). This implicates that the methylation in chorionic villi using the *EagI* site present in the *FMR1* promoter is not completed yet .

In the second case the somatic tissue of the female fetus showed a 2.8 kb and 5.2 kb band representing the normal X chromosome, approximately 50-50%, indicating random X-inactivation (Mx) (figure 4 lane 1 to 3). The full mutation allele was almost completely methylated (Mf). In CVS hypomethylation was observed for both the normal and the full mutation allele. This shows that also in this second full mutation female fetus methylation of the *EagI* site is less abundant than in fetal somatic tissues of the same age.

Since only one methylation sensitive site was tested in these fetuses, we tested a number of other methylation sensitive restriction enzymes in CVS as well. To test the methylation of these sites we pooled 10 non-fragile X female chorionic villi samples. All these samples were of more than 10 weeks of gestational age; X-inactivation is known to have occurred by that time (Tan *et al.* 1993). As a positive control the methylation status of these restriction enzymes was tested on pooled DNA isolated from blood of 10 non-fragile X adult females, of which it is known that X-inactivation (Ix) and the methylation (Mx) should be present. For the pooled female DNA, DNA methylation could be observed for all methylation sensitive enzymes tested (Mx). Methylation was very weak or not detectable in the pooled female CVS. These results indicate that although X-inactivation has occurred very early in development, the methylation is not fully present in non-fragile X female chorionic villi. In addition testing full-term placenta material of two newborn girls (non-fragile X) showed that the methylation of the *EagI* site in the *FMR1* promoter is not present (data not shown).

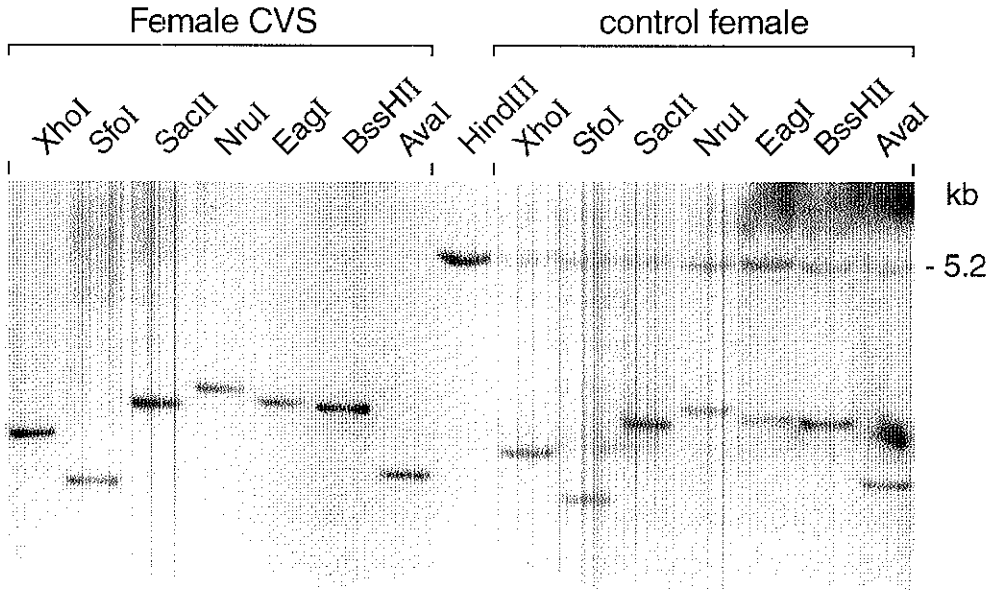

Figure 4. The methylation of DNA pooled from female (not fragile X) CVS of more than 10.0 weeks and DNA pooled from adult female was tested with a number of methylation sensitive enzymes in combination with *Hind*III. pP2 was used as a probe. For all tested enzymes CVS was clearly hypomethylated.

Discussion

Recently, a new method for prenatal diagnosis for the fragile X syndrome has been described. This method is based on the presence of FMRP in the cytotrophoblasts of control male fetuses and the absence of FMRP in the cytotrophoblast of male fetuses with a full mutation. However, the gestational age has been considered important, because of the timing of inactivation of the full mutation allele in chorionic villi. Until 10 weeks of gestational age FMRP is expressed in full mutation male CVS similar to control male CVS. This is thought to occur due to the lack of inactivation and methylation of full mutation alleles at this age. The inactivation of *FMR1* is complete at 12.5 weeks of gestational age resulting in total lack of FMRP (Figure 1). Until now this immunohistochemical test has been applied only to male CVS. Here we report the results of this method for the CVS of two female fetuses with a fully expanded CGG repeat in the *FMR1* gene. The results of the protein test were compared with the DNA analysis as routinely performed for CVS.

Immunohistochemistry of chorionic villi (gestational age of more than 13.0 weeks) of two female fetuses with the full mutation showed that chorionic villi were either completely positive or completely negative labeled for FMRP. Since both positive and negative villi were present this indicates that X-inactivation is a random event in human chorionic villi. In rodents X-inactivation in extra embryonic tissues is different from the embryonic tissues in that random X-inactivation is observed in embryonic tissues whereas in the extra embryonic tissues the paternal X chromosome is inactivated (Takagi *et al.* 1975). In the human situation either X chromosomes of CVS can be subject to X-inactivation. The villi were either completely positive or completely negative for FMRP, there were no villi found with some cells expressing protein and other cells not. This observation indicates that the development of villi is a clonal process. Apparently, X-inactivation has occurred very early in development, before the chorionic villi started to proliferate. Because X-inactivation is maintained upon proliferation and differentiation, villi become either completely positive or completely negative upon proliferation. All cells in a single villus originate from one progenitor cell in which X-inactivation already has occurred.

The available data of the methylation status of chorionic villi of the *FMR1* promoter in literature are conflicting. Some cases have been described showing methylation of the *FMR1* promoter region; other cases report hypomethylation of the CGG repeat region (Sutherland *et al.* 1991; Devys *et al.* 1992; Suzumori *et al.* 1993; Iida *et al.* 1994; Grasso *et al.* 1996). Because the exact age of the chorionic villi is not always documented, it is difficult to compare those studies. For FMRP expression in full mutation male CVS, the gestational age proved to be important for the inactivation of the *FMR1* gene. In full mutation male CVS of less than 12.5 weeks of gestational age, FMRP expression can be observed despite the presence of a full mutation allele. When the gestational age is more than 12.5 weeks the FMRP expression is totally absent in full mutation male CVS. CVS of two full mutation female fetuses showed villi negative for FMRP expression indicating that both alleles, the normal as well as the full mutation allele were inactivated. Inactivation due to both X-inactivation (Ix) of the normal allele and due to inactivation of the full mutation (If) had occurred in these negative chorionic villi.

By performing Southern blot analysis on the female chorionic villi, hypomethylation of the *Eag*I site present in the *FMR1* promoter was observed (Figure 3). For the *FMR1* locus this *Eag*I site is the most tested methylation sensitive restriction site, and absence of FMRP expression was always found to coincide with the methylation of this *Eag*I site in full

mutation male patients. The correlation between the methylation of the *FMR1* promoter and FMRP expression has been found for If as well as Ix. In chorionic villi the situation appears to be different. Hypomethylation of both the inactive X chromosome and of the full mutation allele was observed in CVS. The pattern of Southern blot analysis of CVS showed that the methylation of the X chromosomes in CVS was not complete at a time point where X-inactivation appeared to be present based on the FMRP expression. Assuming a random X-inactivation 50 % of the normal allele and 50% of the full mutation is expected to be methylated. In both case 1 and case 2 hypomethylation was observed for the normal allele. If skewed X-inactivation caused this hypomethylation of the normal allele, it was expected that more than 50 % of the expanded allele would be methylated due to X-inactivation. However, also for the full mutation allele hypomethylation was observed.

If for extra embryonic tissues only DNA methylation is considered one might conclude that X-inactivation is not completed (yet), but the FMRP expression pattern clearly shows that X-inactivation of the *FMR1* locus has occurred already very early in embryonic development. These data support the hypothesis that methylation of the X chromosome is a result of the X-inactivation. The methylation is considered as one of the later events in the X-inactivation process; it is thought to play a role in the maintenance or stabilization of inactivation.

In cells of fragile X males, methylation of the full mutation always results in the absence of FMRP. This correlation between the absence of FMRP and methylation of the *Eag*I site is not confirmed in female CVS. Chorionic villi are clearly hypomethylated although the full mutation *FMR1* allele is inactivated based on the results of immunocytochemistry. A small percentage of full mutation males lack methylation of the *FMR1* promoter region, and they express FMRP (Smeets *et al.* 1995; De Vries *et al.* 1996; Taylor *et al.* 1999). Because of this finding it was thought that methylation of the promoter might be even more important in determining FMRP expression than the repeat length itself. The observation that transcriptional reactivation could be achieved by inducing demethylation with 5-azadeoxycytidine supported this idea (Chiurazzi *et al.* 1998; Chiurazzi *et al.* 1999; Coffee *et al.* 1999). The correlation between FMRP expression and methylation is observed in the examined somatic fetal tissues, only in CVS this correlation could not be confirmed. The reason why the extra embryonic tissues are hypomethylated is unknown. The easiest explanation would be to assume that apparently the tested *Eag*I site is not an important site in determining whether or not there is protein expression. For the methylation due to X-inactivation more methylation sensitive sites in the *FMR1* locus were tested in pooled

female CVS and the results were similar. In CVS clearly hypomethylation was observed for the *FMR1* promoter region when compared to female blood samples. Hypomethylation was also found in full term female placenta.

The only described exception to the rule is in the lung tumor cells of a male fragile X patient (De Graaff *et al.* 1995). This fragile X patient with a full mutation in 99% of his blood cells and a premutation in 1% of his blood cells developed a lung tumor with a premutation. Southern blot analysis of tumor material showed the premutation allele size, indicating that the tumor most likely originated from a cell with this premutation. However, this premutation showed methylation of the *Eag*I and *Bss*HII site. This was the first time methylation of a premutation allele has been observed. In the tumor FMRP expression was observed indicating that in this particular case methylation of the *Eag*I site is not correlated with FMRP expression. We have to take into account that the type of methylation in tumors is different from what is observed in normal cells. Most regions become hypomethylated in tumor cells and some specific regions become methylated. The cause for this specific methylation pattern might be different than the cause involved in X-inactivation and/or inactivation of the full mutation allele. However, this finding shows that the methylation of the *Eag*I site by itself is not an absolute determinant whether or not there is FMRP expression.

For X-inactivation it is known that methylation is just a consequence of X-inactivation. Methylation is thought to play a role in maintenance of the inactive state. So far, it was thought that for fragile X the most important factor in determining FMRP expression was the methylation status, supported by the finding that unmethylated full mutation males express FMRP (Smeets *et al.* 1995) and that methylated full mutation cells when treated with 5-azadeoxycytidine, a demethylating agent, start to express protein (Chiurazzi *et al.* 1998; Chiurazzi *et al.* 1999; Coffee *et al.* 1999). Based on the results from these full mutation female fetuses and the discrepancy found between methylation and FMRP expression in the chorionic villi, we hypothesize that inactivation of the *FMR1* promoter region precede methylation and the methylation is just a consequence of inactivation. The expanded CGG repeat induces the inactivation of the locus and this inactivation somehow results in methylation of the promoter region. Assuming a kind of two-step mechanism X-inactivation and inactivation of the full mutation allele might be more related than so far assumed.

The inactivation mechanism might only be active during a certain period of development. For an unknown reason unmethylated full mutation males escape this inactivation signal,

and thus the subsequent methylation in the somatic tissues. Once the methylation is established it seems to be maintained upon proliferation and differentiation. These results clearly show that the methylation status of CVS is not indicative for the methylation status of the embryo. Therefore prenatal diagnosis should only be based on the repeat length and not on the methylation status of the CVS. The (lack of) FMRP expression in the cytothrophoblasts from male fetuses reflects the inactivation of the full mutation, whereas the (lack of) FMRP expression in cytotrophoblasts from female fetuses reflects both the random inactivation of the X chromosome and the inactivation of the full mutation.

Acknowledgments

The authors like to thank Dr. A. Smits for referring fragile X cases, B. Anar for technical assistance and Ton De Vries-Lentsch for excellent photography.. This study was supported by the Dutch Science Foundation (NWO) through CB-MW (Medical Sciences) and the National Institute of Health (NIH 5 RO1 HD38038).

References

Chiurazzi, P., M. G. Pomponi, et al. (1999). "Synergistic effect of histone hyperacetylation and DNA demethylation in the reactivation of the FMR1 gene." Hum Mol Genet 8(12): 2317-2323.

Chiurazzi, P., M. G. Pomponi, et al. (1998). "In vitro reactivation of the FMR1 gene involved in fragile X syndrome." Hum Mol Genet 7(1): 109-113.

Coffee, B., F. Zhang, et al. (1999). "Acetylated histones are associated with FMR1 in normal but not fragile X-syndrome cells." Nat Genet 22(1): 98-101.

De Graaff, E., R. Willemsen, et al. (1995). "Instability of the CGG repeat and expression of the *FMR1* protein in a male fragile X patient with a lung tumour." Am J Hum Genet 57: 609-618.

De Vries, B. B. A., C. A. M. Jansen, et al. (1996). "Variable FMR1 gene methylation leads to variable phenotype in 3 males from one fragile X family." J Med Genet 33: 1007-1010.

Devys, D., V. Biancalana, et al. (1992). "Analysis of full fragile X mutations in fetal tissues and monozygotic twins indicate that abnormal methylation and somatic heterogeneity are established early in development." Am J Med Genet 43(1-2): 208-216.

Devys, D., Y. Lutz, et al. (1993). "The FMR-1 protein is cytoplasmic, most abundant in neurons and appears normal in carriers of a fragile X premutation." Nature Genet 4(4): 335-340.

Grasso, M., L. Perroni, et al. (1996). "Prenatal diagnosis of 30 fetuses at risk for fragile X syndrome." Am J Med Genet 64(1): 187-190.

Hagerman, R. J. (1996). Physical and behavioral phenotype. Fragile X syndrome: diagnosis, treatment and research. R. J. Hagerman and A. C. Silverman. Baltimore and London, The John Hopkins University Press: 3-87.

Hagerman, R. J., C. E. Hull, et al. (1994). "High functioning fragile X males: Demonstration of an unmethylated fully expanded FMR-1 mutation associated with protein expression." Amer J Med Genet 51(4): 298-308.

Iida, T., Y. Nakahori, et al. (1994). "The CpG island of the FMR-1 gene is methylated differently among embryonic tissues: Implication for prenatal diagnosis." Human Reproduction 9(8): 1471-1473.

Kooy, R. F., R. Willemsen, et al. (2000). "Fragile X syndrome at the turn of the century." Mol Med Today 6(5): 193-198.

McConkie-Rosell, A., A. Lachiewicz, et al. (1993). "Evidence that methylation of the FMR1 locus is responsible for variant phenotypic expression of the fragile X syndrome." Am J Hum Genet 53: 800-809.

Miller, S. A., D. D. Dykes, et al. (1988). "A simple salting out procedure for extracting DNA from human nucleated cells." Nucleic Acids Res 16: 1214.

Oberlé, I., F. Rousseau, et al. (1991). "Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome." Science 252(5010): 1097-1102.

Oostra, B. A., P. B. Jacky, et al. (1993). "Guidelines for the diagnosis of fragile X syndrome." J Med Genet 30(5): 410-413.

Pieretti, M., F. P. Zhang, et al. (1991). "Absence of expression of the FMR-1 gene in fragile X syndrome." Cell 66(4): 817-822.

Rousseau, F., D. Heitz, et al. (1991). "Direct diagnosis by DNA analysis of the fragile X syndrome of mental retardation." N Engl J Med 325(24): 1673-1681.

Sambrook, J., E. F. Fritsch, et al. (1989). Molecular cloning: A laboratory manual. Cold Spring Harbor, Cold Spring Harbor Laboratory press.

Smeets, H., A. Smits, et al. (1995). "Normal phenotype in two brothers with a full FMR1 mutation." Hum Mol Genet 4(11): 2103-2108.

Sutherland, G. R., A. Gedeon, et al. (1991). "Prenatal diagnosis of fragile X syndrome by direct detection of the unstable DNA sequence." N Engl J Med 325(24): 1720-1722.

Suzumori, K., M. Yamauchi, et al. (1993). "Prenatal diagnosis of a hypermethylated full fragile X mutation in chorionic villi of a male fetus." J Med Genet 30(9): 785-7.

Takagi, N. and M. Sasaki (1975). "Preferential expression the paternally derived X chromosome in the extraembryonal membranes of the mouse." Nature 256: 640-643.

Tan, S. S., E. A. Williams, et al. (1993). "X-chromosome inactivation occurs at different times in different tissues of the post-implantation mouse embryo." Nature Genetics 2: 170-174.

Taylor, A. K., F. Tassone, et al. (1999). "Tissue heterogeneity of the FMR1 mutation in a high-functioning male with fragile X syndrome." Am J Med Genet 84(3): 233-9.

Turner, G., T. Webb, et al. (1996). "The prevalence of the fragile X syndrome." Am J Med Genet 64: 196-197.

Verheij, C., C. E. Bakker, et al. (1993). "Characterization and localization of the FMR-1 gene product associated with fragile X syndrome." Nature 363(6431): 722-724.

Verkerk, A. J., M. Pieretti, et al. (1991). "Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome." Cell 65(5): 905-914.

Willemsen, R., J. C. Oosterwijk, et al. (1996). "Prenatal diagnosis of fragile X syndrome." The Lancet 348: 967-968.

Willemsen, R., A. Smits, et al. (1997). "Rapid antibody test for diagnosing fragile X syndrome: A validation of the technique." Hum Genet 99(3): 308-311.

Wohrle, D., U. Salat, et al. (1998). "Unusual mutations in high functioning fragile X males: apparent instability of expanded unmethylated CGG repeats." J Med Genet 35(2): 103-11.

Wohrle, D., S. Schwemmle, et al. (1996). "DNA methylation and triplet repeat stability: New proposals addressing actual questions on the CGG repeat of fragile X syndrome - Letter to the editor." Am J Med Genet 64(2): 266-267.

Yu, S., M. Pritchard, et al. (1991). "Fragile X genotype characterized by an unstable region of DNA." Science 252(5010): 1179-1181.

Chapter 5.4

Fxr2 Knockout Mouse: a Model for Mental Retardation

Carola J.M. Bontekoe*, Kellie L. McIlwain*[1], Ingeborg M Nieuwenhuizen, Lisa A. Yuva-Paylor[1], Anna Nellis[1], Rob Willemsen, Laura Kirkpatrick[1], Zhe Fang[1], Cathy E. Bakker, Robin McAninch[1] , Ngan Cheng, Michelle Merriweather[1], Andre T. Hoogeveen, David Nelson[1], Ben A. Oostra, Richard Paylor[1],

*Contributed equally

Department of Clinical Genetics, Erasmus University Rotterdam, The Netherlands

[1]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, USA.

submitted

Abstract

Fragile X syndrome is a common form of mental retardation caused by the absence of the FMR1 protein, FMRP. *Fmr1* knockout mice exhibit a phenotype with some similarities to humans, such as macroorchidism and behavioral abnormalities. Two homologues of FMRP have been identified, FXR1P and FXR2P. These proteins show high sequence homology, including all functional domains identified in FMRP, such as RNA binding domains. They overlap in tissue distribution with FMRP and interactions between the three FXR proteins have been found. FXR2P shows high expression in brain and testis, like FMRP. To study the function of FXR2P, we generated an *Fxr2* knockout mouse model. No pathological differences between knockout and wild type mice were found in brain or testis. Given the behavioral phenotype in fragile X patients and the phenotype previously reported for the *Fmr1* knockout mouse, we performed a thorough evaluation of the *Fxr2* knockout phenotype using different behavioral assays. Interestingly, behavioral tests showed that the phenotype found in *Fxr2* knockout mice have some similarities and differences compared to the behavioral responses of *Fmr1* knockout mice.

Introduction

Fragile X syndrome is the most common form of inherited mental retardation affecting one in 4000 males (Turner *et al.* 1996; De Vries *et al.* 1997). The main characteristics of the syndrome are mental retardation and macroorchidism in males (Hagerman 1996). This disorder results from an expanded CGG trinucleotide repeat in the 5' untranslated region (UTR) of the FMR1 gene (Fu *et al.* 1991; Oberlé *et al.* 1991; Verkerk *et al.* 1991).This X-linked disorder is caused by the absence of the fragile X mental retardation 1 protein (FMRP). Since the cloning of *FMR1* in 1991 much effort has been put in unraveling the function of FMRP (Verkerk *et al.* 1991; Ashley *et al.* 1993; Feng *et al.* 1997; Kanamori *et al.* 1998). Although the precise function of FMRP has not been elucidated several characteristics of the protein have been described. Two homologues of *FMR1*, *FXR1* and *FXR2* have been identified from numerous species (Coy *et al.* 1995; Siomi *et al.* 1995; Zhang *et al.* 1995). Together the three proteins form a small family of fragile X related (FXR) proteins. Since these proteins show a high sequence homology and overlap in tissue distribution analogous functions are suggested (Tamanini *et al.* 1996; Bakker *et al.* 2000). It

has been suggested that that proteins might partly complement each. Therefore studies concerning FMRP are extended to study the three proteins and their possible interactions, to unravel the function of these proteins and to understand how the absence of FMRP is causing the fragile X phenotype.

In the sequence of *FMR1* a number of domains could be identified. FMRP contains conserved sequence motifs, two KH domains and an RGG box (Siomi *et al.* 1993b) that are present in many RNA binding proteins. FMRP has also been found to be associated with ribosomes (Eberhart *et al.* 1996; Khandjian *et al.* 1996) and to be present in neuronal dendrites (Tamanini *et al.* 1996; Feng *et al.* 1997). The RNA binding domains appear to be functional, since *in vitro* binding of FMRP to homopolymeric RNA was shown and FMRP showed a selectivity for a fraction of mRNAs expressed in brain including its own mRNA (Ashley *et al.* 1993; Siomi *et al.* 1993a; Brown *et al.* 1998). FMRP is mainly cytoplasmatic, but the presence of a nuclear export signal (NES) as well as a nuclear localization signal (NLS) suggest that FMRP is capable of shuttling between the nucleus and the cytoplasm (Eberhart *et al.* 1996; Fridell *et al.* 1996; Bardoni *et al.* 1997).

The three FXR proteins are very homologous by amino acid sequence, especially throughout the amino-terminal and central regions (for FXR1P and FXR2P 86% and 70% identity, respectively). Consistent with this, the functional domains characterized in FMRP are also found in FXR1P and FXR2P. Both FXR1P and FXR2P contain an NES and an NLS (Eberhart *et al.* 1996). Recently, a nucleolar-targeting signal has been identified in both FXR1P and FXR2P (Tamanini *et al.* 2000). This signal is present in FXR2P, and in a minority of FXR1P isoforms due to alternative splicing (Kirkpatrick *et al.* 1999). The carboxy-termini of the three proteins are, however, highly divergent, sharing only 6% similarity. In tissue distribution similarities are found as well. Like FMRP, FXR1P and FXR2P are expressed in the organs affected in the fragile X syndrome. In brain FMRP, FXR1P and FXR2P are found in the cytoplasm of neurons. A minority of FXR1P was demonstrated in the nucleolus of some neurons. In testis the expression of the three proteins is different. For adult testis FMRP is expressed in early spermatogonia, together with FXR1P and FXR2P. For FXR1P and FXR2P high expression is also found in the more maturing spermatogenic cells (Tamanini *et al.* 1997; Khandjian *et al.* 1998).

The similarities found between the three proteins at the amino acid level as well as the overlap in tissue distribution led to the suggestion that the FXR proteins have analogous functions. It has been shown that the FXR proteins can interact suggesting that the three

proteins perform similar functions (Zhang *et al.* 1995). The relatively mild phenotype observed in fragile X patients is difficult to reconcile with the fundamental properties of FMRP. If these genes are capable of partially complementing one another's functions, this might explain the relatively mild phenotype observed in fragile X patients, who lack a highly conserved protein with functions in RNA metabolism

Like FMRP, FXR1P and FXR2P are conserved in other species. All three proteins are also found in mice. A knockout mouse for *Fmr1* was generated in order to study the function of Fmrp (Bakker *et al.* 1994). The phenotype of these mice shows similarities to the phenotype observed in fragile X patients. The knockout mice show a learning deficit and macroorchidism is observed. A knockout model for Fxr1 has yet to be fully reported, but unpublished evidence suggests early postnatal lethality from respiratory difficulties (Siomi and Dreyfuss, personal communication). Fxr1p is highly expressed in muscle and the lack of Fxr1p could influence the function of the lungs and the heart, which are composed of predominantly muscle tissue. Here we describe the generation of a knockout mouse for *Fxr2*. Fxr2p is also relatively highly expressed in brain and testis. If the FXR proteins indeed have analogous functions our *Fxr2* knockout mouse model might shed more light on the function of the FXR proteins and how these proteins interact or influence each other. Double knockouts of *Fmr1* and *Fxr2* might provide data about the likelihood that Fxr2p can compensate partly the absence of Fmrp.

Material and Methods

Targeting construct and ES cell electroporation

A mouse cosmid library CCE was screened to identify the mouse homologue of the *FXR2* gene using human*FXR2* cDNA as probe. A number of overlapping cosmids almost completely covering the *Fxr2* gene was identified. A 11 kb *Bam*HI fragment containing intron 3 to intron 13 was used for the targeting construct. This fragment was cloned into pBluescript. The *Neo* cassette was cloned antisense into *Bgl*II sites thereby removing exon 7. The *tk* cassette was cloned in the *Sal*I site of pBluescript. The targeting construct pCB33 is depicted in figure 1. pCB33 is used for ES cell electroporation.

The plasmid was linearized by a *Cla*I digest to enhance homologous recombination. Plasmid DNA was purified and used to transfect ES-cells. Electroporation was performed with 10^7 ES cells in 400 µl PBS using a Progenetor II Gene Pulser (1200 µF and 117V

during 10 ms). Selection of transfectants was performed by selection for the presence of the *neo* cassette by G418 (200 μg/ml) and absence of the *tk* cassette by Fiau (2 μM). After culturing the cells to allow colony forming, colonies were picked and cultured separately. From each clone one batch was frozen and one batch was used for DNA isolation.

DNA analysis

To check for homologous recombination PCR and Southern blot analysis was performed at all isolated ES cell clones. PCR was performed using primers neof (5' CCTGCGTGTACCCCACAGGTCC 3') and fxr2AB (5' CTGTAAGGATTGCTGTCTGGATCC 3'). Cycle conditions were 2' 94°C, 18x(10" 94°C, 30" 60°C, 3' 68°C), 19x(10" 94°C, 30" 60°C, 3'+20" per cycle 68°C), 15' 68°C using the Expand Long Template PCR System (Boehringer Mannheim, buffer 3). For Southern blot analysis ES cell DNA was digested with *Bgl*II. As a probe for Southern blot hybridization the PCR fragment generated with primers fxr2e4 (5' GCGGATGATGAAGGGAGATG 3') and fxr2iR (5' GGACAGAGCTGGCACTGTG 3') was used. Cycle conditions were 5' 94°C, 30x(30" 95°C, 30" 60°C, 1'30" 72°C), 10' 72°C using standard PCR buffers.

Positive ES cells were used for injection into C57/Bl6J blastocysts.

Immunohistochemistry

Fxr2 knockout and wt mice (n=4) were anaesthetized and sacrificed by perfusion fixation with 3% paraformaldehyde for 10 minutes. Several organs were dissected, including brain, testes, liver, spleen, heart, muscle, kidney and lung and post fixation was performed in 3% paraformaldehyde for overnight. Tissues were embedded in paraffin according to standard protocols. Paraffin sections (5 μm) were examined for gross abnormalities using haematoxylin/eosin staining.

For immunocytochemistry, paraffin sections from wt and *Fxr2* knockout tissues were deparaffinized followed by microwave treatment in 0.01 M sodium citrate solution. Endogenous peroxidase activity was inhibited by 30 min. incubation in PBS-hydrogen peroxide-sodium azide solution. Sections were incubated with rabbit polyclonal antibodies (1937) against FXR2P to study the presence of cross-reactive material in the *Fxr2* knockout mice (Tamanini *et al.* 1997). In addition, the expression patterns of the two homologues Fmrp and Fxr1p were studied using monoclonal and polyclonal antibodies, respectively (Devys *et al.* 1993; Tamanini *et al.* 1997). Furthermore, to study pathological abnormalities

within the brain in more detail antibodies against specific markers, including neurofilament-H (SMI31), microtubule associated protein-2 (MAP-2), Tau protein and synaptophysin were used. Subsequently, an indirect immunoperoxidase technique was performed using 3,3'-diaminobenzidine-HCl as a substrate. Finally, sections were counterstained with haematoxylin, dehydrated with ethanol and mounted with Entellan.

Behavioral studies

Animals: A total number of 56 F2 (129/Ola: FVB: C57BL6/J) male mice were used for the experiments. The mice were tested to ensure the absence of the FVB rd/rd mutation. Two distinct replicate batches of mice were tested. In Batch A there were 11 wild-type and 14 mutant mice. In Batch B there were 16 wild-type and 15 mutant mice. The mice were housed in a room with a 12:12 hr light: dark schedule with lights on at 0600. Mice had access to food and water ad libitum. Mice were housed and tested in accordance with NIH policies on use of animals in research and all behavioral testing procedures have been approved by the Animal Protocol Review Committee at Baylor College of Medicine.

Behavioral tests:

Male mice were subjected to a standard test battery originally reported by Crawley and Paylor, (Crawley *et al.* 1997) using procedures described in McIlwain *et al* (in press). The order of testing has been designed to progress from the least invasive to the most invasive (see McIlwain *et al*, in press). Mice began testing at 10 weeks of age with a general neurological screen.

Locomotor activity in the open-field. One week after the neurological screen, locomotor activity was evaluated by placing mice into an open-field arena. Each subject was placed in the centre of a clear Plexiglas (40 cm x 40 cm x 30 cm) chamber and allowed to explore for 30 minutes. Room lighting was provided by overhead incandescent light bulbs that provided 800 lux inside the test chamber. In addition, white noise was present in the room such that it was approximately 55dB inside the test chamber. Activity in the open-field was quantified by a computer-operated Digiscan optical animal activity system (RXYZCM (16), Omnitech Electronics) containing 16 photoreceptor beams on each side of the arena, dividing the arena into 256 equally-sized squares. Total distance (locomotor activity) vertical activity (rearing measured by number of photobeam interruptions), and centre distance (i.e. the distance traveled in the centre of the arena) were recorded. The centre distance was also divided by the total distance to obtain a centre distance / total distance ratio. The centre

distance / total distance ratio can be used as an index of "anxiety-related responses" (Peier *et al.* 2000a). Data were collected in two-min intervals over the 30 min test session.

Rotarod test. Motor coordination and balance were tested one week later using an accelerating rotarod (UGO Basile Accelerating Rotarod). The rotarod test was performed by placing a mouse on a rotating drum and measuring the time each animal was able to maintain its balance walking on top of the rod. Some mice hold on to the rotating drum as they begin to fall and ride completely around the rod. For these mice, the latency to the first complete revolution was recorded. The speed of the rotarod accelerated from 4 to 40 rpm over a five-minute period. Mice were given four trials with a 30-60 min intertrial rest interval.

Startle and prepulse inhibition of the startle. One week after rotarod testing, prepulse inhibition of acoustic startle responses was measured using the SR-Lab System (San Diego Instruments, San Diego, CA) as previously described (Paylor *et al.* 1997). A test session began by placing a mouse in the Plexiglas cylinder where it was left undisturbed for 5 min. A test session consisted of seven trial types. One trial type was a 40 ms, 120 dB sound burst used as the startle stimulus. There were five different acoustic prepulse plus acoustic startle stimulus trials. The prepulse sound was presented 100 ms before the startle stimulus. The 20 ms prepulse sounds were 74, 78, 82, 86, or 90 dB. Finally, there were trials where no stimulus was presented to measure baseline movement in the cylinders. Six blocks of the seven trial types were presented in pseudorandom order such that each trial type was presented once within a block of seven trials. The average intertrial interval was 15 sec (ranged from 10 to 20 sec). The startle response was recorded for 65 ms (measuring the response every 1-ms) starting with the onset of the startle stimulus. The background noise level in each chamber was 70 dB. The maximum startle amplitude recorded during the 65-ms sampling window was used as the dependent variable.

The following formula was used to calculate % prepulse inhibition of a startle response: 100 - [(startle response on acoustic prepulse plus startle stimulus trials / startle response alone trials) X 100]. Thus, a high % prepulse inhibition value indicates good prepulse inhibition, i.e. the subject showed a reduced startle response when a prepulse stimulus was presented compared to when the startle stimulus was presented alone. Conversely, a low % prepulse inhibition value indicates poor prepulse inhibition, i.e. the startle response was similar with and without the prepulse.

Habituation of the acoustic startle response. One week later, habituation of the acoustic startle response was measured. One hundred startle stimuli (120 dB, 40ms) were presented

to each mouse. The average interstimulus interval was 15 sec. The maximum response to each stimulus was recorded.

Pavlovian conditioned fear. Two to 3-weeks later performance in a conditioned fear task was measured as described by Paylor *et al.* (Paylor *et al.* 1994) using a Freeze Monitor system (San Diego Instruments). The test chamber (26 cm X 22 cm X 18 cm high) was made of clear Plexiglas and surrounded by a photobeam detection system (12 beams X 10 beams). The bottom of the test chamber was a grid floor used to deliver an electric shock. The test chamber was placed inside a sound attenuated chamber (Med Associates, internal dimensions: 56 cm X 38 cm X 36 cm). Mice were observed through windows in the front of the sound attenuated chamber. A mouse was placed in the test chamber (house lights "ON") and allowed to explore freely for two min. A white noise (80 dB), which served as the conditioned stimulus (CS), was presented for 30 sec followed by a mild (2 sec, 0.5 mA) foot-shock, which served as the unconditioned stimulus (US). Two min later another CS-US pairing was presented. The mouse was removed from the chamber 15-30 sec later and returned to its home cage. Freezing behavior was recorded using the standard interval sampling procedure every 10 sec. Responses (run, jump, and vocalize) to the foot-shock were recorded.

Twenty-four hours later, the mouse was placed back into the test chamber for 5 min and the presence of freezing behavior was recorded every 10 sec (context test). Two hours later, the mouse was tested for its freezing to the auditory CS. For the auditory CS test, an empty mouse cage replaced the test chamber. In addition, a container with 1.0 ml of vanilla extract was placed in the sound attenuated chamber. Finally, the sound attenuated chamber was illuminated with red lights. There were two phases during the auditory CS test. In the first phase (pre-CS), freezing was recorded for 3 min without the auditory CS. In the second phase, the auditory CS was turned on and freezing was recorded for another 3 min. The number of freezing intervals was converted to a % freezing value. For the auditory CS test, the % freezing value obtained during the pre-CS period was subtracted from the % freezing value when the auditory CS was present.

Spatial learning in the Morris water task. Two weeks later, mice were trained in the Morris water task (Morris 1981) to locate a hidden escape platform in a circular pool (1.38 m diameter) of water (Upchurch *et al.* 1988). Each mouse was given 8 trials a day, in blocks of four trials for four consecutive days. The time taken to locate the escape platform (escape latency) and the distance traveled were determined. After trial 32, each animal was given a

probe trial. During the probe trial, the platform was removed and each animal was allowed 60 sec to search the pool. The amount of time that each animal spent in each quadrant was recorded (quadrant search time). The number of times a subject crossed the exact location of the platform during training was determined, and compared with crossings of the equivalent location in each of the other quadrants (platform crossing).

Selective search data in the probe trial were analyzed by individual one-way (quadrants) repeated ANOVAs and Newman-Keels post-hoc comparison tests. Two-way ANOVA (genotype X gender) were used to compare the quadrant search time and platform crossing data for the training quadrant only between mutant and wild type mice.

Hot plate test. Two weeks later, the hot-plate test was used to evaluate the sensitivity to a painful stimulus. Mice were placed on a 55.0 °C (+/- 0.3) hot-plate, and the latency to the first hind-paw response (either a foot shake or a paw lick) was recorded.

Data analyses: Data were analyzed using two-way (genotype X batch) or three-way (genotype X batch X repeated measure such as time) Analysis of Variance (ANOVA). Post hoc comparisons were made using either LSD or simple effects tests. Significance was set at $p < 0.05$.

Results

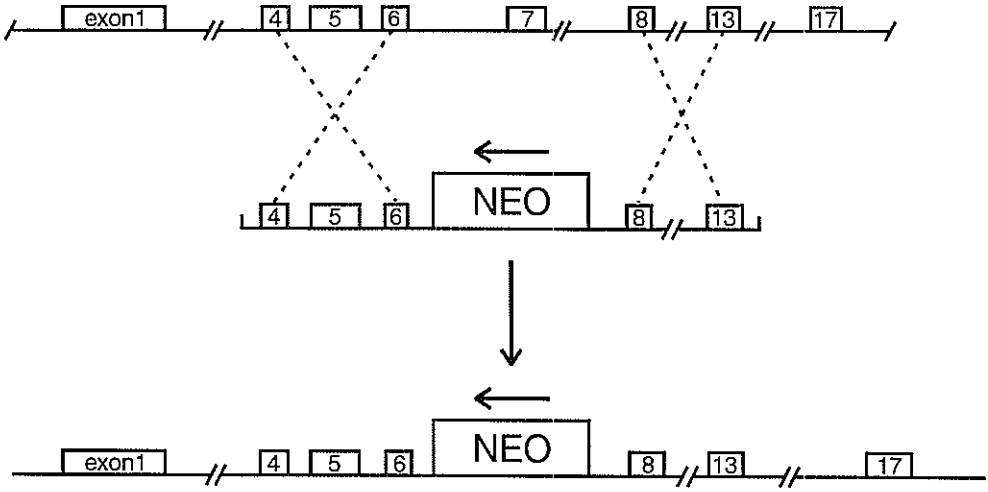

Figure 1. Targeting construct disrupting *Fxr2* by removing exon 7. An antisense *neo* cassette was cloned to allow selection of the transfected ES-cells.

After ES cell electroporation using plasmid pCB33 (Fig. 1) three homologous recombinant clones were identified by PCR and Southern blot analysis. Two clones were used for blastocyst injection. Chimearic animals and subsequent germ line transmissions were obtained for both clones. Littermates were crossed to homozygosity in order to generate *Fxr2* knockout mice.

To test whether the mice homozygous for the mutated *Fxr2* lacked Fxr2p, Western blotting and immunohistochemistry were performed using standard procedures. For Western blotting two antibodies were used, monoclonal A42 and polyclonal 1937, for immunohistochemistry only polyclonal 1937 was used. Western blotting (Fig. 2) as well as immunohistochemistry (Fig. 3B) demonstrated the absence of Fxr2p in the knockout mice.

Mice homozygous for the null allele developed normally and no macroscopic differences could be identified. The mice are viable and fertile. As *Fmr1* knockout mice showed enlarged testicles the testis weight of *Fxr2* mice was determined. The *Fxr2* knockout mice did not show enlarged testicles when compared with normal littermates (data not shown).

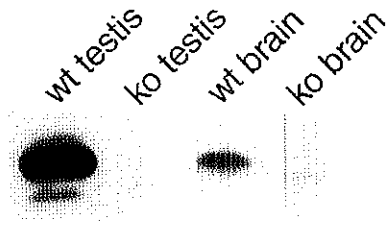

Figure 2. Western blotting showed that the *Fxr2* knockout mice did not express Fxr2p.

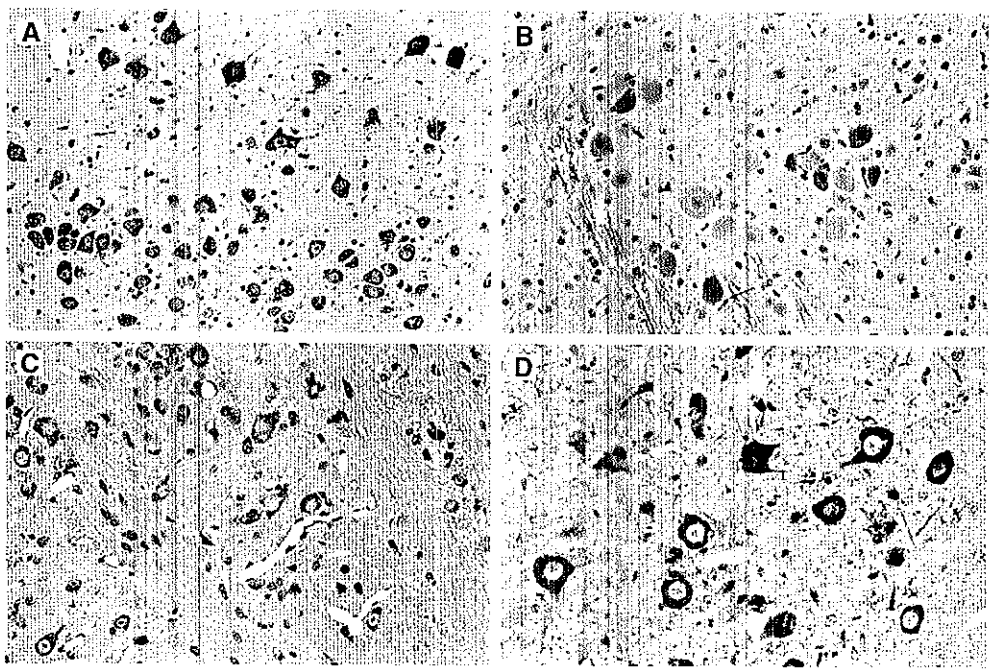

Figure 3. Fxr2p expression in brainstem of wt (A) and *Fxr2* knockout mice (B). The expression of Fmrp (C) and Fxr1p (D) in *Fxr2* knockout mice was not changed compared to wt (data not shown).

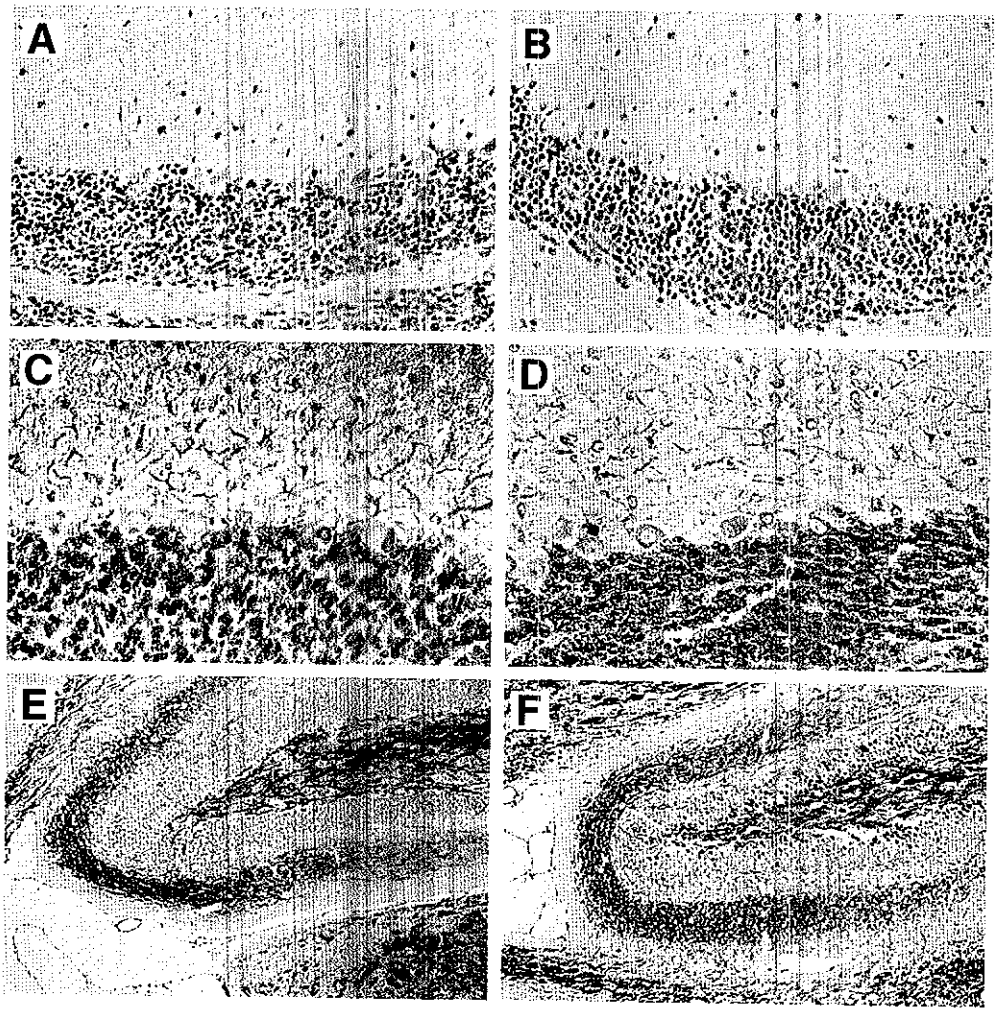

Figure 4. Immunohistochemical analysis of brain sections using the markers synaptophysin (A-B), MAP-2 (C-D) and SMI 31 (E-F) did not show differences between wt (A,C and E) and *Fxr2* knockout mice (B,D and F).

Immunohistochemistry

Light microscopic examination, using routine staining techniques showed no gross abnormalities between wt and *Fxr2* knockout mice (data not shown). Immunocytochemical analysis of sagittal brain sections (wt and *Fxr2* knockout mice), using antibodies against specific brain markers, including SMI31, MAP-2, TAU protein and synaptophysin revealed no pathological abnormalities in the brain compared to wt brain sections (Fig. 4). Fxr2p was absent from brain tissue (neurons) in *Fxr2* knockout mice, using monospecific antibodies against FXR2P (Fig 3), while, sections of brain tissue from wt mice showed the presence of Fxr2p in the cytoplasm of neurons. Interestingly, the expression patterns of Fmrp and Fxr1p in neurons of *Fxr2* knockout mice could not be distinguished from their expression patterns in neurons of wt mice (Fig 3C and 3D.)

Behavioral tests

Fxr2 knockout mice and wt littermates were subjected to a battery of behavioral assays, that included tests for locomotor activity, anxiety related responses, motor coordination and skill learning, sensorimotor gating, sensory adaptation, conditioned fear, spatial learning, and analgesic-related responses. These tests were performed to characterize a range of domains of CNS function, and to compare to the recent behavioral results (Peier *et al.* 2000a) from the *Fmr1* knockout mice.

To insure that there were no severe neurological abnormalities that would potentially interfere with subsequent testing, mice in Batch A were evaluated on a neurological screen (Crawley *et al.* 1997). There were no differences in the baseline responses between *Fxr2* knockout and wt mice in this simple neurological screen assessment suggesting that they had no gross, overt behavioral anomalies.

For each of the tests discussed below the overall main effect of Genotype are presented. However, only the Batch effects that were statistically significant are reported below. Finally, there were no significant Genotype X Batch interactions (p's.0.3), therefore the interaction terms are not presented.

Locomotor activity in the open-field.

The open-field test is used to assess locomotor activity and anxiety-related responses. *Fxr2* knockout mice were significantly more active in the open-field compared with their wild-type littermates. *Fxr2* knockout mice traveled a greater distance [$F(1,52)=22.67$, $p=0.00002$] (Fig. 5A), spent more time moving [$F(1,52)=22.18$, $p=0.00002$] (Fig. 5B), and traveled faster

[F(1,52)=7.6378, p=0.007] (Fig. 5C). The *Fxr2* mutant mice did not rear any more often than wild-type mice [F(1,52) =0.816, p=0.37]. Together these results show that the *Fxr2* knockout show characteristics of hyperactivity compared to wild type littermates.

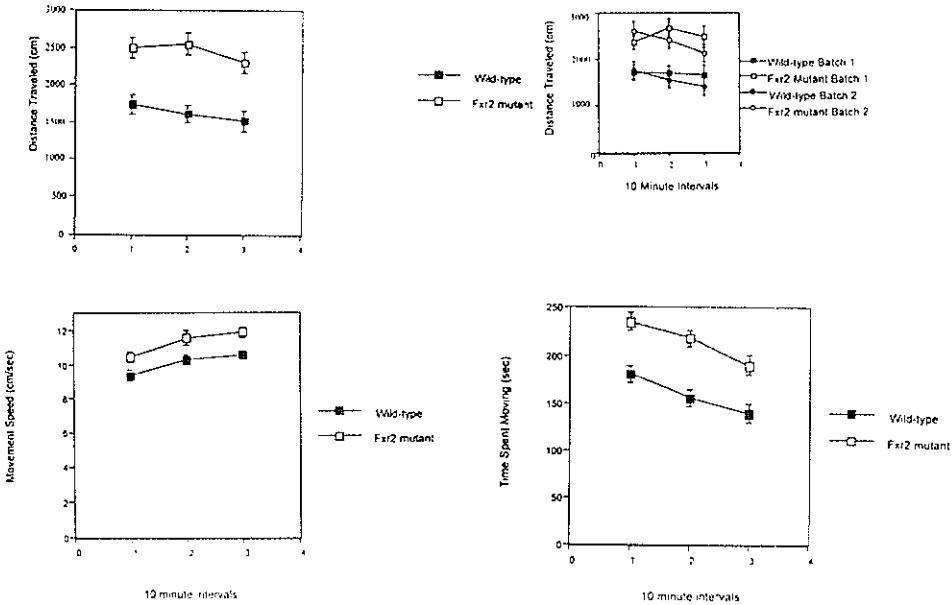

Figure 5.Open-field activity test. The total distance (in centimeters), time spend moving and movement speed for *Fxr2* knockout and wt genotypes. The means (±SEM) are presented.

Using the open-field test one can also assess anxiety-related response using a center : total distance ratio score. Mice prefer to explore the perimeter of the open-field and tend to avoid the center of the arena. The avoidance of the center of the open-field, therefore, is believed to reflect an animal's level of anxiety-related responses. There were no differences in the levels of anxiety-related responses between *Fxr2* knockout and wild-type mice as measured using this center : total distance ratio from the open-field test [F(1,52) = 0.00003, p = 0.996]. Consistent with this finding is the fact that there were also no differences between *Fxr2* knockout and wild-type mice on the Light-Dark test, which is another assay for anxiety-related responses (data not shown).

Rotarod test.

The rotarod test is used to study motor coordination and skill learning. Figure 6 shows that the *Fxr2* knockout mice performed significantly worse than their wild type littermates [$F(1,53) = 7.725$, $p = 0.007$]. An important aspect of these data is the fact that the performance of both the *Fxr2* mutants and wild-types improved significantly during training [$F(7, 371) = 57.629$, $p < 0.00001$], and that there was no genotype X trial interaction [$F(1,53) = 1.634$, $p = .206$]. Thus the rate for the improved performance was similar for both *Fxr2* mutant and wild-type mice. Therefore, this pattern of data suggests that the overall impaired rotarod performance in *Fxr2* mutants may be related to impaired motor coordination, and not to skill learning.

Acoustic startle and prepulse inhibition of the acoustic startle response.

Prepulse inhibition is used to assay sensorimotor gating by quantifying the normal suppression of a startle response by a preceding weak, non-startling prestimulus. The maximum startle amplitude and prepulse inhibition data are presented in Figure 7. In all of our studies we eliminate mice that do not show a startle response since it is not possible to evaluate prepulse inhibition in an animal that does not startle. In this study three mice did not show the minimum startle response (two wild-type and one mutant) and their data were not included in the overall analyses. There was no significant difference in the acoustic startle response between *Fxr2* knockout and wild-type mice [$F(1,49) = 0.365$, $p = 0.0548$]. Overall, the level of prepulse inhibition increased as the sound level of the prepulse increased for both wild-type and *Fxr2* deficient mice [$F(4, 196) = 217.19$, $p < 0.00001$]. However, the overall levels of prepulse inhibition were significantly lower in *Fxr2* knockout mice compared with their wild-type littermates [$F(1,49) = 4.595$, $p = 0.037$]. These finding suggest that the basic startle response is normal in *Fxr2* mutants, but sensorimotor gating as assessed using the prepulse inhibition paradigm is impaired.

Acoustic Startle Habituation. Consistent with the startle response only trials during the prepulse inhibition test, there was no significant difference in the overall startle response between the *Fxr2* knockout and wild-type mice during the startle habituation test [$F(1,49) = 0.103$, $p = 0.748$]. Importantly, both genotypes did show a significant decrease (i.e. habituated) to the startle response across the 100 startle presentations [$F(9,441) = 14.168$, $p = < 0.00001$]. There was a significant batch effect [$F(1,49) = 8.107$, $p = 0.006$], which resulted from the overall levels of startle in the first batch being lower than the responses of the second batch.

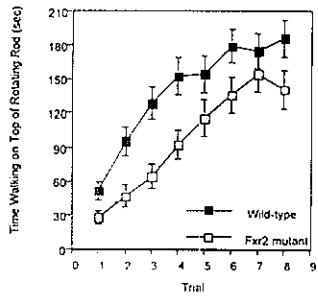

Figure 6. Motor coordination and skill learning was tested in the rotorod assay. Time spent on the rotorod is presented for *Fxr2* knockout mice and wt animals. The means (±SEM) are presented

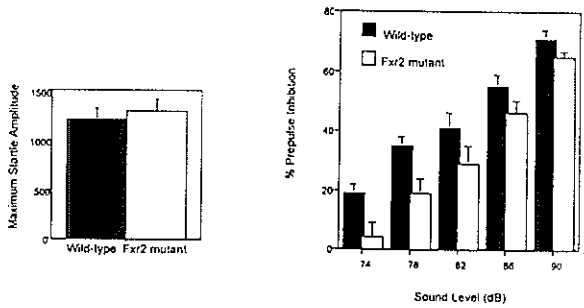

Figure 7. Sensorimotor adaptation was measured in the habituation of the acoustic startle response and sensorigating. The average (±SEM) startle responses to repeated presentations of a 120 dB startle stimulus are given for *Fxr2* knockout anf wt genotypes.

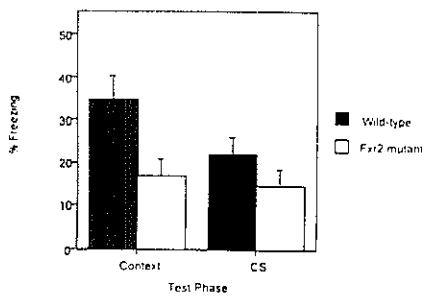

Figure 8.Performance in the conditioned fear test. Percentage freezing assessed during the context and cued tests (CS) are presented.

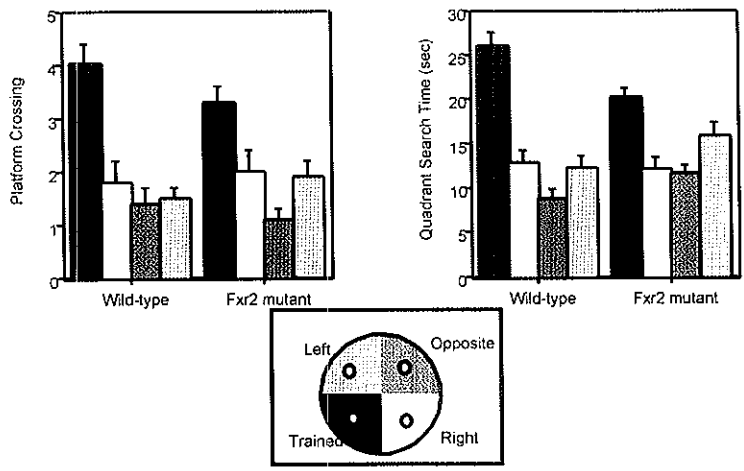

Figure 9. Morris water maze task: Number of platform crossing and Quadrant Search Times for *Fxr2* and wt genotypes are given.

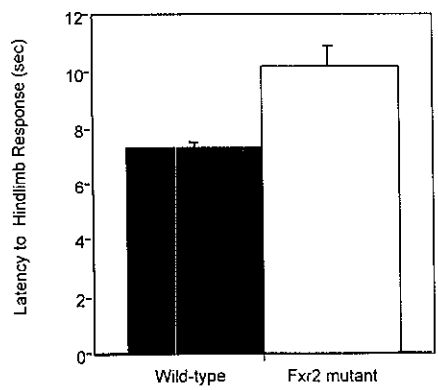

Figure 10. Hotplate test for analgesia- related response. Sensitivity to painful stimuli is shown for *Fxr2* knockout and wt genotypes

Conditioned Fear.

The conditioned fear test assess a fear-based response using a Pavlovian learning and memory setting. Levels of freezing for the context and auditory cued conditioned fear tests are shown in Figure 8. Two mice (one of each genotype) had to be excluded since they did not respond to the footshock during training. During the context test, *Fxr2* knockout mice displayed significantly fewer freezing responses than their wild-type littermates [($F(1,50) = 6.078$, $p = 0.017$]. In contrast, the *Fxr2* knockout mice showed similar levels of freezing during the CS test [$F(1,50) = 0.329$, $p = 0.568$]. These findings suggest that *Fxr2* knockout mice have a selective impairment in fear conditioning that is associated with the context or environment where the shock occurred, but not to a single cue that is associated with the footshock.

Morris water maze.

The hidden-platform version of the Morris water task is widely accepted for examining spatial learning and memory in mice. During Morris water-task training some mice will float or jump off the platform. Since these responses may be incompatible with the behaviors needed to locate and learn the platforms position, we routinely eliminate mice from the experiment if they float or jump from the platform. Therefore, we had to eliminate three mice (one knockout and two wild-type) from this experiment. In addition, one mutant mouse died of undetermined causes prior to the Morris testing.

The swim distance to the escape platform was significantly longer [$F(1,48) = 9.244$, $p = 0.003$] in the *Fxr2* knockout mice compared to the wild-type controls (data not shown). Similarly, the time to locate the platform (data not shown) was also significantly longer in the knockout mice compared to the wild-type mice. [$F1,48] = 16.33$, $p = 0.0002$]. However, there was no overall difference [$F(1,48) = 0.815$, $p = 0.37$] in swim speed (data not shown) between *Fxr2* knockout and wild-type mice.

During the probe trial wild-type mice displayed a spatially biased search pattern. The wild-type mice spent significantly more time in the training quadrant compared to the other quadrants [$F(3,69) = 21.04$, $p < 0.00001$; LSD posthoc Training > other quadrants, $p < 0.0001$] and they crossed the exact place where the platform had been located during training more often than equivalent locations in the other quadrants [$F(3,69) = 9.4$, $p = 0.00003$; LSD posthoc Training > other quadrants, $p < 0.0003$] (Fig.9). Similarly, the *Fxr2* knockout mice spent significantly more time in the training quadrant compared to the other quadrants [$F(3,81) = 6.49$, $p = 0.005$; LSD posthoc Training > Quadrant to Left and

Opposite, $p < 0.0003$, Training > Quadrant to Right, $p = 0.053$], and crossed the training site more often that the comparable site in the other quadrants [$F(3,81) = 6.91$, $p = 0.0003$; LSD posthoc Training > other quadrants, $p < 0.006$]. The wild-type mice did spend significantly more time in the training quadrant compared to the time that the *Fxr2* mutants spent in the training site [$F(1,48) = 9.9067$, $p = 0.004$], but the number of times the two genotypes crossed the training site was equivalent [$F(1,48) = 2.147$, $p = 0.149$].

Taken together it appears that the *Fxr2* mutant mice are not as proficient at locating the platform, but they do use a spatially biased search strategy. Comparing the quadrant search time data from the training quadrant only during the probe trial it appears that the wild-type mice may have a stronger spatial bias for the training site compared to the *Fxr2* knockout mice. However, this difference is not supported by the platform crossing data, which may be a better indicator for spatial search accuracy for the platform position. Thus, *Fxr2* knockout mice do show spatial search behavior in the Morris water task.

Hotplate.

The time to the first hind-limb response is shown in Figure 10. The *Fxr2* knockout mice took significantly longer to shake or lick their hind-paw in response to the heat stimulus [$F(1,51) = 9.86$, $p = 0.0028$]. These findings suggest that the *Fxr2* knockout mice are less sensitive to this type of painful stimuli.

Discussion

It has been hypothesized that FXR1P and FXR2P have a function analogous to FMRP, based on sequence similarity and partly overlapping tissue distribution. Furthermore, it was suggested that the three homologues can interact with each other as homomers and heteromers. The physiological relevance of these interactions is not known. Although the expression levels of FXR1P and FXR2P are not changed in fragile X patients, it is possible that FXR1P and/or FXR2P can complement in part the function of FMRP. Absence of FMRP is sufficient to cause the fragile X phenotype. In order to unravel the function of the FXR proteins and to study their possible interactions it is useful to study the three proteins together. The knockout mouse model for the fragile X syndrome proved to be highly valuable to study the function of FMRP (Bakker *et al.* 1994; Peier *et al.* 2000b). In order to extend these studies to the whole FXR family knockout mouse models will be generated for FXR1P and FXR2P. Here we describe the generation of a knockout mouse for *Fxr2*.

The *Fxr2* knockout mice appear viable and fertile. Macroscopically no gross abnormalities were found. The expression of Fxr2p is relatively high in brain and testis. Therefore special attention was paid to those organs. In brain SMI, MAP-2 and synaptophysin were used as markers, but no differences in the organization of the brain was observed. Likewise immunohistochemistry in testes revealed no differences between *Fxr2* knockout and wild-type mice.

Responses on the behavioral test battery demonstrated several differences between *Fxr2* knockout and wild-type mice. *Fxr2* knockout mice were hyperactive (i.e. traveled a greater distance, spent more time moving, and moved faster) in the open-field test, impaired on the rotarod test, had reduced levels of prepulse inhibition, displayed less contextual conditioned fear, were impaired at locating the hidden platform in the Morris water task, and were less sensitive to a hot pain stimulus. In contrast, anxiety-related responses, acoustic startle, startle habituation, and auditory-cued conditioned fear were similar between the *Fxr2* knockout and wild-type mice. In addition, although the *Fxr2* knockout mice were less proficient at locating the hidden platform during training in the Morris water task, they did appear to use a spatially-biased search strategy similar to the wild-type controls. These behavioral findings support a role for *Fxr2* in domains of CNS function that regulate behavioral responses.

Some, but not all, of the behavioral phenotypes of the *Fxr2* knockout mice resemble the phenotypes of the *Fmr1* knockout mice (summarized in figure 11) evaluated on the same behavioral test battery (Peier *et al.* 2000a). Both *Fxr2* and *Fmr1* knockout mice are hyperactive in the open-field test. Both *Fxr2* and *Fmr1* knockout mice are impaired on the rotarod test. In contrast, *Fmr1* knockout mice, but not *Fxr2* knockout mice, display less anxiety-related responses in the open-field and light-dark test. In addition, the *Fmr1* knockout mice, but not *Fxr2* knockout mice, have a reduced acoustic startle response. From the table one can also see that *Fxr2*, but not *Fmr1*, knockout mice are impaired on the context-dependent conditioned fear test, and show poor performance during training of the Morris water task. In the study of Peier *et al* (Peier *et al.* 2000a), *Fmr1* knockout mice were not tested for levels of sensorimotor gating. Recently, McIlwain and Paylor (unpublished observations) have shown that like *Fxr2* knockout mice, *Fmr1* knockout mice have an altered prepulse inhibition response. However, *Fmr1* knockout mice have enhanced prepulse inhibition, which is opposite of the impaired prepulse inhibition that is seen in the *Fxr2* knockout mice.

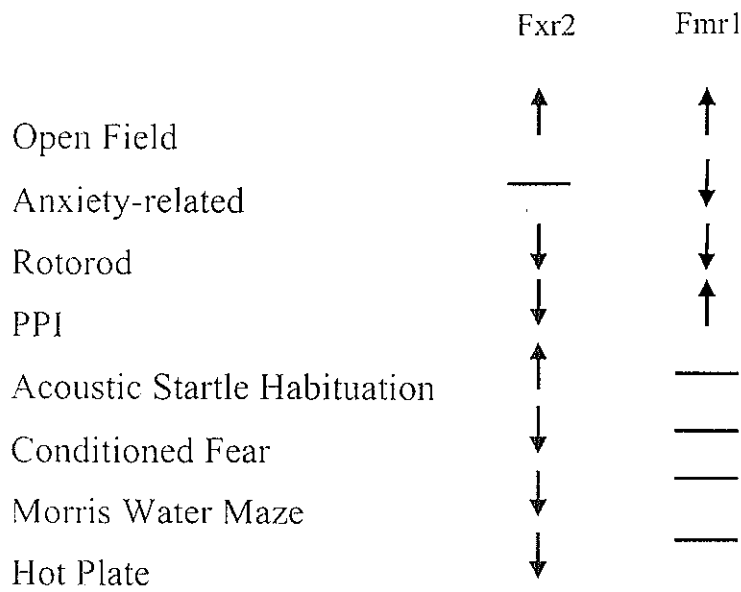

Figure 11. Comparison of the results of the behavioral tests performed for Fmr1 knockout mice and *Fxr2* knockout mice.

It is important to be cautious when comparing the behavioral results from the *Fmr1*- and *Fxr2* knockout mice. Although the behavioral experiments were performed in the same lab using the same equipment, they were not performed simultaneously. There may be environmental differences between the two studies that may have contributed differentially to the outcomes. It is clear that behavioral responses vary among mice of different genetic backgrounds (see review by Crawley *et al* (Crawley *et al.* 1997)). The behavioral studies of Peier *et al* (Peier *et al.* 2000a) were performed using *Fmr1* knockout mice on a C57BL/6 genetic background, while the current studies with the *Fxr2* knockout mice were performed using mice that were on a mixed genetic background (see Methods and Procedures). It is extremely difficult to compare the responses of different mutant mice on different genetic backgrounds. While these methodological differences between the studies with the *Fmr1* knockout and the *Fxr2* knockout mice do warrant caution, they do indicate that *Fxr2* and *Fmr1* contribute to several behavioral responses.

We have recently generated *Fmr1/Fxr2* double knockout mice. We have used a breeding strategy that will generate *Fmr1*, *Fxr2*, and the *Fmr1/Fxr2* double knockouts that are littermates and therefore all on the same type of genetic background. The *Fmr1/Fxr2* double knockout mice will be useful for testing the hypothesis that the gene products from *Fmr1* and *Fxr2* complement/interact with each other. If this hypothesis is accurate, then we

should see some behavioral responses that are exaggerated in the *Fmr1/Fxr2* double knockouts when compared to the individual *Fmr1* and *Fxr2* knockout mice (e.g. open-field activity). In contrast, there may be some responses that are not present in the *Fmr1/Fxr2* double knockouts because the two individual gene knockout phenotypes are in opposing directions (e.g. prepulse inhibition). It is also possible that a deficiency in both *Fmr1* and *Fxr2* will produce behavioral responses that are not observed in either single-gene knockout. Regardless of the outcome of the behavioral studies with the *Fmr1/Fxr2* double knockout mice, we believe that the findings from these double knockout mice will provide important insight into the relationship between the function of Fmrp and the Fxr proteins, and increase our understanding of how these interactions might contribute to fragile X syndrome.

FXR2P has not been associated with any known human disease. Based on the phenotype found for the *Fxr2* knockout mice, it will be interesting to test whether in humans an autosomal recessive form of mental retardation is caused by the absence of FXR2P.

Acknowledgements

We like to thank N. Galjard for providing the CCE cosmid library, and R. van Haperen for blastocyst injections. This study was supported by the Netherlands Organization for Scientific Research (NWO) and by the National Institute of Health (NIH 5 RO1 HD38038 and the BCM Mental Retardation Research Center.

References

Ashley, C., Jr., K. D. Wilkinson, et al. (1993). "FMR1 protein: conserved RNP family domains and selectiveRNA binding." Science 262(5133): 563-568.

Bakker, C. E., Y. de Diego Otero, et al. (2000). "Immunocytochemical and biochemical characterization of FMRP, FXR1P, and FXR2P in the mouse." Exp Cell Res 258(1): 162-70.

Bakker, C. E., C. Verheij, et al. (1994). "Fmr1 knockout mice: A model to study fragile X mental retardation." Cell 78: 23 33.

Bardoni, B., A. Sittler, et al. (1997). "Analysis of domains affecting intracellular localization of the FMRP protein." Neurobiol Dis 4(5): 329-36.

Brown, V., K. Small, et al. (1998). "Purified recombinant Fmrp exhibits selective RNA binding as an intrinsic property of the fragile X mental retardation protein." J Biol Chem 273(25): 15521-15527.

Coy, J. F., Z. Sedlacek, et al. (1995). "Highly conserved 3'UTR and expression pattern of FXR1 points to a divergent gene regulation of FXR1 and FMR1." Hum Mol Genet 4: 2209-2218.

Crawley, J. N. and R. Paylor (1997). "A proposed test battery and constellations of specific behavioral paradigms to investigate the behavioral phenotypes of transgenic and knockout mice." Horm Behav 31(3): 197-211.

De Vries, B. B., A. M. van den Ouweland, et al. (1997). "Screening and diagnosis for the fragile X syndrome among the mentally retarded: an epidemiological and psychological survey. Collaborative Fragile X Study Group." Am J Hum Genet 61(3): 660-667.

Devys, D., Y. Lutz, et al. (1993). "The FMR-1 protein is cytoplasmic, most abundant in neurons and appears normal in carriers of a fragile X premutation." Nature Genet 4(4): 335-340.

Eberhart, D. E., H. E. Malter, et al. (1996). "The fragile X mental retardation protein is a ribosonucleoprotein containing both nuclear localization and nuclear export signals." Hum Mol Genet 5: 1083-1091.

Feng, Y., C. A. Gutekunst, et al. (1997). "Fragile X mental retardation protein: Nucleocytoplasmic shuttling and association with somatodendritic ribosomes." J Neurosci 17(5): 1539-1547.

Fridell, R. A., R. E. Benson, et al. (1996). "A nuclear role for the fragile X mental retardation protein." EMBO J 15(19): 5408-5414.

Fu, Y. H., D. P. Kuhl, et al. (1991). "Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox." Cell 67(6): 1047-1058.

Hagerman, R. J. (1996). Physical and behavioral phenotype. Fragile X syndrome: diagnosis, treatment and research. R. J. Hagerman and A. C. Silverman. Baltimore and London, The John Hopkins University Press: 3-87.

Kanamori, H., R. E. Dodson, et al. (1998). "In vitro genetic analysis of the RNA binding site of vigilin, a multi-KH-domain protein." Mol Cell Biol 18(7): 3991-4003.

Khandjian, E., F. Corbin, et al. (1996). "The fragile X mental retardation protein is associated with ribosomes." Nature Genet 12: 91-93.

Khandjian, E. W., B. Bardoni, et al. (1998). "Novel isoforms of the fragile X related protein FXR1P are expressed during myogenesis." Hum Mol Genet 7(13): 2121-2128.

Kirkpatrick, L. L., K. A. McIlwain, et al. (1999). "Alternative splicing in the murine and human FXR1 genes." Genomics 59(2): 193-202.

Morris, R. G. M. (1981). "Spatial localization does not require the presence of local clues." Learn Motiv 12: 239-260.

Oberlé, I., F. Rousseau, et al. (1991). "Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome." Science 252(5010): 1097-1102.

Paylor, R. and J. N. Crawley (1997). "Inbred strain differences in prepulse inhibition of the mouse startle response." Psychopharmacology (Berl) 132(2): 169-80.

Paylor, R., R. Tracy, et al. (1994). "DBA/2 and C57BL/6 mice differ in contextual fear but not auditory fear conditioning." Behav Neurosci 108(4): 810-7.

Peier, A. M., K. L. McIlwain, et al. (2000a). "(Over)correction of FMR1 deficiency with YAC transgenics: behavioral and physical features." Human Molecular Genetics 9(8): 1145-1159.

Peier, A. M., K. L. McIlwain, et al. (2000b). "(Over)correction of FMR1 deficiency with YAC transgenics: behavioral and physical features." Hum Mol Genet 9(8): 1145-1159.

Siomi, H., M. J. Matunis, et al. (1993a). "The pre-mRNA binding K protein contains a novel evolutionarily conserved motif." Nucleic Acids Res 21(5): 1193-1198.

Siomi, H., M. C. Siomi, et al. (1993b). "The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein." Cell 74(2): 291-298.

Siomi, M. C., H. Siomi, et al. (1995). "FXR1, an autosomal homolog of the fragile X mental retardation gene." EMBO J 14(11): 2401-2408.

Tamanini, F., L. L. Kirkpatrick, et al. (2000). "The fragile X-related proteins FXR1P and FXR2P contain a functional nucleolar-targeting signal equivalent to the HIV-1 regulatory proteins [In Process Citation]." Hum Mol Genet 9(10): 1487-93.

Tamanini, F., N. Meijer, et al. (1996). "FMRP is associated to the ribosomes via RNA." Hum Mol Genet 5: 809-813.

Tamanini, F., R. Willemsen, et al. (1997). "Differential expression of FMR1, FXR1 and FXR2 proteins in human brain and testis." Hum Mol Genet 6: 1315-1322.

Turner, G., T. Webb, et al. (1996). "The prevalence of the fragile X syndrome." Am J Med Genet 64: 196-197.

Upchurch, M. and J. M. Wehner (1988). "Differences between inbred strains of mice in Morris water maze performance." Behav Genet 18(1): 55-68.

Verkerk, A. J., M. Pieretti, et al. (1991). "Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome." Cell 65(5): 905-914.

Zhang, Y., J. P. Oconnor, et al. (1995). "The fragile X mental retardation syndrome protein interacts with novel homologs FXR1 and FXR2." EMBO J 14(21): 5358-5366.

Summary

The fragile X syndrome is the most common form of X-linked inherited mental retardation. The prevalence of the syndrome is 1 in 4000 males. Other characteristics of the syndrome are macroorchidism in males, a long face and everted ears. In 1991, the gene responsible for the fragile X syndrome is cloned and called Fragile X Mental Retardation gene1, *FMR1*. In the 5' UTR of *FMR1* a highly polymorphic CGG trinucleotide repeat is found. Repeat alleles are divided into three groups: normal alleles, premutations and full mutations. Normal alleles range from 5 to 50 CGG triplets and they behave stably upon transmission to the next generation. Premutations range from 50 to 200 CGG triplets. Upon transmission both expansions and contractions are observed. Full mutations (>$(CGG)_{200}$) can coincide with methylation of the promoter region, including the CGG repeat. This methylation blocks the transcription of *FMR1* and results in the absence of the fragile X mental retardation protein (FMRP). The absence of FMRP is causing the fragile X phenotype. This was the first association between expansion of a trinucleotide repeat and human disease, up to now more than a dozen of these so-called trinucleotide repeat diseases have been identified.

Research concerning the fragile X syndrome has focused into two directions. The first direction is related to the mutational mechanism of CGG repeat instability. The other direction involves research to the function of FMRP, and how the absence of FMRP causes mental retardation.

To study the transmission of the CGG repeat through the germ line, a transgenic mouse model was generated. The transgene contained an $(CGG)_{81}$ allele, including two interruptions. Both maternal and paternal transmissions were studied, but no instabilities were observed. Possible explanations for CGG repeat stability were suggested, like the stabilizing effect of the interruptions in the repeat. Flanking sequences, such as cis-acting factors, can also influence the behaviour of the CGG repeat. Biologically, the expanded CGG repeat is located on the X chromosome, for unknown reasons this might be important in repeat instability.

A knock-in mouse model was generated by introducing an expanded $(CGG)_{98}$ of human origin in the endogenous mouse *Fmr1* promoter. Therefore, the changes to the flanking sequences were minimal. In this model, CGG repeat instability was observed upon transmission through the germ line. The frequency of instability was higher in paternal transmission than maternal transmission. Only small instabilities were observed. These results differ from what is observed in the human situation. In humans, upon germ line transmission of an $(CGG)_{98}$ premutation almost exclusively larger expansions occur. This

$(CGG)_{98}$ knock-in mouse model enables us to study the influence of certain factors on repeat instability in a controlled experiment. Furthermore, the timing of the repeat expansion can be determined in these mice. Eventually, it might be that even small expansions result in a repeat size that will coincide with methylation of the mouse *Fmr1* promoter. If this occurs, this knock-in mouse model can be used to study the methylation of the *Fmr1* promoter.

Methylation of the promoter region occurs by repeat expansions above 200 triplets, and blocks the transcription of *FMR1*, resulting in the absence of FMRP. Males with an unmethylated full mutation express FMRP, indicating that the methylation plays an important role in this inactivation process. Chorionic villi, obtained from full mutation female fetuses were studied with immunohistochemistry. Both FMRP-negative and FMRP-positive villi were found. FMRP-negative villi can only be explained by inactivation of both *FMR1* alleles. This means that the normal allele will be subjected to X-inactivation and the full mutation allele will be inactivated as a result of the full mutation. FMRP-positive villi represent those cells in which the normal allele is active, and thus expressing FMRP. Because both FMRP-negative and FMRP-positive villi are present, X-inactivation in chorionic villi is a random event. The fact that the villi were either completely FMRP-negative or completely FMRP-positive shows that each villus originates from one progenitor cell in which X-inactivation was already present. At the FMRP level, this indicates that X-inactivation in chorionic villi is a random and clonal event. In these chorionic villi samples, methylation of the *FMR1* locus was less than expected. Hypomethylation was observed for the normal allele as well as the full mutation allele. Full term (non-fragile X) female placentas were tested for methylation due to X-inactivation, but also in these samples hypomethylation was found. This led to the hypothesis that inactivation, both X-inactivation and inactivation of the full mutation allele, takes place in two steps. First, inactivation occurs (as an yet unknown event). Subsequently, this inactivation leads to methylation. In chorionic villi the inactivation occurs, but is not followed by methylation. The reason for this hypomethylation is not known.

The last study presented in this thesis describes the generation of a knockout mouse model for *Fxr2*. *Fxr2* is a homolog of *Fmr1*. *Fxr2* knockout mice appear viable and fertile. Behavioural studies revealed differences between *Fxr2* knockout mice and their littermates. *Fxr2* knockout mice display hyperactivity. This was also observed for *Fmr1* knockout mice. *Fxr2* knockout mice showed reduced sensitivity to painful stimuli, whereas *Fmr1* knockout mice did not. As the results of the *Fxr2* knockout mice were obtained from a different experiment than the results of the *Fmr1* knockout mice, it is difficult to compare *Fxr2*

knockout mice with *Fmr1* knockout mice. However, the results of the behavioural studies support the hypothesis that *Fxr2* plays a role in domains of the CNS involved in behaviour. The *Fxr2* knockout mouse, together with the *Fmr1* knockout mouse, will be characterized in more detail to compare the data of the *Fxr2* and *Fmr1* knockout mice. *Fxr2/ Fmr1* double knockout mice will provide insight into the interactions of and between the different FXR proteins. In this respect, it will also be useful to generate a *Fxr1* knockout mouse model. At this moment, FXR2P is not known to be associated with any human disease. Based on the results of the *Fxr2* knockout mice it is possible to check an association between FXR2 and an autosomal, recessive form of mental retardation.

Samenvatting

Het fragiele X syndroom is een X-gebonden erfelijke vorm van mentale retardatie. Dit syndroom komt voor bij één op vierduizend mannen. Naast mentale retardatie zijn er een aantal andere kenmerken die in veel patienten worden gevonden waaronder macro-orchidisme, een lang gezicht en wijduitstaande oren. Het gen verantwoordelijk voor het fragiele X syndroom is in 1991 geïdentificeerd, het Fragile X Mental Retardation gen 1(*FMR1*). Dit gen codeert voor FMRP, het fragile X mental retardation proteïne. In het 5' ongetransleerde deel van *FMR1* bevindt zich een trinucleotide CGG repeat. Deze repeat is hoog polymorf, d.w.z. dat er van deze CGG repeat vele allelen voorkomen, verschillende lengten. De allelen worden onderverdeeld in drie groepen: normale allelen, premutaties en volledige mutaties. Normale allelen variëren van 5 tot 50 CGG triplets. Normale allelen vertonen geen instabiliteit bij overerving. Premutaties variëren van 50 tot 200 CGG triplets. Bij overerving naar de volgende generatie kan de repeat instabiel zijn; er kunnen zowel verlengingen als regressies optreden. Volledige mutaties (> $(CGG)_{200}$) kunnen samen gaan met methylering van het *FMR1* promoter gebied, waaronder de CGG repeat. Deze methylering blokkeert de *FMR1* transcriptie en resulteert daardoor in afwezigheid van FMRP, hetgeen het fragiele X syndroom veroorzaakt. Dit was de eerste keer dat een associatie is gevonden tussen trinucleotide repeat verlenging en een humane erfelijke ziekte. Op dit moment zijn er circa veertien van deze "trinucleotide repeat ziekten" bekend.
Na het kloneren van *FMR1* heeft het moleculair biologische onderzoek zich in twee richtingen ontwikkeld. De eerste richting betreft het mutatie mechanisme: Hoe en wanneer verlengt de CGG repeat in het 5' ongetransleerde deel van *FMR1*? De andere richting van het onderzoek richt zich op FMRP: Wat is de functie van FMRP en hoe leidt de afwezigheid van FMRP tot mentale retardatie?
Om repeat instabiliteit te bestuderen zoals deze optreedt in de mens, is een diermodel noodzakelijk. Alleen dan is het mogelijk transmissie van de CGG repeat door de kiembaan te bestuderen. Hiertoe is een transgeen muismodel ontwikkeld, wat wordt beschreven in dit proefschrift. Het transgen bevatte een $(CGG)_{81}$ repeat waarin zich twee interrupties bevonden. In dit muismodel is geen instabiliteit waargenomen. Er zijn meerdere mogelijke oorzaken daarvoor genoemd, zoals het mogelijk stabiliserende effect van de interrupties in de CGG repeat in het transgen. Daarnaast kunnen flankerende sequenties (bijvoorbeeld door cis-acting factoren) een belangrijke rol spelen in CGG repeat instabiliteit. Het feit, dat het transgen willekeurig integreert in het muizengenoom en niet op het X chromosoom waar

de fragiele X CGG repeat zich normaal gesproken bevindt, kan zodoende de CGG repeat instabiliteit beïnvloeden.

De ontwikkeling van een knock-in muismodel stelt ons in staat een verlengde CGG repeat te bestuderen in een omgeving die het dichtst bij de humane stiuatie staat. Daartoe is de $(CGG)_8$ repeat in het muizen-*Fmr1* vervangen door een verlengde humane $(CGG)_{98}$ repeat. Met dit knock-in muismodel is wel instabiliteit waargenomen bij transmissie van de verlengde repeat door de kiembaan. De waargenomen frequentie van instabiliteit was groter bij paternale transmissie dan bij maternale transmissie. Dit resultaat verschilt ten opzichte van de instabiliteit waargenomen in de humane situatie, waar premutaties zowel bij maternale als paternale transmissie instabiliteit laten zien. Daarnaast zijn de verschillen in repeat lengte bij mensen van de ene generatie naar de volgende groter dan waargenomen bij deze muizen. Het genereren van het knock-in CGG trinucleotide repeat muismodel maakt het mogelijk de invloed van bepaalde factoren, die van invloed zijn op repeat instabiliteit, gecontroleerd te bestuderen. Daarnaast kan het tijdstip van instabiliteit bestudeerd worden. Door kleine verlengingen in de CGG repeat is het mogelijk dat deze verlengde repeat uiteindelijk ook in de muis resulteert in methylering van het promoter gebied, waardoor dit muismodel ook informatief kan worden voor de bestudering van de inactivatie van *FMR1*.

Methylering van het promoter gebied ten gevolge van de aanwezigheid van een volledige mutatie blokkeert de transcriptie van *FMR1* en resulteert dus in de afwezigheid van FMRP. Mannen met een ongemethyleerde volledige mutatie brengen wel FMRP tot expressie. Blijkbaar speelt de methylering een cruciale rol in de inactivatie van *FMR1*. Aan de hand van chorion villi afkomstig van vrouwelijke foetussen met een volledige mutatie konden we de interactie tussen methylering en FMRP expressie bestuderen. Met immunohistochemie zijn zowel FMRP-negatieve als FMRP-positieve villi aangetoond. Villi zijn alleen negatief voor FMRP als beide allelen geïnactiveerd zijn; dat houdt in dat het volledige mutatie allel geïnactiveerd zal zijn door de aanwezigheid van de volledige mutatie, en het normale allel door X-inactivatie. De FMRP-positieve villi zijn die cellen waarin het volledige mutatie allel ligt op het inactieve X chromosoom; het normale allel op het actieve X chromosoom brengt FMRP tot expressie. De aanwezigheid van zowel FMRP-negatieve als FMRP-positieve villi is een bewijs voor willekeurige X-inactivatie in chorion villi. De bestudeerde villi waren of helemaal negatief voor FMRP, of helemaal positief. Dit bewijst dat iedere villus ontstaat uit één cel waarin de X-inactivatie al bepaald was. Op eiwit niveau is hiermee aangetoond dat X-inactivatie in chorion villi een willekeurig en klonaal proces is. De methylering in de chorion villi bleek niet volledig. Zowel voor het normale allel als voor het volledige mutatie

allel was er minder methylering dan verwacht op basis van de FMRP expressie. Om te controleren of deze observatie niet afhankelijk was van de leeftijd van de chorion villi, is ook de methylering bestudeerd in (niet fragiele X) placenta's. Ook in placenta werd minder methylering waargenomen dan verwacht op grond van X-inactivatie. Hieruit volgt de hypothese dat de inactivatie zowel X-inactivatie als inactivatie van een volledige mutatie allel plaatsvindt in twee stappen. Er treedt eerst, op nog onbekende wijze, inactivatie op; deze inactivatie wordt vervolgens gevolgd door methylering. De reden waarom deze methylering in chorion villi niet optreedt en wel altijd wordt gevonden in somatische weefsels is niet bekend.

De laatste studie in dit proefschrift beschrijft de ontwikkeling van een knockout muismodel voor *Fxr2*. *Fxr2* is een homoloog van *Fmr1*, met gedeeltelijk hetzelfde expressie patroon. *Fxr2* knockout muizen ontwikkelen zich goed en zijn vruchtbaar. Gedragsproeven laten verschillen zien tussen *Fxr2* knockout en wildtype muizen. De *Fxr2* knockout muizen vertonen hyperactiviteit, vergelijkbaar met de resultaten van de *Fmr1* knockout muizen. Niet alle observaties waren gelijk tussen *Fxr2* en *Fmr1* knockout muizen. In *Fxr2* knockout muizen is een verhoogde pijndrempel waargenomen, in *Fmr1* knockout muizen is de pijndrempel onveranderd. Aangezien de proeven van de *Fxr2* knockout muizen en *Fmr1* knockout muizen in verschillende experimenten zijn uitgevoerd, moeten de gevonden verschillen voorzichtig geïnterpreteerd worden. De resultaten van de gedragsproeven met de *Fxr2* knockout muizen ondersteunen de hypothese dat *Fxr2* een rol speelt in domeinen van het centrale zenuwstelsel betrokken bij gedragsbepaling. Het *Fxr2* knockout muismodel, in combinatie met het *Fmr1* knockout muis model zal in verdere studies beter gekarakteriseerd worden, zodat vergelijkingen tussen de verschillende modellen gemaakt kunnen worden. Daarnaast zullen dubbel knockouts inzicht geven in de interactie van en tussen de verschillende FXR eiwitten. In de toekomst zal mogelijk ook een *Fxr1* knockout muismodel ontwikkeld worden om een zo compleet mogelijk beeld te krijgen van de functies van deze eiwitten. Op dit moment is betrokkenheid van *FXR2* in een humane ziekte niet bekend. Aan de hand van de resultaten van deze gedragsproeven met de *Fxr2* knockout muizen is het mogelijk te onderzoeken of een vorm van autosomaal recessieve mentale retardatie in de mens geassocieerd is met de afwezigheid van FXR2P.

Abbreviations

Asn	asparagine
CVS	chorionic villi sample
DM	myotonic dystrophy
DRPLA	dentatorubro-pallidoluysian atrophy
DNA	deoxyribonucleic acid
ES	embryonic stemcell
FEN1	flap endonuclease 1
FMR1	fragile X mental retardation gene1
FMRP	fragile X mental retardation protein
FRDA	Friedreich's ataxia
FXR1(2)	fragile X related gene 1 (2)
HD	Huntington's disease
Ile	isoleucine
kb	kilobase
kD	kilodalton
KH	K-homology
MMR	mismatch repair
mRNA	messenger ribonucleic acid
mRNP	messenger ribonucleoprotein
NES	nuclear export signal
NLS	nuclear localisation signal
PCR	polymerase chain reaction
SCA	spinocerebellar ataxia
UTR	untranslated region
wt	wild type
YAC	yeast artifical chromosome

Curriculum Vitae

Naam:	Carola Jacoba Maria Ubink-Bontekoe
Geboren:	23 november 1970, Poortugaal
1983-1989	VWO, R.K.S.G. Sint-Montfort, Rotterdam
1989-1995	studie Biomedische Wetenschappen, Universiteit Leiden hoofdvak: Moleculaire Biologie, NKI, Amsterdam
1995-2001	promotie onderzoek, afdeling Klinische Genetica, EUR
februari 2001	wetenschappelijk medewerker Fornix Biosciences N.V.

Publications

Bontekoe, C. J., E. de Graaff, G. J. Breedveld, B. A. Oostra and P. Heutink (1993). "Dinucleotide repeat polymorphism at D11S994 locus." Hum Mol Genet 2(10): 1747.

Heutink, P., T. Haitjema, G. J. Breedveld, B. Janssen, L. A. Sandkuijl, C. J. Bontekoe, C. J. Westerman and B. A. Oostra (1994). "Linkage of hereditary haemorrhagic telangiectasia to chromosome 9q34 and evidence for locus heterogeneity." J Med Genet 31(12): 933-6.

Willemsen, R., C. Bontekoe, F. Tamanini, H. Galjaard, A. Hoogeveen and B. Oostra (1996). "Association of FMRP with ribosomal precursor particles in the nucleolus." Biochem Biophys Res Commun 225(1): 27-33.

Bontekoe, C. J., E. de Graaff, I. M. Nieuwenhuizen, R. Willemsen and B. A. Oostra (1997). "FMR1 premutation allele (CGG)81 is stable in mice." Eur J Hum Genet 5(5): 293-8.

Tamanini, F., R. Willemsen, L. van Unen, C. Bontekoe, H. Galjaard, B. A. Oostra and A. T. Hoogeveen (1997). "Differential expression of FMR1, FXR1 and FXR2 proteins in human brain and testis." Hum Mol Genet 6(8): 1315-22.

Tamanini, F., C. Bontekoe, C. E. Bakker, L. van Unen, B. Anar, R. Willemsen, M. Yoshida, H. Galjaard, B. A. Oostra and A. T. Hoogeveen (1999). "Different targets for the fragile X-related proteins revealed by their distinct nuclear localizations." Hum Mol Genet 8(5): 863-9.

Bakker, C. E., Y. de Diego Otero, C. Bontekoe, P. Raghoe, T. Luteijn, A. T. Hoogeveen, B. A. Oostra and R. Willemsen (2000). "Immunocytochemical and biochemical characterization of FMRP, FXR1P, and FXR2P in the mouse." Exp Cell Res 258(1): 162-70.

Tamanini, F., L. L. Kirkpatrick, J. Schonkeren, L. van Unen, C. Bontekoe, C. Bakker, D. L. Nelson, H. Galjaard, B. A. Oostra and A. T. Hoogeveen (2000). "The fragile X-related proteins FXR1P and FXR2P contain a functional nucleolar-targeting signal equivalent to the HIV-1 regulatory proteins." Hum Mol Genet 9(10): 1487-93.

Nawoord

"Lieve hart,.....mijn boek is af."
Hoewel er geen siddering door het land zal gaan, en ik in vergelijking met de auteur van dit citaat geen geschiedenis zal schrijven, komt dit citaat me voor als het juiste. Ik ben vast niet de enige promovendus die een zekere verwantschap voelt met grote hoeveelheden koffie en de frustratie gelezen te willen worden.

Na mijn studie heb ik geen moment getwijfeld verder te willen "in de wetenschap". Een promotieplaats was dan ook een voor de hand liggende keus. Een keus waar ik nog steeds achter sta en welke ik weer zou maken. Evenzo voelt het nu als een goede keus mijn blik te verruimen buiten de universiteit. Proefondervindelijk vaststellen of er en wat het verschil is tussen "een baan en een carrière", tussen "financiële en wetenschappelijke belangen". Bij deze de belofte dat ik over enkele jaren de balans opmaak.

Vrienden, familie en collega's wil ik bedanken voor alle getoonde belangstelling voor wat mij de afgelopen jaren bezig heeft gehouden, zowel binnen als buiten de klinische genetica. Jullie vriendschap, vertrouwen, en kritiek heeft mij geholpen een aantal zaken binnen het juiste perspectief te plaatsen.

Mijn ouders wil ik bedanken voor hun grenzeloze aandacht en liefde. Als kind ervaar je een veilige thuishaven als vanzelfsprekend. Nu realiseer ik me dat dit geenszins vanzelfsprekend is en ben ik er dankbaar voor.
Lieve André en Freek, mijn nieuwe thuishaven, en voor Freek hoop ik even veilig en vanzelfsprekend. André, het spreekt voor zich dat de wetenschap dat "het wel maar werk is", "beta's geen managers zijn", "gelijk hebben er niet toe doet" en "je onvoorwaardelijk van me houdt", mij gesteund en gesterkt heeft. Mijn dank daarvoor.