Spring 1997

Molecular Diagnosis of Fragile X Syndrome and AGG Interspersion Analysis

Scott Hemmer
Carroll College, Helena, MT
Molecular Diagnosis of Fragile X Syndrome and AGG Interspersion Analysis

Submitted in Partial Fulfillment of the Requirements for Graduation with Honors to the Department of Biology and Chemistry at Carroll College, Helena, Montana

Scott A. Hemmer
April 11, 1997
This thesis for honors recognition has been approved for the Department of Biology and Chemistry by

Dr. Marilyn Schendel, Director
Linda Beischel
Dr. Grant Hokit

4-11-97
Date
ACKNOWLEDGEMENTS

I would like to express my sincere thanks to Linda Beischel for her gracious support and guidance of this thesis project and for offering me such a beneficial educational experience. I would also like to thank Marilyn Schendel for all her advice, help, and considerable contribution to this project.

I would also like to thank Dr. Grant Hokit and Dr. John Johnson without whom completion of this project would not have been possible. Finally, I would like to thank Shodair Hospital and everyone in the Genetics laboratory for their help.
ABSTRACT

Two molecular techniques used for studying Fragile X syndrome were tested in this study. The first technique determines the number of CGG repeats in the FRAXA repeat region. The second examines the number of CGG repeats without AGG interspersions. The first technique used PCR (polymerase chain reaction) to specifically amplify DNA from the repeat region that had been cut by the restriction enzyme Hind III. Then the amplified DNA size was determined using agarose gels stained with ethidium bromide. The first technique demonstrated that it could accurately determine the number of CGG repeats in most individuals. Thus, it shows promise in identifying individuals that do not need further fragile X testing. The second technique also amplified Hind III cut DNA. This DNA was then purified and cut with the restriction enzyme MnlI. These fragments were then also separated on an agarose gel and stained with ethidium bromide. The second technique demonstrated an ability to identify large regions of pure CGG repeats at the FRAXA site. This technique may in the future provide a method of determining premutation stability.
# TABLE OF CONTENTS

Acknowledgments .............................................i
Abstract ..................................................................ii
Table of Contents...........................................iii
List of Illustrations.......................................iv
List of Tables................................................v
Introduction and Literature Review....................1
  Background..................................................1
  Diagnostic Methods....................................4
  AGG Interspersion Analysis..............................11
Materials and Methods......................................16
  DNA Isolation.............................................16
  Hind III Digestion.......................................18
  PCR Amplification.......................................18
  Test Gel....................................................20
  Qia Quick Purification..................................20
  MnlI digestion.............................................22
  High Resolution Vertical Gel............................22
Results..........................................................24
Discussion.....................................................30
Literature Cited................................................40
LIST OF ILLUSTRATIONS

Fig. 1 Photograph of the fragile site (FRAXA) on the X chromosome.......................5
Fig. 2 DNA marker (MspI digest of pBR322 DNA).................23
Fig. 3 Agarose gel of alleles differing by only one repeat.................................26
Fig. 4 Agarose gel of a female sample with only one allele...............................31
Fig. 5 Primer location and MnlI restriction sites in the region of DNA studied............35
Fig. 6 Sample agarose gel showing MnlI digested samples....................................36
LIST OF TABLES

Table 1. CGG repeat numbers..........................27
Table 2. Analysis of Pure CGG repeats...............29
Introduction and Literature Review

Background

Fragile X syndrome is the most common inheritable cause of mental and developmental disability among males and females. It affects approximately one out of every 1250 males and one out of every 2000 females. Some of the phenotypic characteristics of this disease are developmental delays, mental impairment, a long face, macroorchidism, and large ears (Tarleton and Saul, 1993). Fragile X is an X-linked dominant disorder that shows reduced penetrance. Thirty to fifty percent of carrier females demonstrate mental impairment while more than 80% of males with a fragile X chromosome demonstrate signs of the syndrome (Warren and Nelson, 1994). Fragile X syndrome is caused by the absence of expression of the fragile X mental retardation 1 gene (FMR-1). Mutations in this gene prevent the formation of the fragile X mental retardation protein (FMRP). (Pieretti et al. 1991) FMRP is a known RNA binding protein. It is believed to be involved in nuclear export within the brain. This suggests that the symptoms of fragile X may be due to altered translation of transcripts that normally bind to FMRP. (Eberhart et al. 1996) The FMR-1 gene contains a region at its 5' end that is rich in CGG triplet repeats. This triplet repeat region appears to be the primary culprit in preventing expression of FMR-1 (Jacobs, et al. 1993). In many fragile X patients the
number of these CGG repeats has greatly expanded. When the number of repeats reaches a minimum of about 200-250 repeats (vs. 6-50 for normal individuals), hypermethylation occurs. This increased methylation is found in both the repeat region and a CpG island (a class of regulatory sequences characterized by a high density of cytidine phosphate guanosine dinucleotides) that is upstream of the repeat. It is believed that the methylation prevents transcription of the FMR-1 gene which results in the fragile X phenotype (Rousseau et al. 1991).

The reproductive fitness of individuals with fragile X syndrome is usually greatly reduced due to the relatively serious effects of the syndrome. This suggests that a large portion of the fragile X mutations arise de novo (Macpherson et al. 1994). However, fragile X is a relatively common genetic disease. If most of the mutations are de novo, this would suggest that the mutation rate is abnormally high. The high mutation rate was explained once the mutation process was understood (Macpherson et al. 1994). As was stated earlier, normal individuals usually have between 6 and 50 repeats and people with fragile X have greater than 200 repeats. It was later discovered that there is a third class of repeats known as premutations. The range of repeats for this class usually starts around 50 and goes up to 200 repeats (Tarleton and Saul, 1993). Individuals with premutations do not experience hypermethylation as do individuals with full mutations. However the number of CGG
repeats in premutations is unstable during meiosis of female gametes (Fu et al. 1991). The number of repeats can increase or decrease in these unstable alleles. However, there is a much greater tendency to increase (Warren et al. 1994). This explains why fragile X is relatively common. An individual can carry a premutation with no adverse effects and pass it on to his or her offspring. The premutation will continue to change in size as it is passed to the successive generations. Eventually, the number of repeats in the premutation will expand until it becomes a full mutation.

This disease also demonstrates a parent of origin effect similar to genetic imprinting. Transition from premutation to full mutation only occurs in female gametes (Warren et al. 1994; Fu et al. 1991) Therefore, males carrying a premutation can pass that premutation to their immediate offspring without expansion to a full mutation. These individuals are known as "normal transmitting males". However, this is not true of females. Premutations can experience large expansions to full mutations in female gametes (Fu et al. 1991). Therefore, offspring of females who have a premutation are at risk for fragile X syndrome.

Fragile X syndrome also exhibits a form of genetic anticipation that has been named the Sherman paradox. Genetic anticipation is said to occur when the severity of a disease increases with successive generations. Sherman (1984, 1985) noticed there is a 9% risk of fragile X
syndrome in the brothers of normal transmitting males, 40% in their grandsons, and 50% risk in their great grandsons (Warren and Nelson, 1994). As the premutation is passed on it slowly expands and becomes increasingly unstable. Once the number of repeats becomes greater than 90 there is a 100% chance of a transition to a full mutation when passed through a female (Reiss et al. 1994).

Diagnostic Methods

The first method for diagnosing fragile X syndrome utilized cytogenetic analysis of the chromosomes of cells grown in special media. These media were either deficient in folate and the deoxynucleotide thymidine or contained compounds that inhibited folate metabolism or thymidine synthesis. Under these conditions, the X chromosome demonstrated a constricted region termed a fragile site (FRAXA) at a band of the long arm of the X chromosome called Xq27.3 (Tarleton and Saul, 1993; Warren and Nelson, 1994) (see fig. 1). These conditions resulted in the expression of a fragile site because of the high GC content of the region. Individuals with the fragile X full mutation have a large number of CGG repeats. Therefore, DNA replication of this area requires large amounts of deoxycytidine and deoxyguanosine. It has been hypothesized that culturing the cells in conditions which decrease thymidine also led to insufficient deoxyguanosine pools.
Culturing the cells in folate deficient media was believed to disrupt the function of tetrahydrofolate reductase which resulted in inhibition of the production of deoxycytidine. Insufficient amounts of the nucleotides deoxycytidine and deoxyguanosine led to incomplete replication of the repeat region. (Tarleton and Saul, 1993) This incomplete replication showed up visibly as the fragile site seen in cytogenetic testing.

Figure 1. Fragile site. X chromosome exhibiting fragile site (arrow) compared to a normal X chromosome.

The cytogenetic test for FRAXA has many shortcomings and is seldom used today to diagnose fragile X syndrome. As
was stated above, culturing cells under certain conditions can induce expression of the fragile site in those with full mutations, but only in a minority of cells. A positive diagnosis of fragile X syndrome was usually made if an individual expressed the fragile site in 5% or more of their cells. Individuals with 1% to 4% expression of the fragile site were left in a gray area (Tarleton and Saul, 1993). This is why cytogenetic testing led to either false positives or negatives. Cytogenetic methods also were ineffective at identifying premutations in transmitting males or carrier females (Tarleton and Saul, 1993). However, some individuals referred for fragile X testing may have other chromosomal abnormalities. Therefore, a routine chromosome analysis (without testing for the FRAXA site) is still part of the diagnostic evaluation in many genetics laboratories.

Linkage analysis has also been used to diagnose fragile X syndrome. Linkage analysis uses genetic markers that are close to the FMR-1 gene. These markers vary greatly in the population. Restriction enzymes that recognize specific sequences in the DNA cut these marker regions into fragments. The sizes of these fragments will vary because each individual will have slight differences in their DNA sequence. The fragments are called restriction fragment length polymorphisms (RFLPs). Fragments from the marker region can be detected by DNA probes that are specific for
DNA in this region. Analysis of RFLPs was helpful in determining the risk of inheritance of fragile X.

Linkage analysis uses markers that are close to the gene of interest (e.g., FMR-1). The close proximity of the markers is intended to cut down on the odds of recombination occurring between the marker and the gene of interest. Therefore, the markers should be inherited with the associated gene. This method enabled researchers to predict which members of a fragile X family may be carrying a premutation. If a family member had the same genetic markers as the affected individual, then it was likely they were a carrier. However, this method also has problems. To perform a linkage analysis, someone in the family must have already been diagnosed with fragile X syndrome. Researchers also need many individuals of a family to participate in the study so that reference chromosomes can be identified and statistical risk assessments can be made. Finally, with linkage analysis, the possibility of genetic recombination can never be ruled out (Tarleton and Saul, 1993).

The next advancement in fragile X diagnosis was the development of techniques that utilize methylation-sensitive restriction enzymes and Southern blot techniques. As was stated before, individuals with fragile X syndrome exhibit a hypermethylation of the repeat region and of a CpG island upstream of the FMR-1 mutation region (Rousseau et al. 1991). Treating DNA with a methylation-sensitive enzyme (e.g., Eag I) whose restriction site is located in an area that has
become methylated can provide accurate diagnosis of fragile X. If the DNA is first treated with EcoRI it will generate approximately a 5.2 kb fragment using Southern blot analysis. If this non-methylated fragment is also treated with Eag-I the fragment will be cut into two smaller fragments. However, if the CpG island is methylated EagI will not cut the fragment. (Rousseau et al. 1991; Tarleton and Saul, 1993). The size of the fragments formed will also vary if there has been an expansion of the repeat region. The restriction fragments are identified with labeled probes specific for the DNA. Southern blot analysis not only identifies methylation status but also often identifies individuals that have had an expansion but do not exhibit hypermethylation (Tarleton and Saul 1993). Southern blot analysis has become a very common technique for fragile X diagnosis, but has a limited ability to identify small premutations. Smaller premutations are close in size to normal alleles and often analysis cannot distinguish between them, especially in females who may have two normal alleles. Therefore, females sometimes exhibit blurring or three bands upon analysis. (Snow et al. 1993; Tarleton and Saul 1993)

The next advance in diagnosis occurred once the gene sequence of the FMR-1 gene was deciphered. This information facilitated the development of primers for polymerase chain reaction (PCR) techniques. PCR could then be used to determine the number of CGG repeats in the FMR-1 gene. PCR is a method that uses DNA primers and polymerases to copy
specific regions of DNA multiple times. PCR has been primarily used to complement Southern blot analysis. PCR has many advantages over Southern blot. It requires much less DNA and can more accurately determine the number of CGG repeats (Snow et al. 1996). PCR is also relatively fast and inexpensive. Therefore, it is a more effective method of detecting normal alleles and diagnosing premutations.

Originally, there were some difficulties in using PCR to amplify the fragile X region (Papp et al. 1996). For the DNA to be duplicated, it must first be denatured by raising the temperature. In PCR a polymerase is used which is stable at these high temperatures. However, the high CG content of the repeat region interfered with the denaturing process. Apparently the repeat region formed secondary structures that were extremely stable (Papp et al. 1996). Therefore, to denature the DNA, the temperature had to be raised to a point above the temperature at which the polymerase was stable. This problem was first overcome by using 7-deazaguanosine triphosphate (7-deazaGTP). 7-deazaGTP lowers the "melting" temperature of the DNA (Baskaran et al. 1996). This allows the polymerase chain reaction to be run at a lower temperature that is compatible with the polymerase. Non-radioactive techniques usually identify DNA in gels by staining with ethidium bromide. Ethidium bromide is an intercalation agent that is incorporated in-between the stacked bases of DNA. 7-deazaGTP interferes with ethidium bromide staining and

Ethidium bromide is an intercalation agent that is incorporated in-between the stacked bases of DNA.
therefore cannot be used in labs that utilize ethidium bromide detection techniques (Baskaran et al. 1996). Although the use of 7-deazaGTP solves the denaturation problem, modifications in the procedure must occur if it is to be used without radioactive markers.

PCR techniques do not usually allow detection of individuals who have full mutations due to insufficient amplification of the target DNA or incomplete entry of the DNA into the gel (Snow et al. 1993). This is why it is used along with Southern blot analysis.

PCR can be effective in ruling out individuals who are negative for fragile X syndrome (unless the syndrome is due to a mutation other than an expansion). Detection of one normal allele (<50 repeats) in males or two in females is usually sufficient evidence to assume that the individual does not have a full mutation. There have been examples of individuals who are "mosaic" and have both a full mutation in some cells and a premutation in others. However, mosaic individuals with both normal alleles and full mutations have not been found (Rosseau et al. 1991). It is believed this is true because a normal allele is mitotically stable and will not expand during mitosis to form a mosaic individual. Therefore, if PCR analysis shows a male individual with one normal band or a female with two it is relatively safe to assume the individual is negative for fragile X syndrome (unless they have a mutation other than an expansion).
The lower range for premutations at the FMR-1 locus is 40-55 repeats (Eichler et al. 1994). A range is given because the exact point at which alleles become unstable is unknown. Alleles having as many as 52 repeats have been found to be stable while at the same time unstable alleles with only 34 repeats have been identified. For this reason alleles in the range of 40-55 repeats are usually referred to as "gray-zone" alleles (Eichler et al. 1994). The fact that some alleles in this range are stable while others are susceptible to expansion suggests that length is not the only factor affecting instability. It is unknown at this time exactly why or how the repeat region becomes unstable and expands. However, theories have developed that try to explain why the expansion occurs.

The most accepted theory at this time is that slippage of the Okazaki fragment of DNA may occur during replication (Eichler et al. 1994; Rinh and Sinden, 1991; Gacy et al. 1995) In DNA replication the "lagging strand" is replicated in small fragments which are later joined. These fragments are called Okazaki fragments. It is believed that in the large stretches of CGGs one or both ends of the fragment may slip which will lead to an expansion in the number of CGG repeats.
The cause for Okazaki fragment slippage is still unknown. Some researchers believe that in normal sized repeat regions only one end of the Okazaki fragment is capable of slipping because the other end is anchored in an area outside the repeat region (Eichler et al. 1994). Therefore, the number of repeats can only change slightly. As the number of repeats slowly grows it eventually gets large enough so that the entire Okazaki fragment is within the repeat region. At this point it is possible for both ends of the fragment to slip at the same time which could explain the large expansions seen during the transition of premutations to full mutations (Eichler et al. 1994).

A second explanation for Okazaki fragment slippage may be the formation of hairpin structures in the Okazaki fragment. This theory suggests that as the size of the repeat increases it becomes more thermodynamically stable for the lagging strand of DNA to form a hairpin loop (Trinh and Sinden, 1991; Gacy et al. 1995). When a hairpin loop forms the unattached end of the Okazaki fragment slips. Normally slippage of one end can only cause small changes in repeat number because any greater changes would be thermodynamically unfavorable. However, hairpin loop formations can make it thermodynamically possible for larger expansions to occur (Gacy et al. 1995).

These two theories suggest slightly different methods for expansion, but they do agree on the point that non CGG interspersions found in the repeat region have a stabilizing
effect that can prevent expansion. These interspersions have been found as AGG segments. The first theory suggests that AGGs act as anchors. These "anchors" help prevent slippage of the Okazaki fragment thus preventing expansion (Eichler et al. 1994). The hairpin loop theory suggests that AGGs prevent expansion because they inhibit formation of the hairpin loop structure. AGGs can make hairpin loop formation thermodynamically unfavorable and consequentially prevent expansion (Gacy et al. 1995; Trinh and Sinden 1991).

A common pattern seen in many alleles is (CGG)$_n$-AGG-(CGG)$_n$-AGG-(CGG)$_{10}$ (Eichler et al. 1994, 1996). The actual number of AGGs in a repeat region and the repeat pattern can vary greatly. This is why AGG interspersion could be important in determining the difference between normal alleles and premutations in the gray-zone. It appears that the number of pure CGG repeats (repeats without any AGGs) may be more important in determining instability than the total number of repeats. Recent tests have shown smaller premutations that expand usually have only one AGG interspersion (Eichler et al. 1994). If an allele only has one AGG it means that there are a larger number of "pure" CGG repeats (not interrupted by AGGS). Research has also shown that the larger stable alleles usually have 3 or 4 repeats which decreases their number of pure repeats (Zhong et al. 1995; Eichler et al. 1994).
The purpose for this current study is twofold. First, this study tests a new non-radioactive method for analyzing the size of the FRAXA repeat region and determines the success of this technique in identifying normal alleles. This method uses betaine and dimethylsulfoxide (DMSO) (instead of 7-deazaGTP) in the PCR reaction to help in denaturing the DNA. Betaine and DMSO are agents that inhibit secondary structure and help promote melting of the DNA. However, they do not inhibit ethidium bromide staining like 7-deazaGTP (Baskaran et al. 1996). If successful, this method could be used to identify fragile X negative individuals.

Second, this study looks at the number of pure CGG repeats using PCR followed by MnlI digestion. The restriction enzyme MnlI recognizes the sequence GAGG and cuts the DNA, and thus it will recognize AGG interspersions (CGG-AGG-CGG) and cut the DNA in this area. Therefore, the size of fragments produced by MnlI digestion of PCR products from the CGG repeat region may be used to determine the number of pure CGG repeats. Repeats with few or no AGG interspersions will produce larger fragments after MnlI digestion. The various sizes of these fragments can be determined by agarose gel electrophoresis and comparison to DNA fragments of known size. This method was specifically tested on DNA from two individuals from the same family. These individuals both had repeat numbers in the upper end of the gray zone, but one of these alleles was stable and the other eventually expanded. Ultimately, this technique could also
be used diagnostically to provide information on the stability of small premutations and especially mutations in the gray zone.
MATERIALS AND METHODS

DNA Isolation

Briefly, a low salt solution was added to whole blood to lyse the red blood cells. White cells were lysed by a solution containing detergent. High salt precipitated proteins from the solution. The remaining DNA was precipitated from solution with alcohol and redissolved in a dilute salt solution for storage.

The vacutainer tube containing the subject's blood sample was opened in a laminar flow hood. "Universal precautions" were used in handling blood in the lab. Then the sample was poured into a 15 or 50-ml centrifuge tube and the volume of blood was estimated to the nearest 1 ml. This volume was considered "one volume" for all subsequent steps. A total of 3 volumes of RBC lysis solution (Puregene DNA Isolation Kit D-5000, Gentra Systems) was added to the blood sample. A small portion of this volume was used to rinse out the vacutainer tube. Then the tube was inverted to mix, incubated for 10 minutes at room temperature, and centrifuged for 10 minutes at 3200 rpm. The supernate was poured off leaving behind a visible white cell pellet and 0.2-0.3 ml of residual liquid. The tube was vortexed vigorously to resuspend the pellet and next one volume of "cell lysis solution" (Puregene) was added. Then 0.005 ml
"RNAse solution" (Puregene) per 1 ml of original volume was added to the cell lysate. The sample was then inverted 25 times in order to mix and incubated in a 37°C water bath for 15 minutes.

After the incubation, the sample was cooled to room temperature and a 1/3 volume of "protein precipitation solution" (Puregene) was added to the lysate which was then vortexed vigorously for 20 seconds. Then the sample was centrifuged at 3200 rpm for 10 minutes. One volume of isopropanol (Fisher) was poured into new tubes. The supernate from the preceding step was then poured into the isopropanol, which left the precipitated protein pellet behind to be discarded. The supernate was mixed with the isopropanol by gently inverting the sample tube 50 times until white threads of DNA became visible. Using a sterile plastic inoculating loop, the precipitated DNA was transferred into a 1.5-ml microcentrifuge tube containing 70% ethanol in water. The sample was centrifuged (Eppendorf 5415) for 1 minute at 14,000 rpm to pellet the DNA. The supernate was poured off and a sterile swab was used to remove any remaining drops of supernate. The sample was allowed to air dry for 10 minutes.

Next 1/10 volume "DNA hydration solution" (Puregene) was added. The solution was mixed and heated to approximately 50°C for a maximum of 1 hour in order to dissolve the pellet. After the pellet was dissolved, the DNA
concentration was estimated using a fluorometer. Finally, the DNA was stored at 4° C.

**Hind III Digestion**

Prior to PCR amplification, the sample DNA was cut by the restriction enzyme Hind III (20 u/μl) (New England Biolabs). The reaction mixture for the digestion was distilled water, 1X NEB 2 buffer (New England Biolabs), sample DNA, and a excess (10 units/μg DNA) of Hind III. The samples were allowed to digest at 37° C for at least one hour. This predigestion has been reported to increase the yield of specific PCR product from the FMR-1 CGG repeat region. It is hypothesized that this is accomplished because Hind III cuts other areas of CGG repeats in the genome and thus prevents their amplification (Papp et al. 1996).

**PCR Amplification**

The first set of PCR amplifications was performed on DNA from 24 "normal" females. "Normal" individuals were not believed to have fragile X. PCR was also performed on several male DNA samples as controls. The DNA was pre-cut with Hind III before PCR. The Hind III cut DNA (400ng) was added to a reaction mixture containing: 1X cloned Pfu
buffer (Stratagene), 0.2 mM dNTPs, the primers FXD and FXE, 2.5 M fresh betaine, 2% DMSO, and Rediload dye (Research Genetics) in a total reaction volume of 25 μl. Cloned Pfu polymerase (0.5 μl, 1.25 units) (Stratagene) was also added to the top side of the tube and mixed in after the tubes were heated to 98°C to achieve a "hot start". This PCR reaction was adapted from Papp et al. 1996.

In the second set of experiments, two sets of primers were used. These primer sets were FXD/FXE and FXF/FXR. The samples amplified using the FXD/FXE primers used the same reaction mixture as the first set of experiments. However, the amount of reaction mixture and DNA was doubled so the total sample volume was 50 μl. Samples were also amplified with a separate set of primers (FXF/FXR) which created a larger PCR product than the FXD/FXE primers. The reaction mixture for this primer set was 0.1 mM dNTPs, 2.5 M betaine, Rediload, 2% DMSO, 1X Pfu buffer, water and primers FXF/FXR (total reaction volume 50 μl). One μl of cloned Pfu polymerase was placed on the side of the tube for all samples (FXD/FXE and FXF/FXR) and mixed in after 15 seconds of heating.

The two experiments were amplified using the same PCR algorithm. First, the samples were heated to 98°C for one minute. The tubes were then vortexed to mix in the polymerase. After the polymerase was mixed in, the tubes were placed back into the thermocycler and heated at 98°C
for 15 seconds. Next the mixtures were cooled to 65° C for two minutes to allow the primers to anneal. Finally, the mixtures were heated to the elongation temperature of 75° C for three minutes. The mixtures went through 28 of these cycles. On the last cycle the samples were kept at 75° C for 13 minutes and then cooled to 4° C.

**Agarose Test Gel**

After the PCR and MnII digestion a horizontal test gel was run to confirm the presence of product and to help estimate the amount of product present. These gels were made from 1.5% LE/TAE (0.04M Tris, 11.4 mL/L acetic acid, 0.002 M EDTA, pH 8.5) agarose. Approximately 35 ml of agarose was used for each gel. The electrophoresis tray was filled with 1X TAE buffer. 5μl of sample was added to each well. The gels were often run with a 100 base pair marker for size comparison. The marker was mixed in a 6:1 ratio with loading dye. The gel was electrophoresed at 80V for approximately 30 minutes. Then the gel was dyed with ethidium bromide (2μl/ml) and viewed under UV light.

**Qia Quick Purification**

Prior to MnII digestion the PCR products had to be purified. If the DNA was not purified reagents in the PCR
mixture, such as betaine and DMSO, would inhibit the enzyme's ability to digest the DNA. Purification was performed using the QIA Quick system (Qiagen) which binds DNA to a solid matrix in a high salt buffer while impurities are washed away, and then elutes concentrated DNA in a low salt buffer.

First, 5 volumes PB buffer (QIA purification kit) was added per 1 volume of the PCR reaction (e.g. 250 µl buffer & 50 µl sample or 150 µl buffer to 30 µl sample). Then the mixture was placed in a QIA quick spin column which rested inside a 2-ml collection tube. Next the sample was applied to the center of the filter and microcentrifuged at 14000 rpm for 30-60 seconds. After centrifugation the flow through was discarded and the column was reinserted in the 1.5-ml microcentrifuge tube. In the following step 0.75 ml PE buffer (QIA purification kit) was added to each QIA quick column and centrifuged (14000 rpm) for 30-60 sec. After centrifugation the flow through was discarded and the column was again placed in the same 1.5-ml microcentrifuge tube. The tube was centrifuged for an additional 1 min at maximum speed. Then the QIA quick columns were placed in clean 1.5-ml microcentrifuge tubes. Thirty µl of distilled water was added to the center of the QIA quick column and centrifuged for 1 minute at maximum speed in order to elute the sample. The flow through resulting from this centrifugation was the purified DNA.
MNLI Digestion

During analysis of "gray zone" alleles, DNA was cut by the restriction enzyme MnII (5u/μl) after the PCR and purification. The reaction mixture contained 1X NEB2 buffer (New England Biolabs), 1X BSA (bovine serum albumin), purified, amplified DNA, and an excess of MnII. The amount of MnII used depended on which primer set was used during PCR. Samples amplified using primers FXD/FXE were digested with MnII at a 1 unit enzyme to 6 μl volume ratio. Samples amplified with FXF/FXR were digested with 1 unit enzyme per 3 μl volume. The samples were digested at 37° C overnight.

High Resolution Vertical Gel

Once the test gel confirmed the presence of the expected product, a higher resolution vertical gel was run to determine the number of repeats. These vertical agarose gels were 2mm thick. They were composed of 3.5% metaphor/TBE (0.089M Tris, 0.089 M boric acid, 0.002 M EDTA, pH 8) agarose. These gels were cooled for 30 minutes prior to electrophoresis. The amount of sample added to each lane varied depending on how robust the PCR amplification was and on the size of the products being viewed. Usually about 10-15 μl were added. The samples were mixed with loading dye
in approximately a 4:1 sample to dye ratio. A marker (MspI digest of pBR322 DNA) (fig. 2) was run with the sample DNA in a separate lane. This DNA was mixed with 1X TBE and loading dye in a 1:10:5 ratio. The gels were electrophoresed at 4°C in 1X TBE buffer for 2-2.5 hours at 350V. After electrophoresis the gels were stained in ethidium bromide for 20 minutes followed by destaining in distilled water. Once the gel was destained, the migration distance of the DNA marker fragments was measured (see fig 2.). Using this distance and the log of the fragments size, a standard curve was created. This standard curve was then used to extrapolate the sizes of the PCR products.

**Figure 2.** DNA Marker (MspI digest of pBR322 DNA)
The size of each fragment is labeled on the right in base pairs.
RESULTS

The Polymerase Chain Reaction (PCR) is a molecular genetic technique that imitates natural DNA replication. It uses a thermocycling machine (Perkin Elmer) to raise and lower the temperature of the samples. First the machine raises the temperature to a "melting temperature" in order to denature the DNA. Then the temperature is lowered to the annealing temperature. At this temperature oligonucleotide primers specific for certain region of DNA can bind to the DNA. These primers act as a starting point for the DNA polymerase. In the third step the sample is again heated. This time it is heated to an elongation temperature. During this stage the polymerase begins replication at the primers. The denaturation temperature will denature ordinary polymerases. Therefore, heat stable polymerases are used. These polymerases are often isolated from bacteria that live in hot springs. These bacterial polymerases are stable at the higher temperatures needed to perform PCR.

The first objective of this study was to test a new non-radioactive PCR technique. The goal was to determine whether or not this method could be used to rule out individuals for further fragile X syndrome testing. In order to do this, samples of DNA were amplified using PCR and then analyzed to determine the number of repeats. If both of a female's alleles showed a repeat number that was within the normal range (0-40) then the result was negative. Twenty-four
"normal" females were tested using this technique. The test gave negative test results for 17 of the 24 females. These 17 females all had two different normal sized alleles (<40). Each individual was tested at least two times and the negative results were confirmed. The two sets of PCR results agreed on both the number of alleles and the size of the alleles usually only varying by 0-2 repeats. These minor differences can be easily explained by human error in interpreting the data. The remaining 7 females that this testing did not rule out all exhibited only one allele band. These individuals would require further testing to rule out fragile X syndrome. This testing is needed in order to rule out a full mutation. A full mutation would not be visible because of problems in amplification or due to incomplete entry into the gel. Therefore, it is possible these females have a full mutation which is why they only exhibit one allele. It is also possible that these females are homozygous for an allele and therefore exhibit only one visible band.

The amplification using betaine and DMSO was robust and appeared to effectively replicate the CGG region. Amplification of the repeat region is often hindered by its high GC content. This is why substances must be added to inhibit the secondary structure of the DNA and promote melting. 7-deazaGTP has previously been used to accomplish this objective. The procedure tested in this study used betaine and DMSO instead of 7-deazaGTP because of their
compatibility with ethidium bromide staining. The results indicated that betaine and DMSO facilitated melting of the DNA strands. This was demonstrated by the fact that amplification of alleles containing at least 100 repeats was successfully achieved. The technique was also able to differentiate between two alleles that differ by only one repeat (three base pairs) (See fig. 3). The sample population demonstrated an increased number of alleles between 28-30 and a smaller increase at 18-19. The largest allele seen was 39 repeats while the smallest was 17 repeats.

Figure 3. Two samples having two alleles with 29 and 30 repeats.
<table>
<thead>
<tr>
<th>SAMPLE</th>
<th># Repeats</th>
<th># Repeats</th>
<th># Repeats</th>
<th># Repeats</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(7/10)</td>
<td>(8/5)</td>
<td>(8/13)</td>
<td>(8/15)</td>
</tr>
<tr>
<td>1</td>
<td>17,29</td>
<td>19,30</td>
<td>18,28</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>30,39</td>
<td>31,40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>19,29</td>
<td>18,28</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>21,28</td>
<td>22,29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5*</td>
<td>28</td>
<td>29</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>24,29</td>
<td>19,29</td>
<td>18,29</td>
<td></td>
</tr>
<tr>
<td>7*</td>
<td>24</td>
<td>19</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>8*</td>
<td>30</td>
<td>29</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>30,32</td>
<td>30,32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>28,39</td>
<td>30,40</td>
<td>29,39</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>28,35</td>
<td></td>
<td>30,36</td>
<td>29,35</td>
</tr>
<tr>
<td>12*</td>
<td></td>
<td></td>
<td>30</td>
<td>29</td>
</tr>
<tr>
<td>13</td>
<td>23,31</td>
<td>21,29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>24,32</td>
<td>23,31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15*</td>
<td></td>
<td></td>
<td>30</td>
<td>29</td>
</tr>
<tr>
<td>16</td>
<td>29,32</td>
<td>29,32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>29,30</td>
<td>29,30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19*</td>
<td></td>
<td></td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>25,29</td>
<td>25,29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21*</td>
<td></td>
<td>29</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>28,30</td>
<td>28,30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>22,28</td>
<td>22,28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>29,30</td>
<td>28,29</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Non-negative test

This table demonstrates the repeat size determined for each sample. Each sample was tested at least twice. The samples were analyzed on four separate dates which are shown.
A second objective of this study was to look at the number of pure repeats using MnII digestion. This portion of the study was not as effective. The FXF/FXR and FXD/FXE primers both appeared to amplify "gray zone" alleles adequately. The FXF/FXR primer worked especially well and exhibited very robust amplification. The MnII digestion also appeared to have the desired effect. The test gels run after MnII digestion showed that the amplified DNA had been completely digested. A problem did occur when identifying the alleles after digestion. Staining of the agarose gels was done with ethidium bromide which did not effectively stain the small DNA fragments created by MnII digestion. The FXF/FXR primer set exhibited greater amplification than the FXD/FXE set. This made it possible to identify smaller fragments with the FXF/FXR primer set. Eight DNA samples (seven male, one female) were tested to determine the number of pure CGG repeats. The largest number of pure repeats found was approximately 54 repeats. Based on the size of the fragments left after MnII digestion it was possible to hypothesize on the structure of the repeat region. The structures are hypothetical and many are incomplete. (see Table 2) However, it did allow determination of any large pure repeat regions which was the primary goal for this study.
<table>
<thead>
<tr>
<th>Sample #</th>
<th># of pure repeats</th>
<th>possible CGG repeat pattern</th>
<th>Total # of Repeats</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>53</td>
<td>?-AGG-CGG(54)</td>
<td>64</td>
</tr>
<tr>
<td>2</td>
<td>29</td>
<td>CGG(8)-AGG-???-AGG-CGG(30)</td>
<td>50</td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>CGG(9)-AGG-???-CGG(14)</td>
<td>38</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>CGG(9)-AGG-CGG(26)</td>
<td>35</td>
</tr>
<tr>
<td>5</td>
<td>30</td>
<td>CGG(8)AGG-???-AGG-CGG(31)</td>
<td>48</td>
</tr>
<tr>
<td>6</td>
<td>24</td>
<td>CGG(9)-AGG-CGG(25)</td>
<td>36</td>
</tr>
<tr>
<td>7</td>
<td>22</td>
<td>CGG(8)-AGG-???-AGG-CGG(23)</td>
<td>41</td>
</tr>
</tbody>
</table>

These samples were amplified, purified, and then digested. This was done in order to look for any large pure repeats (CGG repeats without AGG interspersion). Samples 1 & 2 are the two alleles from the same family who were the primary subjects in this study. (Female not shown)
DISCUSSION

The vast majority of patients who are tested for fragile X syndrome do not have an expansion of the CGG repeat region. This means the majority of tests for a fragile X mutation are negative. Using nonradioactive PCR and agarose gels to eliminate individuals with normal sized alleles is time and cost effective relative to radioactive PCR and polyacrylamide gels or traditional Southern blots.

Normal sized alleles are mitotically stable (Fu et al. 1994). Therefore, identifying males with one normal allele or females with two normal alleles will eliminate the need for any further testing for a full mutation. The nonradioactive PCR method can effectively rule out all individuals with normal numbers of CGG repeats except females who are homozygous for alleles of equal size, males with greater than 50 repeats, or individuals whose DNA does not amplify. In this study, approximately 71% of the females tested were heterozygous and had two normal alleles. This is comparable to previous figures that suggest the frequency of heterozygous females in the population to be approximately 80% (Wang et al. 1993). The remaining females exhibited only one band and would require further testing to rule out a full mutation. However, it seems probable that these females are actually homozygous and do not have a full mutation. There are several reasons for this. The first is the fact that full mutations are rare, but homozygous
individuals are not (20%) (Wang et al. 1993). A second reason is that in these females their one allele was usually around 29-30 or 19 repeats. These are the two most frequent alleles in the population (Fu et al. 1991; Jacob et al. 1993). This therefore means that females are more likely to be homozygous for one of these alleles. A third reason is the increased fluorescence seen in these females alleles (See fig. 4). Homozygous females have only one band but this one band should have a greater amount of DNA which should lead to brighter bands. Male samples which fail to amplify, or which show an allele of greater than 50 repeats, and female samples which show only one allele less than 50 repeats must be analyzed by Southern blot.

Lanes 1 2 3

Figure 4. Two female sample with single bands that show increased fluorescence (lanes 1 & 2) compared to a an individual with two bands (lane 3).
7-deazaGTP has had success in amplifying fragile X premutations. 7-deazaGTP destabilizes the DNA secondary structure which is necessary when amplifying GC rich regions (Baskaran et al. 1996). If the DNA secondary structure is not destabilized then complete melting will not occur and amplification will not work effectively. However, amplification of larger alleles has not been as successful (Warren et al. 1994). Another problem with 7-deazaGTP is that at higher concentrations it also inhibits ethidium bromide staining. DMSO and betaine (N,N,N-trimethylglycine, an isostabilizing quaternary amine) do not inhibit ethidium bromide staining, but are effective at lowering the melting temperature of GC rich DNA (Baskaran et al. 1996). This new non-radioactive method appears to be capable of amplifying premutations that are at least 100 repeats in length and probably even greater. The range of alleles in our sample study of "normal" individuals was from 17-39 repeats. Our sample population demonstrated an increased number of alleles with 18-19, 21, and 28-30 repeats. These results correlate well with other studies of CGG repeats (Fu et al. 1991) (Jacob et al. 1993).

The procedure used in our study appears to be a promising method of ruling out normal individuals. It is relatively cost efficient, reliable, and fast. It may prove to be a favorable alternative to other fragile X diagnostic techniques. One limitation of this technique is that it only looks at the FRAXA site. Recent studies suggest
expansion at other sites (e.g., FRAXE) may result in the fragile X phenotype (Snow et al. 1993). This method would not identify these individuals. It also would not identify individuals who exhibited fragile X syndrome due to deletions or other mutations in the FMR-1 gene.

The results of the AGG interspersion analysis were limited. The MnII restriction enzyme cuts DNA at the sequence 3'...GGAG(N)₆...5'/5'...CCTC(N)₇...3' (see fig. 5). This means that MnII should cut the repeat region at points 7 bases to the 5' side of the AGG. Therefore, the DNA that was replicated using primers FXD/FXE would form a 15 bp fragment that would not vary in size, and other variable size fragments. The number of these variable fragments would depend on the number of AGGs found in the repeat region. If a repeat region has two AGGs it will have three variable sized fragments; if it has only one AGG it will form two fragments. If the sample had two AGGs, then the first variable fragment should be 30 base pairs (bp) long + (3) X (the number of repeats to the first AGG). A second would be 26bp + (3) X (the number of repeats to the last AGG), the size of the third middle fragment would be equal to the number of pure repeats between the two AGGs + 3. (see fig. 5) The invariant 15 bp repeat will not be visible because it is too small. The variable regions may not be visible either depending on the number of pure repeats (especially the middle fragment). Digestion of FXF/FXR DNA should form five small invariable fragments.
which were not visible. It should also form two larger invariable fragments (74,78) which are visible (see fig 6). This primer set forms variable fragments in a like manner to the FXD/FXE primer set. The variable fragments would all be identical in size to the FXD/FXE fragments except that the first fragment would be one base pair longer.

The digestion of PCR products by MnI appeared to be complete since there were no large bands visible on the agarose test gel following digestion. Previous studies with MnI digestion of the repeat region have been done, but in these experiments only partial MnI digestion was accomplished (Eichler et al. 1994) (Zhong et al. 1995). In our experiments, incomplete digestion was not a problem. The main problem was that amplification with the FXD/FXE primers did not produce sufficient product to allow identification of the small products that remained after MnI digestion using ethidium bromide staining. The FXF/FXR amplification was more robust and worked better for identifying these digestion products. In some of the subjects it was even possible to reconstruct a hypothetical model of the AGG interspersion pattern. Many of these models are incomplete. A possible reason for this may be that the individuals had two AGG interspersions and after digestion the middle fragment was too small to detect. This is because unless this middle segment had greater than 15 pure CGG repeats the fragment formed would be smaller than 50 base pairs which was below the limit to detect.
Figure 5. Partial DNA sequence (Genebank) of the FMR-1 gene.
Figure 6. Agarose gel of PCR products. Lanes 2, 7, and 11 are FXF/FXR amplified samples digested with MnI. All exhibit two repeats at 74 and 78 base pairs. Lanes 1, 6, and 15 are the DNA marker (MspI digest of pBR322 DNA). Lanes 4, 9, and 13 are FXD/FXE amplified samples which were digested with MNLI. Lanes 3, 8, and 12 are FXF/FXR samples which were not digested with MnI. Lanes 5, 10, and 14 are FXD/FXE amplified samples which were not digested with MnI.
The procedure did work well for identifying any large CGG repeat regions which was one of the primary goals of this study. The smaller fragments often were not visible on the test gel, but this was not necessary. The important factor when looking at allele stability is the presence of any large pure repeat regions. This procedure did allow the identification of large regions of pure CGG repeats. Specifically, it demonstrated a significant difference in pure repeat numbers between the two subjects (1 & 2) studied, who were from the same family. Both had alleles in the "premutation" range. One of the individuals had approximately 54 pure CGG repeats in an allele of approximately 64 total repeats while the other had a much smaller pure CGG repeat number around 30, in an allele of approximately 50 total repeats. This difference in the number of pure repeats may help explain why the second allele was stable while the first expanded on transmission.

A female sample was also used in this study. Analysis of her AGG interspersion pattern proved to be much more difficult than in the male samples. Females have two X chromosomes. This can greatly increase the number of fragments. It also makes it hard to identify which chromosome the individual fragments came from. However, it is still possible to identify large regions of pure CGG repeats in females even if the AGG pattern can't be identified and this is critical to determining the
propensity for expansion. However, the AGG interspersion pattern was not able to be extrapolated from the data.

Overall, the MnlI digestion appeared to work as was hoped. The enzyme appeared to recognize the AGG interspersion and MnlI digestion was complete. The problem was that the vertical agarose gels used did not allow identification of all the expected products. However, many of the larger fragments were identifiable. These large fragments allowed us to identify any large regions of pure CGG repeats. This was important because large pure CGG repeats may have an important role in determining the risk of hyperexpansion. Specifically, this procedure identified a large difference in the number of pure CGG repeats between the two samples, from the same family, which demonstrated different stabilities. Identification of an allele with a large number of pure CGG repeats may help explain why this allele was unstable. This technique could be improved if a way to identify these smaller fragments was found. It may be possible to do this by improving the amplification process. Although, it is more likely that the problem could be solved by using a more sensitive identification technique or DNA sequencing.

Based on the results of this study it seems apparent that PCR will be used as a primary diagnostic tool in fragile X testing. PCR should greatly reduce the number of individuals sent on for further testing. Analysis of AGG interspersion may also play a vital role in the future of
fragile X diagnosis. Research in this area should allow more confident prediction of the potential for expansion of "gray-zone" alleles and small premutations (40-60 repeats). It may become possible to define a range of pure CGG repeats at which alleles become unstable. This is why AGG interspersion analysis, either indirectly by restriction enzyme digestion or by direct sequencing of the region itself, may become part of the standard analysis of CGG repeat alleles in the gray-zone and premutation range. This will improve genetic counseling for families about the risk that these types of alleles will expand to a full fragile X mutation in future offspring.
Literature Cited


