Short Report

Allele distribution at the FMR1 locus in the general Chinese population


Fragile X syndrome is an important disease of hereditary mental retardation. Its prevalence in the Chinese population is not clear. We amplified FMR1 CGG repeats from male newborns' blood spots. Approximately 45% of the males had 28 CGG repeats and another 19% had 29 repeats. Besides this major peak, there was a second peak at 34 and 35 repeats. From the 1000 males studied, 3 were found to have repeat numbers in the high borderline range (each with 50, 52 and 53 repeats). This result provides a low but significant risk of fragile X syndrome in the Chinese population.

Fragile X syndrome is one of the most important forms of inherited mental retardation (1). Its prevalence has been estimated to be one in 4000 for males and one in 8000 for females (2, 3). Fragile X syndrome is associated with a fragile site located at the end of X chromosome (Xq27.3). The mutation responsible for fragile X syndrome involves an expansion of a CGG repeat sequence at the 5' untranslated region of the FMR1 gene (4, 5).

The prevalence of fragile X syndrome in the Chinese population is not clear. The population in Taiwan comes from different areas of China. As a medical center in Taiwan, we have encountered only a few cases of fragile X syndrome in the genetic clinic. In a previous study on the variation of the CGG repeat in the FMR1 gene in normal Southern Chinese subjects by Chen et al. (6), the repeat size assessed was smaller when compared with data collected in the USA. The largest repeat number detected in that study was only 37, compared with 54 in populations in North America (7, 8).

In this paper, we analyzed the distribution of FMR1 CGG repeat size in the general Chinese population in Taiwan. We used newborn screen filters, since they were non-biased samples. Owing to the recent advances in polymerase chain reaction (PCR) technique, FMR1 CGG repeats with sizes in the premutation range could be easily amplified (9).

Materials and methods

Samples and patients

The newborn screen center in the National Taiwan University Hospital screens half of the newborns all over the island. Consecutive, non-selected filters from male newborns were used in this study. The FRAXA locus is located on the X chromosome, so males have one copy, but females have two copies of the gene. Since large expansion of CGG repeat gives no PCR product, it may be overlooked in females. In order not to miss any allele, only filters from males were used.

Nine fragile X syndrome patients, one lymphoblast cell line from a fragile X patient (GM3200, ATCC), and 2 females with premutations (81 and 69 repeats) were used as positive controls. Peripheral blood was obtained from the patients and their family members after informed consent.
Allele distribution at the FMR1 locus

Extract DNA from filter paper blood spots

Newborn screen filters were stored at −20°C. A paper disc, 3 mm in diameter, was punched from one blood spot for each filter. The paper disc was incubated in 1 ml red blood cell (RBC) lysis buffer (0.32 M sucrose, 10 mM Tris–HCl, pH 7.4, 5 mM MgCl₂, 1% Triton X-100) for 15 min at room temperate with shaking in a 1.5-ml tube. The supernatant was decanted after centrifugation at 100 g for 1 min. This RBC lysis procedure was repeated three times, then 50 µl of 1.5 × PCR buffer (0.68% Tween 20, 0.68% Nonidet P40, 100 µg/ml proteinase K, 40 mM Tris–HCl, pH 8.3, 3.75 mM MgCl₂) was added. The tubes were vortexed for 1 min, incubated at 52°C for 1 h, vortexed again for 1 min, boiled for 10 min, and cooled on ice. Supernatant was retrieved and stored at −20°C after centrifugation at 50 g for 5 min.

PCR and sequencing

PCR reaction was performed with exo−pfu polymerase (Strategen, CA, USA) on a Perkin-Elmer DNA Thermal Cycler (9). A 25-µl reaction mixture contains 100 µg of each of the primers, 5 µl of DNA, 2.5 µl of 10 × buffer supplied by the manufacturer, 100 µM each of dNTPs (half of the dGTP was replaced by 7-deaza-dGTP), 3.75 µl of dimethylsulfoxide (DMSO) and 1.25 U of exo−pfu. Temperature cycles were 98°C for 5 min, followed by 35 cycles of 65°C for 3 min and 98°C for 1 min, and a final cycle of 65°C for 3 min and 75°C for 10 min. PCR products were analyzed on a 1.5% agarose gel.

Samples were analyzed for exact CGG repeat numbers by polyacrylamide gel analysis. For this purpose, 0.05 µl of α³²P-dCTP (10 µCi/ml) was added to the PCR mixture. Isotope-labeled PCR products were analyzed by a 5% denaturing polyacrylamide gel with 7 M urea. Polyacrylamide gels containing 7 M urea and 40% formamide were also used (10). Sequencing reactions were performed with a CircumVent Thermal Cycle Dideoxy DNA Sequencing Kit (BioLabs, MA, USA). The size marker was pUC19 DNA sequenced from control 24-mer DNA sequencing primer (BioLabs, USA). Images were analyzed by a Phosphoimager (Molecular Dynamics, CA, USA). PE5.1 was obtained from Dr Ben A. Oostra.

Statistical analysis

Statistical analyses were carried out using the Microsoft EXCEL software. The distribution of CGG repeats was presented as mean ± SD. Confidence interval (CI) was calculated for the percentage of prevalence. The convergence of CGG repeat allele frequency was analyzed by Poisson test, and a probability was calculated to determine the differences between each group.

Results

In order to analyze FMR1 CGG repeat size distribution in the Chinese population, 1000 consecutive and non-selected filters from male newborns were obtained from the newborn screen center in the National Taiwan University Hospital. DNA was extracted from the blood spots, and FMR1 CGG repeats were amplified. The results revealed that all samples gave visible PCR products by agarose gel electrophoresis and ethidium bromide staining (Fig. 1A). No FMR1 full mutation was found in these 1000 samples. On the contrary, PCR for fragile X patients or lymphoblast cell line GM3200, either from blood spots or from purified genomic DNA, up to 1 µg per reaction, gave no visible product. PCR for females with premutations (81 and 69 repeats) showed very faint signals on top of the normal PCR products (picture not shown).

PCR products of the first 100 samples and samples that showed larger repeat size in agarose gel analysis were subjected to polyacrylamide gel analysis to determine the numbers of CGG repeats

![Fig. 1. Analysis of FMR1 CGG repeat size in male newborns.](image1)

A) Agarose gel analysis of repeat size which shows one sample (lane 7) with a larger repeat. B) Denaturing polyacrylamide gel analysis of CGG repeat size for 12 consecutive cases. All cases except one (lane 5) in the picture show either 28 or 29 repeats. C) Samples with either 43 (lane 4), 44 (lane 8), or 53 (lane 9) CGG repeats.
Distribution of $FMR_1$ CGG repeat size in normal male newborns. One hundred consecutive cases were studied by denaturing polyacrylamide gel analysis. The chart has been further refined by screening 1000 cases for larger repeats. (Fig. 1B). Both sequencing ladders and the PCR product of pE5.1 (containing 16 CGG repeats, as confirmed by direct sequencing) were run together with the samples. Among the first 100 samples, 45 had 28 CGG repeats and 19 had 29 repeats (mean 29.87 repeats, SD = 4.00) (Fig. 1B), representing 64% (95% CI = 59.2–68.8%) of all cases (Fig. 2). Besides this major peak, there was a second peak with 34 and 35 repeats (Fig. 2). The PCR product of pE5.1 and those with 28 repeats were further analyzed by gels with higher denaturing power (picture not shown). This gel contained 7 M urea and 40% formamide, in order to resolve aberrant motions of repeated sequences (10). CGG repeat numbers calculated from this formamide-containing gel were the same as those from a regular sequencing gel. Direct sequencing of 3 individuals with 28 CGG repeats all showed an interruption by an AGG after nine CGG repeats (picture not shown).

Among those samples who showed larger repeats in agarose gel analysis, 15 had repeat numbers equal to or larger than 40. Four of them had 40 repeats, two each had either 41, 42, or 44 repeats, and one each had 43, 47, 50, 51 or 53 repeats (Fig. 1C). Therefore, the prevalence for repeat size in the high borderline range (50–60 repeats) was three in 1000 (95% CI = 0.3–6.3 in 1000). These data have been used to refine the distribution chart in Fig. 2.

**Discussion**

In this study, 1000 Chinese males were studied for $FMR_1$ CGG repeat size by using newborn screen filters. Around two-thirds of the population in Taiwan comes from South-eastern China, and one-third from other parts of China. Because our newborn screen center received samples all over the island, this data should currently be the most representative for the Chinese population.

In a previous study on the variation of the CGG repeat in $FMR_1$ gene in normal Southern Chinese subjects by Chen et al. (6), the largest repeat number detected was only 37, which was smaller than the data collected in the USA. In the present study, at least 15 alleles were found to have repeat numbers equal to or larger than 40. Compared with the study of Fu et al. (7), both the size of the largest allele (53 repeats in this study and 54 in the study of Fu et al.) and the shape of distribution curves are quite similar. However, the distribution of $FMR_1$ CGG repeat size was more convergent in our study. The most common allele (28 repeats) represented 45% of all samples in this study, as compared with around 30% (deduced from their figure) for 29 repeats in the study of Fu et al. ($p < 0.001$). Statistical analyses for the comparisons of the sums of the most common two or three alleles of the major peak both showed significant differences ($p < 0.01$ and $p < 0.02$, respectively) (7). Therefore, this is compatible with our clinical experience that fragile X syndrome is rare, but still found in Taiwan.

Most early data for the prevalence of fragile X syndrome came from studies for the mentally retarded by chromosome analysis. The cut-off value for fragility in chromosome analysis is arbitrary, and fragile sites other than $FRAXA$ are present on the end of Xq. After the cloning of the $FMR_1$ gene, DNA diagnosis was found to be more reliable than chromosome analysis. The prevalence of fragile X syndrome was now estimated to be one in 4000 for males and one in 8000 for females (2, 3). It has also been claimed that there may be a decline in prevalence due to the effect of genetic counseling (11).

Although the prevalence of fragile X syndrome may be not as high as we previously believed, studies on carrier frequency of $FMR_1$ premutation in females were intriguing. Rousseau et al. (12) found a premutation frequency of 1/259 in females, and the study of Spence et al. (13) showed similar results. Doctors have started to examine pregnant women for carrier status to prevent the occurrence of fragile X syndrome (13). Since the examination of females who have alleles with similar sizes (e.g., 45% of the Chinese population have 28 repeats and another 19% have 29 repeats) requires Southern blot analysis, this approach could not be widely adopted.

In this paper, we demonstrated that the analysis of males for fragile X syndrome employing newborn screen filter is feasible. PCR products with more than 40 CGG repeats could be determined.
Furthermore, CGG repeats in the premutation range gave signals that both increased in size and decreased in intensity. Carrier mothers and other family members at risk of fragile X syndrome could be identified. This information would be potentially of importance to particular families. Further procedures improving DNA extraction are necessary prior to implementing broad-scale population screening.

Acknowledgements
This work was supported by a grant from National Taiwan University Hospital No. NTUH S861503-A37.

References