ORIGINAL ARTICLE

Molecular analysis of fragile X syndrome (FXS) among Malaysian patients with developmental disability

Ernie Zuraida ALI MSc, Yusnita YAKOB MSc, Norsiah MD DESA MSc, Taufik ISHAK BSc, Zubaidah ZAKARIA* MBBS, DCP, Lock-Hock NGU** MBBS, MRCP and Wee-Teik KENG** MBBS, MRCP

Molecular Diagnostics and Protein Unit, Specialised Diagnostics Centre, Institute for Medical Research, *Cancer Research Centre, Institute for Medical Research, **Medical Genetic Department, Kuala Lumpur Hospital, Kuala Lumpur, Malaysia

Abstract

Fragile X syndrome (FXS) is a neurodevelopmental disorder commonly found worldwide, caused by the silencing of fragile X mental retardation 1 (FMR1) gene on the X-chromosome. Most of the patients lost FMR1 function due to an expansion of cytosine-guanine-guanine (CGG) repeat at the 5' untranslated region (5'UTR) of the gene. The purpose of this study is to identify the prevalence of FXS and characterize the FMR1 gene CGG repeats distribution among children with developmental disability in Malaysia. Genomic DNA of 2201 samples from different ethnicities (Malays, Chinese, Indian and others) of both genders were PCR-amplified from peripheral blood leukocytes based on specific primers at 5’UTR of FMR1 gene. Full mutations and mosaics were successfully identified by triple methylation specific PCR (ms-PCR) and subsequently verified with FragilEase kit. The findings revealed for the first time the prevalence of FXS full mutation in children with developmental disability in Malaysia was 3.5%, a slightly higher figure as compared to other countries. Molecular investigation also identified 0.2% and 0.4% probands have permutation and intermediate alleles, respectively. The CGG repeats length observation showed 95% of patients had normal alleles within 11 to 44 CGG repeats; with 29 repeats found most common among Malays and Indians while 28 repeats were most common among Chinese. In conclusion, this is the first report of prevalence and characterisation of CGG repeats that reflects genetic variability among Malaysian ethnic grouping.

Keywords: fragile X syndrome, FXS, CGG repeat, FMR1 gene, prevalence

INTRODUCTION

Fragile X syndrome (FXS, OMIM 300624) is caused by the expansion of CGG repeat in the 5' untranslated region (UTR) of the fragile X mental retardation 1 (FMR1) gene (OMIM 309550), located at the fragile site (FRAXA) on the X chromosome(Xq27.3). FXS is the most common genetic cause of intellectual disability (ID), and one of the leading genetic causes of autism. FXS is characterised by a broad spectrum of behavioural and emotional impairment, psychological problem, learning disabilities and physical abnormalities such as large or prominent ears, long and narrow face, joint laxity, macrocephaly and macroorchidism after puberty.

The CGG repeats are highly polymorphic and can be divided into four categories based on the repeat length: normal (6-44 repeats), intermediate (45-54 repeats), premutation (55-200 repeats) and full mutation (>200 repeats). Normal alleles are usually transmitted in a stable manner from parents to the child whereas intermediate alleles may or may not reveal intergenerational instability. Intermediate alleles are considered a possible risk factor for the CGG repeat expansion. Premutation alleles do not cause FXS phenotype but are unstable and may expand to full mutation in subsequent generations. Expansion to a full mutation arises when the premutation alleles are transmitted from the mothers, while fathers who are carriers of pre-mutation alleles will transmit the premutation stably to their daughters. Full mutation alleles of more than 200 repeats cause hypermethylation of the CpG island within the FMR1 promoter.
region, leading to transcriptional silencing of FMR1 gene and loss of the fragile X mental retardation protein (FMRP). This is believed to cause the FXS phenotype. Nevertheless, methylation of full mutation expansions (>200 CGG), can be incomplete, and may be associated with methylation mosaicism. Generally, FXS phenotype presents in male and less frequent in females. However, 50% of females with full mutation will have FXS symptoms, probably because of random X-inactivation.

Molecular analysis of FXS has been studied primarily in Caucasian populations of North America and Europe with prevalence in general Caucasian population ranges from 1:3717 to 1:8919. The frequency of full mutation in male Caucasian with intellectual disability ranges from 2.6% to 8.7%. FXS is ubiquitous and several molecular analyses have been performed in the North Asia populations of Japan, China, Hong Kong and Taiwan, Southeast Asia populations of Indonesia and Thailand; and South Asia populations of India and Pakistan. Results from previous studies suggested that the prevalence of FXS may in fact differ across the populations.

Malaysia has a multiethnic population. About 67.4% of its population is Malay, 24.6% Chinese, 7.3% Indian and 0.7% others. To date, FMR1 CGG repeat length has been studied in 31 normal Kadazans, which is one of the small ethnic groups in Malaysia. The purpose of this study is to determine the prevalence of FXS among Malaysians with developmental disability and to illustrate the distribution of CGG repeat length among various ethnic groups in Malaysia.

**MATERIALS AND METHODS**

**Samples**

From 2004 to 2013, we received 2201 (2057 males and 144 females) peripheral blood samples with written consent. These samples came from all individuals who have developmental disability with or without physical features or behavioural problems suggestive of FXS. Samples from mothers and affected siblings of FXS children were also included. The subjects were mostly paediatric patients but their ages ranged from 1 month to 20 years old. About 2.5 ml of blood sample was collected into EDTA tubes from each subject. Genomic DNA from peripheral blood was extracted either manually or by the semi-automated system (QIAcube) (Qiagen, Valencia, USA), using QIAamp DNA Blood Mini Kit (Qiagen).

**Polymerase chain reaction (PCR) and cycle sequencing**

The PCR amplification of CGG repeat region of the FMR1 gene was performed with specific primers: FRAXAF (forward) 5’ AGG CGC TCA GCT CGT TTT CGG TTT CAC TTC 3’ and FRAXAR (reverse) 5’ AGC CCC GCA CTT CCA CCA CCA GCT CCT CCA 3’. Briefly, a total volume of 35 µl containing 200 ng of genomic DNA, 0.2 mM of each of dATP, dTTP, dCTP, dGTP and 7-Deaza-2-dGTP; 0.6 µM of primers FRAXAF and FRAXAR, 1x of 10x ViBufferS (Vivantis, CA, USA), 8% DMSO, 1.7 M betaine and 2.5 U of Pfu DNA polymerase (5 U/µl; Vivantis) was pipetted into a PCR tube. The PCR reaction was carried out in ABI thermal cycler (Applied Biosystems, Foster City, USA), with initial denaturing at 94°C for 1 min followed by 35 cycles of denaturation at 94°C for 1 min, annealing at 65°C for 1 min and extension at 72°C for 2 min and final extension at 72°C for 10 min. The PCR product was sized on a 1.8% agarose gel electrophoresis in 1X TBE buffer at 80V, 60 min and visualised in gel documentation system (Syngene GeneGenius Bio Imaging System, Cambridge, UK). PCR product of female and male samples with band of large sizes (>450 bp) was subjected to a re-run in micro-capillary electrophoresis chip to determine the number of CGG repeats.

The PCR product of small size band (<450 bp) was then purified by QIAquick PCR purification kit (Qiagen). Cycle sequencing was performed according to standard protocol at half strength of 2.5X Big Dye V3.1. The cycle sequencing product was purified using DyeX2.0 kit (Qiagen). This product was then loaded onto Avant 3130 Genetic Analyser (Applied Biosystems) for sequencing and the sequences obtained were analysed by SeqScape software v2.0 and v2.6 (Applied Biosystems) to determine the number of CGG repeats.

DNA samples that failed to be amplified with conventional PCR were analysed by triple methylation specific PCR (ms-PCR) as previously described by Zhou et al. Mosaicism cases could be detected as well by using ms-PCR. The FragilEase Kit (PerkinElmer, Turku, Finland) was used to verify results identified in ms-PCR.

**Sodium bisulphite treatment and triple methylation specific PCR (ms-PCR)**

For bisulfite conversion, 1000 ng of genomic DNA was treated with the EpiTect® Bisulfite
Kit (Qiagen, Hilden, Germany). The bisulfite-converted DNA was eluted from the column in 20 µL of elution EB buffer. Three sets of primer were used to amplify modified DNA by using triplet ms-PCR, developed by Zhou et al.14 One set of primer was used to target the non-methylated allele (non-Met-PCR) while the other two sets were for targeting the methylated allele (Met-PCR and mTP-PCR). The PCR reaction was carried out in a thermocycler (Eppendorf, Hamburg, Germany). The PCR amplification conditions and analyses followed the approach described by Zhou et al.14

Verification of full mutation and mosaic cases
For verification of ms-PCR results, a commercial FragilEase kit (PerkinElmer) was used. All samples were PCR amplified according to the manufacturer’s protocol and subsequently run in the micro-capillary electrophoresis chip by using Agilent DNA 7500 kit (PN 5067-1506) to determine the size of the bands. Fragile X reference samples (two female samples) with well-known different repeats sizes were included in each run as quality control. FraxSoft software was used to construe result for all samples.

RESULTS

Ethnic distribution and frequency of fragile X syndrome cases
A total of 2201 samples were analysed. 2057 (93.5%) samples were males and 144 (6.5%) samples were females. 57% of the samples were from Malays (1169 males and 76 females), 30% from Chinese (609 males and 50 females), 10% from Indians (198 males and 14 females) and 4% from other ethnicities (81 males and 4 females). Out of 144 females, 17 (11.8%) were mothers of affected males while 127 (88.2%) were females with developmental disability. We have excluded 17 mothers of affected patients, thus the sample size of 2184 was used for subsequent data analyses.

PCR method was successful in 2108 (96.5%) DNA samples demonstrating normal, intermediate and premutation alleles. 73 male samples failed to amplify probably due to the high number of CGG repeats and 3 female samples showed only single allele. Of 2108 samples that were successfully amplified, 1667 (79.1%) male samples have been sequenced and the remaining 441 (20.9%) male and female samples with large size bands have been analysed by micro-capillary electrophoresis chip.

In summary as shown in Table 1, 2096 (95.9%) samples were found to have normal alleles (ranging 11-44 CGG triplet repeats), 8 (0.4%) have intermediate alleles (ranging 47-53 CGG triplet repeats) and 4 (0.2%) have premutation alleles (ranging 56-150 CGG triplet repeats). Ms-PCR results showed that 3 (0.1%) samples were mosaics of premutation/full mutation and 73 (3.3%) have a full mutation. Among full mutation, 70 (3.6%) samples were found in males while 3 (2.4%) out of 127 samples were found in females (Table 1). Subsequently, we used FragilEase kit to verify all mosaicism and full mutation samples from ms-PCR. The results showed that FragilEase kit has successfully produced similar findings as in ms-PCR but with better genotype identification especially for mosaic samples. Example of result from FragilEase kit is shown in Figure 1.

<table>
<thead>
<tr>
<th>Allele class (CGG repeats range)</th>
<th>Number of samples (Total)</th>
<th>Number of samples (Male)</th>
<th>Number of samples (Female)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal (6-44 CGG)</td>
<td>2096</td>
<td>1974</td>
<td>122</td>
<td></td>
</tr>
<tr>
<td>Intermediate (45-54 CGG)</td>
<td>8</td>
<td>8</td>
<td>-</td>
<td>1/273</td>
</tr>
<tr>
<td>Premutation (55-200 CGG)</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>1/546</td>
</tr>
<tr>
<td>Full mutation (&gt;200 CGG)</td>
<td>73</td>
<td>70</td>
<td>3</td>
<td>1/30</td>
</tr>
<tr>
<td>Mosaic premutation/full mutation</td>
<td>3</td>
<td>3</td>
<td>-</td>
<td>1/728</td>
</tr>
<tr>
<td>Total</td>
<td>2184</td>
<td>2057</td>
<td>127</td>
<td></td>
</tr>
</tbody>
</table>
FIG. 1: FragilEase kit analysis demonstrating full mutation and mosaic male by using micro-capillary electrophoresis. Lane 1, DNA 7500 ladder; 2, blank; 3, reference sample (20 and 218 CGG repeats); 4, reference sample (30 and 371 CGG repeats); 5, normal male (31 CGG repeats); 6, normal female (29 and 35 CGG repeats); 7, full mutation male (537 CGG repeats); 8, no sample; 9, premutation female (27 and 147 CGG repeats); 10, mosaic male with premutation and full mutation (~110 to 350); 11, full mutation male with size mosaicism (~400 to 600 CGG repeats) and 12, full mutation male with size mosaicism (~230 to 500 CGG repeats)

Distribution of full mutation, premutation, intermediate alleles among ethnic groups

The distribution of the FXS cases among major ethnic groups in Malaysia is displayed in Figure 2. Methylation specific PCR and FragilEase analyses showed that 76 (3.5%) samples harboured full mutation fragile X; 40 (3.2%) were Malays, 27 (4.1%) were Chinese, 8 (3.8%) were Indians and 1 (1.2%) was a male subject from a minority ethnic group. Among Malays, 39 (3.3%) of male samples and 1 (1.5%) of female samples were full mutation. For Chinese, 25 (4.1%) of males and 2 (4.7%) female samples were full mutation. Finally for Indians, 8 (4.0%) of male samples were full mutation. 3 (3.9%) were identified as mosaic premutation/full mutation.

PCR analysis identified 19 female premutation carriers, of which 12 were Malays (2 were probands; 10 were mothers of the FXS full mutation patients) and the remaining 7 were Chinese mothers. Only 2 males had premutation alleles and both were Malays. The intermediate allele was detected in 8 male samples; 5 were Malays and 3 were Indians.

Variation of normal and intermediate CGG repeats allele among major ethnic groups

Normal alleles in our study population ranged from 11 to 53 CGG repeats. Figure 3 shows the distribution of normal allele CGG repeat among different ethnic groups and between male and female. Among Malays, the CGG repeats ranged from 11 to 53 repeats while the range among Chinese and Indian were 15-43 and 18-49 respectively (Fig. 3a). Among females, the normal alleles CGG repeats ranged from 18-41 (Fig. 3b).

The most frequent allele size among Malays and Indians was 29 CGG repeats (29.8% and 34.4% respectively) while 28 CGG repeats (26.0% and 17.9% respectively) was the next frequent repeat size. The reverse was found among Chinese where the 28 CGG repeats (34.0%) was the most frequent allele, followed by 29 CGG repeats (26.7%). 30 CGG repeats was the third most common allele in our study among all major ethnic groups.

DISCUSSION

For the first time, we report the prevalence of FXS cases among children of different ethnic groups with developmental disability in Malaysia. The
overall prevalence estimation of full mutation cases in our study was 3.5%, in which 3.6% was in males and 2.4% was in female. Such results are compatible with the fact that males are more affected by FXS as compared to females.\footnote{9}

This report demonstrates that FXS is present in 3.24%, 4.14% and 3.77% of people with developmental disability among Malays, Chinese and Indian, respectively. However, this assessment was a minimal estimation of the true prevalence in Malaysia as not all hospitals sent their samples even though the testing is offered to all hospitals. These cases were non-selective and testing was requested by primary care doctors and pediatricians. Patients included in the analysis could well have other diagnoses such as chromosomal abnormality or epilepsy syndromes that were unknown to the authors. Nevertheless, these data should be a good start for providing primary insights of actual FXS cases.

Meanwhile, in previous studies they have used different approaches to investigate the FMR1 triplet repeat distributions among different target populations such as children with mental retardation (MR), learning disability (LD), attention deficit/hyperactivity disorder (ADHD), attention deficit (AD), children clinically referred with MR and also children in special education or special school. Therefore, we have compared our findings with those adopting similar recruitment approaches as ours i.e. clinical referral for individuals with intellectual disability. Our study has shown that the FXS prevalence was higher than previously reported in Japan (0.8\%),\footnote{15} Indonesia (1.9\%)\footnote{4} and Turkey (3.0\%)\footnote{16}. However, our figure was lower than previously described in other countries such as India (5.3\%),\footnote{17} Thailand (5.3\%),\footnote{18} Netherlands (5.1\%),\footnote{19} Finland (4.9\%),\footnote{20} Mexico (3.8\%)\footnote{21} and USA (3.7\%).\footnote{22} Our prevalence seems lower when compared to the western countries, and that could be probably due to the sample sizes used. The western studies worked on smaller sample size than ours, which may have affected the data generated.

Study of the normal CGG repeats distribution is important in order to determine the repeat variability among ethnics in Malaysian population. Allelic diversity of FMR1 gene in Malaysian population showed the CGG repeat allele frequencies were different among

![FIG. 2: Distribution of FXS cases in different categories and between ethics. Abbreviation: FMM: full mutation male; FMF: full mutation female; PMM: premutation male; PMF: premutation female; IM: intermediate](image)
FIG. 3: Distribution of the normal and intermediate CGG repeats in FMR1 gene between ethnics/races and genders. A) Allele size for male and B) Allele size for female

ethnicities as reported from previous studies (Table 2). Nevertheless, alleles with 29 and 30 repeats were among the commonest in Malaysia, similar to many populations in other parts of the world. Previous study by Kunst et al (1996) found 29 and 30 repeats were equally common among Kadazans, who belong to the ethnic minority found mainly in East Malaysian particularly in Sabah, while 28 repeat allele was not detected at all since this was only a small study. Malaysian Indian showed similar allele distribution to India and Pakistan. However, the Malaysian Chinese had allele distribution pattern that was different from mainland China and Taiwan or even the neighbouring Singaporean Chinese. The Malay also showed different allele distribution especially for second and third common CGG repeat size comparing to Singapore Malays and Indonesian. These differences could be accounted partly by
intermarriages among different ethnics. This was supported by previous population genetic structure studies among Malays, which showed that there were distinct subgroups among Malays with some subgroups clustered more closely with Indonesian ethnic groups and Chinese while other did not.23

**Conclusion**

The present study has shown for the first time the prevalence of FMR1 full mutation among patients of different ethnic groups with developmental disability in Malaysia. We also have successfully identified the permutation and intermediate cases as well as characterised the lengths of CGG repeats in normal patients among ethnicity. Our study support the finding that 29 and 30 CGG repeats are the commonest repeats in the Malaysia population. Both Malay and Chinese showed genetic variability compared to similar ethnics in the region.

### COMPETING INTERESTS

The authors declare that they have no competing interests.

### ACKNOWLEDGEMENTS

We thank the Director General of Health Malaysia for permission to publish this paper. Our special thanks to Director of IMR and Head of SDC for critical reading of the manuscripts and valuable comments. We would like to extend our gratitude to Rosmalaili Kassim, Norhayati Musa, Mohd Heikkal Ismail and Nurul Farahana Rosli for their excellent technical assistance. We are also grateful to all clinicians for their contributions in this study. This study was funded by National Institute of Health (NIH), Malaysia (JPP-IMR 06-043).

### REFERENCES