Simultaneous Screening of the FRAXA and FRAXE Loci for Rapid Detection of FMR1 CGG and/or AFF2 CCG Repeat Expansions by Triplet-Primed PCR


From the Department of Paediatrics,* the Department of Biochemistry,† and the Department of Obstetrics and Gynaecology,** Yong Loo Lin School of Medicine, National University of Singapore, Singapore; the Khoo Teck Puat—National University Children’s Medical Institute,† National University Health System, Singapore; the Department of Molecular and Human Genetics,‡ Baylor College of Medicine, Houston, Texas; the Adelaide Medical School and Robinson Research Institute,§ The University of Adelaide, Adelaide, South Australia, Australia; the Department of Paediatric Medicine,¶ KK Women’s and Children’s Hospital, Singapore; the Cancer and Stem Cell Biology Program,## Duke—National University of Singapore Graduate Medical School, Singapore; the Division of Cellular and Molecular Research,### National Cancer Center Singapore, Singapore; the South Australian Health and Medical Research Institute,#### Adelaide, South Australia, Australia; and the Department of Laboratory Medicine,##### National University Hospital, Singapore

Accepted for publication April 29, 2021.
Address correspondence to Samuel S. Chong, Ph.D., Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore; the Khoo Teck Puat—National University Children’s Medical Institute, Singapore; the Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas; the Adelaide Medical School and Robinson Research Institute, The University of Adelaide, Adelaide, South Australia, Australia; the Department of Paediatric Medicine, KK Women’s and Children’s Hospital, Singapore; the Cancer and Stem Cell Biology Program, Duke—National University of Singapore Graduate Medical School, Singapore; the Division of Cellular and Molecular Research, National Cancer Center Singapore, Singapore; the South Australian Health and Medical Research Institute, Adelaide, South Australia, Australia; and the Department of Laboratory Medicine, National University Hospital, Singapore

Supported in part by Khoo Teck Puat—National University Children’s Medical Institute grant N-178-000-063-001 (F.S.W.). T.L. was an intern under the National University of Singapore’s Science Research Program for high school students.
T.L. and F.S.W. contributed equally to this work.
Disclosures: The authors submitted a patent application on June 2, 2021, directed to a method of detecting a repeat expansion sequence.
Current address of T.L., Human Genetics Program, Wellcome Sanger Institute, and School of Clinical Medicine, University of Cambridge, Cambridge, United Kingdom.

The FRAAXE folate-sensitive fragile site on the X chromosome has been associated with several medical conditions, including intellectual disability, obsessive–compulsive disorder, and primary ovarian insufficiency (POI). FRAAXE fragility is caused by hyper-expansion of CCG trinucleotide repeats within the 5′ untranslated region in exon 1 of AFF2 (alias FMR2) (ClinVar: VCV000010526.1). Hyper-expansion from a normal allele range of 6 to 30 repeats to >200 repeats is accompanied by CpG methylation

Moderate to hyper-expansion of trinucleotide repeats at the FRAXA and FRAAXE fragile sites, with or without concurrent hypermethylation, has been associated with intellectual disability and other conditions. Unlike molecular diagnosis of FMR1 CGG repeat expansions in FRAXA, current detection of AFF2 CCG repeat expansions in FRAAXE relies on low-throughput and otherwise inefficient techniques combining Southern blot analysis and PCR. A novel triplet-primed PCR assay was developed for simultaneous screening for trinucleotide repeat expansions at the FRAXA and FRAAXE fragile sites, and was validated using archived clinical samples of known FMR1 and AFF2 genotypes. Population samples and FRAAXE-affected samples were sequenced for the evaluation of variations in the AFF2 CCG repeat structure. The duplex assay accurately identified expansions at the FMR1 and AFF2 trinucleotide repeat loci. On Sanger sequencing of the AFF2 CCG repeat, the single-nucleotide polymorphism variant rs868914124(C) that effectively adds two CCG repeats at the 5′-end, was enriched in the Malay population and with short repeats (<11 CGGs), and was present in all six expanded AFF2 alleles of this study. All expanded AFF2 alleles contained multiple non-CCG interruptions toward the 5′-end of the repeat. A sensitive, robust, and rapid assay has been developed for the simultaneous detection of expansion mutations at the FMR1 and AFF2 trinucleotide repeat loci, simplifying screening for FRAXA- and FRAAXE-associated disorders. (J Mol Diagn 2021, 23: 941–951; https://doi.org/10.1016/j.jmoldx.2021.04.015)
of the repeat. This CpG methylation in turn silences expression of AF4/FMR2 family member 2 (AFF2), a subunit of SEC-L2, which regulates the transcription of several genes. AFF2 CCG repeat hyper-expansion is the genetic mutation that gives fragile XE nonsyndromic intellectual disability (FRAXE NSID; Online Mendelian Inheritance in Man no. 309548; http://www.ncbi.nlm.nih.gov/omim), a mild (IQ 50 to 70) to borderline (IQ 70 to 85) intellectual disability affecting an estimated 1 in 50,000 to 100,000 males, and also other cognitive–behavioral abnormalities, including obsessive–compulsive disorder. Paradoxically, alleles with <11 repeats or with microdeletions within or near the repeats are enriched in patients with POI. Studies associating intermediate (31 to 60) repeat sizes and Parkinson disease, a neurodegenerative motor system disorder, have been inconclusive.

Interestingly, FRAXE shares genetic features with the well-studied fragile site on the X chromosome, FRAXA, which has more clearly demonstrated clinical involvement. The FRAXA site contains CGG repeats within the 5′ untranslated region in exon 1 of the FMR1 gene. Hyper-expansion of FMR1 CCG repeats to >200 is accompanied by CpG methylation and FMR1 gene silencing, which results in fragile X syndrome (FXS; Online Mendelian Inheritance of Man no. 309550; http://www.ncbi.nlm.nih.gov/omim) (ClinVar: VCV00000972.1). FXS is the most common inherited monogenic cause of intellectual disability, affecting approximately 1 in 5000 males and 1 in 4000 to 8000 females. Premutation alleles are associated with many behavioral features, including anxiety, obsessive–compulsive traits, and depression. In addition, approximately one fifth of females who carry a premutation allele (55 to 200 CGG repeats) are affected with FX-related POI. FMR1 premutations have been identified in 2% of patients with sporadic POI and in 14% of patients with familial POI, making FXPOI the most common genetic cause of POI in euploid women. Furthermore, a subset of FMR1 premutation carriers eventually develop FX–associated tremor/ataxia syndrome (FXTAS), a neurodegenerative disorder with features of ataxia, tremor, and parkinsonism that affects approximately 1 in 4000 men and 1 in 7800 women over the age of 55 years. FMR1 has also been associated with Parkinson disease.

Although both FRAXA and FRAXE fragile sites have been strongly associated with intellectual disabilities, the mild to borderline phenotype of FRAXE NSID could lead to underascertainment and under-diagnosis compared to FXS. Under-diagnosis may be partly attributable to the lack of a rapid, simple, and inexpensive assay to screen for repeat expansions at the FRAXE site. Although standard PCR methods have been described for repeat expansions at the FRAXE site, either as a standalone or multiplexed assay, they cannot detect large premutation and full-mutation alleles, whose detection continues to rely on Southern blot analysis. Molecular diagnosis is crucial for confirming FRAXE NSID due to its nonsyndromic nature, that is, a lack of a distinctive and consistent clinical presentation, in contrast to, for example, facial dysmorphism (prominent ears, jaw, forehead, and long face) and macroorchidism present in patients with FXS.

Based on a previously described simplified strategy for the detection and sizing of FMR1 CCG repeat expansions in patients with FXS/FXPOI/FXTAS based on triplet-primed (TP)-PCR, an equivalent assay was developed for the detection and sizing of AFF2 CCG repeat expansions in FRAXE NSID and associated disorders using a similar TP-PCR approach. In addition, the FRAXE and FXS TP-PCR assays have been combined into a single-tube duplex TP-PCR assay to enable simultaneous screening for FXS, and FRAXE NSID and associated disorders, minimizing the additional cost of testing for FRAXE when FXS testing is performed.

Materials and Methods

Biological Samples

Genomic DNA was extracted from 408 cord blood samples from 161 Chinese, 158 Malay, and 89 Indian unrelated and anonymized male infants born at the National University Hospital (Singapore). DNA samples from another 44 Caucasian and 14 African American males from the Human Variation DNA panel (HD100CAU and HD100AA-2) were purchased from Coriell Cell Repositories (Camden, NJ). Archived and previously characterized genomic DNAs consisting of 40 normal, 17 FMR1 premutation–positive, 23 FMR1 full-mutation–positive, and six AFF2 expansion–positive samples were utilized in assay validation. The de-identified normal and FMR1 CCG repeat expansion–positive samples were obtained from the KK Women’s and Children’s Hospital (Singapore). The AFF2 CCG repeat expansion–positive samples were de-identified archival samples from Baylor College of Medicine (Houston, TX) and The University of Adelaide (Adelaide, South Australia, Australia). Four of the six AFF2 CCG repeat expansion–positive samples were related. FX0230 and FX0229 were maternal uncle and nephew, respectively, whereas DNA_25926 and DNA_3802 were father and daughter.

Duplex TP-PCR of the FMR1 and AFF2 Trinucleotide Repeats

The duplex screen for AFF2 and FMR1 triplet repeat expansions utilized a TP-PCR approach, involving four primers, fluorescein amidite (FAM)-labeled FMR1-R (5′-AGCCCGCAGCACTTCCACCACAGCTCCTCCA-3′) first described in the report from Fu et al as primer F, hexachloro-fluorescein
(HEX)-labeled AFF2-F (5'-CCATGTCGCGGCTTCTAGCTGTCAGGCTCC-3'), and shared primers TP (5'-TGCT-CTTGACCGTAAATGTGCCCCTGATAACGGCGCGG-GCGCGG-3') and Tail (5'-TGCTCTGGACCTGTGAGTGTGCGCGGCTTCTAG-3'). Each 15-μL PCR reaction contained 100 ng of genomic DNA, 2.5× Q-Solution (Qiagen), 1× PCR buffer containing 1.5 mmol/L MgCl₂ (Qiagen), 2 mmol/L dNTP at a 5:1 ratio of (Qiagen, Hilden, Germany), 0.6 mmol/L each of AFF2-F, FMR1-R, and Tail primers; 0.0006 μmol/L TP primer; and 5 U of HotStarTaq DNA polymerase (Qiagen). An initial 15-minute enzyme activation at 95°C was followed by 40 cycles of 99°C for 45 seconds, 55°C for 45 seconds, and 70°C for 4 minutes, with a 15-second increment per extension cycle, and a final extension of 10 minutes at 72°C.

**Capillary Electrophoresis**

A 2-μL aliquot of the duplex TP-PCR product was mixed with 9 μL of Hi-Di formamide and 0.5 μL of GeneScan 500 ROX dye size standard (Applied Biosystems, Foster City, CA), denatured at 95°C for 5 minutes, cooled to 4°C, and resolved in a 3130xl genetic analyzer using a 36-cm capillary filled with POP-7 polymer (Applied Biosystems). The mixture was electrokinetically injected at 1 kV for 5 seconds and electrophoresed for 40 minutes at 60°C. Analysis was performed using GeneMapper software version 4.0 (Applied Biosystems). If an expanded allele was detected after the initial capillary electrophoresis run, a second capillary electrophoresis run was performed using an electrokinetic injection at 10 kV for 5 seconds and electrophoresed for 40 minutes at 60°C.

A 1-μL aliquot of one-twentieth diluted standard PCR product was mixed with 9 μL of Hi-Di formamide and 0.3 μL of GeneScan 500 ROX dye size standard, denatured at 95°C for 5 minutes, cooled to 4°C, and resolved in a 3130xl genetic analyzer using a 36-cm capillary filled with POP-7 polymer. The mixture was electrokinetically injected at 1.2 kV for 23 seconds and electrophoresed for 20 minutes at 60°C.

Post-capillary electrophoresis analysis was performed using GeneMapper software version 4.0.

**Data Interpretation**

Primer AFF2-F anneals to the AFF2 sequence immediately upstream of its CCG repeat and generates HEX-labeled TP-PCR amplicons, whereas primer FMR1-R anneals to the FMR1 sequence immediately downstream of its CCG repeat and generates FAM-labeled TP-PCR products. Both TP-PCR reactions utilize a common TP primer and Tail primer, and their products can be analyzed separately using different fluorescence detection channels, or together. The TP primer was designed to anneal optimally to any stretch of five CCG trinucleotides on the AFF2 sense strand or the FMR1 anti-sense strand. When there is a continuous uninterrupted stretch of more than five repeats, the electropherogram will display a continuous series of peaks that are set 3 bp apart from each other. If a non-CCG interruption is present, the electropherogram will display a gap of approximately 18 bp between the fluorescence peaks, equivalent to the absence of five fluorescence peaks (Figure 1). The repeat size of an uninterrupted stretch is the total number of continuous fluorescence peaks plus 4 (the first fluorescence peak is generated from the TP primer annealing to the first five repeats). For an allele with a non-CCG interruption(s), the repeat size is the total number of continuous fluorescence peaks added to the total number of missing peaks, plus 4. In heterozygous females, one will see a drop in fluorescence intensity for peaks beyond a certain size; the repeat size of the smaller allele can be derived from the number of peaks before the drop, whereas the repeat size of the larger allele can be derived from the total number of peaks (data not shown). For the AFF2 TP-PCR reaction, the leftmost solitary peak is not generated from within the repeat but instead arises from annealing of the TP primer to a (CCG)₅-CCT sequence immediately upstream of the CCG repeat, which was defined as starting after the CTG trinucleotide (Figure 1). AFF2 allele sizing and structure are further elaborated in the Results.

**Sequencing of the AFF2 CCG Repeat**

AFF2 TP-PCR and standard/conventional PCR were performed as described in the sections on “Duplex TP-PCR of the FMR1 and AFF2 Trinucleotide Repeats” and “Standard PCR across the FMR1 and AFF2 Trinucleotide Repeats,” except that all primers were unlabeled. TP-PCR and standard/conventional PCR products were purified using AMPure beads (Agencourt Bioscience, Beverly, MA) according to the manufacturer’s instructions, and quantitated using the Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA). Each 20-μL sequencing reaction contained 10 to 50 ng of purified standard PCR or TP-PCR product, 1× BigDye Terminator Ready Reaction Mix (Applied Biosystems), 2.5× Q-Solution (Qiagen), and 3.2 pmol AFF2-F primer. An initial denaturation at 96°C for 1 minute was followed by 25 cycles of 98°C for 10 seconds, 60°C for 5 seconds, and 60°C for 4 minutes. Extension products were purified using the Oligo Clean &
Concentrator column (Zymo Research, Irvine, CA) according to the manufacturer’s instructions. Eluted purified extension products were vacuum-dried in a Savant SpeedVac Concentrator (Thermo Fisher Scientific) for 5 minutes, resuspended in 12 μL of Hi-Di formamide, and resolved in a 3130xl genetic analyzer using a 36-cm capillary filled with POP-7 polymer. The mixture was electrokinetically injected at 1.2 kV for 16 seconds and electrophoresed for 20 minutes at 60°C. Post-capillary electrophoresis analysis was performed using Sequencing Analysis software version 6.0 (Applied Biosystems).

Results

Evaluation of Duplex TP-PCR on Normal and FMR1 and AFF2 Triplet Repeat Expansion—Positive Samples

Twelve DNA samples were initially utilized to evaluate the duplex TP-PCR assay. Figures 2 and 3 present the patterns generated on TP-PCR electropherography in male or female individuals with no mutations or FMR1 and AFF2 premutations or full mutations. The leftmost peak generated using the AFF2 TP-PCR reaction did not arise from annealing of the TP primer within the AFF2 CCG repeat, which starts after the last CTG trinucleotide of the 5’ flanking sequence; instead, this solitary peak, which migrates as an apparently 111-bp fragment, arises from annealing of the TP primer to the (CCG)4CCT sequence immediately upstream of the repeat tract (Figure 1). The first true peak that arises from annealing of the TP primer to CCGs 1 to 5 of the AFF2 repeat, appears on electropherography as an apparently 138-bp fragment. TP primer annealing occurs at all other positions in the repeat tract containing five consecutive CCGs (eg, CCGs 2 to 6, 3 to 7), with the final fluorescence peak arising from annealing of the TP primer to the last five CCGs of the repeat tract. The repeat size of an AFF2 allele can therefore be rapidly and easily determined using a count of the

![Figure 1](https://example.com/figure1.png)
The absence of fluorescence peaks beyond 55 repeats indicates a lack of expansion, whereas continuous fluorescence peaks extending between 55 and 200 repeats indicate the presence of a premutation allele, and fluorescence peaks extending beyond 200 repeats indicate the presence of a full mutation. These results indicate that the duplex TP-PCR assay successfully detected and accurately identified the $FMR1$ CGG and $AFF2$ CCG repeat premutations and full mutations in DNA samples from both male and female individuals (Figures 2 and 3).

**Population Allele Distributions and Identification of Repeat Structures**

To determine allele size distribution, duplex TP-PCR was used for genotyping in a total of 466 DNA samples from male Chinese (161), Malay (158), Indian (89), Caucasian (44), and African American (14) individuals (Figure 4 and Supplemental Table S1). Allele distributions differed among ethnic groups, with a modal $AFF2$ repeat size of 18 in the Chinese and Malays (32.3% and 27.8%, respectively), but 15 in Caucasians (40.9%), African Americans (50.0%), and Indians (32.6%). Allele size ranges were wider in the Chinese (5 to 31) and Malays (6 to 37) compared to those in Indians (10 to 27), Caucasians (9 to 26), and African Americans (14 to 28). This finding could be partially attributed to the larger sample sizes of the Chinese and Malay populations. Using conventional categorization, 440 samples had normal alleles (11 to 30 repeats), 24 samples had minimal alleles (<11 repeats), and the remaining 2 samples had intermediate alleles (31 to 60 repeats). No premutation or full-mutation alleles were observed. These results were similar to those from an earlier study of the $AFF2$ allele distribution in Han Chinese from mainland China (mode 18; range, 9 to 26) and in Caucasians from New York, NY (mode 16; range, 8 to 34) (Figure 4, A–C).

Chromosomes carrying the combination of 29 $FMR1$ CGG repeats and 18 $AFF2$ CCG repeats were most common (Figure 4D). Of the 466 DNA samples from the male population screened, three $AFF2$ patterns were observed on TP-PCR electropherography. The most common pattern, as documented in the Genome Reference Consortium Human...
Build 38 (GRCh38) and described previously, \(^{24,25}\) was present in 459 samples (98.1%) representing most normal and all intermediate alleles. However, Sanger sequencing identified that the other two patterns were caused by the presence of two single-nucleotide polymorphism (SNP) variants (Figure 4D).

One of the SNP variants, a T>C substitution at chrX:148,500,637 (rs868914124, GRCh38), converts a CTG trinucleotide to a CCG at the 5’ start of the AFF2 CCG repeat.\(^{28}\) Whereas the AFF2 trinucleotide repeat in the rs868914124(T) common allele starts after the last CTG trinucleotide of the 5’ flanking sequence (Figure 5A), this CTG trinucleotide becomes a CCG trinucleotide in the rs868914124(C) variant allele, thus extending the AFF2 trinucleotide repeat tract two CCGs or 6 bp upstream to start after the last CAG trinucleotide of the 5’ flanking sequence (Figure 5B). As a consequence, the first TP-PCR product arising from within the repeat tract of the rs868914124(C) variant allele appears as an apparently 132-bp fluorescence peak (Figure 5B), in contrast to a 138-bp first fluorescence peak from the rs868914124(T) common allele (Figure 5A). On TP-PCR electropherography, this size difference shows up as a narrower gap of 20 bp between the leftmost solitary peak and the first peak of the AFF2 CCG repeat of the rs868914124(C) variant allele (Figure 5B), compared to a wider, 26-bp gap for the rs868914124(T) common allele (Figure 5, A and C).

The rs868914124(C) variant was observed in 8 of 466 AFF2 alleles (1.72%) (Figure 4D), of which 6 were alleles from Malay individuals. Three of the 8 samples with the rs868914124(C) variant contained minimal alleles (<11 repeats), compared to 21 of 458 samples carrying rs868914124(T). Using the Fisher exact test (two-tailed), this rare variant was observed to be enriched in the Malay group (odds ratio \(Z 6.01; 95\% \text{ CI, } 1.06 \text{ to } 61.7; P = 0.021\)) and with minimal alleles (odds ratio = 12.3; 95\% CI, 1.79 to 68.4; \(P = 0.006\)). Interestingly, all 6 AFF2 expanded alleles (five full mutations and one premutation) carried the rare rs868914124(C) variant, although the expanded alleles of an affected male and his maternal uncle, as well as the expanded alleles from a father—daughter pair, were assumed to be identical by descent (Figure 5D and Supplemental Figures S1 and S2).

Unlike AGG interruptions within the CGG repeats commonly observed in normal FMR1 alleles, only one normal AFF2 allele containing non-CCG interruptions

---

**Figure 3** Electropherography of FMR1 (FAM) and AFF2 (HEX) duplex TP-PCR products from AFF2 premutation and full-mutation DNA samples. Electropherography in both FAM and HEX (left), FAM only (middle), and HEX only (right) fluorescence channels is shown. Blue peaks indicate the FAM-labeled FMR1 TP-PCR product, whereas green peaks indicate the HEX-labeled AFF2 TP-PCR product. The threshold repeat size separating normal from expanded alleles is indicated by an orange line.
Figure 4  
AFF2 CCG repeat size and structure distributions.  
A: Population distribution of AFF2 CCG repeat sizes (x axis) and frequencies (y axis) in African American, Caucasian, Chinese, Indian, and Malay populations.  
B: Comparison of Chinese allele frequency distributions in Zhong et al\textsuperscript{24} (gray) and this study (black).  
C: Comparison of Caucasian allele frequency distributions in Zhong et al\textsuperscript{24} (pink) and this study (red).  
D: Heatmaps depicting the population distribution of X chromosomes with different FMR1 CCG and AFF2 CCG repeat size combinations (top), and the population repeat size distribution and abundance of the common and variant AFF2 alleles (bottom).
within its repeat stretch was observed, namely, a CTG interruption at the fifth repeat position (Figure 4D and Figure 5C). This interruption is caused by a C>T substitution at chrX:148,500,652 (rs1389911365, GRCh38). Annealing of the TP primer at positions that include this interruption will result in mismatched pairing and failure to generate a PCR product from that position. On TP-PCR electropherography, failed PCR at primer mismatch positions appears as gaps of absent peaks (Figure 5C). Interestingly, novel non-CCG interruptions were also observed within all six AFF2 expansion positive samples.

Interruptions within the repeat were initially suspected based on the patterns in four samples from males with AFF2 full-mutation alleles on TP-PCR electropherography, which showed gaps of absent/missing peaks between peak clusters (Figure 3). Sanger sequencing revealed the presence of one or more non-CCG interruptions within the 5' end of the repeat, bearing the sequence CCTGTGCAG, which was identical to a 9-nucleotide stretch immediately 5' upstream of the repeat. The interruptions vary in number from one (Supplemental Figure S1A) to more than four (Supplemental Figure S1B). They also vary in location, such as from the 8th to the 10th repeat positions (Supplemental Figure S1A) or from the 6th to 8th positions (Supplemental Figure S2).

In the father–daughter pair (Supplemental Figure S2), the AFF2 CCG repeat expansion was a full mutation in the father (DNA_25926) and a premutation in the daughter (DNA_3802), indicative of a previously documented contraction on transmission.3,29,30 Although AFF2 full-mutation alleles in the nephew (FX0229) and maternal uncle (FX0230) (Supplemental Figure S1) were assumed to be identical by descent, their repeat structures differed by the presence of additional non-CCG interruptions in the maternal uncle. The exact cause of or mechanism for this difference has not been investigated.

Blinded Validation of Duplex TP-PCR Assay on Previously Characterized Samples

Eighty-two archived and previously characterized genomic DNAs were included in a blinded validation of the duplex TP-
PCR assay. The AFF2 CCG repeats were sized as described in the Materials and Methods. The duplex TP-PCR assay accurately classified all 40 normal, 23 FMR1 full-mutation, 17 FMR1 premutation, and 2 AFF2 full-mutation samples that were included in the test (Supplemental Table S2).

Discussion

In recent years, molecular diagnosis of and screening for disorders caused by hyper-expansion of trinucleotide repeats have been greatly improved by the development of simple and rapid yet robust assays based on the TP-PCR method first described by Warner et al. The advantages of TP-PCR over Southern blot analysis or standard PCR include its capacity to accurately size all normal alleles and even modest expansions, to always detect all expansion mutations regardless of their size, and to identify interruptions within the repeat stretch through distinct patterns on TP-PCR electropherography. In contrast, Southern blot analysis can provide only rough estimates of repeat size, whereas standard PCR cannot detect large expansions. Neither method provides information on interruptions within the repeat. Attempts to detect repeat expansions using second-generation or long-read sequencing technologies have also been made. However, due to the technical difficulties in sequencing the GC-rich regions and in analyzing the reads, the sequencing assays lack both accuracy and accessibility.

The presence of variant AFF2 repeat structures caused by gain/loss of non-CCG interruptions, similar, but not identical, to the FMR1 AGG interruptions, may necessitate a closer look at how the CCG repeats in AFF2 are determined. The canonical AFF2 CCG repeat structure assumes that the repeat begins immediately downstream of the sequence CCTGTGCAGCCGCTG, and that no interruptions are present within the repeat stretch (Figure 1). To date, the AFF2 CCG repeat size has been most commonly determined using standard PCR across the repeat stretch, followed by fragment sizing using molecular markers, and finally subtraction of the repeat-flanking sequences and division of the remaining value by 3. Standard PCR and fragment sizing cannot, however, be used for detecting the rs868914124(C) variant or for determining the actual contiguous CCG repeat size due to the assumption that all alleles contain the rs868914124(T) reference nucleotide and do not contain intra-repeat interruptions.

In contrast, TP-PCR will be able to detect repeats with the rs868914124(C) variant with its two additional CCGs at the 5’-end of the repeat stretch, based on the narrower 20-bp gap between the leftmost solitary peak and the first true repeat-primed peak on electropherography. Given that any FRAXE-associated pathogenicity is assumed to be directly correlated with uninterrupted CCG repeat size, akin to FRAXA, the presence of the rs868914124(C) variant needs to be considered for accurate repeat sizing. Its importance is underlined by the observation that all AFF2 full-mutation alleles in this study carry the rs868914124(C) variant.

In silico RNA secondary structure predictions suggest that rs868914124, together with CCG repeat length and parity (odd/even repeat number) variation, could alter AFF2 mRNA thermodynamics and folding (Supplemental Figure S3), and affect protein expression. Although the findings from a recent study suggest that AFF2 protein stabilizes transcription across GC-rich regions of genes, further investigations are needed for determining how these AFF2 CCG repeat variations alter AFF2 transcriptional regulation of GC-rich genes.

This study also showed non-CCG interruptions (CCTGTGCAG) within the AFF2 CCG repeat, analogous to AGG interruptions within the FMR1 CCG repeat. In FMR1, AGG interruptions are commonly observed in normal alleles but are rarely observed in full-mutation alleles. These interruptions contribute to the overall repeat size and aid in stabilizing intergenerational transmission of the alleles. In marked contrast, non-CCG interruptions are extremely rare among AFF2 normal alleles, but their presence in all expanded alleles (five full mutations and one premutation) in this study suggests that they may be much more common in expanded alleles.

AFF2 may be involved in the splicing of FMR1 premRNA and in the regulation of FOS and JUN, whose proteins form ractivating protein 1, a transcription factor complex involved in cell proliferation and neuron activation. This relationship may explain the overlapping clinical similarities between FRAXA- and FRAXE-associated site expansions. Currently, clinical testing strategies for intellectual disability emphasize the use of chromosome microarray analysis as a first-line screening as it has a diagnostic yield of about 12%. However, due to cost considerations, simple PCR-based screening for FXS is often performed prior to chromosome microarray analysis, especially in males with intellectual disability and autism. This simple and inexpensive assay could encourage screening for other disorders suspected to be caused by FRAXA or FRAXE expansion mutations, and could facilitate the detection of potentially pathogenic variants and repeat structures. Hence, it may expand the phenotypic spectrum associated with these two fragile X sites.

Acknowledgments

We thank Guiping Phang and Simin Wong for technical assistance; and Hilary Martin, Arqum Anwar, Minaam Abbas, and Paulo Amaral for their useful comments.

Authors Contributions

S.S.C. conceptualized and coordinated the project and experimental design, and revised the manuscript. T.L. and F.S.W. performed the experiments, analyzed the data,
analysis and interpretation, and wrote and revised the manuscript. F.S.H.C. performed the blinded validation, sequencing of AFF2 CCG repeat, and analyzed the data. Y.G., M.S., D.L.N., and J.G. provided the AFF2 expansion—positive DNA samples and reviewed the manuscript. H-Y.L. provided the FMR1 expansion—positive DNA samples and reviewed the manuscript. S.K.H.T. and C.G.L. participated in the design of the study and reviewed the manuscript. All of the authors read and approved the manuscript.

Supplemental Data

Supplemental material for this article can be found at https://doi.org/10.1016/j.jmoldx.2021.04.015.

References