Clinica Chimica Acta 551 (2023) 117614

Contents lists available at ScienceDirect

Clinica Chimica Acta

journal homepage: www.elsevier.com/locate/cca

Evaluating the clinical utility of a long-read sequencing-based approach in genetic testing of fragile-X syndrome

Fei Hou [a,1], Aiping Mao [b,1], Shan Shan [a], Yan Li [a], Wanli Meng [b], Jiahan Zhan [b], Wenying Nie [a], Hua Jin [a,*]

[a] Department of Prenatal Diagnosis, Jinan Maternal and Child Health Hospital, Jinan 250001, Shandong Province, China
[b] Berry Genomics Corporation, Beijing 102200, China

ARTICLE INFO

Keywords:
Fragile X syndrome
CGG repeat
AGG interruption
Genetic testing
Long-read sequencing

ABSTRACT

Background: Fragile X syndrome (FXS) arises from the *FMR1* CGG expansion. Comprehensive genetic testing for *FMR1* CGG expansions, AGG interruptions, and microdeletions is essential to provide genetic counseling for females carrying premutation alleles. However, conventional PCR-based *FMR1* assays mainly focus on CGG repeats, and could detect AGG interruption only in males.

Methods: The clinical utility of a long-read sequencing-based assay termed comprehensive analysis of FXS (CAFXS) was evaluated in 238 high-risk samples by comparing to conventional PCR assays.

Results: PCR assays identified five premuation and three full mutation categories alleles in all the samples, and CAFXS successfully called all the *FMR1* CGG expansion. CAFXS identified 24-bp microdeletions upstream to the trinucleotide region with 30 CGG repeats, which was miscalled by the length-based PCR methods. CAFXS also identified a 187-bp deletion in about 1/7 of the sequencing reads in a male patient with mosaic full mutation alleles. CAFXS allowed for precise constructing the *FMR1* CGG repeat and AGG interruption pattern in all the samples, and identified a novel and alternative CGA interruption in one normal female sample.

Conclusions: CAFXS represents a more comprehensive and accurate approach for FXS genetic testing that potentially enables more informed genetic counseling compared to PCR-based methods.

1. Introduction

Fragile X syndrome (FXS; OMIM#300624) is one of the most common inherited cause of intellectual disabilities, affecting about 1/4,000 males and 1/8,000 females [1]. FXS is mainly caused by the expansion of CGG trinucleotide repeats in the 5′ untranslated region of the *FMR1* gene [2]. The normal, intermediate (gray zone), premutation, full mutations alleles have CGG repeats ranging from 5 to 44, 45–54, 55–200 and > 200, respectively. The CGG repeat of *FMR1* is susceptible to meiotic instability, which is reflected by repeat size differences between parents and offspring. Normal alleles remain stable upon transmission [2]. Intermediate alleles do no cause any clinical phenotypes, but display a higher risk for expansion to permutation in the offspring but not an expansion to full mutation within one generation [3]. Females with premutation alleles have approximately a 20 % risk for fragile X-associated primary ovarian insufficiency [4,5]. Older males and females with premutation alleles are at risk for fragile X-associated tremor/ataxia syndrome [6,7]. Adults with premuation alleles can also express a spectrum of neuropsychiatric problems referred to as fragile X-associated neuropsychiatric disorders[8]. The premutation alleles confer risk for expansion to full mutations during female germline transmission. The risk of expansion varies from 3 % for 59–69 repeats, to 60 % and 100 % for 70–80 repeats and > 90 repeats, respectively [3]. AGG trinucleotide interruptions within the *FMR1* CGG repeats can increase the repeat stability and reduce the risk of expansion during germline transmission [3,9,10]. The influence of AGG interruptions for CGG expansion to full mutation is most profound for repeats ranging from 70 to 80, whereby alleles with two or more AGGs reduce the risk by more than 60 % compared to alleles with no AGG [10]. Thus, expansion risk prediction in genetic counseling for female premutation carriers requires accurate analysis of both CGG repeat number and AGG interruptions.

Abbreviations: FXS, fragile X syndrome; CAFXS, comprehensive analysis of FXS; LRS, long-read sequencing; SNV, single-nucleotide variation.

* Corresponding author.
E-mail address: tonyshirly@163.com (H. Jin).

[1] Fei Hou and Aiping Mao contributed equally to this work.

https://doi.org/10.1016/j.cca.2023.117614
Received 5 September 2023; Received in revised form 19 October 2023; Accepted 22 October 2023
Available online 30 October 2023

The *FMR1* CGG repeat expansions can be determined by southern blot and/or conventional PCR-based methods. Southern blot can identify premutation and full mutation, but it has low resolution and accurate sizing of CGG repeat number in premutation alleles requires the use of PCR assay [11,12]. The *FMR1* CGG repeat-flanking regions are prone to microdeletions due to instability [13], which could prevent primer annealing for conventional PCR-based assays. The allele dropout could cause both false-negative and false-positive results for *FMR1* genetic analysis [14,15]. AGG interruption analysis can be achieved by triplet-primed PCR methods, but which only provide an indirect indication of the presence of AGG interruptions [16,17]. Interpreting the results is complicated in females with two *FMR1* alleles. Due to limitation of sequencing read length, next-generation sequencing-based methods could not directly determine large CGG expanded alleles [18,19]. To overcome these problems, long-read sequencing (LRS)-based methods have been developed to analyze *FMR1* CGG repeats and the embedded AGG interruptions [20–24]. The generated single-molecule long reads allow for determining the exact repeat pattern for each allele individually.

In this retrospective blind clinical study, we evaluated the clinical utility of a LRS-based approach termed comprehensive analysis of FXS (CAFXS) for analyzing CGG repeat, AGG interruption, microdeletions, single-nucleotide variations (SNVs) and indels of *FMR1*. CAFXS and conventional PCR-based methods were performed for 238 high risk subjects and the results were compared for clinical effectiveness in FXS genetic testing. The *FMR1* CGG repeat and AGG interruption pattern in each allele was also analyzed in the enrolled cohort.

2. Materials and methods

2.1. Study subjects

A total of 238 subjects from 234 families with suspected clinical features of FXS were retrospectively enrolled for CAFXS testing (Supplementary Table S1). The participants were recruited from Jinan Maternal and Child Health Hospital during Jan 2020 to Dec 2021. All the samples met at least one of the following inclusion criteria: primary ovarian insufficiency, infertility, recurrent miscarriage, developmental delay, autism spectrum disorder, intellectual disabilities, learning and behavioral issues, abnormal physical features and family history of FXS. The subjects had an average age of 30.9 years, ranging from 1 to 62 years. The samples were previously tested by*FMR1*PCR assays, relabeled and sent to Berry Genomics for blind CAFXS genetic testing (Fig. 1). This study adhered to the Declaration of Helsinki and was approved by the Ethics Committee of Jinan Maternal and Child Health Hospital (Approval No. IEC20200301-05). Informed written consent was obtained from all the subjects or their legal guardians.

2.2. FXS genetic testing by PCR assays

Two *FMR1* PCR-based assays were employed to analyze CGG repeats for all the enrolled samples. The first was FragilEase PCR reagent kit supplied by PerkinElmer. The second was *FMR1* PCR kit supplied by BioFast Biotech. The CGG repeat number of *FMR1* was calculated according to the instruction manuals.

2.3. FXS genetic testing by long-read sequencing

The CGG reaction of CAFXS assay was performed as previously described, to analyze CGG repeats, AGG interruptions, microdeletions, single-nucleotide variants and indels in exon 1 of *FMR1* [21]. The genomic DNA was subjected to CGG PCR reaction to amplify exon 1 of *FMR1*.The PCR products were then ligated to unique PacBio barcoded adaptor by one-step end-repair and ligation, digested with exonucleases to removed failed ligation products. The uniquely barcoded pre-libraries were purified, quantified with Qubit dsDNA HS assay kit (ThermoFisher Scientific) and then pooled with equal mass, followed by a second purification with AMPure PB beads (Pacific Biosciences). The pooled library were converted to single-molecule real-time dumbbell (SMRTbell) library using the Sequel Binding Kit 2.0 and Internal Control Kit 1.0 (Pacific Biosciences), and then sequenced on Sequel IIe platform (Pacific Biosciences) for 30 h under circular consensus sequencing (CCS) mode.

The raw sequencing subreads were processed to obtain high-quality CCS reads, debarcoded to individual samples, and then aligned to the hg38 reference build using the software suite (smrtlink 10.1.0.119588, Pacific Biosciences). CCS reads with 100 bp sequences flanking CGG repeats on both sides were retrieved by BLAST. CCS reads were then trimmed of flanking sequences to include only the repeat region and sorted from shortest to longest. The CGG repeats and AGG interruptions were color-coded and visualized as waterfall plots. The CGG repeat number of each CCS read was calculated according to length and then kernel density estimation method was applied to identify peaks of CGG repeat number, and the peaks were correlated to different alleles. Here the number of CGG repeats included both CGG and AGG. The microdeletions in exon 1 of *FMR1* were determined by aligning flanking sequences of CGG repeats to reference hg38. The intragenic SNVs and indels of *FMR1* were identified using $\geq$ 30 CCS reads by FreeBayes1.3.4 (Biomatters, Inc., San Diego, CA).The microdeletions of *FMR1* were displayed and confirmed in the Integrative Genomics Viewer (IGV) program with CCS reads.

2.4. Validation of discordant results

Discordance between CAFXS and PCR assays was defined as: 1) $\geq$ 2 CGG repeats difference in the range of 5–54; 2) $\geq$ 5 CGG repeats difference in the range of 55–199; 3) microdeletions or disease-causing SNVs/indels identified by CAFXS. Samples with microdeletions in

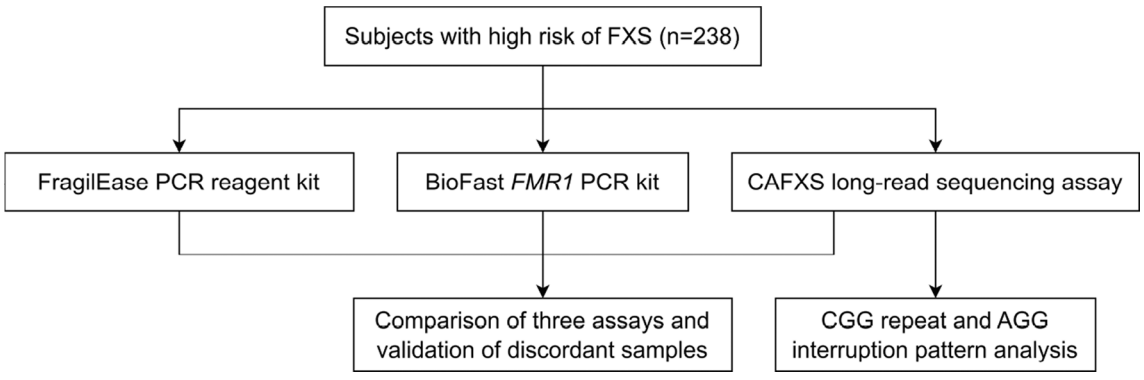

Fig. 1. Flowchart of study design.

exon 1 of *FMR1* identified by CAFXS were validated by specifically designed PCR followed by Sanger sequencing. The primers used for PCR were 5′-GCGCTCAGCTCCGTTTCGGTTTCACTTC-3′ and 5′-TGTAGAAAGCGCCATTGGAGCCCCGCAC-3′.

3. Results

3.1. FXS genetic testing by CAFXS and conventional PCR-based methods

CAFXS assay was performed on the enrolled 238 samples for detecting CGG repeats, AGG interruptions, microdeletion, SNVs and indels in exon 1 of the *FMR1*gene (Supplemental Table S1).*FMR1* CGG expanded alleles were identified in nine samples from seven (2.99 %, 7/234) families (Table 1). Out of the nine samples, five had premutation alleles, three had full mutation alleles, and one had both premutation and full mutation alleles. The AGG interruptions were also precisely analyzed in each allele of all the samples. Moreover, CAFXS revealed microdeletions in the exon1 and upstream region of *FMR1* in three samples. No pathogenic or likely pathogenic SNVs and indels in the *FMR1* target region were identified in these samples.

FagileEase PCR reagent kit and BioFast *FMR1* PCR kit, two conventional PCR-based methods, were employed for CGG repeat analysis.

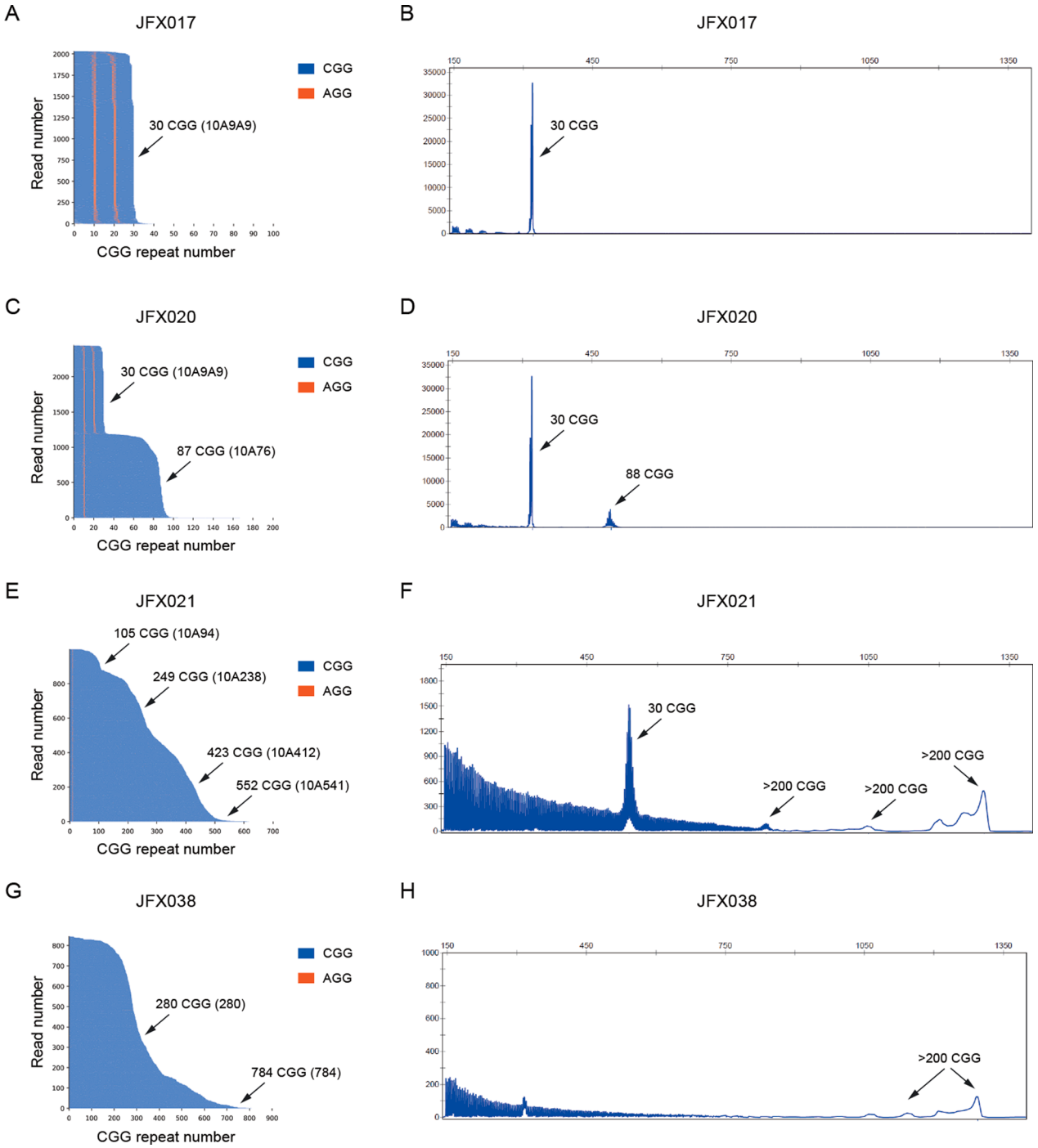

Fig. 2. Display of FMR1 CGG expansions by CAFXS and BioFast PCR assay for representative samples. (A and B) Sample JFX017 with normal allele. (C and D) Sample JFX020 with one normal allele and one premutation allele. (E and F) Sample JFX021 with mosaic premutation and full mutation alleles. (G and H) Sample JFX038 with mosaic full mutation alleles. The arrows highlighted the alleles with different CGG repeats.

These two kits have completely concordant results for all the samples analyzed (Supplemental Table S1). For normal and premutation alleles, the difference of CGG repeat number between two kits was ≤ 1. For full mutation alleles (>200 CGG), the exact CGG repeat number could not be precisely determined by PCR-based methods.

3.2. Comparison between CAFXS and conventional PCR-based methods

CAFXS and PCR assays were concordant in terms of calling normal, premutation and full mutations alleles (Supplemental Table S1, Fig. 2A-H). For CAFXS assay, the CCS reads were displayed in waterfall plots that directly showed the CGG repeat number, with AGG interruptions highlighted (Fig. 2A, 2C, 2E, 2G). The *FMR1* alleles with different CGG repeats amplified by conventional PCR-based method were separated by capillary electrophoresis (Fig. 2B, 2D, 2F, 2H).Of note, CAFXS identified four mosaic alleles with different CGG repeats (105, 249, 423, and 552) in the male proband JFX021, who had intellectual disability (Fig. 2E). CAFXS also identified two mosaic full mutation alleles with CGG repeats of 280 and 784 in the male proband JFX037, who had developmental delay (Fig. 2G).

CAFXS additionally identified microdeletions in three subjects from family FX032 (Table 1, Fig. 3A). CAFXS and conventional PCR assay were discordant for the exact number of CGG repeat in JFX033 (I-2) and JFX035 (II-2). CAFXS showed an allele with 30 CGG repeat in the two samples, while conventional PCR suggested the allele only had 22 CGG repeat (Fig. 3B-E). Further analysis by CAFXS demonstrated that both two samples had 24-bp deletions (chrX:147,911,987–147,912,010) upstream to CGG repeat region (Fig. 3F), which were confirmed by Sanger sequencing (Fig. 3G). Moreover, CAFXS identified mosaic full mutation alleles with different CGG repeats (289, 402, and 563) in the proband JFX036 (III-1) (Fig. 3H). Out of the 999 CCS reads obtained for this sample, 855 (6/7) had intact repeat-flanking sequence, and 144 (1/7) had a 187-bp deletion adjacent to the CGG repeat region (Fig. 3H and 3I). All the other samples did not harbor this 187-bp deletion. Taken together, CAFXS showed better accuracy for determining exact CGG repeats compared to conventional PCR assays.

3.3. CGG repeat and AGG interruption pattern analysis by CAFXS

CAFXS successfully revealed the AGG interruption patterns in all the CGG repeat categories, even in heterozygous and mosaic alleles (Supplementary Table S1, Fig. 4A-I). In total, CAFXS revealed 55 different CGG repeat and AGG interruption patterns from 450 alleles with normal CGG repeats (Table 2). The three most common CGG repeats were 29 (49.1 %), 30 (25.8 %), and 36 (10.4 %). Of the 450 alleles, 335 (74.4 %), 60 (13.3 %), 47 (10.4 %), 5 (1.1 %), and 2 (0.4 %) had 2, 1, 3, 0, and 4 AGG interruptions, respectively. The three most common CGG repeat and AGG interruption patterns $(CGG)_9AGG(CGG)_9AGG(CGG)_9$ (abbreviated as 9A9A9), 10A9A9, and 9A9A6A9) accounted for 77.6 % (349/450) of all X chromosomes. Among the 221 alleles with 29 CGG repeats, 200 (90.5 %) had 2 AGG interruptions, and 21 (9.5 %) lost one or both two AGG interruptions. Five (4.3 %) out of 116 alleles with 30 CGG repeats lost one AGG interruption. Seven (14.9 %) out of 47 alleles with 36 CGG repeats lost one or two AGG interruptions. Sequencing of *FMR1* CGG repeats not only identified AGG interruption, but also any other sequence variation. CAFXS showed that both the two alleles of sample JFX114 had 29 CGG repeats, of which one allele was interrupted by one AGG repeat (9A19), and the other allele was interrupted by two AGG repeats and one CGA repeat (9A9A5C3) (Fig. 4I).

4. Discussion

Accurate analysis of *FMR1* CGG repeat, AGG interruption, microdeletions, and SNVs/indels has been challenging using conventional methods, but this is now overcome by LRS approaches[21,22]. PacBio long-read single-molecule sequencing generates high-fidelity CCS reads and allows for constructing the *FMR1* repeat region unambiguously for both X chromosomes in females. This study evaluated the clinical utility

Table 1
Families with *FMR1* CGG-expanded alleles and AGG interruptions detected by CAFXS.

Family	Sample	Member	Gender	Age (year)	Clinical feature	Type	FragilEase	BioFast	CAFXS		
							CGG repeat	CGG repeat	CGG repeat	AGG pattern	Microdeletion
F020	JFX020	Proband	F	30	One miscarriage, one son with intellectual disability	N / PM	30 / 88	30 / 88	30 / 87	10A9A9* / 10A76	ND
F021	JFX021	Proband	M	7	Intellectual disability	PM / FM	105 / >200	105 / >200	105 / 249 / 423 / 552	10A94 / 10A238 / 10A412 / 10A541	ND
FX032	JFX032	Maternal grandfather	M	62	Normal	N	29	29	29	9A9A9	ND
	JFX033	Maternal grandmother	F	60	Normal	N / PM	22 / 71	22 / 71	30 / 70	10A9A9 / 70	Del (chrX:147911987–147912010)
	JFX034	Mother	F	36	Normal	N / FM	30 / >200	29 / >200	29 / 369 / 749	9A9A9 / 369 /749	ND
	JFX035	Aunt	F	33	Normal	N / N	22 / 29	22 / 29	29 / 30	9A9A9 / 10A9A9	Del (chrX:147911987–147912010)
	JFX036	Proband	M	15	Intellectual disability, long face	FM	>200	>200	289 / 402 / 563	289 / 402 / 563	Del (chrX:147911864–147912050)
F034	JFX037	Proband	M	2	Developmental delay	FM	>200	>200	280 / 784	280 / 784	ND
F040	JFX043	Proband	F	39	Recurrent miscarriage	N / PM	30 / 70	30 / 70	30 / 69	10A19 / 69	ND
F231	JFX235	Proband	F	30	Family history of FXS	N / PM	37 / 84	36 / 83	36 / 83	9A26 / 10A72	ND
F234	JFX238	Proband	F	51	Family history of FXS	N / PM	29 / 72	29 / 72	29 / 71	9A19 / 10A60	ND

F: female; M: male; N, normal allele; PM, premutation; FM, full mutation; ND, not detected; *, 10A9A9 was the abbreviation for $(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$.

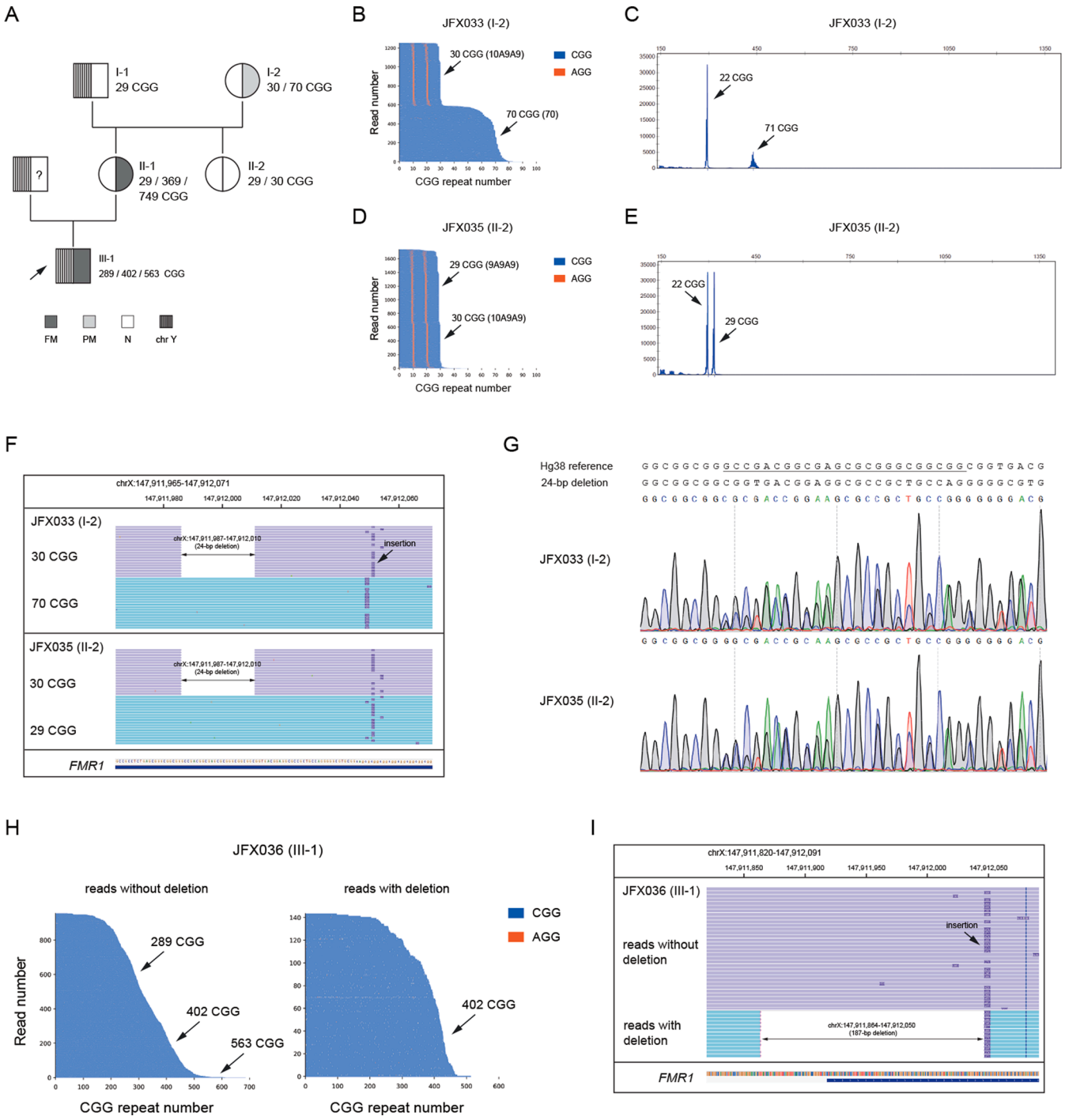

Fig. 3. Identification of *FMR1* microdeletions by CAFXS in family F032. (A) Pedigree chart of family F032. (B) Sample JFX033 with CGG repeat 30 + 70 by CAFXS. (C) Sample JFX033 with CGG repeat 22 + 71 by BioFast PCR assay. (D) Sample JFX035 with CGG repeat 29 + 30 by CAFXS. (E) Sample JFX035 with CGG repeat 22 + 29 by BioFast PCR assay. (F) IGV plot displaying the 24-bp microdeletions in sample JFX033 and JFX035 by CAFXS. The CGG repeat number of *FMR1* in reference build hg38 is 20. Thus, when the CCS reads with 30 or 70 CGG expansion were displayed in IGV, they showed as 30-bp or ~ 150-bp insertions in the repeat region. (G) Sanger sequencing confirmed the 24-bp microdeletions in sample JFX033 and JFX035. (H) Waterfall plots showing the CGG repeats in CCS reads with deletion and without deletion in sample JFX036 by CAFXS. (I) IGV plot showing the 187-bp deletion in sample JFX036 by CAFXS. The arrows highlighted the alleles with different CGG repeats.

of LRS-based CAFXS approach for genetic testing of *FMR1* in 238 high-risk subjects and compared side-by-side with PCR-based methods.

CAFXS successfully called all the normal, premuation and full mutations alleles identified by PCR assays. In addition, CAFXS determined the CGG repeat number of all categories, while PCR assays could only determine the exact repeat number below 200. The *FMR1* CGG repeat flanking region is unstable and prone to microdeletions of a few to hundreds of base pairs [13,25]. Due to limited resolution of capillary electrophoresis, conventional PCR-based methods use primers only a couple of hundred base pairs flanking CGG repeat region and rely on length but not sequence of PCR amplicons to calculate the number of CGG repeat. On one side, the primer annealing can be disrupted by microdeletions, which leads to artificial null allele [15,21]. On the other side, microdeletion within the amplicon could lead to calculation bias.

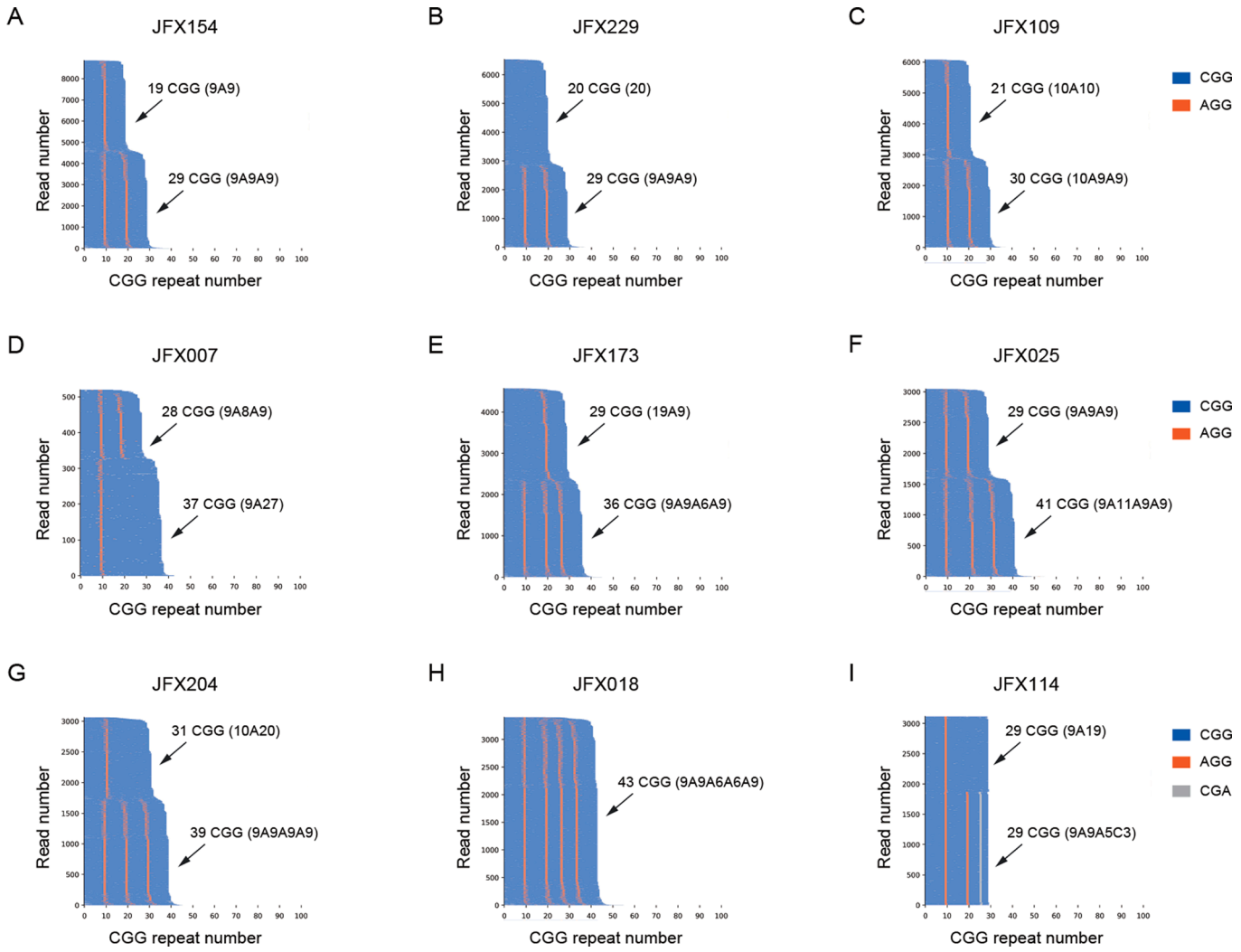

Fig. 4. Waterfall plots showing the different patterns of CGG repeat and AGG interruption in sample JFX154 (A), JFX229 (B), JFX109 (C), JFX007 (D), JFX173 (E), JFX025 (F), JFX204 (G), JFX018 (H), and JFX114 (I).The arrows highlighted the alleles with different CGG repeats.

Here, CAFXS revealed microdeletions in exon 1 of *FMR1* in three subjects from the same family. Due to the 24-bp microdeletions in JFX033 and JFX035, PCR methods miscalled the 30 CGG repeat to 22. CAFXS also identified about 1/7 of the CCS reads with full mutation had 187-bp deletion in sample JFX036. Under some circumstances, the microdeletions could lead to decrease of CGG repeat number, disruption of *FMR1* exon1, allele dropout or change the categories of the alleles identified by PCR assays, and thus are of important clinical significance. CAFXS used long-range primer a few kilobases upstream of CGG repeat and determined the repeat number by actual sequencing, which would be minimally affected by microdeletions.

Incorporating AGG interruption analysis into *FMR1* genetic testing allows accurate risk estimates and greatly improves genetic counseling for woman carrying a premutation allele [9–11,26]. The impact of AGG interruptions is the most profound for females carrying a premutation allele with CGG repeats between approximately 60 s to 80 s. In this study, CAFXS identified premutation alleles in five females with repeat number between 69 and 87.Of the five premutation alleles, three had one AGG interruption and two had no AGG interruption. In family FX032, the premutation allele with 70 CGG repeats and no AGG was expanded to full mutation during transmission. Females with high risk of having a child with FXS might opt for assisted reproduction where one can select for low risk embryos. More accurate risk estimates with LRS approaches would provide high risk couples with appropriate reproductive strategy.

PCR-based assays and Sanger sequencing usually study CGG repeat and AGG interruption pattern were studied in males [27–29], because the two X chromosomes camouflage each other's repeat pattern. LRS provides a direct sequencing of the *FMR1* trinucleotide repeats and hence allows to separate two X chromosomes in females and mosaic alleles. CAFXS allows for successful determination of AGG interruption of all the alleles in the 238 samples. The most common patterns analyzed in the Chinese cohort were 9A9A9, 10A9A9, and 9A9A6A9, which were concordant with previously published report in three Asian populations [29]. Although most interruptions in the CGG repeat are AGG trinucleotide, alternative interruptions like TGG [30], has been discovered. Here, CAFXS identified a novel CGA repeat in sample JFX114. These alternative interruptions might also stabilize the CGG repeat. Systematic research on the transmission of alleles with those alternative and rare interruptions would provide insights in the function of such repeats.

To summarize, LRS-CAFXS has wider detection scope and thus provides better FXS genetic characterization compared to conventional PCR-based assays. With the CGG reaction, the cost per sample of LRS-CAFXS should be comparable to PCR assays. The turnaround time of LRS and PCR assays are approximately 6–8 and 2–3 days, respectively. With the development of more LRS-based assays including thalassemia [31], congenital adrenal hyperplasia [32], and spinal muscular atrophy [33], these tests can be pooled for high-throughput genetic screening

Table 2
Detection of AGG interruption patterns among *FMR1* CGG normal alleles.

No.	CGG repeat number	AGG interruptions	CGG-AGG pattern	Count	Ratio
1	29	2	9A9A9*	199	44.22 %
2	30	2	10A9A9	110	24.44 %
3	36	3	9A9A6A9	40	8.89 %
4	29	1	9A9A5C3	14	3.11 %
5	29	1	19A9	6	1.33 %
6	23	1	9A13	5	1.11 %
7	32	2	9A9A12	5	1.11 %
8	19	1	9A9	4	0.89 %
9	24	1	9A14	3	0.67 %
10	31	2	9A9A11	3	0.67 %
11	31	2	10A9A10	3	0.67 %
12	39	3	9A9A9A9	3	0.67 %
13	20	1	10A9	2	0.44 %
14	30	1	9A20	2	0.44 %
15	30	1	10A19	2	0.44 %
16	31	1	10A20	2	0.44 %
17	32	1	9A22	2	0.44 %
18	33	3	9A6A6A9	2	0.44 %
19	34	1	9A24	2	0.44 %
20	35	1	9A25	2	0.44 %
21	36	2	9A16A9	2	0.44 %
22	36	1	9A26	2	0.44 %
23	37	1	9A27	2	0.44 %
24	38	2	9A9A18	2	0.44 %
25	10	0	10	1	0.22 %
26	11	0	11	1	0.22 %
27	17	0	17	1	0.22 %
28	20	0	20	1	0.22 %
29	21	1	10A10	1	0.22 %
30	22	0	22	1	0.22 %
31	22	1	12A9	1	0.22 %
32	25	1	9A15	1	0.22 %
33	25	1	15A9	1	0.22 %
34	26	2	9A6A9	1	0.22 %
35	28	2	9A8A9	1	0.22 %
36	29	0	29	1	0.22 %
37	29	1	8A20	1	0.22 %
38	30	2	13A6A9	1	0.22 %
39	30	1	20A9	1	0.22 %
40	31	2	10A10A9	1	0.22 %
41	31	2	11A10A8	1	0.22 %
42	32	1	22A9	1	0.22 %
43	33	1	9A23	1	0.22 %
44	34	2	9A14A9	1	0.22 %
45	35	2	9A15A9	1	0.22 %
46	36	2	19A6A9	1	0.22 %
47	36	1	26A9	1	0.22 %
48	36	2	9A9A16	1	0.22 %
49	37	4	9A9A6A6A3	1	0.22 %
50	37	3	12A6A7A9	1	0.22 %
51	40	2	9A9A20	1	0.22 %
52	41	3	9A11A9A9	1	0.22 %
53	42	2	22A9A9	1	0.22 %
54	43	1	33A9	1	0.22 %
55	43	4	9A9A6A6A9	1	0.22 %

* , 9A9A9 was the abbreviation for $(CGG)_9AGG(CGG)_9AGG(CGG)_9$.

like expanded carrier screening. However, conventional PCR-based *FMR1* assays using Genetic Analyzer are not suitable for panel-based genetic screening.

5. Conclusion

Our study demonstrated high clinical utility of LRS-CAFXS for simultaneous characterization of CGG expansions, AGG interruptions, and microdeletions involving the *FMR1* gene. The approach should be easily adapted for comprehensive analysis of other complicated repeat expansion disorders. CAFXS is envisaged to greatly improve FXS risk estimates for better genetic counseling, meanwhile allows genetic screening for potentially *FMR1*-related indications.

CRediT authorship contribution statement

Fei Hou: Conceptualization, Methodology, Resources, Writing – original draft. **Aiping Mao:** Conceptualization, Software, Data curation, Writing – original draft. **Shan Shan:** Validation, Formal analysis. **Yan Li:** Investigation, Writing – review & editing. **Wanli Meng:** Validation, Formal analysis. **Jiahan Zhan:** Investigation, Supervision. **Wenying Nie:** Resources, Supervision. **Hua Jin:** Conceptualization, Resources, Writing – review & editing, Project administration.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: [A.M., W.M., and J.Z. are employees of Berry Genomics Corporation. The authors declare no conflicts of interest].

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by the National Key Research and Development Program of China (2021YFC1005301 and 2021YFC1005303), Shandong Medical and Health Science and Technology Program (202105030122), and Jinan Clinical Medical Science and Technology Innovation Program (202225064). We thank all the subjects who participated in this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cca.2023.117614.

References

[1] J. Hunter, O. Rivero-Arias, A. Angelov, E. Kim, I. Fotheringham, J. Leal, Epidemiology of fragile X syndrome: a systematic review and meta-analysis, Am. J. Med. Genet. A 164A (7) (2014) 1648–1658.
[2] M. Mila, M.I. Alvarez-Mora, I. Madrigal, L. Rodriguez-Revenga, Fragile X syndrome: An overview and update of the FMR1 gene, Clin. Genet. 93 (2) (2018) 197–205.
[3] S.L. Nolin, W.T. Brown, A. Glicksman, G.E. Houck Jr., A.D. Gargano, A. Sullivan, V. Biancalana, K. Brondum-Nielsen, H. Hjalgrim, E. Holinski-Feder, F. Kooy, J. Longshore, J. Macpherson, J.L. Mandel, G. Matthijs, F. Rousseau, P. Steinbach, M.L. Vaisanen, H. von Koskull, S.L. Sherman, Expansion of the fragile X CGG repeat in females with premutation or intermediate alleles, Am. J. Hum. Genet. 72 (2) (2003) 454–464.
[4] M.D. Wittenberger, R.J. Hagerman, S.L. Sherman, A. McConkie-Rosell, C.K. Welt, R.W. Rebar, E.C. Corrigan, J.L. Simpson, L.M. Nelson, The FMR1 premutation and reproduction, Fertil. Steril. 87 (3) (2007) 456–465.
[5] R.D. Hundscheid, A.P. Smits, C.M. Thomas, L.A. Kiemeney, D.D. Braat, Female carriers of fragile X premutations have no increased risk for additional diseases other than premature ovarian failure, Am. J. Med. Genet. A 117A (1) (2003) 6–9.
[6] R.J. Hagerman, B.R. Leavitt, F. Farzin, S. Jacquemont, C.M. Greco, J.A. Brunberg, F. Tassone, D. Hessl, S.W. Harris, L. Zhang, T. Jardini, L.W. Gane, J. Ferranti, L. Ruiz, M.A. Leehey, J. Grigsby, P.J. Hagerman, Fragile-X-associated tremor/ataxia syndrome (FXTAS) in females with the FMR1 premutation, Am. J. Hum. Genet. 74 (5) (2004) 1051–1056.
[7] F. Tassone, R.J. Hagerman, D. Garcia-Arocena, E.W. Khandjian, C.M. Greco, P. J. Hagerman, Intranuclear inclusions in neural cells with premutation alleles in fragile X associated tremor/ataxia syndrome, J. Med. Genet. 41 (4) (2004) e43.
[8] R.J. Hagerman, D. Protic, A. Rajaratnam, M.J. Salcedo-Arellano, E.Y. Aydin, A. Schneider, Fragile X-Associated Neuropsychiatric Disorders (FXAND), Front. Psych. 9 (2018) 564.
[9] C.M. Yrigollen, L. Martorell, B. Durbin-Johnson, M. Naudo, J. Genoves, A. Murgia, R. Polli, L. Zhou, D. Barbouth, A. Rupchock, B. Finucane, G.J. Latham, A. Hadd, E. Berry-Kravis, F. Tassone, AGG interruptions and maternal age affect FMR1 CGG repeat allele stability during transmission, J. Neurodev. Disord. 6 (1) (2014) 24.
[10] C.M. Yrigollen, B. Durbin-Johnson, L. Gane, D.L. Nelson, R. Hagerman, P. J. Hagerman, F. Tassone, AGG interruptions within the maternal FMR1 gene reduce the risk of offspring with fragile X syndrome, Genet. Med. 14 (8) (2012) 729–736.

[11] V. Biancalana, D. Glaeser, S. McQuaid, P. Steinbach, EMQN best practice guidelines for the molecular genetic testing and reporting of fragile X syndrome and other fragile X-associated disorders, Eur. J. Hum. Genet. 23 (4) (2015) 417–425.
[12] S. Filipovic-Sadic, S. Sah, L. Chen, J. Krosting, E. Sekinger, W. Zhang, P. J. Hagerman, T.T. Stenzel, A.G. Hadd, G.J. Latham, F. Tassone, A novel FMR1 PCR method for the routine detection of low abundance expanded alleles and full mutations in fragile X syndrome, Clin. Chem. 56 (3) (2010) 399–408.
[13] R.D. Wells, Mutation spectra in fragile X syndrome induced by deletions of CGG*CCG repeats, J. Biol. Chem. 284 (12) (2009) 7407–7411.
[14] E. de Graaff, P. Rouillard, P.J. Willems, A.P. Smits, F. Rousseau, B.A. Oostra, Hotspot for deletions in the CGG repeat region of FMR1 in fragile X patients, Hum. Mol. Genet. 4 (1) (1995) 45–49.
[15] M.R. Hegde, B. Chong, M. Fawkner, N. Lambiris, H. Peters, A. Kenneson, S. T. Warren, D.R. Love, J. McGaughran, Microdeletion in the FMR-1 gene: an apparent null allele using routine clinical PCR amplification, J. Med. Genet. 38 (9) (2001) 624–629.
[16] B.E. Hayward, K. Usdin, Improved Assays for AGG Interruptions in Fragile X Premutation Carriers, J. Mol. Diagn. 19 (6) (2017) 828–835.
[17] L. Chen, A. Hadd, S. Sah, S. Filipovic-Sadic, J. Krosting, E. Sekinger, R. Pan, P. J. Hagerman, T.T. Stenzel, F. Tassone, G.J. Latham, An information-rich CGG repeat primed PCR that detects the full range of fragile X expanded alleles and minimizes the need for southern blot analysis, J. Mol. Diagn. 12 (5) (2010) 589–600.
[18] M. Bahlo, M.F. Bennett, P. Degorski, R.M. Tankard, M.B. Delatycki, P.J. Lockhart, Recent advances in the detection of repeat expansions with short-read next-generation sequencing, F1000Res 7 (2018) 736.
[19] K. Ibanez, J. Polke, R.T. Hagelstrom, E. Dolzhenko, D. Pasko, E.R.A. Thomas, L.C. Daugherty, D. Kasperaviciute, K.R. Smith, W.G.S.f.N.D. Group, Z.C. Deans, S. Hill, T. Fowler, R.H. Scott, J. Hardy, P.F. Chinnery, H. Houlden, A. Rendon, M.J. Caulfield, M.A. Eberle, R.J. Taft, A. Tucci, C, Genomics England Research, Whole genome sequencing for the diagnosis of neurological repeat expansion disorders in the UK: a retrospective diagnostic accuracy and prospective clinical validation study, Lancet Neurol. 21 (3) (2022) 234–245.
[20] V. Grosso, L. Marcolungo, S. Maestri, M. Alfano, D. Lavezzari, B. Iadarola, A. Salviati, B. Mariotti, A. Botta, M.R. D'Apice, G. Novelli, M. Delledonne, M. Rossato, Characterization of FMR1 Repeat Expansion and Intragenic Variants by Indirect Sequence Capture, Front. Genet. 12 (2021), 743230.
[21] Q. Liang, Y. Liu, Y. Liu, R. Duan, W. Meng, J. Zhan, J. Xia, A. Mao, D. Liang, L. Wu, Comprehensive Analysis of Fragile X Syndrome: Full Characterization of the FMR1 Locus by Long-Read Sequencing, Clin. Chem. 68 (12) (2022) 1529–1540.
[22] S. Ardui, V. Race, T. de Ravel, H. Van Esch, K. Devriendt, G. Matthijs, J. R. Vermeesch, Detecting AGG Interruptions in Females With a FMR1 Premutation by Long-Read Single-Molecule Sequencing: A 1 Year Clinical Experience, Front. Genet. 9 (2018) 150.
[23] S. Ardui, V. Race, A. Zablotskaya, M.S. Hestand, H. Van Esch, K. Devriendt, G. Matthijs, J.R. Vermeesch, Detecting AGG Interruptions in Male and Female FMR1 Premutation Carriers by Single-Molecule Sequencing, Hum. Mutat. 38 (3) (2017) 324–331.
[24] E.W. Loomis, J.S. Eid, P. Peluso, J. Yin, L. Hickey, D. Rank, S. McCalmon, R. J. Hagerman, F. Tassone, P.J. Hagerman, Sequencing the unsequenceable: expanded CGG-repeat alleles of the fragile X gene, Genome Res. 23 (1) (2013) 121–128.
[25] B. Coffee, M. Ikeda, D.B. Budimirovic, L.N. Hjelm, W.E. Kaufmann, S.T. Warren, Mosaic FMR1 deletion causes fragile X syndrome and can lead to molecular misdiagnosis: a case report and review of the literature, Am. J. Med. Genet. A 146A (10) (2008) 1358–1367.
[26] S.L. Nolin, A. Glicksman, N. Ersalesi, C. Dobkin, W.T. Brown, R. Cao, E. Blatt, S. Sah, G.J. Latham, A.G. Hadd, Fragile X full mutation expansions are inhibited by one or more AGG interruptions in premutation carriers, Genet. Med. 17 (5) (2015) 358–364.
[27] E.E. Eichler, J.N. Macpherson, A. Murray, P.A. Jacobs, A. Chakravarti, D.L. Nelson, Haplotype and interspersion analysis of the FMR1 CGG repeat identifies two different mutational pathways for the origin of the fragile X syndrome, Hum. Mol. Genet. 5 (3) (1996) 319–330.
[28] H. Dokic, I. Barisic, V. Culic, B. Lozic, S. Hecimovic, Haplotype and AGG interspersion analysis of FMR1 alleles in a Croatian population: no founder effect detected in patients with fragile X syndrome, Hum. Biol. 80 (5) (2008) 581–587.
[29] Y. Zhou, K. Tang, H.Y. Law, I.S. Ng, C.G. Lee, S.S. Chong, FMR1 CGG repeat patterns and flanking haplotypes in three Asian populations and their relationship with repeat instability, Ann. Hum. Genet. 70 (Pt 6) (2006) 784–796.
[30] C.B. Kunst, S.T. Warren, Cryptic and polar variation of the fragile X repeat could result in predisposing normal alleles, Cell 77 (6) (1994) 853–861.
[31] Q. Liang, W. Gu, P. Chen, Y. Li, Y. Liu, M. Tian, Q. Zhou, H. Qi, Y. Zhang, J. He, Q. Li, L. Tang, J. Tang, Y. Teng, Y. Zhou, S. Huang, Z. Lu, M. Xu, W. Hou, T. Huang, Y. Li, R. Li, L. Hu, S. Li, Q. Guo, Z. Zhuo, Y. Mou, D.S. Cram, L. Wu, A More Universal Approach to Comprehensive Analysis of Thalassemia Alleles (CATSA), J. Mol. Diagn. 23 (9) (2021) 1195–1204.
[32] Y. Liu, M. Chen, J. Liu, A. Mao, Y. Teng, H. Yan, H. Zhu, Z. Li, D. Liang, L. Wu, Comprehensive Analysis of Congenital Adrenal Hyperplasia Using Long-Read Sequencing, Clin Chem 68(7) (2022) 927-939.
[33] S. Li, X. Han, Y. Xu, C. Chang, L. Gao, J. Li, Y. Lu, A. Mao, Y. Wang, Comprehensive Analysis of Spinal Muscular Atrophy: SMN1 Copy Number, Intragenic Mutation, and 2 + 0 Carrier Analysis by Third-Generation Sequencing, J. Mol. Diagn. 24 (9) (2022) 1009–1020.