

Universidade do Estado do Rio de Janeiro

Centro Biomédico Instituto de Biologia Roberto Alcantara Gomes

Camilla Pereira da Silva

Investigação de novos mecanismos genéticos e epigenéticos relacionados à Síndrome do X-frágil

Rio de Janeiro 2022 Camilla Pereira da Silva

Investigação de novos mecanismos genéticos e epigenéticos relacionados à síndrome do X-frágil

Tese apresentada, como requisito parcial para obtenção do título de Doutor, ao Programa de Pós-Graduação em Biociências, da Universidade do Estado do Rio de Janeiro.

Orientadora: Prof.^a Dra. Cíntia Barros Santos-Rebouças Coorientadora: Prof.^a Dra. Adriana Helena de Oliveira Reis

> Rio de Janeiro 2022

CATALOGAÇÃO NA FONTE UERJ/REDE SIRIUS/BIBLIOTECA CB-A

Silva, Camilla Pereira da. Investigação de novos mecanismos genéticos e epigenéticos relacionados à síndrome do X frágil / Camilla Pereira da Silva. - 2022. 161 f.
Orientadora: Prof.ª Dra. Cíntia Barros Santos Rebouças Coorientadora: Prof.ª Dra. Adriana Helena de Oliveira Reis
Doutorado (Tese) - Universidade do Estado do Rio de Janeiro, Instituto de Biologia Roberto Alcântara Gomes. Pós-graduação em Biociências.
1. Genética humana – Teses. 2. Síndrome do Cromossomo do X Frágil.
3. Deficiência intelectual - Teses. 4. Patologia Molecular. I. Rebouças, Cíntia Barros Santos. II. Reis, Adriana Helena de Oliveira. III. Universidade do Estado do Rio de Janeiro. Instituto de Biologia Roberto Alcântara Gomes. IV. Título.

> Bibliotecária: Ana Rachel Fonseca de Oliveira CRB7/6382

Autorizo apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta tese, desde que citada a fonte.

Assinatura

Data

Camilla Pereira da Silva

Investigação de novos mecanismos genéticos e epigenéticos relacionados à síndrome do X-frágil

Tese apresentada, como requisito parcial para obtenção do título de Doutor, ao Programa de Pós-Graduação em Biociências, da Universidade do Estado do Rio de Janeiro.

Aprovada em 28 de abril de 2022.

Coorientadora: Prof.^a Dra. Adriana Helena de Oliveira Reis Instituto de Biologia Roberto Alcantara Gomes – UERJ

Banca Examinadora:

Prof.^a Dra. Cíntia Barros Santos-Rebouças (Orientadora) Instituto de Biologia Roberto Alcantara Gomes - UERJ

Prof.^a Dra. Dayse Aparecida da Silva Instituto de Biologia Roberto Alcantara Gomes - UERJ

Prof.^a Dra. Sheila Coelho Soares Lima Instituto Nacional do Câncer

Prof. Dr. Enrique Medina-Acosta Universidade Estadual do Norte Fluminense DEDICATÓRIA

Às crianças, por sua forma genuína de olhar o mundo.

AGRADECIMENTOS

A Deus por guiar e iluminar meu caminho.

À minha mãe pelo amor, dedicação e paciência em todos os momentos, principalmente nos mais difíceis. A você que me apoia e incentiva a seguir em frente, que me ensina a ser responsável e agir com honestidade e respeito. Devo a você o que sou hoje e todas as conquistas que eu venha a ter.

À prof.^a Cintia Santos-Rebouças pela orientação, dedicação, confiança, disponibilidade e pelos importantes ensinamentos ao longo desses anos de formação.

À prof.^a Adriana Reis pela coorientação e disposição em ajudar e esclarecer minhas dúvidas.

À prof.^a Sheila Coelho, do INCA, pelas importantes contribuições para o desenvolvimento deste trabalho. Ao Diego Camuzi por toda a ajuda, ensinamentos e análises de bioinformática essenciais para este trabalho. À Monique Lopes pela ajuda com o pirosequenciamento.

Ao prof. Miguel Moreira, do INCA, pela colaboração com os sequenciamentos.

Ao prof. Enrique Medina-Acosta (UENF) e ao prof. Filipe Brum Machado (UEMG), por disponibilizarem os primers e protocolo para amplificação de marcadores no cromossomo X e pela ajuda com a análise.

À prof.^a Márcia Pimentel pela orientação durante a minha iniciação científica e mestrado, pela confiança em abrir as portas do SERVGEN para mim e pelos conselhos ao longo da minha trajetória no grupo.

A toda a equipe do SERVGEN pelos muitos ensinamentos. À Andressa, Luma e Jussara pela paciência, pela disposição em ajudar e pelos importantes ensinamentos de coleta, bancada e respeito a cada um dos nossos pacientes. À Sharbilla, Beatriz e Gabriela, pela companhia, pelo agradável convívio, conversas e momentos de descontração. À Mateus e Maria Clara pelas conversas e pelos controles voluntários. Ao André pelas conversas, seja bem-vindo de volta. À Evelyn e Carol pelo apoio e ajuda, saudade do convívio diário com vocês. Um agradecimento especial ao André e à Andressa, que durante a pandemia, me ajudaram muito com apoio logístico e técnico essencial para que tudo funcionasse conforme o planejado e que eu conseguisse concluir em meio às dificuldades do momento.

Ao meu bando, Larissa, Aline, Sidney, Carol, Vanusca, Alexia e Pedro pela amizade e apoio. Obrigada pela alegria nas bagunças e farinhas que me dão força pra seguir adiante e por estarem sempre presentes e transformarem todos os momentos juntos em momentos especiais, até mesmo virtualmente. À Gisele, minha psicóloga, pelo apoio ao longo de todo esse período, principalmente nos momentos mais tensos.

Aos pacientes e seus responsáveis que aceitaram participar dessa pesquisa para que se possa aprimorar o conhecimento sobre a deficiência intelectual e, mais especificamente, sobre a síndrome do X Frágil.

A todos os professores, em especial aos professores e funcionários do PPGB e da UERJ.

Às agências de fomento FAPERJ, CNPq e CAPES pelo apoio financeiro. À CAPES e CNPq, pela concessão de minhas bolsas.

A todos que contribuíram e ajudaram no desenvolvimento deste trabalho.

Faça o teu melhor na condição que você tem, enquanto você não tem condições melhores [...].

Mário Sergio Cortella

RESUMO

SILVA, Camilla Pereira da. *Investigação de novos mecanismos genéticos e epigenéticos relacionados à síndrome do X-frágil.* 2022. 161f. Tese (Doutorado em Biociências) – Instituto de Biologia Roberto Alcantara Gomes, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2022.

A síndrome do X Frágil (SXF) é determinada por uma expansão instável da repetição de trinucleotídeos CGG na região 5' UTR do gene FMR1, sendo reconhecidas quatro categorias alélicas: normal (N, ≤44 CGGs), intermediária (I, 45-54 CGGs), pré-mutação (PM, 55-200 CGGs) e mutação completa (FM, >200 CGGs). Os indivíduos com a SXF apresentam a FM, que resulta no silenciamento do gene por hipermetilação das repetições CGG e do promotor do gene FMR1. A instabilidade nas repetições expandidas pode resultar em mosaicismo de tamanho, formado pela FM com outras categorias de alelos ou com deleções. Com a dificuldade de amplificar essa região, a investigação da SXF é complexa. Ainda, dados recentes apontaram para uma borda de metilação a montante das repetições, que é perdida em pacientes com a SXF, levando à disseminação de uma metilação de novo até o promotor do FMR1. Com isso, o presente estudo teve como objetivos principais reavaliar um grupo de 101 homens com SXF, visando a caracterização de padrões atípicos, analisar a presença da borda de metilação e identificar novos marcadores epigenéticos próximos às repetições. Através da metodologia de reação em cadeia da polimerase direcionada para repetições de trincas (TRP-PCR), a FM "pura" foi confirmada em 52,5% dos indivíduos. Já 27,7% dos pacientes apresentaram mosaicismo de tamanho, sendo 15,8% com PM+FM, 8,9% com N+FM e 3% com N+PM+FM. Seis pacientes (5,9%) ainda apresentavam uma deleção próxima às repetições, evidenciada por metodologias adicionais, podendo observar microhomologia entre os pontos de quebra em alguns deles. A análise de marcadores genéticos no cromossomo X não revelou nenhum caso de síndrome de Klinefelter (47,XXY) entre os casos atípicos de mosaicismo de tamanho. A literatura tem relatado casos de mosaicismo de tamanho com freguências de 12 a 41%. Através de pirosequenciamento de DNA genômico tratado com bissulfito de sódio, foi possível observar uma redução nos níveis de metilação nas CpGs 65-70 à montante do FMR1 em controles, ratificando a presença da borda de metilação, estando esta ausente em pacientes com FM, o que corrobora dados da literatura. Essa borda também parece ser perdida nos indivíduos com mosaicismo de tamanho. Como destaque deste estudo, identificamos uma CpG intrônica com um padrão de metilação inesperado no heatmap, estando hipermetilada em indivíduos controles e hipometilada em pacientes com a SXF. Estudos recentes sugerem que a metilação de elementos intragênicos estejam envolvidos na regulação da expressão gênica. Portanto, esta CpG tem grande potencial para ser usada como um novo marcador epigenético para a SXF, com um valor de corte de 69,5% e especificidade, sensibilidade e acurácia de 100%. Estes resultados promissores podem estimular o desenvolvimento de uma nova metodologia para o diagnóstico para a SXF.

Palavras-chave: Deficiência Intelectual. Cromossomo X. Síndrome do X-frágil. Perfis

atípicos. Marcadores epigenéticos. Diagnóstico molecular.

ABSTRACT

SILVA, Camilla Pereira da. *Investigation of novel genetic and epigenetic mechanisms related to fragile X syndrome.* 2022. 161f. Tese (Doutorado em Biociências) – Instituto de Biologia Roberto Alcantara Gomes, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2022.

Fragile X syndrome (SXF) is determined by an unstable expansion of CGG trinucleotides repetition in the 5'UTR region of the FMR1 gene, with four recognized allele categories: normal (N, ≤44 CGGs), intermediate (I, 45-54 CGGs), premutation (PM, 55-200 CGGs) and full mutation (FM, > 200 CGGs). SXF patients carry the FM allele, which results in gene silencing by hypermethylation of the CGG repeat and FMR1 gene promoter. CGG expansions instability can result in size mosaicism, formed by FM along with other allele categories or deletions. Due to the difficulty in amplifying this sequence, the molecular investigation of SXF is complex. Besides that, a recent study has suggested a methylation boundary upstream of the CGG repeat, which is lost in SXF patients and leads to a *de novo* methylation spreading to the *FMR1* promoter. Thus, the present study aimed to reevaluate a group of 101 SXF males, to characterize atypical profiles, analyze the presence of the methylation boundary and identify novel epigenetic markers close to the CGG repeat. Using Triplet Repeat Primed Polymerase Chain Reaction (TRP-PCR), a "pure" FM profile was confirmed in 52.5% of the SXF patients. 27.7% of the patients had size mosaicism: 15.8% with PM+FM, 8.9% with N+FM and 3% with N+PM+FM. Six patients (5.9%) had a deletion close to the repetitions, evidenced by additional methodologies, and microhomology could be observed between the breakpoints in some of them. The analysis of genetic markers in X chromosome revealed no cases of Klinefelter syndrome (47,XXY) among the atypical cases. Size mosaicism cases have been reported in the literature with frequencies of 12 to 41%. Using pyrosequencing after genomic DNA treatment with sodium bisulfite, it was possible to observe a reduction in methylation levels in CpGs 65-70 upstream of FMR1 in controls, ratifying the presence of the methylation boundary, absent in FM patients, which corroborates literature data. This boundary also seems to be lost in individuals with size mosaicism. As a highlight of this study, we identified an intronic CpG with an unexpected methylation pattern in the *heatmap*, being hypermethylated in controls and hypomethylated in FXS patients. Recent studies suggest that the methylation of intragenic elements is involved in the regulation of gene expression. Therefore, this CpG has great potential to be used as a new epigenetic marker for FXS, with a cutoff value of 69.5% and specificity, sensitivity and accuracy of 100%. These promising results may stimulate the development of a new diagnostic methodology for FXS.

Keywords: Intellectual Disability. X-chromosome. Fragile X syndrome. Atypical profiles. Epigenetic markers. Molecular Diagnosis.

LISTA DE FIGURAS

Figura 1 –	Genes no cromossomo X associados à DI 24		
Figura 2 –	Cariótipo parcial por bandeamento GTG, mostrando o cromossomo X		
	de indivíduo sem a SXF em comparação com um indivíduo com o sítio		
	frágil relacionado à SXF [fra(X)(q27.3)]	25	
Figura 3 –	Localização cromossômica e estrutura do gene FMR1, seu principal		
	transcrito e a proteína FMRP	28	
Figura 4 –	Quatro classes alélicas encontradas na região 5' UTR do gene FMR1		
	e suas consequências funcionais	29	
Figura 5 –	Perfis de metilação na região promotora do gene FMR1 e na região 5'		
	UTR em linhagens celulares de homens com e sem SXF e em		
	homens altamente funcionais	38	
Figura 6 –	Perfis de metilação na região promotora do gene FMR1 e no		
	segmento 5' à montante em linhagens celulares de mulheres		
	saudáveis, homens e mulheres pré-mutados, mulheres com a		
	mutação completa e pacientes com a síndrome de Turner (45,X)	40	
Figura 7 –	Esquema das etapas realizadas na primeira parte deste		
	estudo	45	
Figura 8 –	Esquema das etapas realizadas na segunda parte deste		
	estudo	46	
Figura 9 –	Fluxo da metodologia de amplificação por PCR e análise de metilação		
	na região de repetições CGG, utilizando o kit AmplideX® FMR1		
	mPCR (Asuragen, Inc.)	50	
Figura 10 -	-Princípios da metodologia de TRP-PCR, utilizando o kit AmplideX®		
	FMR1 mPCR (Asuragen, Inc.)	51	
Figura 11 -	-Fluxograma de análise dos dados (tamanho de fragmentos e		
	metilação) da região de repetições CGG no gene FMR1, através do		
	programa GeneMarker v.2.6.3 (SoftGenetics)	52	
Figura 12 -	-Princípio da técnica de pirosequenciamento	68	
Figura 13 -	-Cálculo de hipotenusa a partir de uma curva Roc	70	

Figura 15 – Eletroferograma do sequenciamento referente ao paciente P928/04.... 81

Figura 24 – Box plot com os dados de metilação referentes à primeira CpG analisada no intron 2 (sonda cg22417678)...... 114

Figura 25 – Curva ROC com os dados de metilação referentes à primeira CpG analisada no intron 2 (sonda cg22417678), por pirosequenciamento.... 115

Figura 29 – Curva ROC com os dados de metilação referentes à primeira CpG analisada no intron 2 (sonda cg22417678), por pirosequenciamento,

Figura 30 – incluindo os casos P4210/14 1	127
Resultados da amplificação com o kit AmplideX® FMR1 mPCR de	
amostras de DNA do probando P537/00, ilustrando o padrão onda	
observado na faixa de mutação completa	128

LISTA DE QUADROS

Quadro 1 – Probabilidade de surgimento do alelo de mutação completa no ge		
	FMR1 na prole, de acordo com a extensão das repetições do alelo	
	de pré-mutação materna	31
Quadro 2 –	Contrações observadas no alelo transmitido por vias materna e	
	paterna, de acordo com Nolin e colaboradores (2019)	33
Quadro 3 –	Condições de ciclagem utilizadas no sequenciamento	57
Quadro 4 –	Oligonucleotídeos utilizados nos ensaios de PCR das regiões de	
	interesse para análise de marcadores	59
Quadro 5 –	Condições utilizadas nos ensaios de PCR para amplificação dos	
	fragmentos para análise de marcadores	59
Quadro 6 –	Condições de ciclagem utilizadas nos ensaios de PCR para	
	amplificação dos fragmentos para análise de metilação	59
Quadro 7 –	Sondas de metilação no gene FMR1, utilizadas nos ensaios de	
	array, cujos dados estavam disponíveis publicamente no GEO	
	database	63
Quadro 8 –	Condições utilizadas nos ensaios de PCR para amplificação dos	
	fragmentos para análise de metilação	65
Quadro 9 –	Condições de ciclagem utilizadas nos ensaios de PCR para	
	amplificação dos fragmentos para análise de metilação	66
Quadro 10 -	-Matriz de confusão	70

LISTA DE TABELAS

Tabela 1 –	Dados gerais sobre a amostra de indivíduos estudada		
Tabela 2 –	Dados genotípicos sobre a amostra de indivíduos estudada,		
	avaliados por TRP-PCR	72	
Tabela 3 –	Estimativa aproximada da extensão dos fragmentos observados		
	neste trabalho, sugerindo a presença de deleção	80	
Tabela 4 –	Localização genômica e extensão das deleções identificadas neste		
	estudo e por Gonçalves <i>et al</i> ., 2016	82	
Tabela 5 –	Valores de metilação nas CpGs referentes à borda de metilação na		
	amostra de indivíduos estudada	97	
Tabela 6 –	Valores de metilação nas CpGs referentes à sonda cg22417678 e na		
	CpG à jusante na amostra de indivíduos estudada	111	

LISTA DE ABREVIATURAS E SIGLAS

5-azadC	5-azadesoxicitidina
5' UTR	Região 5' não traduzida <i>(5' Untranslated region)</i>
A	Adenina
APS	Adenosina 5'fosfosulfato
ATP	Trifosfato de adenosina
С	Citosina
chrX	Cromossomo X
CNS/MS	Conselho Nacional de Saúde do Ministério da Saúde
СТ	Controle
del	Deleção
DI	Deficiência intelectual
DILX	Deficiência intelectual ligada ao cromossomo X
DILX-NS	Deficiência intelectual ligada ao cromossomo X não sindrômica
DILX-S	Deficiência intelectual ligada ao cromossomo X sindrômica
DI-NS	Deficiência intelectual não sindrômica
DI-S	Deficiência intelectual sindrômica
DMSO	Dimetilsulfóxido
dNTPs	Desoxirribonucleotídeos Fosfatados
DNA	Ácido desoxirribonucléico
EDTA	Ácido etilenodiamino tetra-acético
FREE1	Elementos Epigenéticos Relacionados ao X-Frágil 1
FREE2	Elementos Epigenéticos Relacionados ao X-Frágil 2
FM	Alelo completamente mutado (Full mutation allele)
FMR1	Gene Fragile X Mental Retardation 1
FMRP	Proteína <u>F</u> ragile X <u>M</u> ental <u>R</u> etardation <u>P</u> rotein
FXPOI	Insuficiência Ovariana Prematura Associada ao X Frágil
FXTAS	Síndrome de Tremor e Ataxia Associada ao X Frágil
G	Guanina
gDNA	DNA genômico
HCI	Ácido clorídrico

HFMs	Homens altamente funcionais (High functioning males)		
HTT	Gene da huntingtina		
HUGG/UNIRIO	Hospital Universitário Gaffreé e Guinle da Universidade Federal		
	do Estado do Rio de Janeiro		
HUPE/UERJ	Hospital Universitário Pedro Ernesto da Universidade do Estado		
	do Rio de Janeiro		
I	Alelo intermediário		
IBGE	Instituto Brasileiro de Geografia e Estatística		
ICX	Inativação do cromossomo X		
IEDE	Instituto Estadual de Diabetes e Endocrinologia Luiz Capriglione		
TF	Fator de transcrição (transcriptional factor)		
IFF/FIOCRUZ	Instituto Fernandes Figueira da Fundação Oswaldo Cruz		
INCa	Instituto Nacional de Câncer		
IPPMG/UFRJ	Instituto de Puericultura e Pediatria Martagão Gesteira da		
	Universidade Federal do Rio de Janeiro		
MAF	Frequência do alelo menos frequente (Minor Allele Frequency)		
MgCl ₂	Cloreto de magnésio		
mPCR	PCR de metilação		
Ν	Alelo normal		
NLS	Domínio de sinal de localização nuclear (nuclear localization		
	signal)		
NES	Domínio do sinal de exportação nuclear (nuclear export signal)		
PBMCs	Células mononucleares de sangue periférico (peripheral blood		
	mononuclear cells)		
PCR	Reação em cadeia da polimerase (Polimerase Chain Reaction)		
PM	Alelo de pré-mutação		
PPi	Pirofosfato		
RNA	Ácido ribonucléico		
RNAi	RNA de interferência		
RNAm / mRNA	RNA mensageiro		
SERVGEN/UERJ	Serviço de Genética Humana da Universidade do Estado do Rio		
	de Janeiro		
SNV	Variante de nucleotídeo único (Single nucleotide variant)		

SXF	Síndrome do X Frágil
Т	Timina
Taq	Enzima DNA polimerase extraída de Thermus Aquaticus
ТВЕ	Tris, ácido bórico e EDTA
TE	Tris, HCI e EDTA
TRP-PCR	Reação em cadeia da polymerase direcionada para repetições
	de trincas (Triplet Repeat Primed Polymerase Chain Reaction)
UENF	Universidade Estadual do Norte Fluminense Darcy Ribeiro
UNIGRANRIO	Universidade do Grande Rio

LISTA DE SÍMBOLOS

- % Porcentagem
- < Menor que
- > Maior que
- ≥ Maior ou igual
- ± Mais ou menos
- ∞ Infinito
- + Soma
- Subtração
- × Multiplicação
- X Vezes
- V Volts
- Marca registrada
- ™ trade mark
- KDa KiloDalton
- pb Pares de bases
- kb Kilobase
- °C Graus Celsius
- H₂O Molécula da água
- mL Mililitro
- µL Microlitro
- mM Milimolar
- µM Micromolar
- ng Nanograma
- min Minutos
- seg Segundos
- U Unidades

SUMÁRIO

	INTRODUÇÃO	20
1	OBJETIVOS	42
1.1	Geral	42
1.2	Específicos	42
2	METODOLOGIA	43
2.1	Pacientes com SXF	43
2.2	Coleta de material biológico	47
2.3	Extração de DNA	47
2.4	Estimativa de integridade e concentração do DNA	48
2.5	Amplificação e análise da metilação nas repetições CGG	49
2.5.1	Reação em cadeia da polimerase (PCR)	49
2.6	Análise de deleções próximas à região de repetições CGG	54
2.6.1	Reação em cadeia da polimerase (PCR)	54
2.6.2	Avaliação da qualidade dos fragmentos amplificados e do rendimento dos	
	ensaios de PCR	55
2.6.3	Purificação dos produtos da PCR	56
2.6.4	Reação de sequenciamento de Sanger	56
2.7	Análise de marcadores nos cromossomos X e Y	58
2.7.1	Reação em cadeia da polimerase (PCR)	58
2.7.2	Avaliação da qualidade dos fragmentos amplificados e do rendimento dos	
	ensaios de PCR	60
2.7.3	Eletroforese capilar	60
2.8	Análise por ferramentas computacionais e produção de heatmaps	61
2.9	Análise de metilação da região da borda de metilação e de parte do	
	intron 2	64
2.9.1	Conversão do DNA genômico com bissulfito de sódio	64
2.9.2	Reação em cadeia da polimerase (PCR)	65
2.9.3	Avaliação da qualidade dos fragmentos amplificados e do rendimento dos	
	ensaios de PCR	66
2.9.4	Pirosequenciamento	67

2.10	Análise por Bioinformática e produção de curvas ROC	69
3	RESULTADOS E DISCUSSÃO	71
3.1	Casuística	71
3.2	Análise para caracterização de perfis mutacionais atípicos	71
3.2.1	Identificação de deleções próximas às repetições CGG	79
3.2.2	Análise para descartar a presença de um cromossomo X extra	82
3.2.3	Casos atípicos na população	85
3.3	Análise de marcadores epigenéticos por Bioinformática	92
3.4	Quantificação de metilação da borda por pirosequenciamento	96
3.5	Análise de quantificação de metilação na posição cg22417678 por	
	pirosequenciamento	109
3.6	Caso P4210/14	118
	CONCLUSÕES	135
	REFERÊNCIAS	136
	APÊNDICE A - Dados de perfis mutacionais da amostra de pacientes e	
	familiares incluídos nesse estudo	151
	APÊNDICE B - Fotodocumentação da eletroforese realizada com os	
	produtos do ensaio de PCR desenvolvido anteriormente pelo grupo	158
	APÊNDICE C – Dados de marcadores nos cromossomos X e Y, internos	
	ao gene SHOX	159
	ANEXO A - Aprovação do projeto pelo Certificado de Apresentação para	
	Apreciação Ética (CAAE)	161
	ANEXO B – Termo de consentimento livre e esclarecido	162

INTRODUÇÃO

A deficiência intelectual (DI) é caracterizada por um funcionamento intelectual significativamente abaixo da média, identificado antes dos 18 anos de idade, com limitações em duas ou mais áreas das habilidades conceituais, sociais e práticas: comunicação (linguagem, leitura, escrita), interação social, empatia, julgamento social, cuidados pessoais, habilidades acadêmicas, de seguir regras e fazer amigos, organização, lazer, trabalho, entre outras (*American Psychiatric Association*, 2017). A prevalência da DI varia entre os estudos, sendo estimada em 2 a 3% da população mundial, constituindo, portanto, um dos mais importantes problemas de saúde pública (Chiurazzi *et al.*, 2008; Ricceri *et al.*, 2014). No Brasil, a prevalência da DI autodeclarada foi de 1,2%, de acordo com a Pesquisa Nacional de Saúde 2019 (IBGE, 2021).

A DI é etiologicamente heterogênea, estando associada a fatores ambientais e genéticos (Chelly et al., 2006). Fatores ambientais que podem causar danos no desenvolvimento neurológico incluem diabetes, hipertensão (pré-eclâmpsia e eclampsia), epilepsia e asma maternas; exposições pré-natais a compostos neurotóxicos; síndrome alcoólica fetal; infecções pré, peri e pós-natais; deficiências nutricionais durante a gestação; complicações de partos prematuros; asfixia peri e pós-natal ou outros traumas e lesões cerebrais (Huang et al., 2016; Kvarnung & Nordgren, 2017). Já dentre as causas genéticas mais comuns estão desde alterações cromossômicas, tais como aneuploidias e alterações estruturais microscopicamente visíveis (inserções, duplicações, deleções e translocações), até variantes monogênicas, variações no número de cópias gênicas submicroscópicas (microdeleções, microduplicações) e alterações epigenéticas (metilação de DNA, modificações pós-traducionais de histonas e regulação por microRNAs, entre outros) (Armatas, 2009; Kvarnung & Nordgren, 2017). Além disso, os aspectos genéticos contribuem para cerca de 50% dos casos, enquanto 25% dos casos estão associados a componentes ambientais e 25% permanecem com causa desconhecida (Marrus & Hall, 2017).

A DI pode ser classificada como branda, moderada e severa, sendo assim categorizada de acordo com a combinação dos níveis de habilidades intelectual e adaptativa com a dependência pessoal necessária (*American Psychiatric*

Association, 2017). A gravidade das manifestações clínicas está também relacionada com o fator etiológico associado à DI. Assim, fatores genéticos, como os desequilíbrios cromossômicos e as síndromes genéticas, e ambientais, como infecções e malformações congênitas, se apresentam em casos mais graves. Por outro lado, causas ambientais mais comuns, como deficiências nutricionais maternas, possuem maior influência em casos de DI branda (Patel *et al.*, 2010; Chiurazzi & Pirozzi, 2016). A influência desses fatores varia muito entre os diferentes países e depende do estilo de vida materno e da qualidade dos cuidados de saúde (Chiurazzi & Pirozzi, 2016). Sendo assim, essa heterogeneidade abrangente dificulta e restringe a identificação das causas envolvidas com a DI (Chelly *et al.*, 2006).

Crianças com DI podem apresentar um amplo espectro de sintomas e sinais clínicos iniciais, dependendo da causa e da severidade da condição (Patel *et al.*, 2010). Dessa forma, crianças com DI mais severa geralmente apresentam sinais mais precocemente reconhecidos, enquanto casos de DI mais branda geralmente se apresentam como um atraso global de desenvolvimento e dificuldade de aprendizado e, com isso, são diagnosticados relativamente mais tarde (Patel *et al.*, 2010). Em geral, os pais relatam alguns dos sinais mais comuns da DI quando observam que seus filhos demoram a sentar, engatinhar ou andar, demoram a falar ou apresentam alterações da fala, têm dificuldades de memória ou compreensão do resultado de suas ações ou em solucionar problemas (Patel *et al.*, 2010; *Centers for Disease Control and Prevention*, 2019).

O diagnóstico clínico é realizado de acordo com os sintomas comportamentais e os marcos de desenvolvimento, seguido de exames físicos, neurológicos, de dismorfologia e de imagem (Patel *et al.*, 2010). Em muitos casos de DI branda, não há uma causa específica clara e, com isso, não há um consenso em relação à necessidade de se estabelecer a etiologia do diagnóstico (Patel *et al.*, 2010). Entretanto, conforme o aumento da severidade da condição, a busca e a possibilidade de se determinar a causa específica também aumenta, visto que a DI moderada e grave podem estar mais relacionadas a síndromes genéticas (Patel *et al.*, 2010).

Clinicamente, a DI pode ser dividida em duas formas: DI sindrômica (DI-S) e não sindrômica (DI-NS) (Kaufman *et al.*, 2010; Chiurazzi & Pirozzi, 2016). Na forma sindrômica, os indivíduos apresentam outras características além da DI, com aspectos clínicos radiológicos, metabólicos ou biológicos associados à condição

(Kaufman *et al.*, 2010; Chiurazzi & Pirozzi, 2016). Por outro lado, as formas não sindrômicas têm sido definidas pela presença da DI como única característica clínica (Kaufman *et al.*, 2010). Entretanto, autismo, epilepsia e déficits neuromusculares são comuns na DI não sindrômica (Chiurazzi & Pirozzi, 2016). Além disso, é possível observar que para alguns genes, algumas variantes podem causar uma síndrome, enquanto outras podem levar a uma forma não sindrômica da DI (Chiurazzi & Pirozzi, 2016). Dessa forma, distinguir as formas DI-S e DI-NS em nível molecular é um grande desafio.

O fato da DI apresentar uma prevalência cerca de 30% maior em indivíduos do sexo masculino levou à suspeita de que alterações em genes localizados no cromossomo X teriam grande contribuição para esta condição (Leonard & Wen, 2002; Patel et al., 2010; Lubs et al., 2012). Isso se dá em função das alterações em genes do cromossomo X em homens, hemizigotos para esses genes, serem mais aparentes fenotipicamente do que em mulheres. Nestas, alterações recessivas podem ser mascaradas pela heterozigose e ainda pode haver compensação de dose favorável à expressão do alelo normal pela inativação desviada do cromossomo X (Lyon, 1961; Stevenson & Schwartz, 2009). De fato, estudos posteriores descreveram casos familiares de DI com padrão de herança ligado ao cromossomo X, reforçando essa hipótese (Lubs et al., 2012; Stevenson et al., 2012). Além disso, apesar desse cromossomo representar apenas 5% do genoma humano, cerca de 15% de seus genes são reconhecidos por estarem associados à DI e atuarem em diversas funções essenciais do Sistema Nervoso Central (SNC), como a plasticidade neuronal, aprendizado, memória e cognição (Ropers, 2006; Neri et al., 2018).

A DI ligada ao cromossomo X (DILX) é definida pela presença de DI ou dificuldade de aprendizado, cuja causa genética está localizada ou há suspeita de que se localize ao longo do cromossomo X (Gécz *et al.*, 2009). A DILX representa aproximadamente 5 a 10% de todas as formas de DI em homens (Bassani *et al.*, 2013).

Clinicamente, a DILX também pode ser dividida em DLIX sindrômica (DILX-S) e DILX não sindrômica (DILX-NS), seguindo os mesmos padrões para DI-S e DI-NS (Mandel & Chelly, 2004; *Greenwood Genetic Center*, 2021). Ou seja, a DILX-S é baseada em características físicas, metabólicas e/ou comportamentais típicas em adição à DI (Gécz *et al.*, 2009), enquanto que na DILX-NS não há sinais dismórficos, metabólicos ou neuromusculares consistentes (Gécz *et al.*, 2009). Atualmente, já foram identificados 162 genes envolvidos na DILX (Greenwood Genetic Center, 2021) (Figura 1). No entanto, acredita-se que grande parte dos mecanismos moleculares subjacentes à DILX ainda seja desconhecida, o que contribui para uma alta taxa de subdiagnóstico (Lubs *et al.*, 2012).



Legenda: (A) genes associados à DILX-NS; (B) genes associados à DILX-S.

Fonte: adaptada de Greenwood Genetic Center, 2021.

Síndrome do X Frágil (SXF)

A forma mais frequente de DILX sindrômica é a síndrome do X Frágil (SXF; OMIM #300624), descrita pela primeira vez em 1943, por Martin e Bell (Martin & Bell, 1943; Saldarriaga *et al.*, 2014). Quase três décadas depois, Lubs descreveu um caso semelhante em uma família com segregação ligada ao cromossomo X (Lubs, 1969). De acordo com Lubs (1969), o cromossomo X dos indivíduos afetados em metáfase possuía um sítio frágil na porção terminal do braço longo do cromossomo X (Figura 2). A condição passou então a ser conhecida como Síndrome do X Frágil. Atualmente, SXF é a segunda causa de DI mais comum (2,4% dos casos de DI), depois da síndrome de Down, e a mais prevalente em homens (Hagerman & Hagerman, 2002; Hunter *et al.*, 2014). A prevalência da SXF na população em geral é estimada em 1 em 4000 homens e 1 em 6000 mulheres (Crawford *et al.*, 2001; *National Fragile X Foundation*, 2021).

Figura 2 – Cariótipo parcial por bandeamento GTG, mostrando o cromossomo X de indivíduo sem a SXF em comparação com um indivíduo com o sítio frágil relacionado à SXF [fra(X)(q27.3)]



Fonte: adaptada de Floriani et al., 2017.

Bebês com a SXF geralmente apresentam hipotonia, com dificuldade de sucção inicial e frequente regurgitação (Hagerman *et al.*, 2010; 2017). Por volta dos dois anos, diversas crianças com essa condição apresentam atraso no desenvolvimento da linguagem, hiperatividade, ansiedade e reatividade sensorial exacerbada (Hagerman *et al.*, 2002; 2017; Berry-Kravis *et al.*, 2010; Cordeiro *et al.*, 2011; Hogan *et al.*, 2017). Geralmente, o diagnóstico clínico da síndrome acontece quando a criança desenvolve atraso ou ausência da fala (Bailey *et al.*, 2009). Outros sinais clínicos que podem estar presentes em crianças pequenas são: atraso motor, hábito de agitar e morder as mãos, pouco contato ocular, otite média frequente e irritabilidade (Sullivan *et al.*, 2006; Cordeiro *et al.*, 2011).

Meninos com SXF apresentam um fenótipo clássico caracterizado por face alongada, orelhas grandes e salientes, fronte proeminente e macroorquidismo póspuberal (Hagerman et al., 2002; Saldarriaga et al., 2014; National Fragile X Foundation, 2021). Os pacientes são frequentemente acometidos por problemas mentais relacionados, como ansiedade e depressão, presentes em um a mais de dois terços dos indivíduos afetados (Bailey et al., 2009; Ciaccio et al., 2017). Além disso, observa-se também problemas de atenção e hiperatividade em mais de 80% dos casos e comportamento autodestrutivo e agressão em mais de 50% de crianças e adolescentes com a síndrome (Sullivan et al., 2006; Hagerman et al., 2009; Ciaccio et al., 2017). Ainda, em muitos casos, crianças e adultos acometidos ainda podem demonstrar deficiência em habilidades funcionais (como em se alimentar, se vestir e se comunicar) (Bailey et al., 2009). Além disso, entre 25 e 50% das crianças e adolescentes com a SXF se enquadram nos critérios de diagnóstico dos transtornos do espectro autista, com habilidades adaptativas reduzidas e desenvolvimento mais lento da linguagem (Philofsky et al., 2004; Ciaccio et al., 2017).

Gene FMR1

O gene *Fragile X Mental Retardation 1 (FMR1*, OMIM #309550) é constituído por 17 éxons, compreendendo aproximadamente 38 kb de extensão, e está localizado em Xq27.3 (Verkerk *et al.*, 1991; Eichler *et al.*, 1993; Handt *et al.*, 2014). Em termos moleculares, a SXF se caracteriza por uma mutação dinâmica, em consequência da expansão instável de uma sequência de trinucleotídeos CGG repetida em *tandem* na região 5' não traduzida (5' UTR) do gene *FMR1*, próxima à região promotora (Oberlé *et al.*, 1991; Cleary & Pearson, 2005; Saldarriaga *et al.*, 2014; Hagerman *et al.*, 2017; Nolin *et al.*, 2019). As mutações dinâmicas ocorrem em regiões do DNA onde há sequências repetidas *em tandem* (Cleary & Pearson, 2005), como a encontrada na região 5' UTR do gene *FMR1*. Normalmente, o número de repetições é polimórfico na população e transmitido de maneira estável ao longo das gerações (Cleary & Pearson, 2005). Entretanto, quando ocorre uma expansão que ultrapassa determinado limiar, elas se tornam instáveis, podendo aumentar de comprimento através de sucessivas gerações até que o número de repetições atinja um tamanho patogênico (Cleary & Pearson, 2005).

Com a variabilidade polimórfica nesta região, podem ser reconhecidas quatro categorias alélicas distintas: (a) normal (até 44 repetições CGG), com os alelos mais comuns nessa faixa sendo os de 29 e 30 repetições e transmissão estável; (b) alelos intermediários (45 a 54 repetições), sem capacidade de se expandir para a mutação completa, mas sendo levemente instáveis e considerados como precursores dos alelos pré-mutados; (c) pré-mutação (55 a 200 CGGs), sendo altamente instáveis quando herdados por via materna; e (d) mutação completa (acima de 200 CGGs), derivados de alelos pré-mutados herdados da mãe (Fu *et al.*, 1991; Nolin *et al.*, 2003; Fernandez-Carvajal *et al.*, 2009; Saldarriaga *et al.*, 2014; Spector *et al.*, 2021).

Nos indivíduos afetados pela SXF, o número de repetições CGG está na faixa de mutação completa, ou seja, acima de 200 CGGs (Wang *et al.*, 2012; Spector *et al.*, 2021). A expansão na faixa de mutação completa tem como consequência o silenciamento epigenético do gene *FMR1*, através da hipermetilação das repetições CGG e da região promotora do gene que resulta em perda completa ou quase completa de seu produto final, a proteína FMRP (Figura 3) (Wang *et al.*,

2012; Ciaccio *et al.*, 2017). Por outro lado, níveis elevados do RNAm transcrito pelo gene *FMR1* são observados em portadores da pré-mutação, o que leva a uma toxicidade em células neurais (Raske & Hagerman, 2009). No entanto, os níveis da proteína FMRP são mais baixos em portadores da pré-mutação do que em indivíduos normais (Figura 4) (Santoro *et al.*, 2012).

Figura 3 – Localização cromossômica e estrutura do gene *FMR1*, seu principal transcrito e a proteína FMRP



Legenda: (a) Representação do cromossomo X. Seta amarela aponta para a posição do sítio frágil e gene *FMR1*; (b) Esquema do gene *FMR1* e seus 17 exons, com os exons codificantes em verde escuro, as regiões não traduzidas em verde claro, e a região de repetições CGG em amarelo. (c) Principal transcrito de RNAm do gene *FMR1*, com todos os exons representados em azul claro e posicionados acima de seus aminoácidos correspondentes. (d) Proteína FMRP, seus principais domínios funcionais e abaixo, a extensão dos aminoácidos. Agenet 1 e 2: domínios Agenet de ligação à cromatina; NLS: sinal de localização nuclear, KH1 e KH2: domínios de homologia K 1 e 2; NES: sinal de exportação nuclear; RGG: box arginina-glicina-glicina de ligação a RNA; S500: serina fosforilada primária.

Fonte: adaptada de Santoro et al., 2012 e U.S. National Library of Medicine, 2020.



Figura 4 – Quatro classes alélicas encontradas na região 5' UTR do gene *FMR1* e suas consequências funcionais

Legenda: (a) Alelo normal (≤44 repetições) permite uma transcrição fisiológica apropriada do gene *FMR1* e tradução da proteína FMRP, (b) assim como o alelo intermediário (45-54 repetições). (c) Alelo de pré-mutação (55-200 repetições) causa um aumento significativo na transcrição do *FMR1*. No entanto, os níveis da proteína são mais baixos que em indivíduos saudáveis. Níveis mais altos do RNAm do gene *FMR1* estão relacionados com a Síndrome de Tremor e Ataxia Associada ao X Frágil (FXTAS) e a Insuficiência Ovariana Prematura Associada ao X Frágil (FXPOI), provavelmente devido ao efeito tóxico causado por pelo acúmulo destas moléculas em células neurais. (d) Alelo de mutação completa (>200 repetições) levando a mudanças epigenéticas nas repetições CGG e no promotor do *FMR1*, assim como ao silenciamento transcricional do gene. Os sintomas da SXF são causados pela falta da proteína FMRP. A classificação das categorias alélicas está de acordo com a *American College of Medical Genetics* (Monaghan *et al.*, 2013; Spector *et al.*, 2021).

Fonte: Adaptada de Santoro et al., 2012.

A proteína FMRP apresenta ligação seletiva a RNA e regula negativamente a tradução dos mRNAs a ela ligados, especialmente nas sinapses neuronais (Wang *et al.*, 2012; Noto *et al.*, 2016; Hagerman *et al.*, 2017). É produzida em vários tecidos, sendo expressa em maiores níveis no início do desenvolvimento embrionário (Kumari *et al.*, 2010). Assim, é uma proteína essencial para o desenvolvimento neuronal e intelectual (Kumari *et al.*, 2010). Além disso, um estudo identificou RNAs mensageiros prováveis de interagir com a FMRP provenientes de genes candidatos ao autismo (Darnell *et al.*, 2011).

A FMRP é uma proteína bastante conservada em mamíferos e sua isoforma principal possui 71 KDa (Blackwell *et al.*, 2010; Healy *et al.*, 2011; Yang *et al.*, 2020). Essa proteína interage com elementos de RNA, se ligando a eles através de sequências específicas mediadas pelos domínios funcionais do tipo KH e um do tipo RGG (Blackwell *et al.*, 2010; Healy *et al.*, 2011). Além disso, ela pode formar homodímeros que interagem com diversas proteínas citoplasmáticas e nucleares envolvidas no metabolismo de RNAm, incluindo RNA de interferência (RNAi) e edição de RNA (Caudy *et al.*, 2002; Ishizuka *et al.*, 2002; Bhogal *et al.*, 2011; Tabolacci *et al.*, 2016a).

A FMRP também apresenta a função de regular a tradução e a estabilidade de RNAm de diversos genes envolvendo os domínios do sinal de localização nuclear (NLS) e do sinal de exportação nuclear (NES), importantes para o transporte do núcleo para o citoplasma e do citoplasma para o núcleo (Brown *et al.*, 2001; Chen *et al.*, 2003; Darnell *et al.*, 2011; Saldarriaga *et al.*, 2014; Tabolacci *et al.*, 2016a). É capaz de regular a síntese de proteínas locais importantes para o desenvolvimento das espinhas dendríticas e na plasticidade sináptica, essenciais para o aprendizado e o desenvolvimento intelectual, através do mecanismo de RNAi (Antar & Bassell, 2003, Antar *et al.*, 2005; Healy *et al.*, 2011; Peprah *et al.*, 2012). Ela ainda tem o papel de remodelar o citoesqueleto através das regiões central e N-terminal, onde estão presentes dois domínios Agenet capazes de se ligar a resíduos de lisina trimetilados, e de formar grandes ribonucleopartículas citoplasmáticas contendo diversas proteínas e RNAs (Maurer-Stroh *et al.*, 2003; Bagni & Greenough, 2005; Adams-Cioaba *et al.*, 2010; Bagni *et al.*, 2012; De Rubeis *et al.*, 2012; Wang *et al.*, 2012).

A mutação completa surge através da expansão de uma pré-mutação materna (Healy *et al.*, 2011). Sendo assim, mães portadoras de pré-mutação em heterozigose podem ter filhos normais, portadores da pré-mutação ou da mutação completa (Healy *et al.*, 2011). De acordo com Nolin e colaboradores (2003), o risco de expansão de uma pré-mutação materna para a mutação completa nos filhos

depende da extensão das repetições CGG e a chance é próxima de 100% para mulheres portadoras de pré-mutação com mais de 99 repetições (Quadro 1).

Quadro 1 – Probabilidade de surgimento do alelo de mutação completa no gene *FMR1* na prole, de acordo com a extensão das repetições do alelo de pré-mutação materna

Número de repetições CGG do alelo materno	Probabilidade de expansão para mutação completa na prole
55-59	3,7%
60-69	5,3%
70-79	31,1%
80-89	57,8%
90-99	80,1%
100-109	100,0%
110-119	98,1%
120-129	97,2%
130-139	94,4%
140-149	100,0%
150-159	100,0%
160-169	100,0%
170-199	100,0%

Fonte: Adaptada de Nolin et al., 2003.

Homens portadores de pré-mutação terão apenas filhas portadoras da prémutação, visto que o homem transmite o cromossomo Y a seus filhos (Healy *et al.*, 2011; Saldarriaga *et al.*, 2014). O alelo da pré-mutação não tende a se expandir quando transmitida por via paterna, devido a uma pressão seletiva sobre os gametas masculinos (Saldarriaga *et al.*, 2014). Há apenas raros relatos de transmissão da mutação completa por via paterna (Zeesman *et al.*, 2004; Alvarez-Mora *et al.*, 2017; Nolin *et al.*, 2019). Apesar de homens com mutação completa geralmente não deixarem descendentes, devido às suas limitações cognitivas, estudos em esperma indicam ainda que as filhas herdariam a pré-mutação e não a mutação completa (Willems *et al.*, 1992; Reyniers *et al.*, 1993; Nolin *et al.*, 2019). Ainda, variantes pontuais ou *frameshift* no *FMR1* também podem acarretar em um fenótipo de SXF, embora representem menos de 1% dos casos já descritos (Handt *et al.*, 2014; Saldarriaga *et al.*, 2014).

Além do tamanho da repetição e do gênero do transmissor do alelo expandido, há mais um fator que contribui para a instabilidade nos alelos prémutados: a ausência de trinucleotídeos AGGs interrompendo as repetições CGG (Nolin et al., 2019). Estudos mostram que essas interrupções funcionam como um fator de proteção, diminuindo o risco de expansão entre gerações (Yrigollen et al., 2012; Nolin et al., 2013; 2015; 2019). No começo dos anos 90, análises revelaram que a estrutura do alelo na categoria normal mais comum inclui uma repetição AGG no 10º ou 11º trinucleotídeo CGG e outra próximo ao 20º CGG (Eichler et al., 1994; Gunter et al., 1998; Hirst et al., 1994; Kunst et al., 1996; Kunst & Warren, 1994; Snow et al., 1994). Entretanto, Eichler e colaboradores (1994) hipotetizaram que longas sequências de repetição sem interrupções em alelos pré-mutados e tratos com mais de cerca de 34 CGGs ininterruptas podem contribuir para a instabilidade em alelos menores também. De fato, estudos recentes demonstraram que alelos de pré-mutação sem AGGs representam um risco para a expansão para a mutação completa na geração seguinte, enquanto alelos que incluem esse trinucleotídeo estão associados com uma maior estabilidade da repetição entre gerações (Yrigollen et al., 2012; Nolin et al., 2013; Nolin et al., 2015; Villate et al., 2020).

Nolin e colaboradores (2019) mostraram ainda que, enquanto muitas transmissões instáveis levam a expansões, contrações para alelos menores também podem ser observadas, mesmo que mais raramente. Em sua pesquisa com mais de 5 mil transmissões, eles constataram que, entre os alelos pré-mutados, contrações transmitidas por via paterna são até 10 vezes mais comuns do que por via materna e, que a frequência de contrações paternas aumenta de acordo com o tamanho da repetição (Quadro 2) (Nolin *et al*, 2019). No entanto, todas as contrações transmitidas por pais pré-mutados resultaram em alelos menores ainda na faixa de pré-mutação, enquanto contrações maternas resultaram em alelos normais ou intermediários em 40% dos casos (Nolin *et al.*, 2019). Além disso, a perda de interrupções AGGs foi observada apenas em contrações de pré-mutação materna, ocorrendo em uma frequência relativamente alta (Nolin *et al.*, 2019). Sendo assim, os autores sugeriram que muitos alelos normais sem interrupções AGGs tenham origem em contrações de alelos pré-mutados maternos (Nolin *et al.*, 2019).

Número de repetições CGG do alelo parental	Contração do alelo materno	Contração do alelo paterno
<45	0,0%	0,08%
45-54	2,0%	0,0%
55-64	1,8%	4,5%
65-74	4,4%	9,4%
75-84	7,1%	26,3%
85-94	4,5%	37,0%
≥95	1,4%	70,8%

Quadro 2 – Contrações observadas no alelo transmitido por vias materna e paterna, de acordo com Nolin e colaboradores (2019)

Legenda: ≤44 CGGs: alelo normal; 45-54 CGGs: alelo intermediário; >55: alelo pré-mutado. Fonte: Adaptada de Nolin *et al.*, 2019.

Portadores da pré-mutação não apresentam o fenótipo clássico da SXF, mas podem exibir outros quadros clínicos, psiquiátricos e neurológicos (Saldarriaga *et al.*, 2014). Estes quadros, chamados de FRAXopatias, incluem a Insuficiência Ovariana Prematura Associada ao X Frágil (FXPOI, OMIM #311360), caracterizada por redução na função ovariana antes dos 40 anos (Sherman, 2000; Man *et al.*, 2017; Hunter *et al.*, 2019) e a Síndrome de Tremor e Ataxia Associada ao X Frágil (FXTAS, OMIM #300623), uma condição neurodegenerativa de manifestação tardia, mais prevalente em homens pré-mutados (Hagerman *et al.*, 2001; Hagerman & Hagerman, 2004; Hagerman *et al.*, 2018; Hall & Berry-Kravis, 2018; Hunter *et al.*, 2019). Além disso, apesar dos alelos intermediários serem considerados como fenotipicamente normais, esses alelos têm sido associados a necessidades especiais em crianças, parkinsonismo em adultos e FXPOI/FXTAS em alguns casos (Bodega *et al.*, 2006; Liu *et al.*, 2013; Hall *et al.*, 2012; 2020).

Perfis Atípicos na SXF

A instabilidade presente nas repetições CGGs expandidas pode gerar mosaicismo de tamanho, ou seja, pode resultar na presença de alelos com diferentes extensões (Nolin *et al.*, 1994; Pretto *et al.*, 2014; Jiraanont *et al.*, 2017).

Com isso, o alelo de mutação completa pode aparecer concomitantemente a alelos de pré-mutação em diferentes tipos celulares (Nolin *et al.*, 1994; Pretto *et al.*, 2014). Além disso, apesar de não serem muito comuns, mosaicismos de tamanho compreendendo alelos de mutação completa e um alelo normal têm sido descrito em diferentes casos de meninos com SXF (Nolin *et al.*, 1994; Todorov *et al.*, 2009; Bonarrigo *et al.*, 2014).

Há alguns anos, sugeriu-se que o mosaicismo de tamanho para a mutação completa e alelo normal poderia se originar durante os estágios iniciais do desenvolvimento, quando alelos pré-mutados transmitidos por via materna podem resultar em expansão para a mutação completa, assim como, menos frequentemente, em contração para o alelo normal (Zhao & Usdin, 2016). Na literatura, há ainda descrição de casos de mosaicismo incluindo alelos normais e intermediários (Jiraanont *et al.*, 2017). O mosaicismo de tamanho representa uma forma de variabilidade fenotípica, em que casos mosaicos apresentam valores de QI superiores aos pacientes com apenas a mutação completa, e pode ser observado em ambos os sexos, com uma incidência maior no sexo masculino (Schneider *et al.*, 2013; Ciaccio *et al.*, 2017).

Além disso, alguns indivíduos ainda exibem o que é chamado de mosaicismo de metilação, que ocorre quando algumas células apresentam alelos metilados e outras carregam alelos não metilados (com número de repetições CGGs variando entre pré-mutação e mutação completa) (Hagerman *et al.*, 1994; Genc *et al.*, 2000). Nesses casos, o nível de metilação pode variar, gerando mosaicismos intra ou inter-tecidos (Hagerman *et al.*, 1994; Genc *et al.*, 2000). Esses alelos não metilados são transcritos e, em muitos casos, superexpressos e potencialmente traduzidos na proteína FMRP (Tassone *et al.*, 2000; Ludwig *et al.*, 2014). No entanto, os níveis de expressão são inversamente relacionados com o número de repetições e os alelos, em particular os de pré-mutação maiores, mostram uma expressão reduzida, possivelmente devido ao déficit na eficiência de tradução da proteína FMRP (Primerano *et al.*, 2002; Allen *et al.*, 2005; Peprah *et al.*, 2010).

O espectro de mosaicismos na SXF ainda pode ser complementado pela presença de deleções no gene *FMR1*, concomitantes a alelos FM (Nolin *et al.*, 1994; Coffee *et al.*, 2008; Han *et al.*, 2006). Nos casos em que deleções menores envolvem a remoção da região promotora ou afetam o sítio de início de transcrição do *FMR1*, a proteína FMRP não será produzida e o indivíduo desenvolve a SXF
(Jiraanont *et al.*, 2017). Em alguns casos, as deleções estão limitadas à região entre o sítio de início da transcrição e o códon de iniciação da tradução e, sendo assim, esses alelos ainda apresentam a capacidade de produzir RNAm (de Graaff *et al.*, 1996; Grønskov *et al.*, 1997; Kenneson *et al.*, 2001).

Devido principalmente à dificuldade de se amplificar produtos repetitivos longos e regiões ricas em CG, a investigação molecular da SXF é complexa. Mesmo diante de metodologias de alto rendimento, a abordagem destas regiões repetitivas ainda representa um grande desafio. Através do uso combinado de diferentes abordagens metodológicas, nosso grupo demonstrou recentemente a existência de padrões mutacionais atípicos e subdetectados, envolvendo mosaicismo somático de mutação completa e deleção no gene FMR1 em indivíduos com a SXF (Gonçalves et al., 2016). Os dados apontaram que as deleções surgiram no início do desenvolvimento embrionário, foram causadas pela presença de microhomologia entre as regiões proximal e distal dos pontos de quebra e eram clinicamente indistinguíveis do padrão fenotípico clássico da SXF (Gonçalves et al., 2016). No entanto, não se sabe ainda que característica da arquitetura genômica faz com que uma mutação completa seja revertida para a deleção durante o início do desenvolvimento embrionário, enquanto outras não (Gonçalves et al., 2016). Além disso, o significado clínico e as implicações terapêuticas destes padrões atípicos ainda são desconhecidos, tendo em vista que em geral apenas a expansão das repetições CGG é interrogada (Gonçalves et al., 2016).

Epigenética do gene FMR1

A expansão das repetições CGG na faixa de mutação completa (>200 repetições) desencadeia uma cascata de eventos epigenéticos, com a propagação da metilação de DNA para a região promotora do gene *FMR1* (Oberlé *et al.*, 1991; Wang *et al.*, 2012; Jiraanont *et al.*, 2017; Kraan *et al.*, 2019), juntamente com modificações de histonas ativas para repressivas (Coffee *et al.*, 1999; 2002; Pietrobono *et al.*, 2005; Tabolacci *et al.*, 2005; 2008a; Kumari & Usdin, 2010). Acredita-se que o evento de silenciamento do gene *FMR1* seja mediado pela formação de estruturas secundárias de DNA e/ou RNA, que afetam o equilíbrio entre

os fatores que favorecem uma configuração de cromatina ativa e os fatores que contribuem para uma configuração de cromatina fechada (Usdin & Kumari, 2015; Kraan *et al.*, 2019).

Em indivíduos não acometidos pela SXF e com alelos de tamanho normal, o promotor se encontra não metilado e a cromatina associada é enriquecida com marcadores de cromatina ativos (como histonas H3 e H4 acetiladas e histona H3 trimetilada na lisina 4) e apresenta poucos marcadores de cromatina inibitórios (como histona H3 trimetilada na lisina 9), permitindo o acesso de fatores de transcrição ao promotor do gene (Kraan *et al.*, 2019). Já em indivíduos com a SXF, a metilação de DNA atinge o promotor e as repetições CGG e a cromatina adota uma conformação compactada enriquecida com marcadores de cromatina inativa (como histona H3 hipometilada na lisina 4, histona H3 dimetilada e trimetilada na lisina 9 e na lisina 27, histona H4 trimetilada na lisina 20), impossibilitando o acesso ao promotor (Kumari & Usdin, 2010; Kraan *et al.*, 2019).

A metilação do alelo de mutação completa ocorre nos estágios iniciais do desenvolvimento embrionário, aproximadamente na 11^a semana de gestação, e pode apresentar um papel na estabilidade das repetições expandidas (Willemsen *et al.*, 2002). Nas células germinativas iniciais de fetos com mutação completa, as repetições expandidas não estão metiladas, enquanto que nas vilosidades coriônicas desses fetos são metiladas em um grau crescente à medida que o desenvolvimento progride, ou seja, há uma relação com a diferenciação celular (Malter *et al.*, 1997). Além disso, uma ilha CpG, isto é, uma região rica em CpGs, localizada a aproximadamente 250 pares de bases à montante das sequências repetidas também se mostra metilada na presença da expansão (Coffee *et al.*, 1999).

Ainda, RNAs não codificantes (ncRNA) participam da fisiopatologia da SXF através de múltiplos aspectos (Zhou *et al.*, 2019). Os microRNAs (miRNA), por exemplo, facilitam principalmente o papel do FMRP como repressor traducional, moldando a plasticidade sináptica, controlando o crescimento do axônio e regulando o nível de FMRP endógeno (Zhou *et al.*, 2019). Entretanto, a forma precisa como a adição de repetições resulta na indução de marcadores de repressão epigenética e a sequência de eventos envolvidos ainda permanecem por serem elucidadas e, assim, o gatilho para o silenciamento do gene *FMR1* segue sendo uma questão não resolvida (Usdin *et al.*, 2014; Epsztejn-Litman & Eiges, 2019).

Borda de metilação no gene FMR1

Mais um ponto que também não está claro é o porquê de um número de repetições de 200 ser requerido para que o processo de silenciamento epigenético seja engatilhado. Neste sentido, dados recentes da literatura têm apontado para a presença de uma borda de metilação, localizada 650-800 nucleotídeos e 65-70 CpGs à montante das repetições CGG e detectada por sequenciamento de DNA tratado com bisulfito de sódio (Naumann et al., 2009; 2014). Essa zona de transição, presente em indivíduos normais, separa um segmento de genoma fortemente ou completamente metilado de uma região livre de metilação, a qual abrange o promotor do gene FMR1 (Naumann et al., 2009; 2014). Já em pacientes com a SXF, com a instabilidade provocada pelo alto número de repetições CGG, essa borda é perdida. Com isso, uma metilação de novo se dissemina até a região promotora do gene, levando à sua inativação (Naumann et al., 2009; 2014). Em indivíduos do sexo masculino altamente funcionais (HFMs; high functioning males), ou seja, que apresentam expansão CGG acima de 200 sem metilação na região de repetição e sem comprometimento intelectual, a borda de metilação se apresenta inalterada (Figura 5) (Naumann et al., 2014; Tabolacci et al., 2016b).

Figura 5 – Perfis de metilação na região promotora do gene *FMR1* e na região 5' UTR em linhagens celulares de homens com e sem SXF e em homens altamente funcionais



Legenda: (a) Ideograma do cromossomo X humano. (b) Mapa parcial mostra os primeiros 10 exons (barras verticais) e introns do gene *FMR1*, além do segmento do genoma à montante, incluindo as repetições CGG. (c) Mapa da região 5' a montante do gene *FMR1*. O esquema apresenta todos os dinucleotídeos CpGs (1 a 104) na região: a seta indica o sítio de início da transcrição. O promotor e a origem da replicação de DNA também estão indicados. A borda entre os dinucleotídeos CpG metilados e não metilados está marcada pelos símbolos • e o, respectivamente, e está designada por uma barra vermelha. Região da borda de metilação CpG 5' à montante do promotor do gene *FMR1* em DNA extraído de (d) fibroblastos transformados pelo gene da telomerase de homens sem SXF, (e) células mononucleares de sangue periférico (PBMCs) não transformadas de paciente com SXF e (f) fibroblastos transformados pelo gene da telomerase de homens HFM (*high functioning male*), ou seja, homens com expansão CGG acima de 200 sem metilação na região de repetição e sem comprometimento intelectual. Os números 20-104 designam os dinucleotídeos CpG na região: ■ metilados e □ não metilados.

Fonte: Adaptada de Naumann et al., 2009; 2014.

Em amostras de mulheres saudáveis, a borda de metilação estava aparente na CpG 69 ou 70 em metade das moléculas de DNA (Figura 6) (Naumann *et al.*, 2009). Na região entre os pares CpG 20 e 68, alguns cromossomos X analisados estavam metilados, enquanto outros permaneceram não metilados, com uma razão próxima a 1:1 (Naumann *et al.*, 2009). Esse padrão de mosaicismo provavelmente reflete a diferença nos níveis de metilação nos dois cromossomos X femininos, sendo um altamente metilado e outro não metilado ou hipometilado, em função da compensação de dose atrelada à inativação de um dos cromossomos X (Naumann *et al.*, 2009).

Os achados de Naumann e colaboradores (2014) ainda revelaram que a borda se encontra estável em genomas com risco de expansão das repetições CGG, tanto em homens quanto em mulheres portadores da pré-mutação (Figura 6). A metilação em cromossomos pré-mutados femininos exibe um perfil composto pelos dois padrões alélicos (Naumann *et al.*, 2014). Em mulheres portadoras da mutação completa, a região está 59% metilada e a borda ainda encontra-se vagamente discernível, enquanto que em portadoras da pré-mutação, a metilação das CpGs à jusante da borda tem uma média de 34,5% de metilação (Naumann *et al.*, 2014). Sendo assim, mulheres portadoras da pré-mutação exibem um perfil de metilação semelhante ao padrão observado em mulheres saudáveis (Naumann *et al.*, 2014). Já o perfil de metilação de DNA distinta na localização esperada e uma região promotora do gene *FMR1* não metilada, como observado em homens saudáveis (Naumann *et al.*, 2014).

Figura 6 – Perfis de metilação na região promotora do gene FMR1 e no segmento 5' à montante em linhagens celulares de mulheres saudáveis, homens e mulheres pré-mutados, mulheres com a mutação completa e pacientes com a síndrome de Turner (45,X)

40



Legenda: Em amostras com risco de expansão das repetições de trinucleotídeos CGG, a borda de metilação na região 5' UTR do gene *FMR1* parece permanecer intacta. Amostras analisadas por sequenciamento de DNA tratado com bissulfito de sódio, derivadas de células mononucleares de sangue periférico (PBMCs) de (a) mulher saudável, mostrando um mosaicismo de metilação entre a CpG 20 e a 70; (b) homem e (c) mulher pré-mutados, mostrando que, entre as CpGs 20 e 65, 45% das CpGs estão metiladas no genoma feminino; (d) e mulher com mutação completa com 59% das CpGs metiladas. (e) Já o perfil de metilação observado em PBMCs de pacientes com síndrome de Turner (45,X) revela uma borda de metilação de DNA estável. Os números 20-104 designam os dinucleotídeos CpG na região: ■ metilados e □ não metilados.

Segundo Naumann e colaboradores (2009), a borda de metilação está presente em todos os tipos e linhagens celulares humanas, independentemente da idade, do gênero e do estágio de desenvolvimento. Essa sequência de borda, tanto quando as CpGs se encontram metiladas quanto não metiladas, se liga especificamente a proteínas nucleares de células humanas (Naumann *et al.*, 2009). Sendo assim, os autores interpretam essa borda como relacionada a uma estrutura de cromatina específica que distingue uma área hipermetilada no genoma de uma região promotora do *FMR1* não metilado e, com isso, protege o gene da disseminação da metilação (Naumann *et al.*, 2009). Esse achado sugere então mudanças na sequência de nucleotídeos e na estrutura da cromatina nesta região de borda em indivíduos com SXF (Naumann *et al.*, 2009).

A estabilidade da borda de metilação pode, portanto, ser crucial para proteger a região promotora do gene *FMR1* contra a metilação (Naumann *et al.*, 2009). Por se tratar de uma condição epigenética e teoricamente reversível (Bar-Nur *et al.*, 2012; Torrioli *et al.*, 2010), a elucidação de perfis de metilação em pacientes com a SXF pode abrir oportunidades de intervenção terapêutica. Isso é corroborado pela existência dos homens com expansões CGG acima de 200 sem metilação sem comprometimento intelectual (*HFMs*) (Tabolacci *et al.*, 2016). Contudo, a região de borda foi pouco explorada e o seu significado ainda é desconhecido, requerendo maiores investigações.

1. OBJETIVOS

1.1 Objetivo geral

Tendo em vista a grande heterogeneidade genética e a alta taxa de subdiagnóstico observados na DILX, o presente estudo tem como objetivo geral a investigação de novos mecanismos genéticos e epigenéticos relacionados à SXF. Para isso, este trabalho é dividido em duas vertentes principais:

1.2 Objetivos específicos

Reavaliar através da metodologia de reação em cadeia da a) polimerase direcionada para repetições de trincas (TRP-PCR, do inglês Triplet Repeat Primed Polymerase Chain Reaction) um grupo de 101 homens com DI, previamente detectados em nosso laboratório como portadores da SXF ou com resultados inconclusivos, de modo a caracterizar padrões mutacionais atípicos subdetectados. dando continuidade ao que foi observado previamente pelo nosso grupo e corroborado pela literatura recente;

b) Identificar novos marcadores epigenéticos em regiões próximas às repetições CGG, incluindo a borda de metilação citada, que possam ser utilizados para inferência de padrões diferencialmente metilados entre indivíduos com a SXF e indivíduos controles e desenvolvimento futuro de metodologias diagnósticas direcionadas, tendo em vista que a metilação tem maior relação com a severidade da condição do que propriamente a expansão de repetições CGG.

2. METODOLOGIA

2.1 Pacientes com SXF

Os pacientes incluídos neste estudo foram encaminhados ao Serviço de da Universidade Estado Rio Genética Humana do do de Janeiro (SERVGEN/UERJ) durante os anos de 1995 a 2020 a partir de diferentes instituições públicas do Estado do Rio de Janeiro, dentre elas o Hospital Universitário Pedro Ernesto (HUPE/UERJ), Instituto de Puericultura e Pediatria Martagão Gesteira (IPPMG/UFRJ), Instituto Fernandes Figueira (IFF/FIOCRUZ), Hospital Universitário Gaffreé e Guinle (HUGG/UNIRIO), Universidade do Grande Rio (UNIGRANRIO), Hospital Municipal Jesus, Hospital Federal de Bonsucesso, e Instituto Estadual de Diabetes e Endocrinologia Luiz Capriglione (IEDE).

Foram selecionados 101 pacientes do sexo masculino com SXF, com idade entre 1 a 30 anos (10,86 ± 6,00 anos) e alguns de seus familiares, incluindo três avôs maternos sem distúrbio neurológico (Apêndice A). Esse grupo amostral foi definido a partir da presença da mutação completa no gene *FMR1* ou de um resultado inconclusivo, através das metodologias de rotina utilizadas no SERVGEN. Além disso, a amostra controle foi composta por 18 voluntários, não pareados, representados por alunos da Universidade do Estado do Rio de Janeiro.

Este estudo faz parte de um projeto de pesquisa mais amplo, coordenado pela Prof^a Dr^a Cíntia Barros Santos-Rebouças, aprovado pelo Comitê de Ética em Pesquisa da Universidade do Estado do Rio de Janeiro (CAAE 46769315.5.0000.5259) (Anexo A). Todos os procedimentos adotados seguiram as condutas descritas na Resolução 466/2012 do Conselho Nacional de Pesquisa (CNS/MS) e em suas complementares, que regem as pesquisas envolvendo seres humanos.

As análises realizadas ao longo deste estudo foram divididas em dois segmentos principais. A primeira parte é constituída pela caracterização dos diferentes perfis mutacionais referentes à expansão das repetições CGG no gene *FMR1*, através da reavaliação de pacientes previamente detectados em nosso laboratório como portadores da SXF. Já o segundo segmento é formado pela

identificação de novos marcadores epigenéticos em regiões próximas às repetições CGG no gene *FMR1*, incluindo a borda de metilação descrita na literatura (Naumann *et al.*, 2009; 2014), que possam ser utilizados para inferir padrões diferencialmente metilados entre os indivíduos com a SXF e os indivíduos controles. Com isso, as análises seguiram as etapas representadas nas Figuras 7 e 8.



Figura 7 – Esquema das etapas realizadas na primeira parte deste estudo

Nota: Após a coleta de sangue venoso dos pacientes, foi realizada a extração de DNA, seguida da avaliação da região de repetições CGG por metodologias de rotina baseadas em PCR. Com a sugestão da presença da mutação completa ou de padrão atípico de repetições, deu-se a investigação da região. O DNA dos pacientes foi amplificado por TRP-PCR, utilizando-se o kit AmplideX® *FMR1* mPCR, e os fragmentos gerados foram separados por eletroforese capilar para análise do número de repetições CGG. Para a identificação de deleções próximas às repetições CGG, foi realizado o ensaio de PCR e sequenciamento do produto. Dentre os casos atípicos, foi realizada a análise de marcadores genéticos no cromossomo X, através do ensaio de PCR, para descartar a possibilidade de síndrome de Klinefelter (47,XXY) nestes casos.

Fonte: A autora, 2022.



Figura 8 – Esquema das etapas realizadas na segunda parte deste estudo

Nota: Para a definição das melhores regiões alvo, foi realizada uma análise prévia por bioinformática com dados de metilação do gene *FMR1* obtidos em bancos de dados públicos. Após a identificação do número de repetições CGG por TRP-PCR, com o kit AmplideX® *FMR1* mPCR, o DNA dos pacientes foi tratado com bissulfito de sódio. O DNA tratado foi amplificado por PCR para cada região alvo definida, sendo uma reação em duplicata para a borda de metilação e uma reação única para o intron 2. Os produtos passaram por uma etapa de pirosequenciamento, seguido da quantificação de metilação nas posições CpG de cada região. A região de borda de metilação foi segmentada em dois fragmentos devido à sua extensão, sendo analisada em duas leituras, utilizando dois oligonucleotídeos.

Fonte: A autora, 2022.

2.2 Coleta de Material Biológico

Amostras de DNA de cada paciente foram obtidas a partir de sangue periférico para a análise molecular. Cerca de 5 mL de sangue foram coletados em

tubos *vacutainer* contendo anticoagulante EDTA (ácido etilenodiamino tetraacético), sob condições de plena assepsia. A coleta de material biológico foi realizada após a autorização da participação na pesquisa, através da assinatura, pelos respectivos responsáveis, do termo de consentimento livre e esclarecido (Anexo B). As amostras utilizadas neste estudo encontram-se armazenadas no banco de amostras biológicas do Serviço de Genética Humana da Universidade do Estado do Rio de Janeiro (SERVGEN/UERJ), sob aprovação da Conselho Nacional de Ética em Pesquisa (CONEP).

2.3 Extração de DNA

Considerando que o banco de amostras de DNA é constituído por amostras extraídas ao longo de mais de duas décadas de atividade do grupo de pesquisa, a extração de DNA foi realizada por diferentes metodologias, incluindo *salting out* (Miller *et al.*, 1988) e diferentes kits comerciais (QIAamp Mini kit, Qiagen; Gentra Puregene Blood kit, Qiagen; Blood Extraction kit, GE Healthcare; Wizard® Genomic DNA Purification Kit, Promega; entre outros), seguindo as instruções contidas nos protocolos descritos pelos fabricantes. Ao término deste procedimento, uma alíquota da solução final contendo o DNA de cada paciente foi armazenada em microtubo de 1,5 mL em geladeira a 4ºC e as demais alíquotas estocadas em tubos de 1,5 mL em freezer a -20ºC.

2.4 Estimativa de integridade e concentração do DNA

A integridade das amostras de DNA foi avaliada através da técnica de eletroforese em gel de agarose à 0,8% (*Bioagency*) diluída em tampão TBE 1X [Tris 89 mM (*GE Healthcare*), ácido bórico 89 mM (*MERCK*), EDTA 2 mM (*GE Healthcare*)]. O preparo das amostras foi feito com 1 µL da alíquota de DNA, 4 µL de água deionizada (MilliQ) e 1,5 µL da solução contendo 2 µL de GelRed[™] (Biotium Inc.) previamente diluído em 500 µL de água MilliQ e 250 µL de corante

de corrida [azul de bromofenol 0,25% (*GE Healthcare*); xileno cianol 0,25% (*GE Healthcare*); glicerol 30% (*ISOFAR*) e 8 µL de água deionizada (MilliQ)]. Essa etapa se deu em cuba horizontal [MultiSUB Midi (Uniscience) e MS 250V Power Supply (*Major Science*)], a 100 V por aproximadamente 35 minutos, utilizando tampão de corrida TBE 1X [Tris 89 mM (*GE Healthcare*), ácido bórico 89 mM (*MERCK*), EDTA 2 mM (*GE Healthcare*)]. Após a eletroforese, o gel foi visualizado, com a exposição à luz ultravioleta, em um sistema de fotodocumentação L-Pix® Ex (*Loccus Biotecnologia*).

As amostras de DNA tiveram sua concentração estimada através de um espectrofotômetro modelo NanoDrop Lite (*Thermo Fisher Scientific*). Essa estimativa, assim como, o grau de pureza das amostras (como contaminação por proteínas, fenóis e carboidratos) foi calculado pelo equipamento com a configuração para DNA de fita dupla (modo *double strand*). Essa etapa foi realizada com 1 µL da alíquota de DNA extraído. Foi utilizado 1 µL da solução de reidratação do kit comercial como branco na calibração do equipamento para amostras de DNA que foram extraídas por kit comercial e o tampão TE [Tris 10 mM (*GE Healthcare*), HCI (*MERCK*), EDTA 1 mM (*GE Healthcare*); pH 7,4] para as amostras de DNA extraídas a partir do método de *salting out*.

2.5 Amplificação e análise da metilação nas repetições CGG

Para a investigação de perfis atípicos em pacientes com SXF, realizou-se a amplificação e a análise de metilação da região de interesse contendo as repetições CGG, através da metodologia de reação em cadeia da polimerase direcionada para repetições de trincas (TRP-PCR, do inglês *Triplet Repeat Primed Polymerase Chain Reaction*).

2.5.1 Reação em cadeia da polimerase (PCR)

Nesta etapa, utilizou-se o kit AmplideX® *FMR1* mPCR (Asuragen, Inc.), seguindo protocolo do fabricante, conforme mostra a Figura 9. Essa metodologia segue o princípio descrito na Figura 10, baseada na utilização de um

oligonucleotídeo híbrido que pareia com as repetições CGG, além de um par de oligonucleotídeos convencional que flanqueia as repetições CGG. Com isso, ele permite uma amplificação da região de repetição CGG sensível e capaz de distinguir entre alelos normais, intermediários, pré-mutados, mutação completa e os casos de mosaicismo de tamanho, como, por exemplo, na presença da mutação completa concomitante ao alelo normal ou à pré-mutação. Além disso, a PCR de metilação (mPCR), baseada na digestão com uma endonuclease sensível à metilação, anterior à etapa de amplificação, permite o cálculo da porcentagem de metilação na região analisada. O kit inclui, ainda, controles para a digestão e para a amplificação das repetições, atestando o funcionamento de ambas as etapas da metodologia. Todas as reações foram realizadas em tubos de 0,2 µL, dentro de um fluxo laminar sob condições de plena assepsia. A ciclagem foi realizada no termociclador Veriti modelo 9902 (Thermo Fisher Scientific Inc), seguindo as condições descritas no protocolo do fabricante. Ao término da ciclagem as reações foram acondicionadas em freezer a -20°C. As sequências dos oligonucleotídeos utilizados nestas reações não foram disponibilizadas pela empresa fabricante do kit.

Figura 9 – Fluxo da metodologia de amplificação por PCR e análise de metilação na região de repetições CGG, utilizando o kit AmplideX® *FMR1* mPCR (Asuragen, Inc.)



Nota: Cada amostra de DNA genômico é previamente diluída a 20 ng/µL e misturada à um DNA controle, disponível no kit. São realizadas então duas reações simultaneamente, uma com oligonucleotídeos marcados com FAM e outra com HEX. A metodologia conta com uma digestão enzimática para análise da metilação na reação HEX, um ensaio de PCR para amplificação e uma separação dos fragmentos por eletroforese capilar.

Fonte: adaptada do protocolo de Asuragen, Inc, disponível em https://asuragen.com/wpcontent/uploads/2016/05/AmplideX-PCR-CE-FMR1-mPCR-RUO-IFU.pdf. Figura 10 – Princípios da metodologia de TRP-PCR, utilizando o kit AmplideX® *FMR1* mPCR (Asuragen, Inc.)



Nota: A amostra de DNA genômico é previamente misturada ao mix de DNA controle, incluindo o Controle de Digestão e o Controle CGG. Cada alíquota é então submetida ao tratamento com a enzima DNase sensível à metilação ou com a enzima controle. Como resultado, apenas os alelos metilados são retidos na reação de digestão e seguem as reações subsequentes, enquanto a reação controle reflete o espectro completo dos alelos da amostra. O controle de digestão na amostra de DNA controle confirma o funcionamento da reação de digestão.

Os produtos da reação foram submetidos à eletroforese capilar para separação dos fragmentos. Para essa etapa, 9 µL de formamida Hi-Di (Thermo Fisher Scientific Inc.), 2 µL do padrão de tamanho de fragmento ROX1000 *size ladder* e 2 µL dos produtos FAM e HEX do ensaio de PCR foram aplicados em uma placa (MicroAmp® Optical 96-Well Reaction Plate, Thermo Fisher Scientific Inc.). A desnaturação das amostras foi realizada a 95°C por 2 minutos e resfriamento a 4°C por 2 minutos, em termociclador Veriti modelo 9902 (Thermo Fisher Scientific Inc.). Após, a placa foi então injetada no equipamento ABI3130 (Thermo Fisher Scientific Inc.) para a separação dos fragmentos.

Fonte: adaptada do protocolo da Asuragen, Inc, disponível em https://asuragen.com/wp-content/uploads/2016/05/AmplideX-PCR-CE-FMR1-mPCR-RUO-IFU.pdf.

Os dados obtidos foram processados no *software* GeneMarker v.2.6.3 (SoftGenetics), conforme a Figura 11. A presença de oligonucleotídeos marcados com corantes diferentes para DNA digerido e controle possibilita a análise de ambas as informações simultaneamente no programa. A reação de digestão é amplificada com oligonucleotídeos marcados com corante HEX, enquanto a reação controle com oligonucleotídeos marcados com FAM. Dessa forma, o número de repetições é visualizado em FAM e a porcentagem de metilação é acessada em HEX.

Figura 11 – Fluxograma de análise dos dados (tamanho de fragmentos e metilação) da região de repetições CGG no gene *FMR1*, através do programa GeneMarker v.2.6.3 (SoftGenetics)



Nota: A análise envolve uma série de passos para obtenção dos tamanhos das repetições CGG no canal FAM e da porcentagem de metilação correspondente a cada alelo. Fonte: adaptada do protocolo de Asuragen, Inc. O programa disponibilizado pelo fabricante do kit calcula a quantidade de repetições CGG presentes em cada amostra, utilizando a função de análise específica para os kits AmplideX®, da seguinte maneira:

$$CGG_i = \frac{Peak_i - c_0}{m_0} \tag{1}$$

Onde *Peak*_i é a altura do pico de um dado produto, C_0 é o fator de correção do tamanho e m_0 é o fator de correção de mobilidade para cada repetição CGG. O fator de correção do tamanho se refere à região em comum de DNA incluída pelos oligonucleotídeos, excluindo as repetições CGG. Enquanto o fator de correção de mobilidade representa a mobilidade aparente de cada unidade de repetição. Os fragmentos amplificados possuem pequenas diferenças em relação aos fatores de correção, dependendo das condições de corrida e o equipamento utilizado. No caso do equipamento ABI3130 (Thermo Fisher Scientific Inc.), utilizado nesse trabalho, o valor de C_0 considerado foi 229,4 e o de m_0 foi 2,965.

Já para a análise de metilação, o programa calcula a razão entre a altura do pico do fragmento digerido em HEX e do não digerido em FAM, ambos normalizados pela altura do pico do fragmento controle CGG (Referência), da seguinte maneira:

$$\%Me_i = 100 \times \frac{[Peak_i/Peak_{REF}]_{HEX}}{[Peak_i/Peak_{REF}]_{FAM}}$$
(2)

Onde *Peak*i corresponde à altura do pico de um produto no canal HEX ou FAM e *Peak_{REF}* é a altura do pico do produto controle CGG no canal correspondente. Os resultados em porcentagem de metilação são classificados da seguinte forma: abaixo de 20% são considerados não metilados, de 20% a 80% são parcialmente metilados e acima de 80%, hipermetilados. As porcentagens de metilação calculados em excesso, ou seja, acima de 100%, são escalonados em 100%, conforme o protocolo do fabricante.

2.6 Análise de deleções próximas à região de repetições CGG

Durante as análises de rotina para a avaliação da região de repetições, através das amplificações por PCR (Fu *et al.*, 1991; Haddad *et al.*, 1996) seguida de eletroforese, algumas amostras levantaram suspeitas da presença de deleções próximas à região CGG. Em busca de melhor caracterizar esses indivíduos, suas amostras foram amplificados novamente através da PCR (Haddad *et al.*, 1996; Araújo, 2018). Os fragmentos de interesse foram extraídos do gel e purificados. As sequências foram então acessadas por sequenciamento de Sanger dos produtos purificados. Nessas análises, foram incluídas também os pacientes com mosaicismo de tamanho envolvendo o alelo normal.

2.6.1 Reação em cadeia da polimerase (PCR)

As amostras com suspeita de deleção foram amplificadas pelo ensaio de PCR desenvolvido por nosso grupo (Araújo,2018) (Apêndice B), em um trabalho anterior do grupo, incluída em sua monografia de conclusão de curso de Graduação em Ciências Biológicas (IBRAG/UERJ). Essa reação é capaz de amplificar fragmentos no gene *FMR1* até pequenas mutações completas. Nessas reações, foram utilizados dois oligonucleotídeos, que amplificam fragmentos até 362 pares de base para alelos normais, entre 365 e 392 pares de base para alelo intermediário, entre 393 e 824 pares de base para pré-mutação, e acima disso já abrange a região de mutação completa. Quando identificado um fragmento de tamanho menor que 362 pares de base, o que pode representar uma possível deleção, esses fragmentos foram extraídos do gel e purificados para sequenciamento. As sequências dos oligonucleotídeos utilizadas nestas reações não tiveram sua divulgação autorizada.

Outra amplificação que também foi utilizada para acessar essa região foi o ensaio de PCR convencional para o rastreamento de expansões CGG (Haddad *et al.*, 1996), rotineiramente utilizada para análise da SXF. Todas as reações foram realizadas em tubos de 0,2 µL, dentro de um fluxo laminar sob condições de plena

assepsia. Essa análise é baseada na ausência de amplificação do fragmento quando expandido a um número elevado de repetições, como no caso de mutação completa e pré-mutações muito grandes (em geral, a amplificação ocorre em alelos com tamanhos menores ou iguais a 100 ou até 120 repetições). Nesse caso, quando identificado um fragmento de interesse de tamanho menor ao fragmento do controle normal, ou seja, menor que 400 pb, o fragmento foi extraído do gel e purificado para sequenciamento.

2.6.2 <u>Avaliação da qualidade dos fragmentos amplificados e do rendimento dos</u> ensaios de PCR

Os produtos dos ensaios de PCR foram visualizados através da eletroforese em gel de agarose 1% (*Thermo Fisher Scientific*) diluída em tampão TBE 1X [Tris 89 mM (*GE Healthcare*), ácido bórico 89 mM (*MERCK*), EDTA 2 mM (*GE Healthcare*)]. O preparo das amostras foi feito com 5 µL do produto dos ensaios de PCR e 1,5 µL da solução contendo 2 µL de GelRedTM (Biotium Inc.) previamente diluído em 500 µL de água MilliQ e 250 µL de corante de corrida [azul de bromofenol 0,25% (*GE Healthcare*); xileno cianol 0,25% (*GE Healthcare*); glicerol 30% (*ISOFAR*) e 8 µL de água deionizada (MilliQ)]. Essa etapa se deu em cuba horizontal [MultiSUB Midi (Uniscience) e MS 250V Power Supply (*Major Science*)], a 100 V por aproximadamente 45 minutos (para o ensaio de PCR desenvolvido por Araújo, 2018) e a 80 V por aproximadamente 80 minutos (para o ensaio de PCR de Haddad *et al.*, 1996), utilizando como tampão de corrida TBE 1X [Tris 89 mM (*GE Healthcare*), ácido bórico 89 mM (*MERCK*), EDTA 2 mM (*GE Healthcare*)].

Para confirmar o tamanho do fragmento amplificado, as bandas dos produtos foram comparadas com um padrão de peso molecular 1 kb DNA *Ladder* Plus (*Life Technologies*). Após a eletroforese, o gel foi visualizado, com a exposição à luz ultravioleta, em um sistema de fotodocumentação L-Pix® Ex (*Loccus Biotecnologia*). A análise do rendimento dos ensaios de PCR e da qualidade dos fragmentos amplificados foi realizada através de estimativa visual da intensidade das bandas.

Os fragmentos extraídos do gel foram purificados com o kit PureLink Quick Gel Extraction (*Thermo Fisher Scientific*), seguindo o protocolo descrito pelo fabricante.

Antes da reação de sequenciamento, os produtos dos ensaios de PCR foram purificados com a enzima ExoSAP-IT® (*Affymetrix*), seguindo o protocolo descrito pelo fabricante. A purificação teve por objetivo a remoção de resíduos resultantes dos reagentes dos ensaios de PCR.

2.6.4 Reação de sequenciamento de Sanger

Para o sequenciamento, utilizou-se os mesmos oligonucleotídeos dos ensaios de PCR (Haddad *et al.*, 1996; Araújo, 2018). Para essa reação, utilizou-se o Kit Big Dye Terminator v3.1 (*Thermo Fisher Scientific*). As amostras foram preparadas com 1 μ L do oligonucleotídeo senso ou anti-senso a 3,2 μ M, 1 μ L de tampão de sequenciamento 5X (*Thermo Fisher Scientific*), 1,5 a 2 μ L de DNA amplificado e purificado, 0,5 μ L do Kit e água deionizada (MilliQ) suficiente para um volume final de 10 μ L. Essas reações foram conduzidas em um termociclador Veriti modelo 9902 (*Thermo Fisher Scientific*), seguindo as condições de ciclagem apresentadas na Quadro 3.

Etapas	
96°C – 1 min	
96ºC – 10 seg	
50°C – 15 seg	215x
60°C – 1 min 15 seg	
96ºC – 10 seg	
50°C – 15 seg	5x
60°C – 1 min 30 seg	
96ºC – 10 seg	
50°C – 15 seg	5 x
60°C – 2 min	

Quadro 3 - Condições de ciclagem utilizadas no sequenciamento

Fonte: A autora, 2022.

Os produtos foram, então, precipitados para a remoção do excesso dos reagentes não incorporados na reação. Com esta finalidade, 80 µL de isopropanol 75% foram adicionados ao produto da reação e, em seguida, a placa de 96 poços [MicroAmp® Optical 96-Well Reaction Plate (*Thermo Fisher Scientific*)] foi centrifugada por 50 minutos a 3000 rpm (*SIGMA*, 2-16). O sobrenadante foi descartado cuidadosamente e a placa com as amostras precipitadas foi centrifugada em posição invertida por 1 minuto a 900 rpm, sendo em seguida levada ao termociclador Veriti 9902 (*Thermo Fisher Scientific*) a 75°C por 5 minutos para secagem.

O material foi ressuspenso em 10 µL de solução de formamida Hi Di (*Thermo Fisher Scientific*). A placa com as amostras foi levada ao termociclador Veriti 9902 (*Thermo Fisher Scientific*) por 5 minutos a 95°C para desnaturação e, imediatamente, incubada em gelo para manter a desnaturação do DNA. A placa foi novamente centrifugada a 900 rpm por 1 minuto e inserida em um sequenciador ABI3130 (*Life Technologies*) para processamento.

Na análise das sequências obtidas pela reação de sequenciamento, utilizou-se o programa BioEdit Sequence Alignment Editor versão 7.0.9.0 (Isis Pharmaceuticals, Inc) e essas sequências foram alinhadas e comparadas ao fragmento correspondente à sequência do gene *FMR1* sem alterações, acessado no banco de dados Ensembl genome browser (em hg38). As deleções foram identificadas, analisando-se as sequências obtidas no sequenciamento com a utilização da ferramenta online Human BLAT Search (em hg38), disponível em

https://genome.ucsc.edu/cgi-bin/hgBlat. Os tamanhos dos fragmentos deletados foram estimados, utilizando-se a ferramenta online Fragment Size Calculator, disponível em http://biotools.nubic.northwestern.edu/SizeCalc.html.

2.7 Análise de marcadores nos cromossomos X e Y

Com o objetivo de avaliarmos se a presença de mais de um alelo de repetições CGGs poderia ser devido à presença de mais de um cromossomo X nos pacientes com padrões atípicos, ou seja, para avaliar a possibilidade dos pacientes com mosaicismo de tamanho apresentarem a síndrome de Klinefelter (47,XXY) além da SXF, foi realizada a análise de marcadores internos ao gene *SHOX*, localizado na região pseudoautossômica 1 nos cromossomos X e Y. Essa análise se deu por PCR, seguida de eletroforese capilar.

2.7.1 Reação em cadeia da polimerase (PCR)

As regiões de interesse para a análise de marcadores foram amplificadas por PCR, utilizando um par de oligonucleotídeos iniciadores para cada fragmento, cedidos pelos professores Enrique Medina-Acosta (UENF) e Filipe Brum Machado (UEMG) (Quadro 4). Todas as reações foram realizadas em tubos de 0,2 µL, seguindo as condições detalhadas no Quadro 5, dentro de um fluxo laminar sob condições de plena assepsia. Os ensaios de PCR foram realizados em um termociclador Veriti 9902 (*Thermo Fisher Scientific*) com os parâmetros de ciclagem descritos no Quadro 6. As sequências dos oligonucleotídeos utilizados nestas reações tiveram sua divulgação preservada.

Quadro 4 – Oligonucleotídeos utilizados nos ensaios de PCR das regiões de interesse para análise de marcadores

	Marcador	Amplicon	Microssatélite	Fluorocromo
1	NUDIM SHOX1	263 bp	TAAA	VIC
2	NUDIM SHOX2	251 bp	CTTT	FAM
3	NUDIM SHOX3	203 bp	TCC	NED
4	NUDIM SHOX4	243 bp	CTTT	PET

Fonte: Adaptado de Medina e Brum, 2021 (cedido pelo autor).

Quadro 5 – Condições utilizadas nos ensaios de PCR para amplificação dos fragmentos para análise de marcadores

Reagentes	
Tampão de reação 10x (Uniscience) *	1x
MgCl ₂ 50 mM (Uniscience)	2,5 mM
dNTPs 5 mM (<i>GE Healthcare</i>)	200,0 µM
Oligonucleotídeos F e R 2 µM	0,1 µM
<i>Taq</i> DNA polimerase 5 U/µL (Uniscience)	1,0 U
DNA ~60/100 ηg	1,0 µL
Volume final	25,0 µL

Nota: Ambos os oligonucleotídeos F e R são reunidos em uma única alíquota. Legenda: F: senso, do inglês *forward*; R: anti-senso, do inglês *reverse*. * Tampão de reação 10x [KCI 5 mM, Tris HCI 7,5 mM - pH 9,0, (NH₄)₂ SO₄ 2 mM]. Fonte: A autora, 2022.

Quadro 6 – Condições de ciclagem utilizadas nos ensaios de PCR para amplificação dos fragmentos para análise de metilação

Oligonicleotídeos 1, 2, 3	Oligonucleotídeos 4
95ºC – 11 min	95⁰C – 11 min
94°C – 1 min	94ºC – 1 min
59°C – 1 min 🔓 29x	61°C – 1 min > 29x
72°C – 1 min	72°C – 1 min 丿
60°C – 60 min	60°C – 60 min
4ºC – ∞	4ºC – ∞

Fonte: A autora, 2022.

2.7.2 <u>Avaliação da qualidade dos fragmentos amplificados e do rendimento dos</u> ensaios de PCR

Os produtos dos ensaios de PCR foram visualizados através da eletroforese em gel de agarose 1% (Thermo Fisher Scientific) diluída em tampão TBE 1X [Tris 89 mM (GE Healthcare), ácido bórico 89 mM (MERCK), EDTA 2 mM (GE Healthcare)]. O preparo das amostras foi feito com 5 µL do produto dos ensaios de PCR e 1,5 µL da solução contendo 2 µL de GelRed™ (Biotium Inc.) previamente diluído em 500 µL de água MilliQ e 250 µL de corante de corrida [azul de bromofenol 0,25% (GE Healthcare); xileno cianol 0,25% (GE Healthcare); glicerol 30% (ISOFAR) e 8 µL de água deionizada (MilliQ)]. Essa etapa se deu em cuba horizontal [MultiSUB Midi (Uniscience) e MS 250V Power Supply (Major Science)], a 100 V por aproximadamente 20 minutos, utilizando como tampão de corrida TBE 1X [Tris 89 mM (GE Healthcare), ácido bórico 89 mM (MERCK), EDTA 2 mM (GE Healthcare)]. Para confirmar o tamanho do fragmento amplificado, as bandas dos produtos foram comparadas com um padrão de peso molecular 1 kb DNA Ladder Plus (Life Technologies). Após a eletroforese, o gel foi visualizado, com a exposição à luz ultravioleta, em um sistema de fotodocumentação L-Pix® Ex (Loccus Biotecnologia). A análise do rendimento do ensaio de PCR e da qualidade dos fragmentos amplificados foi realizada através de estimativa visual da intensidade das bandas e da presença de uma única banda.

2.7.3 Eletroforese capilar

Uma amostra de cada produto dos ensaios de PCR foi submetida à eletroforese capilar no sistema de sequenciamento automático ABI3130 (*Thermo Fisher Scientific*). Foram aplicados numa placa (MicroAmp® Optical 96-Well Reaction Plate – *Thermo Fisher Scientific*), 0,5 a 1,5 µL de cada produto dos ensaios de PCR juntamente com uma mistura composta de 9,0 µL de formamida *Hi-Di* (*Thermo Fisher Scientific*) e 0,5 µL de LIZ 500 (*Thermo Fisher Scientific*). A

placa foi desnaturada a 95°C por 5 minutos, colocada imediatamente no gelo e inserida no sequenciador automático. Os resultados foram analisados com o auxílio do programa GeneMarker V2.6.3 para genotipagem de microssatélites (Softgenetics).

2.8 Análise por ferramentas computacionais e produção de *heatmaps*

Para a identificação de marcadores epigenéticos em regiões próximas às repetições CGG, incluindo a borda de metilação, foi realizada uma análise por Bioinformática, utilizando-se o programa R (disponível em https://www.rproject.org/). Para essa análise, foram acessados dados públicos de metilação por array em plataforma Illumina HumanMethylation450 BeadChip no banco de dados online Expression Omnibus Gene (GEO Datasets, disponível em https://www.ncbi.nlm.nih.gov/geo/). Na literatura, havia um único trabalho com valores de metilação referentes a pacientes com a SXF, publicado por Alisch e colaboradores (2013) (Dataset Series GSE41273). Foram incluídos nessa análise todos os valores referentes ao gene FMR1 fornecidos neste trabalho. Os dados de controles selecionados para esta etapa foram disponibilizados por Alisch et al. (2012) (Dataset Series GSE27097) e Gross et al. (2016) (Dataset Series GSE67705). O heatmap gerado contou com a comparação dos valores de todas as sondas de metilação disponíveis em regiões à montante e ao longo do gene FMR1 em pacientes com a SXF e em indivíduos controles (Quadro 7). Essa análise foi realizada em colaboração com a Prof^a Dr^a Sheila Coelho Soares Lima e o pesquisador Diego Camuzi, do Instituto Nacional de Câncer (INCa).

Para identificar a localização de cada elemento genético e epigenético, a sequência de DNA completa do gene *FMR1*, incluindo as ilhas CpG, foi obtida no banco de dados UCSC *Genome Browser*, utilizando a versão de referência do genoma humano hg19. As sequências referentes aos exons e aos introns do gene foram identificadas no banco de dados Ensembl, a partir do transcrito ENSG00000102081.14. A sequência referente à região promotora do gene *FMR1* foi acessada no banco de dados *Eukaryotic Promoter Database*. As sequências

referentes aos Elementos Epigenéticos Relacionados ao X-Frágil 1 e 2 (FREE1 e FREE2), que consistem em duas regiões reconhecidas como marcadores epigenéticos ao redor das repetições, e as sequências das repetições CGG foram visualizadas a partir do material complementar publicado por Godler e colaboradores (2011).

Quadro 7 – Sondas de metilação no gene *FMR1,* utilizadas nos ensaios de *array*, cujos dados estavam disponíveis publicamente no GEO database

Sonda	Localização no gene <i>FMR1</i> (hg19)	Localização genômica (hg19)	Observação
cg20583833	3491 pb antes do promotor	chrX:146989930	CpG 95
cg07147350	512 pb antes do promotor	chrX:146992908	CpG 65. Integra a região de borda analisada por pirosequenciamento
cg22625568	411 pb antes do promotor	chrX:146993010	CpG 63. Integra a região FREE1
cg05288927	329 pb antes do promotor	chrX:146993092	CpG 60. Integra a região FREE1
cg14470409	296 pb antes do promotor	chrX:146993125	CpG 57. Integra a região FREE1 e a ilha CpG
cg04744025	245 pb antes do promotor	chrX:146993175	CpG 55. Integra a região FREE1 e a ilha CpG
cg04552106	Promotor	chrX:146993426	CpG 24. Integra a região de ilha CpG
cg02921434	Promotor	chrX:146993434	CpG 23. Integra a região de ilha CpG
cg03513471	5'UTR (entre promotor e repetição)	chrX:146993486	CpG 18. Integra a região de ilha CpG
cg16783314	5'UTR/Exon1 (após a repetição)	chrX:146993722	Integra a região FREE2 e ilha CpG
cg08434396	Intron 1	chrX:146993779	Integra a região FREE2 e ilha CpG
cg21274274	Intron 1	chrX:146994129	Integra a região de ilha CpG
cg19741073	Intron 1	chrX:146994167	Integra a região de ilha CpG
cg21378039	Intron 1	chrX:146994369	-
cg27064928	Intron 1	chrX:146996911	Integra o gene FMR1-AS
cg22417678	Intron 2	chrX:147004697	Integra a região do intron 2 analisada por pirosequenciamento

Fonte: A autora, 2022.

2.9 Análise de metilação da região da borda de metilação e de parte do intron2

Com o objetivo de se obter uma quantificação precisa da metilação entre indivíduos com diferentes perfis mutacionais para a SXF e em indivíduos controles, a região da borda de metilação, descrita a 650-800 nucleotídeos e 65-70 CpGs à montante das repetições CGG por Naumann e colaboradores (2009; 2014), foi avaliada por pirosequenciamento nos pacientes selecionados. Nessa análise, foi incluída também a CpG 71 para comparação, visto que se trata de uma zona de transição em indivíduos controles, separando um segmento de genoma fortemente ou completamente metilado de uma região livre de metilação (Naumann *et al.*, 2009; 2014). Além disso, foi analisada a região da sonda de metiloma cg22417678, localizada no intron 2 do gene *FMR1*, devido ao fato dos valores de metilação apresentarem-se diferentes das demais sondas observadas no *heatmap*. Para isso, cada amostra de DNA passou por uma etapa de conversão com bissulfito de sódio, um ensaio de PCR para cada uma das regiões citadas e foi então submetida ao pirosequenciamento.

2.9.1 Conversão do DNA genômico com bissulfito de sódio

Somente amostras de DNA íntegras que apresentaram concentração superior a 25 ng/µL foram tratadas utilizando-se o kit EpiTect Fast DNA Bisulfite (*Qiagen*), seguindo o protocolo do fabricante. O tratamento do DNA com bissulfito de sódio resulta na conversão das citosinas não metiladas em uracila, mantendo as citosinas metiladas inalteradas. Em seguida, o rendimento dessa conversão foi conferido através de nova quantificação em um espectrofotômetro modelo NanoDrop Lite (*Thermo Fisher Scientific*).

Após a conversão com bissulfito de sódio, os fragmentos das regiões de interesse para a análise de metilação foram amplificados por PCR, utilizando um par de oligonucleotídeos iniciadores para cada fragmento, formulados com o programa PyroMark Assay Design Software (*Qiagen*). Para o pirosequenciamento, os oligonucleotídeos antisenso (*reverse*) utilizados nos ensaios de PCR foram marcados com biotina na extremidade 5'. Todas as reações foram realizadas em tubos de 0,2 µL, dentro de um fluxo laminar sob condições de plena assepsia, seguindo as condições detalhadas no Quadro 8. Os ensaio de PCR ocorreram em um termociclador Veriti 9902 (*Thermo Fisher Scientific*) com os parâmetros de ciclagem descritos no Quadro 9.

Quadro 8 – Condições utilizadas nos ensaio de PCR para amplificação dos fragmentos para análise de metilação

	Regiões analisadas		
Reagentes	Borda de metilação	Intron 2	
Tampão de reação 10x Platinum (<i>Thermo Fisher</i> Scientific)*	1X	1X	
MgCl ₂ 50 mM Platinum (<i>Thermo Fisher Scientific</i>)	2,0 mM	1,5 mM	
dNTPs 5 mM (<i>GE Healthcare</i>)	200,0 µM	200,0 µM	
Oligonucleotídeo F 10 µM (IDT ou INVITROGEN/Life Technologies)	0,2 µM	0,1 µM	
Oligonucleotídeo R 10 µM (IDT)	0,2 µM	0,1 µM	
<i>Taq</i> DNA polimerase Platinum 5 U/µL (<i>Thermo Fisher Scientific</i>)	2,0 U	1,0 U	
Dimetilsulfóxido (DMSO) (<i>MERCK</i>)	10%	-	
DNA ~50 ηg	1,0 µL	1,0 µL	
Volume final	50,0 µL	50,0 µL	

Nota: Ambos os oligonucleotídeos R são marcados com biotina na região 5'. F: senso, do inglês *forward*; R: anti-senso, do inglês *reverse*. * Tampão de reação 10x [KCl 5 mM, Tris HCl 7,5 mM - pH 9,0, (NH₄)₂ SO₄ 2 mM].

Fonte: A autora, 2022.

Quadro 9 – Condições de ciclagem utilizadas nos ensaios de PCR para amplificação dos fragmentos para análise de metilação



Fonte: A autora, 2022.

2.9.3 <u>Avaliação da qualidade dos fragmentos amplificados e do rendimento dos</u> ensaio de PCR

Os produtos dos ensaios de PCR foram visualizados por meio da eletroforese em gel de agarose 1% (Thermo Fisher Scientific) diluída em tampão TBE 1X [Tris 89 mM (GE Healthcare), ácido bórico 89 mM (MERCK), EDTA 2 mM (GE Healthcare)]. O preparo das amostras foi feito com 5 µL do produto dos ensaios de PCR e 1,5 µL da solução contendo 2 µL de GelRed[™] (Biotium Inc.) previamente diluído em 500 µL de água MilliQ e 250 µL de corante de corrida [azul de bromofenol 0,25% (GE Healthcare); xileno cianol 0,25% (GE Healthcare); glicerol 30% (ISOFAR) e 8 µL de água deionizada (MilliQ)]. Essa etapa se deu em cuba horizontal [MultiSUB Midi (Uniscience) e MS 250V Power Supply (Major Science)], a 100 V por aproximadamente 20 minutos, utilizando como tampão de corrida TBE 1X [Tris 89 mM (GE Healthcare), ácido bórico 89 mM (MERCK), EDTA 2 mM (GE Healthcare)]. Para confirmar o tamanho do fragmento amplificado, as bandas dos produtos foram comparadas com um padrão de peso molecular 1 kb DNA Ladder Plus (Life Technologies). Após a eletroforese, o gel foi visualizado, com a exposição à luz ultravioleta, em um sistema de fotodocumentação L-Pix® Ex (Loccus Biotecnologia). A análise do rendimento dos ensaios de PCR e da qualidade dos fragmentos amplificados foi realizada através

de estimativa visual da intensidade das bandas e da presença de uma única banda.

2.9.4 Pirosequenciamento

Os produtos dos ensaios de PCR com sinais fortes de amplificação após a eletroforese foram submetidos ao pirosequenciamento. Essa análise foi realizada em colaboração com a Prof^a Dr^a Sheila Coelho Soares Lima, do Instituto Nacional do Câncer (INCa). O pirosequenciamento utiliza o sequenciamento por síntese e trata cada sítio CpG como um polimorfismo C/T, gerando um dado quantitativo da proporção relativa do alelo metilado versus o alelo não metilado.

Para isso, o fragmento de DNA convertido é amplificado e a fita utilizada como molde é marcada com biotina. Após a desnaturação, o amplicon da fita biotinilada é isolado e se hibridiza com o oligonucleotídeo aplicado à reação. O primeiro dNTP é então adicionado à reação, de acordo com a sequência de interesse informada. A DNA polimerase catalisa a adição do dNTP ao oligonucleotídeo, se for complementar à próxima base na fita molde. Cada evento de incorporação de dNTP à fita é acompanhada da liberação de um pirofosfato (PPi), em quantidade equimolar à quantidade de nucleotídeo incorporado. A enzima ATP sulfurilase converte PPi em ATP na presença de adenosina 5'fosfosulfato (APS). Esse ATP leva à conversão, mediada por luciferase, de luciferina em oxiluciferina, gerando luz visível em quantidades proporcionais à quantidade de ATP. A luz produzida é então detectada pelos sensores CCD e vista como um pico no pirograma. A altura de cada pico (sinal de luz) é proporcional ao número de nucleotídeos incorporados. A enzima apirase continuamente degrada nucleotídeos não incorporados e ATP. Quando a degradação é completa, outro nucleotídeo é então incorporado. A adição de dNTPs é realizada sequencialmente (Figura 12).



Figura 12 – Princípio da técnica de pirosequenciamento

Legenda: (a) Ligação do oligonucleotídeo para sequenciamento à fita simples marcada com biotina do produto gerado no ensaio de PCR e incorporação de dNTP à nova fita sendo sintetizada, com consequente liberação de pirofosfato (PPi). (b) A enzima sulfurilase converte adenosina 5'fosfosulfato (APS) e PPi em ATP que é posteriormente utilizado na conversão de luciferina em oxiluciferina. Nesta etapa, há a liberação de luz, que é captada e representada como um pico no gráfico de luz x tempo (Pirograma). (c) Durante todo o procedimento, a enzima apirase degrada os nucleotídeos não incorporados, assim como o ATP.

Fonte: adaptada de Manual do usuário PyroMark Q96 ID (Qiagen), disponível em https://www.qiagen.com/ie/resources/resourcedetail?id=93a25fc1-80c9-461f-82ab-7250557d6dc0&lang=en.

As reações de pirosequenciamento foram realizadas na plataforma PyroMark Q96 ID (*Qiagen*), com o conjunto de reagentes PyroMark Q96 (*Qiagen*), seguindo o protocolo do fabricante. Os dados foram analisados com o programa PyroMark Q96 Software 2.0 (*Qiagen*). As sequências marcadas com biotina, ou seja, os fragmentos de DNA gerados com o oligonucleotídeo R foram selecionados como fita molde na análise. A região da borda de metilação foi analisada por pirosequenciamento de forma segmentada em duas leituras, devido à sua extensão, com o oligonucleotídeo A abrangendo as CpGs 71 a 69 e o B acessando as CpGs 68 a 65. Os oligonucleotídeos utilizados nesta etapa foram desenhados com o programa PyroMark Assay Design Software (*Qiagen*).

2.10 Análise por Bioinformática e produção de curvas ROC

Essa etapa foi realizada utilizando o programa GraphPad Prism versão 5 (GraphPad Software Inc.). Variáveis contínuas referentes aos grupos de análise foram avaliadas para distribuição normal através do teste de Kolmogorov-Smirnov. Diferenças entre as variáveis que passaram no teste de normalidade foram testadas pela análise de variâncias (ANOVA), seguida de pós-teste de Tukey para múltiplas comparações. Enquanto as diferenças entre as variáveis que não seguem a distribuição normal foram aferidas utilizando o teste não paramétrico de Kruskal-Wallis (Kruskal & Wallis, 1952), seguida pelo pós-teste de Dunn para múltiplas comparações (Dunn, 1964).

Para descrever o poder de discriminação do marcador analisado por pirosequenciamento, utilizamos os valores de sensibilidade, especificidade e acurácia, definidos por uma matriz de confusão a partir da produção de uma curva ROC (Quadro 10). A sensibilidade de um marcador se refere à capacidade de um sistema predizer a existência da condição quando ela realmente está presente, ou seja, a capacidade de apontar verdadeiros positivos e é dada pela fórmula:

Sensibilidade =
$$\frac{\text{acertos de predição de positivos}}{\text{total de positivos}}$$
 (1)

A especificidade de um marcador trata-se da capacidade de um sistema predizer a inexistência da condição quando ela realmente está ausente, ou seja, a capacidade de apontar verdadeiros negativos e é dada pela fórmula:

Especificidade =
$$\frac{\text{acertos de predição de negativos}}{\text{total de negativos}}$$
 (2)

Já a acurácia é a proporção de predições corretas totais que um sistema é capaz de realizar e é dada pela fórmula:

Acurácia =
$$\frac{\text{total de acertos de predição}}{\text{total}}$$
 (3)

O valor de corte de metilação entre os grupos positivos (com a SXF) e negativos (indivíduos sem a condição) foi dado após a definição do valor de

metilação no gráfico de curva ROC com melhor acurácia possível definido através do cálculo de hipotenusa (Figura 13).

		Valor verdadeiro (Confirmado por análise)	
		Positivo	Negativo
Valor previsto	Positivo	Verdadeiro Positivo	Falso Positivo
(Predito pelo teste)	Negativo	Falso Negativo	Verdadeiro Negativo

Quadro 10 - Matriz de confusão

Fonte: A autora, 2022.

Figura 13 – Cálculo de hipotenusa a partir de uma curva Roc



Fonte: Camuzi, 2019 (cedido pelo autor).

3. RESULTADOS E DISCUSSÃO

3.1 Casuística

Nesse estudo, foram analisados 101 pacientes do sexo masculino com diagnóstico de SXF (n=75) ou com resultados moleculares inconclusivos para a síndrome (n=26), previamente investigados por metodologias de rotina em nosso laboratório. Dados sobre os indivíduos participantes da pesquisa e os resultados das análises moleculares estão disponíveis no Apêndice A. Na Tabela 1, apresentamos um resumo da coorte estudada.

Dados	Coorte analisada
Número de casos (n)	101
Faixa etária (anos)	1 a 30
Média e desvio das idades (anos)	$10,86 \pm 6,00$
Número de casos familiares (%)	60 (59,4%)
Número de casos esporádicos (%)	24 (23,8%)
Número de casos com recorrência familiar desconhecida (%)	17 (16,8%)

Tabela 1 – Dados gerais sobre a amostra de indivíduos estudada

Fonte: A autora, 2022.

3.2 Análise para caracterização de perfis mutacionais atípicos

A partir da análise por TRP-PCR dos 101 indivíduos, cujo resultado prévio sugeriu a presença da mutação completa ou foi inconclusivo, foi observada apenas a mutação completa (FM "pura") em 53 deles (52,5%). Em 28 (27,7%) meninos, foi identificado mosaicismo de tamanho, sendo 16 (15,8%) pacientes com o alelo da pré-mutação concomitante ao da mutação completa (PM+FM), 9
(8,9%) indivíduos com alelos nas faixas normal e de mutação completa (N+FM) e três (3%) portadores de alelos nas faixas normal, de pré-mutação e de mutação completa (N+PM+FM). Além disso, em 13 indivíduos (12,9%), os resultados foram inconclusivos ou as reações não funcionaram pela metodologia de TRP-PCR (Tabela 2). Os eletroferogramas gerados a partir da amplificação do DNA para os diferentes perfis encontrados estão exemplificados na Figura 14. Algumas amostras apresentaram alelos com baixa amplitude, por TRP-PCR. Nesses casos, é possível que os alelos estejam presentes em baixas proporções ou que este resultado represente um artefato da metodologia. Nesses casos, ainda vale destacar que a reação pode ter apresentado uma amplificação preferencial de algum alelo em detrimento de outro(s).

Tabela 2 – Dados genotípicos sobre a amostra de indivíduos estudada, avaliados por TRP-PCR

Genótipo	Número de indivíduos
FM	53 (52,5%)
PM	1 (1%)
I	3 (3%)
Ν	3 (3%)
PM+FM	16 (15,8%)
N+FM	9 (8,9%)
N+PM+FM	3 (3%)
Inconclusivo / Não funcionou	13 (12,9%)
Total	101

Legenda: FM: alelo de mutação completa; PM: alelo de pré-mutação; N: alelo normal; I: alelo intermediário; PM+FM: mosaicismo de alelos de pré-mutação e de mutação completa; N+FM: mosaicismo de alelos normal e mutação completa; N+PM+FM: mosaicismo de alelos na faixa normal, de pré-mutação e de mutação completa.

Fonte: A autora, 2022.



Figura 14 – Resultados da amplificação de amostras de DNA de pacientes, ilustrando os diferentes perfis encontrados (Continua)











Legenda: Eletroferogramas referentes aos pacientes (A) P4093/13, com o alelo normal de 42 CGGs não metilado (2%); (B); P4257/14, portador de um alelo intermediário de 52 CGGs não metilado (0%) e (C) P4101/13, que apresenta a mutação completa (FM) com mais de 200 repetições CGGs completamente metilada. Além disso, podemos observar mosaicismo de tamanho nos pacientes (D) P3446/11, portador de uma pré-mutação não metilada (0%) de 80 repetições e de uma mutação completa completamente metilada; (E) P1259/06, com ambos os alelos normal não metilados (2%) de 39 CGGs, e de mutação completa completamente metilada e (F) P370/99, portador de um alelo normal parcialmente metilado (36%) de 30 repetições, uma pré-mutação não metilada (9%) de 59 CGGs e uma possível mutação completa com baixo rendimento de amplificação no final do quadro de leitura. Dig e Dig_g: controle de digestão; Ref e Ref_g: controle de referência; NOR e NOR_g: alelo normal; INT e INT_g: alelo intermediário; PM e PM_g: alelo de pré-mutação; FM e FM_g: alelo de mutação completa. A reação de digestão é amplificada com oligonucleotídeos marcados com corante HEX (verde), enquanto a reação controle com oligonucleotídeos marcados com FAM (azul). Dessa forma, o número de repetições é visualizado em FAM e a porcentagem de metilação é acessada em HEX.

Fonte: A autora, 2022.

F

3.2.1 Identificação de deleções próximas às repetições CGG

Em busca de deleções próximas à região de repetições CGG, foi realizada um ensaio de PCR, seguido da eletroforese. Os fragmentos de interesse foram então extraídos do gel, purificados e sequenciados. Nessa análise, foram incluídos probandos com suspeita de deleção nas reações de rotina e portadores de mosaicismo, envolvendo alelos na faixa normal.

Por limitações metodológicas, dos 9 pacientes nos quais detectamos uma possível deleção, somente foi possível acessar a sequência do menor fragmento referente ao paciente P928/04. Os eletroferogramas dos demais pacientes apresentaram muitos ruídos e sobreposições. Entretanto, foi possível estimar o tamanho aproximado dessa redução na extensão do fragmento, a partir de sua migração na eletroforese, considerando que a sequência selvagem descrita nos bancos de dados online contabiliza 20 repetições CGGs (Ensembl genome browser em hg38, Human BLAT Search) (Tabela 3).

Paciente	Resultado da TRP-PCR	Redução estimada na extensão do fragmento na eletroforese	Condições da PCR
P420/99	N+FM	~50 pb e ~80 pb	Araújo, 2018
P718/03	N+FM	~10 pb, ~70 pb e ~100 pb	Araújo, 2018
P928/04	N+FM	~120pb e ~20 pb	Araújo, 2018
P2315A/09	FM	~100 pb	Haddad <i>et al</i> ., 1996
P3274/11	FM	~50 pb e ~100 pb	Haddad <i>et al</i> ., 1996
P3446/11	PM+FM	~70 pb	Haddad <i>et al</i> ., 1996
P4055/13	N+PM+FM	~100 pb	Haddad <i>et al</i> ., 1996
P4101/13	FM	~40 pb, ~70 pb e ~90 pb ~100 pb e ~120 pb	Araújo, 2018; Haddad <i>et al</i> ., 1996
P4210A/14	PM+FM	~50 pb	Araújo, 2018

Tabela 3 – Estimativa aproximada da extensão dos fragmentos observados neste trabalho, sugerindo a presença de deleção

Legenda: Na coluna de resultado da TRP-PCR, temos os alelos referentes às repetições CGG para cada indivíduo, observados por TRP-PCR utilizando o kit AmplideX® *FMR1* mPCR. Na coluna de redução estimada na extensão do fragmento na eletroforese, temos a estimativa do tamanho de sequência perdida em cada amostra, quando comparados ao controle normal. Nessa coluna, podemos observar que alguns probandos apresentaram mais de um fragmento reduzido, visualizados na eletroforese. Já na coluna PCR, temos a reação utilizada para amplificar cada amostra. Aqui, podemos perceber que a amostra do paciente P4101/13 foi amplificada por ambos os ensaios de PCR e apresentou fragmentos com tamanhos brevemente diferentes. Esses fragmentos podem ser os mesmos que, devido à diferença nas condições das amplificações, apresentaram pequenas diferenças nas estimativas de tamanho. FM: alelo de mutação completa; PM: alelo de pré-mutação; N: alelo normal; PM+FM: mosaicismo de alelos de pré-mutação completa; N+PM+FM: mosaicismo de alelos normal e mutação completa.

Fonte: A autora, 2022.

Através do sequenciamento pelo método de Sanger, o paciente P928/04 (N+FM) apresentou uma deleção de 38 pb imediatamente à montante das repetições CGG, além de uma região de repetições contendo seis CGGs restantes. Foi observada também microhomologia, ou seja, foram observadas sequências idênticas ou complementares reversas, nas posições que flanqueiam a deleção, constituída por uma sequência de 10 pb (5' GGCGGCGGCG 3') (Figura 15).



Figura 15 – Eletroferograma do sequenciamento referente ao paciente P928/04

Legenda: (A) Eletroferograma do sequenciamento de Sanger do menor fragmento referente ao paciente P928/04 e (B) esquema da deleção de 38 pb, representada pela linha cinza tracejada. A sequência sublinhada em vermelho aponta para a microhomologia composta por um fragmento de 10 pb (5' GGCGGCGGCG 3') no ponto de quebra. Fonte: A autora, 2022.

Além disso, em nossa casuística, mais 5 pacientes com a FM acompanhada de uma deleção próxima à região de repetição já haviam sido identificados em um trabalho anterior do grupo (Gonçalves *et al.*, 2016) (Tabela 4). Inclusive este estudo foi o ponto de partida para outras publicações de diferentes grupos de pesquisa, buscando casos atípicos envolvendo deleções próximas às repetições CGG, no gene *FMR1*.

Tabela 4 – Localização genômica e extensão das deleções identificadas neste estudo e por Gonçalves *et al.*, 2016

Paciente	CGG	Posição (GRCh38/hg38)	Extensão	Microhomologia	Estudo
P928/04	N+FM	147912013- 147912100*	38 pb	10 pb	Este estudo
P1033/05	FM	147912023– 147912150*	128 pb	2 pb	Gonçalves <i>et al</i> ., 2016
P1234/06	FM	147912043– 147912135*	93 pb	Não há	Gonçalves <i>et al.</i> , 2016
P1337/13	FM	147912039– 147912110*	72 pb	3 pb	Gonçalves <i>et al</i> ., 2016
P1513/07	FM	147911695– 147912736**	1042 pb		Gonçalves <i>et al.</i> , 2016
P1629/07	FM	147912021– 147912113*	93 pb	2 pb	Gonçalves <i>et al</i> ., 2016

Legenda: * A posição das deleções é aproximada, devido à presença de microhomologia nos pontos de quebra. ** A extensão precisa dessa deleção é desconhecida. Fonte: A autora, 2022; Adaptado de Gonçalves *et al.*, 2016.

3.2.2 Análise para descartar a presença de um cromossomo X extra

Com o intuito de excluirmos a possibilidade da presença de um cromossomo X extra em pacientes portadores de mosaicismo, envolvendo alelos na faixa normal ou de uma possível deleção, foi realizada a análise de quatro marcadores localizados no gene *SHOX*.

A análise dos marcadores internos ao gene *SHOX* englobou quatro regiões. Por se localizarem na região pseudoautossômica 1 nos cromossomos X e Y, ou seja, apresentar genes homólogos no braço curto do cromossomo X (Xp) e no braço curto do cromossomo Y (Yp) (Binder *et al.*, 2018), é esperado que essas reações amplifiquem dois segmentos com amplitudes semelhantes, se o indivíduo for heterozigoto, ou um segmento com amplitude maior, se o indivíduo for homozigoto para aquela dada posição. A amplificação de três segmentos com amplitudes semelhantes ou de dois segmentos, em que um apresenta amplitude dobrada em relação a outro, sugere a presença de mais de um cromossomo X.

Os resultados referentes aos quatro pares de oligonucleotídeos não evidenciaram a presença de mais um cromossomo X em nenhum dos pacientes testados (Apêndice C). Um exemplo de resultado referente ao par de oligonucleotídeos 2 é apresentado na Figura 16. Figura 16 – Resultados da análise de marcador no gene SHOX, ilustrando a ausência de um segundo cromossomo X nas amostras de pacientes atípicos



Legenda: Eletroferogramas referentes à amplificação por PCR com os oligonucleotídeos nº 2 das amostras de (A) controle normal (Controle 07) e (B) paciente P420/99 (N+FM), mostrando dois alelos amplificados, e (C) controle positivo (P2771/10), mostrando três alelos amplificados, o que representa a presença do segundo cromossomo X. Fonte: A autora, 2022. Um ponto de grande relevância neste estudo é a busca por perfis atípicos relacionados à expansão do gene *FMR1*, como os casos de mosaicismo de tamanho, principalmente envolvendo alelos na faixa normal e deleções. Na SXF, a instabilidade das expansões de repetição CGG pode gerar alelos de tamanhos distintos em diferentes populações de células (Jiraanont *et al.*, 2017). A frequência para casos de mosaicismo de tamanho em nossa amostra foi de 28%. No entanto, Nolin e colaboradores (1994) observaram uma frequência mais alta por *Southern blotting*, visto que quase metade dos pacientes portadores de SXF estudados (41%) apresentaram mosaicismo. Dentro desses casos, 82% apresentaram mosaicismo de mutação completa, em casos bem mais raros há a presença de mosaicismo de mutação completa e alelos normais, assim como alelos de mutação completa e deleções. Em um outro trabalho, Merenstein e colaboradores (1996) relataram um grupo de 27 casos de mosaicismo de mutação completa com pré-mutação em sua amostra de 218 meninos com SXF (12,4%), analisada por sonda direta de DNA genômico.

Já entre os 12 probandos analisados por Jiraanont e colaboradores (2017), através de uma combinação de *Southern blotting* e PCR, oito (66,7%) foram identificados com mosaicismo de tamanho, abrangendo a mutação completa e alelo normal ou intermediário não metilado, dois (16,7%) apresentaram a mutação completa com a pré-mutação e quatro (33,3%) apresentaram alelos mosaicos com deleção localizada na região de repetição CGG e na região flanqueadora.

Além disso, Pandelache e colaboradores (2019) relataram dois gêmeos monozigóticos portadores de mosaicismo de tamanho, identificados através de uma combinação de PCR e *Southern blott* de amostras de DNA extraído a partir de saliva. Enquanto o gêmeo 1 apresentou um mosaicismo de alelo normal de 30 CGGs, pré-mutação com 99 CGGs e mutação completa com rastro entre 388 e 1025 CGGs, o gêmeo 2 apresentou um mosaicismo de pré-mutação com 99 CGGs e mutação com 99 CGGs e mutação de pré-mutação com 99 CGGs e mutação com 99 CGGs e mutação de pré-mutação com 99 CGGs e mutação com 99 CGGs e mutação de pré-mutação com 99 CGGs e mutação com 99

alelo normal de 30 CGGs e uma pré-mutação de 108 CGGs não metilados e uma mutação completa (>200 CGGs) metilada (Pandelache *et al.*, 2019).

Em geral, é esperado encontrar um percentual significativo de casos de mosaicismo envolvendo a pré-mutação e a mutação completa, devido à maior instabilidade desses alelos (Cleary & Pearson, 2005; Nolin *et al.*, 2019). Estima-se que 38% dos portadores de FM são mosaicos para alelos FM adicionais, alelos PM ou ambos (Latham *et al.*, 2014). Todavia, torna-se bastante relevante a presença de mosaicimo de tamanho envolvendo alelos na faixa normal e deleções em nossa casuística, visto que esses casos são mais raramente encontrados e descritos na literatura. Chama também a atenção a identificação de casos de mosaicismo envolvendo três alelos (N+PM+FM), pois estes são ainda mais raros.

O mecanismo que leva à instabilidade na expansão das repetições CGG ainda não está completamente esclarecido (Jiraanont et al., 2017). Embora contrações de alelos mais extensos para alelos menores sejam muito menos frequentes do que a expansão (Quadros 1 e 2) (Nolin et al., 2019), hipóteses foram propostas para tentar explicar o mecanismo por trás das contrações. Uma das hipóteses mais aceitas é um deslize da polimerase no DNA ou desalinhamento de fita mal pareada (SSM) que pode levar a uma redução no comprimento da região repetida, de uma pré-mutação para um alelo de tamanho normal, ou a uma contração de uma mutação completa para um alelo normal ou para deleções (Schmucker & Seidel, 1999; Tabolacci et al., 2008). Esse fenômeno pode ocorrer no início da embriogênese em uma porcentagem de células, onde a presença de longas sequências repetidas pode induzir o emparelhamento incorreto de novas fitas sintetizadas durante a replicação do DNA (Chiurazzi et al., 1994; Arocena et al., 2000). Foi demonstrado que a metilação de CpGs e SSM podem facilitar a ocorrência de eventos de deleção e, com isso, a região rica em CpG pode ser propensa a rearranjos, podendo resultar em deleções que se estendam para a seguência do promotor a montante da repetição CGG (Edamura & Pearson, 2005). A deleção pós-zigótica pode resultar em alelos de tamanho normal, como observado em indivíduos com SXF que apresentam mosaicismo de tamanho (Orrico et al., 1998; Maia et al., 2017). Além disso, um estudo em quatro meninos de nacionalidade portuguesa apresentando mosaicismo de mutação completa e alelo normal sugeriu a presença de um haplótipo que predispõe a grandes contrações (Maia et al., 2017).

Além das expansões, as contrações das repetições e deleções estão também associadas a alelos expandidos (Hayward & Usdin, 2021). Em alguns casos, a contração gera um alelo na faixa normal (Grasso *et al.*, 1999; Maia *et al.*, 2017; Manor *et al.*, 2017; Prawer *et al.*, 2018), mas em outros, ocorrem deleções envolvendo as repetições e um segmento variável em uma ou em ambas as sequências flanqueadoras (Gonçalves *et al.*, 2016; Hayward & Usdin, 2021). As consequências dessas deleções dependem de quão longe elas se estendem nas regiões flanqueadoras (Hayward & Usdin, 2021). Alguns casos envolvem uma perda mínima de sequência flanqueadora, resultando em um fenótipo normal se o número de repetições estiver agora na faixa normal e sem a presença de outros alelos (Tabolacci *et al.*, 2020; Erbs *et al.*, 2021). Enquanto outros casos envolvem uma deleção de todo o gene *FMR1* ou regiões críticas do exon 1, o que pode resultar em um déficit de FMRP com os mesmos sintomas observados em portadores da FM (Grasso *et al.*, 1999; Hayward & Usdin, 2021).

Eventos de contração que geram um único alelo presente em todas as células provavelmente ocorrem pré-zigoticamente (Manor *et al.*, 2017; Prawer *et al.*, 2018; Tabolacci *et al.*, 2020), mas contrações também podem ocorrer pószigoticamente, resultando em indivíduos mosaicos para alelos de tamanhos diferentes (Ferreira *et al.*, 2013; Gonçalves *et al.*, 2016; Maia *et al.*, 2017; Prawer *et al.*, 2018; Hayward *et al.*, 2019). Além disso, diversas alterações no gene *FMR1* associadas à sintomas da SXF, como variantes *missense, nonsense, frameshift,* inserções e deleções na região promotora, em sequência codificante e nas bordas intron/exon, já foram descritas na literatura (Sitzmann *et al.*, 2018)

Ainda, a maioria dos pontos de quebra de deleções maiores que foram caracterizados a nível de sequência estão associados a microhomologias de 2 a 9 nucleotídeos (Grasso *et al.*, 1999; Fan *et al.*, 2005; Han *et al.*, 2006; Luo *et al.*, 2014; Gonçalves *et al.*, 2016; Tabolacci *et al.*, 2020; Erbs *et al.*, 2021). Apesar da riqueza em CG na região do gene *FMR1*, é possível que as microhomologias estejam envolvidas no mecanismo que gera essas deleções, visto que essas microhomologias são marcas de um processo conhecido como junção de extremidade mediada por microhomologia (MMEJ, do inglês, *Microhomology-mediated end joining*) (Hayward & Usdin, 2021). Esse processo envolve a ressecção do ponto de quebra da fita dupla (DSB) seguido por uma busca por microhomologias curtas nas duas caudas de fita simples de DNA, hibridização das

regiões de microhomologia, remoção de sequências homólogas, preenchimento da lacuna e, finalmente, ligação (Figura 17) (Hayward & Usdin, 2021).

Figura 17 - Representação de potenciais vias para geração de contrações e

А В Deslize da fita Junção mediada por microhomologia Grampo na fita molde promove dissociação da polimerase DSB Realinhamento de Ressecção de extremidade 5'-3' primer / fita molde Mre11+ CtIP Ligação de microhomologias Pol0/ Realinhamento de prime Polß TTTTT $\Pi\Pi$ à montante na fita molde Remoção de seguências Lig3/ e reparo de lacunas Lig1 Nova replicação da DNA fita molde contraída polimerase Ligação ____ TITT

Pequena contração

Grande contração

Legenda: (A) Deslize da fita molde durante a replicação pode ser exacerbada devido à formação de grampos na fita molde. A ligação de primer mais à montante na fita molde pode levar a novas fitas com menos repetições do que a fita molde. As novas replicações podem gerar o alelo contraído. (B) Reparo de quebra de fita dupla (DSB) por junção de extremidade mediada por microhomologia (MMEJ) é iniciado pela ressecção de extremidade para revelar microhomologia em ambas as extremidades do ponto de quebra. O anelamento dessas microhomologias é seguido pela remoção das sequências não homólogas, reparo de lacunas e ligação.

Fonte: Adaptado de Hayward et al., 2021.

O ensaio de PCR convencional para o rastreamento de expansões CGG (Fu *et al.*, 1991; Haddad *et al.*, 1996), rotineiramente utilizada para diagnóstico da SXF, tem a vantagem de ser uma técnica com bom custo benefício. Contudo, essa análise é baseada na ausência de amplificação do fragmento expandido na mutação completa e pré-mutações muito grandes (em geral, a amplificação ocorre em alelos com tamanhos menores ou iguais a 100 até 120 repetições). Assim, apenas alelos normais, intermediários e pré-mutações menores podem ser visualizados. Logo, ainda gera muitos resultados inconclusivos. Além disso, para os casos de mosaicismo, essa técnica não se mostra eficaz, devido à amplificação

deleções

preferencial do menor alelo, seja um alelo normal, intermediário, uma pré-mutação pequena ou uma deleção.

Já a técnica de TRP-PCR, embora tenha alto custo, possui vantagens, como a amplificação de alelos normais, intermediários, pré-mutados, algumas mutações completas menos extensas e casos de mosaicismo de tamanho. A versão dessa técnica em que é possível acessar o status de metilação de cada alelo apresentado pelas amostras analisadas tornou essa técnica ainda mais vantajosa como método de identificação para a SXF. Porém, essa técnica não abrange a detecção de alelos com deleções, já que podem apresentar diferentes extensões e há uma região única amplificada pelos oligonucleotídeos do kit. Outra limitação desta metodologia envolve a amplificação com baixo rendimento de alguns alelos em casos de mosaicismos de tamanho, gerando dúvidas na análise. Por isso, o emprego de apenas uma metodologia molecular para esclarecimento do espectro de variantes relacionadas à SXF ainda não é ideal. Portanto, para melhor caracterizar os perfis mutacionais de SXF atípicos, uma combinação de abordagens, como aconselhado por Gonçalves e colaboradores (2016), e PCRs mais robustas, como o kit AmplideX®, seriam o mais indicado para garantir que esses tipos de mosaicos não sejam subdiagnosticados e potencialmente perdidos pelos testes de rotina para SXF (Jiraanont et al., 2017).

Além disso, a determinação dos perfis mutacionais atípicos subdetectados pode ter implicações terapêuticas, visto que, em geral, as pesquisas para novos fármacos potenciais para o tratamento da SXF são baseadas na presença de hipermetilação, que é encontrada em portadores da mutação completa (Bar-Nur *et al.*, 2012). A estratégia farmacêutica de restauração da atividade do gene *FMR1* tem como alvo modificações potencialmente reversíveis, dentre elas a própria metilação de DNA. A primeira droga testada em células de pacientes com SXF foi a 5-azadesoxicitidina (5-azadC), um inibidor de metiltransferase que conseguiu restaurar a transcrição do gene *FMR1 in vitro* (Chiurazzi *et al.*, 1998). Estudos recentes mostraram um efeito duradouro do 5-azadC na reativação do gene *FMR1* mutante, sem afetar a integridade da borda de metilação identificada por Naumann *et al* (2009), o que sugere que esse efeito não seja randômico, mas restrito a regiões específicas (Figura 18) (Tabolacci *et al.*, 2016).

Figura 18 – Perfis de metilação na borda de metilação próxima à região promotora do gene *FMR1* em linhagens celulares de homens com e sem SXF antes e após o tratamento com o agente desmetilante 5azadesoxicitidina (5-azadC)



	 	 	••	•••	 	 		 	•••	 	•••	 	•••	 		 	•••	 	•••	 	 		
							يعطك				_											í – I	

Legenda: Borda de metilação de Naumann *et al* (2009) no gene *FMR1* após tratamento com 5azadC. (a) O sequenciamento de DNA tratado com bissulfito de sódio de linhagens celulares de controles antes (CT1) e após o tratamento (CT1+5-azadC) por 7 dias não revelou modificação substancial do perfil de metilação, (b) mas foi observada desmetilação quase completa da região promotora em pacientes com SXF após o tratamento (SXF1+5azadC), não afetando a borda de metilação. A borda de metilação é indicada pela seta. Os números 45-82 designam os dinucleotídeos CpG na região: ■ metilados e □ não metilados.

Fonte: Adaptada de Tabolacci et al., 2016.

O ácido valpróico já foi testado em um ensaio de segurança em 10 meninos com SXF e resultou em uma melhora geral da hiperatividade e déficit de atenção (Torrioli *et al.*, 2010). Entretanto, como esses tratamentos são baseados na reversão do silenciamento da mutação completa hipermetilada e em uma sequência intacta do gene, casos de SXF atípicos não estariam sendo cobertos por esses tratamentos. Assim, metodologias que buscam a identificação apenas da mutação completa podem levar a um subdiagnóstico de casos de mosaicismo e o tratamento futuro voltado para agentes desmetilantes poderia não obter os mesmos efeitos nesses indivíduos.

Além da terapia farmacológica, avanços na tecnologia de modificação de DNA, como a CRISPR/Cas9, já foram utilizados para proporcionar a desmetilação do gene *FMR1* em células-tronco pluripotentes induzidas (Park *et al.*, 2015; Liu *et al.*, 2018). Park e colaboradores (2015) demonstraram que, após a correção da expansão da repetição CGG em células-tronco, a ilha CpG a montante da região promotora do gene *FMR1* apresentou desmetilação extensa, um estado de cromatina ativa e início da transcrição, com a expressão do gene restaurada e mantida em células precursoras neurais e neurônios maduros. Liu e colaboradores (2018) relataram que a expressão do gene nos neurônios editados foi mantida *in vivo* após inseridas no cérebro de rato e que a desmetilação das repetições nos neurônios pós-mitóticos também reativou o *FMR1*. Todavia, obstáculos ainda existem até que essas terapias possam ser clinicamente praticadas para o tratamento de pacientes com SXF ou outros distúrbios (Hagerman *et al.*, 2017). Além disso, não se sabe se pacientes com mosaicismos poderiam ser contemplados por essa terapia.

Ainda, a presença de uma pré-mutação com a mutação completa nos casos de mosaicismo pode levar ao desenvolvimento da SXF e da FXTAS, já que a presença da pré-mutação leva a um aumento tóxico nos níveis de RNAm (Pretto *et al.*, 2013). Já que, em geral, a pré-mutação não é metilada, esses pacientes também não seriam beneficiados totalmente com o uso de agentes desmetilantes. Assim, a determinação dos perfis atípicos se torna também importante nesses casos para o direcionamento terapêutico.

3.3 Análise de marcadores epigenéticos por Bioinformática

A figura 19 ilustra a complexidade da arquitetura genética, sob regulação epigenética, no início do gene *FMR1*, desenvolvida a partir de dados da literatura e de bancos de dados públicos.

Figura 19 – Localização de elementos genéticos e epigenéticos no início do gene *FMR1*



Legenda: O mapa parcial mostra o gene *FMR1* até o exon 3 em verde, além do segmento do genoma 5' a montante em cinza, incluindo as repetições CGG em amarelo. A seta indica a região promotora do gene *FMR1*. Todos dinucleotídeos CpGs na região estão marcados pelos símbolos 1. O esquema apresenta as posições das sondas de metiloma em azul e as regiões analisadas neste estudo (borda de metilação e intron 2) em lilás. Os Elementos Epigenéticos Relacionados ao X-Frágil 1 e 2 (FREE1 e FREE2) estão em laranja. Além disso, ainda estão demarcadas as regiões do gene antisenso *FMR1-AS1* e da ilha CpG.

Fonte: A autora, 2022.

O *heatmap* gerado por Bioinformática para a identificação de marcadores epigenéticos no gene *FMR1* contou com a comparação de dados de metilação de todas as sondas disponíveis em banco de dados online (GEO database) em região à montante e ao longo do gene *FMR1* em pacientes com a SFX e em indivíduos controles (Figura 20). Nessa análise, foram incluídos dados de amostras referentes a indivíduos de diferentes etnias e de ambos os sexos.



Figura 20 – Heatmap com dados de metilação no gene FMR1 gerado a partir de bancos de dados online

Fonte: A autora e Camuzi, 2019 (cedido pelo autor).

A partir do *heatmap*, pudemos observar níveis aumentados de metilação na maioria das sondas referentes aos pacientes com SXF, em comparação com as amostras controles. Esses dados estão relacionados ao padrão diferencial de metilação na região promotora do gene *FMR1*, corroborando a literatura. A imagem mostra também valores medianos de metilação nas sondas das amostras femininas, o que está relacionado com a inativação de um dos cromossomos X. Entretanto, a partir dos dados das sondas utilizadas não foi possível observar com clareza as CpGs que constituem a borda de metilação, relatada por Naumann *et al* (2009, 2014).

3.4 Quantificação de metilação da borda por pirosequenciamento

A quantificação de metilação por pirosequenciamento se estendeu à borda de metilação, descrita por Naumann e colaboradores (2009; 2014). Essa região foi analisada de forma segmentada em duas leituras independentes, devido à sua extensão, com o oligonucleotídeo A abrangendo as CpGs 71 a 69 e o oligonucleotídeo B acessando as CpGs 68 a 65. Na Tabela 5, podemos observar os valores de metilação obtidos.

Registro	Genótipo	Oligo	nucleotío	deo A		Oligonucleotídeo B				
		CG 71	CG 70	CG 69	CG 68	CG 67	CG 66	CG 65		
Controle 01	N	82%	72%	55%	29%	55%	42%	31%		
Controle 02	N	78%	68%	51%	28%	48%	37%	25%		
Controle 03	N	83%	71%	56%	37%	53%	45%	31%		
Controle 04	N	86%	71%	46%	26%	55%	42%	31%		
Controle 05	N	83%	69%	47%	28%	53%	40%	27%		
Controle 06	N	82%	71%	57%	32%	59%	50%	34%		
Controle 07	N	78%	63%	44%	24%	40%	32%	22%		
Controle 08	N	82%	69%	38%	27%	51%	41%	28%		
Controle 09	N	76%	65%	47%	31%	51%	39%	24%		
Controle 10	N	81%	74%	53%	30%	54%	41%	29%		
Controle 11	N	85%	73%	46%	29%	56%	40%	32%		
Controle 12	Ν	78%	67%	49%	27%	43%	38%	25%		
Controle 13	Ν	81%	72%	54%	28%	50%	43%	31%		
Controle 14	Ν	82%	72%	42%	29%	52%	41%	30%		
Controle 15	N	77%	62%	44%	25%	40%	29%	22%		
Controle 16	N	75%	64%	43%	25%	39%	31%	22%		
Controle 17	Ν	83%	72%	53%	37%	54%	48%	35%		
Controle 18	Ν	72%	59%	38%	33%	52%	36%	37%		
P4257/14	I	86%	73%	53%	25%	51%	44%	28%		
P4259/14	I	87%	76%	50%	38%	60%	49%	39%		
P4468/15	I	80%	64%	44%	30%	47%	37%	27%		
P1348A/06	PM	84%	72%	57%	NF	NF	NF	NF		
P3946B/16	PM	84%	70%	49%	30%	51%	49%	35%		
P4484D/16	PM	75%	65%	43%	29%	46%	35%	31%		
P076/95	FM	91%	94%	100%	79%	92%	93%	86%		
P138/96	FM	93%	92%	95%	42%	89%	84%	83%		
P248/97	FM	91%	90%	100%	66%	89%	90%	82%		
P248/02	FM	93%	93%	100%	71%	91%	93%	85%		
P312/98	FM	88%	100%	100%	60%	91%	89%	83%		
P344/98	FM	87%	100%	100%	56%	94%	100%	89%		
P400/99	FM	84%	94%	100%	55%	92%	94%	88%		
P442/99	FM	94%	93%	100%	66%	93%	100%	89%		
P457/99	FM	91%	92%	99%	55%	85%	88%	83%		
P457A/99	FM	91%	90%	100%	58%	91%	100%	88%		
P495/00	FM	87%	91%	100%	68%	92%	100%	88%		
P504/00	FM	91%	92%	100%	57%	91%	100%	87%		
P562/01	FM	93%	100%	100%	57%	92%	90%	88%		
P632/01	FM	100%	100%	100%	76%	93%	95%	89%		
P663/02	FM	90%	100%	100%	75%	90%	100%	88%		
P672/02	FM	90%	95%	100%	62%	94%	93%	91%		
P702/02	FM	84%	88%	100%	66%	100%	100%	100%		
P850/04	FM	93%	92%	98%	53%	92%	91%	89%		

Tabela 5 – Valores de metilação nas CpGs referentes à borda de metilação na amostra de indivíduos estudada (continua)

Registro	Genótipo	Oligo	onucleotí	deo A	(Oligonuc	leotídeo E	3
		CG 71	CG 70	CG 69	CG 68	CG 67	CG 66	CG 65
P873/04	FM	93%	95%	100%	66%	87%	88%	86%
P908/04	FM	86%	88%	100%	62%	86%	87%	86%
P1013/05	FM	87%	89%	100%	65%	90%	100%	84%
P1101A/05	FM	100%	100%	100%	50%	89%	93%	86%
P1180/06	FM	88%	100%	100%	49%	92%	91%	87%
P1201/06	FM	89%	89%	100%	56%	89%	100%	76%
P1348/06	FM	92%	92%	100%	72%	100%	87%	88%
P1371/07	FM	90%	90%	100%	60%	90%	92%	87%
P1467/07	FM	96%	100%	100%	63%	92%	92%	88%
P1628/07	FM	100%	100%	100%	63%	100%	100%	86%
P2149/09	FM	86%	100%	100%	63%	91%	94%	87%
P2273/09	FM	100%	89%	100%	60%	90%	91%	87%
P2315A/09	FM	93%	95%	100%	72%	90%	100%	87%
P3207/11	FM	83%	100%	100%	54%	91%	100%	90%
P3274/11	FM+del?	83%	92%	100%	57%	93%	94%	89%
P3384/11	FM	88%	88%	100%	58%	83%	81%	79%
P3946/12	FM	89%	91%	100%	61%	92%	89%	88%
P4101/13	FM	85%	89%	100%	54%	91%	100%	87%
P4112/13	FM	89%	87%	100%	54%	86%	86%	81%
P4194/14	FM	82%	90%	100%	57%	100%	100%	83%
P4484/16	FM	93%	95%	100%	63%	100%	100%	86%
P4644/19	FM	91%	91%	100%	63%	92%	93%	88%
P1033/05	FM+del	86%	97%	100%	71%	100%	100%	85%
P1234/06	FM+del	94%	94%	100%	63%	91%	92%	88%
P1337/13	FM+del	84%	88%	98%	67%	88%	90%	87%
P1629/07	FM+del	83%	84%	100%	67%	89%	90%	86%
P138A/96	PM(?)+FM	81%	84%	91%	52%	79%	81%	71%
P400A/99	PM+FM	91%	88%	100%	61%	89%	100%	82%
P400B/99	PM+FM	90%	100%	100%	50%	89%	100%	83%
P452/99	PM+FM	92%	90%	92%	60%	88%	88%	82%
P1101C/05	PM+FM	92%	90%	96%	65%	88%	90%	83%
P3199/11	PM+FM	89%	91%	95%	54%	91%	91%	85%
P3274B/13	PM+FM	84%	81%	NF	52%	NF	NF	NF
P3350/11	PM+FM	92%	94%	100%	53%	91%	92%	86%
P3446/11	PM+FM+del?	87%	91%	94%	60%	83%	84%	77%
P4121/13	PM+FM	100%	87%	99%	44%	84%	89%	80%
P4210/14	PM+FM+del?	75%	60%	56%	34%	54%	44%	36%
P4210A/14	PM+FM+del?	77%	62%	49%	30%	52%	44%	37%
P420/99	N+FM	94%	91%	93%	71%	84%	80%	76%
P435/99	N+FM	93%	100%	100%	78%	100%	93%	90%
P476/00	N+FM	93%	94%	91%	65%	93%	94%	88%
P537/00	N+FM	88%	88%	100%	50%	91%	91%	85%

Tabela 5 – Valores de metilação nas CpGs referentes à borda de metilação na amostra de indivíduos estudada (continuação)

Registro	Genótipo	Oligo	onucleotí	deo A	Oligonucleotídeo B						
		CG 71	CG 70	CG 69	CG 68	CG 67	CG 66	CG 65			
P668/02	N+FM	92%	91%	100%	51%	90%	100%	86%			
P718/03	N+FM+del?	92%	92%	100%	66%	93%	90%	88%			
P928/04	N+FM	93%	100%	100%	72%	93%	95%	89%			
P981/05	N+FM	88%	93%	100%	57%	92%	100%	85%			
P1259/06	N+FM	89%	100%	100%	59%	93%	100%	89%			
P4442/15	N+FM	90%	91%	93%	55%	88%	86%	84%			
P237/97	N+PM+FM	94%	89%	94%	46%	88%	91%	83%			
P370/99	N+PM+FM	93%	100%	100%	66%	90%	93%	85%			

Tabela 5 – Valores de metilação nas CpGs referentes à borda de metilação na amostra de indivíduos estudada (conclusão)

Legenda: CG: CpG; FM: alelo de mutação completa; PM: alelo de pré-mutação; N: alelo normal; I: alelo intermediário; PM+FM: mosaicismo de alelos de pré-mutação e mutação completa; N+FM: mosaicismo de alelos normal e mutação completa; N+PM+FM: mosaicismo de alelos normal, pré-mutação e mutação completa; del?: possível deleção; del: deleção confirmada; NF: não funcionou.

Fonte: A autora, 2022.

Nas Figuras 21 e 22, podemos observar que no grupo de amostras controles, as porcentagens de metilação sofreram uma redução a partir da CpG 70. Esse achado sugere a presença da borda de metilação proposta por Naumann e colaboradores (2009; 2014), que separa a região à montante da CpG 71 de uma região livre de metilação à jusante da sequência analisada, incluindo o promotor do gene *FMR1*. No grupo de pacientes com o alelo de mutação completa, os valores de metilação se mantiveram superiores aos do grupo controle. Isso sugere que há a perda dessa borda de metilação em indivíduos com mais de 200 repetições CGG.

Figura 21 – Gráfico de linhas com os dados de metilação referentes às CpGs 65 a 71, avaliada por pirosequenciamento



Legenda: CG: CpG; CT: grupo controle; FM: alelo de mutação completa; PM: alelo de pré-mutação; N: alelo normal; I: alelo intermediário; PM+FM: mosaicismo de alelos de pré-mutação e mutação completa; N+FM: mosaicismo de alelos normal e mutação completa; N+PM+FM: mosaicismo de alelos normal, pré-mutação e mutação completa; del: deleção. Entre parênteses, o número amostral. Valores descritos como média ± desvio padrão. As distâncias entre as CpGs 71 e 70 é de 4 pb; entre as CpGs 70 e 69 é de 44 pb; entre as CpGs 69 e 68 é de 45 pb; entre as CpGs 68 e 67 é de 22 pb; entre as CpGs 67 e 66 é de 4 pb; e entre as CpGs 66 e 65 é de 18 pb.

Fonte: A autora, 2022.



Figura 22 – Box plots com os dados de metilação referentes às CpGs 65 a 71, avaliada por pirosequenciamento (continua)



Figura 22 – Box plots com os dados de metilação referentes às CpGs 65 a 71, avaliada por pirosequenciamento (continua)



Figura 22 – Box plots com os dados de metilação referentes às CpGs 65 a 71, avaliada por pirosequenciamento (continua)



Figura 22 – Box plots com os dados de metilação referentes às CpGs 65 a 71, avaliada por pirosequenciamento (continua)

Legenda: Gráficos para cada CpG avaliada, sendo (A) CG 71, (B) CG 70, (C) CG 69, (D) CG 68, (E) CG67, (F) CG66 e (G) CG65. CG: CpG; CT: grupo controle; FM: alelo de mutação completa; PM: alelo de pré-mutação; N: alelo normal; I: alelo intermediário; PM+FM: mosaicismo de alelos de pré-mutação e mutação completa; N+FM: mosaicismo de alelos normal e mutação completa; N+PM+FM: mosaicismo de alelos normal, prémutação e mutação completa. Entre parênteses, o número amostral. Valores descritos como média ± desvio padrão. Acima das chaves constam os valores de p, obtidos através do teste não paramétrico de Kruskal-Wallis (Kruskal & Wallis, 1952), seguida pelo pós-teste de Dunn para múltiplas comparações (Dunn, 1964). Diferença estatística obtida guando p<0.05.

Apesar dos pequenos números amostrais em cada subgrupo alélico, os indivíduos com mosaicismos de tamanho parecem seguir o mesmo padrão dos pacientes com o perfil clássico de FM. Ou seja, a borda de metilação parece ser perdida na presença do mosaicismo quando um alelo FM está presente.

Ainda, as amostras com o alelo pré-mutado apresentam padrão semelhante aos controles. Esses indivíduos são avôs maternos de pacientes com SXF, mostrando que as mães desses meninos com a síndrome herdaram por via paterna a pré-mutação, que por sua vez deu origem à mutação completa na prole. Esses homens pré-mutados não apresentam DI.

Além disso, um achado que chama a atenção é a redução nos níveis de metilação na CpG 68. Esse padrão acontece em todos os indivíduos analisados, tanto pacientes como controles. A razão para essa queda ainda não está clara. De acordo com o UCSC Genome Browser, não há nenhuma variante de nucleotídeo único (SNV; *single nucleotide variant*) próxima a essa posição ou elementos repetitivos na região, mas há uma possível região promotora para o gene *FMR1* na região (em hg19) (UCSC Genome Browser, 2020).

Naumann e colaboradores (2009) identificaram que sequências do promotor e da região 5' a montante do gene *FMR1* completamente metiladas *de novo* apresentam CpGs isoladas não metiladas em amostras de DNA de pacientes com SXF. Estudos relataram que, na região ao redor da CpG 30, algumas das CpGs isoladas não metiladas coincidem com sítios de ligação para fatores de transcrição (*USF1* e *USF2* ou *NRF1*) (Schwemmle *et al.*, 1997; Pietrobono *et al.*, 2002). CpGs isoladas não metiladas também ocorrem no genoma metilado *de novo* de adenovírus tipo 12 em linhagens celulares animais transformadas e mantidas em cultura (Hochstein *et al.*, 2007). De acordo com Naumann e colaboradores (2009), essas CpGs isoladas não metiladas parecem surgir durante a metilação *de novo* em organismos vivos, mas não durante a metilação *in vitro*. Isso ocorre porque uma estrutura de cromatina específica ou a ligação de proteínas pode impedir a metilação *de novo* nesses locais específicos (Naumann *et al.*, 2009).

Em nosso estudo, foi possível observar a borda de metilação, localizada de 650 a 800 nucleotídeos e dos pares CpGs 65 a 70 à montante da região de repetições CGG no gene *FMR1*, identificada por Naumann e colaboradores (2009; 2014). Essa borda é definida por uma sequência à montante dos dinucleotídeos

CpGs completamente metilados e um trecho à jusante desprovido de metilação (Naumann *et al.*, 2009). Além disso, a região da borda de metilação é conservada em humanos, independente da idade e da linhagem celular e está presente na mesma localização tanto em tecido fetal quanto adulto (Naumann *et al.*, 2009). No entanto, nos meninos com SXF, essa borda é perdida e quase todos os 88 dinucleotídeos CpGs no fragmento de 2.260 pb até a região de repetição estão metilados (Naumann *et al.*, 2009).

Ainda não se sabe como as expansões de trinucleotídeos CGG no FMR1 e a perda da borda de metilação na região 5' se relacionam (Naumann et al., 2014). De acordo com Naumann e colaboradores (2009), a borda de metilação possivelmente coincide com uma estrutura de cromatina específica que, quando desestabilizada, permite que a metilação se espalhe, o que acaba resultando na metilação completa do promotor do gene e em seu silenciamento. Inclusive, já foram relatadas mudanças na estrutura da cromatina na região do FMR1 nos cromossomos de pacientes com SXF (Datta et al., 2011; Kumari et al., 2012; Lanni et al., 2013). No entanto, existe a hipótese de que a metilação pode fluir da região de repetição expandida e atingir o promotor e a borda de metilação (Naumann et al., 2014). Além disso, Colak e colaboradores (2014) relataram a formação de um duplex no RNAm do FMR1, incluindo a transcrição da região repetida, e a própria repetição CGG, que pode estar envolvida nas etapas iniciais de silenciamento epigenético do gene, embora possa não explicar a manutenção desse silenciamento. Mas, ainda segundo Naumann e colaboradores (2009), a identificação de complexos específicos entre a sequência da borda e proteínas nucleares pode fornecer alguma pista de como essa borda de metilação é mantida.

O controle epigenético no gene *FMR1* conta com uma complexa regulação transcricional, incluindo a ação de transcritos de RNA (Nobile *et al.*, 2021). Os transcritos de RNA não codificante (ncRNA) são particularmente abundantes no cérebro (Djebali *et al.*, 2012). Alguns ncRNAs longos (IncRNAs) estão envolvidos no controle epigenético do *locus* de onde surgem e sua transcrição pode ocorrer a partir de ambas as fitas (Nobile *et al.*, 2021). O transcrito anti-senso mais abundante para o *locus FMR1*, conhecido como FMR1-AS1, é expresso em altos níveis no cérebro de indivíduos não afetados (Lanni *et al.*, 2013; Nobile *et al.*, 2021). Ele tem início no intron 2 do gene *FMR1* e está ausente em pacientes com

SXF, sendo superexpresso em células com a PM, exatamente como o transcrito senso (Ladd *et al.*, 2007; Lanni *et al.*, 2013).

Um ncRNA adicional, conhecido como FMR4, também foi identificado. Ele surge na orientação anti-senso em relação ao *locus FMR1* e é particularmente expresso no córtex frontal adulto e hipocampo (Khalil *et al.*, 2008). Novamente, em pacientes com SXF não há transcrição do FMR4, enquanto os alelos PM apresentam uma leve superexpressão em relação aos alelos normais (Khallil *et al.*, 2008; Nobile *et al.*, 2021). FMR5 e FMR6 são ainda dois ncRNAs adicionais que surgem do *locus FMR1* (Pastori *et al.*, 2014). FMR5 é transcrito 1 kb a montante do sítio de início da transcrição do *FMR1* na orientação senso, com seu sítio de início de transcrição mapeando na região metilada a montante da borda de metilação (Pastori *et al.*, 2014; Nobile *et al.*, 2021). Esse transcrito é expresso em várias áreas do cérebro humano de indivíduos não acometidos e portadores da PM (Pastori *et al.*, 2014; Nobile *et al.*, 2021). Já o FMR6 é transcrito da fita antisenso, sobrepondo-se à região 3' UTR e aos três últimos exons do gene *FMR1* (Pastori *et al.*, 2014; Nobile *et al.*, 2021). Esse transcrito é silenciado em portadores da SXF e da PM (Pastori *et al.*, 2014; Nobile *et al.*, 2021). Esse transcrito é silenciado em

Ainda que os transcritos anti-senso possam apresentar papéis relevantes na regulação epigenética da transcrição do *FMR1*, suas funções exatas ainda não foram esclarecidas (Nobile *et al.*, 2021). Em outros *loci*, esses transcritos não codificantes induzem a montagem de proteínas envolvidas na organização da heterocromatina ou direcionam o recrutamento de maquinaria para silenciar o *locus* (Wang *et al.*, 2011). Ainda, levantou-se a hipótese de que as proteínas do grupo Polycomb (PcGs) são direcionadas para a repetição CGG do gene *FMR1* através de ncRNAs provenientes do próprio *locus* (Rinn *et al.*, 2007; Pandey *et al.*, 2008). Essa hipótese é apoiada pela evidência de que a maioria das proteínas PcG tem como alvo regiões ricas em GC, e com isso, as repetições CGG do gene *FMR1* podem ser silenciadas por esses complexos (Rinn *et al.*, 2007; Pandey *et al.*, 2008; Nobile *et al.*, 2021).

Recentemente, foi demonstrado que ncRNAs específicos se ligam a DNA metiltransferase 1 (DNMT1) (Di Ruscio *et al.*, 2013). A DNMT1 atua na manutenção da metilação do DNA e apresenta maior afinidade por dinucleotídeos CpG hemimetilados do que por DNA não metilado (Hermann *et al.*, 2004). Estudos sugerem que a DNMT1 é sequestrada por RNAs que atuam como

escudo, impedindo a metilação do *locus* (Di Ruscio *et al.*, 2013; Nobile *et al.*, 2021).

Além disso, a proteína CTCF se liga à borda de metilação e essa interação é perdida em promotores metilados nos genomas de indivíduos com SXF (Lanni *et al.*, 2013). Essa proteína pode desempenhar um papel complexo na regulação da atividade do gene *FMR1* (Lanni *et al.*, 2013). Além disso, o tratamento com agente desmetilante (5-azadC) não restaura a ligação dessa proteína ao gene *FMR1* reativado (Lanni *et al.*, 2013). Dados *in sílico* afirmam que uma alça de cromatina mediada por homodímeros de CTCF pode existir entre o intron 2 e a região da borda de metilação ou promotor em alelos normais e não metilados (Lanni *et al.*, 2013).

De acordo com os dados obtidos por Naumann e colaboradores (2014), no sítio de ligação da proteína CTCF, próximo à borda de metilação na região 5' do gene *FMR1*, há uma SNV 3,6 vezes mais frequente em genomas expandidos e 2,4 vezes mais frequente em genomas que perderam a borda de metilação. Inclusive, um estudo anterior identificou 54 SNVs comuns na região do gene *FMR1* (Brightwell *et al.*, 2002). Além disso, dois RNAs longos não codificantes foram apontados como transcritos preferencialmente na SXF e na FXTAS (Pastori *et al.*, 2014).

Vários trabalhos têm investigado o significado funcional das bordas de metilação de DNA no genoma humano e têm observado instabilidade dessas bordas associada a doenças humanas específicas (Naumann *et al.*, 2014). Castel e colaboradores (2010) detectaram uma repetição de trinucleotídeos CTG livre de CpG que define uma borda de metilação CpG anormal em tecidos de pacientes com distrofia miotônica. Também foi relatada uma borda de metilação de DNA no sítio CpG 44, definida por uma repetição 5'(ATAAA)₁₉₋₂₄-3', na região 5' a montante do gene da glutationa S-transferase (*GSTP1*) em vários tecidos humanos, que é perdida em células de câncer de próstata (Millar *et al.*, 2000). Esse evento também é acompanhado de metilação completa do promotor e silenciamento do gene (Millar *et al.*, 2000). Já a borda de metilação descrita no segmento 5' a montante do promotor do gene da huntingtina (*HTT*) se mantém estável tanto em pacientes com a doença de Huntington quanto em indivíduos saudáveis e em ambos os alelos no cromossomo 4 (Naumann *et al.*, 2014).
Embora pareça haver uma relação entre a expansão da região de repetição CGG e a perda da borda de metilação CpG na região 5' a montante do gene *FMR1*, há exemplos de expansões que devam ser apontados. Portadores de pré-mutação e indivíduos HFMs apresentam uma borda estável, que confere uma proteção contra a disseminação da metilação, a inativação do promotor e o fenótipo da SXF (Lani *et al.*, 2013; Naumann *et al.*, 2014; Tabolacci *et al.*, 2016). A razão para essas divergências na borda de metilação é ainda desconhecida, mas demonstra que expansões de repetições de trinucleotídeos próximas a bordas de metilação podem ser compatíveis com uma borda estável e regiões promotoras não metiladas (Naumann *et al.*, 2014). Além disso, cabe destacar que a borda de metilação ainda não havia sido investigada em indivíduos com mosaicismo de tamanho.

3.5 Análise de quantificação de metilação na posição cg22417678 por pirosequenciamento

Durante o desenvolvimento do *heatmap* (Figura 20), a sonda cg22417678 chamou a nossa atenção. De acordo com os dados, essa sonda apresenta o padrão inverso da maioria das sondas. Nesta posição, as amostras de pacientes com SXF mostram níveis reduzidos de metilação, quando comparadas com as amostras controles. Esse padrão também pode ser observado no gráfico de dispersão (Figura 23A). Considerando essa posição como um possível marcador epigenético, a curva Roc realizada com esses dados mostra valores altos de especificidade e sensibilidade (Figura 23B). Essas observações impulsionaram análises detalhadas dessa posição por pirosequenciamento no presente trabalho. Figura 23 – Distribuição e curva Roc com os dados de metilação para a sonda cg22417678, disponíveis na literatura



Legenda: (a) Gráfico de distribuição, no qual podemos observar valores de metilação nas amostras de pacientes e controles. (b) Curva Roc, mostrando a possibilidade de definição dessa posição como marcador epigenético com altos valores de especificidade e sensibilidade. Fonte: A autora, 2022.

Devido a essa observação no *heatmap* (Figura 16), a CpG referente à sonda cg22417678 foi analisada por pirosequenciamento. Nessa análise, incluiuse também a CpG à jusante mais próxima para comparação. Essa posição se tornou de interesse durante o desenvolvimento deste trabalho por se mostrar um possível marcador epigenético para a SXF e para o desenvolvimento de novas ferramentas de diagnóstico molecular da síndrome. Na Tabela 6, podemos observar os valores de metilação obtidos.

Registro Genótipo Oligonucleotídeo A Posição 1 Posição 2 cg22417678 **Controle 01** Ν 83% 82% Controle 02 Ν 80% 89% **Controle 03** Ν 78% 86% **Controle 04** Ν 82% 77% **Controle 05** Ν 81% 76% **Controle 06** Ν 82% 86% Ν Controle 07 76% 82% 87% **Controle 08** Ν 83% **Controle 09** Ν 81% 88% Ν **Controle 10** 83% 89% **Controle 11** Ν 78% 87% **Controle 12** 77% Ν 88% **Controle 13** Ν 82% 88% **Controle 14** Ν 82% 88% **Controle 15** Ν 80% 89% **Controle 16** Ν 81% 85% **Controle 17** Ν 79% 85% **Controle 18** Ν 87% 88% P4257/14 L 81% 87% P4259/14 Т 80% 75% P4468/15 L 76% 86% PM 78% NF P1348A/06 P3946B/16 PΜ 75% 86% P4484D/16 69% 87% ΡM P076/95 FM 48% 86% P138/96 FΜ 43% 85% P248/97 FM 53% 84% P248/02 FΜ 40% 88% P312/98 FΜ 49% 82% P344/98 FM 23% 87% P400/99 51% 87% FΜ P442/99 FM 42% 87% P457/99 41% FΜ 86% FΜ P457A/99 35% 87% P495/00 FΜ 45% 90% P504/00 FΜ 33% 88% P562/01 83% FM 40% P632/01 FM 41% 85% P663/02 FM 38% 87%

Tabela 6 – Valores de metilação nas CpGs referentes à sonda cg22417678 e na CpG à jusante na amostra de indivíduos estudada (continua)

Registro Genótipo Oligonucleotídeo A Posição 1 Posição 2 cg22417678 P672/02 FM 88% 32% P702/02 FΜ 40% 88% P850/04 FM 47% 87% P873/04 FM 39% 83% P908/04 FM 30% 85% P1013/05 FM 39% 83% P1101A/05 FΜ 38% 88% P1180/06 35% 83% FM P1201/06 FΜ 46% 89% 49% P1348/06 85% FΜ P1371/07 FΜ 44% 88% P1467/07 FM 43% 83% P1628/07 FM 53% 85% P2149/09 FΜ 33% 83% P2273/09 FM 22% 83% P2315A/09 FM+del? 33% 87% P3207/11 FM 20% 72% P3274/11 FM+del? 34% 86% P3384/11 FM 44% 85% P3946/12 FM 61% 84% P4101/13 FΜ 36% 86% P4112/13 FΜ 55% 86% P4194/14 FΜ 29% 82% FM 33% P4484/16 86% P4644/19 FM 48% 87% FM+del P1033/05 29% 84% P1234/06 FM+del 36% 87% P1337/13 FM+del 40% 86% P1629/07 FM+del 41% 76% P138A/96 PM(?)+FM 54% 83% P400A/99 PM+FM 60% 90% P400B/11 PM+FM 51% 85% P452/99 PM+FM 49% 88% P1101C/05 PM+FM 56% 86% P3199/11 PM+FM 34% 85% P3274B/13 PM+FM 63% 86% P3350/11 PM+FM 39% 85% P3446/11 PM+FM+del? 51% 88%

P4121/13

PM+FM

52%

82%

Tabela 6 – Valores de metilação nas CpGs referentes à sonda cg22417678 e na CpG à jusante na amostra de indivíduos estudada (continuação)

Registro	Genótipo	Oligonucleotídeo A	
		Posição 1 cg22417678	Posição 2
P4210/14	PM+FM+del?	81%	87%
P4210A/14	PM+FM+del?	83%	85%
P420/99	N+FM	50%	83%
P435/99	N+FM	42%	83%
P476/00	N+FM	43%	87%
P537/00	N+FM	39%	88%
P668/02	N+FM	52%	73%
P718/03	N+FM+del?	26%	88%
P928/04	N+FM	51%	87%
P981/05	N+FM	39%	79%
P1259/06	N+FM	41%	88%
P4442/15	N+FM	43%	87%
P237/97	N+PM+FM	48%	87%
P370/99	N+PM+FM	47%	88%

Tabela 6 – Valores de metilação nas CpGs referentes à sonda cg22417678 e na CpG à jusante na amostra de indivíduos estudada (conclusão)

Legenda: FM: alelo de mutação completa; PM: alelo de pré-mutação; N: alelo normal; I: alelo intermediário; PM+FM: mosaicismo de alelos de pré-mutação e mutação completa; N+FM: mosaicismo de alelos normal e mutação completa; N+PM+FM: mosaicismo de alelos normal, pré-mutação e mutação completa; del?: possível deleção; del: deleção confirmada; NF: não funcionou.

Fonte: A autora, 2022.

Como podemos observar na Tabela 6, o *status* de metilação para a segunda CpG na região analisada no intron 2 se mostraram estáveis em todos os grupos analisados, revelando valores entre 79 e 91%. Entretanto, os níveis de metilação obtidos por pirosequenciamento para a primeira CpG, referente à sonda cg22417678, mostraram diferenças significativas entre os grupos. Nessa posição, o grupo de pacientes com mais de 200 repetições CGG apresentou valores inferiores ao grupo controle (p < 0,0001). Além disso, assim como na região da borda de metilação, as amostras com o alelo pré-mutado apresentam padrão semelhante aos controles, como podemos ver na Figura 24. Os grupos de indivíduos com mosaicismo de tamanho incluindo um alelo de FM também parecem seguir o mesmo padrão dos pacientes com o padrão clássico da FM.

Figura 24 – *Box plot* com os dados de metilação referentes à primeira CpG analisada no intron 2 (sonda cg22417678)



Intron 2

Legenda: CT: grupo controle; FM: alelo de mutação completa; PM+FM: mosaicismo de alelos de pré-mutação e mutação completa; N+FM: mosaicismo de alelos normal e mutação completa; N+PM+FM: mosaicismo de alelos normal, pré-mutação e mutação completa; Casos: grupos FM, PM+FM, N+FM e N+PM+FM somados. Entre parênteses, o número amostral. Valores descritos como média ± desvio padrão. Acima das chaves constam os valores de *p*, obtidos através do teste não paramétrico de Kruskal-Wallis (Kruskal & Wallis, 1952), seguida pelo pós-teste de Dunn para múltiplas comparações (Dunn, 1964). Diferença estatística obtida quando *p*<0,05.

Fonte: A autora, 2022.

Utilizando apenas os dados de metilação na primeira posição para os grupos controle e casos, ou seja, os grupos FM, FM+del, PM+FM, N+FM e

N+PM+FM somados, foi realizada a curva ROC (Figura 25). Essa análise nos mostrou um valor de corte entre os grupos em 69,5% de metilação, ponto esse em que os valores de especificidade, sensibilidade e acurácia para esse marcador se encontram em 100%. Além disso, foi observada diferença estatística entre todos os grupos analisados, inclusive entre o grupo de casos somados, quando comparados ao grupo controle (Figura 24).

Figura 25 – Curva ROC com os dados de metilação referentes à primeira CpG analisada no intron 2 (sonda cg22417678), por pirosequenciamento



Curva ROC do Intron 2

Legenda: Nessa análise, foram incluídos apenas os dados de metilação referentes à primeira posição analisada no intron 2 para os grupos controle e o casos somados, incluindo os grupos FM, FM+del, PM+FM, N+FM e N+PM+FM. A curva ROC mostrou um valor de corte em 69,5%, onde a especificidade, a sensibilidade e a acurácia do marcador epigenético atingem 100%.

Fonte: A autora, 2022.

De acordo com o UCSC Genome Browser (em hg19), não há nenhum elemento regulatório ou repetitivo na região, mas esta posição está localizada em uma região no RNAm que se liga a proteínas relacionadas com *splicing* alternativo. Além disso, de acordo com banco de dados GnomAD (https://gnomad. broadinstitute.org/), há uma variante na guanina logo à jusante da citosina analisada. Trata-se de uma SNV de baixa frequência (rs183760491 G>A), sendo identificada com uma *minor allele frequency* (MAF) global de 0,003. Essa variante é mais observada em europeus, com MAF de 0,009, enquanto que em latinos, a MAF é de 0,001.

A metilação de DNA é um dos principais mecanismos epigenéticos para a regulação da expressão gênica (Lowdon *et al.*, 2016). A metilação das CpGs nas regiões promotoras dos genes recebeu mais atenção, pois na maioria das vezes está associada ao silenciamento transcricional, seja diretamente, bloqueando o acesso de fatores de transcrição, ou indiretamente, recrutando outras proteínas com domínio de ligação a grupamentos metil (Moore *et al.*, 2013; Anastasiadi *et al.*, 2018). Ilhas CpGs que geralmente permanecem não metiladas, se sobrepõem aos promotores de genes e estão associadas à regulação da transcrição gênica (Illingworth & Bird, 2009; Straussman *et al.*, 2009). Estudos recentes revelam que a metilação da região à jusante do local do início da transcrição é também altamente informativa da expressão gênica e, além dos promotores, os *enhancers* (ou seja, intensificadores) também se ligam aos fatores de transcrição, interagem com o promotor e exibem hipometilação durante o desenvolvimento (Bogdanović *et al.*, 2016; Anastasiadi *et al.*, 2018).

Além disso, estudos mostraram diferenças nos níveis de metilação entre o primeiro e o restante dos exons e sugerem que os níveis de expressão gênica sejam melhor inversamente correlacionados com a metilação do primeiro exon do que com o promotor (Brenet *et al.*, 2011). Ainda, foi demonstrada uma correlação positiva entre a metilação do corpo gênico e a expressão gênica (Ball *et al.*, 2009). Esses estudos sugerem que a metilação de DNA de elementos distais ou intragênicos com diferentes graus de densidade de CpGs estejam envolvidos na regulação da expressão gênica e que essa metilação desempenhe papéis duplos, seja inibindo ou permitindo a expressão, dependendo da região genômica (Anastasiadi *et al.*, 2018).

Alisch e colaboradores (2013) analisaram nove pacientes com SXF e 53 controles saudáveis, usando Infinium HumanMethylation450 BeadChips. Os autores mostraram uma hipermetilação importante associada à SXF em toda a ilha CpG que abrange a região 5' UTR do gene *FMR1*, sugerindo que algumas dessas

posições CpGs, cujo perfil de metilação é distinto poderiam melhorar os métodos diagnósticos atuais (Alisch *et al.*, 2013). Ainda, Alisch *et al* (2013) revelaram que essa metilação está significativamente diminuída em uma CpG no corpo do gene e é consistente com trabalhos anteriores indicando que a hipermetilação do corpo do gene está associada à expressão gênica ativa (Hellman & Chess, 2007; Alisch *et al.*, 2013). Os dados de Alisch *et al* (2013) foram incluídos na análise de heatmap realizada neste estudo e a posição CpG no corpo do gene relatada pelos autores foi analisada por pirosequenciamento em nosso trabalho.

Ademais, Godler e colaboradores (2010) relataram a ocorrência de duas regiões reconhecidas como marcadores epigenéticos na SXF, nomeadas de Elemento Epigenético Relacionado ao X-Frágil 1 e 2 (FREE1 e FREE2). Essas regiões estão localizadas ao redor da repetição CGG (Godler *et al.*, 2010; Pietrobono *et al.*, 2002), sendo a região FREE1 a cerca de 400 pb à montante das repetições CGG (Godler *et al.*, 2011). Embora a ilha CpG esteja localizada na região 5' não traduzida do gene, a região FREE2 está na porção traduzida que segue o ATG no exon 1, o que qualifica a FREE2 como uma região intragênica entre o exon 1 e o intron 1 do *FMR1* (Godler *et al.*, 2012). A metilação de ambas as regiões FREE foi correlacionada com a da ilha CpG no *FMR1* e negativamente relacionada com a expressão da proteína FMRP em amostras sanguíneas de homens HFMs (Godler *et al.*, 2010). Com isso, Godler e colaboradores (2010) sugerem que o padrão de metilação na região FREE1 poderia ser usado para diferenciar homens e mulheres com SXF e controles.

Além disso, vale destacar que o pequeno número amostral de alguns subgrupos até o momento pode ser o responsável por algumas das observações neste estudo. Ainda assim, são resultados promissores e podem estimular o desenvolvimento de uma nova metodologia de diagnóstico molecular menos custosa e/ou mais informativa do que as atuais. Visto que, muitos ensaios realizados para determinar casos de SXF são baseados em enzimas sensíveis à metilação (como *BssH*II, *Eag*I, *Nru*I ou *Sac*II) que não clivam amostras de DNA de portadores da SXF, pois clivam apenas amostras sem metilação na região de repetições CGG (Rousseau *et al.*, 1991; Hayward *et al.*, 2017). Portanto, essas metodologias podem estar sujeitas a falhas no processo de digestão enzimática. Ao desenhar um ensaio baseado em uma posição que se encontra não metilada em pacientes com SXF, haveria maior segurança nas análises.

3.6 Caso P4210/14

Um casos de gêmeos monozigóticos com SXF em específico nos chamou bastante a atenção. Ambos os meninos apresentaram o perfil mosaico PM+FM na metodologia de TRP-PCR . No entanto, o alelo de mutação completa mostrou um padrão em formato de "onda" no eletroferograma (Figura 26). O caso P4210/14 (gêmeo 1) apresentou uma pré-mutação com 181 CGGs parcialmente metilada (54% de metilação) e uma mutação completa não metilada (0% de metilação), enquanto seu irmão P4210A/14 (gêmeo 2) apresentou uma pré-mutação também com 181 CGGs completamente metilada e uma mutação completa não metilada (1% de metilação).

Além disso, foi coletado e analisado também o DNA da mãe desses meninos (P4210B/14) e pode ser observado que ela apresenta um alelo normal com 29 CGGs e uma pré-mutação com 84 CGGs, ambos parcialmente metilados (56% e 67% de metilação, respectivamente) (Figura 26).

Esse ensaio foi repetido e obtivemos resultados semelhantes. Esses resultados foram discutidos com a assistência científica da empresa revendedora do kit, mas não se chegou a uma conclusão de interpretação biológica. Esse achado ainda não foi relatado na literatura e requer maiores investigações.

Figura 26 – Resultados da amplificação com o kit AmplideX® *FMR1* mPCR de amostras de DNA da família P4210/14, ilustrando o padrão "onda" observado na faixa de mutação completa nos meninos (continua)



Figura 26 – Resultados da amplificação com o kit AmplideX® *FMR1* mPCR de amostras de DNA da família P4210/14, ilustrando o padrão onda observado na faixa de mutação completa nos meninos (continuação)



Figura 26 – Resultados da amplificação com o kit AmplideX® *FMR1* mPCR de amostras de DNA da família P4210/14, ilustrando o padrão onda observado na faixa de mutação completa nos meninos (conclusão)



Legenda: Eletroferogramas referentes aos pacientes gêmeos (A) P4210/14, que apresenta um pico baixo de pré-mutação (181 CGGs) parcialmente metilada (54% de metilação) e um padrão onda na faixa de mutação completa (>200 CGGs) não metilada (0% de metilação) e (B) P4210A/14, portador de um pico baixo de pré-mutação (181 CGGs) completamente metilada (100% de metilação) e um padrão onda na faixa de mutação completa (>200 CGGs) não metilada (1% de metilação). Além disso, podemos observar a presença de um alelo normal (29 CGGs) e uma pré-mutação (84 CGGs) ambos parcialmente metilados (56% e 67% de metilação) em sua mãe (C) P4210B/14. Dig e Dig_g: controle de digestão; Ref e Ref_g: controle de referência; NOR e NOR_g: alelo normal; INT e INT_g: alelo intermediário; PM e PM_g: alelo de pré-mutação; FM e FM_g: alelo de mutação completa. A reação de digestão é amplificada com oligonucleotídeos marcados com corante HEX (verde), enquanto a reação controle com oligonucleotídeos marcados com FAM (azul). Dessa forma, o número de repetições é visualizado em FAM e a porcentagem de metilação é acessada em HEX.

Fonte: A autora, 2022.

С

Além disso, foi analisada a borda de metilação dos gêmeos e os valores de metilação dessa região sofrem uma redução a partir da CG 70, o que sugere a presença da borda como nos controles (Figuras 27 e 28). Ainda, foi analisada a posição de interesse no intron 2 e podemos observar um status de metilação elevado, podendo ser também comparado aos controles (Figura 29). Embora esse achado seja divergente dos demais pacientes acometidos pela SXF, condiz com a ausência de metilação nas repetições, como ocorre nos controles, mesmo na presença do fenótipo da SXF em ambos os gêmeos.

Figura 27 – Gráfico de linhas com os dados de metilação referentes às CpGs 65 a 71, avaliada por pirosequenciamento



Legenda: CG: CpG; FM: alelo de mutação completa; PM: alelo de pré-mutação; N: alelo normal; I: alelo intermediário; PM+FM: mosaicismo de alelos de pré-mutação e mutação completa; N+FM: mosaicismo de alelos normal e mutação completa; N+PM+FM: mosaicismo de alelos normal, pré-mutação e mutação completa; del: deleção confirmada; Entre parênteses, o número amostral. Valores descritos como média ± desvio padrão. Fonte: A autora, 2022.



Figura 28 – Box plots com os dados de metilação referentes às CpGs 65 a 71, avaliada por pirosequenciamento (continua)



Figura 28 – Box plots com os dados de metilação referentes às CpGs 65 a 71, avaliada por pirosequenciamento (continuação)



Figura 28 – Box plots com os dados de metilação referentes às CpGs 65 a 71, avaliada por pirosequenciamento (continuação)



Figura 28 – Box plots com os dados de metilação referentes às CpGs 65 a 71, avaliada por pirosequenciamento (conclusão)

Legenda: Gráficos para cada CpG avaliada, sendo (A) CG 71, (B) CG 70, (C) CG 69, (D) CG 68, (E) CG67, (F) CG66 e (G) CG65. CG: CpG; CT: grupo controle; FM: alelo de mutação completa; PM: alelo de pré-mutação; N: alelo normal; I: alelo intermediário; PM+FM: mosaicismo de alelos de pré-mutação e mutação completa; N+FM: mosaicismo de alelos normal e mutação completa; N+PM+FM: mosaicismo de alelos normal, pré-mutação e mutação completa. Entre parênteses, o número amostral. Valores descritos como média ± desvio padrão. Acima das chaves constam os valores de *p*, obtidos através do teste não paramétrico de Kruskal-Wallis (Kruskal & Wallis, 1952), seguida pelo pós-teste de Dunn para múltiplas comparações (Dunn, 1964). Diferença estatística obtida quando *p*<0,05.

Fonte: A autora, 2022.

Figura 29 – Box Plot com os dados de metilação referentes à primeira CpG analisada no intron 2, por pirosequenciamento, incluindo os casos P4210/14



Legenda: FM: alelo de mutação completa; PM: alelo de pré-mutação; N: alelo normal; I: alelo intermediário; PM+FM: mosaicismo de alelos de pré-mutação e mutação completa; N+FM: mosaicismo de alelos normal e mutação completa; N+PM+FM: mosaicismo de alelos normal, pré-mutação e mutação completa; del?: possível deleção; Entre parênteses, o número amostral. Valores descritos como média ± desvio padrão. Acima das chaves constam os valores de p, obtidos através do teste não paramétrico de Kruskal-Wallis (Kruskal & Wallis, 1952), seguida pelo pós-teste de Dunn para múltiplas comparações (Dunn, 1964). Diferença estatística obtida quando p<0,05.</p>

Fonte: A autora, 2022.

Devido aos padrões atípicos observados em todas as metodologias testadas, as análises de curva ROC apresentadas na Figura 25 não contam com os dados de metilação referentes aos gêmeos P4210/14. Como podemos observar na Figura 30, a inclusão do caso dos gêmeos nas análises gera uma leve redução nos valores de sensibilidade para 97,06% e de acurácia para 98%. Entretanto, o fato deles não apresentarem uma mutação completa hipermetilada,

como é o esperado para os pacientes com SXF, leva a acreditar que o diagnóstico por metilação não seja o indicado para esse caso. Com isso, o caso segue com inconclusão nos resultados das análises moleculares e requer maiores investigações para somar ao que já foi observado até o momento. Esse achado pode representar uma limitação para o diagnóstico molecular por metilação, assim como alerta para a importância na melhor caracterização de casos atípicos.

Figura 30 – Curva ROC com os dados de metilação referentes à primeira CpG analisada no intron 2 (sonda cg22417678), por pirosequenciamento, incluindo os casos P4210/14



Curva ROC: Intron 2

Legenda: Nessa análise, foram incluídos apenas os dados de metilação referentes à primeira posição analisada no intron 2 para os grupos controle e o casos somados, incluindo os grupos FM, FM+del, PM+FM, N+FM e N+PM+FM e os gêmeos P4210/14. A curva ROC mostrou um valor de corte em 69,5%, onde a sensibilidade, a especificidade e a acurácia do marcador epigenético atingem 100%, 97% e 98%, respectivamente. Fonte: A autora, 2022.

Além dos gêmeos P4210/14, apenas mais um paciente apresentou um padrão nesse formato inesperado no eletroferograma da reação de TRP-PCR, utilizando o kit AmplideX® *FMR1* mPCR. O probando P537/00 apresentou uma

mutação completa completamente metilada em formato de "onda" (Figura 31). Entretanto, os resultados das reações de pirosequenciamento da região da borda de metilação e da posição de interesse no intron 2 (sonda cg22417678) referentes à amostra deste indivíduo não foram distintos dos resultados para os demais pacientes com mutação completa hipermetilada (Tabelas 5 e 6). Figura 31 – Resultados da amplificação com o kit AmplideX® *FMR1* mPCR de amostras de DNA do probando P537/00, ilustrando o padrão onda observado na faixa de mutação completa (continua)



Figura 31 – Resultados da amplificação com o kit AmplideX® *FMR1* mPCR de amostras de DNA do probando P537/00, ilustrando o padrão onda observado na faixa de mutação completa (conclusão)

В



Legenda: Eletroferogramas referentes ao paciente P537/00, que apresenta (A) um padrão onda na faixa de mutação completa (>200 CGGs) completamente metilada (87% de metilação), acompanhado (B) de um possível alelo na faixa normal parcialmente metilado (78%). Aqui (B) podemos observar que, embora o fragmento referente à mutação completa tenha amplificado com pico mais baixo, possivelmente devido à competição entre os alelos na amplificação, o padrão onda apareceu no quadro de leitura referente ao corante ROX (verde). Além disso, podemos observar também que o alelo normal pode ter apresentado alta porcentagem de metilação devido ao baixo rendimento da reação. Dig e Dig_g: controle de digestão; Ref e Ref_g: controle de referência; NOR e NOR_g: alelo normal; INT e INT_g: alelo intermediário; PM e PM_g: alelo de pré-mutação; FM e FM_g: alelo de mutação completa. A reação de digestão é amplificada com oligonucleotídeos marcados com corante HEX (verde), enquanto a reação controle com oligonucleotídeos marcados com FAM (azul). Dessa forma, o número de repetições é visualizado em FAM e a porcentagem de metilação é acessada em HEX. Fonte: A autora, 2022.

Embora bastante raro, casos como o dos gêmeos P4210/14 foram relatados na literatura nos últimos anos. Em 2018, Fernández e colaboradores relataram dois casos atípicos, envolvendo mosaicismo de tamanho e de metilação. O caso 1 é um menino de 10 anos com hiperatividade e sinais de déficit de atenção e aprendizado (Fernández et al., 2018). Ele apresentou, através de Southern blotting de DNA extraído de sangue periférico, uma mutação completa não metilada e pré-mutações metilada e não metilada (Fernández et al., 2018). Já a PCR de metilação (mPCR) em DNA de leucócitos mostrou uma pré-mutação de 170 CGGs completamente metilada (100% de metilação), uma pré-mutação de 193 CGGs não metilada (3% de metilação) e uma mutação completa não metilada (8% de metilação) (Fernández et al., 2018). Já o caso FXS 2, com fenótipo claro de SXF, apresentou, também através de PCR de metilação (mPCR) em DNA extraído de amostra oral, uma pré-mutação de 135 CGGs completamente metilada (100% de metilação) e uma mutação completa parcialmente metilada (44% de metilação), embora esse achado não tenha sido observado em amostras de outros tecidos do paciente (Fernández et al., 2018).

Outro caso semelhante foi relatado por Hayward e colaboradores (2019). Os autores identificaram um mosaicismo de tamanho, envolvendo duas prémutações de aproximadamente 165 e 175 CGGs, ambas metiladas, além de uma mutação completa não metilada, em um homem de 20 anos, diagnosticado com SXF através de *Southern blotting* (Hayward *et al.*, 2019). O mosaicismo foi acessado através de PCR de DNA genômico, extraído de amostra de saliva e de sangue, pré-digerido com enzima sensível à metilação (*Hpall*) (Hayward *et al.*, 2019).

O mecanismo que leva à instabilidade na expansão das repetições CGG ainda não está claro (Jiraanont et al., 2017). Entretanto, a existência de perfis atípicos, como os mosaicismos descritos nos últimos anos, sugerem que variações genéticas e epigenéticas no gene FMR1 podem estar altamente relacionadas 0 compromentimento cognitivo distúrbios com em do neurodesenvolvimento (Fernández et al., 2018). Além disso, esses casos ainda destacam a importância da investigação molecular em pacientes com comprometimento cognitivo através de uma combinação de abordagens para perfis metodológicas garantir que esses atípicos não sejam subdiagnosticados e potencialmente perdidos pelos testes de rotina para SXF (Gonçalves *et al.*, 2016; Jiraanont *et al.*, 2017).

Em experimentos com CRISPR-Cas9, todo o trecho da repetição CGG foi deletado em células-tronco, mas a metilação não foi perdida em todas as linhagens celulares e foi sugerido que, em células que se dividem de forma relativamente mais lenta, a metilação é mais provável de ser mantida (Xie et al., 2016; Hayward et al., 2019). Esse efeito pode ser aumentado em células diferenciadas onde a desmetilação de DNA ativo não é tão eficaz e a metilação é então mais provável de ser transmitida para as células adiante (Hayward et al., 2019). Assim, Hayward e colaboradores (2019) propuseram uma interpretação para esse perfil atípico em que o paciente tenha herdado um alelo de FM que só se tornou metilado em um subconjunto de células embrionárias e, então, em algum ponto no início do desenvolvimento, a contração do alelo de FM metilado gerou os dois alelos de PM diferentes. Dessa forma, esses alelos permaneceram metilados e, com isso, não seriam propensos a expansão adicional, enquanto os alelos de FM não metilados mantiveram a propensão a serem instáveis (Hayward et al., 2019). Ainda, pode ser que os alelos de PM metilados contribuam significativamente para o quadro clínico do paciente (Hayward et al., 2019).

Além dos gêmeos P4210/14 serem interessantes casos adicionais à lista de achados atípicos na SXF, eles podem ter importantes implicações diagnósticas e terapêuticas. A estratégia farmacêutica de restauração da atividade do gene *FMR1* tem como alvo modificações potencialmente reversíveis, como a metilação de DNA (Bar-Nur *et al.*, 2012). Entretanto, como esses tratamentos são baseados na reversão do silenciamento da mutação completa hipermetilada, casos de SXF atípicos, como o P4210/14, não estariam sendo atendidos por esses tratamentos e o tratamento voltado para agentes desmetilantes poderia não obter os mesmos efeitos nesses indivíduos. Assim, a determinação dos perfis atípicos.

CONCLUSÕES

a) A partir da análise por TRP-PCR dos 101 indivíduos deste estudo, a FM foi confirmada em 53 deles (52,5%). Em 28 meninos (27,7%), foi identificado mosaicismo de tamanho, sendo 16 pacientes (15,8%) com PM+FM, 9 (8,9%) com N+FM e três portadores (3%) de N+PM+FM. Em um paciente com N+FM foi possível caracterizar uma deleção de 38 pb próxima às repetições CGG, contendo microhomologia de 10 pb nos pontos de quebra;

b) A partir do *heatmap* gerado a partir de dados de bancos de dados públicos envolvendo sondas para CpGs metiladas no gene *FMR1*, observamos níveis de metilação aumentados na maioria das sondas em casos com SXF, se comparados aos controles, corroborando a literatura. Mas, estes dados não confirmaram a existência da borda de metilação relatada por Naumann *et al* (2009, 2014);

- c) A análise por pirosequenciamento da região da borda de metilação mostrou valores reduzidos a partir da CpG 70 nas amostras controles, ratificando a presença da borda de metilação proposta. Já no grupo de casos, os valores de metilação se mantiveram superiores, o que sugere a perda da borda nesses indivíduos;
- d) O heatmap mostrou que a sonda cg22417678, no intron 2, apresentou o padrão inverso das demais. Através da análise por pirosequenciamento dessa CpG, os casos mostraram valores de metilação inferiores aos controles com diferença significativa entre os grupos (*p* < 0,0001). Esse marcador apresentou um valor de corte em 69,5% de metilação, com altos valores de especificidade, sensibilidade e acurácia;</p>
- e) Esses resultados são promissores e podem estimular o desenvolvimento de uma nova metodologia de diagnóstico molecular, complementando as metodologias atualmente disponíveis.

REFERÊNCIAS

ADAMS-CIOABA MA, GUO Y, BIAN CB, *et al.* Structural studies of the tandem Tudor domains of fragile X mental retardation related proteins FXR1 and FXR2. *PLoS ONE* 2010; 5: e13559.

ALISCH RS, BARWICK BG, CHOPRA P, et al. Age-associated DNA methylation in pediatric populations. *Genome Res* 2012; 22(4): 623-632.

ALISCH RS, WANG T, CHOPRA P, *et al.* Genome-wide analysis validates aberrant methylation in fragile X syndrome is specific to the FMR1 locus. *BMC Med Genet* 2013; 14: 18.

ALLEN EG, SHERMAN S, ABRAMOWITZ A, *et al.* Examination of the effect of the polymorphic CGG repeat in the FMR1 gene on cognitive performance. *Behav Genet* 2005; 35(4): 435–445.

ALVAREZ-MORA MI, GUITART M, RODRIGUEZ-REVENGA L, *et al.* Paternal transmission of a FMR1 full mutation allele. *Am J Med Genet* A 2017; 173(10): 2795–2797.

AMERICAN PSYCHIATRIC ASSOCIATION. American Psychiatric Association, 2017. www.psychiatry.org. Acessado em 17/10/2021.

ANASTASIADI D, ESTEVE-CODINA A, PIFERRER F. Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species. *Epigenetics Chromatin* 2018; 11(1): 37.

ANTAR LN, BASSELL GJ. Sunrise at the Synapse: The FMRP mRNP Shaping the Synaptic Interface. *Neuron* 2003; 37(4): 555–558.

ANTAR LN, DICTENBERG JB, PLOCINIAK M, *et al.* Localization of FMRPassociated mRNA granules and requirement of microtubules for activity-dependent trafficking in hippocampal neurons. *Genes Brain Behav* 2005; 4(6): 350–359.

ARAÚJO, FM. Estabelecimento de uma nova ferramenta molecular para investigação de expansões no gene *FMR1*. 2018. 63 f. Monografia (Graduação em Ciências Biológicas) – Instituto de Biologia Roberto Alcantara Gomes, Universidade do Estado do Rio de Janeiro, Rio de janeiro, 2018.

ARMATAS V. Mental retardation: definitions, etiology, epidemiology and diagnosis. *J Sport Health Res* 2009; 1: 112-122.

AROCENA DG, DE DIEGO Y, OOSTRA BA, *et al.* A fragile X case with an amplification/deletion mosaic pattern. *Hum Genet* 2000; 106(3): 366–369.

BAGNI C, GREENOUGH WT. From mRNP trafficking to spine dysmorphogenesis: the roots of fragile X syndrome. *Nat Rev Neurosci* 2005; 6(5): 376–387.

BAGNI C, TASSONE F, NERI G, *et al.* Science in medicine Fragile X syndrome: causes, diagnosis, mechanisms, and therapeutics. *J Clin Invest* 2012; 122(12): 4314-4322.

BAILEY DB, RASPA M, HOLIDAY D, *et al.* The functional skills of individuals with fragile X syndrome: A lifespan, cross-sectional analysis. *Am J Intellec Dev Disabil* 2009; 114: 289-303.

BALL MP, LI JB, GAO Y, *et al.* Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nat Biotechnol* 2009; 27: 361–368.

BAR-NUR O, CASPI I, BENVENISTY N, *et al.* Molecular analysis of FMR1 reactivation in fragile-X induced pluripotent stem cells and their neuronal derivatives. *J Mol Biol* 2012; 4: 180-183.

BASSANI S, ZAPATA J, GEROSA L, *et al.* The neurobiology of X-linked intellectual disability. *Neuroscientist* 2013; 19(5): 541-552.

BERRY-KRAVIS E, RASPA M, LOGGIN-HESTER L, *et al.* Seizures in fragile X syndrome: characteristics and comorbid diagnoses. *Am J Intellect Dev Disabil* 2010; 115: 461–472.

BHOGAL B, JEPSON JE, SAVVA YA, *et al.* Modulation of dADAR-dependent RNA editing by the Drosophila fragile X mental retardation protein. *Nat Neurosci* 2011; 14(12): 1517–1524.

BINDER G, RAPPOLD GA. SHOX Deficiency Disorders. 2005 [Updated 2018 Jun 28]. In: ADAM MP, ARDINGER HH, PAGON RA, *et al.* GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2022.

BLACKWELL E, ZHANG X, CEMAN S. Arginines of the RGG box regulate FMRP association with polyribosomes and mRNA. *Hum Mol Genet* 2010; 19: 1314-1323.

BOGDANOVIĆ O, SMITS AH, DE LA CALLE MUSTIENES E, *et al.* Active DNA demethylation at enhancers during the vertebrate phylotypic period. *Nat Genet* 2016; 48: 417–426.

BODEGA B, BIONE S, DALPRA L, *et al.* Influence of intermediate and uninterrupted FMR1 CGG expansions in premature ovarian failure manifestation. *Hum Repro* 2006; 21(4): 952–957.

BONARRIGO FA, RUSSO S, VIZZIELLO P, *et al.* Think about it: FMR1 gene mosaicism. *J Child Neurol* 2014; 29(9): NP74–NP77.

BRENET F, MOH M, FUNK P, *et al.* DNA methylation of the first exon is tightly linked to transcriptional silencing. *PLoS ONE* 2011; 6: e14524.

BRIGHTWELL G, WYCHERLEY R, POTTS G, *et al.* A highdensity SNP map for the FRAX region of the X chromosome. *J Hum Genet* 2002; 47: 567–575.

BROWN V, JIN P, CEMAN S, *et al.* Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome. *Cell* 2001; 107(4): 477–487.

CASTEL AL, NAKAMORI M, TOMÈ S, *et al.* Expanded CTG repeat demarcates a boundary for abnormal CpG methylation in myotonic dystrophy patient tissues. *Hum Mol Genet* 2010; 20: 1–15.

CAUDY AA, MYERS M, HANNON GJ, *et al.* Fragile X-related protein and VIG associate with the RNA interference machinery. *Genes Dev* 2002; 16(19): 2491–2496.

CENTERS FOR DISEASE CONTROL AND PREVENTION. Facts About Intellectual Disability, 2019. www.cdc.gov. Acessado em 30/07/2020.

CHELLY J, KHELFAOUI M, FRANCIS F, *et al.* Genetics and pathophysiology of mental retardation. *Eur J Hum Genet* 2006: 14: 701-713.

CHEN L, YUN SW, SETO J, *et al.* The fragile X mental retardation protein binds and regulates a novel class of mRNAs containing U rich target sequences. *Neuroscience* 2003; 120(4): 1005–1017.

CHIURAZZI P, SCHWARTZ CE, GECZ J, et al. XLMR genes: update 2007. Eur J Hum Genet 2008; 16: 422-434.

CHIURAZZI P, KOZAK L, NERI G. Unstable triplets and their mutational mechanism: size reduction of the CGG repeat vs. germline mosaicism in the fragile X syndrome. *Am J Med Genet* 1994; 51(4): 517–521.

CHIURAZZI P, PIROZZI F. Advances in understanding – genetic basis of intellectual disability. *F1000Research* 2016; 5(0): 599.

CHIURAZZI P, POMPONI MG, WILLEMSEN R, *et al.* In vitro reactivation of the FMR1 gene involved in fragile X syndrome. *Hum Mol Genet* 1998; 7: 109-113.

CIACCIO C, FONTANA L, MILANI D, *et al.* Fragile X syndrome: a review of clinical and molecular diagnoses. *Ital J Pediatr* 2017; 43(1): 39.

CLEARY JD, PEARSON CE. Replication fork dynamics and dynamic mutations: the fork-shift model of repeat instability. *Trends Genet* 2005; 21: 272-280.

COFFEE B, ZHANG F, CEMAN S, *et al.* Histone modifications depict an aberrantly heterochromatinized FMR1 gene in fragile x syndrome. *Am J Hum Genet* 2002; 71: 923–932.

COFFEE B, IKEDA M, BUDIMIROVIC DB, *et al.* Mosaic FMR1 deletion causes fragile X syndrome and can lead to molecular misdiagnosis. *Am J Med Genet* A 2008; 146A(10): 1358-1367.

COFFEE B, ZHANG F, WARREN ST, *et al.* Acetylated histones are associated with FMR1 in normal but not fragile X syndrome cells. *Nat Genet* 1999; 22: 98-101.

COLAK D, ZANINOVIC N, COHEN MS, *et al.* Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome. *Science* 2014; 343: 1002–1005.

CORDEIRO L, BALLINGER E, HAGERMAN R, *et al.* Clinical assessment of DSM-IV anxiety disorders in fragile X syndrome: prevalence and characterization. *J Neurodev Disord* 2011; 3: 57–67.

CRAWFORD DC, ACUÑA JM, SHERMAN SL. FMR1 and the fragile X syndrome: human genome epidemiology review. *Genet Med* 2001; 3: 359-371.

DARNELL JC, VAN DRIESCHE SJ, ZHANG C, *et al.* FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* 2011; 146(2): 247–261.

DATTA S, ALAM MP, MAJUMDAR SS, *et al.* Nucleosomal occupancy and CGG repeat expansion: a comparative analysis of triplet repeat region from mouse and human fragile X mental retardation gene 1. *Chromosome Res* 2011; 19: 445–455.

DE GRAAFF E, DE VRIES BB, WILLEMSEN R, *et al.* The fragile X phenotype in a mosaic male with a deletion showing expression of the FMR1 protein in 28% of the cells. *Am J Med Genet* 1996; 64(2): 302–308.

DE RUBEIS S, FERNANDEZ E, BUZZI A, *et al.* Molecular and Cellular Aspects of Mental Retardation in the Fragile X Syndrome: From Gene Mutation/s to Spine Dysmorphogenesis. *Adv Exp Med Biol* 2012; 970: 517–551.

DI RUSCIO A, EBRALIDZE AK, BENOUKRAF T, *et al.* DNMT1-interacting RNAs block gene-specific DNA methylation. *Nature* 2013; 503: 371–376.

DJEBALI S, DAVIS CA, MERKEL A, *et al.* Landscape of transcription in human cells. *Nature* 2012; 489: 101–108.

DUNN OJ. Multiple Comparisons Using Rank Sums. *Technometrics* 1964; 6(3): 241–252.

EDAMURA KN, PEARSON CE. DNA methylation and replication: implications for the "deletion hotspot" region of FMR1. *Hum Genet* 2005; 118(2): 301–304.

EICHLER EE, RICHARDS S, GIBBS RA, *et al.* Fine structure of the human FMR1 gene. *Hum Molec Genet* 1993; 2: 1147-1153. Nota: Erratum: *Hum Molec Genet* 1994; 3: 684-685.

EICHLER EE, HOLDEN JA, POPOVICH BW, *et al.* Length of uninterrupted CGG repeats determines instability in the FMR1 gene. *Nat Genet* 1994; 8: 88–94.

EPSZTEJN-LITMAN S, EIGES R. Monitoring for Epigenetic Modifications at the FMR1 Locus. *Methods Mol Biol* 2019; 1942: 29-48.

ERBS E, FENGER-GRON J, JACOBSEN CM, *et al.* Spontaneous rescue of a FMR1 repeat expansion and review of deletions in the FMR1 non-coding region. *Eur J Med Genet* 2021; 64: 104244.

FAN H, BOOKER JK, MCCANDLESS SE, *et al.* Mosaicism for an FMR1 gene deletion in a fragile X female. *Am J Med Genet* A 2005; 136: 214–217.

FERNÁNDEZ E, GENNARO E, PIROZZI F, *et al.* FXS-Like Phenotype in Two Unrelated Patients Carrying a Methylated Premutation of the *FMR1* Gene. *Front Genet* 2018; 9: 442.

FERNANDEZ-CARVAJAL I, LOPEZ POSADAS B, PAN R, *et al.* Expansion of an FMR1 grey-zone allele to a full mutation in two generations. *J Mol Diagn* 2009; 11(4): 306-310.

FERREIRA SI, PIRES LM, FERRAO J, *et al.* Mosaicism for FMR1 gene full mutation and intermediate allele in a female foetus: A postzygotic retraction event. *Gene* 2013; 527: 421–425.

FLORIANI MA, VILAS BOAS MR, ROSA RFM, *et al.* Report of a patient with fragile X syndrome unexpectedly identified by karyotype analysis. *J Bras Patol Med Lab* 2017; 53(2): 108-109.

FU YH, KUHL DP, PIZZUTI A, *et al.* Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* 1991; 67(6): 1047-1058.

GÉCZ J, SHOUBRIDGE C, CORBETT M. The genetic landscape of intellectual disability arising from chromosome X. *Trends Genet* 2009; 25: 308-316.

GENC B, MULLER-HARTMANN H, ZESCHNIGK M, *et al.* Methylation mosaicism of 5'-(CGG)(n)-3' repeats in fragile X, premutation and normal individuals. *Nucleic Acids Res* 2000; 28(10): 2141–2152.

GODLER DE, SLATER HR, BUI QM, *et al.* FMR1 intron 1 methylation predicts FMRP expression in blood of female carriers of expanded FMR1 alleles. *J Mol Diagn* 2011;13: 528–536.

GODLER DE, SLATER HR, BUI QM, *et al.* Fragile X mental retardation 1 (FMR1) intron 1 methylation in blood predicts verbal cognitive impairment in female carriers of expanded FMR1 alleles: evidence from a pilot study. *Clin Chem* 2012; 58(3): 590-598.

GODLER DE, TASSONE F, LOESCH DZ, *et al.* Methylation of novel markers of fragile X alleles is inversely correlated with FMRP expression and FMR1 activation ratio. *Hum Mol Genet* 2010; 19: 1618–1632.

GONÇALVES TF, DOS SANTOS JM, GONÇALVES AP, *et al.* Finding FMR1 mosaicism in Fragile X syndrome. *Expert Rev Mol Diagn* 2016; 16: 501-507.

GRASSO M, FARAVELLI F, LO NIGRO C, *et al.* Mosaicism for the full mutation and a microdeletion involving the CGG repeat and flanking sequences in the FMR1 gene in eight fragile X patients. *Am J Med Genet* 1999; 85: 311–316.

GREENWOOD GENETIC CENTER. XLID Genetic Research, 2021. https://www.ggc.org/xlid-genetic-research. Acessado em 11/10/2021.

GRØNSKOV K, HJALGRIM H, BJERAGER MO, *et al.* Deletion of all CGG repeats plus flanking sequences in FMR1 does not abolish gene expression. *Am J Hum Genet* 1997; 61(4): 961–967.

GROSS AM, JAEGER PA, KREISBERG JF, *et al.* Methylome-wide Analysis of Chronic HIV Infection Reveals Five-Year Increase in Biological Age and Epigenetic Targeting of HLA. *Mol Cell* 2016;62(2):157-168.

GUNTER C, PARADEE W, CRAWFORD DC, *et al.* Re-examination of factors associated with expansion of CGG repeats using a single nucleotide polymorphism in FMR1. *Hum Mol Genet* 1998; 7(12): 1935–1946.

HADDAD LA, MINGRONI-NETTO RC, VIANNA-MORGANTE MA, *et al.* A PCRbased test suitable for screening for fragile X syndrome among mentally retarded males. *Hum Genet* 1996; 197: 808-812.

HAGERMAN RJ, HAGERMAN PJ. Fragile X Syndrome: Diagnosis, Treatment and Research. Baltimore: The Johns Hopkins Univ Press 2002; 3–109.

HAGERMAN RJ, BERRY-KRAVIS E, HAZLETT H, *et al.* Fragile X syndrome. *Nat Rev Dis Primers* 2017; 3: 17065.

HAGERMAN PJ, BERRY-KRAVIS E, KAUFMANN WE, *et al.* Advances in the treatment of fragile X syndrome. *Pediatr* 2009; 123: 378-390.

HAGERMAN PJ, HAGERMAN RJ. Fragile X-associated tremor/ataxia syndrome (FXTAS). *Ment Retard Dev Disabil Res Rev* 2004; 10(1): 25-30.

HAGERMAN RJ, HULL CE, SAFANDA JF, *et al.* High functioning fragile X males: demonstration of an unmethylated fully expanded FMR-1 mutation associated with protein expression. *Am J Med Genet* 1994; 51(4): 298-308.

HAGERMAN RJ, LEEHEY M, HEINRICHS W, *et al.* Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X. *Neurology* 2001; 57(1): 127–130.

HAGERMAN RJ, PROTIC D, RAJARATNAM A, *et al.* Fragile X-Associated Neuropsychiatric Disorders (FXAND). *Front Psychiatry* 2018; 9: 564.

HALL DA, BERRY-KRAVIS E. Fragile X syndrome and fragile X-associated tremor ataxia syndrome. *Handb Clin Neurol* 2018; 147: 377-391.

HALL DA, NAG S, OUYANG B, *et al.* Fragile X Gray Zone Alleles Are Associated With Signs of Parkinsonism and Earlier Death. *Mov Disord* 2020; 35(8): 1448-1456.

HALL D, TASSONE F, KLEPITSKAYA O, *et al.* Fragile X associated tremor ataxia syndrome in FMR1 gray zone allele carriers. *Mov Disord* 2012; 27(2): 296–300.

HAN XD, POWELL B, PHALIN JL, *et al.* Mosaicism for a full mutation, premutation, and deletion of the CGG repeats results in 22% FMRP and elevated FMR1 mRNA levels in a high-functioning fragile X male. *Am J Med Genet* 2006; 140(13): 1463–1471.

HANDT M, EPPLEN A, HOFFJAN S, *et al.* Point mutation frequency in the FMR1 gene as revealed by fragile X syndrome screening. *Mol Cell Probes* 2014; 28: 279–283.

HAYWARD BE, KUMARI D, USDIN K. Recent advances in assays for the Fragile X-related disorders. *Hum Genet* 2017; 136(10): 1313–1327.

HAYWARD B, LOUTAEV I, DING X, *et al.* Fragile X syndrome in a male with methylated premutation alleles and no detectable methylated full mutation alleles. *Am J Med Genet* A 2019;179(10): 2132-2137.

HAYWARD BE, USDIN K. Mechanisms of Genome Instability in the Fragile X-Related Disorders. *Genes* (Basel) 2021; 12(10): 1633.

HEALY A, RUSH R, OCAIN T. Fragile X syndrome: An update on developing treatment modalities. *ACS Chemical Neuroscience* 2011; 2(8): 402-410.

HELLMAN A, CHESS A. Gene body-specific methylation on the active X chromosome. *Science* 2007; 315(5815): 1141–1143.

HERMANN A, GOYAL R, JELTSCH A. The Dnmt1 DNA-(cytosine-C5)methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J Biol Chem* 2004; 279: 48350–48359.

HIRST MC, GREWAL PK, DAVIES KE. Precursor arrays for triplet repeat expansion at the fragile X locus. *Hum Mol Genet* 1994; 3(9): 1553–1560.

HOCHSTEIN N, MUIZNIEKS I, MANGEL L, *et al.* Epigenetic status of an adenovirus type 12 transgenome upon long-term cultivation in hamster cells. *J Virol* 2007; 81: 5349–5361.

HOGAN AL, CARAVELLA KE, EZELL J, *et al.* Autism spectrum disorder symptoms in infants with fragile X syndrome: a prospective case series. *J Autism Dev Disord* 2017; 47: 1628–1644.

HUANG J, ZHU T, QU Y, *et al.* Prenatal, Perinatal and Neonatal Risk Factors for Intellectual Disability: A Systemic Review and Meta-Analysis. *PLoS One* 2016; 11(4): e0153655.

HUNTER JE, BERRY-KRAVIS E, HIPP H, *et al.* FMR1 Disorders. *GeneReviews*® [Internet]. Seattle (WA): University of Washington, Seattle; 1993–2021 [updated 2019 Nov 21].

HUNTER J, RIVERO-ARIAS O, ANGELOV A, *et al.* Epidemiology of fragile X syndrome: a systematic review and meta-analysis. *Am J Med Genet* A 2014; 164A(7): 1648–1658.

ILLINGWORTH RS, BIRD AP. CpG islands— 'a rough guide'. *FEBS Lett* 2009; 583: 1713–1720.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). Pesquisa Nacional de Saúde 2019. https://agenciadenoticias.ibge.gov.br/agencia-sala-deimprensa/2013-agencia-de-noticias/releases/31445-pns-2019-pais-tem-17-3milhoes-de-pessoas-com-algum-tipo-de-deficiencia. Acessado em 17/10/2021

ISHIZUKA A, SIOMI MC, SIOMI H. A Drosophila fragile X protein interacts with components of RNAi and ribosomal proteins. *Genes Dev* 2002; 16(19): 2497–2508.

JIRAANONT P, KUMAR M, TANG HT, *et al.* Size and methylation mosaicism in males with Fragile X syndrome. *Expert Rev Mol Diagn* 2017; 17(11): 1023–1032.

KAUFMAN L, AYUB M, VINCENT JB. The genetic basis of non-syndromic intellectual disability: a review. *J Neurodev Disord* 2010; 2 (4): 182-209.

KENNESON A, ZHANG F, HAGEDORN CH, *et al.* Reduced FMRP and increased FMR1 transcription is proportionally associated with CGG repeat number in intermediate-length and premutation carriers. *Hum Mol Genet* 2001; 10(14): 1449–1454.

KHALIL AM, FAGHIHI MA, MODARRESI F, *et al.* A novel RNA transcript with antiapoptotic function is silenced in fragile X syndrome. *PLoS ONE* 2008; 3: e1486.

KRAAN CM, GODLER DE, AMOR DJ. Epigenetics of fragile X syndrome and fragile X-related disorders. *Dev Med Child Neurol* 2019; 61(2): 121-127.

KRUSKAL WH, WALLIS WA. Use of Ranks in One-Criterion Variance Analysis. *J Am Stat Assoc* 1952; 47(260): 583–621.

KUMARI D, LOKANGA R, YUDKIN D, *et al.* Chromatin changes in the development of the fragile X-associated disorders and Friedreich ataxia. *Biochim Biophys Acta* 2012; 1819: 802–810.

KUMARI D, USDIN K. The distribution of repressive histone modifications on silenced FMR1 alleles provides clues to the mechanism of gene silencing in fragile X syndrome. *Hum Mol Genet* 2010; 19(23): 4634–4642.

KUNST CB, WARREN ST. Cryptic and polar variation of the fragile X repeat could result in predisposing normal alleles. *Cell* 1994; 77(6): 853–861.

KUNST CB, ZERYLNICK C, KARICKHOFF L, *et al.* FMR1 in global populations. *Am J Hum Genet* 1996; 58(3): 513–522.

KVARNUNG M, NORDGREN A. Intellectual Disability & Rare Disorders: A Diagnostic Challenge. In: Posada de la Paz M, Taruscio D, Groft S (eds) Rare Diseases Epidemiology: Update and Overview. *Adv Exp Med Biol* 2017; 1031:39-54.

LADD PD, SMITH LE, RABAIA NA, *et al.* An antisense transcript spanning the CGG individuals. *Hum Mol Genet* 2007; 16: 3174–3187.

LANNI S, GORACCI M, BORRELLI L, *et al.* Role of CTCF Protein in Regulating FMR1 Locus Transcription. *PLoS Genet* 2013; 9(7): e1003601.

LATHAM GJ, COPPINGER J, HADD AG, *et al.* The role of AGG interruptions in fragile X repeat expansions: A twenty-year perspective. *Front Genet* 2014; 5: 244.

LEONARD H, WEN X. The epidemiology of mental retardation: challenges and opportunities in the new millennium. *Ment Retard Dev Disabil Res Rev* 2002; 8: 117-134.

LIU Y, WINARNI T, ZHANG L, *et al.* Fragile X-associated tremor/ataxia syndrome (FXTAS) in grey zone carriers. *Clin Genet* 2013; 84(1): 74–77.

LIU XS, WU H, KRZISCH M, *et al.* Rescue of Fragile X Syndrome Neurons by DNA Methylation Editing of the FMR1 Gene. *Cell* 2018;172(5): 979-992.

LOWDON RF, JANG HS, WANG T. Evolution of epigenetic regulation in vertebrate genomes. *Trends Genet TIG* 2016; 32: 269–283.

LUBS HA. A marker X chromosome. Am J Hum Genet 1969; 21(3): 231-244.

LUBS HA, STEVENSON RE, SCHWARTZ CE. Fragile X and X-linked intellectual disability: four decades of discovery. *Am J Hum Genet* 2012; 90: 579-590.

LUDWIG AL, ESPINAL GM, PRETTO DI, *et al.* CNS expression of murine fragile X protein (FMRP) as a function of CGG-repeat size. *Hum Mol Genet* 2014; 23(12): 3228–3238.

LUO S, HUANG W, CHEN C, *et al.* An novel deletion to normal size in the sperm of a fragile X full mutation male. *Clin Genet* 2014; 86: 295–297.

LYON MF. Gene action in the X-chromosome of the mouse (*Mus musculus L.*). *Nature* 1961; 190: 372–373.

MAIA N, LOUREIRO JR, OLIVEIRA B, *et al.* Contraction of fully expanded FMR1 alleles to the normal range: predisposing haplotype or rare events? *J Hum Genet* 2017; 62: 269-275.

MALTER HE, IBER JC, WILLEMSEN R, *et al.* Characterization of the full fragile X syndrome mutation in fetal gametes. *Nat Genet* 1997; 15: 165-169.

MAN L, LEKOVICH J, ROSENWAKS Z, *et al.* Fragile X-Associated Diminished Ovarian Reserve and Primary Ovarian Insufficiency from Molecular Mechanisms to Clinical Manifestations. *Front Mol Neurosci* 2017; 10: 290.

MANDEL JL, CHELLY J. Monogenic X-linked mental retardation: is it as frequent as currently estimated? The paradox of the ARX (Aristaless X) mutations. *Eur J Hum Genet* 2004; 12(9): 689-693.

MANOR E, JABAREEN A, MAGAL N, *et al.* Prenatal Diagnosis of Fragile X: Can a Full Mutation Allele in the FMR1 Gene Contract to a Normal Size? *Front Genet* 2017; 8: 158.

MARRUS N, HALL L. Intellectual Disability and Language Disorder. *Child Adolesc Psychiatr Clin N Am* 2017; 26(3): 539–554.

MARTIN JP, BELL J. A pedigree of mental defect showing sexlinkage. *J Neurol Psychiatry* 1943; 6(3-4): 154–157.

MAURER-STROH S, DICKENS NJ, HUGHES-DAVIES L, *et al.* The Tudor domain 'royal family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. *Trends Biochem Sci* 2003; 28: 69-74.

MERENSTEIN SA, SOBESKY WE, TAYLOR AK, *et al.* Molecular-clinical correlations in males with an expanded FMR1 mutation. *Am J Med Genet* 1996; 64: 388-394.

MILLAR DS, PAUL CL, MOLLOY PL, et al. A distinct sequence (ATAAA)n separates methylated and unmethylated domains at the 50-end of the GSTP1 CpG island. J Biol Chem 2000; 275: 24893–24899.

MILLER SA, DYKES DD, POLESKY HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988; 16(3): 1215.

MONAGHAN KG, LYON E, SPECTOR EB, *et al.* ACMG standards and guidelines for fragile X testing: a revision to the disease-specific supplements to the Standards and Guidelines for Clinical Genetics Laboratories of the American College of Medical Genetics and Genomics. *Genet Med* 2013; 15: 575–586.
MOORE LD, LE T, FAN G. DNA methylation and its basic function. *Neuropsychopharmacology* 2013; 38: 23–38.

NATIONAL FRAGILE X SYNDROME. Fragile X 101, 2021. https://fragilex.org/. Acessado em 31/10/2021.

NAUMANN A, HOCHSTEIN N, WEBER S, *et al.* A distinct DNA methylation boundary in the 5'-upstream sequence of the FMR1 promoter binds nuclear proteins and is lost in fragile X syndrome. *Am J Hum Genet* 2009; 85: 606–616.

NAUMANN A, KRAUS C, HOOGEVEEN A, *et al.* Stable DNA Methylation Boundaries and Expanded Trinucleotide Repeats: Role of DNA Insertions. *J Mol Biol* 2014; 426: 2554–2566.

NERI G, SCHWARTZ CE, LUBS HA, *et al.* X-linked intellectual disability update 2017. *Am J Med Genet* A 2018; 176 (6): 1375-1388.

NOBILE V, PUCCI C, CHIURAZZI P, *et al.* DNA Methylation, Mechanisms of FMR1 Inactivation and Therapeutic Perspectives for Fragile X Syndrome. *Biomolecules* 2021, 11(2): 296.

NOLIN SL, BROWN WT, GLICKSMAN A, *et al.* Expansion of the fragile X CGG repeat in females with premutation or intermediate alleles. *Am J Hum Genet* 2003; 72 (2): 454-464.

NOLIN SL, GLICKSMAN A, ERSALESI N, *et al.* Fragile X full mutation expansions are inhibited by one or more AGG interruptions in premutation carriers. *Genet Med* 2015; 17(5): 358–364.

NOLIN SL, GLICKSMAN A, HOUCK GE Jr, *et al.* Mosaicism in fragile X affected males. *Am J Med Genet* 1994; 51(4): 509–512.

NOLIN SL, GLICKSMAN A, TORTORA N, *et al.* Expansions and contractions of the FMR1 CGG repeat in 5,508 transmissions of normal, intermediate, and premutation alleles. *Am J Med Genet* A 2019; 179(7):1148-1156.

NOLIN SL, SHA S, GLICKSMAN A, *et al.* Fragile X AGG analysis provides new risk predictions for 45-69 repeat alleles. *Am J Med Genet* A 2013; 161: 771–788.

NOTO V, HARRITY C, WALSH D, *et al.* The impact of FMR1 gene mutations on human reproduction and development: a systematic review. *J Assist Reprod Genet* 2016; 33(9): 1135-1147.

OBERLÉ I, ROUSSEAU F, HEITZ D, *et al.* Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science* 1991; 252: 1097-102.

ORRICO A, GALLI L, DOTTI MT, *et al.* Mosaicism for full mutation and normalsized allele of the FMR1 gene: a new case. *Am J Med Genet* 1998; 78(4): 341– 344.

PANDELACHE A, BAKER EK, ALIAGA SM, *et al.* Clinical and Molecular Differences between 4-Year-Old Monozygous Male Twins Mosaic for Normal, Premutation and Fragile X Full Mutation Alleles. *Genes (Basel)* 2019; 10(4): 279.

PANDEY RR, MONDAL T, MOHAMMAD F, *et al.* Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol Cell* 2008; 32: 232–346.

PARK CY, HALEVY T, LEE DR *et al.* Reversion of FMR1 methylation and silencing by editing the triplet repeats in fragile X iPSC-derived neurons. *Cell Rep* 2015; 13: 234–241.

PASTORI C, PESCHANSKY VJ, BARBOUTH D, *et al.* Comprehensive analysis of the transcriptional landscape of the human FMR1 gene reveals two new long noncoding RNAs differentially expressed in Fragile X syndrome and Fragile X associated tremor/ataxia syndrome. *Hum Genet* 2014; 133(1): 59-67.

PATEL DR, GREYDANUS DE, CALLES JL Jr, *et al.* Developmental disabilities across the lifespan. *Dis Mon* 2010; 56: 304-397.

PEPRAH E. Fragile X syndrome: the FMR1 CGG repeat distribution among world populations. *Ann Hum Genet* 2012; 76(2): 178–191.

PEPRAH E, HE W, ALLEN E, *et al.* Examination of FMR1 transcript and protein levels among 74 premutation carriers. *J Hum Genet* 2010; 55(1): 66–68.

PHILOFSKY A, HEPBURN SL, HAYES A, *et al.* Linguistic and cognitive functioning and autism symptoms in young children with Fragile X syndrome. *Am J Ment Retard* 2004; 109: 208-218.

PIETROBONO R, TABOLACCI E, ZALFA F, *et al.* Molecular dissection of the events leading to inactivation of the FMR1 gene. *Hum Mol Genet* 2005; 14: 267–277

PIETROBONO R, POMPONI MG, TABOLACCI E, *et al.* Quantitative analysis of DNA demethylation and transcriptional reactivation of the FMR1 gene in fragile X cells treated with 5-azadeoxycytidine. *Nucleic Acids Res* 2002; 30: 3278–3285.

PRAWER Y, HUNTER M, CRONIN S, *et al.* Prenatal Diagnosis of Fragile X Syndrome in a Twin Pregnancy Complicated by a Complete Retraction. *Genes* 2018; 9: 287.

PRETTO D, YRIGOLLEN CM, TANG HT, et al. Clinical and molecular implications of mosaicism in FMR1 full mutations. *Front Genet* 2014; 5: 318.

PRETTO DI, HUNSAKER MR, CUNNINGHAM CL, *et al.* Intranuclear inclusions in a fragile X mosaic male. *Transl Neurodegener* 2013; 2: 10.

PRIMERANO B, TASSONE F, HAGERMAN RJ, *et al.* Reduced FMR1 mRNA translation efficiency in fragile X patients with premutations. *RNA (New York, NY)* 2002; 8(12): 1482–1488.

RASKE C, HAGERMAN PJ. Molecular pathogenesis of fragile X-associated tremor/ataxia syndrome. *J Investig Med* 2009; 57(8): 825–829.

REYNIERS E, VITS L, DE BOULLE K, *et al.* The full mutation in the FMR-1 gene of male fragile X patients is absent in their sperm. *Nat Genet* 1993; 4: 143–146.

RICCERI L, CATANIA MV, BARDONI B. Editorial: neural and behavioral biology of intellectual disability (ID). *Neurosci Biobehav Rev* 2014; 46 Pt 2:159-160.

RINN JL, KERTESZ M, WANG JK, *et al.* Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 2007; 129: 1311–1323.

ROPERS HH. X-linked mental retardation: many genes for a complex disorder. *Curr Opin Genet Dev* 2006; 16: 260-269.

ROUSSEAU F, HEITZ D, BIANCALANA V, *et al.* Direct diagnosis by DNA analysis of the fragile X syndrome of mental retardation. *N Engl J Med* 1991; 325: 1673–1681.

SALDARRIAGA W, TASSONE F, GONZÁLEZ-TESHIMA LY, *et al.* Fragile X Syndrome. *Colomb Med* 2014; 45(4): 190-198.

SANTORO MR, BRAY SM, WARREN ST. Molecular Mechanisms of Fragile X Syndrome: A Twenty-Year Perspective. *Annu Rev Pathol Mech* 2012; 7: 219-245.

SCHMUCKER B, SEIDEL J. Mosaicism for a full mutation and a normal size allele in two fragile X males. *Am J Med Genet* 1999; 84(3): 221–225.

SCHWEMMLE S, DE GRAAF E, DEISSLER H, *et al.* Characterization of FMR1 promoter elements by in vivo-footprinting analysis. *Am J Hum Genet* 1997; 60: 1354–1362.

SCHNEIDER A, SERITAN A, TASSONE F, *et al.* Psychiatric features in highfunctioning adult brothers with fragile X spectrum disorders. *Prim Care Companion CNS Disord* 2013; 15: 2.

SHERMAN SL. Premature ovarian failure in the fragile X syndrome. *Am J Med Genet* 2000; 97(3): 189-194.

SITZMANN AF, HAGELSTROM RT, TASSONE F, *et al.* Rare FMR1 gene mutations causing fragile X syndrome: A review. *Am J Med Genet* A 2018; 176(1): 11-18.

SNOW K, TESTER DJ, KRUCKEBERG KE, *et al.* Sequence analysis of the fragile X trinucleotide repeat: Implications for the origin of the fragile X mutation. *Hum Mol Genet* 1994; 3(9): 1543–1551.

SPECTOR E, BEHLMANN A, KRONQUIST K, *et al.* Laboratory testing for fragile X, 2021 revision: a technical standard of the American College of Medical Genetics and Genomics (ACMG). *Genet Med* 2021; 23(5): 799–812.

STEVENSON RE, SCHWARTZ CE. X-linked intellectual disability: unique vulnerability of the male genome. *Dev Disabil Res Rev* 2009; 15(4): 361-368.

STEVENSON RE, SCHWARTZ CE, ROGERS RC. Atlas of X-Linked Intellectual Disability Syndromes. 2nd ed. *Estados Unidos da América: Oxford University Press*, 2012.

STRAUSSMAN R, NEJMAN D, ROBERTS D, *et al.* Developmental programming of CpG island methylation profiles in the human genome. *Nat Struct Mol Biol* 2009; 16: 564–571.

SULLIVAN K, HATTON D, HAMMER J, *et al.* ADHD symptoms in children with FXS. *Am J Med Genet* A 2006; 140A: 2275-2288.

TABOLACCI E, MANCANO G, LANNI S, *et al.* Genome-wide methylation analysis demonstrates that 5-aza-2-deoxycytidine treatment does not cause random DNA demethylation in fragile X syndrome cells. *Epigenetics Chromatin* 2016b; 9: 12.

TABOLACCI E, MOSCATO U, ZALFA F, *et al.* Epigenetic analysis reveals a euchromatic configuration in the FMR1 unmethylated full mutations. *Eur J Hum Genet* 2008a; 16: 1487–1498.

TABOLACCI E, PALUMBO F, NOBILE V, *et al.* Transcriptional reactivation of the FMR1 Gene. A possible approach to the treatment of the fragile X syndrome. *Genes* 2016a; 7(8): 1–16.

TABOLACCI E, PIETROBONO R, MANERI G, et al. Reversion to Normal of FMR1 Expanded Alleles: A Rare Event in Two Independent Fragile X Syndrome Families. Genes 2020; 11: 248.

TABOLACCI E, PIETROBONO R, MOSCATO U, *et al.* Differential epigenetic modifications in the FMR1 gene of the fragile X syndrome after reactivating pharmacological treatments. *Eur J Hum Genet* 2005; 13: 641–648.

TABOLACCI E, POMPONI MG, PIETROBONO R, *et al.* A unique case of reversion to normal size of a maternal premutation FMR1 allele in a normal boy. *Eur J Hum Genet* 2008b; 16(2): 209–214.

TASSONE F, HAGERMAN RJ, LOESCH DZ, *et al.* Fragile X males with unmethylated, full mutation trinucleotide repeat expansions have elevated levels of FMR1 messenger RNA. *Am J Med Genet* 2000; 94(3): 232–236.

TODOROV T, TODOROVA A, KIROV A, *et al.* Fragile X mosaic male full mutation/normal allele detected by PCR/MS-MLPA. *BMJ Case Rep* 2009; 2009: bcr.06.2008.0139.

TORRIOLI M, VERNACOTOLA S, SETINI C, *et al.* Treatment with valproic acid ameliorates ADHD symptoms in fragile X syndrome boys. *Am J Med Genet* A 2010; 152A: 1420-1427.

U.S. NATIONAL LIBRARY OF MEDICINE. FMR1 Gene, 2020. https://ghr.nlm.nih.gov/gene/ FMR1#location. Acessado em 04/04/2020.

USDIN K, HAYWARD BE, KUMARI D, *et al.* Repeat-mediated genetic and epigenetic changes at the FMR1 locus in the Fragile X-related disorders. *Front Genet* 2014; 5: 226.

USDIN K, KUMARI D. Repeat-mediated epigenetic dysregulation of the FMR1 gene in the fragile X-related disorders. *Front Genet* 2015; 6: 192.

VERKERK AJ, PIERETTI M, SUTCLIFFE JS, *et al.* Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* 1991; 65: 905-914.

VILLATE O, IBARLUZEA N, MAORTUA H, *et al.* Effect of AGG Interruptions on FMR1 Maternal Transmissions. *Front Mol Biosci* 2020; 7: 135.

WANG T, BRAY SM, WARREN ST. New perspectives on the biology of fragile X syndrome. *Curr Opin Genet Dev* 2012; 22(3): 256-263.

WANG X, SONG X, GLASS CK, *et al.* The long arm of long noncoding RNAs: Roles as sensors regulating gene transcriptional programs. *Cold Spring Harb Perspect Biol* 2011; 3: a003756.

WILLEMS PJ, VAN ROY B, DE BOULLE K, *et al.* Segregation of the fragile X mutation from an affected male to his normal daughter. *Hum Mol Genet* 1992; 1(7): 511–515.

WILLEMSEN R, BONTEKOE CJ, SEVERIJNEN LA, *et al.* Timing of the absence of FMR1 expression in full mutation chorionic villi. *Hum Genet* 2002; 110(6): 601-605.

XIE N, GONG H, SUHL JA, *et al.* Reactivation of FMR1 by CRISPR/Cas9mediated deletion of the expanded CGG-repeat of the fragile X chromosome. *PLoS One* 2016; 11(10): e0165499.

YANG W, YAN A, XU Y, *et al.* Further identification of a 140bp sequence from amid intron 9 of human FMR1 gene as a new exon. *BMC Genet* 2020; 21: 63.

YRIGOLLEN CM, DURBIN-JOHNSON B, GANE L, *et al.* AGG interruptions within the maternal FMR1 gene reduce the risk of offspring with fragile X syndrome. *Genet Med* 2012; 14(8): 729–736.

ZEESMAN S, ZWAIGENBAUM L, WHELAN DT, et al. Paternal transmission of fragile X syndrome. *Am J Med Genet* A 2004; 129A(2): 184–189.

ZHAO XN, USDIN K. UPS and downs: mechanisms of repeat instability in the fragile X-related disorders. *Genes (Basel)* 2016; 7(9):1–14.

ZHOU Y, HU Y, SUN Q, *et al.* Non-coding RNA in Fragile X Syndrome and Converging Mechanisms Shared by Related Disorders. *Front Genet* 2019;10:139.

Nº	Registro	Procedência	ldade (anos)	Cariótipo	HF	Southern Blotting	PCR de rotina	PCR* Eletroforese em agarose	PCR* Eletroforese Capilar	AmplideX® (Alelo / nº CGG / Metilação)	Obs
1	P076/95	HUPE/UERJ	7	46,XY	F	Rastro 6,5-7,2 + rastro 7,5-8,0		NF		FM / >200 CGG / HMe	
2	P138/96	IFF/FIOCRUZ	13	46,XY	F	5,8+6,6+7,1		NF		FM / >200 CGG / HMe	
3	P138A/96	IFF/FIOCRUZ	26	Xf+	F	6,0+6,6		Rastro 350-400pb		PM (?) / 62 CGG / PMe FM / >200 CGG / HMe	Irmão de P138/96
4	P141/95	IFF/FIOCRUZ	17	46,XY	F	5,8+6,9		NF		Não Funcionou	
5	P162/96	IFF/FIOCRUZ	8	46,XY	F	6,5+7,3					
6	P176/96	IFF/FIOCRUZ	25	46,XY Xf+	F	6,3+7,3				Não Funcionou	
7	P237/97	IPPMG/UFRJ	9	46,XY Xf+	F	6,3		NF		N (?) / 39 CGG / PMe PM / 64 CGG / NMe FM / >200 CGG / PMe	No AmplideX®, pico baixo de N
8	P248/97	IFF/FIOCRUZ	13	46,XY	F	5,9	Positivo	NF		FM / >200 CGG / HMe	
9	P248/02	HUPE/UERJ	3	46,XY [25]	F		Positivo			FM / >200 CGG	Primo de P248/97
10	P290/98	IFF/FIOCRUZ	13	46,XY	Е	6,4		NF		Não Funcionou	
11	P312/98	HMJESUS	11	46,XY	Е	6,1		NF		FM/>200CGG/ HMe	
12	P344/98	HUPE/UERJ	2	46,XY	Е	6,0+6,6		NF		FM / >200 CGG / HMe	
13	P369A/99	UFRJ	11	NI	F		Positivo				Irmão de P369/99
14	P370/99	UFRJ	15	46,XY Xf-	F	6,0 + 6,6		~300pb ~400pb ~500pb	N / 29 CGG PM/ 58 CGG	N / 30 CGG / PMe PM / 59CGG / NMe FM(?)	No AmplideX®, pico baixo de FM
15	P400/99	IFF/FIOCRUZ	11	46,XY	F	6,3		NF		FM / >200 CGG / HMe	
16	P400A/99	IFF/FIOCRUZ	8	NI	F	6,2+6,5		~600pb	PM / 119 CGG	PM / 93 CGG / HMe PM / 122 CGG / NMe FM / >200 CGG / HMe	Paciente falecido
17	P400B/11	HUPE/UERJ	9	NI	F		Positivo	NF		PM + FM	Primo de P400/99
18	P405/99	HUPE/UERJ UFRJ	9	NI	F	5,8				PM / 102 CGG / PMe FM / > 200CGG / HMe	No AmplideX®, 78% digestão

APÊNDICE A – Dados de perfis mutacionais da amostra de pacientes e familiares incluídos nesse estudo (continua)

N٥	Registro	Procedência	ldade (anos)	Cariótipo	HF	Southern Blotting	PCR de rotina	PCR* Eletroforese em agarose	PCR* Eletroforese Capilar	AmplideX® (Alelo / nº CGG / Metilacão)	Obs
19	P420/99	HUPE/UERJ	7	46,XY	F	5,8	Positivo	NF		N / 31 CGG / NMe FM / >200 CGG / HMe	
20	P423/99	IFF/FIOCRUZ	10	NI	F	5,8	Positivo	NF		PM / 164 CGG / HMe PM / 191 CGG / NMe FM / >200 CGG / HMe	
21	P435/99	IFF/FIOCRUZ	5	46,XY	F	6,8	Positivo	~300pb ~500pb		N / 29 CGG / PMe FM / >200 CGG / PMe	No AmplideX®, picos baixos
22	P442/99	ND	ND	ND	ND	6,4		~300pb	N / 29 CGG	FM / >200CGG/ NMe	No AmplideX®, pico baixo de FM
23	P443/00	CEPUERJ	ND	ND	ND	7,2	Positivo	Rastro 300-400pb	N / 26 CGG	FM / >200CGG/ HMe	
24	P452/99	IFF/FIOCRUZ	9	46,XY Xf-	Е	6,0	Positivo	NF		PM / 70 CGG / NMe FM / >200 CGG / HMe	
25	P457/99	HUPE/UERJ	11	46,XY	F	6,2+6,5+7,2	Positivo	NF		FM / >200 CGG / HMe	
26	P457A/99	HUPE/UERJ	11	NI	F			Rastro 250-350pb		FM / >200 CGG / HMe	Primo de P457/99
27	P495/00	CEPUERJ Dra Raquel	23	NI	F	6,4		Rastro 300-400pb	N / 31 CGG	FM / >200 CGG / HMe	
28	P476/00	Dr Ricardo CEPUERJ	14	Xf+	F	6,1		~300pb ~500 pb		N / 29 CGG / NMe FM / >200 CGG / HMe	No AmplideX®, 55% de digestão
29	P504/00	IFF/FIOCRUZ	9	NI	F	7,0	Positivo		N / 28 CGG	FM / >200 CGG / HMe	
30	P537/00	HUGG/UNIRIO	15	46,XY	Е	5,8		Rastro 300-400pb ~800pb ~1500pb	N / 23 CGG I / 48 CGG PM/144CGG	N(?) / 34 CGG / PMe FM / >200CGG / HMe	No AmplideX®, picos baixos de N e padrão onda na FM
31	P562/01	PAM Dra Ruth Rissin CEPUERJ	19	NI	F	5,8+7,5				FM / >200 CGG / HMe	No AmplideX®, 77% de digestão
32	P583/01	HMJesus	4	NI	F	5,8		NF		FM / > 200CGG / HMe	

APÊNDICE A – Dados de perfis mutacionais da amostra de pacientes e familiares incluídos nesse estudo (continuação)

Nº	Registro	Procedência	Idade (anos)	Cariótipo	HF	Southern Blotting	PCR de rotina	PCR* Eletroforese em agarose	PCR* Eletroforese Capilar	AmplideX® (Alelo / nº CGG / Metilação)	Obs
33	P607/01	APABB Dra Ruth Rissin	22	46,XY [30]	F	6,4		Rastro 300-500pb	l / 51 CGG FM / 252 CGG	FM / >200 CGG / HMe	
34	P616/01	CEPUERJ	ND	ND	ND	3,9+4,0+7,2 +7,9+8,9		<350pb	PM / 59 CGG	FM / >200 CGG / PMe FM / >200 CGG / HMe	
35	P632/01	HUPE/UERJ	16	46,XY [25]	NI	5,8		NF		FM / >200 CGG / HMe	
36	P663/02	CEPUERJ	ND	46,XY	ND	6,8		NF		FM / >200 CGG / HMe	
37	P668/02	CEPUERJ	ND	ND	ND	8		NF		N(?) / 31 CGG / NMe	
38	P672/02	IPPMG/UFRJ CEPUERJ	7	NI	Е	8				FM / >200 CGG / HMe	No AmplideX®, 48% digestão
39	P702/02	HUGG/UNIRIO	30	46,XY Xf+ [30]	F			Rastro	FM / 236 CGG FM / 245 CGG	FM / >200 CGG / HMe FM / >200 CGG / HMe	
40	P718/03	IPPMG/UFRJ	5	NI	F		Positivo	~300pb ~500pb	PM /146 CGG FM / 325 CGG	N(?) / 31 CGG / NMe FM / >200 CGG / HMe	Possível deleção na rotina
41	P850/04	IFF/FIOCRUZ	ND	ND	ND		Positivo	Rastro 300-500pb	N / 26 CGG PM /199 CGG	FM / >200 CGG / HMe	
42	P873/04	HUPE/UERJ	7	46,XY Xf+	F		Positivo	NF		FM / >200 CGG / HMe	
43	P908/04	HUPE/ UERJ CEPUERJ	5	NI	Е		Positivo			FM / >200 CGG / HMe	
44	P928/04	HUPE/UERJ	7	46,XY	Е		Positivo	~1500pb ~2000pb		N(?) / 29 CGG / PMe FM / >200 CGG / HMe	
45	P981/05	HUPE/UERJ	13	NI	F		Positivo	Rastro		N / 21 CGG / NMe FM / >200 CGG / HMe	
46	P1013/05	HUPE/UERJ	5	NI	F		Positivo	NF		FM / >200 CGG / HMe	
47	P1033/05	HUPE/UERJ	24	ND	Е		Del ~396 pb			FM / >200 CGG / HMe	
48	P1090/05	HUPE/UERJ	16	NI	F		Positivo			Não Funcionou	
49	P1101/05	IPPMG/UFRJ	ND	ND	F		Positivo	NF		Não Funcionou	

APÊNDICE A – Dados de perfis mutacionais da amostra de pacientes e familiares incluídos nesse estudo (continuação)

Nº	Registro	Procedência	ldade (anos)	Cariótipo	HF	PCR de rotina	PCR* Eletroforese em agarose	PCR* Eletroforese Capilar	AmplideX® (Alelo / nº CGG / Metilação)	Obs
50	P1101A/05	IPPMG/UFRJ	ND	ND	F	Positivo	NF		FM / >200 CGG / HMe	Irmão de P1101/05
51	P1101C/05	IPPMG/UFRJ	ND	ND	F				PM / 117 CGG / NMe FM / >200 CGG / HMe	Irmão de P1101/05
52	P1103/05	IFF/FIOCRUZ	13	46,XY [20]	Е	Positivo	NF		Não Funcionou	
53	P1113/05	HUPE/UERJ	15	NI	F		NF		Não Funcionou	
54	P1135/06	CEPUERJ	ND	46,XY	ND	Positivo			Não Funcionou	
55	P1180/06	HUGG/UNIRIO	4	46,XY	F	Positivo			FM / >200 CGG / HMe	No AmplideX®, 72% de digestão
56	P1201/06	CEPUERJ	ND	46,XY	ND	Positivo			FM / >200 CGG / HMe	
57	P1234/06	IPPMG/UFRJ	9	ND	Е	Del 123 pb			FM / >200 CGG / HMe	
58	P1259/06	IFF/FIOCRUZ	9	NI	F	Positivo	~300pb	N / 38 CGG	N / 39 CGG / NMe FM / >200 CGG / HMe	
59	P1337/13	IPPMG/UFRJ	5	46,XY	Е	Del ~72 pb			FM / >200 CGG / HMe	
60	P1348/06	IFF/FIOCRUZ	8	NI	F	Positivo	~550pb	N / 27 CGG N / 29 CGG N / 35 CGG PM / 56 CGG PM / 86 CGG PM / 92 CGG	FM / >200 CGG / HMe	
61	P1348A/06	IFF/FIOCRUZ	56	NI	NI		~550pb	PM / 106 CGG	PM / 111 CGG	Avô materno de P1348/06. Não apresenta distúrbio neurológico.
62	P1371/07	CEPUERJ	ND	ND	ND	Positivo		NF	FM / >200 CGG / HMe	
63	P1467/07	HUPE/UERJ	6	NI	F	Positivo	~300pb	N / 29 CGG	FM / >200 CGG / HMe	
64	P1513/07	CEPUERJ	5	ND	F	Del 1042pb			FM / >200 CGG / HMe	
65	P1594/07	IPPMG/UFRJ	9	46,XY Xf-	F	Positivo	~300pb	N / 29 CGG N / 35 CGG I / 50 CGG	FM / >200 CGG / HMe	

APÊNDICE A – Dados de perfis mutacionais da amostra de pacientes e familiares incluídos nesse estudo (continuação)

APÊNDICE A – Dados de perfis mutacionais da amostra de pacientes e familiares incluídos nesse estudo (continuação)

Nº	Registro	Procedência	ldade (anos)	Cariótipo	HF	PCR de rotina	PCR* Eletroforese em agarose	PCR* Eletroforese Capilar	AmplideX® (Alelo / nº CGG / Metilação)	Obs
66	P1628/07	HGBonsucesso	13	NI	NI	Positivo	~300pb Rastro 500-600pb	N / 29 CGG PM / 100 CGG PM 117 CGG	FM / >200 CGG / HMe	
67	P1629/07	IFF/FIOCRUZ	18	ND	Е	Del 93 pb			FM / >200 CGG / HMe	
68	P1664/07	IFF/FIOCRUZ	7	NI	F	Positivo			Não Funcionou	
69	P1975/08	UNIGRANRIO / CEPUERJ	11	NI	Е	Positivo			PM / 116 CGG / NMe FM / >200 CGG / HMe	
70	P2115/09	IFF/FIOCRUZ	9	NI	F	Positivo	NF		Não Funcionou	
71	P2149/09	UNIGRANRIO	15	NI	F	Positivo	NF		FM / >200 CGG / HMe	No AmplideX®, 69% de digestão
72	P2273/09	IFF/FIOCRUZ	1	NI	Е	Positivo	NF		FM / >200 CGG / HMe	
73	P2315/09	IFF/FIOCRUZ	15	NI	F	Positivo	NF		FM / >200 CGG / HMe	Gêmeo de P2315A/09
74	P2315A/09	HUGG/UNIRIO	15	NI	F	Del?	~1000pb		FM / >200 CGG / HMe	Gêmeo de P2315/09
75	P2551/10	CEPUERJ	ND	ND	ND	Positivo	NF		PM / 140 CGG / PMe PM / 159 CGG / HMe	
76	P2773/10	M.Público	11	NI	NI	Positivo			FM / >200 CGG / HMe	
77	P3199/11	IPPMG/UFRJ	7	46,XY Xf+	F	Positivo	~600pb	PM / 143 CGG	PM / 147 CGG / NMe FM / >200 CGG / HMe	No AmplideX®, pico baixo de N
78	P3207/11	IPPMG/UFRJ	3	NI	E	Inconclusivo	~900pb	FM / 239 CGG	FM / >200 CGG / HMe	
79	P3231/11	HGBonsucesso	12	NI	F	Inconclusivo	NF		PM / 59 CGG / NMe FM / >200 CGG / HMe	
80	P3274/11	UNIGRANRIO	12	46,XY	F	Inconclusivo	Rastro 300-400pb ~1500pb	I / 49 CGG	FM / >200 CGG / HMe	Possível deleção no ensaio de PCR de rotina
81	P3274B/13	HUGG/UNIRIO	29	NI	F	Positivo	~400pb ~500pb	PM / 56 CGG PM / 97 CGG	PM / 58 CGG / NMe PM / 99 CGG / NMe FM / >200 CGG / HMe	Irmão de P3274/11. No AmplideX®, pico baixo de FM
82	P3350/11	IFF/FIOCRUZ	3	NI	Е	Positivo	~650pb	PM/ 117CGG PM/ 133CGG	PM / 138 CGG / NMe FM / >200 CGG / HMe	
83	P3384/11	HUPE/UERJ	9	NI	F	Inconclusivo			FM / >200 CGG / HMe	

PCR* PCR* AmplideX® PCR de Idade Cariótipo HF (Alelo / nº CGG / Procedência Eletroforese Eletroforese Obs Nº Registro (anos) rotina em agarose Capilar Metilação) PM / 80 CGG / NMe P3446/11 **UNIGRANRIO** 9 NI NI Inconclusivo ~400pb Possível deleção na rotina 84 FM / >200 CGG / HMe P3459/11 HUPE/UERJ 4 NI Е 85 Positivo Е NF 86 P3946/12 HUPE/UERJ 7 NI Positivo FM / >200 CGG Avô materno de P3946/12. HUPE/UERJ 73 NI Е PM / 66 CGG / NMe No AmplideX®, digestão e 87 P3946B/16 ~400pb PM / 64 CGG referência não funcionaram ~300pb N / 30 CGG / NMe ~500pb NI F 88 P4055/13 **IFF/FIOCRUZ** 4 Del? PM / 124 CGG / NMe Possível delecão na rotina ~600pb FM / >200 CGG / HMe ~700pb P4093/13 **IFF/FIOCRUZ** 7 46,XY [20] F Т NF N/42 CGG/NMe 89 ~500pb Е Possível deleção na rotina 90 P4101/13 HUPE/UERJ 14 NI Del? FM / >200 CGG / HMe ~2000pb NI F 91 P4112/13 HUGG/UNIRIO 17 Inconclusivo FM / >200 CGG / HMe PM / 162 CGG / HMe 46,XY [20] F HUGG/UNIRIO NF PM / 197 CGG / HMe 92 P4121/13 13 Inconclusivo FM / >200 CGG / HMe Rastro N / 30 CGG **IFF/FIOCRUZ** 46,XY [20] F FM / 214CGG 93 P4181/13 11 Positivo 300-550pb FM / >200 CGG / HMe FM / 272CGG ~900pb 46,XY Xf-F 6 94 P4183/13 HUPE/UERJ Т N / 45 CGG [100] **UNIGRANRIO** 6 Е P4194/14 NI Positivo ~900pb FM / >200 CGG / HMe 95 No AmplideX®, padrão PM(?) / 182CGG / PMe onda e pico alto próximo à 96 46,XY [22] F P4210/14 IEDE 11 Positivo NF FM / >200 CGG / NMe região controle. Será deleção? Gêmeo de P4210/14. No AmplideX®, padrão onda e PM / 182 CGG / PMe IEDE 97 P4210A/14 11 46,XY [22] F ~800pb pico alto próximo à região FM / >200 CGG / NMe controle. Será deleção?

APÊNDICE A – Dados de perfis mutacionais da amostra de pacientes e familiares incluídos nesse estudo (continuação)

Nº	Registro	Procedência	ldade (anos)	Cariótipo	HF	PCR de rotina	PCR* Eletroforese em agarose	PCR* Eletroforese Capilar	AmplideX® (Alelo / nº CGG / Metilação)	Obs
98	P4257/14	IFF/FIOCRUZ	8	46,XY	F	I			I / 52 CGG / NMe	
99	P4259/14	UNIGRANRIO	16	NI	Е	I			I / 46 CGG / NMe	
100	P4442/15	HUPE/UERJ	13	NI	Е	Positivo	~700pb	N / 29 CGG PM/ 156CGG	N / 30 CGG / PMe FM / >200 CGG / HMe	
101	P4468/15	HUPE/UERJ	6	46,XY [20]	NI		~400pb	I / 47 CGG	I / 48 CGG / NMe	
102	P4484/16	HUPE/UERJ	5	46,XY [20]	NI	Positivo			FM / >200 CGG	
103	P4484D/16	HUPE/UERJ	69						РМ	Avô do P4484/16, sem distúrbio neurológico
104	P4644/19	HUPE/UERJ	8	46,XY	F	Inconclusivo	Inconclusivo		FM / >200 CGG / HMe	

APÊNDICE A – Dados de perfis mutacionais da amostra de pacientes e familiares incluídos nesse estudo (conclusão)

Legenda: *Em um trabalho anterior, o grupo desenvolveu um ensaio de PCR, capaz de amplificar fragmentos no gene *FMR1* até pequenas mutações completas, descrita por Felipe Martins Araújo (2018) em sua monografia de conclusão de curso de Graduação em Ciências Biológicas (IBRAG/UERJ). HF: história familiar de DI, relatada pela família; F: caso familiar; E: caso esporádico; NI: não informado; ND: não disponível; NF: não funcionou; N: alelo normal (<44 CGGs); I: alelo intermediário (45-54 CGGs); PM: pré-mutação (55-200 CGGs); FM: mutação completa (>200 CGGs); NMe: não metilado (<20% de metilação); PMe: parcialmente metilado (20-80% de metilação); HMe: hipermetilado (>80% de metilação). Fonte: A autora, 2022.

APÊNDICE B – Fotodocumentação da eletroforese realizada com os produtos do ensaio de PCR desenvolvida anteriormente pelo grupo



Legenda: Em um trabalho anterior do grupo, foi desenvolvida uma PCR, capaz de amplificar fragmentos no gene *FMR1* até pequenas mutações completas, descrita por Felipe Martins Araújo (2018) em sua monografia de conclusão de curso de Graduação em Ciências Biológicas (IBRAG/UERJ). Nessas reações, foram utilizados dois oligonucleotídeos, que amplificam fragmentos até 362 pares de base para alelos normais, entre 365 a 392 pares de base para alelo intermediário, entre 393 a 824 pares de base para pré-mutação, e acima disso já abrange a região de mutação completa. Essas reações foram aqui realizadas em busca de fragmentos menores do que 362 pares de base, o que pode representar uma possível deleção. Na imagem B, na quinta *lane*, foi aplicado o produto de PCR realizada com o produto de PCR aplicada na nona lane da imagem A. Além disso, ainda na imagem B, nas *lanes* 6, 9, 11, e 13, foram aplicados os produtos de PCR realizadas com 200 ng de DNA, enquanto nas *lanes* 7, 10, 12 e 14, foram aplicados produtos de PCR realizadas com 65 ng de DNA. pb: pares de bases; ng: nanogramas; 1kb Plus: marcador de peso molecular 1 kb Plus. Aqui podemos obervar fragmentos de tamanhos menores do que o fragmento do controle normal nas lanes referentes aos pacientes P718/03, P4101/13, P4210A/14, P420/99, P928/04, o que pode representar uma possível deleção.

Registro	Genótipo	Oligonucle	otídeos 1	Oligonucle	otídeos 2	Oligonucle	otídeos 3	Oligonucleotídeos 4	
		Tamanho do	Altura do	Tamanho do	Altura do	Tamanho do	Altura do	Tamanho do	Altura do
		Fragmento	pico	Fragmento	pico	Fragmento	pico	Fragmento	pico
Controle 19	Controle feminino	236,4	7538	230,5 263,5	5209 4660	202,7 205,8	9496 9222	265,1	8290
Controle 07	Controle masculino	236,4	9782	244,5 272,1	5282 4541	205,9 207,0	8270 8634	251,9 265,1	7434 7302
P507/00	Controle Klinefelter	236,4	7387	230,5 249,9 251,9	6442 5994 5579	202,6 205,7	2912 7922	255,7 265,2	2243 1127
P2771/10	Controle Klinefelter	236,3 263,8	11013 3662	244,8 259,8 272,3	6029 5294 4748	205,7 206,9	4465 2368	255,6 265,0	6011 3225
P237/97	N+PM+FM	236,4 246,5	1753 2491	264,9 272,7	6166 5828	205,9	6697	255,9	655
P370/99	N+PM+FM	236,4	8701	230,9 252,3	9334 7162	202,5 205,6	7900 9244	255,7	8689
P420/99	N+FM	326,5 241,5	4813 4482	244,7 264,6	7423 6682	205,7 206,8	6460 6343	NF	NF
P435/99	N+FM	NF	NF	230,6 252,0	3449 2926	202,6 205,8	2948 4296	265,2	2311
P476/00	N+FM	246,5	8232	234,8	7793	202,7 206,0	4671 5953	NF	NF
P537/00	N+FM	236,4 246,6	5505 6410	230,5 247,8	6246 5937	202,6 205,7	3628 4643	259,4 265,2	981 1016
P668/02	N+FM	NF	NF	247,9 250,0	3603 3479	205,9 207,0	5582 5386	NF	NF
P718/03	N+FM+del?	236,4	6082	250,0 268,3	8421 7862	205,8	7626	265,2	4029
P928/04	N+FM	236,4	8060	230,6 252,1	7577 6843	202,7 205,8	7455 9484	255,7 265,2	2096 2212
P981/05	N+FM	236,4	5750	259,8	9323	205,9	8844	251,9 265,1	2256 2075
P1234/06	FM+del	259,4	3614	231,0	8494	202,8 203,9	3321 3141	255,8	1941

APÊNDICE C – Dados de marcadores nos cromossomos X e Y, internos ao gene SHOX (continua)

Registro	Genótipo	Oligonucl	eotídeos 1	Oligonucl	eotídeos 2	Oligonucl	eotídeos 3	Oligonucleotídeos 4	
		Tamanho do Fragmento	Altura do pico						
P1259/06	N+FM	236,4 246,5	5835 4961	244,6 255,9	6390 5686	205,9 206,9	12220 12036	255,7 265,1	2920 3076
P1337/13	FM+del	259,4	3316	231,1 256,3	6574 5751	205,8	8130	252,1 255,9	686 676
P1629/07	FM+del	236,5	7183	245,4 248,7	5790 5655	205,9 207,1	1907 1824	265,4	1563
P2315A/09	FM+del?	236,5	5917	240,1 256,7	5245 4675	206,2	10174	252,3 256,1	263 235
P3274/11	FM+del?	236,4	1452	252,8 269,1	3027 2829	206,0	2810	NF	NF
P3446/11	PM+FM+del?	236,3	9884	230,6 264,5	8813 8268	202,6 205,7	6188 8729	255,8	2462
P4101/13	FM+del?	236,4 246,5	5319 4413	269,0 272,9	8068 7367	205,9	9997	256,0 265,4	2051 2125
P4210/14	PM+FM+del?	236,4	13859	252,0 259,8	8320 8251	205,8	4484	251,9 255,7	1567 1537
P4210A/14	PM+FM+del?	236,3	11444	252,0 259,7	2537 2358	205,7	7932	252,0 255,8	3454 3385
P4442/15	N+FM	NF	NF	239,2 268,4	9261 10488	205,8	6453	252,0 265,2	1775 1845

APÊNDICE C – Dados de marcadores nos cromossomos X e Y, internos ao gene SHOX (conclusão)

Legenda: NF: não funcionou; N: alelo normal (≤44 CGGs); I: alelo intermediário (45-54 CGGs); PM: pré-mutação (55-200 CGGs); FM: mutação completa (>200 CGGs). Fonte: A autora, 2022.

ANEXO A – Aprovação do projeto pelo Certificado de Apresentação para Apreciação Ética (CAAE)

Plataforma				
		lnforme o E-mail	Informe a Senha	LOGIN
			Esqueceu a senha?	Cadastre-se v3.2
	Você está em: Público > Confirmar Aprovação pelo CAAE ou Parecer			
	CONFIRMAR APROVAÇÃO PELO CAAE OU PARECER			
	Informe o número do CAAE ou do Parecer: Número do CAAE: Número do Parecer 46769315.5.0000.5259 Esta consulta retorna somente pareceres aprovados. Caso não aprese aprovado. DETALHAMENTO	: Pesquisar ente nenhum resultado, o número do parecer informado não é válido ou não corre	esponde a um parecer	*
	Título do Projeto de Pesquisa:			
	Investigação de fatores genéticos e epigenét	icos associados à deficiência intelectual 🏑		
	Número do CAAE:	Número do Parecer:		
	46769315.5.0000.5259	1524485		
	Quem Assinou o Parecer:	Pesquisador Responsável:		
	Jorge Aives de Aimeida Venancio Data Início do Cronograma: 01/08/2016 01/08/2026	Contato Público: Cintia Barros SAntos Rebouças		
	Voltar			٢

ANEXO B – Termo de consentimento livre e esclarecido (continua)



Universidade do Estado do Rio de Janeiro Instituto de Biologia Roberto Alcantara Gomes Departamento de Genética

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

TÍTULO DO PROJETO: INVESTIGAÇÃO DE FATORES GENÉTICOS E EPIGENÉTICOS ASSOCIADOS À DEFICIÊNCIA INTELECTUAL E FENÓTIPOS RELACIONADOS

INSTITUIÇÃO RESPONSÁVEL: UNIVERSIDADE DO ESTADO DO RIO DE JANEIRO PESQUISADORA RESPONSÀVEL: PROF^A DR^A CÍNTIA BARROS SANTOS-REBOUÇAS NOME DO PACIENTE:______DN:_____ REGISTRO: PROCEDÊNCIA:

O projeto "Investigação de fatores genéticos e epigenéticos associados à deficiência intelectual e fenótipos relacionados", realizado pelo Serviço de Genética Humana da Universidade do Estado do Rio de Janeiro (SERVGEN), tem como objetivo caracterizar aspectos de natureza genética que podem levar à deficiência intelectual em homens e mulheres, assim como a outras condições familiares relacionadas a esta condição. Através do conhecimento gerado a partir deste projeto, pretendemos melhorar o conhecimento sobre as causas da deficiência intelectual e doenças a ela relacionadas, gerando dados científicos que possam ser utilizados para o desenvolvimento de metodologias de diagnóstico e para a elaboração futura de estratégias voltadas para o tratamento/prevenção destas condições.

Os indivíduos selecionados para participarem deste projeto são pessoas diagnosticadas clinicamente como portadoras de deficiência intelectual de causa desconhecida ou de doenças relacionadas e seus familiares. As condutas adotadas nesta pesquisa seguirão as normas de Bioética descritas na resolução 466/2012 do Conselho Nacional de Saúde. Todo participante tem garantido o direito de pensar sobre a sua participação e/ou conversar com outras pessoas, caso possa interessar, antes de decidir participar desta pesquisa. Qualquer indivíduo selecionado poderá se recusar a participar ou retirar o seu consentimento em qualquer momento da pesquisa, sem que haja penalização ou prejuízo ao seu atendimento.

O exame gratuito será realizado a partir da coleta de uma pequena amostra de sangue (5 mL), podendo levar em alguns pacientes, ao aparecimento de uma leve marca avermelhada no local da punção. Todos os procedimentos serão acompanhados por uma equipe de geneticistas e biólogos. O material biológico ficará armazenado em um Biorrepossitório sob o encargo da pesquisadora responsável nas dependências no Serviço de Genética Humana da Universidade do Estado do Rio de Janeiro, sendo utilizado somente para fins acadêmico-científicos. Cada amostra de material biológico será identificada por números específicos e somente poderá ser utilizada para investigações futuras, mediante o consentimento dos doadores (ou de seus responsáveis) e/ou da aprovação do Comitê de Ética em Pesquisa (CEP) da Instituição e ainda da Comissão Nacional de Ética em Pesquisa (CONEP), quando for o caso. Cabe ao doador ou ao seu responsável legal o direito de retirar as amostras obtidas do Biorrepossitório a qualquer momento, se assim o desejar, através de manifestação por escrito assinada. A eventual transferência do material biológico humano armazenado para um

ANEXO B – Termo de consentimento livre e esclarecido (conclusão)

outro Biorrepositório ou Biobanco, da própria ou de outra instituição, será comunicada ao sujeito da pesquisa, sempre que possível ou, na impossibilidade, o CEP/CONEP serão consultados. Além disso, o participante da pesquisa ou seu responsável será informado sobre eventual perda ou destruição de suas amostras biológicas, bem como, sobre o encerramento do Biorrepositório ou Biobanco, quando for o caso.

Você pode querer tomar conhecimento ou não das informações genéticas geradas por esta pesquisa. Para isso, a qualquer momento, comunique a um dos pesquisadores o seu interesse nos resultados das análises.

Desejo ter acesso a todos os resultados das análises realizadas com a amostra biológica coletada. () SIM NÃO ()

Autorizo a divulgação de doenças e/ou mutações acidentalmente encontradas provenientes dos resultados obtidos a partir da amostra coletada.

() SIM NÃO ()

As condições de sigilo quanto aos resultados dos exames, assim como, a identidade dos pacientes será preservada, impedindo qualquer tipo de discriminação ou estigmatização, seja individual ou coletiva. Os resultados somente serão divulgados ao paciente ou a seu responsável legal. Ao responsável será fornecido todo e qualquer esclarecimento que se fizer necessário antes, durante e após a realização da pesquisa. Os dados desta pesquisa podem vir a ser divulgados e publicados científicamente em congressos e revistas científicas, respeitando a privacidade e o anonimato dos participantes, visando evitar a exposição e a discriminação dos doadores e de suas famílias.

Caso sejam encontradas alterações genéticas, o aconselhamento genético será oferecido de forma conjunta pela coordenadora do projeto de pesquisa e os geneticistas clínicos que acompanham o paciente, visando a orientação sobre os riscos de recorrência da condição, possível curso da condição e as condutas terapêuticas disponíveis, quando for o caso.

Eu, , responsável pelo paciente acima citado, declaro que li e compreendi o que me foi explicado e, desta forma, autorizo voluntariamente a participação no estudo, bem como, a coleta, o depósito, o armazenamento e a utilização do material biológico coletado.

Rio de Janeiro, ____ de ____ de ____

identidade

n°

Assinatura do responsável

Assinatura do pesquisador

🕾 Endereço e telefone para contatos (Horário de atendimento: 9:00 – 17:00 horas): Rua São Francisco Xavier, 524 - Pavilhão Haroldo Lisboa da Cunha - sala 501F - Maracanã - Rio de Janeiro - RJ - CEP 20550-013 - Tel: (21) 2334-0039 - Email: cbs@uerj.br.

Em caso de dúvidas sobre os aspectos éticos envolvidos nesta pesquisa, consulte o Comitê de Ética Institucional responsável (CEP/HUPE)- Tel: (21) 2868-8253. E-mail: cep-hupe@uerj.br