

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA VEGETAL



**Síndrome de X-Frágil em Portugal: Caracterização
molecular por SNPs e STRs**

Joana Maria Geraldés da Rocha Loureiro

Mestrado em Biologia Molecular e Genética

2010

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA VEGETAL



**Síndrome de X-Frágil em Portugal: caracterização
molecular por SNPs e STRs**

Joana Maria Galdes da Rocha Loureiro

Dissertação orientada por:

Doutora Paula Maria Vieira Jorge

Técnica Superior de Saúde do ramo de Genética – Assistente Principal, UMO-P, INSA-Porto

Professor Doutor Pedro João Neves Silva

Professor Auxiliar da Faculdade de Ciências da Universidade de Lisboa

Mestrado em Biologia Molecular e Genética

2010

Agradecimentos

Quero agradecer a todos aqueles que me apoiaram e que de alguma forma contribuíram para a realização deste trabalho:

A toda a Unidade de Genética Molecular, Centro de Genética Médica Dr Jacinto Magalhães do INSA Porto, pela atenção que demonstraram ao longo do meu trabalho. Em especial à Dr^a Rosário Santos, por ter proporcionado a elaboração da minha Tese na Unidade de Genética Molecular, à Luisa pela separação dos DNAs usados na População Controlo, ao Eurico pelos “truques” dicas importantes para ter bons resultados na técnica de hrMCA. À Bárbara, por partilha de resultados. Quero também deixar um especial obrigado à Isabel pela amizade, apoio e preciosa ajuda durante todo o trabalho.

Ao Professor António Amorim pelas excelentes contribuições na interpretação dos resultados e proveitosas discussões científicas.

Ao Doutor Paulo Pinho e Costa pela disponibilidade e discussões científicas.

Ao Miguel e à Sara Reis, por resolverem os meus problemas burocráticos na FCUL.

Ao meu irmão, pelas explicações nas “formulas reboscadas” do excel.

Aos meus pais, por todo o apoio, suporte, motivação e por terem sido o meu porto de abrigo ao longo desta etapa como de todas as outras da minha vida.

Ao Professor Pedro Silva pela sua orientação e críticas ao trabalho.

E principalmente à minha Orientadora Doutora Paula Jorge, pela sua grande dedicação e por colocar ao meu alcance todos os materiais, condições para elaboração do trabalho e pelos contactos efectuados com o Professor António Amorim e Doutor Paulo Pinho e Costa. E pela amizade demonstrada ao longo do ano lectivo.

Lista de abreviaturas

AMPA - *α-amino-3-hydroxi-5-metil-4-isoxazole-propionico*

CJM - Centro Genética Médica Dr Jacinto Magalhães, INSA

DNA - Ácido desoxiribonucleico

FMR1- *Fragile-X mental retardation 1gene*

FMRP- *Fragile-X mental retardation protein*

FXS- Fragile-X syndrome/ Síndrome de X-Frágil

gDNA - DNA genómico

HRM- *High Resolution Melting*

hrMCA - high resolution Melting Curve Analysis

kb - Kilobases

PCR - *polimerase chain reaction*/ reacção de polimerização de cadeia

pb - par de base

RISC - *RNA Induced Silencing Complex/ Complexo de Indução do Silenciamento do RNA*

SNP - *Single Nucleotide Polymorphism*/ Polimorfismo de Base Única

STR - *Short Tandem Repeat*

UMO-P - Unidade de Genética Molecular-Porto

UTR - *Untranslated region*/ Região não traduzida

Resumo

A Síndrome de X-Frágil (FXS) é a causa de atraso mental hereditária mais comum, com penetrância reduzida em mulheres e uma prevalência na Europa estimada de 1/4000 homens e 1/8000 mulheres. A mutação responsável pela FXS resulta da expansão do número de repetições [CGG] acima de 200, presentes na região 5'UTR do gene *FMR1* (Xq27.3, cromossoma X).

Com o objectivo de caracterizar a população X-Frágil Portuguesa e eventualmente identificar cromossomas fundadores, foram estudados quatro *Short Tandem Repeats* (STRs) e três *Single Nucleotide Polymorphisms* (SNPs) em 213 indivíduos normais e 123 indivíduos com Síndrome X-Frágil, perfazendo duas amostras representativas de todo o País. Os STRs usados DXS998, DXS548, FRAXAC1 e FRAXAC2, flanqueiam a região repetitiva [CGG] e foram escolhidos com base em estudos elaborados noutras populações. Estes STRs foram amplificados por PCR Multiplex, com primers fluorescentes e os produtos analisados no programa *GeneMapper v4.0*. Para aumentar a informatividade da análise haplotípica, foram construídas Linhagens com três SNPs: rs971000, rs29282 e rs25715, que foram detectados por *High Resolution Melting* e nunca tinham sido utilizados na construção de haplótipos ligados ao *FMR1*.

Na análise dos resultados, utilizou-se o programa *Arlequin v3.5*, para calcular frequências alélicas e haplotípicas nas duas populações e haplótipos partilhados. Recorreu-se a um teste de Chi-Quadrado, onde se verificou que a frequência de cada alelo é significativamente diferente nas populações Controlo e X-Frágil. Foram desenhadas redes de haplótipos STRs nas Linhagens estabelecidas com SNPs no programa *Network v4.5*. para inferir a genealogia haplotípica e identificar haplótipos fundadores.

Foi observada uma diferente distribuição haplotípica entre as populações e apenas oito haplótipos são partilhados. Na população Controlo é possível observar um haplótipo representativo da população (17%), o A1, que nunca se encontra na população FXS. A população FXS, por sua vez mostrou-se muito heterogénea não tendo sido possível identificar nenhum cromossoma fundador.

Palavras-chave: Síndrome de X-Frágil, Haplótipo, Linhagem, *Single Nucleotide Polymorphism*, *Short Tandem Repeat*

Abstract

Fragile-X syndrome (FXS) is the most common cause of hereditary mental retardation, with a reduced penetrance in women, and an estimated prevalence in Europe of 1 / 4000 males and 1 / 8000 women. The FXS responsible mutation results from the expansion of the number of [CGG] repeats above 200, in the *FMR1* gene 5'UTR region (Xq27.3, X chromosome).

In an attempt to characterize X-fragile Portuguese population and identify a possible founder effect, genomic DNA of 213 normal individuals and 123 fragile X syndrome individuals was studied. Haplotype background around *FMR1* locus was constructed using four Short Tandem Repeats (STRs), namely DXS998, DXS548, FRAXAC1 and FRAXAC2 flanking the [CGG] repeat region, that were chosen based on studies in other populations. These STRs were amplified by multiplex PCR with fluorescent primers and the size of the alleles was analyzed in GeneMapper software v4.0. In order to increase the informativity three Single Nucleotide Polymorphisms (SNPs): rs971000, rs29282 and rs25715 that had never applied in *FMR1* linked haplotypes were used to construct Lineages and were detected by high resolution melting.

Arlequin v3.5 was used for allele and haplotype frequency calculation in both populations as well as shared haplotypes. A chi-square test, enable showed that the frequency of each allele is significantly different in control populations and fragile X syndrome. Haplotype networks were designed for SNPs Lineages with software *Network* v4.5 in order to infer about haplotypes genealogy and to identify founder chromosomes.

A different haplotype distribution was found among these two populations. Only eight haplotypes were common to both populations. Furthermore, no founder haplotype was identified. Indeed in the Control population the most represented haplotype (17%), A1, was never found in any FXS individual, whereas in the FXS population a high diversity of haplotypes was found.

Key words: Fragile-X Syndrome, Haplotype, Lineage, Single Nucleotide Polymorphism, Short Tandem Repeat

Índice

Lista de abreviaturas.....	i
Resumo.....	iii
Abstract	v
I. Introdução.....	1
I.1 A Síndrome de X-Frágil	1
I.1.1 Características moleculares do gene <i>FMR1</i>	1
I.1.2 Fénótipos associados ao aumento do número de repetições [CGG]	2
I.1.3 A proteína FMRP na Síndrome de X-Frágil	3
I.2 Marcadores Moleculares.....	4
I.2.1 SNPs.....	4
I.2.2 STRs	4
I.3 Análise haplotípica na Síndrome de X-Frágil – efeito fundador?	6
II. Objectivos.....	6
III. Materiais e Métodos	7
III.1 Características da amostra	7
III.2 Recolha de Amostras.....	7
III.3 Marcadores SNPs e STRs	7
III.4 Desenho de primers.....	9
III.5 Determinação número de tripletos [CGG]	9
III.6 Reacção de sequenciação	9
III.7 PCR Multiplex	10
III.8 High Resolution Melting Curve Analysis - hrMCA.....	10
III.9 Análise Estatística.....	11
III.10 Construção de Linhagens e haplótipos	12
IV. Resultados	13
IV.1 Alelos encontrados nos marcadores STR e respectiva nomenclatura	13
IV.2 Distribuição dos Alelos STR na população Controlo e X-Frágil.....	14
IV.3 Distribuição dos alelos SNP na população Controlo e X-Frágil.....	17
VI.4 Linhagens.....	18
IV.5 Haplótipos	19
IV.6 Quantificação do número de repetições [CGG] na população Controlo	22
IV.7 Polimorfismos nos amplicões STRs.....	22
V. Discussão de Resultados.....	24
V.1 Linhagens SNP	24

V.2 Haplótipos STR	25
V.3 Distribuição dos haplótipos STR pelas Linhagens SNP	25
V.4 A controversa origem da mutação completa.....	26
VI. Conclusões e Perspectivas futuras	28
VII. Referências.....	29
VII.1 Referências Bibliográficas	29
VII.2 Referências Electrónicas	31
Anexos	I
Anexo 1- Primers utilizados	I
Anexo 2- Determinação do número de tripletos [CGG]	II
Anexo 3- Amplificação dos STRs por PCR Multiplex	II
Anexo 4- <i>High Resolution Melting Curve Analysis</i> - hrMCA	III
Anexo 5 – Haplótipos e frequências haplotípicas na população Controlo	III
Anexo 6 – Haplótipos e frequências haplotípicas na população com Síndrome de X-Frágil	V

I. Introdução

I.1 A Síndrome de X-Frágil

A Síndrome de X-Frágil (FXS - *Fragile X Syndrome*) é a causa de atraso mental hereditária mais comum, com uma penetrância reduzida em mulheres (Hagerman, 1996), e uma prevalência na população europeia estimada de 1/4000 homens e 1/8000 mulheres (Tuner *et al*, 1996). O seu nome advém da observação, por técnicas de citogenética, de um sítio frágil, localizado no cromossoma X na região Xq27.3 (locus FRAXA) (Sutherland, 1977) (Figura 1). Este sítio frágil é consequência da expansão de uma região polimórfica repetitiva [CGG], localizada na região 5'UTR do gene *FMR1* (Yu *et al*, 1991).

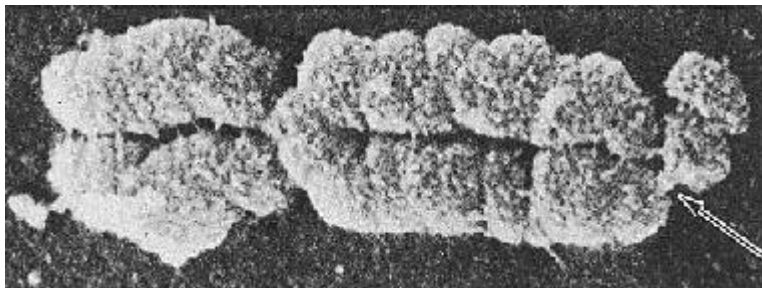


Figura 1 Sítio frágil na região Xq27.3. Adaptado de *Strachan and Read, 2005, Capítulo 11.5.*

I.1.1 Características moleculares do gene *FMR1*

O número de repetições do triplete [CGG] no gene *FMR1* é polimórfico na população. Os alelos deste gene podem-se dividir-se em 3 classes consoante o número de tripletos [CGG], denominados: (1) – normais, com 5-55 repetições, (2) - pré-mutados, que contêm entre 55-200 repetições e (3) - mutação completa, que possuem mais de 200 tripletos [CGG] (Fu, Kuhl e Pizzuti, 1991). Há ainda autores que consideram a existência de um outro grupo de alelos denominados intermédios (*grey zone*) com repetições entre 50 e 55 [CGG], (ou 40 e 55, dependendo dos grupos de investigação). Por norma os alelos de tamanho normal são considerados estáveis, isto é, geralmente não ocorre alteração do número de [CGG] quando transmitidos às gerações seguintes, ao contrário dos alelos pré-mutados ou com mutação completa cujo número de repetições [CGG] pode expandir massivamente numa meiose. Já a instabilidade na transmissão dos alelos intermédios, ainda não está definida (Eichler *et al*, 1994). Acima das 200 repetições ocorre uma inactivação do gene *FMR1* devido à metilação não só desta região polimórfica como do promotor do gene e ilhas CpG adjacentes, levando ao silenciamento transcricional do gene.

Apesar das causas moleculares da expansão dos [CGG] não estarem totalmente esclarecidas existem alguns factores de risco associados à instabilidade dos alelos expandidos (pré-mutados e com mutação completa). Sabe-se que os alelos de tamanho normal não correspondem a cadeias de [CGG] completamente “puras”, isto é, são normalmente interrompidos por [AGG] que geralmente ocorrem a cada 9 ou 10 [CGG]. Assim, alguns autores sugerem que a instabilidade dos alelos expandidos possa estar associada quer ao número, quer ao padrão e localização das interrupções [AGG]. Na maioria dos alelos pré-mutados e expandidos não se observa nenhuma interrupção [AGG], ou apenas uma ou duas interrupções em toda a sequência polimórfica, localizando-se normalmente na região mais proximal, uma vez que esta perda de [AGG] se dá na região 5' da sequência. A perda destas interrupções a 5' pode ocorrer por fenómenos de *slippage*, via transversão de A→C ou por *crossing-over* desigual.

As mutações completas têm origem a partir de alelos pré-mutados, que são meioticamente instáveis e podem expandir “massivamente” para mutação completa. A transição de alelos pré-mutados para alelos mutados está positivamente associada com a transmissão pela linha germinal feminina, onde os alelos pré-mutados podem expandir para mutação completa, resultando numa descendência com Síndrome de X-Frágil (Fu *et al*, 1991). As expansões [CGG] são altamente instáveis pelo que alguns indivíduos podem exibir uma heterogeneidade somática, tanto ao nível do número de repetições como quanto ao seu estado de metilação, estes indivíduos portadores de alelos com graus de metilação e/ou número de repetições [CGG] diferente nas células somáticas, denominam-se mosaicos.

1.1.2 Fénótipos associados ao aumento do número de repetições [CGG]

Os doentes com a mutação completa (> 200 [CGG]) causadora da Síndrome de X-Frágil caracterizam-se essencialmente por atraso mental, orelhas proeminentes, face longa, comportamento autista e macroorquidismo (este último no caso dos indivíduos de sexo masculino). Há ainda indivíduos, designados por *High Functional Males*, que possuem mutação completa não metilada, nestes casos não há manifestação clínica da Síndrome, uma vez que não ocorre a inactivação do gene *FMR1* (Hagerman, 1996).

Até recentemente não era atribuído qualquer significado clínico aos indivíduos portadores de alelos pré-mutados, para além do risco de expansão nas gerações futuras. Presentemente sabe-se que podem ocorrer diferentes manifestações fenotípicas nos portadores de alelos pré-mutados e que estas manifestações fenotípicas variam consoante se trata de indivíduos do sexo masculino ou feminino. No primeiro caso está descrita a Síndrome de ataxia e tremor ligada à Síndrome de X-Frágil, (FXTAS), presente

em 50% dos indivíduos com mais de 70 anos. No caso de ~20% das mulheres portadoras de pré-mutações pode-se observar uma insuficiência ovárica precoce ligada à Síndrome de X-Frágil, ou seja, falência ovárica antes dos 40 anos, (FXPOI) (Hagerman e Hagerman, 2004).

1.1.3 A proteína FMRP na Síndrome de X-Frágil

Apesar de ainda não serem conhecidos os mecanismos moleculares que levam à expansão dos [CGG] sabe-se que a consequente ausência da proteína codificada pelo gene *FMR1*, a FMRP, é a responsável pelo aparecimento da sintomatologia da Síndrome de X-Frágil. A proteína FMRP liga-se especificamente a RNAs mensageiros alvo, fazendo parte do complexo RISC¹ e sendo responsável pela regulação da síntese proteica nas dendrites dos neurónios. Os indivíduos com FXS possuem uma morfologia neuronal característica, com espinhas detriticas muito finas e compridas, devido à incorrecta regulação da síntese proteica ao nível dos neurónios. Pensa-se que a FMRP regule negativamente a síntese proteica estimulada pela activação do receptor de glutamato metabotrópico 1 (mGluR1) (Bear, Huber e Warren, 2004). A FXS resulta então, pelo menos parcialmente, de uma resposta exagerada da estimulação do mGluR. Uma das consequências primárias da ausência da FMRP parece ser a excessiva internalização do receptor α AMPA² em resposta à estimulação do mGluR1, que conduz a alterações da actividade sináptica (Figura 2).

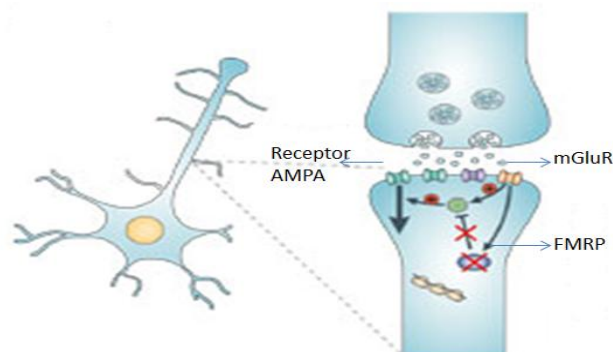


Figura 2 Inibição da síntese proteica resultante da estimulação do receptor de glutamato metabotrópico1 (mGluR1) e consequente excessiva internalização do receptor α AMPA. (Adaptado de Gatchel e Zoghbi, 2005)

- ¹ **RISC** – *RNA Induced Silencing Complex*, complexo multiproteico responsável pelo silenciamento de mRNAs específicos.
- ² **Receptor α AMPA** – Receptor *α -amino-3-hydroxi-5-metil-4-isoxazole-propionico*, receptor transmembranar ionotrópico do glutamato que medeia a rápida transmissão sináptica no sistema nervoso central.

Baseado na descrição dos mecanismos e vias moleculares envolvidas nos sintomas da Síndrome de X-Frágil, tem-se vindo a tentar desenvolver terapias moleculares, tomando como alvo o mGluR1 ou o receptor AMPA com o objectivo de prevenir alguns sintomas da Síndrome de X-Frágil. O diagnóstico precoce da Síndrome desempenha um papel importante na eficácia destas terapias.

I.2 Marcadores Moleculares

Um marcador molecular é um elemento genético, identificável por técnicas de biologia molecular e que permite o estudo comparativo de indivíduos. Um marcador molecular pode estar em *linkage* com um determinado locus, e pode ser associado a uma doença ou característica genética. Normalmente são escolhidos conjuntos de marcadores associados ao *locus*/gene de interesse e aplicados na construção de mapas de *linkage* e estudos de análise de segregação. Dadas as suas características são consideradas ferramentas essenciais em estudos evolutivos e populacionais (Jarne e Lagoda, 1996).

I.2.1 SNPs

Os *Single Nucleotide Polymorphisms*, SNPs, são o tipo de variação genética mais comum no genoma humano, ocorrendo aproximadamente um a cada Kilobase (Wang *et al*, 1998). Os SNPs têm vantagens sobre outros marcadores moleculares, nomeadamente no facto de terem uma baixa taxa de mutação, aproximadamente 10^{-9} por geração (Satta *et al*, 1993) e de serem maioritariamente bialélicos, o que facilita a análise estatística e bioinformática. Os SNPs permanecem então na população sem sofrerem mutação recorrente muito mais tempo que os alelos dos *Short Tandem Repeats*, STRs, tendo potencial para revelar características muito ancestrais das populações.

I.2.2 STRs

Os STRs ou marcadores microssatélite como são vulgarmente designados, são sequências polimórficas de repetições até 6 pares de bases (Weber e Wong, 1993) que estão extremamente difundidos no genoma humano e podem estar localizados tanto em regiões codificantes como em regiões não codificantes (Tautz e Renz, 1984). No caso de STRs localizados em regiões não codificantes, geralmente não há pressão selectiva a ser exercida sobre o número de repetições. Há contudo excepções como no caso da Síndrome de X-Frágil em que a manifestação clínica da doença está relacionada com o número de repetições de uma sequência microssatélite [CGG] (Calabrrese e Sainudiin, 2005) localizada numa região não codificante, sobre a qual ocorre selecção.

Os STRs mutam a uma taxa aproximadamente de 10^{-3} por geração (Heyer, 1993), pelo mecanismo de *replication slippage*. Neste processo, é o deslizamento (*slippage*) da DNA polimerase que provoca uma alteração do número de repetições da sequência (Figura 3). Segundo este modelo, após o início da replicação há uma dissociação das cadeias de DNA e o posterior re-emparelhamento da nova cadeia com a cadeia molde. Dado o número de unidades repetitivas idênticas, este re-emparelhamento pode ocorrer forma errada levando a um aumento ou diminuição do número de repetições (Hellegren *et al*, 2004). Este mecanismo faz com que os STRs tenham um elevado grau de variabilidade.

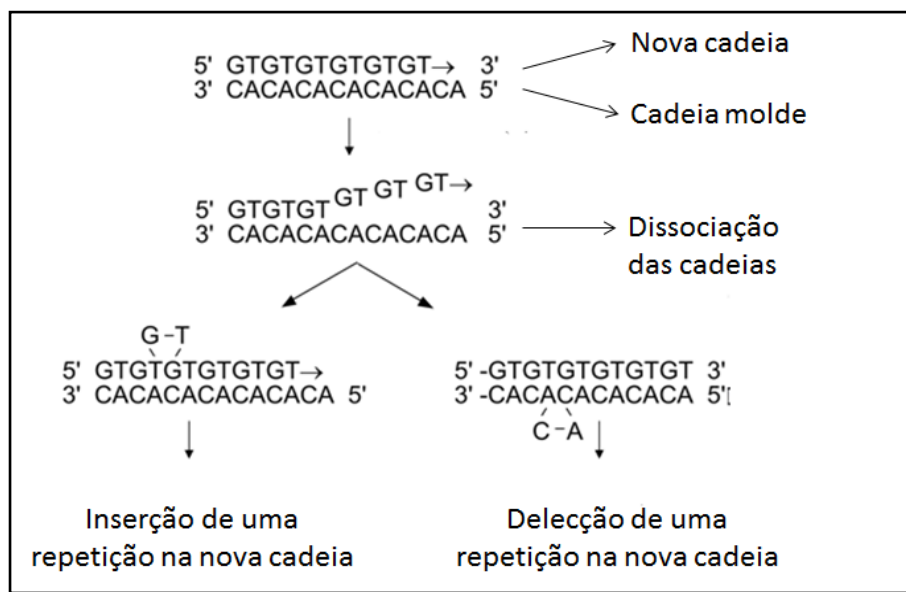


Figura 3 Aumento ou diminuição do número de repetições de marcadores STRs por *replication slippage*. (Adaptado de Hawk *et al*, 2005).

Existem vários modelos de evolução de STRs, o mais antigo é o *Stepwise Mutation Model* (Ota e Kimura, 1973). Neste modelo, as unidades de repetição têm uma probabilidade equivalente de aumentar ou diminuir, a uma taxa independente do número de repetições, sendo a variação sempre de uma unidade. Este modelo, no entanto, não tem em conta o tamanho do microssatélite. Mais tarde, surge um novo modelo denominado *Proportional Slippage Model*, que já tem em conta o tamanho da região repetitiva, sendo a taxa de mutação proporcional ao número de repetições do marcador STR (Kruglyak *et al*, 1998).

I.3 Análise haplotípica na Síndrome de X-Frágil – efeito fundador?

Segundo estudos anteriores em diversas populações, há evidências da existência de efeitos fundadores na Síndrome de X-Frágil. O exemplo mais flagrante é o da Finlândia onde 80% dos indivíduos com Síndrome de X-Frágil partilham o mesmo haplótipo, o qual apenas se observa em 8% dos indivíduos normais (Haataja *et al*, 1994). Estes resultados não são reprodutíveis em todas as populações estudadas, dado que na grande maioria é encontrado mais do que um possível perfil molecular fundador (Zhong *et al*, 1994).

II. Objectivos

As implicações genéticas e epigenéticas bem como os mecanismos que conduzem à expansão da região repetitiva [CGG] do gene *FMR1*, não estão completamente clarificados. As estratégias comumente utilizadas, para a clarificação destes mecanismos, incluem a determinação do número de tripletos repetitivos, do respectivo padrão de interrupções [AGG] e o estabelecimento de perfis genéticos associados ao *FMR1* recorrendo quer a STRs quer a SNPs em indivíduos normais e com Síndrome de X-Frágil. Paralelamente, a aplicação das mesmas metodologias em populações de controlo permite reconhecer factores envolvidos na estabilidade daquela região. A informação respeitante à população portuguesa é relativamente escassa uma vez que o único estudo existente envolveu apenas a análise de dois marcadores microssatélite (DXS548 e FRAXAC1) num total de 128 cromossomas X (normais e com mutação completa no gene *FMR1*) (Peixoto *et al*, 1998).

Neste trabalho, foram estudados 213 indivíduos de controlo e 123 com Síndrome de X-Frágil com o objectivo de (1) caracterizar a população portuguesa com Síndrome de X-Frágil através da construção de haplótipos por uma abordagem alargada que inclui 5 marcadores STRs e 3 SNPs intragénicos e (2) tentar identificar cromossoma(s) fundadore(s) para a Síndrome de X-Frágil em Portugal.

III. Materiais e Métodos

III.1 Características da amostra

Foram incluídas neste estudo amostras de DNA genómico (gDNA) de 213 indivíduos do sexo masculino em que foi excluída a presença de mutação completa (população Controlo) e 123 indivíduos do sexo masculino com Síndrome de X-Frágil (população FXS); ambas as *pools* de amostras foram anonimizadas, mantendo-se apenas a informação sobre a localização geográfica de origem de forma a perfazer amostras representativas de todo o País.

III.2 Recolha de Amostras

As amostras de gDNA dos doentes com Síndrome de X-Frágil, utilizadas neste trabalho, tinham dado entrada durante a última década, no Centro de Genética Médica Jacinto Magalhães (UMO-P CJM), tendo sido realizado o estudo molecular do gene *FMR1* por Southern blot em todas as amostras FXS em que foram identificados alelos com > 200 tripletos [CGG]. Os casos de mosaicismo (de tamanho e metilação) foram excluídos deste estudo.

As amostras (Controlo e FXS) utilizadas extraídas previamente para diagnóstico, pelo método *salting in- salting out* a partir de leucócitos de sangue periférico e posteriormente armazenadas, tendo-se recorrido a estas amostras para realização, no presente ano, deste estudo.

O gDNA foi quantificado por Nanodrop e a concentração ajustada a $300\text{ng}\cdot\mu\text{L}^{-1}$, $50\text{ng}\cdot\mu\text{L}^{-1}$ e $5\text{ng}\cdot\mu\text{L}^{-1}$ para aplicação das técnicas de PCR Multiplex e hrMCA, respectivamente.

III.3 Marcadores SNPs e STRs

Para a caracterização dos haplótipos, inicialmente, foram utilizados vários marcadores moleculares nomeadamente cinco *short tandem repeats* e três *single nucleotide polymorphisms*. Os cinco STR que flanqueiam a região polimórfica [CGG]: DXS8313, DXS548, DXS998, FRAXAC1 e FRAXAC2 foram escolhidos tendo em conta estudos anteriores que tiveram como objectivo a identificação de cromossomas fundadores para a Síndrome de X-Frágil.

A unidade de repetição de cada STR é o dinucleótido CA, à excepção dos marcadores DXS8313, em que a unidade de repetição é o trinucleotídeo [CCG] e FRAXAC2, que consiste numa sequência complexa com três regiões polimórficas concatenadas $(A)_n(TC)_n(CA)_n$ (Zhong *et al*, 1993).

Uma vez que os SNPs são *loci* com uma taxa de mutação inferior à dos STRs é expectável que a sua combinação possa contribuir para aumentar a informatividade da caracterização populacional. Contrariamente aos STRs os marcadores SNPs escolhidos nunca tinham sido estudados noutras populações com Síndrome de X-Frágil, não havendo referências para estudos comparativos. Os SNPs rs971000, rs29282 e rs25715 são intragénicos e descritos como tendo resultado de transição C↔T.

A localização dos oito marcadores moleculares seleccionados encontra-se na Figura 4, todos os marcadores moleculares estão em regiões não codificantes, os STRs DXS998, DXS548 e FRAXAC1 localizam-se a montante do gene enquanto que os restantes marcadores são intragénicos. O rs971000 localiza-se no intrão 1, o FRAXAC2 no intrão 2, o rs29282 no intrão 7 e o rs25715 no intrão 9. O DXS8313 está localizado a jusante do gene *FMR1*.

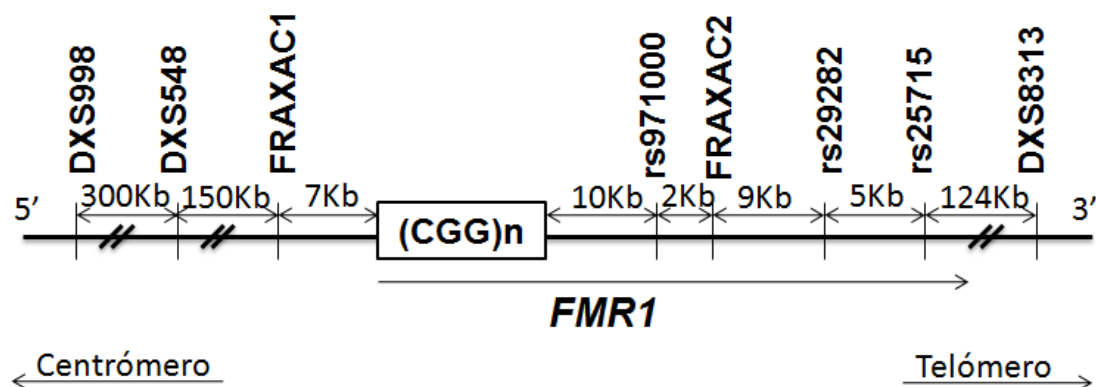


Figura 4 Localização dos marcadores moleculares relativamente à região polimórfica [CGG].

A posição nucleotídica de cada STR está de acordo com a sequência de referência AC_000066.1 do gene *FMR1* do GeneBank (<http://www.ncbi.nlm.nih.gov/RefSeq.html>), segundo a qual os marcadores STRs encontram-se nas seguintes posições:

- DXS998 → 146959992,
- DXS548 → 147155519,
- FRAXAC1 → 147338298,
- FRAXAC2 → 147357868,
- DXS8313 → 147480407.

e a posição nucleotídica de cada SNP está de acordo com a sequência de referência NT_011681.16:

- rs971000 → NT_011681.16:g.3444469A>G,
- rs29282 → NT_011681.16:g.3455280T>C,
- rs25715 → NT_011681.16:g.3459968T>C.

III.4 Desenho de primers

As sequências dos primers utilizados neste trabalho encontram-se no Anexo1.

Na amplificação de alguns STRs foram utilizados primers já descritos como o DXS998; DXS548 (Richard *et al*, 1991) e STR DXS8313 (Schlessinger, 1995 - Não publicado, <http://www.ncbi.nlm.nih.gov/genome/sts/>) Devido à existência de um polimorfismo no STR DXS548 foi necessário o recurso a um terceiro primer: DXS548 P3 (Chiurazzi *et al*, 1996).

Para amplificar a região flanqueadora de cada um dos SNPs e dos STRs FRAXAC1 e FRAXAC2 foram desenhados novos primers no programa *Primer express* (Applied Biosystems) e a estabilidade dos mesmos foi testada virtualmente no programa *Fast PCR* (Kalendar, Lee e Schulman, 2009).

O par de primers utilizado na determinação do número de tripletos [CGG] foi anteriormente desenhado para utilização na rotina laboratorial de diagnóstico.

III.5 Determinação número de tripletos [CGG]

No caso dos indivíduos da população Controlo foi necessário determinar o número de repetições [CGG] para confirmar que realmente são portadores de um alelo de tamanho normal. O número de tripletos [CGG] foi determinado por PCR convencional, no *veriti thermalcycler* segundo a metodologia de rotina usado no diagnóstico (Anexo 2). Posteriormente, no âmbito de outro estudo, determinou-se o padrão de interrupções [AGG] na mesma amostra Controlo que foi usada neste trabalho, pelo que não se mostram aqui os materiais e métodos utilizados.

III.6 Reacção de sequenciação

Para a reacção de sequenciação procedeu-se ao PCR assimétrico dos fragmentos a amplificar com BigDye® (mistura de PCR com enzima Taq polimerase e nucleotidos marcados e não marcados). Os fragmentos foram sujeitos a electroforese capilar no 3130xl Genetic Analyser (Applied Biosystems) e por fim a sequência analisada no Sequencing Analysis® (Applied Biosystems).

III.7 PCR Multiplex

Para determinação do tamanho dos alelos dos STRs, optimizou-se a técnica PCR Multiplex, de modo a amplificar os diversos STRs numa única reacção de PCR. Foi utilizado o termociclador *Veriti Thermal Cycler* e o programa e Mistura em anexo (Anexo 3). Para se poder aplicar esta técnica numa primeira fase foram desenhados *primers* para que os diferentes amplicões pudessem ser distinguidos pelo tamanho. Para facilitar a análise dos resultados um primer de cada par possui um fluorocromo ligado à extremidade 5' HexTM ou 6-FAMTM (Anexo1), permitindo que a fluorescência emitida por esta molécula seja detectada pelo sistema óptico do sequenciador durante a electroforese capilar. O conjunto dos diferentes tamanhos dos amplicões e cores associadas são observados no electroferograma dado pelo programa GeneMapper® v4. (Applied Biosystems) (Figura 5). Na determinação do tamanho dos amplicões foi usado o padrão de pesos moleculares ROX500 (Applied Biosystems). A partir do tamanho dos amplicões, foi inferido o número de repetições de cada marcador de cada amostra por comparação com uma amostra padrão sequenciada previamente.

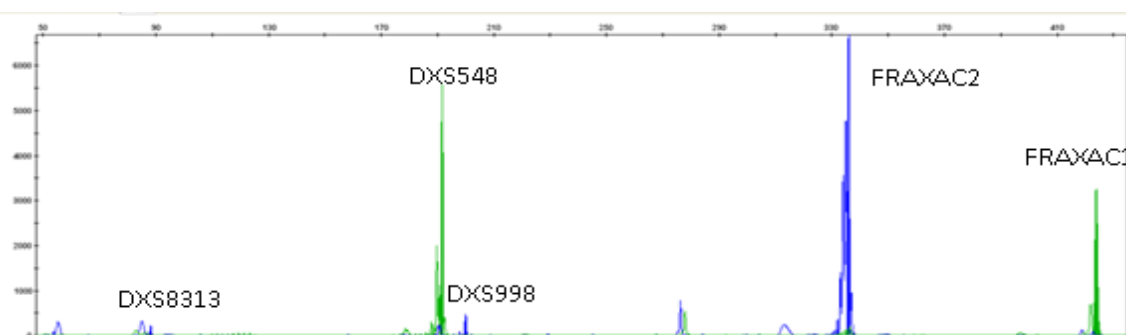


Figura 5 Electroferograma observado através do programa *Gene Mapper* dos produtos de PCR resultantes da amplificação dos 5 marcadores moleculares em Multiplex.

III.8 High Resolution Melting Curve Analysis - hrMCA

A técnica de hrMCA foi optimizada para determinar em cada indivíduo os alelos SNP. Esta técnica baseia-se na diferença de temperatura de *melting* para identificação dos diferentes SNPs, detectada fluorescência emitida durante a abertura das cadeias, pelo fluorocromo intercalante EvagreenTM (Biotium). A fluorescência é detectada pelo sistema óptico do rotorgene6000 (corbett life sciences) e a diferente temperatura de *melting*, traduz-se em curvas diferentes em que cada tipo de curva corresponde a um SNP (Figura 6).

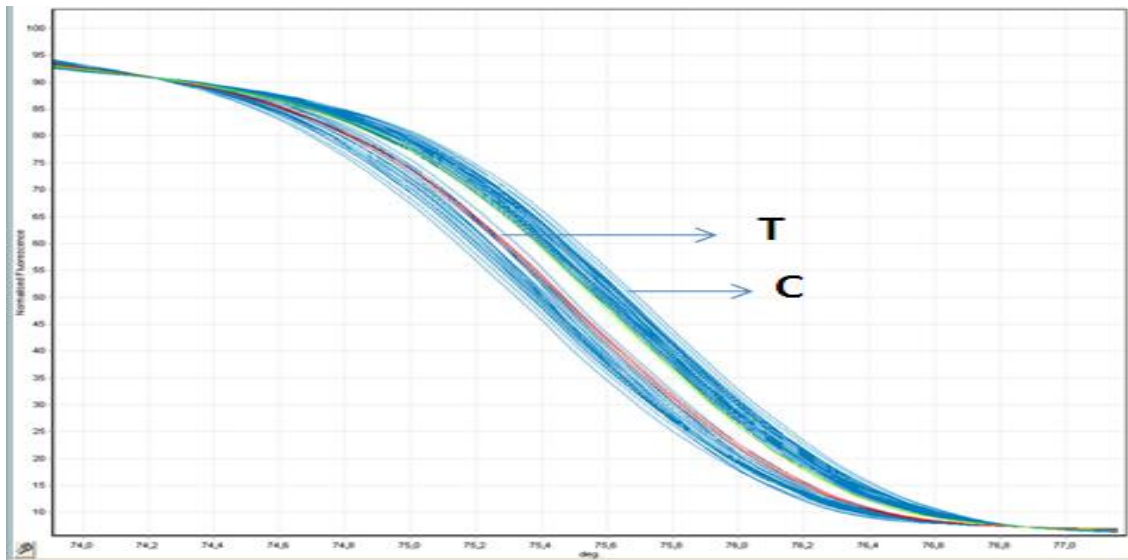


Figura 6 Curvas de *melting* observadas a partir do HRM do SNP rs25715. As curvas de *melting* distribuem-se por dois grupos e a cada um dos grupos corresponde um nucleótido cuja presença é responsável pela diferença de temperatura de *melting* que se pode observar.

Uma das desvantagens desta técnica é a sua grande sensibilidade à concentração de DNA e de reagentes entre os tubos de reacção de uma mesma corrida. Por este motivo, para validação dos resultados, foram utilizadas três réplicas para cada amostra. Em cada corrida utilizavam-se Controlos previamente sequenciados para assim se fazer corresponder as três curvas de *melting* resultantes de uma amostra a um SNP. As curvas de *melting* analisadas por normalização no programa associado ao rotorgene6000. As condições e programa necessários em cada corrida encontram-se no Anexo 4.

III.9 Análise Estatística

Para análise estatística dos resultados obtidos, recorreu-se ao Programa Arlequin ver3.5.1 (Excoffier e Lischer, 2010) com o intuito de determinar as frequências alélicas e haplotípicas em cada população, bem como os haplótipos partilhados. Aplicou-se o teste do Chi-quadrado na distribuição dos alelos em cada STR e na distribuição das Linhagens construídas pela conjugação dos SNPs.

Para calcular a diversidade génica de cada STR e SNP utilizou-se a seguinte fórmula (Nei, 1975):

$$\text{Diversidade génica} = 1 - \sum (\text{Frequência alélica relativa})^2$$

Após construção dos haplótipos, recorreu-se ainda à aplicação Network ver4.5 (<http://www.fluxus-engineering.com>) para inferir as relações genealógicas entre os haplótipos STR pertencentes à mesma Linhagem, segundo o método *Median joining*.

III.10 Construção de Linhagens e haplótipos

As Linhagens foram construídas pela conjugação dos três SNPs escolhidos, devido à sua baixa taxa de mutação e foram designadas de acordo com a localização de cada SNP (de 5' para 3') relativamente ao gene *FMR1*, ou seja, pela concatenação dos nucleótidos presentes nos SNPs rs971000, rs29282 e rs25715 respectivamente. Os haplótipos STR foram construídos da mesma forma, pela concatenação dos alelos presentes em cada STR de acordo com a sua posição de 5' para 3'.

A designação dos haplótipos completos foi atribuída pela junção da nomenclatura atribuída à Linhagem SNP com a nomenclatura dos haplótipos STR.

IV. Resultados

IV.1 Alelos encontrados nos marcadores STR e respectiva nomenclatura

Os alelos dos marcadores DXS8313, DXS998, DXS548, FRAXAC1 e FRAXAC2 foram estudados em amostras de gDNA das populações Controlo e com Síndrome de X-Frágil (FXS). O DXS8313, apenas tinha sido referido na literatura num estudo de deleções na região Xq-27-28 ligadas ao cancro da próstata (Kibel *et al*, 2003), nunca tendo sido associado ao gene *FMR1*. Durante a caracterização dos diferentes STRs, o marcador DXS8313, não se mostrou polimórfico nas amostras utilizadas durante a optimização da metodologia, pelo que não foi utilizado na construção dos haplótipos.

A partir do tamanho de cada alelo inferiu-se o número de repetições correspondente a cada marcador por comparação entre o tamanho do amplicão (pares de bases) e uma amostra sequenciada previamente. Uma vez determinado o número de repetições dos alelos, foi-lhes atribuída a nomenclatura sugerida por Chiurazzi, 1999. Esta nomenclatura foi proposta com o objectivo de padronizar a designação dos haplótipos construídos com STRs localizados no *locus* do gene *FMR1*. Segundo esta nomenclatura, cada alelo designa-se pelo tamanho, em pares de bases, da sua região repetitiva.

Os alelos observados nas duas populações estudadas, bem como o número de repetições correspondentes encontram-se na Tabela 1.

Uma vez que o STR FRAXAC2 é constituído por três regiões polimórficas contínuas ($A_nTC_nCA_n$), os alelos deste marcador apenas são designados pelo tamanho dos fragmentos resultantes (pb) da reacção de PCR.

Nos alelos STR da população Portuguesa não foi encontrado nenhum alelo “novo” quando se comparam os alelos obtidos com os que foram observados noutras populações.

Síndrome de X-Frágil em Portugal

DXS998		DXS549		FRAXAC1		FRAXAC2
Alelo	Repetições	Alelo	Repetições	Alelo	Repetições	Alelo (pb)
32	16	38	19	34	17	329
34	17	40	20	36	18	330
36	18	42	21	38	19	331
38	19	44	22	40	20	332
40	20	46	23	42	21	333
42	21	48	24	44	22	334
44	22	50	25	46	24	335
46	23	52	26			336
48	24	54	27			337
		56	28			338
						339
						340
						341
						342

Tabela 1 Nomenclatura atribuída aos alelos encontrados em cada marcador. No DXS998, DXS548 e FRAXAC1 é feita a correspondência entre a nomenclatura do alelo e o número de repetições CA correspondentes.

IV.2 Distribuição dos Alelos STR na população Controlo e X-Frágil

Após determinação dos alelos STR presentes em cada amostra, obtiveram-se as seguintes frequências absolutas (Tabela 2):

DXS998			DXS548			FRAXAC1			FRAXAC2		
Alelo	Frequência Absoluta		Alelo	Frequência Absoluta		Alelo	Frequência Absoluta		Alelo	Frequência Absoluta	
	Controlo	FXS		Controlo	FXS		Controlo	FXS		Controlo	FXS
32	31	10	38	20	0	34	3	0	329	1	0
34	119	0	40	143	0	36	26	44	330	0	0
36	43	88	42	23	26	38	164	39	331	4	0
38	17	2	44	2	24	40	3	1	332	5	7
40	0	16	46	1	22	42	17	33	333	7	14
42	1	0	48	5	5	44	0	5	334	9	10
44	2	5	50	12	1	46	0	1	335	37	35
46	0	0	52	6	43	Total	213	123	336	116	29
48	0	2	54	0	1				337	16	23
Total	213	123	56	1	1				338	14	2
			Total	213	123				339	0	3
									340	1	0
									341	2	0
									342	1	0
									Total	213	123

Tabela 2 Frequências alélicas absolutas observadas nas populações Controlo e FXS para os quatro STRs estudados.

Os alelos dos marcadores STRs estudados revelaram uma distribuição diferente entre a população Controlo e a população X-Frágil (Figura 7).

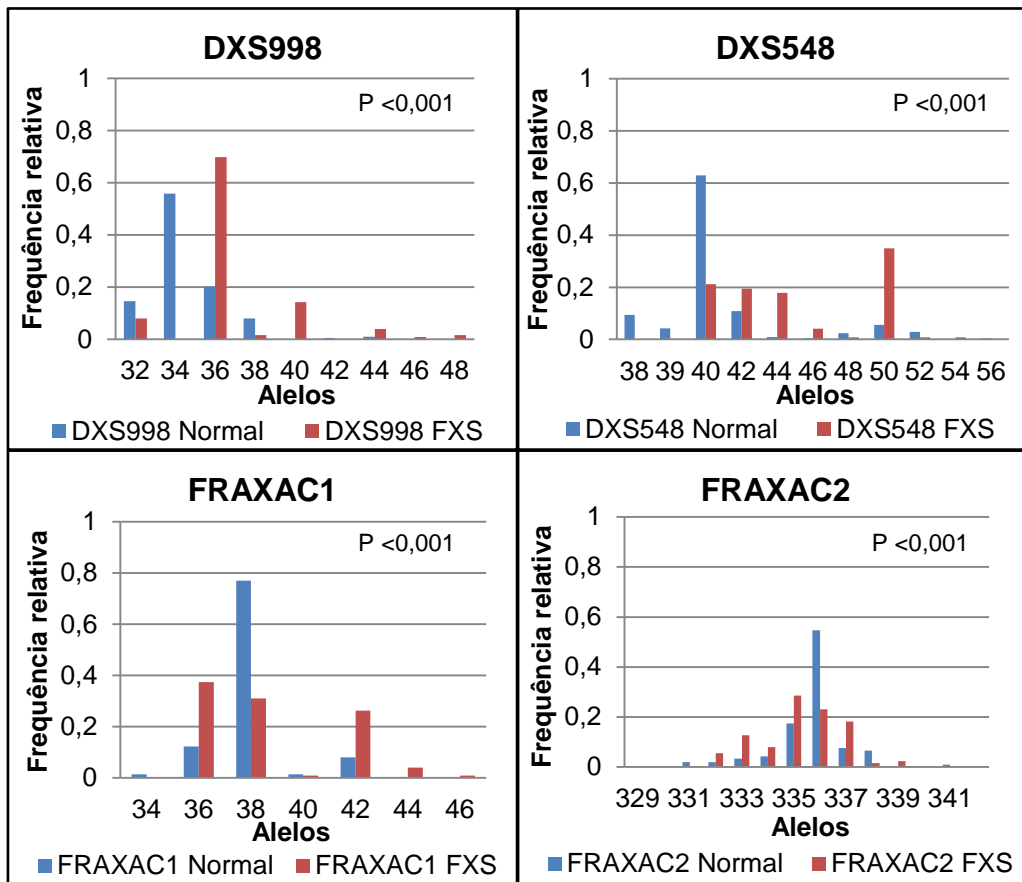


Figura 7 Distribuição dos alelos de cada marcador STR na população Controlo e população X-Frágil

Foi observada na população Controlo uma distribuição unimodal nos STRs DXS998 e FRAXAC2 e uma distribuição bimodal nos STRs DXS548 e FRAXAC1, o mesmo também acontece na população FXS nos STR DXS998 e FRAXAC2. Nos alelos dos marcadores DXS548 e FRAXAC1 da população FXS apresentam uma distribuição pouco característica, em que são observados vários alelos modais.

A frequência da maioria dos alelos dos vários STRs em estudo, quando estudados individualmente, também é significativamente diferente entre as populações (Tabela 3).

Síndrome de X-Frágil em Portugal

DXS998		DXS548		FRAXAC1		FRAXAC2	
Alelo	Valor p	Alelo	Valor p	Alelo	Valor p	Alelo	Valor p
32	0,1195	38	0,0005	34	0,1809	329	0,4437
34	0,0000	40	0,0000	36	0,0000	330	-
36	0,0000	42	0,0097	38	0,0000	331	0,1239
38	0,0201	44	0,0000	40	0,6124	332	0,2176
40	0,0000	46	0,0000	42	0,0000	333	0,0008
42	0,4598	48	0,3721	44	0,0034	334	0,1436
44	0,0429	50	0,0273	46	0,1929	335	0,0271
46	0,1747	52	0,0000			336	0,0000
48	0,0546	54	0,1875			337	0,0080
		56	0,6933			338	0,0382
						339	0,0229
						340	0,4437
						341	0,2780
						342	0,4437

Tabela 3 Teste do Chi-quadrado que mostra a significante diferente distribuição dos alelos encontrados para cada marcador molecular entre a população Controlo e X-Frágil.

Quando se compara a diversidade génica de cada STR entre as duas populações (Tabela 4), verificamos que à excepção do DXS998, todos os STR são mais variáveis na população FXS.

	População Controlo	População FXS
DXS998	0,619	0,483
DXS548	0,577	0,761
FRAXAC1	0,386	0,695
FRAXAC2	0,656	0,806

Tabela 4 Diversidade génica de cada marcador STR na população Controlo e FXS

Tal como em estudos anteriores, o STR FRAXAC2 apresenta uma maior diversidade génica relativamente aos restantes STR em estudo e o FRAXAC1 é o marcador que apresenta menor diversidade génica. Estes resultados revelam uma grande heterogeneidade alelica na população FXS.

Em todos os marcadores, pode-se observar que a população Controlo é homogénea, sendo sempre representada por um alelo modal, contrariamente a população FXS é bastante heterogénea quanto à distribuição dos alelos. Nesta população, nos marcadores DXS548 e FRAXAC1, é visível uma grande variabilidade que se pode ver na ampla distribuição dos alelos ao invés sobressair um alelo pela sua elevada frequência.

IV.3 Distribuição dos alelos SNP na população Controlo e X-Frágil

Foram analisados os três SNPs: rs971000, rs29282 e rs25725 nas amostras das populações Controlo e FXS (Tabela 5). Todos os SNPs seleccionados para análise apresentam alelos C ou T e localizam-se em regiões intrónicas do gene *FMR1*. Não foram observadas mutações/alterações na sequência dos amplicões dos SNPs em nenhuma das amostras estudadas.

Alelo	rs971000		rs29282		rs25715	
	Frequência Absoluta		Frequência Absoluta		Frequência Absoluta	
	Controlo	FXS	Controlo	FXS	Controlo	FXS
C	77	104	17	28	83	107
T	136	19	196	95	130	16
Total	213	123	213	123	213	123

Tabela 5 Frequências absolutas dos alelos SNPs na população Controlo e FXS

A frequência e distribuição dos SNPs nas duas populações observa-se na Figura 8.

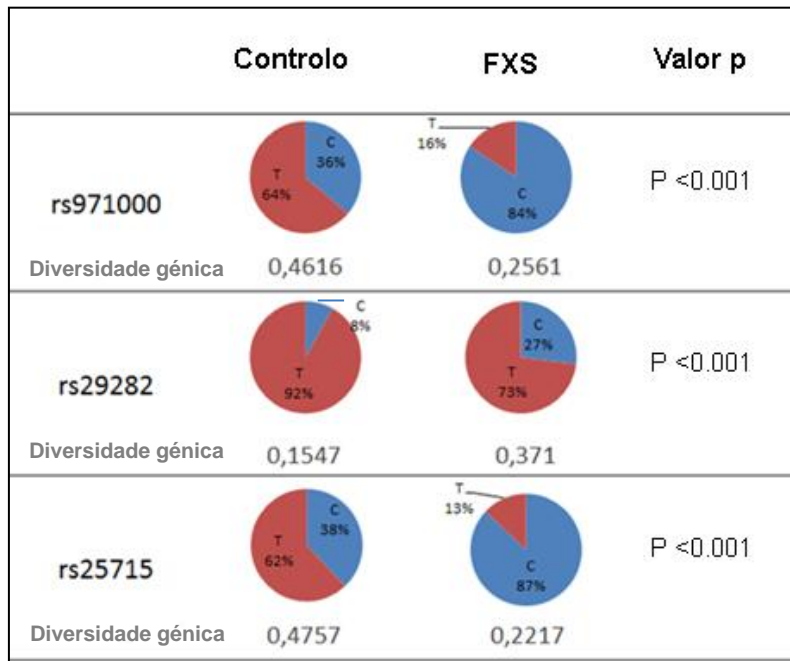


Figura 8 Distribuição dos SNPs rs971000, rs29282 e rs25715 nas Populações Controlo e X-Frágil, diversidade génica e respectivo valor p < 0.001.

As frequências observadas na população portuguesa Controlo estão de acordo com a informação disponível em refSNP do NCBI para a população Europeia do HapMap-CEU (<http://www.ncbi.nlm.nih.gov/projects/SNP/>)

Na população Portuguesa dos três SNPs estudados, só o rs29282 é que se mostrou pouco variável, contudo na população FXS a sua variabilidade aumenta. Nos outros dois SNPs estudados verifica-se um nucleótido bastante representado pela população com Síndrome de X-Frágil, apesar da baixa variabilidade nesta população, podemos verificar que há uma associação do nucleótido C dos SNPs rs971000 e rs25715 à Síndrome de X-Frágil ($p < 0,001$).

VI.4 Linhagens

Os SNPs têm uma baixa taxa de mutação comparativamente com os STRs, sendo então mais estáveis no tempo, assim, a partir da conjugação dos três SNPs, pretende-se aumentar a escala de tempo e obter informações sobre a ancestralidade da expansão da região repetitiva [CGG] e respectivo fundo genético.

Após construção das Linhagens SNPs, foi-lhes atribuída uma nomenclatura específica (Tabela 6). Verifica-se que as duas populações em estudo partilham a mesma diversidade, isto é, as Linhagens que se observam na população FXS, também estão presentes na população Controlo, embora com uma distribuição diferente (Figura 9).

Linhagem	Nomenclatura
TTT	A
CTC	B
CCC	C
CTT	D
TTC	E
TCC	F

Tabela 6 Nomenclatura atribuída a cada Linhagem SNP

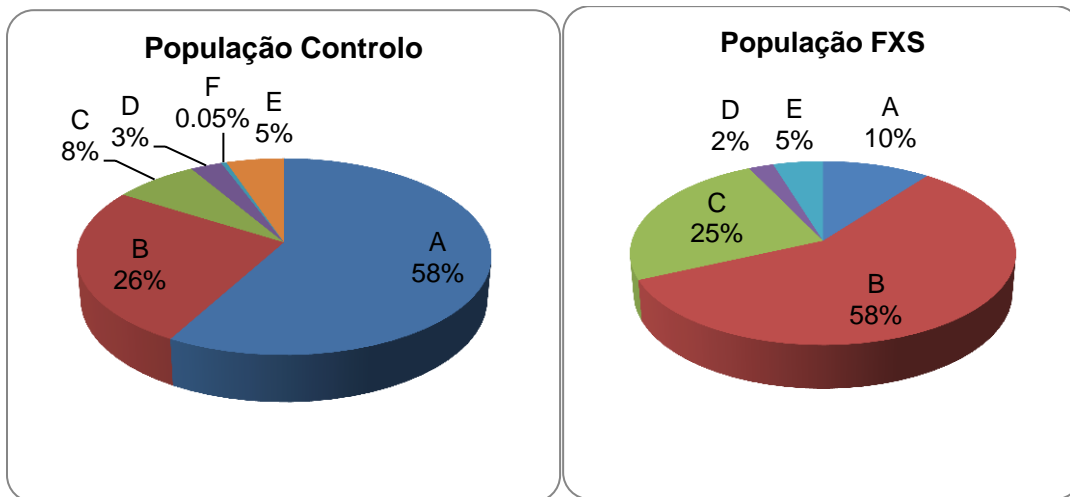


Figura 9 Frequência das Linhagens SNPs nas populações Controlo e X-Frágil.

Das seis Linhagens observadas na população Controlo, a mais frequente é a A representada por 58% dos indivíduos. Seguida da B e C com 26% e 8% respectivamente. A Linhagem F é considerada rara uma vez que só apareceu na população Controlo com uma frequência de 0.05% (apenas um indivíduo da amostra).

Na população X-Frágil, foram observadas cinco Linhagens. A Linhagem mais frequente nesta população é a B presente em 58% dos indivíduos, seguida da C com 25% dos indivíduos e a Linhagem A com apenas 10%.

Quando se compara a distribuição das Linhagens entre as duas populações verifica-se que esta é significativamente diferente ($p < 0,001$).

Já a diversidade génica das Linhagens nas duas populações, é igual, 0,59, ou seja, embora a distribuição das Linhagens seja diferente, as duas populações são iguais no que diz respeito à sua variabilidade.

IV.5 Haplótipos

Após construção dos haplótipos STR-SNP, observam-se 92 haplótipos na população Controlo (Anexo 5) e 62 haplótipos na população FXS (Anexo 6). Os haplótipos são designados pela junção da letra correspondente à Linhagem a que pertencem (Tabela 6) e de um número que é atribuído a cada haplótipo STR dentro de cada Linhagem (Anexos 5 e 6). Destes haplótipos apenas 8 são partilhados pelas duas populações (Tabela 7) embora que por uma baixa frequência de indivíduos.

Haplótipos	Controlo	FXS
A2	0,0751	0,0244
A3	0,0516	0,0081
B26	0,0047	0,0488
B11	0,0047	0,0325
B10	0,0094	0,0244
B9	0,0094	0,0081
C6	0,0047	0,0244
C4	0,0081	0,0081

Tabela 7 Haplótipos partilhados pelas duas populações e respectivas frequências relativas

O haplótipo mais frequente na população Controlo é o A1, representado por 17% dos indivíduos normais, seguindo-se o A2 (7.5%), o A3 (5.2%) e o B1 (5.2%). Na população FXS, o haplótipo mais frequente é o B32, presente em 9% dos indivíduos estudados.

Na tentativa de analisar as relações genealógicas entre os haplótipos encontrados nas duas populações, foram desenhadas redes de haplótipos individualmente para cada Linhagem (Figura 10). Isto porque as Linhagens SNP permitem “mergulhar” mais profundamente no tempo, dando informações sobre divergências mais ancestrais. O que permite avaliar os haplótipo STR que divergiram de um ancestral (Linhagem) comum. Nas redes construídas, os haplótipos STR estão representados por círculos, cujos tamanhos estão de acordo com a sua frequência na população, nestas redes todas as relações possíveis entre haplótipos adjacentes estão representadas por linhas cujo tamanho está de acordo com o número de mutações ocorridas.

Não foi desenhada uma rede de haplótipos para a Linhagem F, uma vez que apenas se encontrou um haplótipo com este fundo genético.

A Linhagem A é a mais representada na população Controlo do nosso País. Na sua rede, é possível observar, um haplótipo central bastante frequente na população Controlo, a partir do qual surgiram vários haplótipos que se dispõem em estrela, por mutação em apenas um STR. Nas redes construídas a partir das restantes Linhagens, não é possível observar esta estrutura característica.

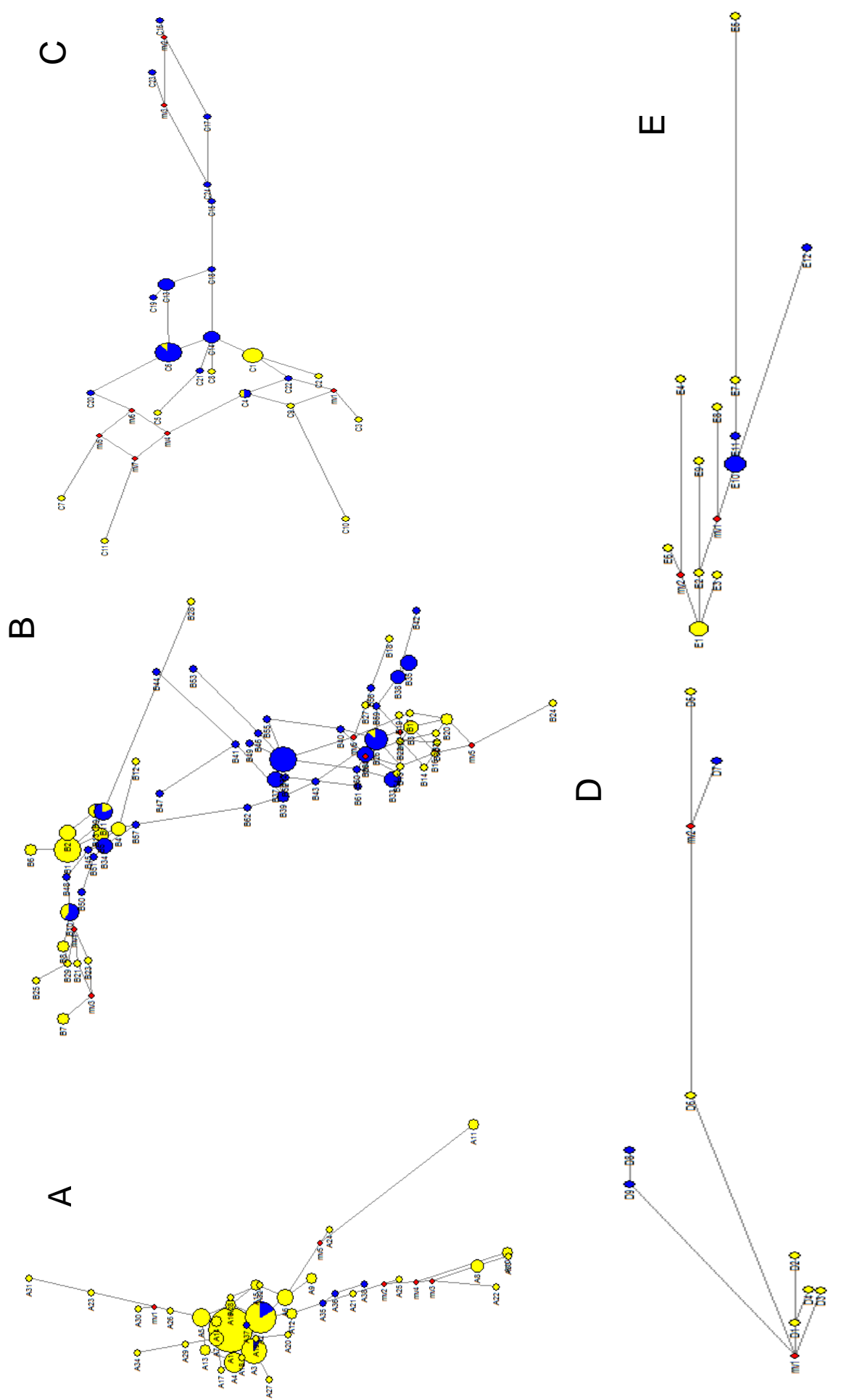


Figura 10 Redes de haplótipos desenhadas para as Linhagens A, B, C, D e E. A amarelo encontram-se os haplótipos da população Controle e a azul os haplótipos da população FXS.

IV.6 Quantificação do número de repetições [CGG] na população Controlo

Nos 213 indivíduos que da população Controlo foi feita a determinação do número de repetições [CGG] cuja frequência dos alelos [CGG] pode ser observada no Gráfico 1.

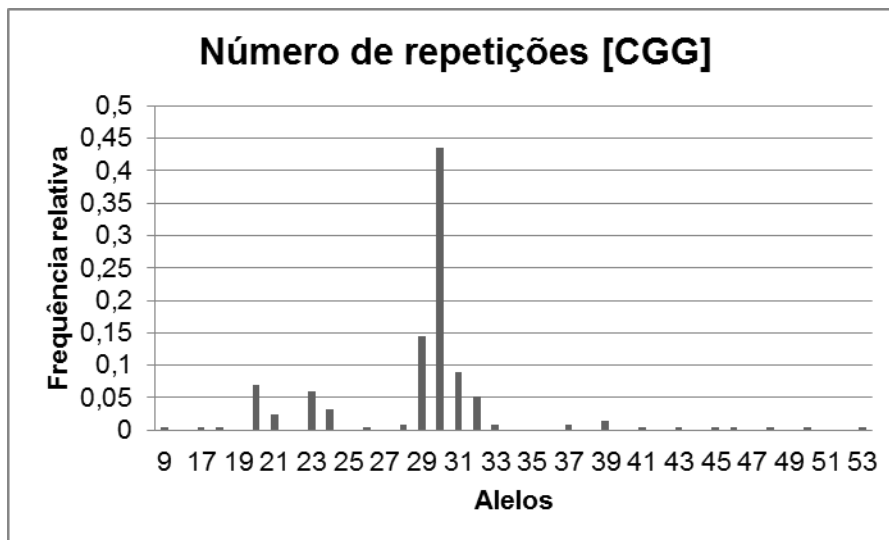


Gráfico 1 distribuição dos alelos [CGG] na população Controlo

Na população Portuguesa com alelos de tamanho considerado normal, os alelos variam entre um mínimo de 9 repetições [CGG] e um máximo de 53 repetições. É de notar que o alelo com 9 repetições apenas estava presente num único indivíduo, sendo um alelo raro, e entre este e o alelo com 17 repetições não foi encontrado nenhum alelo na amostra estudada. O alelo mais frequente contém 30 repetições [CGG], está presente em 43% da população e está fortemente associado aos haplótipos: A1, A2 e A3 com 60% 79% e 100% dos indivíduos respectivamente. Estes haplótipos são os mais frequentes na população Controlo e diferem apenas no alelo do marcador DXS998 aquele que se encontra mais afastado do número de repetições [CGG].

Na população X-Frágil, não foi estabelecida uma relação entre o número de repetições [CGG] e os respectivos haplótipos uma vez que a mutação completa é altamente instável, o que em células de sangue periférico de indivíduos com Síndrome de X-Frágil se traduz num número muito elevado e variável de repetições ($200 < \text{CGG} < 750$).

IV.7 Polimorfismos nos amplicões STRs

Dentro das regiões amplificadas pelos primers dos marcadores DXS548 e FRAXAC1 foram identificados outros polimorfismos para além da região CA repetitiva.

No caso do STR DXS548, foi encontrado um polimorfismo localizado 86 pb a montante da região polimórfica repetitiva CA: $(C)_4(G)_n$. Este polimorfismo já estava descrito no

refSNP do NCBI, com valores de n de 9 e 10, designando-se por rs113191334 (<http://www.ncbi.nlm.nih.gov/projects/SNP/>). No entanto na população Portuguesa o n pode assumir valores entre 9 e 11, esta variabilidade de n já tinha sido observado na população Caucasiana Americana (Crawford *et al*, 2000). Este polimorfismo quase nunca é referido em publicações onde o marcador DXS548 é utilizado, estes estudos, apontam para a existência de alelos intermédios, em que o número de repetições não é um número inteiro. Uma vez que o objectivo deste trabalho, era a correcta caracterização dos alelos do STR optou-se pelo uso de um terceiro primer que não abrange o rs113191334.

No marcador FRAXAC1, observou-se uma duplicação de 7pb localizada 100pb a montante da região repetitiva CA. Está presente tanto na população Controlo (13% dos cromossomas estudados) como na população X-Frágil (18% dos cromossomas), o que foi provado por sequenciação de um subgrupo de 15 amostras, onde se pode ver um exemplo na Figura 11 e os restantes portadores da duplicação foram inferidos pelo tamanho dos produtos de PCR.

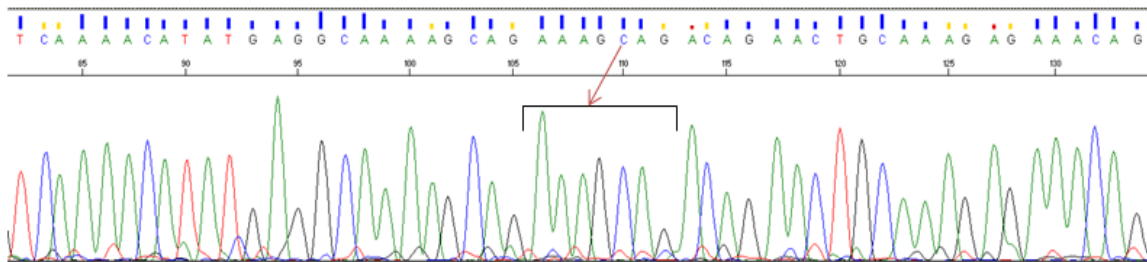


Figura 11 Electroferograma onde se observa a sequenciação de um amplicão resultante da amplificação do marcador FRAXAC1 onde está presente a duplicação de 7pb AAAGCA.

Esta duplicação não estava descrita, tendo sido submetida na base de dados dbSNP do NCBI, à qual foi atribuída a designação temporária de ss263192808, sendo a designação definitiva, publicada nesta base de dados, na próxima actualização em Janeiro-Março de 2011.

A presença do ss263192808 está fortemente associada à Linhagem B, dado que 100% dos indivíduos X-Frágil portadores desta duplicação pertencem à Linhagem B e 79% dos indivíduos Controlo portadores da duplicação também pertencem a esta Linhagem. O que sugere que a duplicação de 7pb seja recente e tenha surgido na linhagem B.

V. Discussão de Resultados

Desde o início da década de 90 que vários grupos de investigação tentam encontrar um cromossoma fundador para a mutação causadora da Síndrome de X-Frágil, baseando-se na significativa diferença na distribuição de haplótipos construídos com STRs, na população Controlo e X-Frágil. Nesses trabalhos, foram essencialmente utilizados dois marcadores moleculares, o DXS548 e o FRAXAC1 e em alguns casos foram também introduzido o marcador FRAXAC2 ou o DXS998.

A Finlândia foi o único país onde foi encontrado um único cromossoma fundador (Haataja *et al*, 1994). O aparecimento de um haplótipo fundador nesta região, representado por 80% dos indivíduos afectados, pode-se justificar aparentemente pela própria história e estrutura populacional. Na história do País existiu um pequeno número de colonos e posterior isolamento das populações, o que leva a uma deriva genética que é favorável ao aparecimento de efeitos fundadores (Chapelle, 1993). Nos restantes estudos, em cada país é observado mais do que um cromossoma fundador representados por uma baixa percentagem na população.

A heterogeneidade dos haplótipos STR é observada noutros estudos efectuados em várias populações com Síndrome de X-Frágil até à data (Oudet *et al*, 1993; Zhong *et al*, 1994; Chiurazzi *et al*, 1996; Enis, Murray e Morton 2001, Arrieta *et al*, 2003), não tem qualquer sustentação teórica. Na tentativa de aprofundar estes estudos, recentemente tem-se vindo a alterar o tipo de abordagem a este tema introduzindo polimorfismos mais estáveis como os SNPs. O presente estudo representa então a primeira análise de haplótipos STR-SNP na população Portuguesa. Neste trabalho, numa primeira fase de análise procedeu-se à divisão populacional por Linhagens e posteriormente à distribuição dos haplótipos STR observados nas duas populações de acordo com a Linhagem correspondente.

V.1 Linhagens SNP

A definição de Linhagens na População Portuguesa, pela conjugação de três SNPs, mostrou que a Síndrome de X-Frágil está representada em cinco dos seis fundos genéticos observados na população Controlo, em vez de estar apenas fortemente associada a uma das Linhagens e/ou apresentar uma conjugação de SNPs diferente da população Controlo. Esta distribuição da mutação completa pelas respectivas Linhagens sugere que ao longo do tempo houve uma múltipla fundação.

V.2 Haplótipos STR

Tal como noutras populações, a população Portuguesa com Síndrome de X-Frágil, é dotada de uma grande heterogeneidade nos marcadores STRs tanto individualmente como em haplótipo. Quando se compararam os haplótipos STRs da população com Síndrome de X-Frágil em Portugal com as populações X-Frágéis de outros países, não se encontram semelhanças significativas tal como acontece quando se comparam os haplótipos desses países entre eles. De facto, os alelos STR que flanqueiam a região repetitiva [CGG], são mais variáveis na população com FXS do que na população Controlo, havendo uma correlação entre a mutação completa e a diversidade nos *loci* cromossómicos adjacentes. Para além disso, os haplótipos FXS não aparentam ter uma relação genealógica próxima como era de esperar. Isto sugere que haja um mecanismo celular responsável pela instabilidade dos STR tanto na região [CGG] repetitiva, como dos *loci* adjacentes. A instabilidade que é observada na região repetitiva [CGG] que se estende aos *loci* vizinhos pode ser devida à existência de um mecanismo ainda não identificado responsável por gerar esta instabilidade (Ennis *et al*, 2001). Este mecanismo poderá ser um factor *trans* ou *cis*, tal como uma DNA polimerase, a alteração nos mecanismos de reparação do DNA ou uma diferente conformação de DNA entre outros.

Contrariamente ao que se observa na população com Síndrome de X-Frágil, a população Controlo, mostrou-se homogénea, sendo o haplótipo mais frequente o A1 representado por 17% dos indivíduos. Quando comparamos a nossa população Controlo com a utilizada nos outros estudos (Itália, França, País Basco, Estados Unidos, África Sub-sariana), apenas usando haplótipos STR, verificamos que todas elas são muito semelhantes e homogéneas. Na população Portuguesa, este haplótipo (A1) nunca apareceu na população FXS e está fortemente associado a alelos com 30 repetições [CGG] e com um padrão de interrupções AGG(CGG)₁₀AGG(CGG)₉AGG(CGG)₉. Este padrão é aquele que se considera como o mais estável, evitando os fenómenos de *slippage* e conseqüente instabilidade alélica devido à expansão do número de tripletos [CGG] (Oliveira *et al*, comunicação em forma de poster, SPGH 2010)

V.3 Distribuição dos haplótipos STR pelas Linhagens SNP

Após a construção de redes haplotípicas para cada Linhagem, seria de esperar a observação de uma forte relação genealógica entre os haplótipos FXS encontrados, ou seja, um haplótipo central bastante frequente a partir do qual tiveram origem vários haplótipos devido à mutação num STR.

Na Linhagem A, representada por mais de metade da população Controlo, os haplótipos associados a alelos com tamanho normal, tem uma estrutura característica.

Nesta Linhagem, existe um haplótipo central que corresponde ao mais frequente na população e teoricamente o mais ancestral, a partir do qual divergem vários haplótipos, que resultaram de uma mutação num STR. Ao nível da população FXS, a Linhagem A quando comparada com as restantes, mostra-se refractária, para além da baixa incidência de FXS (10% da população FXS), não mostra uma múltipla origem da mutação como as restantes Linhagens, os haplótipos FXS estão apenas num ramo.

As Linhagens C,D e E, mostram uma evidente distância genética entre os haplótipos FXS, no entanto estão pouco representadas em ambas as populações.

Tendo em conta o variado aspecto das redes haplotípicas, a Linhagem B é aquela que se mostra mais complexa e bastante representada nas duas populações ($\frac{1}{4}$ da população Controlo e mais de metade da população FXS). O facto desta Linhagem se dividir em dois grupos, aponta para a existência de duas origens, talvez por recombinação entre os *loci* STR ou por mutação recorrente. A elevada frequência populacional e complexidade na distribuição haplotípica desta Linhagem indica que provavelmente que a Linhagem B será a mais ancestral.

V.4 A controversa origem da mutação completa

A partir dos resultados obtidos e respectiva análise, não é possível identificar um cromossoma fundador. Os haplótipos STR com mutação completa têm o mesmo fundo genético (SNP) que os haplótipos STR normais, a duplicação de 7pb encontrada neste trabalho está sempre associada à Linhagem B e ao 38 do marcador FRAXAC1, apontando para que esta duplicação seja recente, encontra-se presente tanto em indivíduos normais como indivíduos FXS. Para além disto, não há qualquer relação genealógica entre os haplótipos STRs encontrados na população FXS. Mais do que a não existência de um efeito fundador os resultados permitem pensar que apesar de nunca se terem observado mutações *de novo* em estudos familiares (Smits *et al*, 1992), estas possam ter surgido várias vezes na população. As várias migrações a que a Península Ibérica foi sujeita ao longo da História, podem ter trazido novos cromossomas com mutação completa, o que levou a um aumento da variabilidade genética da população, no entanto, não explica a heterogeneidade populacional observada.

A Síndrome de X-Frágil possui duas propriedades contraditórias: Apesar dos indivíduos portadores de um alelo mutado terem uma baixa taxa de reprodução, a Síndrome de X-Frágil tem uma elevada prevalência na população. Dadas estas duas controvérsias, tiveram de surgir várias mutações completas ao longo do tempo para explicar a elevada presença da mutação na população. Segundo o *multistep model* de Morton e Macpherson (1992), um alelo mutado S (> 50 repetições), pode ser originado a

partir de um alelo estável N. Os alelos S, podem ter uma elevada prevalência na população e podem ser mantidos na população durante 90 gerações. Estes alelos, S, podem originar alelos pré-mutados e instáveis (Z) a partir dos quais podem surgir mutações completas, (L) com uma taxa elevada na linha germinal feminina. De acordo com este modelo, um alelo S, pode gerar um número elevado de mutações completas, pelo que parece plausível a existência de vários cromossomas fundadores, no entanto, na população Portuguesa não é possível identifica-los uma vez que os haplótipos não têm uma clara relação genealógica entre eles.

A região repetitiva [CGG], é uma região muito dinâmica e após a perda das interrupções [AGG], a probabilidade de ocorrer aumento do número de repetições do alelo aumenta, no entanto, segundo o *stepwise mutation model*, este aumento/diminuição deveria ser de uma unidade de repetição, que não se observa. Os alelos Z pela perda de [AGG], tornam-se altamente instáveis e podem aumentar/diminuir bastante o número de repetições numa única meiose, evidenciando a grande dinâmica da mutação. Pelos resultados obtidos, também é possível verificar que nos STRs associados à Síndrome, o *stepwise mutation model* também não se aplica.

VI. Conclusões e Perspectivas futuras

A partir da caracterização da População Portuguesa com SNPs e STRs que flanqueiam a região [CGG] repetitiva do gene *FMR1*, podemos verificar a existência de uma população Controlada homogénea e uma população com Síndrome de X-Frágil bastante heterogénea. Com a nova abordagem que envolve a inclusão de SNPs na análise haplotípica, foi possível verificar que não existe um haplótipo fundador para a Síndrome de X-Frágil em Portugal, o que pode ser explicado pela dinâmica da região repetitiva [CGG], que leva ao aparecimento de mutações *de novo*, o que consequentemente pode levar a heterogeneidade haplotípicas nesta população. Para além disso, não é observável uma relação genealógica entre os haplótipos da população X-Frágil, o que pode ser explicado pela existência de um mecanismo celular que gera instabilidade tanto ao nível da região [CGG] repetitiva como dos STRs franqueadores, este mecanismo pode também estar a “mascarar” haplótipos fundadores dentro de cada Linhagem.

Ainda existem muitas dúvidas a esclarecer no que diz respeito aos haplótipos associados à Síndrome de X-Frágil e à origem da mutação. Será que a heterogeneidade alélica é “herdada”, os haplótipos STR variam em cada geração? E então, de que forma é que a instabilidade dos alelos é transmitida de geração em geração? Em primeiro lugar, é necessário fazer estudos familiares que englobem mães preferencialmente com mais de um filho com Síndrome de X-Frágil, para se estudar de que forma é que a instabilidade associada aos alelos STR é transmitida de uma geração para a geração seguinte e posteriormente em que fase é que surge esta instabilidade, se na meiose ou na fase zigótica. Será também interessante estudar os padrões e número de interrupções [AGG] nos indivíduos com Síndrome de X-Frágil, dado que são um factor importante na instabilidade da região repetitiva [CGG].

Com o objectivo de caracterizar outros factores (*cis/trans*) implicados na instabilidade desta região, seria importante caracterizar o haplótipo de autossomas, que numa situação normal são segregados de modo independente do cromossoma X. Será também importante estudar marcadores STR no DNA mitocondrial em pares mãe-filho em que ambos são portadores de um alelo instável. Este tipo de DNA apenas é transmitido pela mãe, permitindo inferir se a instabilidade associada ao *locus FMR1*, se deve a um factor *trans*, e então se poder avançar no sentido de identificar o factor envolvido na instabilidade.

VII. Referências

VII.1 Referências Bibliográficas

- Arrieta I, Peñagarikano O, Téllez M, Ortega B, Flores P, Criado B, Veiga I, Peixoto AL, Lostao CM (2003), *The FMR1 CGG repeat and linked microsatellite markers in two Basque valleys*, *Heredity* 90, 206-211.
- Bear MF, Huber KM, Warren ST (2004), *The mGluR theory of fragile X mental retardation*. *Trends Neurosci* 27, 370–377.
- Calabrese P, Sainudiin R (2005), *Models of Microsatellite evolution* In, *Statistics for Biology and Health*, Part III, Springer (Ed), 290-305.
- Chapelle A (1993), *Disease gene mapping in isolated human populations, the example of Finland*, *J Med Genet* 30, 857-865.
- Chiurazzi P (1999), *DNA Panel for Interlaboratory Standardization of Haplotype Studies on the Fragile X Syndrome and Proposal for a New Allele Nomenclature*, *Am J Med Genet* 83, 347-349.
- Chiurazzi P, Genuardi M, Kozak L, Giovannucci-Uzielli ML, Bussani C, Dagna-Bricarelli F, Grasso M, Perroni L, Sebastio G, Sperandeo MP, Oostra BA, Neri G (1996), *Fragile X Founder Chromosomes in Italy, a few Initial Events and Possible Explanation for Their Heterogeneity*, *Am J Med Genet* 64, 209-215.
- Crawford C, Schwartz EC, Meadows KL, Newman JL, Taft LF, Gunter C, Brown WT, Carpenter NJ, Howard-Peebles PN, Monaghan KG, Nolin SL, Reiss AL, Feldman GL, Rohlf EM, Warren ST, Cherman SL (2000), *Survey of the Fragile X Syndrome CGG Repeat and the Short- Tandem-Repeat and Single-Nucleotide-Polymorphism Hplotypes in an African American Population*, *Am J Hum Genet* 66, 480-493.
- Eichler HH, Holden JJ, Propovich BW, Reiss AI, Snow K, Thibodeau SN, Richards CS, Ward PA, Nelson DL (1994), *Length of uninterrupted CGG repeats determines instability in the FMR1 gene*. *Nat Genet* 8, 88-94.
- Ennis S, Murray A, Brightwell G, Morton E, Jacobs PA (2007), *Closely Linked Cis-Acting Modifier of Expansion of the CGG Repeat in High Risk FMR1 Haplotypes*, *Hum Mutat* 28, 1216-1224.
- Ennis S, Murray A, Morton NE (2001), *Haplotypic Determinants of Instability in the FRAX Region, Concatenated Mutation of Founder Effect?*, *Hum Mutat* 18, 61-69.
- Excoffier L, Lischer HEL (2010), *Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows*. *Mol Ecol Resour* 10, 564-567.
- Fu YH, Kuhl DPA, Pizzuti A (1991), *Variations of the CGG repeat at the Fragile x site results in genetic instability, Resolution of the Sherman paradox*, *Cell* 67,1047-1058.
- Gatchel JR, Zoghbi HY (2005), *Disease of Unstable Repeat Expansion, Mechanism and Common Principals*, *Nat Rev Genet* 6, 743-745.

- Haataja R, Vaisanen ML, Li M, Ryyanen M, Leisti J (1994), *The fragile syndrome in Finland, demonstration of a founder effect by analysis of microsatellite haplotypes*, Hum Genet 94, 479-483.
- Hagerman PJ, Hagerman RJ (2004), *The fragile-X premutation, a maturing perspective*, Am J Hum Genet 74, 805–816.
- Hagerman RJ (1996), *Physical and behavioral phenotype* In, Hagerman RJ, Cronister A (eds). *Fragile X Syndrome, Diagnosis, Treatment and Research*. Johns Hopkins University Press Baltimore/London, 3-87.
- Hawk J, Stefanovic L, Boyer J, Petes T, Farber R (2005), *Variation in efficiency of DNA mismatch repair at different sites in the yeast genome*, Proc Natl Acad Sci U S A 102, 8639-8643.
- Heyer E, Puymirat J, Dieltjes P, Bakker E, De Knijff P (1997), *Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees*, Hum Mol Genet 6,799–803.
- Jarne P, Lagoda PJJ (1996), *Microsatellites, from molecules to populations and back*, Trends Ecol Evol 11, 424-429.
- Kalendar R, Lee D, Schulman AH (2009), *FastPCR Software for PCR Primer and Probe Design and Repeat Search*. Genes Genome Genomics 1, 1-14.
- Kibel AS, Faith DA, Bova GS, Isaacs WB (2003), *Xq27-28 deletions in prostatic carcinoma*, Genes Chromosomes Cancer 4, 381-388.
- Kruglyak S, Durrett RT, Schug MD, Aquadro CF (1998), *Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations*, Proc Natl Acad Sci U S A 18, 10774-10778.
- Morton NE, Macpherson JN (1992), *Population genetics of the fragile-x syndrome, Multiallelic model for the FMR1 locus*, Proc Natl Acad Sci U S A 9, 4215-4217.
- Nei M (1975), *Molecular Population Genetics and Evolution*, American Elsevier (Ed), 288.
- Oliveira B, Marques I, Loureiro J, Santos R, Jorge P (2010), *FMR1 repeat region instability assessment by Triplet primed PCR*, Comunicação em forma de Poster, 14^a reunião Sociedade Portuguesa de Genética Humana (SPGH), 67.
- Ota T, Kimura M (1993), *A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population*. Genet Res 22, 201–204.
- Oudet C, Mornet E, Serre JL, Thomas F, Lentès- Zengerling S, Kretz C, Deluchat C, Tejada I, Boué J, Boué A, Mandel E (1993), *Linkage Disequilibrium between the Fragile X Mutation and Two Closely Linked CA Repeats Suggests That Fragile X Chromosomes are Derived from a Small Number of Founder Chromosomes*, Am J Hum Genet 52, 297-304.
- Peixoto A, Santos MR, Seruca R, Amorim A, Castedo S (1998), *Analysis of FMR1 and flanking microsatellite markers in normal and fragile X chromosomes in Portugal: evidence for a “protector” haplotype*, Eur J Hum Genet 6, 518-522.

- Peprah EK, Allen EG, Williams SM, Woodard LM, Sherman SL (2010), *Genetic Diversity of the Fragile X Syndrome Gene (FMR1) in a Large Sub-Saharan West African Population*, *Ann Hum Genet* 74, 316-325.
- Richards RI, Shen Y, Holman K, Kozman H, Hyland VJ, Mulley JC, Sutherland GR (1991), *Fragile X Syndrome, Diagnosis using highly polymorphic microsatellite markers*, *Am J Hum Genet* 48, 1051-1057.
- Satta Y, Ohuigin C, Takahata N, Klein J (1993), *The synonymous substitution rate of the major histocompatibility complex loci in primates*. *Proc Natl Acad Sci USA* 90, 7480–7484.
- Smits A, Smeets D, Hamel B, Dreesen J, Oost B (1992), *High prevalence of the fra(X) syndrome cannot be explain by a high mutation rate*, *Am J Med Genet* 43, 345-352.
- Strachan T, Read A (2005), *Human Molecular Genetics*, Chapter 11, 3th edition, Garland Science, 279-282.
- Sutherland GR (1977), *Fragile sites on human chromosomes, Demonstration of their dependence on the type of tissue culture medium*, *Science* 197, 265–266.
- Tautz D, Renz M (1984), *Simple sequences are ubiquitous repetitive components of eukaryotic genomes*, *Nucleic Acids Res* 12, 4127-4138.
- Turner G, Webb T, Wake S, Robinson H (1996), *Prevalence of fragile X syndrome*, *Am J Med Genet* 64, 196–197.
- Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J, Kruglyak L, Stein L, Hsie L, Topaloglou T, Hubbell E, Robinson E, Mittmann M, Morris M S, Shen N, Kilburn D, Rioux J, Nusbaum C, Rozen S, Hudson TJ, Lipshutz R, Chee MS, Lander ES (1998), *Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome*, *Science* 280, 1077-1082.
- Weber JL, Wong C (1993), *Mutations of human short tandem repeats*, *Hum Mol Genet* 2,1123-1128.
- Yu S, Pritchard M, Kremer E, Lynch M, Nancarrow J, Baker E, Holman K, Mulley JC, Warren ST, Schlessiger D, Sutherland GR, Richards RI (1991), *Fragile X genotype characterized by an unstable region of DNA*, *Science* 252, 1179-1181.
- Zhong N, Ye L, Dokkin C, Brown WT (1993), *A complex mutable polymorphism located within fragile X gene*, *Nat Genet* 5, 248-252.
- Zhong N, Ye L, Dokkin C, Brown WT (1994), *Fragile X founder chromosome effects, linkage disequilibrium or microsatellite heterogeneity?*, *Am J Med Genet* 51, 405-411.

VII.2 Referências Eletrônicas

<http://www.fluxus-engineering.com>, Agosto de 2010

<http://www.ncbi.nlm.nih.gov/genome/sts/>, Outubro de 2009

<http://www.ncbi.nlm.nih.gov/projects/SNP/>, Setembro de 2010

<http://www.ncbi.nlm.nih.gov/RefSeq.html>, Setembro de 2010

Anexos

Anexo 1- Primers utilizados

STRs

Designação	Primer foward	Primer reverse	Concentração em stock
DXS8313	6-Fam-5'-CTTCTCTCAATCTCAGT TTAGG-3'	5'-AATAGATCAAGAACCTGAAGG-3'	100 pmol.µl ⁻¹
DXS548	5'-TACATTAGAGTCACCTGTGGTG C-3'	Hex-5'-AGAGCTTCACTATGCAATGG AATC-3'	100 pmol.µl ⁻¹
DXS998	5'-AGACATCACAGCAATTTTTCAA AG-3'	6-Fam-5'-AATAGTTTGTATTTTGA TCAATGCAA-3'	100 pmol.µl ⁻¹
FRAXAC1	Hex-5'-CAGGGCAAGGACATTA TCAGG-3'	5'-CATGATGAGAGTCACTTGAAGC TG-3'	100 pmol.µl ⁻¹
FRAXAC2	5'- ACTGCTCCGGAAGTTGAAT CC-3'	6-Fam-5'-CCCAGCTACTCGGGAGG C- 3'	100 pmol.µl ⁻¹

DXS548-P3	5'- GAATAGTCTCTGGGGTGGATCTC-3'	100 pmol.µl ⁻¹
DXS548-R	Hex-5'-AGAGCTTCACTATGCAA TGAATC-3'	100 pmol.µl ⁻¹

SNPs

Designação	Primer Foward	Primer reverse	Concentração em stock
rs971000	5'-AGTGGAGGATGGAGAG GCATTC-3'	5'-TATGCAGTATGAGGTAAAGTT GAAACG-3'	100 pmol.µl-1
rs29282	5'-CTTGATGTATGATG CTGAACTCTTAGC-3'	5'-ATGAACACAGGGACATCAAGA AAG-3'	100 pmol.µl-1
rs25715	5'-TCATGATGCAAATGAAA GCAGTC-3'	5'-CCCACCTTAAAGATAACTCAAAA CCAC-3'	100 pmol.µl-1

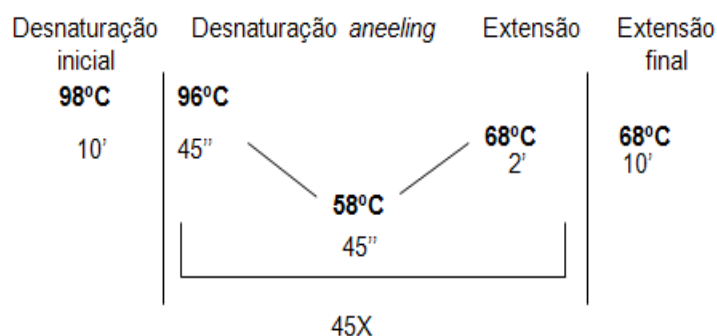
Região repetitiva [CGG] no gene *FMR1*

	Primer Foward	Primer reverse	Concentração em stock
g.FMR1-CGG	5'-TTCGGTTTCACTTCCGGTG- 3'	6-Fam-5'-CCATCTTCTTTCAGC CCTGC-3'	100 pmol.µl-1

Anexo 2- Determinação do número de tripletos [CGG]

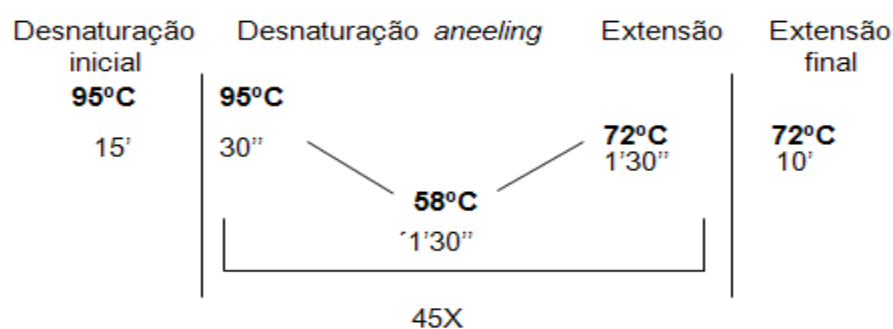
Preparação da mistura para PCR e programa utilizado no termociclador

Mistura		
dH ₂ O	6,49µl	
Betaina 5M	5µl	
DNTPs (10mM)	A	0,3µl
	T	0,3µl
	C	0,3µl
	G	0,06µl
7-Deaza	1µl	
DMSO	1,2µl	
PrimerF 10 pmol.µl ⁻¹	1µl	
PrimerR 10 pmol.µl ⁻¹	1µl	
Accutaq buffer 10x	2µl	
Accutaq	0,35µl	
gDNA 300-500ng.µl ⁻¹	1µl	
Volume final	20µl	



Anexo 3- Amplificação dos STRs por PCR Multiplex

Programa de termociclador utilizado e preparação da mistura para PCR

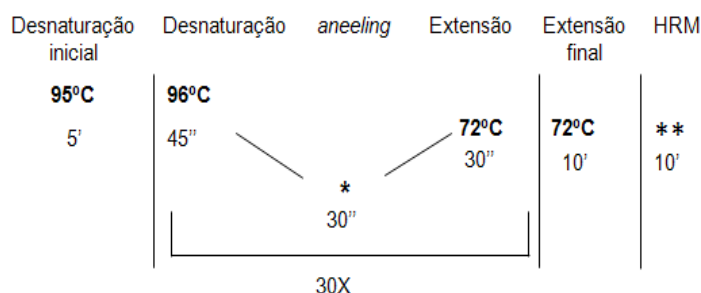


Mistura para PCR Multiplex	
dH ₂ O	4,5µl
Master Mix Multiplex Qiagen 2x	7,5µl
Mistura de Primers	1,5µl
gDNA 300ng.µl ⁻¹	1,5µl
Volume final	15µl

Mistura de Primers		
Primers	Concentração	Volume
DXS8313-F	2 pmol.µl ⁻¹	1,5µl
DXS8313-R		
DXS548-F	2 pmol.µl ⁻¹	4µl
DXS548-R		
DXS998-F	3 pmol.µl ⁻¹	1,5µl
DXS998-R		
FRAXAC1-F	5 pmol.µl ⁻¹	2µl
FRAXAC1-R		
FRAXAC2-F	5 pmol.µl ⁻¹	2,5µl
FRAXAC2-R		

Anexo 4- High Resolution Melting Curve Analysis- hrMCA

Programas e preparação de misturas



Temperatura de <i>annealing</i>		
*	rs971000	60°C
	rs29282	58°C
	rs25715	56°C

Intervalo de temperatura de HRM			<i>melting</i>
**	rs971000	77°C - 87°C	~ 84,5°C
	rs29282	75°C - 85°C	~ 79°C
	rs25715	70°C - 80°C	~ 75,5 °C

Mistura para rs25715	
Água	2µl
Master Mix Promega 2x	7,5µl
Primer-F 5pmol.µl-1	1,5µl
Primer-R 5pmol.µl-1	
Evagreen 20x	1µl
gDNA 50ng.µl-1	1,5µl
Volume final	15µl

Mistura para rs971000 e rs29282	
Água	1µl
Master Mix Promega 2x	7,5µl
Primer-F 5pmol.µl-1	1,5µl
Primer-R 5pmol.µl-1	
Evagreen 20x	1µl
gDNA 5ng.µl-1	2,5µl
Volume final	15µl

Anexo 5 – Haplótipos e frequências haplotípicas na população Controlo

Haplótipo	Frequência Relativa	Frequência Absoluta	STRs				SNPs		
			DXS998	DXS548	FRAxAC1	FRAxAC2	rs971000	rs929282	rs25715
A1	0,1737	37	34	40	38	336	T	T	T
A2	0,0751	16	36	40	38	336	T	T	T
A3	0,0516	11	32	40	38	336	T	T	T
A4	0,0376	8	34	38	38	336	T	T	T
A5	0,0282	6	34	40	38	338	T	T	T
A6	0,0235	5	38	40	38	336	T	T	T
A7	0,0188	4	34	40	38	335	T	T	T
A8	0,0141	3	36	50	38	336	T	T	T
A9	0,0094	2	38	38	38	336	T	T	T
A10	0,0094	2	34	42	38	336	T	T	T
A11	0,0094	2	44	40	38	336	T	T	T
A12	0,0094	2	36	42	38	336	T	T	T
A13	0,0094	2	34	38	38	338	T	T	T
A14	0,0094	2	34	40	38	337	T	T	T
A15	0,0094	2	38	40	38	338	T	T	T
A16	0,0094	2	36	40	38	338	T	T	T
A17	0,0047	1	34	38	38	335	T	T	T
A18	0,0047	1	32	40	38	335	T	T	T

Síndrome de X-Frágil em Portugal

Haplótipo	Frequência Relativa	Frequência Absoluta	STRs				SNPs		
			DXS998	DXS548	FRAxAC1	FRAxAC2	rs971000	rs929282	rs25715
A19	0,0047	1	36	40	38	334	T	T	T
A20	0,0047	1	36	42	38	334	T	T	T
A21	0,0047	1	34	44	38	336	T	T	T
A22	0,0047	1	32	40	38	338	T	T	T
A23	0,0047	1	34	40	34	341	T	T	T
A24	0,0047	1	42	40	38	335	T	T	T
A25	0,0047	1	34	46	38	335	T	T	T
A26	0,0047	1	34	40	38	340	T	T	T
A27	0,0047	1	32	42	38	336	T	T	T
A28	0,0047	1	34	40	40	336	T	T	T
A29	0,0047	1	34	42	38	335	T	T	T
A30	0,0047	1	34	40	38	342	T	T	T
A31	0,0047	1	38	40	34	341	T	T	T
A32	0,0047	1	36	40	40	336	T	T	T
A33	0,0047	1	38	50	38	336	T	T	T
A34	0,0047	1	34	42	36	336	T	T	T
B1	0,0516	11	34	40	38	335	C	T	C
B2	0,0188	4	34	40	38	334	C	T	C
B3	0,0141	3	34	50	42	337	C	T	C
B4	0,0141	3	36	40	38	336	C	T	C
B5	0,0141	3	36	40	38	335	C	T	C
B6	0,0094	2	34	38	38	335	C	T	C
B7	0,0094	2	38	42	36	333	C	T	C
B8	0,0094	2	34	40	36	332	C	T	C
B9	0,0094	2	32	40	38	334	C	T	C
B10	0,0094	2	36	40	36	332	C	T	C
B11	0,0047	1	32	40	38	336	C	T	C
B12	0,0047	1	38	38	38	336	C	T	C
B13	0,0047	1	32	40	38	335	C	T	C
B14	0,0047	1	34	52	42	336	C	T	C
B15	0,0047	1	36	52	42	337	C	T	C
B16	0,0047	1	34	52	42	337	C	T	C
B17	0,0047	1	32	50	40	337	C	T	C
B18	0,0047	1	32	48	36	333	C	T	C
B19	0,0047	1	38	52	42	336	C	T	C
B20	0,0047	1	32	50	42	337	C	T	C
B21	0,0047	1	38	40	36	331	C	T	C
B22	0,0047	1	34	50	42	336	C	T	C
B23	0,0047	1	36	42	36	331	C	T	C
B24	0,0047	1	32	56	42	338	C	T	C
B25	0,0047	1	34	38	36	331	C	T	C
B26	0,0047	1	36	50	42	337	C	T	C
B27	0,0047	1	32	48	38	335	C	T	C
B28	0,0047	1	34	40	42	337	C	T	C
B29	0,0047	1	34	40	36	331	C	T	C
B30	0,0047	1	34	48	42	336	C	T	C
B31	0,0047	1	32	50	42	336	C	T	C
C1	0,0282	6	34	42	36	335	C	C	C
C2	0,0047	1	34	42	38	336	C	C	C

Síndrome de X-Frágil em Portugal

Haplótipo	Frequência Relativa	Frequência Absoluta	STRs				SNPs		
			DXS998	DXS548	FRAXAC1	FRAXAC2	rs971000	rs929282	rs25715
C3	0,0047	1	32	40	38	335	C	C	C
C4	0,0047	1	32	42	36	333	C	C	C
C5	0,0047	1	36	40	38	334	C	C	C
C6	0,0047	1	36	42	36	333	C	C	C
C7	0,0047	1	34	48	36	337	C	C	C
C8	0,0047	1	36	42	36	337	C	C	C
C9	0,0047	1	32	40	36	333	C	C	C
C10	0,0047	1	34	38	34	332	C	C	C
C11	0,0047	1	32	40	36	329	C	C	C
D1	0,0047	1	34	40	38	336	C	T	T
D2	0,0047	1	32	40	38	336	C	T	T
D3	0,0047	1	34	38	38	335	C	T	T
D4	0,0047	1	34	40	38	337	C	T	T
D5	0,0047	1	36	52	42	337	C	T	T
D6	0,0047	1	34	48	38	335	C	T	T
E1	0,0141	3	34	40	38	336	T	T	C
E2	0,0047	1	36	40	38	336	T	T	C
E3	0,0047	1	34	38	38	336	T	T	C
E4	0,0047	1	32	42	36	333	T	T	C
E5	0,0047	1	34	52	42	336	T	T	C
E6	0,0047	1	32	40	38	337	T	T	C
E7	0,0047	1	34	44	38	335	T	T	C
E8	0,0047	1	38	42	36	336	T	T	C
E9	0,0047	1	36	40	42	336	T	T	C
F1	0,0047	1	34	40	38	336	T	C	C

Anexo 6 – Haplótipos e frequências haplotípicas na população com Síndrome de X-Frágil

Haplótipo	Frequência Relativa	Frequência Absoluta	STRs				SNPs		
			DXS998	DXS548	FRAXAC1	FRAXAC2	rs971000	rs29282	rs25715
A35	0,0488	6	36	44	38	336	T	T	T
A2	0,0244	3	36	40	38	336	T	T	T
A36	0,0081	1	36	44	38	335	T	T	T
A3	0,0081	1	32	40	38	336	T	T	T
A37	0,0081	1	36	40	38	335	T	T	T
A38	0,0081	1	36	44	36	335	T	T	T
B32	0,0894	11	36	50	38	335	C	T	C
B26	0,0488	6	36	50	42	337	C	T	C
B33	0,0325	4	36	50	44	335	C	T	C
B34	0,0325	4	36	40	36	337	C	T	C
B35	0,0325	4	40	50	42	336	C	T	C
B36	0,0325	4	36	50	42	336	C	T	C
B11	0,0325	4	32	40	38	335	C	T	C

Síndrome de X-Frágil em Portugal

Haplótipo	Frequência Relativa	Frequência Absoluta	STRs				SNPs		
			DXS998	DXS548	FRAxAC1	FRAxAC2	rs971000	rs929282	rs25715
B37	0,0325	4	36	44	42	336	C	T	C
B10	0,0244	3	36	40	36	332	C	T	C
B38	0,0244	3	40	50	42	337	C	T	C
B39	0,0163	2	36	44	42	337	C	T	C
B40	0,0081	1	36	50	42	339	C	T	C
B41	0,0081	1	40	44	42	337	C	T	C
B42	0,0081	1	44	50	42	337	C	T	C
B43	0,0081	1	36	46	42	337	C	T	C
B44	0,0081	1	40	46	46	337	C	T	C
B9	0,0081	1	32	40	38	334	C	T	C
B45	0,0081	1	36	40	36	336	C	T	C
B46	0,0081	1	36	50	36	335	C	T	C
B47	0,0081	1	40	40	42	336	C	T	C
B48	0,0081	1	36	40	36	334	C	T	C
B49	0,0081	1	38	50	38	335	C	T	C
B50	0,0081	1	36	42	36	338	C	T	C
B51	0,0081	1	36	40	36	338	C	T	C
B52	0,0081	1	36	50	38	334	C	T	C
B53	0,0081	1	36	52	36	332	C	T	C
B54	0,0081	1	36	54	42	339	C	T	C
B55	0,0081	1	36	50	38	339	C	T	C
B56	0,0081	1	32	48	38	334	C	T	C
B57	0,0081	1	36	40	38	337	C	T	C
B58	0,0081	1	32	50	42	337	C	T	C
B59	0,0081	1	38	50	42	337	C	T	C
B60	0,0081	1	36	50	42	335	C	T	C
B61	0,0081	1	36	50	42	334	C	T	C
B62	0,0081	1	36	44	40	337	C	T	C
C12	0,0488	6	36	42	36	333	C	C	C
C13	0,0325	4	40	42	36	333	C	C	C
C14	0,0325	4	36	42	36	335	C	C	C
C6	0,0244	3	36	42	36	334	C	C	C
C15	0,0081	1	44	42	36	335	C	C	C
C16	0,0081	1	48	46	36	332	C	C	C
C17	0,0081	1	44	46	36	334	C	C	C
C18	0,0081	1	40	42	36	335	C	C	C
C19	0,0081	1	40	42	36	332	C	C	C
C20	0,0081	1	36	46	36	333	C	C	C
C21	0,0081	1	36	40	36	335	C	C	C
C22	0,0081	1	32	42	36	335	C	C	C
C23	0,0081	1	48	40	36	333	C	C	C
C24	0,0081	1	44	42	36	334	C	C	C
C4	0,0081	1	32	42	36	333	C	C	C
D7	0,0081	1	36	50	44	335	C	T	T
D8	0,0081	1	36	40	36	332	C	T	T
D9	0,0081	1	36	40	36	333	C	T	T
E10	0,0325	4	36	44	38	336	T	T	C
E11	0,0081	1	36	44	38	335	T	T	C
E12	0,0081	1	44	44	38	336	T	T	C