

Modificadores Genéticos da Instabilidade da Repetição (CAG)_n na Doença de Machado-Joseph/Ataxia Espinocerebelosa do Tipo 3

Dissertação de Mestrado

Sara Raquel Rebelo Pavão

Mestrado em

Ciências Biomédicas



**Modificadores Genéticos da Instabilidade da Repetição (CAG)_n na
Doença de Machado-Joseph/Ataxia Espinocerebelosa do Tipo 3**

Dissertação de Mestrado

Sara Raquel Rebelo Pavão

Orientador (es): Prof.^ª Doutora Maria Manuela de Medeiros Lima

Doutora Mafalda Sofia Bastos Raposo

Dissertação submetida como requisito parcial para obtenção do grau de

Mestre em Ciências Biomédicas



**“VALEU A PENA? TUDO VALE A PENA
SE A ALMA NÃO É PEQUENA.”
FERNANDO PESSOA**

AGRADECIMENTOS

Em primeiro lugar à prof^a Doutora Manuela Lima, por ter aceitado ser minha orientadora e por ter acreditado sempre nas minhas capacidades. Agradeço-lhe a oportunidade que me deu de poder trabalhar na área da Genética Humana, assim como a sua disponibilidade em ajudar e partilhar comigo os seus conhecimentos. Agradeço cada momento que me motivou a continuar! É para mim uma grande referência de profissionalismo, tanto como investigadora como docente. Obrigada!

À doutora Mafalda Raposo, também minha orientadora, que sempre mostrou disponibilidade em todas as etapas da elaboração deste trabalho. Obrigada por ter partilhado comigo o seu conhecimento e, sobretudo, obrigada pela paciência comigo em momentos que eu própria não teria.

À doutoranda Ana Rosa Melo, que me acompanhou nas etapas laboratoriais, no processo de escrita e análise de resultados e pela motivação nos momentos mais difíceis. Agradeço as conversas nas horas de almoço e a sua amizade, assim como a partilha de conhecimentos.

À doutoranda Ana Ferreira, a disponibilidade para me ajudar em determinados momentos deste percurso, bem como a partilha dos seus conhecimentos. Agradeço, também a amizade e a boa disposição que sempre tornaram os dias melhores.

À minha amiga Filipa Medeiros, pelo carinho, pela motivação que me deu e pela amizade que sempre nos acompanharam. Agradeço, também a ajuda no laboratório que me deu em determinados momentos deste trabalho.

Ao senhor Pedro Mântua, meu amigo, que tem vindo a acompanhar o meu percurso académico. Pelas suas palavras, pelos seus conselhos e por também acreditar em mim. A sua ajuda tem sido fundamental no meu percurso.

Aos meus melhores amigos Francisco Santos, Horácio Dutra e Pedro Carvalho, obrigada pela paciência, por terem acreditado em mim e pelo apoio incondicional.

Ao João, pelo tempo que não passámos juntos, pela sua paciência, pelo carinho, pelo amor e apoio incondicionais que sempre recebi. Obrigada por não teres desistido de mim!

Às minhas irmãs, Rita e Teresa, pelo carinho, pela amizade e cumplicidade que temos. Pela vossa paciência quando precisei conversar, pelo vosso apoio na minha caminhada.

Por último, e não menos importantes, aos meus pais, a eles que são o pilar da minha vida e que sempre me motivaram a continuar. Pela educação que pude ter, pelo esforço que fizeram para que eu pudesse terminar. Foi um percurso difícil, mas sem eles jamais teria conseguido.

| | |
|--|-----------|
| Índice | |
| Índice de Figuras..... | 6 |
| Índice de Tabelas..... | 7 |
| Resumo..... | 8 |
| Abstract..... | 10 |
| 1. Introdução..... | 13 |
| Repetições Trinucleotídicas..... | 13 |
| Doenças de Poliglutamina..... | 15 |
| Doença de Machado-Joseph/ Ataxia Espinocerebelosa do tipo 3..... | 16 |
| Mecanismos de Reparação do DNA e Instabilidade das Repetições..... | 20 |
| Objectivos do trabalho..... | 23 |
| 2. Sujeitos e Métodos..... | 25 |
| Sujeitos..... | 25 |
| Amostras..... | 26 |
| Trabalho Experimental..... | 26 |
| Extração de DNA de mucosa bucal..... | 26 |
| Determinação do Número de Repetições CAG do Gene <i>ATXN3</i> em amostras de Sangue e de Mucosa Bucal..... | 27 |
| PCR Convencional..... | 27 |
| Triplet Repeat Primed PCR (TP-PCR)..... | 27 |

| | |
|---|-----------|
| Análise de Fragmentos por Electroforese Capilar | 28 |
| Análise das Variantes Alélicas dos Genes <i>PMS2</i> , <i>ERCC6</i> e <i>FAN1</i> por Sequenciação | 28 |
| Desenho dos Primers | 28 |
| Otimização das Condições da PCR..... | 29 |
| Sequenciação por Eletroforese Capilar | 30 |
| Tratamento dos Dados | 31 |
| 3. Resultados e Discussão..... | 34 |
| Caracterização da População de Estudo..... | 34 |
| Caracterização do Padrão de Instabilidade Somática | 35 |
| Caracterização do Padrão de Instabilidade Intergeracional..... | 37 |
| Caracterização das Variantes Alélicas dos Genes <i>ERCC6 g.19868 G>A</i> , <i>FAN1 g.3228531 C>G</i> e <i>PMS2 g.26796 C>A</i> | 41 |
| Estudo Populacional das Variantes..... | 41 |
| Relação entre as Variantes Alélicas e a Instabilidade Somática..... | 42 |
| Relação entre as Variantes Alélicas e a Instabilidade Intergeracional | 43 |
| Correlações Genótipo-Fenótipo | 43 |
| 4. Conclusão | 47 |
| 5. Referências | 49 |

ÍNDICE DE FIGURAS

| | |
|--|----|
| Figura 1. Diferença entre o número de repetições CAG para o alelo expandido obtida nas amostras de sangue e nas amostras de mucosa bucal para os 26 indivíduos seleccionados. Os resultados estão apresentados por frequência absoluta do número de indivíduos. | 35 |
| Figura 2. Comparação do índice de mosaicismo entre amostras de mucosa bucal e de sangue..... | 36 |
| Figura 3. Classificação e direcção da instabilidade da repetição CAG no alelo expandido do gene <i>ATXN3</i> . A: classificação das 31 transmissões estudadas. B: direcção das 21 transmissões instáveis..... | 38 |
| Figura 4. Variação do número de repetições CAG entre progenitor e descendente, no alelo expandido do gene <i>ATXN3</i> | 40 |
| Figura 5. Relação entre a idade de início da doença e a frequência alélica dos genes <i>ERCC6</i> e <i>FAN1</i> . A média da idade de início dos sintomas foi estimada tendo em consideração o número de repetições CAG no alelo expandido (67,94 repetições). | 44 |
| Figura 6. Relação entre a idade de início da doença e os genótipos obtidos para o gene <i>FAN1</i> ; $\neq p < 0,05$. A média da idade de início dos sintomas foi estimada tendo em consideração o número de repetições CAG no alelo expandido como co-variável (67,94 repetições)..... | 45 |

ÍNDICE DE TABELAS

| | |
|--|----|
| Tabela 1. Doenças de poliglutamina, respectivos loci e proteínas, bem como número de repetições CAG (adaptado de Fan <i>et al.</i> , 2014)..... | 16 |
| Tabela 2. Sequências dos primers utilizados na reacção de PCR e sequenciação, para análise das variantes alélicas nos genes <i>PMS2</i> , <i>ERCC6</i> e <i>FAN1</i> . As temperaturas de <i>annealing</i> estão representadas na tabela. As sequências a negrito indicam os primers utilizados na reacção de sequenciação. | 29 |
| Tabela 3. Descrição demográfica, molecular e clínica dos 54 indivíduos DMJ seleccionados para o presente estudo. | 34 |
| Tabela 4. Dados genéticos e idade média de concepção das transmissões utilizadas neste estudo..... | 39 |
| Tabela 5. Frequências alélicas e genotípicas das variantes <i>PMS2</i> g.26796 C>A, <i>ERCC6</i> g.19868 G>A e <i>FAN1</i> g.3228531 C>G. | 41 |
| Tabela 6. Resultados do teste exacto de diferenciação populacional entre a coorte DMJ açoriana (N=108) e as restantes populações descritas na tabela. N: número de cromossomas; (-): diferenças não significativas; (+): p<0,05. | 42 |
| Tabela 7. IM para cada tecido de acordo com as variantes para os genes estudados.. | 42 |

RESUMO

A doença de Machado-Joseph (DMJ), também conhecida por ataxia espinocerebelosa do tipo 3 (SCA3), é uma doença autossómica dominante causada pela expansão de motivos repetitivos CAG no exão 10 do gene *ATXN3*, localizado em 14q32.1. O tamanho das repetições é o principal condicionante da idade de manifestação e severidade da doença. Os alelos normais têm um comportamento estável, mantendo-se inalterado o número de repetições aquando da mitose e da meiose. Todavia, na DMJ, à semelhança do que acontece noutras doenças de poliglutamina (poliQ), os alelos expandidos apresentam instabilidade somática e intergeracional. O papel dos mecanismos de reparação do DNA na instabilidade das repetições permanece por esclarecer. Os principais objectivos da presente dissertação foram: (1) caracterizar a instabilidade somática e intergeracional em sujeitos DMJ açorianos; (2) analisar a associação entre as variantes *PMS2* g.26796 C>A (rs1805323), *ERCC6* g.19868 G>A (rs2228528) e *FAN1* g.3228531 C>G (rs3512) e a instabilidade somática e intergeracional; (3) analisar a associação entre as variantes referidas em (2) e a idade de início da DMJ. No presente estudo foram utilizadas amostras de DNA de 54 sujeitos DMJ. Para o estudo da instabilidade intergeracional usaram-se 50 amostras de DNA de sangue (num total de 31 transmissões). Para o estudo do mosaïcismo somático utilizaram-se 26 amostras de DNA de sangue e 26 amostras de DNA de mucosa bucal, representando 26 indivíduos. A determinação das repetições CAG no alelo expandido do gene *ATXN3* foi efectuada utilizando PCR convencional e TP-PCR (*Triplet repeat primed PCR*), seguidas de análise de fragmentos por electroforese capilar. As amostras foram genotipadas para as variantes alélicas dos genes *PMS2*, *ERCC6* e *FAN1* por sequenciação de Sanger. Para a análise da instabilidade somática, após o cálculo dos genótipos obtidos por AF, foi calculada a diferença entre o tamanho da repetição CAG obtido a partir do DNA extraído das 26 amostras de sangue e do DNA extraído de mucosa bucal, para cada indivíduo (CAGexp (sangue) – CAGexp (mucosa bucal)). A partir da análise do padrão de picos obtido na TP-PCR, determinou-se o índice de mosaïcismo (IM). O

genótipo obtido por PCR convencional foi utilizado para classificar as transmissões em estáveis e instáveis, calculando-se a diferença no número de repetições CAG entre o descendente e o progenitor ($CAG_{exp}(\text{descendente}) - CAG_{exp}(\text{progenitor})$). A presença/ausência de instabilidade somática e intergeracional foi correlacionada com as variantes alélicas *ERCC6* g.19868 G>A, *FAN1* g.3228531 C>G e *PMS2* g.2696 C>A. Cinquenta por cento dos indivíduos não apresentam mosaïcismo somático nos 2 tecidos estudados. O valor médio de IM para cada um dos tecidos foi semelhante, apesar de ligeiramente superior no sangue, indicando uma maior instabilidade neste tecido relativamente ao epitélio da mucosa bucal. Das transmissões estudadas, 70% foram instáveis e 30% estáveis, sendo que nas instáveis 66,7% corresponderam a expansões e 33,8% a contracções. Não foi observada nenhuma associação entre as variantes *ERCC6* g.19868 G>A e *FAN1* g.3228531 C>G e a instabilidade somática e intergeracional. De acordo com a correlação genótipo-fenótipo efectuada entre o gene *FAN1* e a idade de início da DMJ observou-se que os doentes *FAN1**GG apresenta uma idade significativamente mais tardia que os doentes *FAN1**CC. Este resultado está de acordo com o previamente reportado numa série extensa de doentes com várias doenças de poliQ. A confirmação do papel do *FAN1* na patogénese da DMJ permitirá compreender a interligação da via de reparação de *Fanconi Anemia*, na qual o gene *FAN1* está envolvido, e a patogénese molecular da DMJ, assim evidenciando um potencial papel desta via enquanto alvo terapêutico.

ABSTRACT

Machado-Joseph disease (MJD), also known as spinocerebellar ataxia type 3 (SCA3), is an autosomal dominant disease and is caused by an abnormal number of CAG repeats in exon 10 of the *ATXN3* gene, located on 14q32.1. Age at onset and severity of the disease is mainly explained by the size of the CAG motifs in the expanded allele. Normal alleles present a stable behavior, in which the number of repeats remains unchanged in mitotic and meiotic divisions. However, in MJD, similarly to the observed in other poliQ diseases, alterations in the size of the expanded alleles present mitotic and meiotic instability. The role of DNA repair in the instability of CAG repeats remains to be clarified. The main goals of the present dissertation were: (1) characterize the somatic and intergenerational instability in Azorean MJD subjects; (2) analyze the association between *PMS2* g.26796 C> A (rs1805323), *ERCC6* g.19868 G> A (rs2228528) and *FAN1* g.3228531 C> G (rs3512) and somatic and intergenerational instability; (3) analyze the association between the variants mentioned in (2) and the age at onset in MJD. In the present study, a total of DNA samples from 54 MJD subjects were used; DNA from 50 blood samples (for a total of 31 transmissions) and from 26 blood and buccal swab samples were used in the intergenerational and somatic instability analysis, respectively. The number of CAG repeats in the *ATXN3* gene expanded allele was determined using conventional PCR and TP-PCR (Triplet repeat primed PCR), followed by fragment analysis by capillary electrophoresis (FA). Allelic variants of the *PMS2*, *ERCC6* and *FAN1* genes were genotyped by Sanger sequencing. The difference between the size of the CAG repeats in blood and in buccal swab samples ($CAG_{exp}(\text{blood}) - CAG_{exp}(\text{buccal swab})$) as well as the calculation of the mosaicism index (MI) by TP-PCR were used to characterize somatic instability. The difference between the number of CAG repeats in the offspring and in the transmitting parent ($CAG_{exp}(\text{offspring}) - CAG_{exp}(\text{transmitting parent})$) was calculated to characterize intergenerational instability. The presence/absence of the allelic variants *ERCC6* g.19868 G> A, *FAN1* g.3228531 C> G and *PMS2* g.26796 C> A was correlated with somatic and

intergeracional instability. Somatic instability was not observed in 50% of the individuals analyzed. Furthermore, although slightly higher in the blood, MI was similar in both tissues. Seventy percent of the transmissions were unstable and from this 66,7% were expansions and 33,8% were contractions. These findings are in accordance with previous reports for MJD. Associations between *ERCC6* g.19868 G> A and *FAN1* g.3228531 C> G variants and somatic and intergenerational instability were not observed. Genotype/phenotype analysis revealed that patients presenting *FAN1**GG have a significantly late onset than patients with *FAN1**CC, a result in accordance with a previous larger study including patients with polyQ diseases. The confirmation of *FAN1* as MJD modifier will allow to understand the link between Fanconi Anemia repair mechanisms, in which *FAN1* is involved, and MJD pathogenesis, thus evidencing the potential of this pathway as a therapeutic target.