



MAFALDA SOFIA BASTOS RAPOSO

Predicting and tracking Machado-Joseph disease:  
biomarkers of diagnosis and prognosis

UNIVERSIDADE DOS AÇORES  
FACULDADE DE CIÊNCIAS E TECNOLOGIA  
PONTA DELGADA

2016

MAFALDA SOFIA BASTOS RAPOSO

# Predicting and tracking Machado-Joseph disease: biomarkers of diagnosis and prognosis

Dissertation for PhD degree  
in Biology presented to the  
University of the Azores

Supervisor: Maria Manuela de Medeiros Lima

Co-supervisor: Maria da Conceição Félix Bettencourt

UNIVERSIDADE DOS AÇORES  
FACULDADE DE CIÊNCIAS E TECNOLOGIA  
PONTA DELGADA

2016



União Europeia  
Fundo Social Europeu



PROEMPREGO



Governo dos Açores

**AÇORES**  
2020  
PROGRAMA OPERACIONAL  
FEDER FSE



**FRC**

FUNDO REGIONAL PARA A CIÊNCIA

## ACKNOWLEDGMENTS

The development of the present dissertation has been a great adventure, both personally and professionally, and I must express my eternal gratitude to all the people who shared with me this time:

- To my supervisors. To Professor Manuela Lima, which has been always present, for her unconditional support, for her ability of teaching me, for allowing me to grow up (after hitting often with my head on the walls, which, usually, were in the format of reviewers), by our scientific conversations (often mixed with talks less scientific) and her way of living science, being an example to follow (my example). I'm afraid that these words not thank you enough. To Dr. Conceição Bettencourt, who despite the distance has always been there when needed, by their exclamation points (usually above 5 was not a good sign), for your friendship and because we share the same birthday day (which itself says a lot of things).
- To the members of GAIN: Prof. Manuela, Dr. Carlos, Dr. Teresa and Dr. João. Our memorable trips to the islands, the team spirit and the capacity of joint efforts for a unique benefit (the MJD patients). The biggest laughs of this journey were undoubtedly shared with this group, which in turn made most of the shared moments, unforgettable. A special word to Dr. João, who allowed me to share with him, all these years, the clinical evaluation of the MJD patients, and that allowed me to live, closely, the reality of MJD. I must also mention Mariazinha, who received me with open arms and was always there when needed (a lot of times).
- To the MJD patients and their families, for their cooperation and availability. Often I shared their suffering, but, also, their hope.
- To my family. To my uncles, cousins, godmother, Andreia and Sofia that despite the distance we have always been together and support me unconditionally. My godson, who was born in the

same month and year of the beginning of this PhD project, and I hope that one day, he will be proud of me.

- To my friends, from Moncorvo and São Miguel. To my friends from Moncorvo for being my friends ever and forever (Dininha, Marisa, Tânia, Paula, Sandra, the Leonardo brothers, Julinho, Nuno, Zé Tó). To my friends from São Miguel, which are, currently, my support, who put me up all these years, and helping me to avoid missing home. To Zulmira, for her unconditional friendship, and to the other members of the Teves family (my Azorean family that received me with open arms), to Duarte and Ana (I will never be able to thank you enough for your friendship), to my housemates (especially to Claudia, for sharing with me many evenings), to Rita for her endless friendship, to my “lunch-box, trails and anything else UAc group” (Elizabete, Marta, Ana, Zita, Claudia, Ana Rosa).

- To the past and present members of the Human Genetics group. To Cristina (by skype and emails), Amanda (my favourite Catalan) and Nadiya for the helpful brainstormings during these years. To Pedro Mantua (the only man in the group, God help him), for given me laboratorial support and for listened my “thoughts” in the second floor.

- To Filipe, for all the moments we share, for being the other half. Of all the good things that sharing live brings love is the greatest of all.

- To my mammy for all the love and unconditional support, for being the best mother, for always being there, for being an example of life, for everything.

# TABLE OF CONTENTS

<b>ABSTRACT .....</b>	<b>vii</b>
<b>RESUMO .....</b>	<b>ix</b>
<b>INTRODUCTION.....</b>	<b>1</b>
••• MACHADO-JOSEPH DISEASE (MJD)/SPINOCEREBELLAR ATAXIA TYPE 3 (SCA3): AN OVERVIEW ••• .....	3
<i>Epidemiology</i> .....	4
<i>Clinical features</i> .....	5
<i>Molecular genetics and genotype-phenotype correlations</i> .....	6
<i>Origins of MJD mutation</i> .....	7
<i>Ataxin-3 and its biological role</i> .....	8
<i>MJD clinical trials and its outcome measures</i> .....	10
••• BIOMARKERS IN SCA3: STATE OF THE ART ••• .....	11
<i>Definitions and validation steps</i> .....	11
<i>Clinical biomarkers</i> .....	13
<i>Imaging biomarkers</i> .....	15
<i>Molecular biomarkers</i> .....	17
Trait biomarkers.....	17
State biomarkers.....	18
••• PERTINENCE AND MAIN GOAL ••• .....	21
<b>CHAPTER I .....</b>	<b>31</b>
<i>THE AZOREAN MACHADO-JOSEPH DISEASE/SPINOCEREBELLAR ATAXIA TYPE 3 COHORT: CLINICAL AND GENETIC CHARACTERIZATION.</i> .....	31
<b>CLINICAL BIOMARKERS</b>	
<b>CHAPTER II .....</b>	<b>49</b>
<i>NYSTAGMUS AS AN EARLY OCULAR ALTERATION IN MACHADO-JOSEPH DISEASE (MJD/SCA3).</i> .....	49
<b>MOLECULAR TRAIT BIOMARKERS</b>	
<b>CHAPTER III .....</b>	<b>63</b>
<i>REPLICATING STUDIES OF GENETIC MODIFIERS IN SPINOCEREBELLAR ATAXIA TYPE 3: CAN HOMOGENEOUS COHORTS AID? ....</i>	63
<b>CHAPTER IV .....</b>	<b>71</b>
<i>PROMOTER VARIATION AND EXPRESSION LEVELS OF INFLAMMATORY GENES IL1A, IL1B, IL6 AND TNF IN BLOOD OF SPINOCEREBELLAR ATAXIA TYPE 3 (SCA3) PATIENTS.</i> .....	71

**MOLECULAR STATE BIOMARKERS**

**CHAPTER V..... 89**  
*NOVEL CANDIDATE BLOOD-BASED TRANSCRIPTIONAL BIOMARKERS OF MACHADO-JOSEPH DISEASE. .... 89*

**CHAPTER VI..... 109**  
*HSPB1 AND BCL2 mRNA LEVELS IN BLOOD OF SPINOCEREBELLAR ATAXIA TYPE 3 SUBJECTS: RESULTS FROM CROSS-SECTIONAL AND LONGITUDINAL ANALYSES. .... 109*

**CHAPTER VII..... 127**  
*MITOCHONDRIAL DNA CONTENT AND COMMON DELETION QUANTIFICATION IN BLOOD OF SPINOCEREBELLAR ATAXIA TYPE 3 PRETAXIC CARRIERS AND PATIENTS. .... 127*

**CONCLUSIONS ..... 137**

## ABSTRACT

Machado-Joseph disease/spinocerebellar ataxia type 3 (MJD/SCA3) is a rare late-onset neurodegenerative disorder, which is currently untreatable. In this polyglutamine (polyQ) disease, a CAG repeat expansion located at the *ATXN3* gene, is the causative mutation. As a progressive disorder, MJD is characterized by different stages throughout its natural course, namely by a long preclinical period, during which molecular and even clinical alterations could be starting, many years before the ataxia onset. MJD onset is partially explained by the size of the CAG tract constituting its main trait biomarker. The incompleteness of this explanation, however, implies the existence of genetic modifiers, which should provide additional trait biomarkers. The identification of molecular trait biomarkers of MJD will allow to: (1) improve the prediction of age at onset and (2) empower interventional trials, by allowing genetic stratification of patients. The characterization of MJD disease stages has so far been performed using mainly clinical scales or imaging data, while few studies investigated molecular biomarkers. The identification of such molecular state biomarkers would allow: (1) the prediction of age at pathological onset (age at which cell damage is already present although a clinical phenotype is not yet evident), (2) monitoring different stages of disease progression, and (3) detecting subtle therapeutic benefits, in emergent clinical trials. The main goal of the present PhD project was to contribute to the development of molecular biomarkers for MJD. RNA and DNA, obtained from blood samples of preataxic subjects and MJD patients of Azorean background, including samples collected in two different moments, constituted the main resource used throughout the present PhD project. Samples from population controls were also used. The determination of the number of CAG repeats at the *ATXN3* was performed for all MJD subjects studied. The homogeneous MJD cluster comprising a total of 88 patients was characterized. The average of the neurological examination score for spinocerebellar ataxia (NESSCA) available for 69 patients was  $12 \pm 5$  (mean  $\pm$  standard deviation); 40% of the variance explanation is provided by the number of the CAG repeats in the expanded allele and disease duration. Ocular alterations, namely nystagmus, are clinical features commonly described in MJD. The evaluation of nystagmus in the early stages of the disease was undertaken using asymptomatic carriers and non-carriers of *ATXN3* mutation. The frequency of nystagmus in asymptomatic carriers (17%) and its absence in non-carriers of the mutation suggests that nystagmus may appear before gait disturbance and can thus be considered an early sign of MJD. A candidate gene approach was used in this work to pursuit

modifiers of the MJD onset. The CAG repeat number in loci associated with polyQ disorders, namely spinocerebellar ataxias (SCA) types 1, 2, 6, 7, 17, Huntington disease (HD), and dentato-rubro-pallidolusian atrophy (DRPLA) was investigated as modulators of the age at onset in the Azorean MJD cohort; a negative correlation was found between the longer allele at *ATXN1* and disease onset, significantly increasing by 1.5% the explanation of onset variance. Furthermore, and because a role of inflammation in MJD has been postulated, promoter variants leading to alterations in cytokines expression were studied, and their impact in the disease onset was analysed. Patients carrying the *IL6*\*C allele had a significant earlier onset. A complex net of mechanisms, which includes transcriptional regulation, ubiquitin-proteasome system, autophagy, apoptosis as well as mitochondrial function, has been implicated in MJD pathogenesis. Results from the transcriptome-wide gene expression analyses, followed by several steps of validation, showed that even in a peripheral tissue, transcription dysregulation was detectable and that *FCGR3B*, *P2RY13* and *SELPLG* were dysregulated in patients. To further investigate dysregulated mRNA levels as a way to identify candidate biomarkers, the expression patterns of nine candidate genes - *HSPB1*, *DNAJB1*, *DNAJB12*, *DNAJB14*, *BAX*, *BCL2*, *SOD2*, *IL1B* and *IL6* genes were quantified. Decreased levels of *IL6* and *BCL2* mRNA were observed in preataxic subjects. Lower *HSPB1* and *BCL2* mRNA levels after the ataxia onset were, furthermore, evidenced. A significant decrease of *BCL2* mRNA levels during disease progression was observed in all patients used in the exploratory longitudinal study. Mitochondrial DNA (mtDNA) damage and depletion has been inconsistently reported in MJD. In this work the mtDNA common deletion was significantly more frequent in patients and preataxic carriers than in controls. In conclusion, although the clinical and physiological phenotype of MJD reflects mainly neuronal damage, novel peripheral molecular alterations were described in the present dissertation. The confirmation of the link between peripheral abnormalities and neurodegeneration offers a new window of opportunity for biomarker discovery. More steps of validation will be crucial to develop a battery of molecular biomarkers able to describe disease progression, as well as to be used as outcome measures in future preventive and/or disease-modifying clinical trials. Data from molecular biomarkers should also be further correlated with imaging and clinical measures. Efforts to test the promising molecular biomarkers in independent cohorts, as well as to investigate their potential in longitudinal studies are warranted.

## RESUMO

A doença de Machado-Joseph/ataxia espinocerebelosa do tipo 3 (DMJ/SCA3) é uma doença neurodegenerativa rara e de início tardio, que permanece sem tratamento. À semelhança do que acontece nas restantes doenças de poliglutamina (poliQ), a mutação na base da DMJ corresponde à expansão de um motivo CAG na região codificante do respetivo gene causal, *ATXN3*. Sendo uma doença progressiva, a DMJ é caracterizada por várias fases, ao longo da sua progressão, nomeadamente por um longo período pré-clínico; durante esse período deverão ocorrer alterações moleculares e até mesmo clínicas, as quais podem surgir muitos anos antes da ataxia. A idade de início da ataxia, usualmente o primeiro sintoma da doença, é maioritariamente explicada pelo tamanho do tracto CAG que constitui assim o principal biomarcador de diagnóstico molecular da DMJ. Genes modificadores deverão, adicionalmente, contribuir para a explicação da idade de início da doença; tais modificadores genéticos, podem também ser considerados biomarcadores de diagnóstico da DMJ. A identificação deste tipo de biomarcadores permitirá: (1) melhorar a previsão da idade de início da doença e (2) aumentar o poder estatístico em ensaios clínicos, permitindo a estratificação dos doentes usando informação proveniente dos modificadores genéticos. As diferentes fases da DMJ são atualmente caracterizadas recorrendo à utilização de informação clínica (principalmente de escalas clínicas) ou de dados obtidos a partir de estudos de imagem, sendo o conhecimento acerca das alterações moleculares que deverão ocorrer ao longo da doença muito incipiente, o que se reflete na ausência de biomarcadores de monitorização da doença. A identificação deste tipo de biomarcadores permitirá: (1) a previsão da idade de início patológica (idade à qual alterações ao nível celular já estão presentes, mas o doente ainda não apresenta sintomas), (2) a monitorização das diferentes fases da progressão da doença e (3) a detecção de benefícios terapêuticos subtis, em futuros ensaios clínicos. O objetivo principal desta dissertação foi contribuir para o desenvolvimento de biomarcadores moleculares da DMJ. RNA e DNA, obtidos a partir de amostras de sangue de indivíduos pré-clínicos (portadores da mutação *ATXN3*, mas sem sintomatologia) e doentes DMJ de origem açoriana, incluindo amostras colhidas em dois momentos diferentes, constituíram o principal recurso para o desenvolvimento deste projeto de doutoramento. Foram também utilizadas amostras de indivíduos saudáveis da mesma população, que foram usados como controlo. O número de repetições CAG no gene *ATXN3* foi determinado para todos os indivíduos DMJ. A coorte DMJ açoriana, constituída por um total de 88 doentes, foi caracterizada. Para 69 dos doentes

avaliados, utilizando a escala clínica NESSCA (*Neurological Examination Score for Spinocerebellar Ataxia*), a média da pontuação foi de  $12 \pm 5$  (média  $\pm$  desvio padrão); 40% da variância desta pontuação é explicada pelo número de repetições CAG no alelo expandido e pela duração da doença. As alterações oculares, nomeadamente a presença de nistagmo, fazem parte do conjunto de características clínicas habitualmente descritas na DMJ. A avaliação do nistagmo em fases iniciais da doença, nomeadamente na fase pré-clínica, foi realizada em portadores assintomáticos e não-portadores da mutação *ATXN3*. A frequência do nistagmo em portadores assintomáticos (17%) e sua ausência em não-portadores da mutação sugere que o nistagmo pode preceder a ataxia e, portanto, ser considerado um sinal precoce da DMJ. Nesta dissertação, foi utilizada uma abordagem de genes candidatos para identificar modificadores da idade de início da DMJ. O número de repetições CAG em *loci* de doenças polyQ, nomeadamente as ataxias espinocerebelosas (SCA) dos tipos 1, 2, 6, 7, 17, a doença de Huntington (HD) e a atrofia dentato-rubro-pálido-luisiana (DRPLA) foi investigado como potencial modificador da idade de início em doentes açorianos; observou-se uma correlação negativa entre o alelo normal maior no gene *ATXN1* e a idade de início da doença, sendo que a presença deste alelo aumentou significativamente em 1,5% a explicação da variância da idade de início. Adicionalmente, e por ter sido descrito na DMJ um papel da inflamação, foram estudadas, como modificadores da idade de início, variantes alélicas na região promotora de genes de citocinas, uma vez que estas variações podem alterar os seus níveis de expressão. Os doentes que possuíam o alelo *IL6\*C* apresentaram uma idade de início significativamente mais precoce. Um conjunto complexo de mecanismos, que inclui a regulação da transcrição, o sistema ubiquitina-proteassoma, a autofagia, a apoptose e a função mitocondrial, tem sido implicado na patogénese da DMJ. Os resultados de uma análise de expressão *whole-genome*, seguida de vários passos de validação, mostraram que mesmo num tecido periférico, a desregulação da transcrição estava alterada e que os genes *FCGR3B*, *P2RY13* e *SELPLG* estavam desregulados em doentes. Utilizando a desregulação dos níveis de mRNA como argumento para identificar novos candidatos a biomarcadores, foram analisados os padrões de expressão de nove genes - *HSPB1*, *DNAJB1*, *DNAJB12*, *DNAJB14*, *BAX*, *BCL2*, *SOD2*, *IL1B* e *IL6*. Observou-se que em indivíduos pré-clínicos os níveis de mRNA dos genes *IL6* e *BCL2* eram mais baixos relativamente aos controlos. Em indivíduos já com sintomatologia, foram observados níveis mais baixos de mRNA dos genes *HSPB1* e *BCL2*, comparativamente aos controlos. No estudo exploratório longitudinal foi observada uma diminuição significativa dos níveis de mRNA do gene *BCL2* durante a progressão da doença. Resultados inconsistentes sobre o dano e a depleção

do DNA mitocondrial (mtDNA) em doentes DMJ têm sido reportados. Neste trabalho, verificou-se que os doentes e os portadores pré-clínicos apresentavam mais deleções do mtDNA (*common deletion*) do que os controlos. Embora as características clínicas e fisiológicas inerentes à DMJ reflitam principalmente alterações do processo neurodegenerativo, alterações moleculares quantificadas em sangue foram descritas nesta dissertação. A confirmação de uma relação entre as alterações periféricas e a neurodegeneração constitui uma nova oportunidade para o desenvolvimento de biomarcadores. Passos adicionais de validação serão cruciais para desenvolver um conjunto de biomarcadores moleculares capazes de monitorizar a progressão da doença, bem como de detectar alterações terapêuticas subtis em futuros ensaios clínicos, incluindo ensaios preventivos. Os biomarcadores moleculares devem também ser capazes de se correlacionar com marcadores clínicos e de imagiologia. Todos os esforços serão realizados no sentido de testar os biomarcadores moleculares aqui propostos em coortes independentes, bem como de investigar o seu potencial em estudos longitudinais.