

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL**

**Avaliação das repetições hexanucleotídicas no gene *C9orf72* como possível modificador de idade de início da doença em pacientes com doença de Machado-Joseph**

Yelena Perevalova

*Dissertação submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da UFRGS como requisito parcial para a obtenção do grau de Mestre em Genética e Biologia Molecular*

Orientadora: Maria Luiza Saraiva-Pereira

**Porto Alegre**

**Abril, 2018**

## **Realização do Trabalho e Financiamento**

Este trabalho foi realizado no Laboratório de Identificação Genética (LIG) do Centro de Pesquisa Experimental (CPE) e no Serviço de Genética Médica (SGM) do Hospital de Clínicas de Porto Alegre (HCPA), com apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e Fundo Incentivo de Pesquisa HCPA (FIPE-HCPA).

## Sumário

<b>REALIZAÇÃO DO TRABALHO E FINANCIAMENTO</b>	<b>i</b>
<b>SUMÁRIO</b>	<b>ii</b>
<b>RESUMO</b>	<b>v</b>
<b>ABSTRACT</b>	<b>vi</b>
<b>I. INTRODUÇÃO</b>	<b>1</b>
<b>1. Doenças Neurodegenerativas</b>	<b>1</b>
<b>2. Ataxias Espinocerebelares</b>	<b>2</b>
<b>3. Doença de Machado-Joseph ou Ataxia Espinocerebelar tipo 3</b>	<b>3</b>
<b>a. Manifestações Clínicas</b>	<b>3</b>
<b>b. Bases Genéticas</b>	<b>4</b>
<b>c. Bases Moleculares</b>	<b>6</b>
<b>d. Efeitos Modificadores de Idade de Início da Doença de Machado-Joseph</b>	<b>7</b>
<b>4. C9orf72</b>	<b>11</b>
<b>a. Estrutura e Função do gene <i>C9orf72</i></b>	<b>12</b>
<b>b. Expansão GGGGCC</b>	<b>12</b>
<b>c. Esclerose Lateral Amiotrófica e Demência Frontotemporal</b>	<b>15</b>
<b>5. Justificativa</b>	<b>16</b>
<b>II. OBJETIVOS</b>	<b>18</b>

<b>III. RESULTADOS</b>	<b>19</b>
<i>C9orf72</i> gene and Cerebellar Ataxia: A Tenuous Relationship	20
Analysis of <i>C9orf72</i> hexanucleotide repeats length as a potential modifying factor of age of onset in Machado-Joseph Disease/Spinocerebellar Ataxia Type 3 patients	34
<b>IV. DISCUSSÃO E CONCLUSÕES</b>	<b>53</b>
<b>V. REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>55</b>

## Lista de Abreviaturas, Símbolos e Unidades

A	adenina
ALS	esclerose lateral amiotrófica ( <i>Amyotrophic lateral sclerosis</i> )
ii	idade de início da doença ( <i>age of onset</i> )
C	citosina
DM1	distrofia miotônica do tipo 1
DNA	ácido desoxirribonucleico ( <i>desoxyribonucleic acid</i> )
ER	retículo endoplasmático ( <i>endoplasmatic reticulum</i> )
FTD	demência frontotemporal ( <i>Frontotemporal dementia</i> )
G	guanina
MJD	doença de Machado-Joseph ( <i>Machado-Joseph disease</i> )
mRNA	RNA mensageiro ( <i>messenger RNA</i> )
polyQ	poliglutamina
RNA	ácido ribonucleico ( <i>ribonucleic acid</i> )
SCA	ataxia espinocerebelar ( <i>spinocerebellar ataxia</i> )
SNP	polimorfismo de nucleotídeo único ( <i>single nucleotide polymorphism</i> )
T	timina
UTR	região não-traduzida ( <i>untranslated region</i> )

## Resumo

A doença de Machado-Joseph (MJD – *Machado-Joseph disease*), também conhecida como ataxia espinocerebelar tipo 3 (SCA3 – *spinocerebellar ataxia type 3*), é uma doença neurodegenerativa, com padrão de herança autossômica dominante, causada por uma expansão do trato de repetições CAG no gene *ATXN3*. Estudos prévios mostram que a idade de início (ii) da doença apresenta uma correlação inversa ao número de repetições CAG no alelo expandido, mas que somente entre 40 e 68% da variação da ii pode ser explicado pelo tamanho da expansão. Considerando que um mecanismo molecular comum para todas as doenças neurodegenerativas foi proposto anteriormente, o restante da variação da correlação pode estar associado ao mecanismo molecular de outras doenças. Expansão da repetição hexanucleotídica (GGGGCC ou G<sub>4</sub>C<sub>2</sub>) no gene *C9orf72* foi identificado como a etiologia em uma proporção significativa dos pacientes de esclerose lateral amiotrófica (ALS – *amyotrophic lateral sclerosis*), demência frontotemporal (FTD – *frontotemporal dementia*) e os afetados da comorbidade ALS-FTD. O presente estudo teve como objetivo investigar o tamanho do trato de hexanucleotídeo no gene *C9orf72* em pacientes com MJD/SCA3 como um potencial modificador de ii. A genotipagem dos alelos do *C9orf72* foi realizada por PCR com primers fluorescentes, seguida de eletroforese capilar, e avaliada no grupo teste e no grupo controle. Foram genotipados 83 pacientes com MJD/SCA3 e 102 controles saudáveis sem história de ataxia. Todos os pacientes e controles foram recrutados no Hospital de Clínicas de Porto Alegre. A expansão patogênica do *C9orf72* foi definida como trato contendo mais de 30 repetições hexanucleotídicas. Os indivíduos foram divididos em duas categorias para fins de análises estatísticas: grupo 1) indivíduos com ambos alelos contendo até 6 repetições, inclusive (alelo “pequeno”) e grupo 2) os indivíduos com pelo menos um alelo de 7 até 30 repetições, inclusive (alelo “intermediário”). Nenhum indivíduo com expansão patogênica foi detectado. A frequência alélica de G<sub>4</sub>C<sub>2</sub> foi estabelecida e nenhuma diferença foi observada entre elas. Além disso, não foi encontrada correlação do tamanho do alelo com pacientes de ii precoce e nem com os pacientes de ii tardia. Portanto, nossos resultados indicam que o tamanho do alelo do *C9orf72* não atua como fator modificador da idade de início da doença em pacientes com MJD/SCA3.

## Abstract

Machado-Joseph disease (MJD), also known as spinocerebellar ataxia type 3 (SCA3) is a neurodegenerative disorder of an autosomal dominant trait caused by an expansion of a CAG repeat tract in the *ATXN3* gene. Previous studies show that age of onset (AO) of the disease is inversely correlated with CAG repeat length in the expanded allele that can explain from 40 to 68% of variation in AO. Considering that a common molecular mechanism for all neurodegenerative diseases has been previously suggested, the key to describing this variation could lie in the molecular biology of other diseases. Hexanucleotide repeat (GGGGCC or G<sub>4</sub>C<sub>2</sub>) expansions in *C9orf72* were identified as the etiology in a significant portion of patients with amyotrophic lateral sclerosis (ALS), patients with frontotemporal dementia (FTD), and those affected by comorbid ALS-FTD. Taking this into consideration, our aim was to investigate *C9orf72* hexanucleotide tract length in MJD/SCA3 patients as a potential modifier of AO. After establishing a reliable protocol to genotype *C9orf72* alleles using PCR with fluorescent primers followed by capillary electrophoresis, 83 MJD/SCA patients and 102 healthy controls with no history of ataxia were genotyped. All individuals were recruited from the Hospital de Clínicas de Porto Alegre. Pathogenic expansions were defined as alleles containing more than 30 hexanucleotide repeats. Individuals were divided into two categories for statistical analysis: group 1) those with both alleles containing up to 6 repeats (“small” alleles) and group 2) those with one or both alleles of 7 up to 30 repeats (“intermediate” alleles). No pathogenic expansions were detected. Allelic distribution was established, and no difference was observed between groups. In addition, neither early-onset nor late-onset MJD/SCA3 patients have shown to be correlated with neither the small nor the intermediary *C9orf72* genotype so far. Therefore, results obtained in this study indicate that *C9orf72* allele length is not a modifier of age of onset in MJD/SCA3 patients.

## I. Introdução

### 1. Doenças Neurodegenerativas

As doenças neurodegenerativas constituem uma família de patologias reconhecidas mais recentemente, sendo que as primeiras publicações descrevendo uma série de doenças como “neurodegenerativas” surgiram só na segunda metade da década de 60 do século passado (Jennekens 2014). Elas representam um conjunto de condições crônicas heterogêneas cuja consequência é a perda progressiva da funcionalidade e/ou integridade estrutural dos neurônios (Przedborski et al. 2003). A grande maioria dessas doenças se manifestam tardiamente, o que dificultou o seu estudo no passado pelo fato que os indivíduos não sobreviviam até a idade de início da doença (ii). Atualmente, com o envelhecimento da população mundial, as doenças neurodegenerativas apresentam um grande interesse para a saúde pública mundial, com previsão da Organização Mundial de Saúde indicando que, até o ano 2040, as doenças neurodegenerativas vão superar o câncer como a segunda maior causa de morte natural (Gammon 2014).

Devido ao fato dos estudos das doenças neurodegenerativas serem recentes, da dificuldade de obtenção de material biológico do sistema nervoso e das limitações de metodologias, a utilização de genética em estudos moleculares das doenças neurodegenerativas só entrou na pauta em 1979 (Young 2009), com uma discussão do uso de marcadores genéticos para tentar localizar o gene responsável pela doença de Huntington. O resultado foi a identificação de uma região polimórfica do cromossomo 4 e, posteriormente, a identificação do gene conhecido atualmente como *HTT* (Gusella et al. 1983), dando início a era da neurogenética. Mesmo assim, a maioria das doenças neurodegenerativas ainda não tem sua etiologia precisamente descrita. O mecanismo exato da neurodegeneração ainda não foi completamente elucidado, mas, a cada ano, surgem mais evidências de uma via comum, ou pelo menos semelhante molecularmente, entre todas as doenças neurodegenerativas (Lin and Beal 2006; Jellinger 2010; Haass 2010; Avila 2010; Durrenberger et al. 2015).

A classificação dessas doenças apresenta uma série de dificuldades, sendo que existe uma grande sobreposição do quadro clínico e patológico entre doenças mesmo de famílias



moleculares bastante diferentes. Existe uma série de critérios utilizados para identificar e classificar a doença neurodegenerativa de um paciente, levando em consideração a topologia da lesão, as características clínicas e, quando conhecida, a etiologia (Kovacs 2016). É a partir dessa classificação, que surgiu um grupo de doenças do trato de poliglutaminas (doenças poliQ). Essas doenças, na sua maioria, apresentam padrão de herança autossômica dominante, com ii na idade média adulta e a taxa progressiva da doença (Orr and Zoghbi 2007). Essas doenças são causadas por uma expansão do trinucleotídeo CAG em um gene causador, o qual será posteriormente traduzido por um trato de poliQ nas proteínas, havendo a formação de aglomerados proteicos tóxicos para células. É importante notar que mesmo com esse fator descrito, há evidências que só a síntese da proteína mutada não é suficiente para causar a doença (Katsuno et al. 2006; Li and Li 2006; Thompson 2008).

## 2. Ataxias Espinocerebelares

As ataxias espinocerebelares (SCA – *spinocerebellar ataxia*) formam uma família constituída por dezenas de doenças heterogêneas molecularmente e clinicamente semelhantes, as quais só podem ser corretamente identificadas através do uso de exames moleculares (Orr and Zoghbi 2007). De acordo com os critérios descritos anteriormente, podemos descrever as SCAs como desordens genéticas hereditárias (na sua maioria autossômica dominantes), resultando em atrofia cerebelar, com o paciente exibindo marcha atáxica e incoordenação nos movimentos dos membros, na fala e na movimentação ocular, entre outros sintomas (Teive 2009; Jayadev and Bird 2013). Apesar da falta de estudos abrangentes sobre a prevalência mundial das SCAs, existe um grande número de estudos regionais (Ruano et al. 2014).

A patogênese das SCAs é variável. Entretanto, podemos dividi-las em quatro categorias: ataxias poliQs, disfunção dos canais iônicos, alteração nos sinais de transdução e repetições não-codificantes (provavelmente levando à toxicidade de RNA nos neurônios) (Shakkottai and Fogel 2013), sendo que as ataxias poliQs são as mais prevalentes. Tipicamente, esse tipo de SCA tem como etiologia uma expansão do tipo (CAG)<sub>n</sub> localizada na parte codificante do gene (Trott and Houenou 2012). Mas existem exceções, como no caso SCA8, onde é uma expansão (CTG)<sub>n</sub> na região não-codificante que resulta em uma proteína poliglutamina pura lida a partir da fita

complementar. De todas as formas de SCAs, somente 16 têm um gene associado conhecido, indicando a dificuldade de identificação dos fatores genéticos nessas doenças, além da sua provável natureza multifatorial (Teive et al. 2012).

Pelo menos seis tipos diferentes de SCAs (SCA1, SCA2, SCA3, SCA6, SCA7, SCA10) foram identificadas no Brasil, além das famílias sem diagnóstico molecular (de Castilhos et al. 2014). A SCA3, também conhecida como doença de Machado-Joseph (MJD – *Machado-Joseph disease*), é a mais prevalente, sendo responsável por aproximadamente 60% dos pacientes afetados por SCAs no Brasil (de Castilhos et al. 2014; Cintra et al. 2014). No nosso estado (Rio Grande do Sul), a prevalência de MJD/SCA3 é ainda mais marcante (Jardim et al. 2001c; Ruano et al. 2014).

### 3. Doença de Machado-Joseph ou Ataxia Espinocerebelar tipo 3

#### a. Manifestações Clínicas

A MJD/SCA3, foi identificada na década de 70 por três grupos independentes, estudando famílias residentes no Estados Unidos com ascendência açoriana. Os patriarcas de duas dessas famílias, William Machado (Nakano et al. 1972) e Antone Joseph (Rosenberg et al. 1976), foram imortalizados no nome da doença. E um terceiro grupo foi descrito como a doença “degeneração nigro-espino-dentatal com oftalmoplegia” (Woods and Schaumburg 1972). O quadro clínico da doença é bastante variável, com a manifestação principal sendo a ataxia cerebelar, mas a sintomatologia inclui também espasticidade (Dürr et al. 1996), dificuldade na deglutição (Isono et al. 2013), dificuldade na articulação de fala (Wolf et al. 2017), comprometimento cognitivo (Kawai et al. 2004), entre outros (Jardim et al. 2001a; Jardim et al. 2001b; Lee et al. 2003; Vale et al. 2010; González-Zaldívar et al. 2015). Essa grande variabilidade contribuiu para a dificuldade na classificação dessa doença, junto com sua identificação inicial.

Em 1978, Coutinho e Andrade fizeram a tentativa mais completa em descrever a doença. Foram identificadas três categorias da MJD/SCA3, as quais se distinguem pelos sinais extra-piramidais e/ou sinais periféricos presentes. Os sinais cerebelares, tais como ataxia, disartria e oftalmoplegia externa progressiva, podem ser encontrados em quase todos os pacientes. Os

pacientes MJD/SCA3 tipo 1 apresentam sinais piramidais e extrapiramidais marcantes, tais como hiperreflexia, sinal de Babinski, distonia. Os pacientes do tipo 3 apresentam, na sua maioria, alterações periféricas, tais como dor nos membros, entorpecimento. Por fim, os pacientes do tipo 2 da doença apresentam uma sintomatologia intermediária, sem sinais neurológicos intensos observados, e podem se transformar no tipo 1 ou no tipo 3 após um período prolongado. Os três tipos apresentam ii diferentes, com o tipo 1 se caracterizando por manifestação mais precoce, ao redor de 25 anos, e os tipos 2 e 3 desenvolvendo os primeiros sintomas na quarta década de vida, mas o com tipo 3 representando o subtipo mais tardio, começando mais perto de cinquenta anos (Coutinho and Andrade 1978).

Com o aumento dos estudos nessas doenças raras, foram identificados mais dois subtipos da doença. Um quarto tipo foi relatado em poucas famílias com sintomas de parkinsonismo, como tremores, rigidez e bradicinesia (Tuite et al. 1995; Gwinn-Hardy et al. 2001; Buhmann et al. 2003; Bettencourt et al. 2011b). Esses pacientes, diferentemente dos outros subtipos da MJD/SCA3, respondem bem ao tratamento com levodopa. Em 2009, um possível quinto tipo da MJD/SCA3 foi escrito por um grupo chinês, com sintomas imitando paraplegia espástica hereditária (Wang et al. 2009). Mas os dados referentes a esse tipo da doença ainda não foram confirmados, nem aceitos de forma abrangente.

#### b. Bases Genéticas

A identificação do mecanismo molecular responsável pela MJD/SCA3 é de grande interesse, tanto científico como para o desenvolvimento de terapias e tratamentos. Estudos realizados por Stevanin e colaboradores em 1993 eliminaram os genes causadores de SCA1 e SCA2 como candidatos etiológicos da MJD, que eles denominaram SCA3, por convenção, em um estudo posterior (Stevanin et al. 1995). Esse precedente, exacerbado pela alta heterogeneidade clínica, gerou uma série de confusões sobre a identidade de SCA3 e MJD, necessitando a investigação na cognição dessas síndromes aparentemente clinicamente distintas, mas molecularmente parecidos. Depois de uma série de publicações, MJD e SCA3 foram confirmados como sendo a mesma doença (Matilla et al. 1995; Haberhausen et al. 1995), mas por convenção, até hoje, os dois nomes estão utilizados para denominar essa doença

neurodegenerativa específica. Tomando como exemplo pesquisa em outras doenças neurodegenerativas, o primeiro passo na identificação do *locus* responsável para o desenvolvimento da MJD/SCA3 foi a utilização do mapeamento genético por análise de ligação, com dois grupos identificando uma região do braço longo do cromossomo 14 como associado à doença (Takiyama et al. 1993; St George-Hyslop et al. 1994). Como várias desordens neurodegenerativas e multisistêmicas, como a SCA1 e a doença de Huntington, foram relacionados com expansões trinucleotídicas, esse mecanismo foi uma das linhas principais para investigação na MJD/SCA3. Em 1994, Kawaguchi e colaboradores localizam um trato de repetições CAG na região previamente associada com MJD/SCA3 (Kawaguchi et al. 1994).

Posteriormente, em 2001, Ichikawa e colaboradores descreveram a estrutura do gene *MJD*, hoje em dia conhecido como *ATXN3*, e a expressão desse gene (Ichikawa et al. 2001). A estrutura do gene foi determinada como sendo composta por 11 éxons, intercalados por 10 íntrons, com a expansão CAG localizada no início do décimo éxon. Essa descoberta abriu as portas para investigações iniciais do mecanismo proteico que causa a doença, sendo que a expansão nucleotídica, quando traduzida, gera um trato poliQ estendido. Em termos da expressão, análises por *Northern Blot* mostraram expressão ubíqua do RNA do *ATXN3*, em espécies de vários tamanhos, indicando possíveis sítios de *splicing* no gene ou poliadenilação diferenciada (Ichikawa et al. 2001).

Geneticamente falando, MJD/SCA3 canonicamente apresentava uma doença autossômica dominante de penetrância completa. Estudo abrangente de 149 pacientes afetados, 204 indivíduos em risco e 320 indivíduos normais delimitou a definição do alelo “normal” como apresentando entre 12 e 44 repetições, enquanto os alelos patogênicos foram descritos como aqueles contendo entre 61 e 87 repetições (Maciel et al. 2001). Nesse mesmo trabalho, foi descrito o caso de um indivíduo com 51 repetições sem manifestações clínicas, indicando que um alelo desse comprimento não expressa o fenótipo MJD/SCA3. Entretanto, a questão dos alelos intermediários gerou resultados inconclusivos e estão ainda mais complicados por serem raros. Ao contrário do trabalho de Maciel e colaboradores, uma série de outros estudos mostraram que os alelos intermediários são capazes de desenvolver a doença (van Alfen et al. 2001; Gu et al. 2004), sendo o alelo com 45 repetições o menor alelo patogênicos descrito (Padiath et al. 2005).

É importante notar que essa expansão CAG não é considerada estável; ou seja, a tendência da expansão nessa faixa é aumentar seu tamanho para a próxima geração (Maruyama et al. 1995). O alelo normal é transmitido de forma conservadora, sem mudança, enquanto o alelo mutante tende a acumular repetições (Igarashi et al. 1996). Esse fenômeno molecular é a causa de antecipação: em uma família, a idade de início da doença é mais precoce na próxima geração (Klausen 2007).

### c. Bases Moleculares

A principal característica marcante das doenças poliQs são as chamadas inclusões proteicas formadas por aglomeração de proteínas poliQ mutantes (Zoghbi and Orr 1999). No caso da MJD/SCA3, a proteína que forma esses aglomerados é a ataxina-3, uma proteína expressa de forma ubíqua nos tecidos do sistema nervoso humano (Nishiyama et al. 1996). A função específica dessa proteína não é conhecida, mas modelos em camundongo *knock-out* e análises estruturais indicam uma provável função relacionada com a ubiquitinação (Chow et al. 2004; Schmitt et al. 2007), especificamente relacionando a ataxina-3 com processos de degradação de proteínas (Burnett et al. 2003). Além disso, a interação de tratos poliQ, e especificamente o trato da ataxina-3, com a maquinaria de regulação de transcrição está bem documentada (Shimohata et al. 2000; McCampbell et al. 2000; Li et al. 2002). Dificultando mais a situação, modelos *knock-out* em *Caenorhabditis elegans* não demonstraram diferença de comportamento, nem na expectativa de vida (Rodrigues et al. 2007). Em termos da localização, a ataxina-3 consegue se deslocar entre o núcleo e o citoplasma da célula, indicando sua alta potência transportadora (Chai et al. 2002).

A primeira indicação que a proteína mutante tem provável papel no desenvolvimento da doença entra nessa mesma questão de localização: nos pacientes afetados, a ataxina-3 preferencialmente desloca para o núcleo, formando inclusões (Schmidt et al. 1998). Esse efeito da proteína mutada parece ser ao mesmo tempo crítico para o desenvolvimento da doença e, ao mesmo tempo, ter grande poder neuroprotetor. Estudos com a proteína huntingtina, uma proteína poliQ que responsável, quando expandida, pela doença de Huntington, demonstra que a introdução de aglomerados de peptídeos poliQ no núcleo de neurônios leva até a degeneração

das células, ou seja, eles são altamente tóxicos para células (Yang et al. 2002). Ao mesmo tempo, existem evidências que as inclusões de proteínas poliQ não levam até a morte celular, mas protegem a célula, provavelmente devido ao sequestro de proteína mutante (Arrasate et al. 2004). Esses resultados aparentemente contraditórios demonstram a complexidade da questão da neurodegeneração, e indicam a provável existência de fatores adicionais atuando nesse processo, além da simples inclusão de agregados poliQ.

#### d. Efeitos Modificadores de Idade de Início da Doença de Machado-Joseph

A idade de início da MJD/SCA3 apresenta grande interesse para indivíduos afetados e para pesquisadores investigando a doença por causa da sua grande variabilidade. As primeiras tentativas de descrever os mecanismos que determinam a ii se baseavam, de novo, nos estudos de outras doenças neurodegenerativas. Essas tentativas foram realizadas inicialmente por dois grupos independentes que consideravam o tamanho da expansão, ou seja, o número de repetições CAG, como o principal fator determinante na ii. Em 1995, Maciel e colaboradores descreveram correlação inversa entre o tamanho do alelo expandido e a ii, sendo que alelos menores correspondiam a ii mais tardias (Maciel et al. 1995). Esse estudo determinou que esse fator é capaz de explicar somente cerca de 45,8% da variabilidade da ii em pacientes com MJD/SCA3 em uma amostra de pacientes de origem portuguesa, norteamericana e brasileira. No mesmo ano, Maruyama e colaboradores demonstraram que, em uma população japonesa, esse mesmo fator explicava ao redor de 76% de variação da ii (Maruyama et al. 1995).

Essa inconsistência da influência de tamanho da expansão continuou a ser observado em estudos posteriores nessa área. O estudo de uma coorte holandesa, em 2002, demonstrou que o número de repetições contribuiu entre 52 até 76% na variabilidade da ii da doença (van de Warrenburg et al. 2002), enquanto que, em uma população cubana, esse mesmo fator contribuiu até 81% (González-Zaldívar et al. 2015). Um estudo recente realizado no estado do Rio Grande do Sul apontou uma influência relativamente menor do tamanho de expansão nos pacientes dessa região, sendo que somente 62,1% da variação foi explicada por esse fator (Saute and Jardim 2015). Essas informações demonstram uma variação da influência da expansão em populações distintas, indicando a provável existência de fatores adicionais que atuam no

desenvolvimento de doença, levando às discussões se esse fator é genético ou ambiental (DeStefano et al. 1996).

Na investigação de fatores genéticos, o primeiro suspeito foi o tamanho do alelo normal do *ATXN3*. Um dos primeiros estudos desse aspecto foi realizado em 1996 por Dürr e colaboradores, utilizando uma grande amostragem de pacientes e controles com origens geográficas globais distintas (Dürr et al. 1996). Em uma coorte de 91 pacientes afetados com ii descrita, o alelo expandido representou uma contribuição de 64% para variação da ii. A inclusão de alelo “normal” na análise acrescentou mais 6%, sendo que o sentido de correlação continuou sendo a mesma, ou seja, alelos não-expandidos maiores corresponderam a uma ii mais precoce. No Brasil, um estudo abrangente confirmou esses resultados, com 59,6% e 6,01% de influência do alelo expandido e alelo normal, respectivamente (França et al. 2012). Mas, mesmo levando em conta esse fator, o tamanho do alelo normal apresenta uma contribuição pequena e ainda permanece uma grande porção da variabilidade sem uma causa específica.

Uma outra área de investigação é a pesquisa de variabilidade nas regiões regulatórias e seu efeito sobre a ii da doença. Long e colaboradores identificaram dois polimorfismos de nucleotídeos únicos (SNP – *single nucleotide polymorphism*) na região do 3’UTR como potenciais modificadores de ii em MJD/SCA3 (Long et al. 2015). Em uma população chinesa, a variação da ii foi pouco influenciada pelo alelo expandido (<50%), direcionando os pesquisadores para a busca de outros fatores que poderiam influenciar a ii nesses pacientes. O alelo A do rs709930 e o alelo T do rs910369 foram detectados mais frequentemente em pacientes com MJD/SCA3 quando comparado com controles saudáveis. Além disso, eles foram capazes em diminuir a ii de 2 até 4 anos, mostrando um efeito danoso para os pacientes e contribuindo com 2% na variação da ii. Análises semelhantes realizadas na região 5’UTR não identificaram SNPs clinicamente significativos (Bettencourt et al. 2012). Esse estudo encontrou um único SNP (rs3814834) nessa região e tentativas de caracterizar uma função molecular para essa variante não apresentaram resultados.

Além dos SNPs, fatores epigenéticos presentes nas regiões reguladoras foram levantados como possíveis modificadores da doença. A hipermetilação das ilhas de CpG na região promotora foi identificada como potencial modificadora da doença, mas as análises sempre ficam prejudicadas pela influência do efeito do alelo expandido. A redução de 10% da metilação

foi associada com redução de até 2,4 anos na ii, mas esse efeito praticamente desaparece quando considerado em conjunto com o tamanho da expansão e do alelo normal (Emmel et al. 2011). Portanto, não foi possível confirmar ou excluir completamente o efeito da metilação na região promotora do *ATXN3* em pacientes com MJD/SCA3. Estudos mais recentes, realizados em uma população de família chinesas, avaliou o grau de metilação da região promotora do *ATXN3* em irmãos com o mesmo tamanho da expansão, mas com ii diferentes. Dessa forma, a influência dos fatores genéticos já conhecidos foi eliminada e o efeito ambiental minimizado. Os resultados obtidos mostraram níveis de metilação significativamente maiores nos indivíduos com ii mais precoce (Wang et al. 2017), contradizendo os resultados do trabalho anterior. Esses fatos indicam que variações populacionais podem atuar de forma distinta como fatores moduladores na MJD/SCA3.

Além do alto número de investigações realizadas em fatores intragênicos, os potenciais efeitos modificadores de outros genes, além do *ATXN3*, estão sendo amplamente avaliados. Os genes associados a outras doenças neurodegenerativas são alvos desses estudos, com evidências crescentes que apontam para o fato que a biologia molecular da neurodegeneração de uma forma geral é um mecanismo compartilhado ou, pelo menos semelhante, em várias doenças neurodegenerativas (Durrenberger et al. 2015).

Os primeiros genes investigados foram os causadores de outras doenças poliQ, devido ao fato da mutação pertencer ao mesmo grupo e o quadro clínico ser bastante semelhante. Um estudo abrangente do EUROSCA, um consórcio europeu dedicado à pesquisa de SCAs, foi realizado com o objetivo de localizar genes modificadores de doença compartilhados entre as SCAs e outras doenças de expansões trinucleotídicas (Tezenas du Montcel et al. 2014). O estudo foi conduzido utilizando amostras de 1255 pacientes e incluiu oito genes que causam doenças neurodegenerativas por expansões trinucleotídicas. As análises realizadas pelo grupo demonstraram que três genes atuaram como modificadores para MJD/SCA3. Um deles foi o gene *ATXN2*, o qual está associado à ataxia espinocerebelar do tipo 2 (SCA2) (Pulst et al. 1996). Os alelos de tamanho intermediário nesse gene, definidos como aqueles contendo entre 27 e 29 repetições CAG, foram associados com ii precoce quando comparados com alelos *ATXN2* menores. O outro modificador identificado foi o gene *ATN1*, cuja expansão causa a atrofia dentatorubro-palidoluisiana (Koide et al. 1994; Nagafuchi et al. 1994). A interação de *ATN1*



com o alelo expandido no *ATXN3* foi associado à diminuição na ii de MJD/SCA3. O terceiro gene identificado como um potencial neuroprotetor foi o gene *HTT*, o qual, quando expandido causa a doença de Huntington (MacDonald et al. 1993). Nesse estudo, a presença de alelos curtos no gene *HTT* foi associada a um aumento na ii de pacientes com MJD/SCA3. Não foi identificado nenhum efeito dos genes *ATXN1*, *ATXN7*, *TBP* e *CACNA1A* sob a ii de pacientes com MJD/SCA3. Uma série de experimentos confirmatórios foram realizados posteriormente em quatro populações (caucasianas ou asiáticas) para reforçar a associação entre essas doenças. Entretanto, os resultados não foram replicados, provavelmente por diferenças nas frequências alélicas entre populações distintas (Tezenas du Montcel et al. 2014).

Essas interações gênicas foram estudadas em outras populações e os resultados obtidos evidenciam novamente diferenças populacionais. A análise de uma coorte de 102 pacientes brasileiros com MJD/SCA3 não encontrou associação dos genes estudados pelo EUROSCA e nem influência dos mesmos no quadro clínico dos pacientes (de Castilhos et al. 2014). Entretanto, um estudo anterior do mesmo grupo, identificou o gene *ATXN2* como modificador de uma manifestação clínica, fasciculações, em pacientes com MJD/SCA3 (Jardim et al. 2003). Importante ressaltar que, apesar da população ter a mesma origem, os pacientes incluídos foram diferentes.

Além dessas relações dentro da família de doenças neurodegenerativas, existem evidências possivelmente relacionando outros genes além dos acima descritos como modificadores da ii. O mais polêmico é a associação do alelo  $\epsilon 2$  do gene *APOE*. Dois grupos independentes, um deles incluindo 192 indivíduos de populações de ascendência portuguesa (brasileiros, açorianos e portugueses) (Bettencourt et al. 2011a) e outro com 155 chineses (Peng et al. 2014), identificaram o alelo  $\epsilon 2$  como fator de risco para ii precoce. No primeiro, a inclusão de alelo  $\epsilon 2$  nas análises de modelos lineares aumentou significativamente o efeito da variabilidade de 59,8% até 60,9%, e se traduziu clinicamente pela diminuição da ii em até 5 anos. No estudo de Peng e colaboradores, o efeito também foi drástico, com o alelo  $\epsilon 2$  diminuindo a ii em até 4 anos, e aumentando a explicação da variabilidade de 44,6% até 45,9% (Peng et al. 2014). Além da influência do *APOE*, foi descrito o efeito agravante do alelo *IL6\*C* em pacientes com MJD/SCA3, com uma diminuição na ii de até 4 anos (Raposo et al. 2017). O

mesmo estudo também começou a relacionar os mecanismos moleculares da influência do gene *APOE* com o sistema imune, especificamente compartilhando a via molecular do gene *IL6*. Por outro lado, um estudo mais antigo, de Zhou e colaboradores, analisou dados de 403 pacientes asiáticos e não encontrou efeito do alelo  $\epsilon 2$ , mesmo com as frequências alélicas sendo parecidas com as dos dois estudos anteriores (Zhou et al. 2014). Nesse mesmo estudo, que exclui as variantes do gene *APOE* como possíveis modificadores da doença, os autores de novo reforçaram o papel modificador do alelo normal na MJD/SCA3, elevando o valor preditivo de 65,9% até 66,6%.

Como citado anteriormente, o papel da ataxina-3, tanto na forma normal quanto a proteína mutante, ainda não foi elucidado de forma definitiva. Mas, existem evidências que, agregados dessa proteína e de outras doenças poliQs, podem ter a sua agregação alterada por proteínas de choque térmico, tais como hsp70 (Robertson et al. 2010). Na ação dessas proteínas é igualmente importante a função das chaperonas e de enzimas que guiam a sua atividade. E um desses componentes, a proteína CHIP, inspirou a investigação na sua possível papel na MJD/SCA3, sendo que ele é capaz de suprimir os agregados *in vivo* e *in vitro* (Miller et al. 2005). Um estudo realizado com pacientes brasileiros investigou o possível efeito neuroprotetor de polimorfismos no gene *CHIP* em 210 pacientes com MJD/SCA3 (França et al. 2012). Entretanto, apesar da fundamentação teórica, não foram encontrados polimorfismos com efeito modificador na coorte estudada.

Portanto, a quantidade de dados produzidos até o momento demonstra a dificuldade da identificação de qualquer modificador na MJD/SCA3, sendo que o efeito de variabilidade populacional e/ou do tamanho amostral do estudo. Desta forma, permanece em aberto a procura por modificadores que possam explicar a variabilidade de ii que não é explicada pela variação do tamanho dos alelos mutantes no gene *ATXN3* em pacientes com MJD/SCA3.

#### 4. C9orf72

C9orf72 (*chromosome 9 open reading frame 72*) é uma proteína em grande parte ignorada pelos estudos científicos até 2011, quando foi descrita como a etiologia molecular para um grande número de pacientes de esclerose lateral amiotrófica (ALS - *amyotrophic lateral*

*sclerosis*), demência frontotemporal (FTD - *frontotemporal dementia*) e comorbidade ALS-FTD por dois grupos independentes (DeJesus-Hernandez et al. 2011; Renton et al. 2011). A partir daí, ocorreu um aumento exponencial de número de estudos conduzidos associados a essa proteína e seu gene.

#### a. Estrutura e Função do gene *C9orf72*

O gene *C9orf72* está localizado no braço curto do cromossomo 9, região 21.2. Atualmente, duas mutações patogênicas foram descritas nesse gene: um região de expansão hexanucleotídica (GGGGCC)<sub>n</sub>, também conhecida como G<sub>4</sub>C<sub>2</sub> (DeJesus-Hernandez et al. 2011; Renton et al. 2011) e uma mutação de mudança de quadro de leitura em um sítio de *splicing*, com essa última tendo sido descrita em um único paciente (Liu et al. 2016). A expansão na região não-codificante desse gene é o principal alvo de estudo do *C9orf72*. O transcrito está sujeito a um mecanismo de *splicing* alternativo complexo, gerando três variantes de pré-RNA mensageiros (mRNA), codificando um total de duas isoformas da proteína C9orf72 (DeJesus-Hernandez et al. 2011).

Estruturalmente falando, a proteína C9orf72 é homóloga a proteínas da classe DENN (Zhang et al. 2012; Levine et al. 2013), as quais têm função de regulamento das Rab GTPases. A função exata da proteína codificada pelo gene ainda não é conhecida, mas há implicações da proteína na regulação do tráfego endossomal devido a sua interação com proteínas da classe Rab GTPases (Farg et al. 2014). Além disso, interações com proteínas Rab implicam o papel da C9orf72 no processo de autofagia (Webster et al. 2016). Como a proteína só recentemente entrou na visão popular da pesquisa, mais estudos são necessários para determinação de todas as interações de C9orf72 nos processos celulares.

#### b. Expansão GGGGCC

Em 2011, DeJesus-Hernandez e colaboradores e Renton e colaboradores descreveram simultaneamente uma repetição hexanucleotídica G<sub>4</sub>C<sub>2</sub>, localizada na região não-codificante no

íntron 1 do gene *C9orf72* (DeJesus-Hernandez et al. 2011; Renton et al. 2011). O aumento do número dessas repetições foi associado com a etiologia de até 20% dos casos de ALS-FTD. A partir daí o papel do *C9orf72* vem sendo investigado por vários grupos em doenças neurodegenerativas. Devido ao fato da região ser rica em GC, a repetição pode formar estruturas do tipo G-quadruplex, uma interação de ligação de hidrogênio de Hoogsteen, que impede a ação da DNA polimerase (Sen and Gilbert 1988).

O mecanismo molecular pelo qual essa mutação se torna patogênica ainda não está definido, mas existem três hipóteses parcialmente descritas e demonstradas experimentalmente. A primeira postula que a mutação produz haploinsuficiência da proteína *C9orf72* no organismo, com estudos mostrando que os níveis de mRNA caem drasticamente em pacientes portadores da expansão hexanucleotídica (DeJesus-Hernandez et al. 2011; Gijssels et al. 2012), provavelmente pelo impedimento da transcrição do gene (Haeusler et al. 2014). Mesmo com essas evidências moleculares, experimentos *in vivo* apresentam uma outra perspectiva: a depleção de mRNA do *C9orf72* nos neurônios dos murinos não resultou em desenvolvimento de uma série de sintomas neuromotores (Lagier-Tourenne et al. 2013; Koppers et al. 2015). Além disso, a herança dominante da doença em humanos é contrária a teoria de perda de função, sendo que doenças autossômicas dominantes são causadas, principalmente, por mutações de ganho de função. Com a expansão no gene fazendo parte de um íntron, é pouco provável que a estrutura da proteína seria modificada pela presença dessa mutação específica.

A segunda hipótese é o papel da toxicidade de *foci* de RNA formados por transcritos do alelo mutante (Gendron et al. 2014). Uma evidência que apoia essa hipótese é proveniente de uma outra doença neurodegenerativa: a distrofia miotônica do tipo 1 (DM1). A mutação responsável por DM1 se localiza numa região não codificante do gene-causador da doença, assim como a expansão G<sub>4</sub>C<sub>2</sub>. Especificamente, uma expansão CTG na região 3'UTR do gene *DMPK* (*myotonic dystrophy protein kinase*) resulta no desenvolvimento da DM1 (Brook et al. 1992). O mecanismo molecular pelo qual essa mutação interfere no metabolismo normal da célula foi caracterizado como interferência nos mecanismos de processamento do RNA, incluindo *splicing* alternativo (Lin et al. 2006), e sequestro de proteína MBNL1 (*muscle blind-like protein 1*) (Miller et al. 2000). Baseando nessa prerrogativa, foram investigados os efeitos dos *foci* de RNA no caso de expansão no gene *C9orf72* e foram encontrados resultados

semelhantes ao que ocorre no produto gênico de *DMPK*. Uma série de proteínas que se ligam ao RNA foram colocalizadas com os *foci* de RNA. Uma delas, por exemplo, é a Pur- $\alpha$ , uma proteína ubíqua e multifuncional, que reconhece a região GGGGCC como ligante de forma específica (Xu et al. 2013). Além disso, outra via molecular foi proposta por Burguete e colaboradores em 2015: os grânulos de RNA tendem a se deslocar, se incorporam com as partículas ribonucleoproteicas e, com isso, desencadeiam a disfunção de ramificação dendrítica (Burguete et al. 2015). Esse novo mecanismo mostrou que a ação neurodegenerativa da expansão hexanucleotídica não se limita ao núcleo e à maquinaria transcriptional, mas provavelmente está também envolvida nos passos posteriores do tráfego de RNA dentro da célula. A validade dessa hipótese está sendo contestada por sua ligação íntima com a terceira hipótese: a toxicidade dos dipeptídeos.

Mesmo sendo localizada fora da região codificante, é possível ocorrer a tradução da região da repetição a partir de um mecanismo denominado tradução associada a repetição iniciada por não-ATG (tradução RAN ou *RNA-translation – repeat-associated non-ATG translation*). Esse mecanismo permite que regiões altamente repetitivas, tais como a expansão G<sub>4</sub>C<sub>2</sub>, sirvam como sequência iniciadora de tradução, levando à produção de proteínas não-canônicas (Koppers et al. 2015). A natureza da mutação permite seis diferentes quadros de leitura, que são capazes de resultar em cinco tipos de proteínas dipeptídicas, as quais são capazes de se aglomerar em inclusões proteicas parecidas com as das doenças poliQ (Ash et al. 2013; Mori et al. 2013). Os mecanismos exatos sobre como essas proteínas levam à degeneração estão sendo amplamente discutidos, com teorias que incluem a falha nas vias de degeneração proteica, a formação de folhas amiloides ou a sobrecarga de estresse celular sobre os sistemas de transporte intracelulares. Proponentes dessa hipótese frequentemente entram em conflito com os da teoria de toxicidade de *foci* de RNA, afirmando que a toxicidade vem das proteínas e não dos transcritos, trazendo como comprovação experimentos com utilização de tratos de expansões interrompidos, que não são capazes de desencadear a tradução RAN (Mizielinska et al. 2014). A dificuldade em apoiar essa hipótese vem da ausência de aglomerados de proteínas dipeptídicas nas amostras pós-morte dos pacientes, confundindo ainda mais o mistério da neurodegeneração mediada por *C9orf72*.

### c. Esclerose Lateral Amiotrófica e Demência Frontotemporal

A esclerose lateral amiotrófica (ALS – *amyotrophic lateral sclerosis*), uma doença neurodegenerativa clássica é caracterizada por ii tardio e uma grande variedade de anomalias relacionadas com o movimento ou a atividade muscular e, na sua grande maioria, de etiologia desconhecida (Zarei et al. 2015). A demência frontotemporal (FTD – *Frontotemporal dementia*) é um termo amplamente utilizado para descrever uma série de síndromes de demência patologicamente heterogêneas e distintas do mal do Alzheimer (Warren et al. 2013). Apesar de ter um quadro clínico altamente variável, uma característica comum da grande maioria dos pacientes com FTD é a presença de inclusões proteicas semelhantes as das doenças neurodegenerativas como MJD/SCA3 (Rabinovici and Miller 2010). A história de comorbidade ALS-FTD é longa, mas só recentemente, com avanços nas tecnologias moleculares, foi proposta a teoria que essas duas doenças possam representar dois extremos de um espectro único de vias moleculares neurodegenerativas comuns (Ferrari et al. 2011). Até 2011, a grande maioria dos casos não tinha nenhuma explicação molecular descrita, com poucos casos de ALS sendo causados por mutações no gene *SOD1*, resultando em inclusões da proteína mutante nos neurônios (Rosen et al. 1993). O grande avanço na área foi a descoberta de expansão G<sub>4</sub>C<sub>2</sub> no gene *C9orf72* como principal agente etiológico no espectro ALS-FTD, podendo explicar até 34,2% dos casos de ALS e 25,9% dos casos de FTD puro (van Blitterswijk et al. 2012).

Em termos de mecanismos específicos às doenças ALS-FTD relacionadas à expansão, estudos mostram resultados contraditórios. Análises *in vitro* dos transcritos de *C9orf72* humano demonstraram a produção de dipeptídeos tóxicos a partir de tradução RAN (Tabet et al. 2018). Por outro lado, modelos funcionais em *zebrafish* elimina esses tratos tóxicos como a etiologia da doença e apontam para a toxicidade de RNA como culpado (Swinnen et al. 2018), sendo que os transcritos de *C9orf72* demonstravam afinidade para subunidades ribossomais. Uma característica importante é que a presença da expansão no *C9orf72* não indica ausência de outros sinais neurodegenerativos nos pacientes. Um exemplo é a presença das inclusões da proteína TDP-43 em grande número de casos com a expansão de G<sub>4</sub>C<sub>2</sub> (Boeve et al. 2012; Mahoney et al. 2012).

Portanto, a ocorrência de expansão do hexanucleotídeo G<sub>4</sub>C<sub>2</sub> no gene *C9orf72* está amplamente descrita, mas o efeito dessa alteração no processo de neurodegeneração ainda precisa ser elucidado.

## 5. Justificativa

O estudo de genes modificadores de doenças neurodegenerativas apresenta grande interesse para maior conhecimento tanto da fisiopatologia de MJD/SCA3 quanto para o desenvolvimento de terapias para essa doença. Conforme dito anteriormente, genes relacionados com outras doenças neurodegenerativas podem interagir e podem atuar como modificadores de outras patologias com mecanismos semelhantes. Além disso, existem uma série de evidências para a possível relação entre MJD/SCA3 e *C9orf72*, que serão descritos abaixo.

A MJD/SCA3 e a ALS compartilham vias moleculares de degradação proteica. Tanto a ataxina-3 quanto a superóxido-dismutase 1 (SOD1) foram associadas com vias proteicas que passam pelo retículo endoplasmático (ER, *endoplasmatic reticulum*). A interação entre essas duas proteínas foi demonstrada e essa interação é regulada por um representante de enzimas do ER: a gp78 (Ying et al. 2009). Esse estudo demonstrou a estreita relação da gp78 com as proteínas ataxina-3 e SOD1, sendo que, quando superexpressa, ela é capaz de regular os aglomerados de proteínas mutantes através de ubiquitinação e, conseqüentemente, envio para vias de degradação.

Além da via de degradação compartilhada, foram documentados casos de indivíduos com expansão hexanucleotídica que compartilhavam quadros clínicos e patológicos com pacientes de MJD/SCA3. Um paciente dinamarquês foi identificado como carregador da expansão, com histórico familiar de ataxia, que não entra no quadro clássico de ALS (Lindquist et al. 2013). Na França, um grupo identificou caso de pura ataxia cerebelar ligada com a expansão G<sub>4</sub>C<sub>2</sub>, indicando que existe um mecanismo molecular do *C9orf72* que atua na degradação de cerebelo (Corcia et al. 2016). O caso de paciente italiano carregador da mutação e demonstrando as sintomas de desordens psiquiátricas e ataxia cerebelar progressiva (Meloni et al. 2017). Esses casos clínicos mostram claramente que o papel de *C9orf72* não está confinado

para as vias de doenças de ALS-FTD, e demonstram a necessidade em investigar a influência desse gene em outras doenças neurodegenerativas.

Levando em consideração o papel do gp78 e esses casos clínicos, a questão da relação de *C9orf72* como potencial modificador da MJD/SCA3 permanecia aberto. Até o momento, um único estudo foi realizado na investigação dessa interação (Wang et al. 2015). Nesse estudo, foi descrito o possível papel dos alelos normais “intermediários” (alelos com 7 a 30 repetições). Foi observado que pacientes com pelo menos um alelo intermediário apresentara ii aproximadamente 3 anos mais cedo do que os pacientes com dois alelos de até 6 repetições, inclusive. Portanto, as evidências apresentadas indicam um possível papel das repetições nucleotídicas no *C9orf72* como modificador na MJD/SCA3.



## II. Objetivos

### 1. Objetivo Geral

Avaliar a repetição hexanucleotídica GGGGCC no gene *C9orf72* como possível modificador da idade de início em pacientes brasileiros com doença do Machado-Joseph/ataxia espinocerebelar tipo 3 (MJD/SCA3).

### 2. Objetivos Específicos

- Determinar a distribuição alélica da repetição GGGGCC no gene *C9orf72* em amostras controles;
- Avaliar a distribuição alélica GGGGCC no gene *C9orf72* em pacientes com MJD/SCA3;
- Comparar a distribuição alélica nos dois grupos (controles e pacientes com MJD/SCA3);
- Dividir os pacientes com MJD/SCA3 em dois grupos de acordo com a presença ou não de alelos “intermediários” no gene *C9orf72* e comparar com a ii da doença de uma forma geral;
- Comparar apenas os pacientes com MJD/SCA3 do grupo de ii precoce e ii tardia de acordo com a expansão no gene *ATXN3* com a distribuição dos alelos normais no gene *C9orf72*.

### **III. Resultados**

Os resultados desse trabalho serão apresentados na forma de dois manuscritos conforme listados abaixo:

- *C9orf72* gene and Cerebellar Ataxia: A Tenuous Relationship – Review
- Analysis of *C9orf72* hexanucleotide repeats length as a potential modifying factor of age of onset in Machado-Joseph Disease/Spinocerebellar Ataxia Type 3 patients – Original Article

Os dois manuscritos serão submetidos para publicação no periódico *Clinical Genetics*.

## ***C9orf72* gene and Cerebellar Ataxia: A Tenuous Relationship**

Yelena Perevalova<sup>1,2</sup>, Maria-Luiza Saraiva-Pereira<sup>1,2,3,4</sup>

1. Laboratório de Identificação Genética – Hospital de Clínicas de Porto Alegre (HCPA);
2. Programa de Pós-Graduação em Genética e Biologia Molecular – Universidade Federal do Rio Grande do Sul (UFRGS);
3. Serviço de Genética Médica – HCPA;
4. Depto. de Bioquímica – UFRGS, Porto Alegre, RS, Brazil.

Running Title: *C9orf72* Gene and Ataxias

Corresponding author:

Maria Luiza Saraiva-Pereira, PhD

Laboratório de Identificação Genética - Centro de Pesquisa Experimental

Hospital de Clínicas de Porto Alegre

Ramiro Barcelos 2350

CEP 90035-903, Porto Alegre, RS, Brazil

e-mail: [mlpereira@hcpa.edu.br](mailto:mlpereira@hcpa.edu.br)

## **Abstract**

*C9orf72* took neurogenetics by storm in 2011 with the discovery of a close relationship of this gene with amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) as well as the whole spectrum in between. The potential role of *C9orf72* in neurodegenerative disorders outside ALS-FTD syndromes remains elusive and controversial. Even considering molecular evidences point to a pathological mechanism similar to that of other diseases, the clinical proof of a role of *C9orf72* gene in neurological diseases proves controversial. The evidence for *C9orf72*'s role in psychiatric disturbances has reached an almost unequivocal acceptance. At the same time, how the expansion could affect other areas governed by the nervous system remains a fairly open question. In this review, we aim to outline the current state of molecular and clinical evidence surrounding *C9orf72* and its role in ataxic movement disorders. We attempt to ascertain if *C9orf72* interferes with neurological function beyond that of psychological symptoms or if the motor functions of the nervous system remain unaffected by this novel mutation.

## **Keywords**

*C9orf72*, ataxia, SCA, neurodegeneration, modifiers, proteinopathy

## Introduction

The discovery of an intronic hexanucleotide GGGGCC (or G<sub>4</sub>C<sub>2</sub>) repeat in *C9orf72* as the causal mutation of the co-morbid amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) took the world of neurogenetics by surprise. The meteoric rise of this discovery into the spotlight led to some researchers dubbing it the “new gene on the block”<sup>1</sup>. After large scale identification of the hexanucleotide mutation as a frequent cause of previously idiopathic ALS, FTD and ALS-FTD<sup>2-6</sup> led to a flurry of investigation into the role of this gene in other neurodegenerative diseases. Rare syndromes of dubious genetic origin, such as Huntington’s disease phenocopies<sup>7</sup>, progressive muscular atrophy, and primary lateral sclerosis<sup>8</sup> were among the first to attempt to correlate *C9orf72* with the development of these diseases. But others, such as parkinsonism<sup>9</sup>, ruled out the hexanucleotide mutation immediately as a potential contributing factor.

Cerebellar ataxias comprise a family of highly clinically and molecularly heterogeneous neurological disorders, mostly progressive in nature, predominantly affecting gait and limb movement. Often these symptoms are accompanied by a progressive cerebellar deterioration. The largest group of hereditary disorders defined as cerebellar ataxias are, with no doubt, the spinocerebellar ataxias (SCAs). These diseases that are predominantly inherited as an autosomal dominant trait affect up to 5 in 100,000 individuals worldwide<sup>10</sup>. Due to the highly diverse clinical nature of these diseases, they are difficult to effectively diagnose without a complete family history or molecular diagnosis. Even within populations of patients affected by the same disease, such as spinocerebellar ataxia type 3 (SCA3) or also known as Machado-Joseph disease (MJD), the symptomology of patients and severity of the disease can vary greatly, suggesting that factors other than the molecular etiology (in the case of SCA3 – a CAG expansion in the *ATXN3* gene) could be at work attenuating or intensifying the progression and development of the disease<sup>11</sup>. In this review, we look at the molecular and clinical case relating the G<sub>4</sub>C<sub>2</sub> expansion region to various ataxic disorders and discuss the future of *C9orf72* in ataxia research.

## Molecular Hypotheses for a Connection

The mechanism of general neurodegeneration remains an open mystery in the field of molecular biology. Elucidating the molecular origins of the *C9orf72* hexanucleotide expansion induced neurological dysfunction has been a challenge by itself, with three proposed molecular pathways, potentially acting alone or in tandem: 1) mutant *C9orf72* RNA transcripts form *foci* that are capable of sequestering proteins vital to cell function and/or survival<sup>12</sup>; 2) the hexanucleotide expansion induces repeat-associated non-ATG translation, leading to expression of toxic dipeptide chains<sup>13</sup>; 3) a haploinsufficiency of normal *C9orf72* transcripts<sup>14</sup>. These particular molecular mechanisms echo within the hereditary ataxias: the predominant proposed theory for the cause of SCA3 being the aggregation of toxic mutated ATXN3 protein within the cell; Friedreich's ataxia is supposedly a consequence of haploinsufficiency of frataxin; one of the main pathophysiological manifestations of SCA8 are RNA *foci* suspected of causing the disease. While simple mechanistic similarities are not evident enough of a concrete connection, it does create a foundation for further investigation.

Determining potential sequestering role of RNA *foci* of *C9orf72* transcripts is made difficult by the large number of interactants of the hexanucleotide repeat sequence. Two promising candidates are pur- $\alpha$  and hnRNP K. Pur- $\alpha$ , a protein vital to the regulation of gene expression in neurons<sup>15</sup>, binds to GC-rich regions, specifically CGG repeats, and this gradual insufficiency has been shown to lead to neurodegeneration in models of Fragile X-associated tremor/ataxia syndrome<sup>16</sup>, with heterozygous mouse models developing ataxia and suffering neuronal loss<sup>17</sup>. Heterogeneous nuclear ribonucleoproteins (hnRNPs) have been shown to co-localize with *C9orf72 foci*<sup>13,18,19</sup>. Of note is hnRNP K, which preferentially binds to antisense hexanucleotide transcripts<sup>19</sup>, and the loss of this particular hnRNP has been associated with SCA10<sup>20</sup>.

The expression of dipeptide chain aggregates may possibly serve as a function beyond simple toxicity. A common characteristic of C9 ALS-FTD patients is the formation of p62 intranuclear inclusions<sup>21</sup>, which possibly form as a component of dipeptide repeat protein inclusions<sup>22</sup>. This form of sequestration could lead to an insufficiency of p62 within the cell, which has been shown to cause a severe early onset form of ataxia<sup>23</sup>. Along with p62, UBQLN2

mutations and proteopathies have been shown to cause fairly early-onset ALS-FTD diseases<sup>24</sup> as well as being associated with the hexanucleotide expansion and co-localizing with p62<sup>25</sup>. Together, these two proteins form part of a proposed aggregate pathway<sup>26,27</sup>, whose inactivation would hinder the clearance of toxic polyglutamines implicated in hereditary ataxia.

Elucidating the role of *C9orf72* within the cell itself is essential to determining the potential role of this protein in the course of the development of hereditary ataxias. The prevailing theory of *C9orf72*'s role in regulating molecular trafficking and autophagy<sup>28,29</sup> offers a promising molecular connection between this new protein on the block and the polyglutamine ataxias. For example, in the case of SCA3, autophagy dysregulation is often observed in the tissues of patients<sup>30,31</sup> and animal models suggest that the activation of autophagic machinery as a potential treatment path, mitigating the neurological symptoms of the disease<sup>32-35</sup>, possibly due to its degradation of mutant ataxin-3<sup>36</sup>. Homozygous of rare missense mutations in the autophagy-related 5 gene display ataxia, among other neurological symptoms, most likely due to a severe loss of autophagy function<sup>37</sup>. Knowing that the expression of the *C9orf72* protein is shown to be inversely correlated with the number of G<sub>4</sub>C<sub>2</sub> repeats present in the carried alleles<sup>38</sup>, even hypothetically “normal” genotypes (alleles with less than 30 to 45 repeats) can have a modifying effect on the development of an ataxic disease. At least one ataxia-causing gene, *ATXN2* has been demonstrated acting in concert with *C9orf72*, with mutant alleles synergizing to induce neuronal death *in vitro* and in a zebrafish model, and display diseased symptomology in the latter<sup>34</sup>, which suggests that *C9orf72* could act as a fatal second blow to milder genotypes in other ataxic diseases.

## The Clinical Evidence of *C9orf72*-induced Ataxia

Despite the extreme focus on ALS-FTD, several cases of seemingly non-orthodox diseases were identified in *C9orf72* expansion carriers. The first case of possible “C9 ataxia” was documented in 2013, when a Danish patient of olivopontocerebellar degeneration, with a familial history of ataxia, was identified as a carrier of the hexanucleotide expansion<sup>39</sup>. A 2014 case report describes two American siblings with a familial history of ALS presenting ataxic symptoms within multiple system atrophy (MSA), while lacking cerebellar ataxia<sup>40</sup>, with both carrying the expansion. A French case in 2015 documented a patient of pure cerebellar ataxia seemingly caused by the *C9orf72* expansion<sup>41</sup>. A 36-year-old patient with a familial history of neurological disorders but *not* ataxia was admitted with symptoms associated classically with the hereditary ataxias, but not ALS-FTD: wide gait, incoordination of the limbs and ataxic dysarthria. Molecular tests for the most common types of ataxia (SCA1, SCA2, SCA3, SCA6, SCA7, SCA12, SCA17), Freidreich’s ataxia and dentatorubral-pallidoluysian atrophy (DRPLA) were all negative. Due to the patient’s mother suffering from FTD, *C9orf72* was tested and a pathogenic expansion was revealed. The 2017 case of a Sardinian patient with a complex mix of psychiatric disorders and progressive ataxia, negative for the common SCA mutations widened the possibilities of *C9orf72* pathogenesis: the patient was found to be a carrier of 21 hexanucleotide repeats, well below the previously described pathogenic limit of 30, which was attributed as the cause of her psychological and neurological degeneration<sup>42</sup>. *C9orf72* could serve as a modifying factor of molecularly classical ataxic disease, like other neurodegenerative disease-causing genes, despite not being the direct cause of ataxia. A 2015 study of Chinese SCA3 patients noted a possible neuroprotective factor of short-length (<7 repeats) *C9orf72* alleles, potentially decreasing age of onset by as much as 3 years<sup>43</sup>.

While these cases are preliminarily the only sample of patients with C9-attributed ataxia, it is important to note that *C9orf72* as a disease-causing agent has only been present on the scene since 2011, with the vast majority of studies aimed at demonstrating either the molecular mechanism of the disease in the ALS-FTD continuum, or statistical studies of patients of these diseases.



## The Clinical Case Against *C9orf72*

The first widescale analysis of patients outside the ALS-FTD spectrum was performed in 2012 with a survey of 209 North American patients diagnosed with ataxia of unknown etiology, negative for the five most common forms of ataxia to the population (SCA1, SCA2, SCA3, SCA6, SCA7)<sup>44</sup>. Only one Caucasian woman, representing 0.5% of the sample was shown to be positive for the expansion, with a familial history of ALS, but no previous record of ataxia. The cerebellar variant of multiple system atrophy (MSA) was next to be analyzed, with 331 patients and an addition of 98 unknown SCA patients of Taiwanese origin tested for the expansion, with no positive result<sup>45</sup>. A 2015 study of generalized MSA confirmed these results, with 100 Caucasian patients identified as non-carriers of the mutation<sup>46</sup>. Study of sporadic ataxia patients continued in 2016 within a Chinese population of patients, with no expansion carriers being detected<sup>47</sup>. Finally, a 2018 study of 440 formally undiagnosed German SCA patients detected only one patient, aggravated with signs of dementia, carrying the expansion (0.2%)<sup>48</sup>.

Further complicating the question of the role of hexanucleotide expansion in hereditary ataxias, a 2017 survey of the Clinical Research Consortium for Spinocerebellar Ataxias (CRC-SCA) brought about new data concerning the frequency of the expansion within SCA patients<sup>49</sup>. While having no effect on ataxic progression in any SCA patients, the expansion was present in 7 patients (3 of SCA1, 2 of SCA2, 1 of SCA3 and 1 of SCA6), with SCA1, SCA2 and SCA6 showing a higher than expected frequency of this mutation, and intermediary alleles served as a marker for the progression of non-motor symptoms, mainly depression. Bringing further support to the hypothesis that *C9orf72* acts as a secondary hit to the clinical phenotype of patients, the 2017 description of a family with co-morbid ataxia and dementia carrying two distinct expansion mutations: the *C9orf72* hexanucleotide expansion and the trinucleotide expansion of *ATXN2*<sup>50</sup>.

## **Final Considerations**

Despite the blossoming role of *C9orf72* as the new golden child of neurogenetics, its role outside of the ALS-FTD continuum remains an elusive one. While some researchers eagerly assign the role of causative agent to the hexanucleotide expansion in neurodegenerative disorders with difficult to ascertain etiologies, others maintain an arm's length distance from the "new gene on the block". Despite several fairly promising molecular theories on a regulating role of *C9orf72* within ataxia, the clinical data, by and large, seems to deny a solid connection. Indeed, with recent studies questioning the effect of *C9orf72* on ataxic progression, while confirming its influence in non-locomotive symptoms, the relationship between the gene and ataxia can best be described as tenuous. While there is no denying the as-of-yet unexplained mechanisms behind the development of ataxia in otherwise genotypically normal patients, it appears that the key to understanding ataxic disorders will not lie within the mechanisms of the hexanucleotide repeat.

## References

1. Bigio EH. C9ORF72, the new gene on the block, causes C9FTD/ALS: new insights provided by neuropathology. *Acta Neuropathol (Berl)*. 2011;122(6):653-655. doi:10.1007/s00401-011-0919-7
2. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, et al. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron*. 2011;72(2):245-256. doi:10.1016/j.neuron.2011.09.011
3. Renton AE, Majounie E, Waite A, et al. A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron*. 2011;72(2):257-268. doi:10.1016/j.neuron.2011.09.010
4. Abramychева NY, Lysogorskaia EV, Stepanova MS, et al. C9ORF72 hexanucleotide repeat expansion in ALS patients from the Central European Russia population. *Neurobiol Aging*. 2015;36(10):2908.e5-9. doi:10.1016/j.neurobiolaging.2015.07.004
5. García-Redondo A, Dols-Icardo O, Rojas-García R, et al. Analysis of the C9orf72 gene in patients with amyotrophic lateral sclerosis in Spain and different populations worldwide. *Hum Mutat*. 2013;34(1):79-82. doi:10.1002/humu.22211
6. Smith BN, Newhouse S, Shatunov A, et al. The C9ORF72 expansion mutation is a common cause of ALS+/-FTD in Europe and has a single founder. *Eur J Hum Genet EJHG*. 2013;21(1):102-108. doi:10.1038/ejhg.2012.98
7. Koutsis G, Karadima G, Kartanou C, Kladi A, Panas M. C9ORF72 hexanucleotide repeat expansions are a frequent cause of Huntington disease phenocopies in the Greek population. *Neurobiol Aging*. 2015;36(1):547.e13-16. doi:10.1016/j.neurobiolaging.2014.08.020
8. van Rheen W, van Blitterswijk M, Huisman MHB, et al. Hexanucleotide repeat expansions in C9ORF72 in the spectrum of motor neuron diseases. *Neurology*. 2012;79(9):878-882. doi:10.1212/WNL.0b013e3182661d14
9. Lesage S, Le Ber I, Condroyer C, et al. C9orf72 repeat expansions are a rare genetic cause of parkinsonism. *Brain*. 2013;136(Pt 2):385-391. doi:10.1093/brain/aws357

10. van de Warrenburg BPC, Sinke RJ, Verschuuren-Bemelmans CC, et al. Spinocerebellar ataxias in the Netherlands: prevalence and age at onset variance analysis. *Neurology*. 2002;58(5):702-708.
11. Tezenas du Montcel S, Durr A, Bauer P, et al. Modulation of the age at onset in spinocerebellar ataxia by CAG tracts in various genes. *Brain J Neurol*. 2014;137(Pt 9):2444-2455. doi:10.1093/brain/awu174
12. Xu Z, Poidevin M, Li X, et al. Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. *Proc Natl Acad Sci*. 2013;110(19):7778–7783.
13. Mori K, Arzberger T, Grässer FA, et al. Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol (Berl)*. 2013;126(6):881-893. doi:10.1007/s00401-013-1189-3
14. Gijselinck I, Van Langenhove T, van der Zee J, et al. A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: a gene identification study. *Lancet Neurol*. 2012;11(1):54-65. doi:10.1016/S1474-4422(11)70261-7
15. Li Y, Koike K, Ohashi S, et al. Pur alpha protein implicated in dendritic RNA transport interacts with ribosomes in neuronal cytoplasm. *Biol Pharm Bull*. 2001;24(3):231-235.
16. Jin P, Duan R, Qurashi A, et al. Pur  $\alpha$  binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a Drosophila model of Fragile X Tremor/Ataxia Syndrome. *Neuron*. 2007;55(4):556-564. doi:10.1016/j.neuron.2007.07.020
17. Barbe MF, Krueger JJ, Loomis R, Otte J, Gordon J. Memory deficits, gait ataxia and neuronal loss in the hippocampus and cerebellum in mice that are heterozygous for Pur-alpha. *Neuroscience*. 2016;337:177-190. doi:10.1016/j.neuroscience.2016.09.018
18. Lee Y-B, Chen H-J, Peres JN, et al. Hexanucleotide Repeats in ALS/FTD Form Length-Dependent RNA Foci, Sequester RNA Binding Proteins, and Are Neurotoxic. *Cell Rep*. 2013;5(5):1178-1186. doi:10.1016/j.celrep.2013.10.049
19. Haeusler AR, Donnelly CJ, Periz G, et al. C9orf72 Nucleotide Repeat Structures Initiate Molecular Cascades of Disease. *Nature*. 2014;507(7491):195-200. doi:10.1038/nature13124

20. White MC, Gao R, Xu W, et al. Inactivation of hnRNP K by Expanded Intronic AUUCU Repeat Induces Apoptosis Via Translocation of PKC $\delta$  to Mitochondria in Spinocerebellar Ataxia 10. *PLoS Genet.* 2010;6(6). doi:10.1371/journal.pgen.1000984
21. Al-Sarraj S, King A, Troakes C, et al. p62 positive, TDP-43 negative, neuronal cytoplasmic and intranuclear inclusions in the cerebellum and hippocampus define the pathology of *C9orf72*-linked FTLN and MND/ALS. *Acta Neuropathol (Berl)*. 2011;122(6):691-702. doi:10.1007/s00401-011-0911-2
22. Mann DM, Rollinson S, Robinson A, et al. Dipeptide repeat proteins are present in the p62 positive inclusions in patients with frontotemporal lobar degeneration and motor neurone disease associated with expansions in *C9ORF72*. *Acta Neuropathol Commun.* 2013;1:68. doi:10.1186/2051-5960-1-68
23. Haack TB, Ignatius E, Calvo-Garrido J, et al. Absence of the Autophagy Adaptor SQSTM1/p62 Causes Childhood-Onset Neurodegeneration with Ataxia, Dystonia, and Gaze Palsy. *Am J Hum Genet.* 2016;99(3):735-743. doi:10.1016/j.ajhg.2016.06.026
24. Deng H-X, Chen W, Hong S-T, et al. Mutations in UBQLN2 cause dominant X-linked juvenile and adult onset ALS and ALS/dementia. *Nature.* 2011;477(7363):211-215. doi:10.1038/nature10353
25. Brettschneider J, Van Deerlin VM, Robinson JL, et al. Pattern of ubiquilin pathology in ALS and FTLN indicates presence of *C9ORF72* hexanucleotide expansion. *Acta Neuropathol (Berl)*. 2012;123(6):825-839. doi:10.1007/s00401-012-0970-z
26. Hjerpe R, Bett JS, Keuss MJ, et al. UBQLN2 Mediates Autophagy-Independent Protein Aggregate Clearance by the Proteasome. *Cell.* 2016;166(4):935-949. doi:10.1016/j.cell.2016.07.001
27. Katsuragi Yoshinori, Ichimura Yoshinobu, Komatsu Masaaki. p62/SQSTM1 functions as a signaling hub and an autophagy adaptor. *FEBS J.* 2015;282(24):4672-4678. doi:10.1111/febs.13540
28. Ugolino J, Ji YJ, Conchina K, et al. Loss of *C9orf72* Enhances Autophagic Activity via Deregulated mTOR and TFEB Signaling. *PLOS Genet.* 2016;12(11):e1006443. doi:10.1371/journal.pgen.1006443

29. Ji YJ, Ugolino J, Brady NR, Hamacher-Brady A, Wang J. Systemic deregulation of autophagy upon loss of ALS- and FTD-linked C9orf72. *Autophagy*. 2017;13(7):1254-1255. doi:10.1080/15548627.2017.1299312
30. Onofre I, Mendonça N, Lopes S, et al. Fibroblasts of Machado Joseph Disease patients reveal autophagy impairment. *Sci Rep*. 2016;6:28220. doi:10.1038/srep28220
31. Sittler Annie, Muriel Marie - Paule, Marinello Martina, Brice Alexis, den Dunnen Wilfred, Alves Sandro. Deregulation of autophagy in postmortem brains of Machado - Joseph disease patients. *Neuropathology*. 2017;0(0). doi:10.1111/neup.12433
32. Watchon M, Yuan KC, Mackovski N, et al. Calpain Inhibition Is Protective in Machado–Joseph Disease Zebrafish Due to Induction of Autophagy. *J Neurosci*. 2017;37(32):7782-7794. doi:10.1523/JNEUROSCI.1142-17.2017
33. Nascimento-Ferreira I, Nóbrega C, Vasconcelos-Ferreira A, et al. Beclin 1 mitigates motor and neuropathological deficits in genetic mouse models of Machado-Joseph disease. *Brain J Neurol*. 2013;136(Pt 7):2173-2188. doi:10.1093/brain/awt144
34. Sellier C, Campanari M, Julie Corbier C, et al. Loss of C9ORF72 impairs autophagy and synergizes with polyQ Ataxin - 2 to induce motor neuron dysfunction and cell death. *EMBO J*. 2016;35(12):1276-1297. doi:10.15252/embj.201593350
35. Ferreira ACLV. Pharmacological activation of autophagy in Machado-Joseph disease. 2012. <https://estudogeral.sib.uc.pt/jspui/handle/10316/22070>. Accessed March 20, 2018.
36. Ou Z, Luo M, Niu X, et al. Autophagy Promoted the Degradation of Mutant ATXN3 in Neurally Differentiated Spinocerebellar Ataxia-3 Human Induced Pluripotent Stem Cells. *BioMed Research International*. doi:10.1155/2016/6701793
37. Kim M, Sandford E, Gatica D, et al. Mutation in ATG5 reduces autophagy and leads to ataxia with developmental delay. *eLife*. 5. doi:10.7554/eLife.12245
38. van der Zee J, Gijssels I, Dillen L, et al. A pan-European study of the C9orf72 repeat associated with FTLN: geographic prevalence, genomic instability, and intermediate repeats. *Hum Mutat*. 2013;34(2):363-373. doi:10.1002/humu.22244

39. Lindquist S, Duno M, Batbayli M, et al. Corticobasal and ataxia syndromes widen the spectrum of C9ORF72 hexanucleotide expansion disease. *Clin Genet*. 2013;83(3):279-283. doi:10.1111/j.1399-0004.2012.01903.x
40. Goldman J, Quinzii CM, Dunning-Broadbent J, et al. Multiple System Atrophy and Amyotrophic Lateral Sclerosis in a Family with Hexanucleotide Repeat Expansions in C9orf72. *JAMA Neurol*. 2014;71(6):771-774. doi:10.1001/jamaneurol.2013.5762
41. Corcia P, Vourc'h P, Guennoc A-M, et al. Pure cerebellar ataxia linked to large C9orf72 repeat expansion. *Amyotroph Lateral Scler Front Degener*. 2016;17(3-4):301-303. doi:10.3109/21678421.2015.1113298
42. Meloni M, Farris R, Solla P, Mascia MM, Marrosu F, Cannas A. C9ORF72 Intermediate Repeat Expansion in a Patient With Psychiatric Disorders and Progressive Cerebellar Ataxia. *The Neurologist*. 2017;22(6):245. doi:10.1097/NRL.0000000000000147
43. Wang C, Chen Z, Yang F, et al. Analysis of the GGGGCC Repeat Expansions of the C9orf72 Gene in SCA3/MJD Patients from China. Gonzalez-Alegre P, ed. *PLOS ONE*. 2015;10(6):e0130336. doi:10.1371/journal.pone.0130336
44. Fogel BL, Pribadi M, Pi S, Perlman SL, Geschwind DH, Coppola G. C9ORF72 Expansion is Not a Significant Cause of Sporadic Spinocerebellar Ataxia. *Mov Disord Off J Mov Disord Soc*. 2012;27(14):1835-1836. doi:10.1002/mds.25245
45. Hsiao C-T, Tsai P-C, Liao Y-C, Lee Y-C, Soong B-W. C9ORF72 repeat expansion is not a significant cause of late onset cerebellar ataxia syndrome. *J Neurol Sci*. 2014;347(1-2):322-324. doi:10.1016/j.jns.2014.10.042
46. Scholz SW, Majounie E, Revesz T, et al. Multiple system atrophy is not caused by C9orf72 hexanucleotide repeat expansions. *Neurobiol Aging*. 2015;36(2):1223.e1-1223.e2. doi:10.1016/j.neurobiolaging.2014.08.033
47. He M, Yan W-Q, Zeng S, et al. C9ORF72 repeat expansion is not detected in sporadic ataxia patients in mainland China. *J Neurol Sci*. 2016;361:181-183. doi:10.1016/j.jns.2015.12.034
48. Aydin G, Dekomien G, Hoffjan S, Gerding WM, Epplen JT, Arning L. Frequency of SCA8, SCA10, SCA12, SCA36, FXTAS and C9orf72 repeat expansions in SCA patients

negative for the most common SCA subtypes. *BMC Neurol.* 2018;18. doi:10.1186/s12883-017-1009-9

49. Figueroa KP, Gan Shi - Rui, Perlman Susan, et al. C9orf72 repeat expansions as genetic modifiers for depression in spinocerebellar ataxias. *Mov Disord.* 2017;33(3):497-498. doi:10.1002/mds.27258

50. Zhang M, Xi Z, Misquitta K, et al. C9orf72 and ATXN2 repeat expansions coexist in a family with ataxia, dementia, and parkinsonism. *Mov Disord Off J Mov Disord Soc.* 2017;32(1):158-162. doi:10.1002/mds.26841



**Analysis of *C9orf72* hexanucleotide repeats length as a potential modifying factor of age of onset in Machado-Joseph Disease/Spinocerebellar Ataxia Type 3 patients**

Yelena Perevalova<sup>1,2</sup>, Gabriel V. Furtado<sup>1,2</sup>, Eduardo P. Mattos<sup>1,2</sup>, Jonas Saute<sup>1,3</sup>, Laura Bannach Jardim<sup>1,2,3,4</sup>, Maria-Luiza Saraiva-Pereira<sup>1,2,3,5</sup>

1. Laboratório de Identificação Genética – Hospital de Clínicas de Porto Alegre (HCPA);
2. Programa de Pós-Graduação em Genética e Biologia Molecular – Universidade Federal do Rio Grande do Sul (UFRGS);
3. Serviço de Genética Médica – HCPA;
4. Departamento de Medicina Interna – UFRGS;
5. Depto. de Bioquímica – UFRGS, Porto Alegre, RS, Brazil.

Running Title: *C9orf72* - Potential Modifier of MJD/SCA3

Corresponding author:

Maria Luiza Saraiva-Pereira, PhD

Laboratório de Identificação Genética – Centro de Pesquisa Experimental

Hospital de Clínicas de Porto Alegre

Ramiro Barcelos 2350

CEP 90035-903, Porto Alegre, RS, Brazil

e-mail: [mlpereira@hcpa.edu.br](mailto:mlpereira@hcpa.edu.br)

## **Abstract**

Machado-Joseph disease, also known as spinocerebellar ataxia type 3 (MJD/SCA3), is the most common form of autosomal dominant ataxia worldwide. While the genetic cause of this disorder has been long been defined, the large discrepancy in age of onset among patients alludes to the influence of additional factors. To date, no such genetic factor has been reliably identified. A hexanucleotide expansion in *C9orf72*, identified in 2011, and established to be the etiology of cases of amyotrophic lateral sclerosis and frontotemporal dementia, raises as a potential candidate for this role. Our aim was to determine whether *C9orf72* contributes to the age of onset in MJD/SCA3 patients from a Brazilian cohort. For this study, 83 MJD/SCA3 patients and 102 healthy controls were recruited. Based on prior studies, we divided detected normal alleles into two groups: small (<7 repeats) and intermediate (7-30 repeats). No pathogenic expansions (>30 repeats) were detected in the samples. We observed no difference in intermediate allele frequency when compared to healthy controls, and there was no difference in allele frequency between the groups of early- and late-onset patients. We conclude that *C9orf72* does not act as a modifier of age of onset in Brazilian MJD/SCA3 patients.

## **Keywords**

Machado-Joseph disease, spinocerebellar ataxia type 3, age of onset, *C9orf72*, modifier

## Introduction

Neurodegenerative diseases comprise a large group of clinically and pathologically heterogeneous disorders resulting from a progressive loss neurons or neural function (degeneration). Those disorders are in general a series of adult onset or advanced-adult onset diseases, often caused by abnormal protein aggregation in tissues of the nervous system. As the worldwide population ages, neurodegenerative diseases present a large interest to health community. Mounting evidence demonstrates that these disorders, which include Alzheimer's disease (AD), Huntington's disease (HD), the amyotrophic-lateral-sclerosis-frontotemporal-dementia spectrum (ALS-FTD) and spinocerebellar ataxias (SCAs), among others, share common molecular pathways for neurodegeneration<sup>1</sup>. Causative genes of one neurodegenerative disease have been shown to act as modifiers in other diseases<sup>2</sup>.

Spinocerebellar ataxias are a group of neurodegenerative disorders defined by progressive degeneration of neural tissue within areas related to the regulation of movement and coordination, primarily the cerebellum. Among them, Machado-Joseph disease or Spinocerebellar Ataxia type 3 (MJD/SCA3) is the most prevalent at both in Brazil<sup>3</sup> and worldwide<sup>4-6</sup>. The etiology of the disease is an unstable CAG trinucleotide expansion in the *ATXN3* gene<sup>7</sup>, with patients showing a wide range of heterogeneous symptoms, including age of onset. A decrease in the latter has been associated to larger number of repeats<sup>8,9</sup>. However, this factor could only explain from 45% to 76% of the overall variation in age of onset. In an effort to determine which other factors influence this characteristic a large scale study of *CHIP* polymorphisms was carried out, but were unable to significantly alter the percentage of variation explained<sup>2</sup>.

A large scale study of the EUROSCA cohort demonstrated that three polyglutamine-disease-causing genes identified *ATXN2*, *ATN1* and *HTT* as possible modifiers, with intermediary and larger alleles of the former two leading to earlier age of onset, and short alleles of the latter delaying the start of the disease<sup>10</sup>. However, these results were not replicated latter studies in 4 different populations. The role of *ATXN2* was again brought into question when two studies conducted by our group on the same cohort showed contradictory results when analyzed one decade later<sup>3,11</sup>. The association of a specific allele of *APOE* produced similar contradictory

results in three separate studies<sup>12-14</sup>. Previous studies undertaken by our group have also tackled this question. The potential effect of hypomethylation as a neuroprotective factor, while promising, was discarded when the size of the normal allele was taken into account<sup>15</sup>. Similarly, expansions within *ATXN6* and *ATN1* were ruled out as a modifying factor in MJD/SCA3 age of onset<sup>11</sup>. Thus, the question of modifying factors in MJD/SCA3 remains an open one.

Due to intense research into the underlying mechanism of neurodegeneration, a molecular link between ALS and MJD/SCA3 was discovered in the form of gp78<sup>15</sup>. This new shared pathway paved the way for a potential investigation into other aspects of ALS which may act as modifiers in MJD/SCA3. In 2011 a hexanucleotide expansion (GGGGCC) in the first intron of *C9orf72* was identified as the causative agent in a substantial number of patients affected by the ALS-FTD spectrum<sup>16,17</sup>. The underlying mechanism by which this mutation causes the development of the C9 ALS-FTD syndrome is unknown, with preliminary experimental results pointing to a dual function of RNA *foci* toxicity<sup>18</sup>, dipeptide aggregation<sup>19</sup>, and loss of *C9orf72* function<sup>20</sup>, mechanisms which are similar to the proposed pathways by which *ATXN3* causes MJD/SCA3. Aside from this molecular foundation, a series of *C9orf72* hexanucleotide expansion carriers develop symptoms rare to the ALS-FTD spectrum, but common among SCA patients<sup>21-23</sup>.

Taking these molecular and clinical observations into consideration, our aim was to investigate the potential effect of different *C9orf72* allele lengths on age of onset in a cohort of MJD/SCA3 Brazilian patients.

## Materials and Methods

### Subjects

A total of 83 MJD/SCA3 patients that were previously confirmed by molecular analysis were included in this study. Those patients were fully described in previous studies conducted by our group<sup>24,25</sup>. We have also included 102 controls with no history of ataxia in this study. Both, patients and controls, were recruited from the Hospital de Clínicas de Porto Alegre (HCPA – Clinical Hospital of Porto Alegre). Due to the relative rarity of the disease, our sample size was limited to a convenience samples which we were able to obtain, as such, some patients were relatives.

### Ethics Statement

This project has been approved by the Ethics Committee for Research at HCPA (Comissão de Ética em Pesquisa do HCPA), project #08-584. This regulatory body follows the guidelines outlined by the Code of Ethics of the World Medical Association in the Declaration of Helsinki. All participants provided written consent prior to material collection.

### Genetic Analysis

Patient gDNA was obtained via the peripheral blood utilizing the salting-out method. DNA samples were stored at -20°C until analysis. Following previously established protocols<sup>16,26</sup>, we first detected the size of non-pathogenic alleles in subjects using a FAM-labelled fluorescent polymerase chain reaction (PCR) system (Tables S1). Selected samples were sequenced using standard Sanger sequencing to confirm allele size. Patients identified as non-heterozygous potential carriers were later screened using a repeat-primed PCR to detect larger alleles outside of the bounds of detection of the first step (Table S1). A positive *C9orf72* expansion sample of a patient with ALS was utilized as a control. For MJD/SCA3 patients, allele size was determined using fluorescent fragment analysis identifying patients of MJD/SCA3. All fragment length

analyses were performed using an ABI 3130xl genetic analyzer, Sequence Analysis software and the GeneMapper *ID* software, v. 3.2. Based on previous literature, we define an expanded allele as one with more than 30 GGGGCC repeats.

### Data Interpretation

Utilizing obtained allelic data, both MJD/SCA3 and control samples were divided into three potential groups based on the *C9orf72* alleles detected: 1) S - “short alleles” – patients who only carried *C9orf72* alleles with less than 7 G<sub>4</sub>C<sub>2</sub> repeats; 2) I - “intermediate allele carriers” – patients who carried at least one non-pathogenic allele with 7 or more G<sub>4</sub>C<sub>2</sub> repeats; and 3) exp – “expansion carriers” – patients carrying the expansion (>30 repeats). This was done in accordance with studies showing that as the number of repeats increases, the expression levels of *C9orf72* will decrease, potentially causing eventual haploinsufficiency<sup>27</sup>.

### Statistical Analysis

Analysis of distribution of *C9orf72* alleles was carried out utilizing a Mann-Whitney U test. CAG repeat length in *ATXN3* in MJD/SCA3 patients was associated with age of onset using a linear regression model. A multiple linear regression analysis was utilized to determine if the inclusion of *C9orf72* alleles would improve the age of onset model. Extreme phenotypes in the age of onset, defined as patients whose residual age of onset was at least 1.5 standard deviations earlier or later than the mean age of onset, were identified and divided into two groups: early- and late-onset. To determine if these extreme phenotypes were correlated with the presence of certain alleles in *C9orf72*, we performed a Mann-Whitney U test comparing the frequencies of these alleles in early- and late-onset MJD/SCA3 patients. All statistical analyses were performed using PASW Statistics 18 (Release 18.0.0).

## Results

Data collected concerning the *C9orf72* alleles is presented in Table 1, along with the sex ratios of the cohorts included in this study. We did not detect pathological expansions *C9orf72* mutations in either cohort. In healthy controls, the range of hexanucleotide length varied from 2 to 20 repeats, whereas MJD/SCA3 patients presented a range from 2 to 19 repeats. Interestingly, despite the influence of European colonization on the Brazilian population, our population did not strictly follow the trimodal European 2-5-8 distribution, with 6, 7, 9 and 10 repeats also being relatively frequent alleles. We did not find a significant difference in the distribution of repeat numbers in the larger, smaller nor average allele lengths in MJD/SCA3 patients and the normal controls (Fig. 1). There was also no difference in the number of individuals attribute to the I group as opposed to the S group among the two cohorts.

Regression analysis showed that *C9orf72* allele size did not correlate with age of onset ( $p = 0.60$ ), nor did an association appear if analysing the small and large alleles separately ( $p = 0.31$  and  $p = 0.83$ , respectively). There was no significant difference between the two extreme phenotype groups in either the number of repeats in either allele, nor in the demographic information (Table 2). When analysing whether *C9orf72* had an effect on extreme phenotypes of MJD/SCA3, we were unable to find a significant correlation between larger allele length and age of onset in neither early- nor late-onset patients (Fig. 2).

## Discussion

The identification of modifying factors of age of onset in patients of neurodegenerative diseases is crucial to investigations into the pathophysiology of MJD/SCA3 and the clinical implications for patients resulting from the variations detected in the underlying molecular mechanisms of the disease. The description of a hexanucleotide expansion in the first intron of *C9orf72* as the causative mutation of the ALS-FTD continuum rippled within neurogenetics research, with investigators failing to identify a link between the mutation and other forms of dementia, such as AD<sup>28</sup> or among other pure neurodegenerative disorders, while also relating the mutation to less well-described disorders, such as HD phenocopies. However, the absence of a potential link between the *C9orf72* hexanucleotide expansion and ataxic disorders remains less conclusive, with select studies showing carriers affected with gait ataxia and pure cerebellar degeneration. Besides, the contradictory evidence concerning expansion length and age of onset in disease, along with implications of intermediate-sized alleles as being risk factors in Parkinson's disease<sup>29</sup> indicate that even wild-type *C9orf72* itself may play a role in the process of neurodegeneration.

A unique study performed in 2015 in a Chinese cohort of MJD/SCA3 patients did not demonstrate a direct correlation between *C9orf72* repeat length and age of onset<sup>30</sup>. However, the authors did not rule out a potential neuroprotective role of shorter (<7 repeats) alleles<sup>30</sup>. Despite this initial study, we cannot apply these conclusions to our population, seeing that the distribution of normal-length alleles differs between Asian and European populations. Considering that Asian populations frequently demonstrate fewer expansion carriers, as well as shorter normal-length alleles, the effect of *C9orf72* may be differently pronounced in our population, seeing as the genetic ancestry in the Rio Grande do Sul region is predominantly European<sup>31,32</sup>.

Here, we identify the allelic distribution of *C9orf72* hexanucleotide repeats to be similar to both European and Asian cohorts. While the 2-repeat allele is universal in predominance worldwide as well as in our study (57%), Brazilian patients did not follow the strict trimodal pattern of 2-5-8 repeats previously observed in European populations<sup>27</sup>. The number of patients



showing 9 repeats (5%) was over half of those showing 8 repeats (9%). Individuals with 6, 7 and 10 repeats were also relatively prevalent within the population (4% each).

We did not observe a correlation between *C9orf72* allele length and age of onset in MJD/SCA3 patients. Differently from Wang, *et al.* (2015), we did not notice any effect of the intermediate length alleles on the age of onset in MJD/SCA3 patients, despite our population showing a similar proportion of intermediate allele carriers. As such, we question the role of *C9orf72* as a neuroprotective factor for age of onset in this disease. To further support our view that *C9orf72* does not modify age of onset in MJD/SCA3 patients, our analysis of extreme early and late onset patients showed no statistical difference in overall average allele length, small allele length, nor large allele length between the two groups. This was done in order to eliminate any mitigating statistical factors within the expected age-of-onset MJD/SCA3 patient group. While there was no significant correlation between the length of the longer *C9orf72* allele, we did note a rather interesting tendency of longer alleles to induce a much more extreme phenotype (i.e. early-onset patients with longer alleles had an earlier AO than early-onset patients with shorter alleles and vice-versa for late-onset patients). However, this could simply be an artifact of the relative small sample size of both extreme phenotype groups, which is to be expected, and this also serves as evidence against a hypothetical role of *C9orf72* as a potential modifier of MJD/SCA3.

In conclusion, we provide evidence that *C9orf72* does not act as a modifying factor of age of onset in MJD/SCA3 in our cohort. We acknowledge that the *C9orf72* hexanucleotide expansion has been implicated in ataxic disease, however we did not find evidence of it being a direct modifying factor in MJD/SCA3.

## References

1. Ross CA, Poirier MA. Protein aggregation and neurodegenerative disease. *Nat Med*. 2004;10 Suppl:S10-17. doi:10.1038/nm1066
2. França MC, Emmel VE, D'Abreu A, et al. Normal ATXN3 Allele but Not CHIP Polymorphisms Modulates Age at Onset in Machado–Joseph Disease. *Front Neurol*. 2012;3. doi:10.3389/fneur.2012.00164
3. de Castilhos RM, Furtado GV, Gheno TC, et al. Spinocerebellar Ataxias in Brazil—Frequencies and Modulating Effects of Related Genes. *The Cerebellum*. 2014;13(1):17-28. doi:10.1007/s12311-013-0510-y
4. Gan S-R, Shi S-S, Wu J-J, et al. High frequency of Machado-Joseph disease identified in Southeastern Chinese kindreds with spinocerebellar ataxia. *BMC Med Genet*. 2010;11:47. doi:10.1186/1471-2350-11-47
5. Ranum LP, Lundgren JK, Schut LJ, et al. Spinocerebellar ataxia type 1 and Machado-Joseph disease: incidence of CAG expansions among adult-onset ataxia patients from 311 families with dominant, recessive, or sporadic ataxia. *Am J Hum Genet*. 1995;57(3):603-608.
6. Schöls L, Amoiridis G, Langkafel M, et al. Machado-Joseph disease mutations as the genetic basis of most spinocerebellar ataxias in Germany. *J Neurol Neurosurg Psychiatry*. 1995;59(4):449-450.
7. Kawaguchi Y, Okamoto T, Taniwaki M, et al. CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. *Nat Genet*. 1994;8(3):221-228. doi:10.1038/ng1194-221
8. Maciel P, Gaspar C, DeStefano AL, et al. Correlation between CAG repeat length and clinical features in Machado-Joseph disease. *Am J Hum Genet*. 1995;57(1):54-61.
9. Maruyama H, Nakamura S, Matsuyama Z, et al. Molecular features of the CAG repeats and clinical manifestation of Machado—Joseph disease. *Hum Mol Genet*. 1995;4(5):807–812.
10. Tezenas du Montcel S, Durr A, Bauer P, et al. Modulation of the age at onset in spinocerebellar ataxia by CAG tracts in various genes. *Brain J Neurol*. 2014;137(Pt 9):2444-2455. doi:10.1093/brain/awu174

11. Jardim L, Silveira I, Pereira ML, et al. Searching for modulating effects of SCA2, SCA6 and DRPLA CAG tracts on the Machado-Joseph disease (SCA3) phenotype. *Acta Neurol Scand.* 2003;107(3):211-214.
12. Peng H, Wang C, Chen Z, et al. APOE  $\epsilon$ 2 allele may decrease the age at onset in patients with spinocerebellar ataxia type 3 or Machado-Joseph disease from the Chinese Han population. *Neurobiol Aging.* 2014;35(9):2179.e15-18. doi:10.1016/j.neurobiolaging.2014.03.020
13. Bettencourt C, Raposo M, Kazachkova N, et al. The APOE  $\epsilon$ 2 allele increases the risk of earlier age at onset in Machado-Joseph disease. *Arch Neurol.* 2011;68(12):1580-1583. doi:10.1001/archneurol.2011.636
14. Zhou Q, Ni W, Dong Y, Wang N, Gan S-R, Wu Z-Y. The Role of Apolipoprotein E as a Risk Factor for an Earlier Age at Onset for Machado-Joseph Disease Is Doubtful. Li X-J, ed. *PLoS ONE.* 2014;9(11):e111356. doi:10.1371/journal.pone.0111356
15. Ying Z, Wang H, Fan H, et al. Gp78, an ER associated E3, promotes SOD1 and ataxin-3 degradation. *Hum Mol Genet.* 2009;18(22):4268-4281. doi:10.1093/hmg/ddp380
16. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, et al. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron.* 2011;72(2):245-256. doi:10.1016/j.neuron.2011.09.011
17. Renton AE, Majounie E, Waite A, et al. A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron.* 2011;72(2):257-268. doi:10.1016/j.neuron.2011.09.010
18. Gendron TF, Belzil VV, Zhang Y-J, Petrucelli L. Mechanisms of toxicity in C9FTLD/ALS. *Acta Neuropathol (Berl).* 2014;127(3):359–376.
19. Tabet R, Schaeffer L, Freyermuth F, et al. CUG initiation and frameshifting enable production of dipeptide repeat proteins from ALS/FTD C9ORF72 transcripts. *Nat Commun.* 2018;9. doi:10.1038/s41467-017-02643-5
20. Gijssels I, Van Langenhove T, van der Zee J, et al. A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: a gene identification study. *Lancet Neurol.* 2012;11(1):54-65. doi:10.1016/S1474-4422(11)70261-7

21. Lindquist S, Duno M, Batbayli M, et al. Corticobasal and ataxia syndromes widen the spectrum of C9ORF72 hexanucleotide expansion disease. *Clin Genet*. 2013;83(3):279-283. doi:10.1111/j.1399-0004.2012.01903.x
22. Corcia P, Vourc'h P, Guennoc A-M, et al. Pure cerebellar ataxia linked to large C9orf72 repeat expansion. *Amyotroph Lateral Scler Front Degener*. 2016;17(3-4):301-303. doi:10.3109/21678421.2015.1113298
23. Meloni M, Farris R, Solla P, Mascia MM, Marrosu F, Cannas A. C9ORF72 Intermediate Repeat Expansion in a Patient With Psychiatric Disorders and Progressive Cerebellar Ataxia. *The Neurologist*. 2017;22(6):245. doi:10.1097/NRL.0000000000000147
24. Donis KC, Saute JAM, Krum-Santos AC, et al. Spinocerebellar ataxia type 3/Machado-Joseph disease starting before adolescence. *Neurogenetics*. 2016;17(2):107-113. doi:10.1007/s10048-016-0473-5
25. Souza GN, Kersting N, Krum-Santos AC, et al. Spinocerebellar ataxia type 3/Machado-Joseph disease: segregation patterns and factors influencing instability of expanded CAG transmissions. *Clin Genet*. 2016;90(2):134-140. doi:10.1111/cge.12719
26. Jiao B, Guo J-F, Wang Y-Q, et al. C9orf72 mutation is rare in Alzheimer's disease, Parkinson's disease, and essential tremor in China. *Front Cell Neurosci*. 2013;7:164. doi:10.3389/fncel.2013.00164
27. van der Zee J, Gijssels I, Dillen L, et al. A pan-European study of the C9orf72 repeat associated with FTLT: geographic prevalence, genomic instability, and intermediate repeats. *Hum Mutat*. 2013;34(2):363-373. doi:10.1002/humu.22244
28. Harms MB, Neumann D, Benitez BA, et al. Parkinson Disease is not associated with C9ORF72 repeat expansions. *Neurobiol Aging*. 2013;34(5):1519.e1-1519.e2. doi:10.1016/j.neurobiolaging.2012.10.001
29. Nuytemans K, Bademci G, Kohli MM, et al. C9ORF72 Intermediate Repeat Copies Are a Significant Risk Factor for Parkinson Disease. *Ann Hum Genet*. 2013;77(5):351-363. doi:10.1111/ahg.12033
30. Wang C, Chen Z, Yang F, et al. Analysis of the GGGGCC Repeat Expansions of the C9orf72 Gene in SCA3/MJD Patients from China. Gonzalez-Alegre P, ed. *PLOS ONE*. 2015;10(6):e0130336. doi:10.1371/journal.pone.0130336

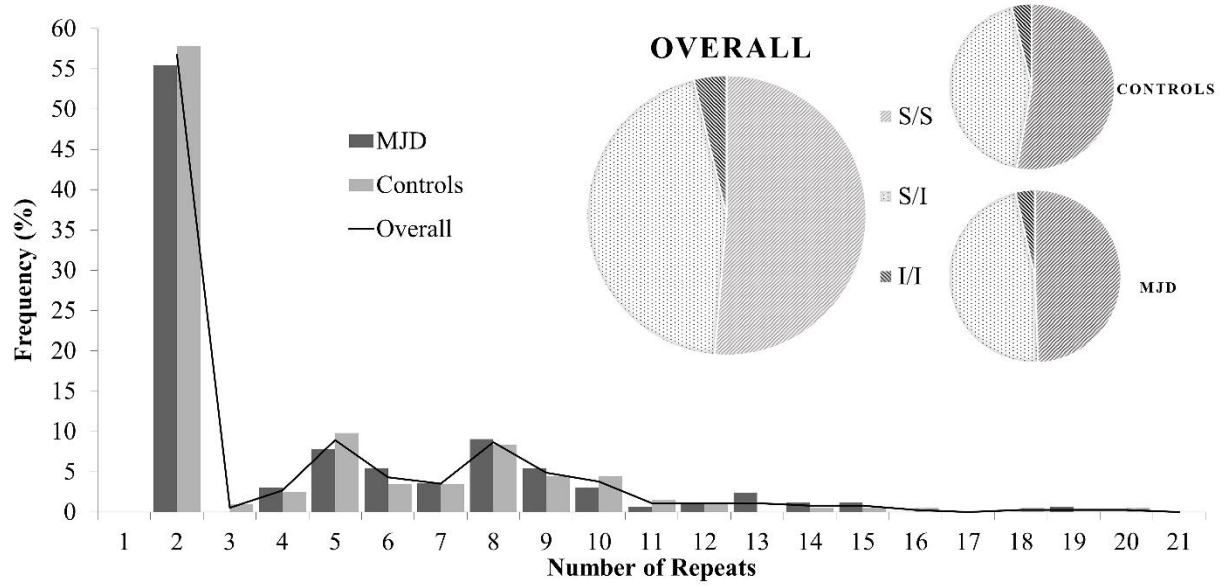
31. Manta FS de N, Pereira R, Vianna R, et al. Revisiting the Genetic Ancestry of Brazilians Using Autosomal AIM-Indels. *PLOS ONE*. 2013;8(9):e75145. doi:10.1371/journal.pone.0075145
32. Marrero AR, Das Neves Leite FP, De Almeida Carvalho B, et al. Heterogeneity of the genome ancestry of individuals classified as White in the state of Rio Grande do Sul, Brazil. *Am J Hum Biol Off J Hum Biol Counc*. 2005;17(4):496-506. doi:10.1002/ajhb.20404

## Legends to Figures

**Figure 1. Frequency of *C9orf72* alleles.** Number of G<sub>4</sub>C<sub>2</sub> repeats observed in the first intron of *C9orf72* in MJD/SCA3 patients (alleles n = 166), healthy controls (alleles n = 204) and a summation representing the overall population of RS, Brazil (alleles n = 370). Inset - proportion of intermediate allele carriers to non-carriers.

**Figure 2. Distribution of G<sub>4</sub>C<sub>2</sub> repeats in early and late onset MJD/SCA3 patients according to deviation from the expected age of onset.** The p-values for correlation between number of repeats and standardized residual age of onset in early- and late-onset patients were 0.35 and 0.80, respectively.

**Figure 1**



**Figure 2**





**Table 1.** Demographic information and average number of repeats in alleles in MJD/SCA3 patients and in controls.

<b>Variable</b>	<b>MJD Patients</b>	<b>Controls</b>	<b><i>p</i> value</b>
Samples, n	83	102	
Male, n (%) <sup>†</sup>	38 (45.78%)	47 (46.08%)	0.97
Number of repeats in the small allele, median ± SD <sup>‡</sup>	2 ± 2	2 ± 2	0.89
Number of repeats in the larger allele, median ± SD <sup>‡</sup>	7 ± 4	6 ± 4	0.33
Intermediate allele group, n (%) <sup>†</sup>	42 (50.60%)	48 (47.06%)	0.63

SD = standard deviation; <sup>†</sup> $\chi^2$ - test; <sup>‡</sup>Mann-Whitney U

**Table 2.** Demographic information and average number of repeats in alleles in early-onset and late-onset MJD/SCA3 patients.

<b>Variable</b>	<b>Early Onset</b>	<b>Late Onset</b>	<b><i>p</i> value</b>
Samples, n	36	16	
Male, n (%)†	16 (44.44%)	7 (43.75%)	0.92
Number of repeats in the small allele, median ± SD‡	2 ± 1	2 ± 2	0.70
Number of repeats in the larger allele, median ± SD‡	8 ± 4	6 ± 4	0.49
Intermediate allele group, n (%)†	19 (52.78%)	7 (43.75%)	0.61

SD = standard deviation; † $\chi^2$ - test; ‡Mann-Whitney U

## Supplementary Material

**Supplementary Table S1.** Primers utilized in this study.

<b>Primer Name</b>	<b>Primer Sequence (5'→3')</b>	<b>Procedure</b>
C9SeqRepFor	CAAGGAGGGAAACAACCGCAGCC	Hexanucleotide
C9SeqRepRev	GCAGGCACCGCAACCGCAG	Repeat Region Sequencing
C9orf72F	FAM-CCCCACTACTTGCTCTCACAG	Allele Size
C9orf72R	ATGAGAGGGAAAGTAAAAATGCGTC	Determination
MRX-F	FAM- TGTA AACGACGGCCAGTCAAGGAGGGAAACAACC GCAGCC	Repeat-primed PCR <sup>26</sup>
MRX-M13R	CAGGAAACAGCTATGACCGGGCCCGCCCCGACCACGCC CCGGCCCCGGCCCCGG	
MRX-R1	CAAGGAAACAGCTATGACC	

#### IV. Discussão e Conclusões

Durante a realização do nosso estudo, efetuamos uma revisão abrangente da literatura, considerando os possíveis papéis de *C9orf72* no desenvolvimento de sintomatologia atáxica. A função principal da proteína *C9orf72* continua desconhecida, o que dificulta a elucidação do mecanismo de desenvolvimento das doenças relacionadas com esse gene. Cada uma das três vias propostas para a ação da proteína *C9orf72* como agente causador da doença podem ser mecanisticamente relacionados com as doenças atáxicas. Um dos mais promissores nessa área de pesquisa é o papel tóxico dos agregados dipeptídicos para a via de agregação (Brettschneider et al. 2012; Hjerpe et al. 2016), que resulta na remoção dos grandes aglomerados proteicos, tais como os que se formam nas doenças poliQ. Mais detalhes podem ser vistos no nosso artigo de revisão.

Além das possíveis vias moleculares de atuação do gene *C9orf72* para o desenvolvimento de ataxias, durante a última década vários estudos clínicos surgiram relacionando esse gene com esse grupo heterogêneo de doenças, com resultados contraditórios. Enquanto relatos de casos individuais mostram uma possível relação entre esse gene e sintomas atáxicos (Goldman et al. 2014; Corcia et al. 2016), estudos abrangentes incluindo um grande número de pacientes negam essa relação (Fogel et al. 2012; He et al. 2016). É importante notar que tipicamente esses estudos de grande escala só consideram a expansão patogênica, em vez de também analisar os alelos normais como fator etiológico.

O único estudo sobre a possível correlação de *C9orf72* com a MJD/SCA3 foi realizado em 2015, numa coorte chinesa. Nesse estudo foi descoberto um possível papel neuroprotetor dos alelos mais curtos nos indivíduos com essa doença (Wang et al. 2015). Levando em consideração a diferença na distribuição dos alelos normais do *C9orf72* nas populações asiáticas e europeias (van der Zee et al. 2013; He et al. 2015), e o fato que populações diferentes podem manifestar ii distintas (Tezenas du Montcel et al. 2014), não seria adequado a aplicação direta desses resultados aos pacientes brasileiros de MJD/SCA3.

Portanto, identificamos inicialmente uma diferença no perfil de frequências alélicas na nossa população comparada com a distribuição trimodal estrita europeia. Isso pode ser devido a influência das outras populações (tanto colonizadoras, como indígenas) na população

brasileira (Manta et al. 2013), mas ainda foi possível observar a grande influência europeia no perfil genético da nossa população. Nenhuma expansão patológica foi detectada nos indivíduos amostrados, conforme esperado.

Em termos da correlação da *C9orf72* com ii nos pacientes de MJD/SCA3, diferentemente dos resultados do Wang, *et al.*, (2015), não observamos influência neuroprotetora dos alelos curtos. Para garantir que isso não seria um resultado que sofre viés do grande número de pacientes com ii estatisticamente esperado, foi utilizada a técnica de amostragem de fenótipos extremos. Levando em consideração só os pacientes com ii muito precoce ou tardia, nenhuma relação do *C9orf72* foi descoberta com nenhum dos grupos. Dentro dos fenótipos extremos, não detectamos a influência do tamanho dos alelos normais sobre a gravidade fenotípica.

Por fim, considerando as informações geradas por esse estudo, na coorte analisada, o tamanho do trato das repetições hexanucleotídicas do gene *C9orf72* não foi identificado como agente modificador da idade de início doença do Machado-Joseph.

## V. Referências Bibliográficas

- Arrasate M, Mitra S, Schweitzer ES, Segal MR and Finkbeiner S (2004) Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* 431:805–810. doi: 10.1038/nature02998
- Ash PEA, Bieniek KF, Gendron TF, Caulfield T, Lin W-L, DeJesus-Hernandez M, van Blitterswijk MM, Jansen-West K, Paul JW, Rademakers R et al. (2013) Unconventional Translation of C9ORF72 GGGGCC Expansion Generates Insoluble Polypeptides Specific to c9FTD/ALS. *Neuron* 77:639–646. doi: 10.1016/j.neuron.2013.02.004
- Avila J (2010) Common mechanisms in neurodegeneration. *Nat Med* 16:1372–1372. doi: 10.1038/nm1210-1372a
- Bettencourt C, Raposo M, Kazachkova N, Cymbron T, Santos C, Kay T, Vasconcelos J, Maciel P, Donis KC, Saraiva-Pereira ML et al. (2011a) The APOE  $\epsilon$ 2 allele increases the risk of earlier age at onset in Machado-Joseph disease. *Arch Neurol* 68:1580–1583. doi: 10.1001/archneurol.2011.636
- Bettencourt C, Raposo M, Kazachkova N, Santos C, Kay T, Vasconcelos J, Maciel P, Donis KC, Saraiva-Pereira ML, Jardim LB et al. (2012) Sequence Analysis of 5' Regulatory Regions of the Machado–Joseph Disease Gene (ATXN3). *The Cerebellum* 11:1045–1050. doi: 10.1007/s12311-012-0373-7
- Bettencourt C, Santos C, Coutinho P, Rizzu P, Vasconcelos J, Kay T, Cymbron T, Raposo M, Heutink P and Lima M (2011b) Parkinsonian phenotype in Machado-Joseph disease (MJD/SCA3): a two-case report. *BMC Neurol*. doi: 10.1186/1471-2377-11-131
- Boeve BF, Boylan KB, Graff-Radford NR, DeJesus-Hernandez M, Knopman DS, Pedraza O, Vemuri P, Jones D, Lowe V, Murray ME et al. (2012) Characterization of frontotemporal dementia and/or amyotrophic lateral sclerosis associated with the GGGGCC repeat expansion in C9ORF72. *Brain* 135:765–783. doi: 10.1093/brain/aws004
- Brettschneider J, Van Deerlin VM, Robinson JL, Kwong L, Lee EB, Ali YO, Safren N, Monteiro MJ, Toledo JB, Elman L et al. (2012) Pattern of ubiquilin pathology in ALS and FTL

- indicates presence of C9ORF72 hexanucleotide expansion. *Acta Neuropathol (Berl)* 123:825–839. doi: 10.1007/s00401-012-0970-z
- Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, Hunter K, Stanton VP, Thirion J-P and Hudson T (1992) Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* 68:799–808.
- Buhmann C, Bussopulos A and Oechsner M (2003) Dopaminergic response in Parkinsonian phenotype of Machado-Joseph disease. *Mov Disord* 18:219–221. doi: 10.1002/mds.10322
- Burguete AS, Almeida S, Gao F-B, Kalb R, Akins MR and Bonini NM (2015) GGGGCC microsatellite RNA is neuritically localized, induces branching defects, and perturbs transport granule function. *eLife* 4:e08881. doi: 10.7554/eLife.08881
- Burnett B, Li F and Pittman RN (2003) The polyglutamine neurodegenerative protein ataxin-3 binds polyubiquitylated proteins and has ubiquitin protease activity. *Hum Mol Genet* 12:3195–3205. doi: 10.1093/hmg/ddg344
- Cannas A, Solla P, Borghero G, Floris GL, Chio A, Mascia MM, Modugno N, Muroi A, Orofino G, Di Stefano F et al. (2015) C9ORF72 intermediate repeat expansion in patients affected by atypical parkinsonian syndromes or Parkinson's disease complicated by psychosis or dementia in a Sardinian population. *J Neurol* 262:2498–2503. doi: 10.1007/s00415-015-7873-6
- Chai Y, Shao J, Miller VM, Williams A and Paulson HL (2002) Live-cell imaging reveals divergent intracellular dynamics of polyglutamine disease proteins and supports a sequestration model of pathogenesis. *Proc Natl Acad Sci* 99:9310–9315.
- Chow MKM, Mackay JP, Whisstock JC, Scanlon MJ and Bottomley SP (2004) Structural and functional analysis of the Josephin domain of the polyglutamine protein ataxin-3. *Biochem Biophys Res Commun* 322:387–394. doi: 10.1016/j.bbrc.2004.07.131
- Cintra VP, Lourenço CM, Marques SE, de Oliveira LM, Tumas V and Marques W (2014) Mutational screening of 320 Brazilian patients with autosomal dominant spinocerebellar ataxia. *J Neurol Sci* 347:375–379. doi: 10.1016/j.jns.2014.10.036
- Corcia P, Vourc'h P, Guennoc A-M, Amador MDM, Blasco H, Andres C, Couratier P, Gordon PH and Meininger V (2016) Pure cerebellar ataxia linked to large C9orf72 repeat

- expansion. *Amyotroph Lateral Scler Front Degener* 17:301–303. doi: 10.3109/21678421.2015.1113298
- Coutinho P and Andrade C (1978) Autosomal dominant system degeneration in Portuguese families of the Azores Islands: A new genetic disorder involving cerebellar, pyramidal, extrapyramidal and spinal cord motor functions. *Neurology* 28:703–703. doi: 10.1212/WNL.28.7.703
- de Castilhos RM, Furtado GV, Gheno TC, Schaeffer P, Russo A, Barsottini O, Pedroso JL, Salarini DZ, Vargas FR, Lima MA de FD d et al. (2014) Spinocerebellar Ataxias in Brazil—Frequencies and Modulating Effects of Related Genes. *The Cerebellum* 13:17–28. doi: 10.1007/s12311-013-0510-y
- DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, Nicholson AM, Finch NA, Flynn H, Adamson J et al. (2011) Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* 72:245–256. doi: 10.1016/j.neuron.2011.09.011
- DeStefano AL, Cupples LA, Maciel P, Gaspar C, Radvany J, Dawson DM, Sudarsky L, Corwin L, Coutinho P and MacLeod P (1996) A familial factor independent of CAG repeat length influences age at onset of Machado-Joseph disease. *Am J Hum Genet* 59:119–127.
- Dürr A, Stevanin G, Cancel G, Duyckaerts C, Abbas N, Didierjean O, Chneiweiss H, Benomar A, Lyon-Caen O, Julien J et al. (1996) Spinocerebellar ataxia 3 and Machado-Joseph disease: clinical, molecular, and neuropathological features. *Ann Neurol* 39:490–499. doi: 10.1002/ana.410390411
- Durrenberger PF, Fernando FS, Kashefi SN, Bonnert TP, Seilhean D, Nait-Oumesmar B, Schmitt A, Gebicke-Haerter PJ, Falkai P, Grünblatt E et al. (2015) Common mechanisms in neurodegeneration and neuroinflammation: a BrainNet Europe gene expression microarray study. *J Neural Transm Vienna Austria* 122:1055–1068. doi: 10.1007/s00702-014-1293-0
- Emmel V, Alonso I, Jardim L, Saraiva-Pereira M and Sequeiros J (2011) Does DNA methylation in the promoter region of the ATXN3 gene modify age at onset in MJD (SCA3) patients? *Clin Genet* 79:100–102. doi: 10.1111/j.1399-0004.2010.01508.x



- Farg MA, Sundaramoorthy V, Sultana JM, Yang S, Atkinson RAK, Levina V, Halloran MA, Gleeson PA, Blair IP, Soo KY et al. (2014) C9ORF72, implicated in amyotrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. *Hum Mol Genet* 23:3579–3595. doi: 10.1093/hmg/ddu068
- Ferrari R, Kapogiannis D, Huey ED and Momeni P (2011) FTD and ALS: a tale of two diseases. *Curr Alzheimer Res* 8:273–294.
- Fogel BL, Pribadi M, Pi S, Perlman SL, Geschwind DH and Coppola G (2012) C9ORF72 Expansion is Not a Significant Cause of Sporadic Spinocerebellar Ataxia. *Mov Disord Off J Mov Disord Soc* 27:1835–1836. doi: 10.1002/mds.25245
- França MC, Emmel VE, D'Abreu A, Maurer-Morelli CV, Secolin R, Bonadia LC, da Silva MS, Nucci A, Jardim LB, Saraiva-Pereira ML et al. (2012) Normal ATXN3 Allele but Not CHIP Polymorphisms Modulates Age at Onset in Machado–Joseph Disease. *Front Neurol*. doi: 10.3389/fneur.2012.00164
- Gammon K (2014) Neurodegenerative disease: brain windfall. *Nature* 515:299–300.
- Gendron TF, Belzil VV, Zhang Y-J and Petrucelli L (2014) Mechanisms of toxicity in C9FTLD/ALS. *Acta Neuropathol (Berl)* 127:359–376.
- Gijssels I, Van Langenhove T, van der Zee J, Slegers K, Philtjens S, Kleinberger G, Janssens J, Bettens K, Van Cauwenberghe C, Pereson S et al. (2012) A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: a gene identification study. *Lancet Neurol* 11:54–65. doi: 10.1016/S1474-4422(11)70261-7
- Goldman J, Quinzii CM, Dunning- Broadbent J, Waters C, Mitsumoto H, Brannagan TH, Cosentino S, Huey E, Nagy P and Kuo S-H (2014) Multiple System Atrophy and Amyotrophic Lateral Sclerosis in a Family with Hexanucleotide Repeat Expansions in C9orf72. *JAMA Neurol* 71:771–774. doi: 10.1001/jamaneurol.2013.5762
- González-Zaldívar Y, Vázquez-Mojena Y, Laffita-Mesa JM, Almaguer-Mederos LE, Rodríguez-Labrada R, Sánchez-Cruz G, Aguilera-Rodríguez R, Cruz-Mariño T, Canales-Ochoa N, MacLeod P et al. (2015) Epidemiological, clinical, and molecular characterization of Cuban families with spinocerebellar ataxia type 3/Machado-Joseph disease. *Cerebellum Ataxias* 2:1. doi: 10.1186/s40673-015-0020-4

- Gu W, Ma H, Wang K, Jin M, Zhou Y, Liu X, Wang G and Shen Y (2004) The shortest expanded allele of the MJD1 gene in a Chinese MJD kindred with autonomic dysfunction. *Eur Neurol* 52:107–111. doi: 10.1159/000080221
- Gusella JF, Wexler NS, Conneally PM, Naylor SL, Anderson MA, Tanzi RE, Watkins PC, Ottina K, Wallace MR and Sakaguchi AY (1983) A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* 306:234–238.
- Gwinn-Hardy K, Singleton A, O'Suilleabhain P, Boss M, Nicholl D, Adam A, Hussey J, Critchley P, Hardy J and Farrer M (2001) Spinocerebellar ataxia type 3 phenotypically resembling parkinson disease in a black family. *Arch Neurol* 58:296–299.
- Haass C (2010) Initiation and propagation of neurodegeneration. *Nat Med* 16:1201–1204. doi: 10.1038/nm.2223
- Haberhausen G, Damian MS, Leweke F and Müller U (1995) Spinocerebellar ataxia, type 3 (SCA3) is genetically identical to Machado-Joseph disease (MJD). *J Neurol Sci* 132:71–75.
- Haeusler AR, Donnelly CJ, Periz G, Simko EAJ, Shaw PG, Kim M-S, Maragakis NJ, Troncoso JC, Pandey A, Sattler R et al. (2014) C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* 507:195–200. doi: 10.1038/nature13124
- He J, Tang L, Benyamin B, Shah S, Hemani G, Liu R, Ye S, Liu X, Ma Y, Zhang H et al. (2015) C9orf72 hexanucleotide repeat expansions in Chinese sporadic amyotrophic lateral sclerosis. *Neurobiol Aging* 36:2660.e1–8. doi: 10.1016/j.neurobiolaging.2015.06.002
- He M, Yan W-Q, Zeng S, Liu Z, Zhou Y, Zeng X-F, Zeng J-S, Jiang H, Shen L, Tang B-S et al. (2016) C9ORF72 repeat expansion is not detected in sporadic ataxia patients in mainland China. *J Neurol Sci* 361:181–183. doi: 10.1016/j.jns.2015.12.034
- Hjerpe R, Bett JS, Keuss MJ, Solovyova A, McWilliams TG, Johnson C, Sahu I, Varghese J, Wood N, Wightman M et al. (2016) UBQLN2 Mediates Autophagy-Independent Protein Aggregate Clearance by the Proteasome. *Cell* 166:935–949. doi: 10.1016/j.cell.2016.07.001
- Ichikawa Y, Goto J, Hattori M, Toyoda A, Ishii K, Jeong SY, Hashida H, Masuda N, Ogata K, Kasai F et al. (2001) The genomic structure and expression of MJD, the Machado-Joseph disease gene. *J Hum Genet* 46:413–422. doi: 10.1007/s100380170060

- Igarashi S, Takiyama Y, Cancel G, Rogaeva EA, Sasaki H, Wakisaka A, Zhou Y-X, Takano H, Endo K and Sanpei K (1996) Intergenerational instability of the CAG repeat of the gene for Machado-Joseph disease (MJD1) is affected by the genotype of the normal chromosome: implications for the molecular mechanisms of the instability of the CAG repeat. *Hum Mol Genet* 5:923–932.
- Isono C, Hirano M, Sakamoto H, Ueno S, Kusunoki S and Nakamura Y (2013) Differences in dysphagia between spinocerebellar ataxia type 3 and type 6. *Dysphagia* 28:413–418. doi: 10.1007/s00455-013-9450-4
- Jardim L, Silveira I, Pereira ML, do Céu Moreira M, Mendonça P, Sequeiros J and Giugliani R (2003) Searching for modulating effects of SCA2, SCA6 and DRPLA CAG tracts on the Machado-Joseph disease (SCA3) phenotype. *Acta Neurol Scand* 107:211–214.
- Jardim LB, Pereira ML, Silveira I, Ferro A, Sequeiros J and Giugliani R (2001a) Neurologic findings in Machado-Joseph disease: relation with disease duration, subtypes, and (CAG)n. *Arch Neurol* 58:899–904.
- Jardim LB, Pereira ML, Silveira I, Ferro A, Sequeiros J and Giugliani R (2001b) Machado-Joseph disease in South Brazil: clinical and molecular characterization of kindreds. *Acta Neurol Scand* 104:224–231.
- Jardim LB, Silveira I, Pereira ML, Ferro A, Alonso I, do Céu Moreira M, Mendonça P, Ferreirinha F, Sequeiros J and Giugliani R (2001c) A survey of spinocerebellar ataxia in South Brazil - 66 new cases with Machado-Joseph disease, SCA7, SCA8, or unidentified disease-causing mutations. *J Neurol* 248:870–876.
- Jayadev S and Bird TD (2013) Hereditary ataxias: overview. *Genet Med Off J Am Coll Med Genet* 15:673–683. doi: 10.1038/gim.2013.28
- Jellinger KA (2010) Basic mechanisms of neurodegeneration: a critical update. *J Cell Mol Med* 14:457–487. doi: 10.1111/j.1582-4934.2010.01010.x
- Jennekens FGI (2014) A short history of the notion of neurodegenerative disease. *J Hist Neurosci* 23:85–94. doi: 10.1080/0964704X.2013.809297
- Katsuno M, Adachi H, Waza M, Banno H, Suzuki K, Tanaka F, Doyu M and Sobue G (2006) Pathogenesis, animal models and therapeutics in spinal and bulbar muscular atrophy (SBMA). *Exp Neurol* 200:8–18. doi: 10.1016/j.expneurol.2006.01.021

- Kawaguchi Y, Okamoto T, Taniwaki M, Aizawa M, Inoue M, Katayama S, Kawakami H, Nakamura S, Nishimura M and Akiguchi I (1994) CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. *Nat Genet* 8:221–228. doi: 10.1038/ng1194-221
- Kawai Y, Takeda A, Abe Y, Washimi Y, Tanaka F and Sobue G (2004) Cognitive impairments in Machado-Joseph disease. *Arch Neurol* 61:1757–1760. doi: 10.1001/archneur.61.11.1757
- Klausen PR (2007) Trends in birth defects research. Nova Biomedical Books, New York
- Koide R, Ikeuchi T, Onodera O, Tanaka H, Igarashi S, Endo K, Takahashi H, Kondo R, Ishikawa A and Hayashi T (1994) Unstable expansion of CAG repeat in hereditary dentatorubral-pallidolusian atrophy (DRPLA). *Nat Genet* 6:9–13. doi: 10.1038/ng0194-9
- Koppers M, Blokhuis AM, Westeneng H, Terpstra ML, Zundel CAC, Vieira de Sá R, Schellevis RD, Waite AJ, Blake DJ, Veldink JH et al. (2015) C9orf72 ablation in mice does not cause motor neuron degeneration or motor deficits. *Ann Neurol* 78:426–438. doi: 10.1002/ana.24453
- Kovacs GG (2016) Molecular Pathological Classification of Neurodegenerative Diseases: Turning towards Precision Medicine. *Int J Mol Sci*. doi: 10.3390/ijms17020189
- Lagier-Tourenne C, Baughn M, Rigo F, Sun S, Liu P, Li H-R, Jiang J, Watt AT, Chun S and Katz M (2013) Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proc Natl Acad Sci* 110:E4530–E4539.
- Lee WY, Jin DK, Oh MR, Lee JE, Song SM, Lee EA, Kim G-M, Chung JS and Lee KH (2003) Frequency analysis and clinical characterization of spinocerebellar ataxia types 1, 2, 3, 6, and 7 in Korean patients. *Arch Neurol* 60:858–863. doi: 10.1001/archneur.60.6.858
- Levine TP, Daniels RD, Gatta AT, Wong LH and Hayes MJ (2013) The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. *Bioinformatics* 29:499–503. doi: 10.1093/bioinformatics/bts725
- Li F, Macfarlan T, Pittman RN and Chakravarti D (2002) Ataxin-3 Is a Histone-binding Protein with Two Independent Transcriptional Corepressor Activities. *J Biol Chem* 277:45004–45012. doi: 10.1074/jbc.M205259200
- Li S and Li X-J (2006) Multiple pathways contribute to the pathogenesis of Huntington disease. *Mol Neurodegener* 1:19. doi: 10.1186/1750-1326-1-19

- Lin MT and Beal MF (2006) Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. *Nature* 443:787–795. doi: 10.1038/nature05292
- Lin X, Miller JW, Mankodi A, Kanadia RN, Yuan Y, Moxley RT, Swanson MS and Thornton CA (2006) Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. *Hum Mol Genet* 15:2087–2097. doi: 10.1093/hmg/ddl132
- Lindquist S, Duno M, Batbayli M, Puschmann A, Braendgaard H, Mardosiene S, Svenstrup K, Pinborg L, Vestergaard K, Hjermind L et al. (2013) Corticobasal and ataxia syndromes widen the spectrum of C9ORF72 hexanucleotide expansion disease. *Clin Genet* 83:279–283. doi: 10.1111/j.1399-0004.2012.01903.x
- Liu F, Liu Q, Lu CX, Cui B, Guo XN, Wang RR, Liu MS, Li XG, Cui L and Zhang X (2016) Identification of a novel loss-of-function C9orf72 splice site mutation in a patient with amyotrophic lateral sclerosis. *Neurobiol Aging* 47:219.e1-219.e5. doi: 10.1016/j.neurobiolaging.2016.07.027
- Long Z, Chen Z, Wang C, Huang F, Peng H, Hou X, Ding D, Ye W, Wang J, Pan Q et al. (2015) Two novel SNPs in ATXN3 3' UTR may decrease age at onset of SCA3/MJD in Chinese patients. *PloS One* 10:e0117488. doi: 10.1371/journal.pone.0117488
- MacDonald ME, Ambrose CM, Duyao MP, Myers RH, Lin C, Srinidhi L, Barnes G, Taylor SA, James M and Groot N (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 72:971–983.
- Maciel P, Costa MC, Ferro A, Rousseau M, Santos CS, Gaspar C, Barros J, Rouleau GA, Coutinho P and Sequeiros J (2001) Improvement in the molecular diagnosis of Machado-Joseph disease. *Arch Neurol* 58:1821–1827.
- Maciel P, Gaspar C, DeStefano AL, Silveira I, Coutinho P, Radvany J, Dawson DM, Sudarsky L, Guimarães J and Loureiro JE (1995) Correlation between CAG repeat length and clinical features in Machado-Joseph disease. *Am J Hum Genet* 57:54–61.
- Mahoney CJ, Beck J, Rohrer JD, Lashley T, Mok K, Shakespeare T, Yeatman T, Warrington EK, Schott JM, Fox NC et al. (2012) Frontotemporal dementia with the C9ORF72 hexanucleotide repeat expansion: clinical, neuroanatomical and neuropathological features. *Brain* 135:736–750. doi: 10.1093/brain/awr361

- Manta FS de N, Pereira R, Vianna R, Araújo ARB de, Gitaí DLG, Silva DA da, Wolfgramm E de V, Pontes I da M, Aguiar JI, Moraes MO et al. (2013) Revisiting the Genetic Ancestry of Brazilians Using Autosomal AIM-Indels. *PLOS ONE* 8:e75145. doi: 10.1371/journal.pone.0075145
- Maruyama H, Nakamura S, Matsuyama Z, Sakai T, Doyu M, Sobue G, Seto M, Tsujihata M, Oh-i T and Nishio T (1995) Molecular features of the CAG repeats and clinical manifestation of Machado—Joseph disease. *Hum Mol Genet* 4:807–812.
- Matilla T, McCall A, Subramony SH and Zoghbi HY (1995) Molecular and clinical correlations in spinocerebellar ataxia type 3 and Machado-Joseph disease. *Ann Neurol* 38:68–72. doi: 10.1002/ana.410380113
- McCampbell A, Taylor JP, Taye AA, Robitschek J, Li M, Walcott J, Merry D, Chai Y, Paulson H, Sobue G et al. (2000) CREB-binding protein sequestration by expanded polyglutamine. *Hum Mol Genet* 9:2197–2202.
- Meloni M, Farris R, Solla P, Mascia MM, Marrosu F and Cannas A (2017) C9ORF72 Intermediate Repeat Expansion in a Patient With Psychiatric Disorders and Progressive Cerebellar Ataxia. *The Neurologist* 22:245. doi: 10.1097/NRL.0000000000000147
- Miller JW, Urbinati CR, Teng-Umuay P, Stenberg MG, Byrne BJ, Thornton CA and Swanson MS (2000) Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J* 19:4439–4448. doi: 10.1093/emboj/19.17.4439
- Miller VM, Nelson RF, Gouvion CM, Williams A, Rodriguez-Lebron E, Harper SQ, Davidson BL, Rebagliati MR and Paulson HL (2005) CHIP suppresses polyglutamine aggregation and toxicity in vitro and in vivo. *J Neurosci Off J Soc Neurosci* 25:9152–9161. doi: 10.1523/JNEUROSCI.3001-05.2005
- Mizielinska S, Gronke S, Niccoli T, Ridler CE, Clayton EL, Devoy A, Moens T, Norona FE, Woollacott IOC, Pietrzyk J et al. (2014) C9orf72 repeat expansions cause neurodegeneration in *Drosophila* through arginine-rich proteins. *Science* 345:1192–1194. doi: 10.1126/science.1256800
- Mori K, Arzberger T, Grässer FA, Gijssels I, May S, Rentzsch K, Weng S-M, Schludi MH, van der Zee J, Cruts M et al. (2013) Bidirectional transcripts of the expanded C9orf72

- hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol (Berl)* 126:881–893. doi: 10.1007/s00401-013-1189-3
- Nagafuchi S, Yanagisawa H, Ohsaki E, Shirayama T, Tadokoro K, Inoue T and Yamada M (1994) Structure and expression of the gene responsible for the triplet repeat disorder, dentatorubral and pallidolusian atrophy (DRPLA). *Nat Genet* 8:177–182. doi: 10.1038/ng1094-177
- Nakano KK, Dawson DM and Spence A (1972) Machado disease A hereditary ataxia in Portuguese emigrants to Massachusetts. *Neurology* 22:49–49.
- Nishiyama K, Murayama S, Goto J, Watanabe M, Hashida H, Kanazawa I, Katayama S, Nakamura S and Nomura Y (1996) Regional and cellular expression of the Machado-Joseph disease gene in brains of normal and affected individuals. *Ann Neurol* 40:776–781.
- Orr HT and Zoghbi HY (2007) Trinucleotide Repeat Disorders. *Annu Rev Neurosci* 30:575–621. doi: 10.1146/annurev.neuro.29.051605.113042
- Padiath QS, Srivastava AK, Roy S, Jain S and Brahmachari SK (2005) Identification of a novel 45 repeat unstable allele associated with a disease phenotype at the MJD1/SCA3 locus. *Am J Med Genet B Neuropsychiatr Genet* 133B:124–126. doi: 10.1002/ajmg.b.30088
- Peng H, Wang C, Chen Z, Sun Z, Jiao B, Li K, Huang F, Hou X, Wang J, Shen L et al. (2014) APOE  $\epsilon$ 2 allele may decrease the age at onset in patients with spinocerebellar ataxia type 3 or Machado-Joseph disease from the Chinese Han population. *Neurobiol Aging* 35:2179.e15–18. doi: 10.1016/j.neurobiolaging.2014.03.020
- Przedborski S, Vila M and Jackson-Lewis V (2003) Series Introduction: Neurodegeneration: What is it and where are we? *J Clin Invest* 111:3–10. doi: 10.1172/JCI200317522
- Pulst SM, Nechiporuk A, Nechiporuk T, Gispert S, Chen XN, Lopes-Cendes I, Pearlman S, Starkman S, Orozco-Diaz G, Lunke A et al. (1996) Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. *Nat Genet* 14:269–276. doi: 10.1038/ng1196-269
- Rabinovici GD and Miller BL (2010) Frontotemporal Lobar Degeneration: Epidemiology, Pathophysiology, Diagnosis and Management. *CNS Drugs* 24:375–398. doi: 10.2165/11533100-000000000-00000

- Raposo M, Bettencourt C, Ramos A, Kazachkova N, Vasconcelos J, Kay T, Bruges-Armas J and Lima M (2017) Promoter Variation and Expression Levels of Inflammatory Genes IL1A, IL1B, IL6 and TNF in Blood of Spinocerebellar Ataxia Type 3 (SCA3) Patients. *NeuroMolecular Med* 19:41–45. doi: 10.1007/s12017-016-8416-8
- Renton AE, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, Schymick JC, Laaksovirta H, van Swieten JC, Myllykangas L et al. (2011) A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron* 72:257–268. doi: 10.1016/j.neuron.2011.09.010
- Robertson AL, Headey SJ, Saunders HM, Ecroyd H, Scanlon MJ, Carver JA and Bottomley SP (2010) Small heat-shock proteins interact with a flanking domain to suppress polyglutamine aggregation. *Proc Natl Acad Sci* 107:10424–10429.
- Rodrigues A-J, Coppola G, Santos C, Costa M do C, Ailion M, Sequeiros J, Geschwind DH and Maciel P (2007) Functional genomics and biochemical characterization of the *C. elegans* orthologue of the Machado-Joseph disease protein ataxin-3. *FASEB J* 21:1126–1136. doi: 10.1096/fj.06-7002com
- Rosen DR, Siddique T, Patterson D, Figlewicz DA, Sapp P, Hentati A, Donaldson D, Goto J, O'Regan JP, Deng H-X et al. (1993) Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* 362:59. doi: 10.1038/362059a0
- Rosenberg RN, Nyhan WL, Bay C and Shore P (1976) Autosomal dominant striatonigral degeneration: A clinical, pathologic, and biochemical study of a new genetic disorder. *Neurology* 26:703–703. doi: 10.1212/WNL.26.8.703
- Ruano L, Melo C, Silva MC and Coutinho P (2014) The Global Epidemiology of Hereditary Ataxia and Spastic Paraplegia: A Systematic Review of Prevalence Studies. *Neuroepidemiology* 42:174–183. doi: 10.1159/000358801
- Saute JAM and Jardim LB (2015) Machado Joseph disease: clinical and genetic aspects, and current treatment. *Expert Opin Orphan Drugs* 3:517–535. doi: 10.1517/21678707.2015.1025747
- Schmidt T, Landwehrmeyer GB, Schmitt I, Trottier Y, Auburger G, Laccone F, Klockgether T, Völpel M, Epplen JT, Schöls L et al. (1998) An Isoform of Ataxin-3 Accumulates in the



- Nucleus of Neuronal Cells in Affected Brain Regions of SCA3 Patients. *Brain Pathol* 8:669–679. doi: 10.1111/j.1750-3639.1998.tb00193.x
- Schmitt I, Linden M, Khazneh H, Evert BO, Breuer P, Klockgether T and Wuellner U (2007) Inactivation of the mouse *Atxn3* (ataxin-3) gene increases protein ubiquitination. *Biochem Biophys Res Commun* 362:734–739. doi: 10.1016/j.bbrc.2007.08.062
- Sen D and Gilbert W (1988) Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* 334:364–366. doi: 10.1038/334364a0
- Shakkottai VG and Fogel BL (2013) Clinical Neurogenetics. *Neurol Clin* 31:987–1007. doi: 10.1016/j.ncl.2013.04.006
- Shimohata T, Nakajima T, Yamada M, Uchida C, Onodera O, Naruse S, Kimura T, Koide R, Nozaki K and Sano Y (2000) Expanded polyglutamine stretches interact with TAFII130, interfering with CREB-dependent transcription. *Nat Genet* 26:29–36.
- St George-Hyslop P, Rogaeva E, Huterer J, Tsuda T, Santos J, Haines JL, Schlumpf K, Rogaev EI, Liang Y and McLachlan DC (1994) Machado-Joseph disease in pedigrees of Azorean descent is linked to chromosome 14. *Am J Hum Genet* 55:120.
- Stevanin G, Cancel G, Dürr A, Chneiweiss H, Dubourg O, Weissenbach J, Cann HM, Agid Y and Brice A (1995) The gene for spinal cerebellar ataxia 3 (SCA3) is located in a region of approximately 3 cM on chromosome 14q24. 3-q32. 2. *Am J Hum Genet* 56:193.
- Stevanin G, Chneiweiss H, Le Guern E, Ravise N, Dürr A, Penet C, Agid Y and Brice A (1993) Genetic heterogeneity of autosomal dominant cerebellar ataxia type I: evidence for the existence of a third locus. *Hum Mol Genet* 2:1483–1485.
- Swinnen B, Bento-Abreu A, Gendron TF, Boeynaems S, Bogaert E, Nuyts R, Timmers M, Scheveneels W, Hersmus N, Wang J et al. (2018) A zebrafish model for *C9orf72* ALS reveals RNA toxicity as a pathogenic mechanism. *Acta Neuropathol (Berl)*. doi: 10.1007/s00401-017-1796-5
- Tabet R, Schaeffer L, Freyermuth F, Jambeau M, Workman M, Lee C-Z, Lin C-C, Jiang J, Jansen-West K, Abou-Hamdan H et al. (2018) CUG initiation and frameshifting enable production of dipeptide repeat proteins from ALS/FTD *C9ORF72* transcripts. *Nat Commun*. doi: 10.1038/s41467-017-02643-5

- Takiyama Y, Nishizawa M, Tanaka H, Kawashima S, Sakamoto H, Karube Y, Shimazaki H, Soutome M, Endo K and Ohta S (1993) The gene for Machado-Joseph disease maps to human chromosome 14q. *Nat Genet* 4:300–304. doi: 10.1038/ng0793-300
- Teive HA (2009) Spinocerebellar ataxias. *Arq Neuropsiquiatr* 67:1133–1142.
- Teive HAG, Munhoz RP, Arruda WO, Lopes-Cendes I, Raskin S, Werneck LC and Ashizawa T (2012) Spinocerebellar ataxias: genotype-phenotype correlations in 104 Brazilian families. *Clin Sao Paulo Braz* 67:443–449.
- Tezenas du Montcel S, Durr A, Bauer P, Figueroa KP, Ichikawa Y, Brussino A, Forlani S, Rakowicz M, Schöls L, Mariotti C et al. (2014) Modulation of the age at onset in spinocerebellar ataxia by CAG tracts in various genes. *Brain J Neurol* 137:2444–2455. doi: 10.1093/brain/awu174
- Thompson LM (2008) Neurodegeneration: A question of balance. *Nature* 452:707. doi: 10.1038/452707a
- Trott A and Houenou LJ (2012) Mini-review: spinocerebellar ataxias: an update of SCA genes. *Recent Pat DNA Gene Seq* 6:115–121.
- Tuite PJ, Rogaeva EA, St George-Hyslop PH and Lang AE (1995) Dopa-responsive parkinsonism phenotype of Machado-Joseph disease: confirmation of 14q CAG expansion. *Ann Neurol* 38:684–687. doi: 10.1002/ana.410380422
- Vale J, Bugalho P, Silveira I, Sequeiros J, Guimarães J and Coutinho P (2010) Autosomal dominant cerebellar ataxia: frequency analysis and clinical characterization of 45 families from Portugal. *Eur J Neurol* 17:124–128. doi: 10.1111/j.1468-1331.2009.02757.x
- van Alfen N, Sinke RJ, Zwarts MJ, Gabreëls-Festen A, Praamstra P, Kremer BP and Horstink MW (2001) Intermediate CAG repeat lengths (53,54) for MJD/SCA3 are associated with an abnormal phenotype. *Ann Neurol* 49:805–807.
- van Blitterswijk M, DeJesus-Hernandez M and Rademakers R (2012) How do C9ORF72 repeat expansions cause ALS and FTD: can we learn from other non-coding repeat expansion disorders? *Curr Opin Neurol* 25:689–700. doi: 10.1097/WCO.0b013e32835a3efb
- van de Warrenburg BPC, Sinke RJ, Verschuuren-Bemelmans CC, Scheffer H, Brunt ER, Ippel PF, Maat-Kievit JA, Dooijes D, Notermans NC, Lindhout D et al. (2002) Spinocerebellar

- ataxias in the Netherlands: prevalence and age at onset variance analysis. *Neurology* 58:702–708.
- van der Zee J, Gijssels I, Dillen L, Van Langenhove T, Theuns J, Engelborghs S, Philtjens S, Vandebulcke M, Sleegers K, Sieben A et al. (2013) A pan-European study of the C9orf72 repeat associated with FTLN: geographic prevalence, genomic instability, and intermediate repeats. *Hum Mutat* 34:363–373. doi: 10.1002/humu.22244
- Wang C, Chen Z, Yang F, Jiao B, Peng H, Shi Y, Wang Y, Huang F, Wang J, Shen L et al. (2015) Analysis of the GGGGCC Repeat Expansions of the C9orf72 Gene in SCA3/MJD Patients from China. *PLOS ONE* 10:e0130336. doi: 10.1371/journal.pone.0130336
- Wang C, Peng H, Li J, Ding D, Chen Z, Long Z, Peng Y, Zhou X, Ye W, Li K et al. (2017) Alteration of methylation status in the ATXN3 gene promoter region is linked to the SCA3/MJD. *Neurobiol Aging* 53:192.e5-192.e10. doi: 10.1016/j.neurobiolaging.2016.12.014
- Wang Y, Du J, Wang J, Chen J, Chen C, Luo Y, Xiao Z, Jiang H, Yan X, Xia K et al. (2009) Six cases of SCA3/MJD patients that mimic hereditary spastic paraplegia in clinic. *J Neurol Sci* 285:121–124. doi: 10.1016/j.jns.2009.06.027
- Warren JD, Rohrer JD and Rossor MN (2013) Clinical review. Frontotemporal dementia. *BMJ* 347:f4827.
- Webster CP, Smith EF, Bauer CS, Moller A, Hautbergue GM, Ferraiuolo L, Myszczyńska MA, Higginbottom A, Walsh MJ, Whitworth AJ et al. (2016) The C9orf72 protein interacts with Rab1a and the ULK1 complex to regulate initiation of autophagy. *EMBO J* 35:1656–1676. doi: 10.15252/embj.201694401
- Wolf AE, Mourão L, França MC, Machado Júnior AJ and Crespo AN (2017) Phonoarticulation in spinocerebellar ataxia type 3. *Eur Arch Oto-Rhino-Laryngol Off J Eur Fed Oto-Rhino-Laryngol Soc EUFOS Affil Ger Soc Oto-Rhino-Laryngol - Head Neck Surg* 274:1139–1145. doi: 10.1007/s00405-016-4240-x
- Woods BT and Schaumburg HH (1972) Nigro-spino-dentatal degeneration with nuclear ophthalmoplegia. A unique and partially treatable clinico-pathological entity. *J Neurol Sci* 17:149–166.

- Xu Z, Poidevin M, Li X, Li Y, Shu L, Nelson DL, Li H, Hales CM, Gearing M and Wingo TS (2013) Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. *Proc Natl Acad Sci* 110:7778–7783.
- Yang W, Dunlap JR, Andrews RB and Wetzel R (2002) Aggregated polyglutamine peptides delivered to nuclei are toxic to mammalian cells. *Hum Mol Genet* 11:2905–2917.
- Ying Z, Wang H, Fan H, Zhu X, Zhou J, Fei E and Wang G (2009) Gp78, an ER associated E3, promotes SOD1 and ataxin-3 degradation. *Hum Mol Genet* 18:4268–4281. doi: 10.1093/hmg/ddp380
- Young AB (2009) Four decades of neurodegenerative disease research: how far we have come! *J Neurosci Off J Soc Neurosci* 29:12722–12728. doi: 10.1523/JNEUROSCI.3767-09.2009
- Zarei S, Carr K, Reiley L, Diaz K, Guerra O, Altamirano PF, Pagani W, Lodin D, Orozco G and China A (2015) A comprehensive review of amyotrophic lateral sclerosis. *Surg Neurol Int* 6:171. doi: 10.4103/2152-7806.169561
- Zhang D, Iyer LM, He F and Aravind L (2012) Discovery of Novel DENN Proteins: Implications for the Evolution of Eukaryotic Intracellular Membrane Structures and Human Disease. *Front Genet*. doi: 10.3389/fgene.2012.00283
- Zhou Q, Ni W, Dong Y, Wang N, Gan S-R and Wu Z-Y (2014) The Role of Apolipoprotein E as a Risk Factor for an Earlier Age at Onset for Machado-Joseph Disease Is Doubtful. *PLoS ONE* 9:e111356. doi: 10.1371/journal.pone.0111356
- Zoghbi HY and Orr HT (1999) Polyglutamine diseases: protein cleavage and aggregation. *Curr Opin Neurobiol* 9:566–570.