

Ana Filipa Bartolomeu Ferreira

Polimorfismo nos *loci ATXN2*,
CACNA1A e *ATN1*: um estudo de
potenciais genes modificadores da
doença de Machado-Joseph em
indivíduos açorianos saudáveis

Ponta Delgada

2012

Ana Filipa Bartolomeu Ferreira

Polimorfismo nos *loci ATXN2, CACNA1A e ATN1*:
um estudo de potenciais genes modificadores da
doença de Machado-Joseph em indivíduos
açorianos saudáveis

Dissertação para obtenção do grau de
Mestre, em Ciências Biomédicas,
apresentada à Universidade dos Açores

Orientadores: Professora Doutora Maria Manuela de Medeiros Lima

Dra. Mafalda Sofia Bastos Raposo



Universidade dos Açores

Departamento de Biologia

Ponta Delgada

2012

“Do not go where the path may
lead, go instead where there is no
path and leave a trail.”

Ralph Waldo Emerson

AGRADECIMENTOS

Antes de mais gostaria de mencionar que a minha caminhada até aqui só foi possível porque tive sempre a meu lado um grupo de pessoas que me ajudou, não só me dando coragem e estímulo para ultrapassar todas as dificuldades, mas também celebrando comigo cada momento de alegria e realização pessoal/profissional vivida ao longo desta jornada.

À minha orientadora Professora Doutora Manuela Lima, por ter acreditado nas minhas capacidades e ter aceitado orientar-me nesta etapa, dando-me a oportunidade de voltar a trabalhar na área de Genética Humana. Agradeço-lhe, também, o facto de ter apresentado a sua total disponibilidade em ajudar e em partilhar comigo o seu conhecimento, assim como, pelas “pancadas sempre educativas”, pelo seu apoio e pela sua amizade. Obrigada por ser uma referência de profissionalismo, mostrando-me “How to do good research”.

À Dra. Mafalda Raposo, também minha orientadora, pela incansável e inestimável ajuda prestada durante a elaboração desta tese. Pela sua total disponibilidade, por aturar as minhas loucuras e pela sua amizade, muito obrigada “chica”.

Ao Dr. Bruno Bettencourt do Serviço Especializado de Epidemiologia e Biologia Molecular do Hospital de Santo Espírito de Angra do Heroísmo pela ajuda prestada na realização prática deste trabalho.

À Doutora Cristina Santos e à Doutora Conceição Bettencourt pela ajuda imprescindível na confirmação dos genótipos.

Ao Ricardo Camarinho, pela boa disposição e bom humor, assim como ajuda prestada sempre que solicitada.

À São, colega de mestrado e de gabinete, pela sua boa disposição e bom humor, pela companhia nas longas horas de produção científica, e pela disponibilidade em ajudar sempre que solicitada. Obrigada por tornares este percurso mais alegre.

A todos os meus amigos, além dos já referidos, pela constante disponibilidade, boa disposição, pelos puxões de orelhas, e amizade verdadeira, muito obrigada!

E por fim...

Mas não por último, quero agradecer às pessoas mais importantes da minha vida, aos meus pais, Júlio e Ana Luísa, pela educação que me deram, que me permitiu ser a pessoa que sou hoje. Por todo o apoio, confiança e amor incondicional, um obrigado muito especial aos meus pais e ao meu irmão Johnny.

ÍNDICE

ÍNDICE DE TABELAS.....	VIII
ÍNDICE DE FIGURAS.....	IX
RESUMO	X
ABSTRACT	XII
CAPÍTULO 1: INTRODUÇÃO	1
1.1. PERTINÊNCIA E OBJECTIVOS.....	2
1.2. REPETIÇÕES TRINUCLEOTÍDICAS.....	2
1.2.1. REPETIÇÕES TRINUCLEOTÍDICAS: ASPECTOS GERAIS	2
1.2.2. REPETIÇÕES TRINUCLEOTÍDICAS ASSOCIADAS A DOENÇAS.....	3
1.3. ATAXIAS ESPINOCEREBELOSAS: ASPECTOS GERAIS	5
1.3.1. ATAXIA ESPINOCEREBELOSAS TIPO 2.....	10
1.3.1.1. CARACTERIZAÇÃO CLÍNICA DA DOENÇA	10
1.3.1.2. GENÉTICA MOLECULAR E PATOGÉNESE	11
1.3.1.3. OUTRAS DOENÇAS ASSOCIADAS AOS ALELOS NORMAIS GRANDES E/OU INTERMEDIÁRIOS NO GENE DA SCA2	12
1.3.2. ATAXIA ESPINOCEREBELOSAS TIPO 6.....	13
1.3.2.1. CARACTERIZAÇÃO CLÍNICA DA DOENÇA	13
1.3.2.2. GENÉTICA MOLECULAR E PATOGÉNESE.....	13
1.3.3. ATROFIA DENTATO-RUBRO-PÁLIDO-LUISIANA.....	14
1.3.3.1. CARACTERIZAÇÃO CLÍNICA DA DOENÇA	14
1.3.3.2. GENÉTICA MOLECULAR E PATOGÉNESE.....	15
1.3.4. ATAXIA ESPINOCEREBELOSAS TIPO 3.....	16
1.3.4.1. CARACTERIZAÇÃO CLÍNICA DA DOENÇA	16
1.3.4.2. GENÉTICA MOLECULAR E PATOGÉNESE.....	17
1.3.4.3. GENES MODIFICADORES DA MJD.....	19
CAPÍTULO 2: SUJEITOS E MÉTODOS.....	21
2.1. AMOSTRA	22
2.2. GENOTIPAGEM.....	22
2.2.1. SEQUENCIAÇÃO	25
2.3. ANÁLISE ESTATÍSTICA	25
CAPÍTULO 3: RESULTADOS	27

3.1. GENOTIPAGEM: COMPARAÇÃO DOS RESULTADOS OBTIDOS POR AF E SEQUENCIAÇÃO	28
3.2. ANÁLISES INTRA-POPULACIONAIS	28
3.2.1. <i>LOCUS</i> DA SCA3	28
3.2.2. <i>LOCUS</i> DA SCA2	31
3.2.3. <i>LOCUS</i> DA SCA6	33
3.2.4. <i>LOCUS</i> DA DRPLA.....	34
3.1. ANÁLISES INTER-POPULACIONAIS.....	35
CAPÍTULO 4: DISCUSSÃO.....	38
4.1. ANÁLISE DE FRAGMENTOS VS. SEQUENCIAÇÃO.....	39
4.2. <i>LOCUS</i> DA SCA3.....	39
4.3. <i>LOCUS</i> DA SCA2.....	42
4.3.1. ASSOCIAÇÃO DAS REPETIÇÕES CAG COM DOENÇAS NEURODEGENERATIVAS	45
4.4. <i>LOCUS</i> DA SCA6.....	46
4.5. <i>LOCUS</i> DA DRPLA	47
CAPÍTULO 5: CONCLUSÕES	50
CAPÍTULO 6: REFERÊNCIAS BIBLIOGRÁFICAS	53

ÍNDICE DE TABELAS

Tabela 1 – Principais características moleculares das doenças de poliQ.....	8
Tabela 2 – Classificação clínica dos três subtipos da doença de Machado-Joseph.....	16
Tabela 3 – <i>Primers</i> e temperatura de <i>annealing</i> usadas na amplificação do motivo CAG nos <i>loci</i> da SCA2, SCA6 e DRPLA.....	24
Tabela 4 – Número de acesso na base de dados (NCBI) da sequência de referência e fórmula utilizada para o cálculo do motivo CAG, em cada <i>locus</i>	25
Tabela 5 – Polimorfismo das repetições CAG e padrão das interrupções CAA e AAG em alelos normais e intermediários em cada <i>locus</i>	31
Tabela 6 – Resultados dos testes exactos de diferenciação populacional e do cálculo da matriz de distância genética linearizadas entre populações (F_{st} Slatkin) ($^+p < 0,05$, $^{++}p < 0,01$, $^{+++}p < 0,005$, $^{++++}p < 0,00005$, $^{+++++}p < 0,00001$, diferenças não significativas (-) e informação não disponível (IND)).	36

ÍNDICE DE FIGURAS

Fig. 1 – Classificação genética das doenças de RTs consoante a localização do motivo repetitivo no gene causal da doença (adaptada de Brouwer <i>et al.</i> , 2009).	5
Fig. 2 – Genótipo de três indivíduos para os <i>loci</i> da SCA2 (A,B), SCA6 (C,D) e DRPLA (E,F) obtidos respectivamente por análise de fragmentos e sequenciação. Cada pico na análise de fragmentos está identificado com o respectivo tamanho de fragmento (pb) e correspondente alelo. Na sequenciação as repetições CTG (por complemento com CAG, uma vez que, foi usado para cada um dos <i>locus</i> o <i>primer</i> reverso) estão sublinhadas. No indivíduo heterozigótico, as repetições CTG do alelo 1 estão delimitadas por um asterisco.....	29
Fig. 3 – Distribuição alélica no <i>locus</i> da MJD/SCA3 em 156 cromossomas de indivíduos saudáveis não aparentados da população dos Açores.	30
Fig. 4 – Distribuição genotípica no <i>locus</i> da MJD/SCA3 de 78 indivíduos saudáveis não aparentados da população dos Açores.	30
Fig. 5 – Distribuição alélica no <i>locus</i> da SCA2 em 156 cromossomas de indivíduos saudáveis não aparentados da população dos Açores.	32
Fig. 6 – Distribuição genotípica no <i>locus</i> da SCA2 de 78 indivíduos saudáveis não aparentados da população dos Açores.	32
Fig. 7 – Distribuição alélica no <i>locus</i> da SCA6 em 156 cromossomas indivíduos saudáveis não aparentados da população dos Açores.	33
Fig. 8 – Distribuição genotípica no <i>locus</i> da SCA6 de 78 indivíduos saudáveis não aparentados da população dos Açores.	34
Fig. 9 – Distribuição alélica no <i>locus</i> da DRPLA em 152 cromossomas de indivíduos saudáveis não aparentados da população dos Açores.	35
Fig. 10 – Distribuição genotípica no <i>locus</i> da DRPLA de 76 indivíduos saudáveis não aparentados da população dos Açores.	35
Fig. 11 – <i>Multidimensional Scaling</i> usando as distâncias genéticas linearizadas (Fst Slatkin), para o <i>locus</i> da SCA2, entre as diferentes populações.	37

RESUMO

Introdução: a doença de Machado-Joseph (MJD) é uma ataxia espinocerebelosa (SCAs) hereditária autossómica dominante, causada por repetições CAG no gene *ATXN3*, que codificam um tracto de poliglutamina, e conduzem a um ganho de função neurotóxica da proteína correspondente, a ataxina-3, por mecanismos que não estão completamente conhecidos. As repetições CAG, porém, explicam apenas parte da variabilidade fenotípica observada nesta doença, sendo que a restante explicação poderá ser fornecida por variação em genes modificadores. As proteínas com um tracto de poliglutamina expandido têm tendência para formar agregados e, portanto, têm potencial para modular o fenótipo das doenças de expansão. De facto, estudos realizados em modelos animais e em doentes mostraram a existência de interacções entre as proteínas nativas codificadas pelos genes causais das diferentes SCAs e a ataxina-3 mutada. Com este trabalho pretendeu-se iniciar a investigação de um estudo caso-controlo da doença de Machado-Joseph, determinando as repetições CAG nos *loci* da SCA2, SCA6 e DRPLA. **Sujeitos e métodos:** analisaram-se 78 indivíduos saudáveis açorianos não aparentados, selecionados por emparelhamento com uma série de doentes de MJD. Os *loci* *ATXN3*, *ATXN2*, *CACNA1A* e *ATN1* foram amplificados por PCR e determinou-se o tamanho dos alelos por análise de fragmentos, num sequenciador automático. Para controlo de qualidade, a confirmação de alguns genótipos foi efectuada por sequenciação. Foram realizadas análises intra-populacionais (frequências alélicas e genotípicas, conformidade com o equilíbrio de Hardy-Weinberg, diversidade genética) e inter-populacionais (diferenciação populacional e distância genética). **Resultados:** verificou-se, para todos os *loci*, uma diferença no tamanho dos alelos obtidos em ambas as técnicas utilizadas e efectuaram-se as correcções do respectivo tamanho do alelo. Na amostra estudada, todos os alelos observados estavam dentro da amplitude normal, tendo-se observado apenas um alelo intermediário no *locus* da SCA2. Todos os *loci* estavam em equilíbrio de Hardy-Weinberg. Os alelos 14 e 23 (29,5%), 22 (89,1%), 11 (35,9%) e 15 (23,7%) foram os mais frequentes nos *loci* da SCA3, SCA2, SCA6 e DRPLA, respectivamente. No *locus* *ATXN2*, os alelos 22 tinham 1-2 interrupções CAA, e os alelos 29 e 33 tinham apenas 1 CAA. Nos *loci* *ATXN2* e *ATN1* observou-se uma frequência relativa elevada de alelos normais grandes (9,6% e 10,5%, respectivamente). **Discussão:** a confirmação, por sequenciação, do tamanho das repetições CAG veio reforçar a noção de que o uso de um marcador alélico na análise de fragmentos é essencial para uma genotipagem correcta. No *locus* *ATXN2*, o alelo intermediário observado poderá implicar uma susceptibilidade acrescida para desenvolver a SCA2; porém, a presença de uma interrupção poderá conferir-lhe uma maior estabilidade. Neste mesmo *locus* os alelos com 24-33 CAGs poderão estar associados com o de risco de desenvolvimento de outras doenças neurodegenerativas. No *locus* *CACNA1A*, a baixa frequência de alelos normais grandes parece ser coerente com a ausência de casos reportados de SCA6 nos

Açores. No *locus ATN1*, a inexistência de correlação entre a elevada frequência de alelos normais grandes e a frequência da DRPLA poderá estar relacionada com o haplótipo associado a estes alelos. **Conclusão:** a genotipagem dos *loci ATXN2, CACNA1A* e *ATN1* em indivíduos açorianos saudáveis permitirá a utilização desta amostra como grupo controlo num estudo de potenciais genes modificadores da MJD, usando doentes da população açoriana.

ABSTRACT

Introduction: Machado-Joseph disease (MJD) is a spinocerebellar ataxia (SCAs) autosomal dominant inherited caused by CAG repeats in the *ATXN3* gene, encoding a polyglutamine tract, and leading to a neurotoxic gain of function of the corresponding protein, ataxin-3, by mechanisms that are not fully known. The size of the CAG repeats, however, only partially explain the phenotypic variability observed in this disease, and the remaining explanation may be provided by variation in modifier genes. Proteins with expanded polyglutamine tracts tend to form aggregates and thus have the potential to modulate the phenotype of expansion diseases. In fact, studies both in animal models and in patients showed the existence of an interaction between the native proteins encoded by different SCAs *loci* and mutant ataxin-3. With this work we intended to begin the investigation of a case control study of Machado-Joseph disease, determining the CAG repeats at the *SCA2*, *SCA6* and *DRPLA* *loci*. **Subjects and methods:** we analyzed 78 healthy, unrelated individuals born in the Azores selected by matching it with a series of patients with MJD. The *ATXN3*, *ATXN2*, *CACNA1A* and *ATN1* *loci* were amplified by PCR and the fragment size was determined by fragments analysis in an automated sequencer. For quality control, a few genotypes were confirmed by sequencing. Intra-population (allele and genotype frequencies, conformity to Hardy-Weinberg expectations, genetic diversity) and inter-population (population differentiation and genetic distance) analyzes were performed. **Results:** for all *loci*, a difference in allele size was obtained with the two techniques used; the respective corrections in allele sizes were made. All alleles observed in this sample were in the normal range, and only one intermediate allele was observed at the *SCA2* *locus*. All *loci* were in conformity with Hardy-Weinberg expectations. Alleles with 14 and 23 (29.5%), 22 (89.1%), 11 (35.9%) and 15 (23.7%) repeats were the most frequent in *SCA3*, *SCA2*, *SCA6* and *DRPLA* *loci*, respectively. In the *ATXN2* *locus*, alleles 22 had 1-2 CAA interruptions, and alleles 29 and 33 had only 1 CAA. A relatively high frequency of large normal alleles (9.6% and 10.5%, respectively) was observed at the *ATXN2* and *ATN1* *loci*. **Discussion:** the confirmation of the CAG repeats size by sequencing reinforces the idea that the use of an allelic ladder in fragment analysis is essential for accurate genotyping. The observed intermediate allele, at the *ATXN2* *locus*, may increase the susceptibility for developing *SCA2*, but the interruption may confer in stability to this allele. In this *locus* the alleles with 24-33 CAGs may be associated with the risk for developing other neurodegenerative diseases. In the *CACNA1A* *locus*, the low frequency of large normal alleles seems to be consistent with the absence of reported cases of *SCA6* in the Azores. In the *ATN1* *locus*, the lack of correlation between *DRPLA* frequency and large-sized normal alleles may be related to the haplotype associated with these alleles. **Conclusion:** the genotyping of the *ATXN2*, *CACNA1A* and *ATN1* *loci* in healthy Azorean individuals will allow the

use of this sample as a control group in a study of potential modifiers gene of MJD, using Azorean patients.

CAPÍTULO 1: INTRODUÇÃO

1.1. PERTINÊNCIA E OBJECTIVOS

A análise do motivo CAG nos genes *ATXN2*, *CACNA1A* e *ATN1*, tendo em vista a determinação do potencial papel destes genes como modificadores do fenótipo da doença de Machado-Joseph (*Machado-Joseph disease*, MJD), está a ser investigada pelo grupo de Epidemiologia e Genética Humana (EpiGene) do Centro de Investigação de Recursos Naturais (CIRN), da Universidade dos Açores. A abordagem metodológica exige o estudo de doentes MJD e dos controlos correspondentes. Neste trabalho, pretende-se iniciar a primeira fase da investigação, analisando a variação nos *loci* da SCA2, SCA6 e DRPLA em indivíduos açorianos saudáveis (não MJD) não aparentados. Os objetivos específicos deste trabalho são os seguintes:

- Confirmar a inexistência da mutação causal da MJD na amostra estudada;
- Optimizar os protocolos de amplificação das repetições CAG nos *loci* *ATXN2*, *CACNA1A* e *ATN1*;
- Determinar a variabilidade genética nos *loci* referidos;
- Comparar os resultados obtidos com os reportados para outras populações.

1.2. REPETIÇÕES TRINUCLEOTÍDICAS

1.2.1. REPETIÇÕES TRINUCLEOTÍDICAS: ASPECTOS GERAIS

O genoma humano é composto por pelo menos 50% de DNA repetitivo, sendo que 3% corresponde a repetições de sequências simples, entre as quais se incluem os microsatélites (Lander *et al.*, 2001). Os microsatélites, ou repetições curtas em tandem (*short tandem repeats*), são constituídos por pequenos motivos de um a seis nucleótidos, repetidos sucessivamente sem interrupções, que ocorrem com elevada frequência no genoma (Hancock, 1999). Estas sequências têm uma taxa de mutação relativamente elevada, estimando-se como sendo da ordem dos que seja de 10^{-6} a 10^{-2} , por *locus* e por geração (Weber & Wong, 1993; Li *et al.*, 2002), o que implica a perda ou ganho de unidades repetitivas, resultando num elevado grau de polimorfismo (Ellegren, 2004). O mecanismo de mutação dos microsatélites é denominado de “mutação dinâmica” (Richard & Sutherland, 1992).

Devido à sua enorme abundância e ao seu elevado grau de polimorfismo, os microsatélites têm sido amplamente utilizados não só como marcadores genéticos no mapeamento do genoma, como também na tentativa de resolver variadas questões

biológicas (Schlotterer, 2000), sendo utilizados em estudos de diversidade genética intra e inter-espécies e populações (Ellegren, 1991; Schlotterer *et al.*, 1991), bem como em investigações forenses (Jeffreys *et al.*, 1991). Mais tarde, a descoberta de que uma classe específica de microsátélites, nomeadamente as repetições trinucleotídicas (RTs), estavam envolvidas em diversas doenças genéticas, veio suscitar um interesse adicional no estudo do processo evolutivo dos microsátélites (ver, entre outros, Pearson & Sinden, 1998; Schlotterer, 2000).

1.2.2. REPETIÇÕES TRINUCLEOTÍDICAS ASSOCIADAS A DOENÇAS

As RTs, assim como outros microsátélites, são muito polimórficas, produzindo um elevado número de alelos e criando um potencial para um elevado grau de heterozigotia. As RTs abaixo de um certo tamanho, variável entre *loci*, são estáveis durante a mitose e meiose, enquanto acima de um certo tamanho tornam-se bastante instáveis (Strachan & Read, 1999). A variabilidade observada nas RTs influencia a sua capacidade de expansão e contracção, ainda que, na maioria dos casos, se observe um viés para a expansão; a amplitude da variação, expansão e contracção, pode ser significativamente diferente entre as sequências dos motivos repetitivos (Jurka & Pethiyagoda, 1995).

As doenças de RTs são causadas por uma expansão instável do motivo repetitivo. Nos alelos normais, as RTs variam normalmente entre apenas algumas e quase 40 unidades, enquanto os alelos patogénicos possuem geralmente mais de 40 repetições (Tsuji, 1997). Existem, no entanto, excepções, como é o caso da SCA6, em que indivíduos homozigóticos para alelos normais apresentam entre 4 a 20 motivos CAG, ao passo que nos doentes, o número de repetições do alelo expandido varia entre 20 e 29 CAGs (revisto em Matos *et al.*, 2011). O número de unidades repetitivas afecta a instabilidade das RTs, todavia, na população normal, as RTs raramente apresentam instabilidade durante a transmissão. Segundo Ohta & Kimura (1973), a evolução dos microsátélites segue um “*stepwise mutation model*”, o qual propõe que a mutação faz aumentar ou diminuir uma unidade de repetição no tamanho do alelo, com igual probabilidade. Além disso, tem sido sugerido que os alelos normais de maior tamanho poderiam construir um “reservatório” a partir do qual os alelos expandidos poderiam ser gerados (revisto em Bettencourt *et al.*, 2007). Segundo vários autores, entre os

quais Cleary & Pearson (2003), a presença de interrupções na sequência repetitiva está também associada com uma maior estabilidade das RTs.

A instabilidade das repetições pode resultar de múltiplos processos que ocorrem individualmente ou em combinação, dependendo do tecido, estado de proliferação e fase de desenvolvimento da célula. No entanto, o processo mutacional que permite que um alelo normal expanda para um tamanho patológico ainda não está totalmente compreendido (revisto em Bettencourt & Lima, 2011). Sabe-se que a instabilidade das RT deverá ser influenciada pela formação de estruturas secundárias e conformações terciárias alternativas do DNA durante vários processos metabólicos do DNA, incluindo a replicação do DNA, reparação, recombinação e transcrição (revisto em Castel *et al.*, 2010).

De entre os vários microsátélites, as RTs associadas a doenças têm um particular interesse, dado que a expansão instável das repetições pode alterar significativamente o controlo da expressão génica ou estrutura e função da proteína codificada, podendo resultar em diferentes doenças neurológicas (Subramanian *et al.*, 2003). Mais de 20 doenças têm sido associadas a este tipo de sequências repetitivas do DNA, incluindo doenças neurodegenerativas e neuromusculares (revisto em Gatchel & Zoghbi, 2005; Pearson *et al.*, 2005; revisto em Orr & Zoghbi, 2007). Embora as doenças de expansão de RTs partilhem um mecanismo de mutação análogo, a localização no gene das RTs no gene causal da doença pode variar (Gomes-Pereira & Monckton, 2006). De acordo com a localização das RTs estas podem ser classificadas em quatro classes: repetições na região 5' não traduzida (5'-UTRs), exões, intrões e região 3' não traduzida (3'-UTRs) (Fig. 1). Na região codificante, as repetições CAG são as mais comuns (revisto em Bettencourt *et al.*, 2007).

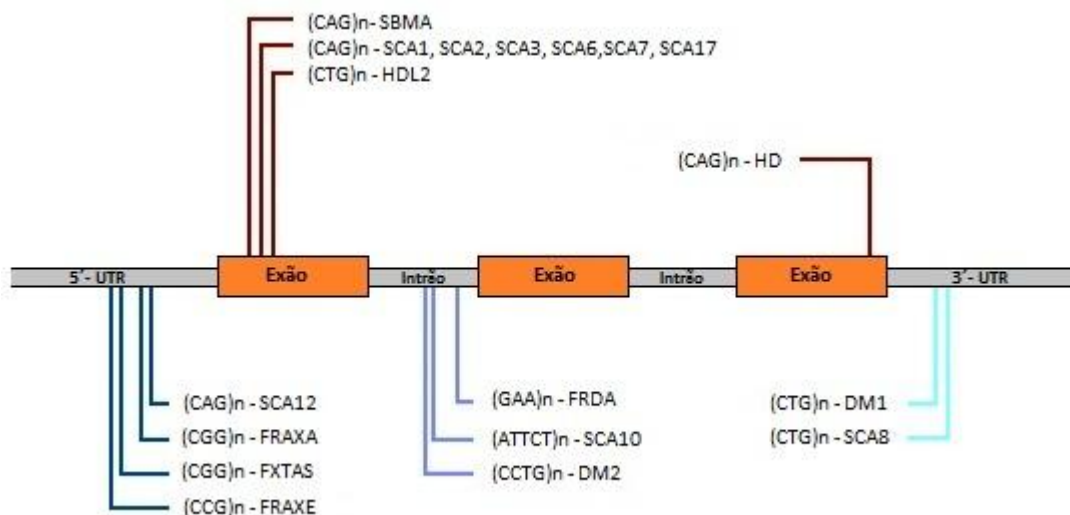


Fig. 1 – Classificação genética das doenças de RTs consoante a localização do motivo repetitivo no gene causal da doença (adaptada de Brouwer *et al.*, 2009).

1.3. ATAXIAS ESPINOCEREBELOSAS: ASPECTOS GERAIS

As ataxias espinocerebelosas (*Spinocerebellar Ataxias, SCAs*) constituem um grupo complexo e heterogéneo de ataxias hereditárias autossómicas dominantes. As SCAs são doenças neurodegenerativas de início tardio e são caracterizadas por uma progressiva disfunção cerebelosa, variavelmente associada com outros sintomas do sistema nervoso central (SNC) e periférico (SNP) (revisto em Bettencourt & Lima, 2011). Na maioria das SCAs as primeiras manifestações ocorrem normalmente na terceira e quarta década, e podem incluir entre outras manifestações cerebelosas, dificuldade na coordenação motora, disartria, demência, sinais piramidais e disfunção óculo-motora (Harding, 1984). A caracterização clínica das SCAs revelou uma grande variabilidade de apresentações sintomáticas, tendo-se observado que os sinais e sintomas apresentados pelos doentes podem variar não só de um subtipo de SCA para outro, como também, dentro do mesmo subtipo, dependendo do número de RTs, tornado assim difícil o diagnóstico realizado exclusivamente por observação clínica (Stevanin *et al.*, 2000). O diagnóstico diferencial das SCAs depende assim da deteção, através de testes moleculares, da mutação no gene causal da doença (ver, entre outros, Stevanin *et al.*, 2000; Stephens *et al.*, 2001).

Avanços nos estudos moleculares de famílias permitiram uma classificação genética das SCAs, substituindo assim as classificações anteriores, que até então eram

baseadas nos sintomas clínicos e idade de início dos mesmos (Sequeiros *et al.*, 2012). Até ao presente cerca de 30 subtipos de SCAs foram descritos, e consoante a natureza subjacente às mutações causais, estes subtipos podem ser divididos em três classes distintas: 1) ataxias de “poliglutamina”, causadas por expansão de repetições CAG que codificam um tracto de poliglutamina (poliQ) expandido na proteína correspondente; 2) ataxias causadas por expansão de repetições fora da região codificante do gene causal da doença e 3) ataxias causadas por mutações convencionais no gene causal da doença (Soong & Paulson, 2007).

Doenças de PoliQ

Entre as doenças causadas por expansões dinâmicas, as doenças de poliQ são a classe mais comum, apesar destas doenças serem globalmente raras (revisto em Zoghbi & Orr, 2000). Até ao presente foram identificadas nove doenças de poliQ: SCA1, 2, 3, 6, 7 e 17, atrofia muscular espino-bulbar (*Spinal and bulbar muscular atrophy*, SBMA), atrofia dentato-rubro-pálido-luisiana (*Dentatorubral-pallidoluysian atrophy*, DRPLA) e doença de Huntington (*Huntington disease*, HD) (revisto em Matos *et al.*, 2011) (Tabela 1). Todas estas doenças são caracterizadas por uma perda neuronal seletiva no SNC, que é acompanhada por um conjunto de alterações físicas associadas, embora cada uma das doenças apresente sintomatologias características e diferentes regiões do cérebro sejam afectadas (revisto em Costa & Paulson, 2011). As doenças poliQ possuem várias características em comum, no entanto, as proteínas associadas a estas doenças não partilham qualquer homologia fora do tracto poliQ, sendo estruturalmente e funcionalmente diferentes (revisto em Matos *et al.*, 2011). O desenvolvimento destas doenças é causado não apenas pela perda da função normal da proteína, devido ao enrolamento incorrecto desta, mas também por um ganho de função neurotóxica, que envolve a formação de agregados nas células afectadas (ver, entre outros, revisto em Gatchel & Zoghbi, 2005). A formação de inclusões neuronais intranucleares (*neuronal intranuclear inclusions*, NIIs) constitui uma das principais características patológicas das doenças poliQ. Estudos neuropatológicos sugeriram que as NIIs, inicialmente consideradas estruturas tóxicas que conduziam à morte das células neuronais, podem ser resultado de uma reacção celular para reduzir o efeito tóxico das proteínas mutadas (revisto em Yamada *et al.*, 2008). A maioria destas doenças apresenta também agregados citoplasmáticos, com excepção da SCA17 e SBMA (revisto em Yamada *et al.*,

2008). As proteínas mutadas expressam-se ubiquamente no SNC, contudo, populações específicas de neurónios são afectadas em cada uma destas doenças, originando um padrão neurodegenerativo e sintomatológico característico (revisto em Gatchel & Zoghbi, 2005).

De modo geral, existe uma correlação positiva entre o número de repetições e a gravidade da doença e uma correlação negativa com a idade de início dos sintomas (revisto em Friedman *et al.*, 2006). Nas doenças poliQ, a variação no número de repetições CAG explica entre 25 a 80% da variabilidade da idade de início (Maciel *et al.*, 1995; Maruyama *et al.*, 1995; Paulson & Fischbeck, 1996; Schöls *et al.*, 1997; Matsuyama *et al.*, 1997; Hayes *et al.*, 2000; Velázquez-Pérez *et al.*, 2009; Bettencourt *et al.*, 2010), dependendo da doença.

As repetições CAG nos genes causais das doenças poliQ podem apresentar instabilidade mitótica e meiótica, sendo que o grau de instabilidade varia entre as diferentes doenças. Muitas destas doenças poliQ apresentam mosaïcismo somático, resultando em diferentes populações de células (dentro do mesmo tecido ou entre tecidos) que têm diferentes tamanhos de repetições (revisto em Bettencourt *et al.*, 2007). As mutações na linha germinativa estão também envolvidas na instabilidade das RTs, e na maioria das doenças poliQ, os alelos patogénicos tendem a ser mais instáveis, o que normalmente leva a uma maior propensão para a expansão nas sucessivas gerações, durante a transmissão paterna (Everett & Wood, 2004). A maioria destas doenças apresenta um fenótipo mais severo e/ou um início dos sintomas mais precoce em sucessivas gerações. Este fenómeno é designado de antecipação e é explicado pelo processo mutacional das RTs. Os alelos normais são geralmente transmitidos para a descendência sem alterações, enquanto a maioria dos alelos expandidos (com algumas excepções, como é o caso da SCA6) são instáveis durante a transmissão, levando a um aumento do número de repetições (revisto em Bettencourt *et al.*, 2007).

Tabela 1 – Principais características moleculares das doenças de poliQ.

Doença	Gene	Locus	Interrupções	Repetições CAG		Proteína	Referência
				Normal	Patológico		
DRPLA	<i>ATN1</i>	12p13.31	-	6-34	49-88	Atrofina-1	Matos <i>et al.</i> (2011) Castel <i>et al.</i> (2010)
SCA1	<i>ATXN1</i>	6p23	CAT	6-39	40-82	Ataxina-1	Matos <i>et al.</i> (2011) Castel <i>et al.</i> (2010)
SCA2	<i>ATXN2</i>	12q24	CAA	13-31	37-75	Ataxina-2	Laffita-Mesa <i>et al.</i> (2012) Castel <i>et al.</i> (2010)
SCA3	<i>ATXN3</i>	14q24.3-q31	CGG, GCC, CAG, CAA	12-44	61-87	Ataxina-3	Bettencourt & Lima (2011) Castel <i>et al.</i> (2010)
SCA6	<i>CACNA1A</i>	19p13	-	4-20	20-29	Subunidade alfa 1A do canal de cálcio	Matos <i>et al.</i> (2011) Castel <i>et al.</i> (2010)
SCA7	<i>ATXN7</i>	3p21.1-p12	-	4-35	37-306	Ataxina-7	Matos <i>et al.</i> (2011) Castel <i>et al.</i> (2010)
SCA17	<i>TBP</i>	6q27	Informação não disponível	25-42	47-63	TBP	Matos <i>et al.</i> (2011) Castel <i>et al.</i> (2010)
HD	<i>HTT</i>	4p16.3	-	6-34	36-121	Huntingtina	Matos <i>et al.</i> (2011) Castel <i>et al.</i> (2010)
SBMA	<i>AR</i>	Xq11-q12	-	9-36	38-62	Receptor andrógeno	Matos <i>et al.</i> , (2011) Castel <i>et al.</i> (2010)

Epidemiologia

Os dados epidemiológicos sobre a prevalência das SCAs estão restritos a alguns estudos realizados em regiões geograficamente isoladas, e a maioria não reflete a real ocorrência destas doenças (revisto em Schöls *et al.*, 2004). A nível mundial, as SCAs são consideradas doenças raras, com uma prevalência estimada entre 0,3 a 2,0 por 100 000 pessoas (Leone *et al.*, 1995; Silva *et al.*, 1997; van de Warrenburg *et al.*, 2002). A prevalência dos diferentes subtipos varia entre diferentes grupos étnicos e populações continentais (revisto em Schöls *et al.*, 2004), porém, a prevalência de algumas SCAs em populações geograficamente isoladas pode ser muito elevada devido à ocorrência de efeito fundador (revisto em Sequeiros *et al.*, 2012). É o que sucede para a SCA2 na província de Holguín, Cuba, em que a prevalência é de 43:100 000 habitantes (revisto em Velázquez-Pérez *et al.*, 2011). Em Portugal, a prevalência das SCAs dominantes é de 5,5 por 100 000 pessoas (revisto em Sequeiros *et al.*, 2012). Entre as SCAs dominantes, a SCA3 é o subtipo mais frequente a nível mundial (21%), seguida da SCA2 (15%) e SCA6 (15%), enquanto a frequência da DRPLA é muito baixa (< 1%). A SCA2 é o segundo subtipo mais prevalente (revisto em Schöls *et al.*, 2004). Todavia existe um considerável número de famílias sem diagnóstico molecular; esta proporção difere entre diferentes populações, variando entre 6% no Brasil (Rio Grande do Sul) (Jardim *et al.*, 2001) e 86,5% no Reino Unido (Leggo *et al.*, 1997).

A SCA2 é mais frequente no nordeste de Cuba, e também na região oriental da Índia (57%), Reino Unido (47%), Sul de Itália (47%), México (43%), Cantabria, Espanha (30%), e Coreia (13%). Em Singapura é a segunda SCA mais frequente (33%) (revisto em Sequeiros *et al.*, 2012). Em Portugal, a frequência da SCA2 varia entre 2% a 4% (Silveira *et al.*, 1998; Silveira *et al.*, 2002; Vale *et al.*, 2010).

A SCA6 é muito comum no Japão, onde representa entre 10 a 30% de todas as ataxias dominantes. É igualmente a forma mais comum na Austrália (17%), Nordeste de Inglaterra, onde a prevalência estimada foi de 1,6:100 000, Alemanha (22%) e Holanda (15%). No resto da Europa, a prevalência parece ser muito baixa (revisto em Sequeiros *et al.*, 2012). Em Portugal, não foi reportado nenhum caso de SCA6 (Silveira *et al.*, 1998; Silveira *et al.*, 2002; Vale *et al.*, 2010).

A DRPLA é relativamente frequente no Japão (7-20%), e representa 6% e 3% de todos os casos de SCAs em Singapura e na Coreia, respectivamente (revisto em Sequeiros *et al.*, 2012). No entanto, embora seja muito rara noutras populações, várias

famílias de descendência europeia têm sido identificadas, incluído cinco famílias em Portugal Continental, correspondendo a um total de 25 doentes (Vale *et al.*, 2010), e sete famílias no Reino Unido (revisto em Sequeiros *et al.*, 2012).

A SCA3 é a mais frequente das SCAs diagnosticadas no Brasil (69-92%), Portugal (58-74%), Singapura (53%), China (48-49%), Holanda (44%), Alemanha (42%) e Japão (28-63%). E é menos frequente no Canadá (24%), Estados Unidos (21%), México (12%), Austrália (12%) e Índia (5-14%), sendo ainda considerada relativamente rara na África do Sul (4%) e Itália (1%) (revisto em Bettencourt & Lima, 2011). Em Portugal, apesar de ser relativamente rara no território continental (1:100,000), com exceção de uma pequena área no vale do rio Tagus (1:1000), é muito frequente nos Açores (1:3472), onde a maior prevalência a nível mundial ocorre na Ilha das Flores (1:239) (revisto em Bettencourt & Lima, 2011). Um estudo recente revelou um aumento da prevalência da SCA3 nesta região, como sendo 1:2597 nos Açores (1:146 na Ilha das Flores) (Araújo, submetida).

1.3.1. ATAXIA ESPINOCEREBELOSA TIPO 2

1.3.1.1. CARACTERIZAÇÃO CLÍNICA DA DOENÇA

A SCA2 (OMIM#183090) caracteriza-se pela presença de um síndrome cerebeloso progressivo, com ataxia da marcha, disartria, dismetria e adiadococinésia. Para além disso, mais de 90% dos doentes apresentam desaceleração dos movimentos oculares sacádicos e 88% movimentos oculares voluntários com amplitude reduzida. Estes doentes exibem ainda dificuldade de deglutição e na realização de algumas funções autónomas, sendo todavia raro apresentarem demência (revisto em Velázquez-Pérez *et al.*, 2011). A SCA2 normalmente ocorre na idade adulta, com uma média de idade de início de 33 anos. Em 80% dos casos foi observado o fenómeno de antecipação (Sanpei *et al.*, 1996). Doentes que possuem expansões CAG grandes mostram sintomas atípicos, enquanto os que possuem repetições CAG menores podem vir a desenvolver sintomas parkinsonianos (Tan *et al.*, 2004; Lu *et al.*, 2004). Segundo Kim *et al.* (2007), tem sido reportada a existência de uma sobreposição entre o fenótipo da SCA2 e o fenótipo de outras ataxias e doenças parkinsonianas (parkinsonismo familiar e esporádico). Estes doentes apresentam, para além de ataxia, sintomas da doença de Parkinson (*Parkinson disease*, PD), paralisia supranuclear progressiva (*progressive supranuclear palsy*, PSP) e

atrofia de múltiplos sistemas (*multiple system atrophy, MSA*), entre outras (Furtado *et al.*, 2002; Kim *et al.*, 2007).

1.3.3.2. GENÉTICA MOLECULAR E PATOGÉNESE

A mutação causal da SCA2 consiste na expansão de repetições CAG no exão 1 no gene *ATXN2*, localizado no cromossoma 12, em 12q24.1 (Imbert *et al.*, 1996; Pulst *et al.*, 1996; Sanpei *et al.*, 1996). Nos indivíduos normais as repetições são polimórficas quer no tamanho, variando entre 13 e 31 CAGs, como também na estrutura interna, apresentando uma ou mais interrupções CAA (Laffita-Mesa *et al.*, 2012). No entanto, as repetições CAG no *locus* da SCA2 não são tão polimórficas como nas outras SCAs, dado que 95% dos alelos têm 22 ou 23 repetições, com duas interrupções CAA. Na maioria das populações, o alelo 22 com a configuração $(CAG)_8CAA(CAG)_4CAA(CAG)_8$ é o mais comum (ver, entre outros, Imbert *et al.*, 1996; Sanpei *et al.*, 1996; Choudhry *et al.*, 2001; Sobczak & Krzyzosiak, 2004; Laffita-Mesa *et al.*, 2012). Os alelos patológicos possuem mais de 34 repetições CAG, embora os alelos com 32 a 34 CAGs, designados de alelos intermediários, estejam associados a uma penetrância incompleta, enquanto os alelos com mais de 37 repetições apresentam penetrância completa (Cancel *et al.*, 1997; Geschwind *et al.*, 1997; Laffita-Mesa *et al.*, 2012). Foram reportados apenas alguns casos de doentes com desenvolvimento tardio da doença, em que o alelo expandido tinha entre 32 e 34 repetições CAGs (Futamura *et al.*, 1998; Santos *et al.*, 1999; Costanzi-Porrini *et al.*, 2000; Fernandez *et al.*, 2000).

A maioria dos alelos normais tem uma a quatro interrupções CAA, enquanto a grande maioria dos alelos patogénicos possui um tracto de repetições puro (Costanzi-Porrini *et al.*, 2000; Laffita-Mesa *et al.*, 2012). As interrupções CAA conferem uma estabilidade genómica ao tracto de repetições CAG, ao passo que a ausência destas interrupções confere uma predisposição para a expansão (Chung *et al.*, 1993; Pearson *et al.*, 1998). Assim, a predisposição para a instabilidade e expansão das repetições na SCA2 parece estar correlacionada não apenas com o tamanho mas também com a pureza (ausência ou presença de interrupções CAA) do tracto de repetições CAG (Choudhry *et al.*, 2001).

O gene *ATXN2* codifica uma proteína designada de ataxina-2, que tem um tracto poliQ localizado perto da região N-terminal (Sahba *et al.*, 1998). Esta proteína é composta por 1312 aminoácidos e tem um peso molecular estimado de 140 KDa, na

sua forma nativa (revisto em Velázquez-Pérez *et al.*, 2011). A ataxina-2 é constituída por dois domínios globulares (Lsm e domínio associado a Lsm) e por duas regiões não globulares flexíveis, a região N-terminal que contém o tracto poliQ e a C-terminal que contém outros sítios de possível interacção com outras proteínas, como é o caso do domínio PAM2 (*PABP interacting Motives*) que interage com a proteína de ligação à cauda de poliadenina (*Poly-(A) Binding Protein*, PABP) (revisto em Orr & Zoghbi, 2007). Até ao momento a função desta proteína permanece incompletamente compreendida. Várias evidências sugerem, que esta esteja envolvida no metabolismo do mRNA e na regulação da tradução (revisto em Velázquez-Pérez *et al.*, 2011). Segundo Satterfield & Pallanck, (2006), a presença dos dois domínio globulares, sugere que está envolvida em vários processos essenciais do RNA, incluído na modificação do RNA (*splicing* do pré-mRNA e degradação do RNA), e por outro lado, a interacção da ataxina-2 com a PABP sugere que poderá ter um papel no transporte, estabilidade e/ou tradução do mRNA.

A ataxina-2 é ubiquamente expressa no citoplasma, essencialmente no complexo de Golgi, da maioria dos tecidos (Huynh *et al.*, 2003); nos pulmões e rins a proteína apresenta uma expressão significativamente menor (Imbert *et al.*, 1996; Pulst *et al.*, 1996; Sanpei *et al.*, 1996). No cérebro é mais abundante nas células de Purkinje, tanto nos indivíduos normais como nos doentes, porém, a expressão da proteína aumenta consideravelmente nestes últimos (Huynh *et al.*, 1999). Segundo Matilla-Dueñas *et al.* (2010), a expansão do tracto poliQ altera a função da ataxina-2 o que leva a um ganho de função tóxica através da formação de agregados citoplasmáticos. As Nlls são raras na SCA2, estando apenas presentes nas células de Purkinje e no tronco cerebral (revisto em Yamada *et al.*, 2008).

1.3.3.3. OUTRAS DOENÇAS ASSOCIADAS AOS ALELOS NORMAIS GRANDES E/OU INTERMEDIÁRIOS NO GENE DA SCA2

O tamanho do alelo normal grande e/ou intermediário no *locus ATXN2* está associado com três doenças neurológicas clinicamente distintas, e cada uma destas doenças está associada a um determinado número de repetições CAG e a uma configuração específica da sequência repetitiva (Yu *et al.*, 2011). Como já referido, o fenótipo da SCA2 é causado por uma sequência pura de mais de 34 repetições CAG no gene *ATXN2* (Imbert *et al.*, 1996; Pulst *et al.*, 1996; Sanpei *et al.*, 1996; Cancel *et al.*, 1997). Por outro lado, os alelos com 28 a 33 repetições estão associados como factor

de risco da esclerose lateral amiotrófica (*amyotrophic lateral sclerosis, ALS*), embora seja provável que o tamanho das repetições varie consoante a população (ver, entre outros, Elden *et al.*, 2010; Chen *et al.*, 2011; Daoud *et al.*, 2011; Ross *et al.*, 2011). Segundo Yu *et al.* (2011), todas as repetições CAG associadas com a ALS possuem entre uma a três interrupções CAA. Além disso, foi também observada a existência de um haplótipo diferente entre os indivíduos que têm três interrupções CAA e os que têm menos interrupções, sendo que a idade de início da doença depende do número de interrupções CAA e do haplótipo associado a estes alelos (Yu *et al.*, 2011). Por sua vez, alelos com 32 a 49 CAGs, com pelo menos uma interrupção CAA, estão associados com a PD (ver, entre outros, Furtado *et al.*, 2002; Silveira *et al.*, 2002; Kim *et al.*, 2007).

1.3.2. ATAXIA ESPINOCEREBELOSA TIPO 6

1.3.2.1. CARACTERIZAÇÃO CLÍNICA DA DOENÇA

Praticamente todos os doentes com SCA6 (OMIM#183086) apresentam, entre outros sintomas, disartria, nistagmo, desequilíbrio progressivo e severa dificuldade de coordenação dos membros (Gomez *et al.*, 1997). Alguns doentes têm ainda episódios ocasionais de vertigens, todavia não apresentam fraqueza muscular ou dano cognitivo (Jen *et al.*, 1998). A SCA6 é uma doença de início tardio, e 60% dos doentes desenvolve a doença após os 50 anos de idade. Normalmente, esta doença é de progressão lenta e não leva à morte precoce, contudo a maioria dos doentes fica dependente de uma cadeira de rodas no fim da sexta década de vida (Schöls *et al.*, 2000; Craig *et al.*, 2004).

1.3.2.2. GENÉTICA MOLECULAR E PATOGÉNESE

Em 1997, Zhuchenko e colaboradores mapearam a subunidade alfa 1A do canal de cálcio neuronal localizada no cromossoma 19, em 19p13, e identificaram o gene causal como uma expansão de uma repetição CAG, localizada no exão 47 do gene *CACNA1A* (*α 1A subunit of the voltage-dependent calcium channel*). As repetições CAG no gene *CACNA1A* são polimórficas, variando entre 4 e 20 repetições nos cromossomas normais e entre 20 e 29 nos expandidos; indivíduos que possuam um alelo 20 podem também desenvolver a doença (Komeichi *et al.*, 2001; revisto em Gatchel & Zoghbi, 2005; Trott *et al.*, 2006) ainda que numa idade mais avançada (van de Warrenburg *et al.*, 2005). Mariotti *et al.* (2001) relataram um efeito de dosagem génica, ao observar sintomas num indivíduo homozigótico para o alelo 19, ao passo que os indivíduos da

mesma família que tinham este alelo, em heterozigotia com um alelo normal, não apresentavam sintomas; tendo sido o mesmo fenómeno anteriormente observado por Ikeuchi *et al.* (1997) em dois doentes japoneses. O fenómeno de antecipação é raro, dado que as repetições CAG nos alelos normais e patológicos, neste *locus*, são menores e mais estáveis ao longo de sucessivas gerações quando comparadas com as outras doenças poliQ (Zhuchenko *et al.*, 1997). O gene *CACNA1A* codifica a subunidade $\alpha 1A$ do canal de cálcio tipo P/Q (*CACNA1A*), uma proteína com um peso molecular aproximado de 282,4 KDa (Uversky, 2009). O mRNA da *CACNA1A* é expresso ubiquamente no SNC, particularmente no cerebelo (Fletcher *et al.*, 1996). Esta proteína é estruturalmente formada por 2505 aminoácidos e por um número variável de glutaminas, estando o tracto poliQ localizado na região distal C-terminal (Uversky, 2009). Na SCA6, a acumulação de agregados no núcleo das células parece não ser necessária para desenvolver a doença, dado que a formação de Nlls é rara, ocorrendo apenas nas células de Purkinje, contrariamente ao observado nas outras doenças poliQ, com excepção da SCA2 (revisto em Yamada *et al.*, 2008). Visto que a região C-terminal do gene *CACNA1A* tem um papel importante na função do canal de cálcio (Walker & DeWaard, 1998), diversos autores têm sugerido que a expansão do tracto poliQ na SCA6 leva a uma alteração ou perda de função do canal de cálcio das células de Purkinje, resultando num nível de cálcio intracelular excessivo, levando à subsequente morte celular (Koh & Cotman, 1992; Rettig *et al.*, 1996; Restituito *et al.*, 2000).

1.3.3. ATROFIA DENTATO-RUBRO-PÁLIDO-LUISIANA

1.3.3.1. CARACTERIZAÇÃO CLÍNICA DA DOENÇA

A DRPLA (OMIM#125370) é caracterizada clinicamente pela presença de ataxia cerebelosa e demência, às quais se associam, em graus variáveis: movimentos involuntários (coreia e mioclonia) e epilepsia (Komure *et al.*, 1995). A doença pode desenvolver-se entre a primeira e a sétima década, com uma média da idade de início de 30 anos. As características clínicas estão associadas à idade de início da doença, sendo que os doentes com um início precoce (menos de 20 anos) apresentam um fenótipo progressivo de epilepsia mioclónica, caracterizado clinicamente por mioclonia, epilepsia, demência e ataxia, enquanto os doentes com um início mais tardio (após os 40 anos) apresentam geralmente ataxia cerebelosa, coreoatetose e demência. Os

doentes com uma idade de início entre os 20 e 40 anos apresentam formas transitórias entre os dois tipos (Naito & Oyanagi, 1982; Takahashi *et al.*, 1988).

1.3.3.2. GENÉTICA MOLECULAR E PATOGÉNESE

A DRPLA é causada por uma expansão (CAG)_n na região codificante do gene *ATN1* no cromossoma 12, em 12p13.2. Os alelos normais têm normalmente entre 6 e 34 repetições, enquanto os alelos patogénicos têm entre 49 a 88 CAGs (Koide *et al.*, 1994; Nagafuchi *et al.*, 1994). Kurohara e colaboradores (1997) observaram uma forma de paraplegia espástica hereditária autossómica recessiva, fenótipo distinto do reportado normalmente em doentes com DRPLA, em indivíduos homocigóticos para os alelos intermediários (40 ou 41 repetições). No *locus* da DRPLA, o fenómeno da antecipação genética é particularmente evidentes em famílias japonesas, em que há um aumento do tamanho das repetições de 8 motivos CAG com um conseqüente decréscimo de 34,4 anos na idade de início dos sintomas, em média, quando transmitidos pelo pai. Porém, segundo Yanagisawa *et al.* (1996) o mecanismo subjacente à instabilidade poderá dever-se não só ao número de repetições nos alelos normais grandes como também ao haplótipo associado a estes alelos, uma vez que, todos os alelos intermediários e patológicos, na população japonesa, estão associados com o haplótipo específico A1-B1 (polimorfismo bialélico intragénico). Martins e colaboradores (2003) observaram que nas populações de descendência caucasiana, todos os alelos patológicos tinham o haplótipo A1-B1, estando este também presente nos alelos normais grandes.

O gene *ATN1* codifica uma proteína, denominada atrofina-1, com um peso molecular aproximado de 190 KDa, na sua forma nativa. Estruturalmente, é formada por 1185 aminoácidos e por um número variável de glutaminas (tracto poliQ) próximo da região N-terminal (Koide *et al.*, 1994; Kanazawa, 1999). Embora a função desta proteína ainda não seja conhecida, um estudo realizado em *Drosophila*, sugere que possa estar envolvida na regulação da transcrição (Zhang *et al.*, 2002). A atrofina-1 expressa-se ubiquamente e está localizada principalmente no citoplasma dos neurónios e células periféricas de indivíduos normais e afectados (Yazawa *et al.*, 1995; Knight *et al.*, 1997). Igarashi *et al.* (1998) observaram inclusões nucleares somente em doentes com DRPLA.

1.3.4. ATAXIA ESPINOCEREBELOSA TIPO 3

1.3.4.1. CARACTERIZAÇÃO CLÍNICA DA DOENÇA

A SCA3, também conhecida como MJD (OMIM#109150), é caracterizada por um fenótipo complexo e pleiotrópico, que envolve os sistemas cerebeloso, piramidal, extrapiramidal, periférico e óculo-motor. A MJD apresenta uma idade média de início de 40,2 anos, tendo o caso mais prematuro reportado sido diagnosticado numa criança de 4 anos (Carvalho *et al.*, 2008) e o mais tardio num indivíduo de 70 anos (Coutinho, 1992). A sobrevida média reportada é de cerca de 21 anos (revisto em Bettencourt & Lima, 2011). A acentuada heterogeneidade clínica e progressão natural da doença levou à descrição de três subtipos clínicos diferentes bem estabelecidos (Tabela 2). Os três tipos clínicos podem mesmo estar ocasionalmente presentes numa única família (revisto em Bettencourt & Lima, 2011). Foram ainda propostos dois tipos clínicos adicionais, todavia, estes são considerados raros. O tipo IV é caracterizado por parkinsonismo associado aos sinais cerebelosos (Tuite *et al.*, 1995). Este fenótipo foi observado por Bettencourt *et al.* (2011) em dois doentes açorianos. O tipo V é caracterizado por paraplegia espástica sem ataxia cerebelosa (Sakai & Kawakami, 1996; Kaneko *et al.*, 1997).

Tabela 2 – Classificação clínica dos três subtipos da doença de Machado-Joseph.

Subtipos	Características clínicas	Referência
I	Início precoce (média de 24,3 anos) associada a uma rápida evolução dos sintomas. Ataxia cerebelosa, oftalmoparésia externa progressiva (OEP), sinais piramidais e extrapiramidais.	
II	Início intermédio (média 40,5 anos). Ataxia cerebelosa e OEP, com ou sem sinais piramidais; quando presentes os sinais extrapiramidais são ligeiros.	Coutinho, 1992
III	Início tardio (média 46,8 anos). Ataxia cerebelosa e OEP, sinais piramidais e extrapiramidais ligeiros ou ausentes, sinais periféricos muito marcados.	

1.3.4.2. GENÉTICA MOLECULAR E PATOGÉNESE

O *locus* da MJD/SCA3 foi mapeado no braço longo do cromossoma 14, em 14q32.1 (Takiyama *et al.*, 1993), um ano mais tarde Kawaguchi e colaboradores (1994) identificaram a expansão de um motivo CAG, localizada na região codificante do gene *ATXN3*, como causa da MJD. As repetições CAG no gene *ATXN3* são interrompidas por uma interrupção AAG e duas interrupções CAA, apresentando a seguinte configuração $(CAG)_2CAAAG(CAG)CAA(CAG)_n$ (Kawaguchi *et al.*, 1994). O gene *ATXN3* abrange uma região genómica de cerca de 48 Kb, estando as repetições CAG localizadas na extremidade 5' do exão 10. Este gene foi inicialmente descrito como apresentando 11 exões (Ichikawa *et al.*, 2001), mas, num estudo recente, Bettencourt *et al.* (2010) descreveram dois exões adicionais (6a e 9a).

No *locus* da MJD, indivíduos não portadores possuem entre 12 e 44 repetições CAG, enquanto nos doentes, normalmente heterozigóticos, o número de repetições no alelo expandido varia entre 61 e 87 (Maciel *et al.*, 2001). Os casos de homozigotia na MJD são extremamente raros, mas os poucos descritos parecem apresentar uma forma mais severa da doença sugerindo a existência de um efeito de dosagem génica (revisto em Costa & Paulson, 2011). A MJD é considerada uma doença de penetrância completa apesar de terem sido identificados alguns casos (2%) de não penetrância (Sequeiros, 1989). Os alelos intermediários são raros, tendo sido somente reportados sete casos (45-56) associados com MJD (revisto em Sequeiros *et al.*, 2012). A maioria dos casos reportados teve um início tardio dos sintomas (> 50 anos) e apresentava diferentes características clínicas, para além dos principais sinais da MJD (revisto em Sequeiros *et al.*, 2012). Por outro lado, foi reportado um alelo com 51 repetições CAG que parece não estar associado à doença, sugerindo que os alelos intermediários na MJD podem apresentar baixa penetrância tal como sucede em outras doenças poliQ (revisto em Bettencourt & Lima, 2011). Num estudo realizado por Lima *et al.* (2005), numa amostra representativa da população portuguesa, não foram encontrados alelos intermediários; estes resultados não suportam a hipótese de que os alelos normais de maior tamanho possam constituir um “reservatório” para a produção contínua de alelos expandidos (revisto em Bettencourt & Lima, 2011).

A questão da origem mundial da mutação na base da MJD tem sido, alvo de vários trabalhos (ver, entre outros, Gaspar *et al.*, 2001; Martins *et al.*, 2007). Nos Açores, Sequeiros & Coutinho (1993) sugeriram que a elevada frequência de MJD

resultaria da ocorrência de efeito fundador. Dois estudos realizados por Gaspar *et al.* (2001) e Martins *et al.* (2007) reportaram a existência de dois eventos mutacionais que seriam responsáveis pela presença da MJD nas famílias açorianas, um facto evidenciado anteriormente em estudos baseados na reconstrução genealógica de famílias afetadas (Lima, 1996; Lima *et al.*, 1998). Estes dois haplótipos também foram encontrados na população de Portugal Continental. Com base nas análises de haplótipos, tem sido sugerido que dois eventos mutacionais podem explicar a presente distribuição geográfica global da MJD (Gaspar *et al.*, 2001; Martins *et al.*, 2007).

O gene *ATXN3* codifica uma proteína, denominada ataxina-3, com um peso molecular aproximado de 40-43 KDa, na sua forma nativa. Estruturalmente, é composta por 339 aminoácidos e por um número variável de poliQs (Kawaguchi *et al.*, 1994). Esta proteína é ubiquamente expressa em tecidos neuronais e não neuronais, encontrando-se presente no citoplasma e no núcleo de diferentes tipos de células. Nos neurónios, a proteína encontra-se principalmente no citoplasma (Paulson *et al.*, 1997a). A ataxina-3 pertencente à família das cisteína-proteases. Esta proteína é composta por um domínio globular Josefina com actividade catalítica, na região N-terminal, seguido de uma região C-terminal não estruturada e flexível, composta por dois ou três motivos de interação com a ubiquitina e pelo tracto poliQ (revisto em Bettencourt & Lima, 2011; revisto em Costa & Paulson, 2011). Foi proposto que a ataxina-3, na sua forma nativa, tem uma função de enzima desubiquitinante, na via proteossoma das ubiquitinas, desempenhando um papel nas vias de controlo celular da qualidade das proteínas (Lima, 2010). Para além disso, tem sido também sugerido o envolvimento da forma normal da proteína na regulação da transcrição (Li *et al.*, 2002; Chou *et al.*, 2008), na regulação da formação de agregados e degradação de proteínas enviadas do retículo endoplasmático (Burnett & Pittman, 2005). A ataxina-3 encontra-se geralmente no citoplasma das células neuronais nos indivíduos saudáveis; nos doentes de MJD encontra-se principalmente no núcleo destas células (Paulson *et al.*, 1997b). Estudos realizados mostraram, todavia, que não existe uma perfeita sobreposição entre as regiões do cérebro afectadas e a presença de Nlls (revisto em Matos *et al.*, 2011). Na sua forma mutada, quando o tracto poliQ atinge um limiar patológico, a proteína adquire um ganho de função neurotóxica, levando à morte neuronal, por um processo que ainda não é totalmente conhecido (revisto em Bettencourt & Lima, 2011). Este ganho de função é causado provavelmente pela alteração conformacional da proteína,

que irá afectar muitas das propriedades da proteína, nomeadamente a estabilidade e degradação, localização subcelular, interações moleculares com outras proteínas e propensão à formação de agregados (revisto em Costa & Paulson, 2011).

Na MJD o tamanho do motivo CAG explica, dependendo da série de doentes em estudo, entre 50% a 75% da variação da idade de início dos sintomas (revisto em Bettencourt & Lima, 2011). Na série de doentes açorianos o número de repetições, determinado no DNA genómico e mRNA, explica 68% e 67% da variação na idade de início, respectivamente (Bettencourt *et al.*, 2010). Além disso, o tamanho dos alelos patogénicos tem sido também associado com a frequência de outras características clínicas, nomeadamente como a retracção palpebral e os sinais piramidais, as quais são mais frequentes em indivíduos que apresentam um maior número de repetições (Takiyama *et al.*, 1995). O facto das repetições apenas explicarem parte da correlação com o fenótipo, observado na MJD e noutras SCAs, confirma o envolvimento de outros factores, que podem ser ambientais ou genéticos (van de Warrenburg *et al.*, 2005).

1.3.4.3. GENES MODIFICADORES DA MJD

Segundo diversos autores uma importante porção da variação na idade de início da MJD poderá ser explicada, para além do número de repetições CAG, por factores familiares. E embora não se possa excluir a influência dos factores ambientais, o facto é que têm sido observada uma menor variabilidade dentro de famílias comparativamente à observada entre famílias, o que suporta a contribuição de outros factores genéticos, nomeadamente genes modificadores, para a restante variação fenotípica (ver, entre outros, DeStefano *et al.*, 1996). Os poucos estudos que se propuseram a identificar modificadores do fenótipo da MJD fizeram-no através da análise de possíveis genes candidatos (revisto em Bettencourt & Lima, 2011). Por exemplo, num estudo recente, Bettencourt *et al.* (2011) mostraram que o gene da APOE, nomeadamente o alelo APOE ϵ 2, modela a variabilidade fenotípica da MJD, estando associado com uma idade de início mais precoce.

A análise do efeito modelador do gene causal das diferentes SCAs no fenótipo da MJD tem adquirido particular interesse, dado que estas doenças partilham o mesmo mecanismo mutacional, e que as proteínas com um tracto poliQ expandido têm tendência para formar agregados com outras proteínas (revisto em Matos *et al.*, 2011).

Vários estudos têm demonstrado a existência de interações entre as proteínas codificadas pelos genes causais das diferentes SCAs (Uchihara *et al.*, 2001; Pulst *et al.*, 2005; Lim *et al.*, 2006). Segundo Okazawa (2003), as proteínas com sequências curtas de poliQ poderão ser levadas a agregar-se a outras proteínas com um tracto poliQ expandido, contribuindo potencialmente para a variação fenotípica da MJD. Num outro estudo, Jardim *et al.* (2003) identificaram em 39 doentes brasileiros, uma associação entre os alelos de maior tamanho (22 e 23 CAG) no *locus* da SCA2 e sinais específicos da MJD, concretamente, fasciculações. Nenhuma associação foi encontrada com outras características fenotípicas, nomeadamente a idade de início, antecipação, e tipo clínico (Jardim *et al.*, 2003); estes resultados poderão, todavia, ser consequência da reduzida amostras de doentes utilizados neste estudo. Segundo Jardim *et al.* (2003), o processo patogénico do tracto poliQ expandido da MJD, poderá ser de alguma forma influenciado pelo tamanho normal do tracto poliQ de outras SCAs. Diversos estudos realizados em modelos animais, nomeadamente em *Drosophila*, demonstraram que a actividade endógena do gene *atx2*, ortólogo do gene *ATXN2*, modula a progressão neurodegenerativa induzida pela proteína Atx3 mutada (Ghosh & Feany, 2004; Derek & Bonini, 2008), suportando a hipótese acima proposta. Os resultados reportados por Pulst *et al.* (2005) parecem também reforçar a hipótese sugerida por Okazawa (2003), ao indicar que os alelos normais de maior tamanho no gene causal da SCA6 estão correlacionados com uma idade de início precoce da SCA2. Os resultados observados nos estudos acima mencionados indicam que os genes causais de outras SCAs serão bons candidatos a genes modificadores da MJD.

CAPÍTULO 2: SUJEITOS E MÉTODOS

2.1. AMOSTRA

A amostra analisada neste trabalho incluiu 78 indivíduos não aparentados da população geral açoriana. Com o objetivo de poder incluir, posteriormente, os resultados deste trabalho num estudo caso-controlo da MJD, os indivíduos estudados pertenciam a famílias sem a doença, tendo sido selecionados com base em critérios que permitiram o emparelhamento com os doentes de MJD, nomeadamente no que se refere ao sexo, naturalidade e idade à data de colheita. Todos os indivíduos analisados foram seleccionados a partir do Repositório Biomédico do grupo EpiGene do CIRN, que incluí as respectivas amostras e também os dados demográficos. O grupo de indivíduos analisados nesta amostra consiste em 31 homens e 39 mulheres, que apresentam um intervalo de idades entre os 22 e 77 anos de idade [$43,83 \pm 13,035$ (média \pm desvio padrão, DP)]; 28 indivíduos são naturais das Flores, 35 de São Miguel, 11 da Graciosa e 4 da Terceira. Todos os indivíduos deram o seu consentimento informado.

2.2. GENOTIPAGEM

A amplificação do fragmento que contém o motivo CAG do gene *ATXN3* foi feita a partir do DNA genómico, previamente extraído a partir de leucócitos de sangue periférico, através do método de *salting out* (Miller *et al.*, 1988). A amplificação foi feita usando os *primers* descritos por Kawaguchi *et al.* (1994): MJD52F (5'-CCAGTGACTACTTTGATTCG-3') e MJD70R (5'-CTTACCTAGATCACTCCCAA-3' marcado com fluorescência 6-FAM). A reacção de amplificação foi feita num volume total de 25 μ l usando as seguintes concentrações de reagentes: 1 μ M *primer forward* e *reverse* (25 μ M), 300 μ M dNTPs, 2,5 mM MgCl₂, 10x tampão [160 mM (NH₄)₂SO₄, 670 mM Tris-HCl (pH 8,8 a 25°C), 0,1% Tween-20], 10% dimetilsulfóxido (DMSO), 1 U de Biotaq™ DNA polimerase (Bioline) e \approx 100 ng de DNA genómico. A amplificação foi feita num termociclador BIORAD C1000 utilizando as seguintes condições: desnaturação durante 5' a 95°C, seguida de 25 ciclos de 1' a 95°C, 1' a 56°C e 1' a 72°C; e por fim uma extensão final de 10' a 72°C. Em seguida confirmou-se a amplificação do fragmento de interesse por electroforese em gel de agarose a 2,5%.

Para determinar o tamanho do fragmento amplificado adicionou-se 1 μ l de produto amplificado (diluído sempre que necessário) a 0,25 μ l de GeneScan 500-LIZ (marcador de tamanho e peso molecular conhecido) e 12,2 μ l de Hi-Di™ Formamida (Applied Biosystems, Foster City, CA, USA), sendo em seguida aquecido durante 2' a

90°C e posteriormente colocado durante 5' em gelo. A determinação do tamanho dos alelos no *locus* da SCA3 foi efectuada no Serviço Especializado de Epidemiologia e Biologia Molecular (SEEBMO) do Hospital de Santo Espírito de Angra do Heroísmo, e foi feita por análise de fragmentos (AF) através de um sequenciador automático ABI-Prism 3100 Genetic Analyzer (Applied Biosystems), usando o Polímero de Performance Optimizada POP-7 (POP-7) (Applied Biosystems).

As condições de amplificação do motivo CAG nos *loci* da SCA2, SCA6 e DRPLA foram modificadas das descritas para o *locus* da SCA3, mas com algumas modificações (Tabela 3). A determinação do tamanho do fragmento amplificado foi feita com as condições previamente descritas, mas a separação dos fragmentos foi efectuada nas instalações do CIRN, na Universidade dos Açores, num sequenciador automático ABI-Prism 310 Genetic Analyzer (Applied Biosystems), usando o Polímero de Performance Optimizada POP-4 (POP-4) (Applied Biosystems).

O número de repetições CAG existentes em cada um dos alelos foi determinado subtraindo ao tamanho do fragmento obtido por análise de fragmentos o número de repetições CAG existentes na sequência que está disponível na base de dados do Centro Nacional de Informação Biotecnológica (NCBI) (<http://www.ncbi.nlm.nih.gov/>), e dividindo por três o resultado. A fórmula usada para calcular o número de repetições CAG para cada *locus* está descrita na tabela 4. Os alelos foram designados segundo a nomenclatura da *Human Genome Variation Society* (HGVS) (<http://www.hgvs.org/>). A determinação dos alelos normais grandes foi feita de acordo com os critérios utilizados por Takano *et al.* (1998): SCA3 (> 27 CAG), SCA2 (> 22 CAG), SCA6 (> 13 CAG) e DRPLA (> 17 CAG).

Tabela 3 – *Primers* e temperatura de *annealing* usadas na amplificação do motivo CAG nos *loci* da SCA2, SCA6 e DRPLA.

Doença/Gene	Primers	Temperatura de <i>annealing</i> (°C)	Referência
SCA2/ATXN2	SCA2-F: 5' FAM-GGGCCCCTCACCATGTGCG-3' SCA2-R: 5' -CGGGCTTGCGGACATTGG-3'	59	Pulst <i>et al.</i> , 1996
SCA6/CACNA1A	S-5-F1: 5' FAM-CACGTGTCCTATTCCCCTGTGATCC-3' S-5-R1: 5' -TGGGTACCTCCGAGGGCCGCTGGTG-3'	63	Zhuchenko <i>et al.</i> , 1997
DRPLA/ATN1	CTGB37.5 F: 5' FAM-CACCAGTCTCAACACATCACCATC-3' CTGB37.5 R: 5' -CCTCCAGTGGGTGGGGAAATGCTC-3'	63	Li <i>et al.</i> , 1993

Tabela 4 – Número de acesso na base de dados (NCBI) da sequência de referência e fórmula utilizada para o cálculo do motivo CAG, em cada *locus*.

Doença/Gene	N.º acesso (NCBI)	Número de repetições CAG
MJD/ATXN3	NG_008198.1	$[(\text{produto AF (pb)} - 199 \text{ pb})/3]$
SCA2/ATXN2	NG_011572	$[(\text{produto AF (pb)} - 61 \text{ pb})/3]$
SCA6/CACNA1A	NM_000068	$[(\text{produto AF (pb)} - 102 \text{ pb})/3]$
DRPLA/ATN1	NG_008047	$[(\text{produto AF (pb)} - 96 \text{ pb})/3]$

2.2.1. SEQUENCIAÇÃO

Amostras de cada um dos *loci* (*ATXN3*, $n = 3$; *ATXN2*, $n = 5$; *CACNA1A*, $n = 5$ e *ATN1*, $n = 4$) foram sequenciadas de modo a confirmar o número de repetições CAG obtidos por AF. Anteriormente à sequenciação, os produtos de PCR foram purificados usando o GeneJet™ PCR Purification Kit (Thermo Scientific). Após a purificação quantificou-se o produto de PCR por espectrofotometria. A reacção de sequenciação foi feita usando o BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) e para cada um dos *loci* foram usados os seguintes *primers* com uma concentração de 5 pmol/μl: MJD52F (*SCA3*), SCA2-R (*SCA2*), S-5-R1 (*SCA6*) e CTGB37.5-R (*DRPLA*). Os produtos de PCR purificados usados na reacção de sequenciação foram diluídos sempre que necessário até uma concentração de ≈ 10 ng/μl. Os produtos da reacção de sequenciação foram depois purificados com o SigmaSpin™ Sequencing Reaction Cleanup kit (Sigma-Aldrich) e quantificados por espectrofotometria. A sequenciação dos produtos de PCR foi feita num volume total de 10 μl, 5 μl de produto de PCR (≈ 40 ng/μl) e 5 μl de Hi-Di™ Formamida (Applied Biosystems, Foster City, CA, USA). A sequenciação foi feita num sequenciador automático ABI-Prism 310 Genetic Analyzer (Applied Biosystems), usando o Polímero de Performance Otimizada POP-4 (POP-4) (Applied Biosystems).

2.3. ANÁLISE ESTATÍSTICA

As frequências alélicas e genóticas foram determinadas para todos os *loci* usando o programa PASW SPSS v. 18.0 (SPSS Inc., 2009), assim como as medidas de distribuição (média, variância, assimetria e curtose). Testou-se a conformidade com o equilíbrio de Hardy-Weinberg (EHW), usando o teste exacto de Guo & Thompson

(1992) e determinou-se a diversidade genética de acordo com Nei (1987) usando o programa Arlequin versão 3.11 (Excoffier *et al.*, 2005). O teste exacto de diferenciação populacional foi usado para analisar a presença de diferenças nas frequências dos alelos, entre pares de populações. Foi calculada a matriz de distâncias genéticas linearizadas (F_{st} Slatkin), usando as frequências alélicas, para analisar a presença de diferenciação entre a população em estudo e outras populações com dados publicados. As análises estatísticas supracitadas foram feitas no programa Arlequin versão 3.11 (Excoffier *et al.*, 2005). Para o *locus* da SCA2, a matriz das distâncias genéticas linearizadas (F_{st} Slatkin) foi representada num espaço bidimensional usando a *Multidimensional Scaling* (SPSS Inc., 2009).

CAPÍTULO 3: RESULTADOS

3.1. GENOTIPAGEM: COMPARAÇÃO DOS RESULTADOS OBTIDOS POR AF E SEQUENCIAÇÃO

Com o objectivo de confirmar o número de repetições CAG no *locus* da SCA3, previamente obtido por AF, sequenciaram-se três amostras: um indivíduo homozigótico para o alelo 14 e dois homozigóticos para o alelo 15. A análise do cromatograma permitiu observar que os alelos aos quais tinham sido atribuídas 15 repetições CAG na AF apresentavam um motivo repetitivo a menos. Em seguida, comparou-se o tamanho do fragmento obtido na AF, para os restantes alelos, com o correspondente obtido para os três alelos sequenciados e efectuou-se a respectiva correcção (± 1 CAG) sempre que necessário (ver discussão).

Em relação aos *loci* da SCA2 e SCA6 observou-se uma discrepância de mais três repetições CAG entre os resultados obtidos por AF e por sequenciação, enquanto no *locus* da DRPLA a diferença observada foi de duas repetições CAG a mais. Assim, todos os alelos sequenciados de cada *locus* evidenciaram uma discrepância relativamente aos resultados obtidos por AF: SCA2 (+3CAG), SCA6 (+3CAG) e DRPLA (+2CAG) (Fig. 2). Em conformidade com estas conclusões efectuou-se a correcção do número de repetições CAG dos alelos observados em cada *locus*: SCA2 [alelo AF (pb) + (CAG)₃], SCA6 [alelo AF (pb) + (CAG)₃] e DRPLA [alelo AF (pb) + (CAG)₂].

3.2. ANÁLISES INTRA-POPULACIONAIS

3.2.1. *Locus* da SCA3

A distribuição alélica mostrou a presença de 17 variantes para o segmento que contém o motivo CAG no gene *ATXN3*. Todos os alelos observados estavam dentro da amplitude normal, variando entre 14 e 35 repetições [$21,13 \pm 5,37$ (DP)]; a variância da frequência alélica foi de 28,86. As frequências alélicas revelaram uma distribuição bimodal, sendo os alelos 14 e 23 os mais frequentes (ambos com 29,5%). A frequência de alelos normais grandes (> 27 CAG) foi de 7,1% (Fig. 3), não tendo sido encontrados alelos intermediários. A composição genotípica revelou 32 combinações diferentes, sendo o genótipo 14/23 o mais representativo (16,7%), seguindo-se os genótipos 14/14 e 23/23 (ambos com 9%) (Fig. 4). As frequências genotípicas estavam em conformidade com o EHW ($P = 0,85063$), e o valor de heterozigotia observado foi de 80,8%. A distribuição alélica mostrou uma assimetria positiva [$0,067 \pm 0,194$ (assimetria \pm erro padrão, EP)] e uma curtose negativa [$- 0,576 \pm 0,386$ (EP)].

Para os seis alelos sequenciados foi possível obter também a configuração do padrão interno das repetições CAG, tendo estes alelos revelado a seguinte disposição, $(CAG)_2CAAAG(CAG)CAA(CAG)_8$ (Tabela 5).

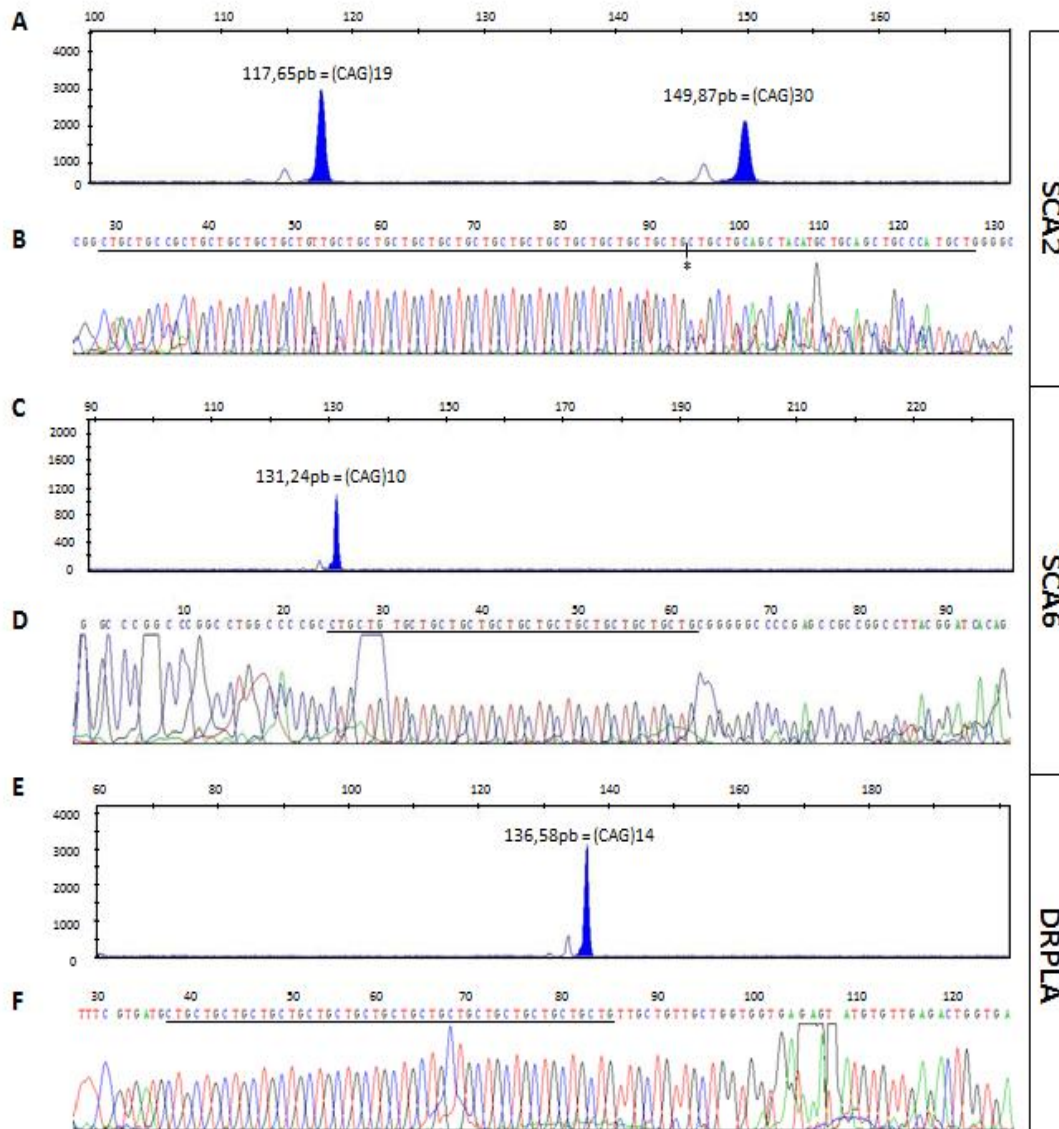


Fig. 2 – Genótipo de três indivíduos para os *loci* da SCA2 (A,B), SCA6 (C,D) e DRPLA (E,F) obtidos respectivamente por análise de fragmentos e sequenciação. Cada pico na análise de fragmentos está identificado com o respectivo tamanho de fragmento (pb) e correspondente alelo. Na sequenciação as repetições CTG (por complemento com CAG, uma vez que, foi usado para cada um dos *locus* o *primer* reverso) estão sublinhadas. No indivíduo heterozigótico, as repetições CTG do alelo 1 estão delimitadas por um asterisco.

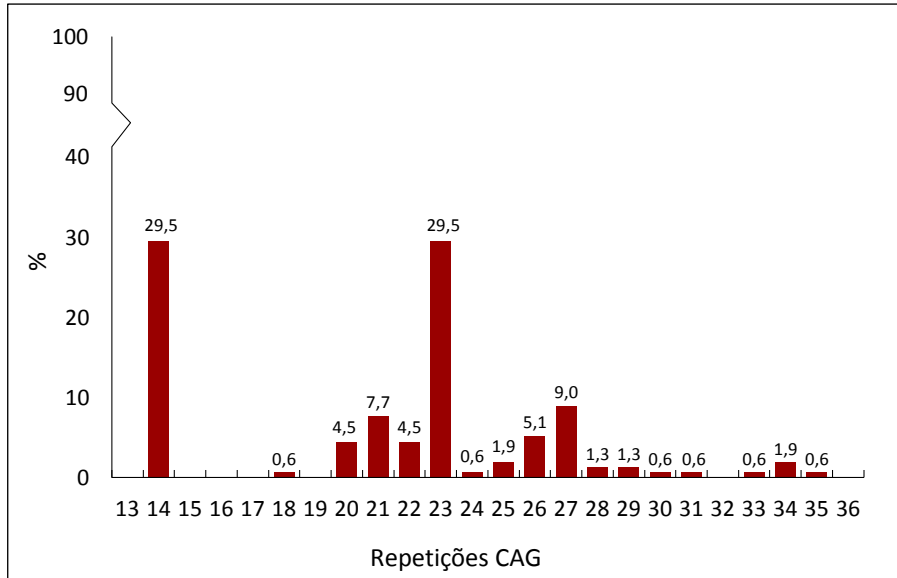


Fig. 3 – Distribuição alélica no *locus* da MJD/SCA3 em 156 cromossomas de indivíduos saudáveis não aparentados da população dos Açores.

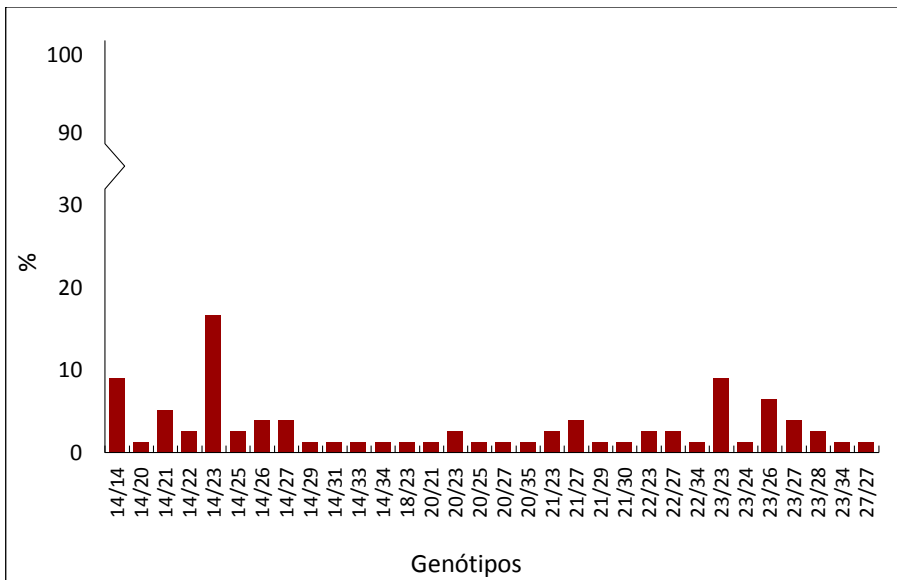












Fig. 4 – Distribuição genotípica no *locus* da MJD/SCA3 de 78 indivíduos saudáveis não aparentados da população dos Açores.

Tabela 5 – Polimorfismo das repetições CAG e padrão das interrupções CAA e AAG em alelos normais e intermediários em cada *locus*.

<i>Locus</i>	N.º de alelos	(CAG)n	Padrão das interrupções CAA e AAG dentro das repetições CAG
SCA3	6	14	2+-1+8 
	6	22	8+4+8 
SCA2	2	22	13+8 
	1	29	20+8 
	1	33	24+8 
SCA6	2	4	4 
	5	11	11 
	3	13	13 
DRPLA	4	8	8 
	4	16	16 

As repetições CAG estão representadas por círculos vazios e as interrupções CAA (+) e AAG (-) por círculos pretos e cinzentos, respectivamente.

3.2.2. *Locus* da SCA2

A análise da composição alélica revelou a presença de sete variantes alélicas para o segmento que contém o motivo CAG no gene *ATXN2*. O número de repetições CAG estava dentro da amplitude normal, variando entre 15 a 33 repetições [média 22,20 ± 1,29 (DP)]; a variância da frequência alélica foi de 1,66. No total de 156 cromossomas analisados 89,1% tinham o alelo 22 sendo a frequência de alelos normais grandes (> 22 CAG) de 9,6% (Fig. 5). Foi encontrado um alelo intermediário (0,6%) com 33 repetições. Existiam oito genótipos, sendo o genótipo 22/22 o mais frequente (80,8%), seguido do genótipo 22/23 (9%) (Fig. 6). A distribuição genotípica demonstrou que existia apenas uma combinação (15/22) à esquerda do genótipo mais frequente e seis à direita deste, indicando uma assimetria positiva na curva de distribuição alélica. As frequências genotípicas estavam em conformidade com o EHW ($P = 0,25394$), com um valor de heterozigotia observado de 16,7%. A distribuição alélica mostrou uma curtose positiva [43,332 ± 0,386 (EP)] e uma assimetria positiva [4,071 ± 0,194 (EP)], com uma clara tendência para os alelos normais de maior tamanho.

A sequenciação dos oito alelos com 22 repetições CAG permitiu observar dois padrões distintos das repetições CAG, $(CAG)_8CAA(CAG)_4CAA(CAG)_8$ e $(CAG)_{13}CAA(CAG)_8$ (Tabela 5). Por sua vez, a sequenciação dos alelos 29 e 33 revelou que estes possuíam apenas uma interrupção CAA, mostrando a seguinte configuração $(CAG)_{20}CAA(CAG)_8$ e $(CAG)_{24}CAA(CAG)_8$, respectivamente (Tabela 5).

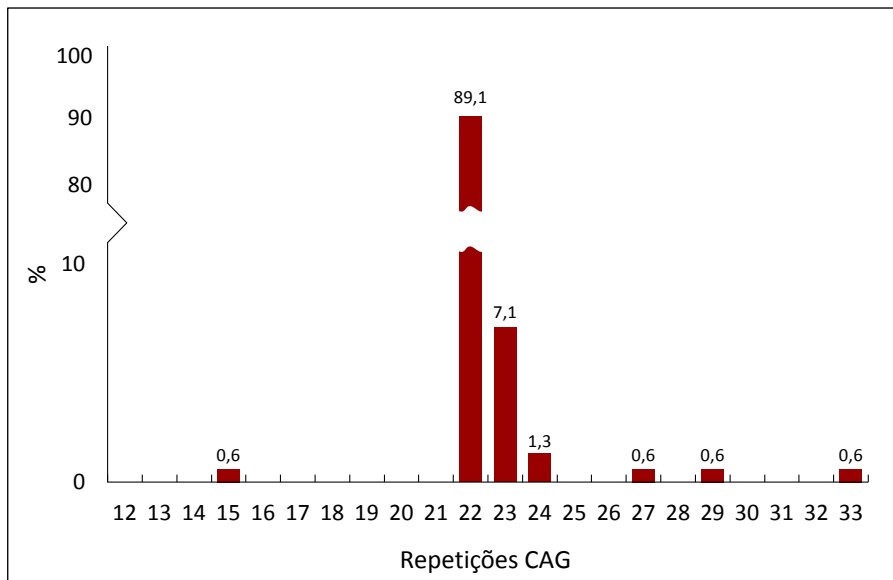


Fig. 5 – Distribuição alélica no *locus* da SCA2 em 156 cromossomas de indivíduos saudáveis não aparentados da população dos Açores.

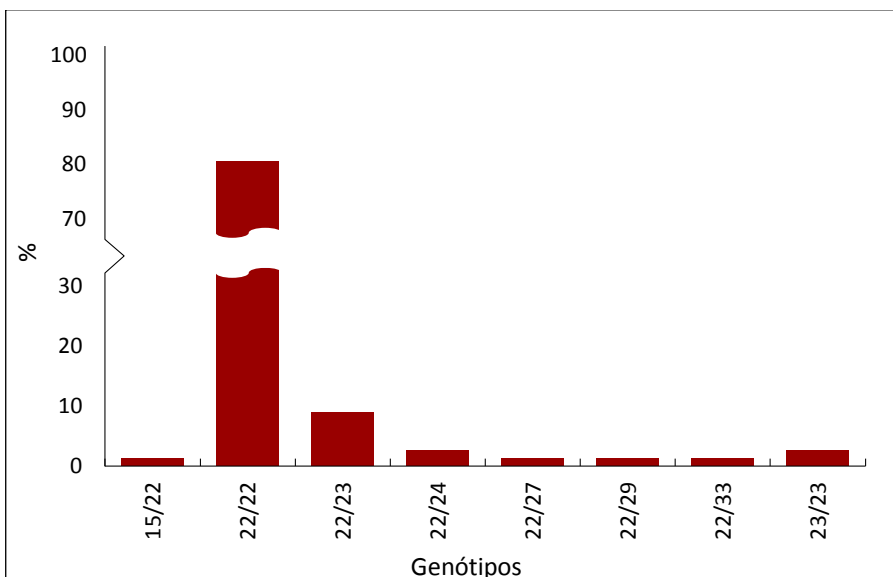


Fig. 6 – Distribuição genotípica no *locus* da SCA2 de 78 indivíduos saudáveis não aparentados da população dos Açores.

3.2.3. Locus da SCA6

A análise da diversidade alélica no gene *CACNA1A* mostrou que as oito variantes alélicas se encontravam dentro da amplitude normal, variando entre 4 e 14 repetições [média $11,03 \pm 2,39$ (DP)]. A variância da frequência alélica foi de 5,70. Nos 156 cromossomas analisados, foram detectados seis alelos com quatro repetições CAG (3,8%). Os alelos mais frequentes foram o alelo 11 (35,9%), seguindo-se os alelos 13 (31,4%), 7 (12,8%) e 12 (11,5%). Os alelos normais grandes (> 13 CAG) representavam 3,2% da amostra. Não foram encontrados alelos intermediários (Fig. 7). A distribuição genotípica revelou que o genótipo 11/13 foi o mais frequente (25,6%), seguido pelos genótipos 7/11, 11/11 (ambos com 11,5%), 12/13 (10,3%) e 13/13 (9%), tendo sido encontrados 18 genótipos diferentes na amostra estudada (Fig. 8). As frequências genotípicas estavam em conformidade com o EHW ($P = 0,20377$). O valor de heterozigotia observado foi de 73,1%. A distribuição alélica revelou uma assimetria negativa [$-1,397 \pm 0,194$ (EP)], sendo que 53,8% dos cromossomas observados tinham um número de repetições CAG inferior ou igual à moda (alelo 11), observando-se, ainda, uma curtose positiva [$1,320 \pm 0,386$ (EP)].

Para os 10 alelos sequenciados foi possível observar também a configuração da estrutura interna das repetições CAG, tendo-se verificado que todos estes alelos tinham uma estrutura de repetições pura (Tabela 5).

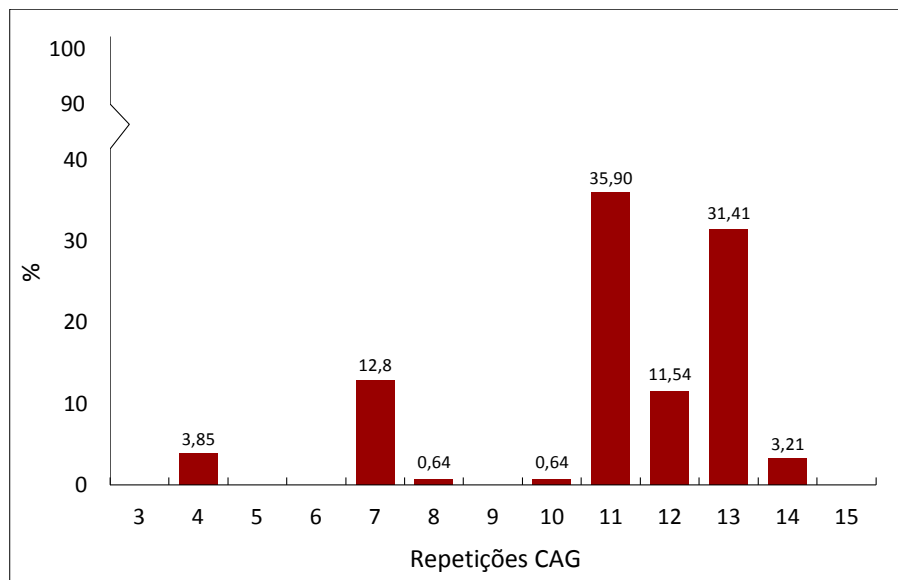


Fig. 7 – Distribuição alélica no locus da SCA6 em 156 cromossomas indivíduos saudáveis não aparentados da população dos Açores.

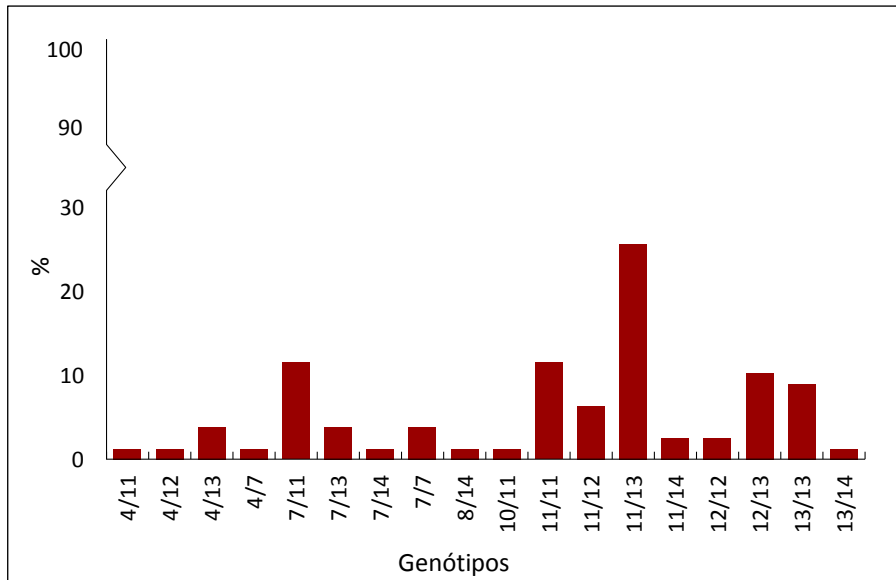


Fig. 8 – Distribuição genotípica no *locus* da SCA6 de 78 indivíduos saudáveis não aparentados da população dos Açores.

3.2.4. *Locus* da DRPLA

A análise da distribuição alélica no gene *ATN1* revelou a presença de 14 variantes. Todos os alelos estavam dentro da amplitude normal, variando entre 8 e 23 repetições [média $13,80 \pm 3,51$ (DP)]; a variância da frequência alélica foi de 12,31. Os alelos 15 e 16 foram os mais frequentes (23,7% e 17,8%, respectivamente). Os alelos normais grandes (> 17 repetições CAG) estavam presentes em 10,5% da população, e não foram encontrados alelos intermediários (Fig. 9). Na amostra estudada foram observadas 38 combinações genotípicas diferentes, sendo os genótipos 8/15, 8/16 e 15/15 os mais frequentes (cada um com 6,6%) (Fig. 10). As frequências genotípicas estavam em conformidade com o EHW ($P = 0,80715$), com uma heterozigotia observada de 84,2%. A distribuição do motivo CAG mostrou uma assimetria negativa [$- 0,271 \pm 0,197$ (EP)], com 67,1% dos cromossomas analisados tendo um número de repetições inferior ou igual à moda (alelo 15), e mostrou, ainda, uma curtose negativa [$- 0,608 \pm 0,391$ (EP)].

A sequenciação dos oito alelos permitiu observar também a estrutura interna das repetições CAG, tendo-se verificado que todos estes alelos tinham uma estrutura de repetições pura (Tabela 5).

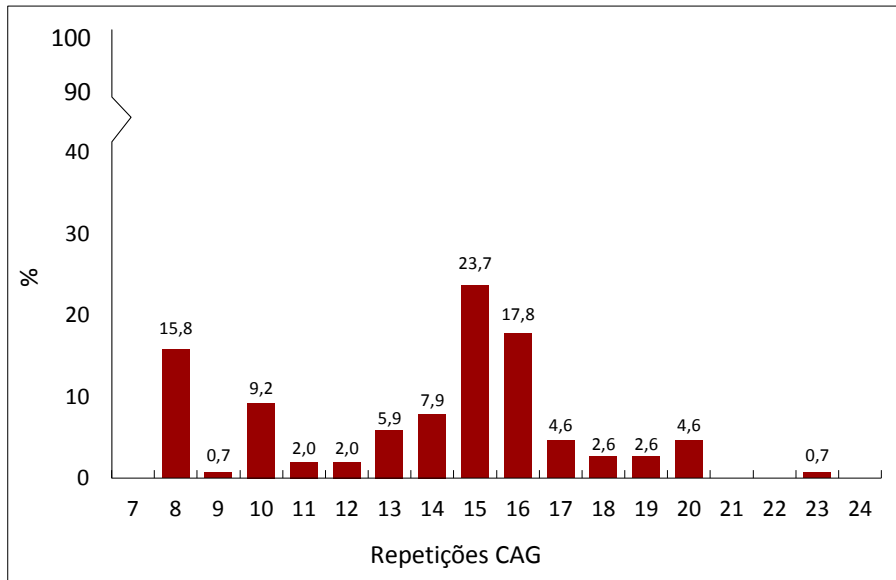


Fig. 9 – Distribuição alélica no *locus* da DRPLA em 152 cromossomas de indivíduos saudáveis não aparentados da população dos Açores.

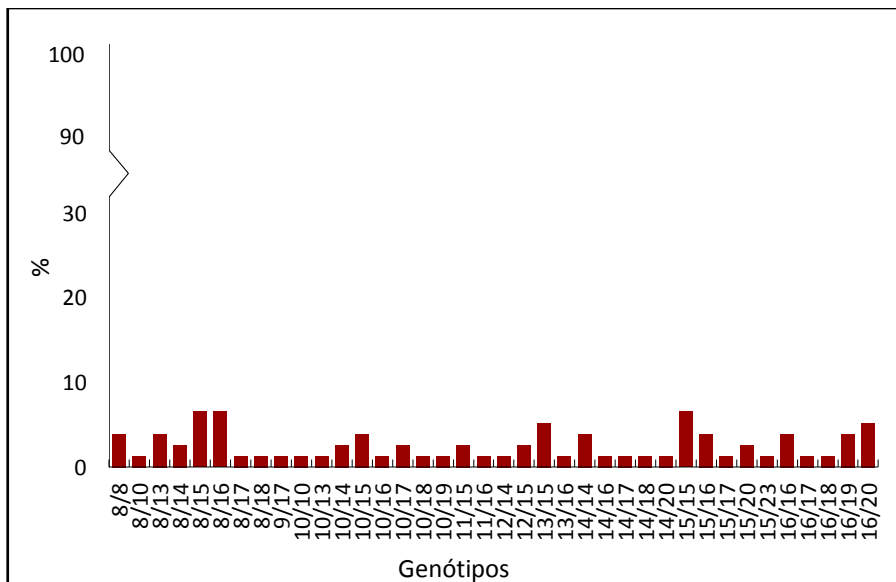


Fig. 10 – Distribuição genotípica no *locus* da DRPLA de 76 indivíduos saudáveis não aparentados da população dos Açores.

3.1. ANÁLISES INTER-POPULACIONAIS

O teste exacto de diferenciação populacional com base nas frequências alélicas no *locus* da SCA3 revelou a ausência de diferenças significativas entre a amostra estudada e os resultados obtidos por Bettencourt *et al.* (2008) para a população dos Açores. Por outro lado, o perfil alélico revelou diferenças significativas com a população

da Finlândia (Juvonen *et al.*, 2005) e do Brasil (Freund *et al.*, 2009). No *locus* da SCA2, não foram encontradas diferenças significativas entre o perfil alélico da amostra estudada e as populações da Polónia (Sobczak & Krzyzosiak, 2004), Índia (Choudhry *et al.*, 2001), Suíça e Alemanha (Gispert *et al.*, 2012), tendo sido encontradas, todavia, diferenças significativas com a população da Finlândia (Juvonen *et al.*, 2005), Brasil (Freund *et al.*, 2009) e Portugal Continental (Ramos *et al.*, 2010). O perfil alélico no *locus* da SCA6 mostrou diferenças significativas com a população da Finlândia (Juvonen *et al.*, 2005) e do Brasil (Freund *et al.*, 2009). O perfil alélico observado na amostra em estudo, no *locus* da DRPLA, revelou diferenças significativas com a população da Finlândia (Juvonen *et al.*, 2005) (Tabela 6). Globalmente os resultados do teste exacto de diferenciação demonstraram ausência de diferenças significativas entre a amostra estudada e as populações europeias (com excepção da população finlandesa), ainda e diferenças significativas com a população brasileira.

Tabela 6 – Resultados dos testes exactos de diferenciação populacional e do cálculo da matriz de distância genética linearizadas entre populações (F_{st} Slatkin) ($^+p < 0,05$, $^{++}p < 0,01$, $^{+++}p < 0,005$, $^{++++}p < 0,00005$, $^{+++++}p < 0,00001$, diferenças não significativas (-) e informação não disponível (IND)).

	Populações	SCA3	SCA2	SCA6	DRPLA
Teste exacto	Finlândia	+++++	+	++++	+++++
	Brasil	+++++	+++++	+++++	IND
	Suíça	IND	-	IND	IND
	Alemanha	IND	-	IND	IND
	Polónia	IND	-	IND	IND
	Índia	IND	-	IND	IND
	Portugal	IND	+++	IND	IND
	Açores	-	IND	IND	IND
F_{st} Slatkin	Finlândia	0,00796 ⁺⁺	0,00710	0,03150 ⁺⁺⁺⁺	0,04017 ⁺⁺⁺⁺
	Brasil	0,19871 ⁺⁺⁺⁺⁺	3,34273 ⁺⁺⁺⁺⁺	0,34761 ⁺⁺⁺⁺⁺	IND
	Suíça	IND	0,00290	IND	IND
	Alemanha	IND	0,00000	IND	IND
	Polónia	IND	0,00000	IND	IND
	Índia	IND	0,00000	IND	IND
	Portugal	IND	0,01484	IND	IND
	Açores	0,00485	IND	IND	IND

A matriz de distâncias genéticas linearizadas (F_{st} Slatkin), construída com base nas frequências alélicas do *locus* da SCA3 demonstrou que a amostra estudada se diferencia das populações da Finlândia (Juvonen *et al.*, 2005) e do Brasil (Freund *et al.*, 2009), sendo estatisticamente semelhante à população dos Açores previamente estudada (Bettencourt *et al.*, 2008). Para melhor compreender a relação entre as populações analisadas, no *locus* da SCA2, a matriz de distâncias genéticas linearizadas (F_{st} Slatkin) foi representada num espaço bidimensional utilizando a *Multidimensional Scaling*, que confirmou que a população do Brasil (Freund *et al.*, 2009) se diferencia das outras populações (Choudhry *et al.*, 2001; Sobczak & Krzyzosiak, 2004; Juvonen *et al.*, 2005; Ramos *et al.*, 2010; Gispert *et al.*, 2012), inclusivamente da amostra estudada (Fig. 11). No *locus* da SCA6, a amostra em estudo diferencia-se das populações da Finlândia (Juvonen *et al.*, 2005) e do Brasil (Freund *et al.*, 2009) (Tabela 6). No *locus* da DRPLA, a amostra estudada diferencia-se da população da Finlândia (Juvonen *et al.*, 2005) (Tabela 6).

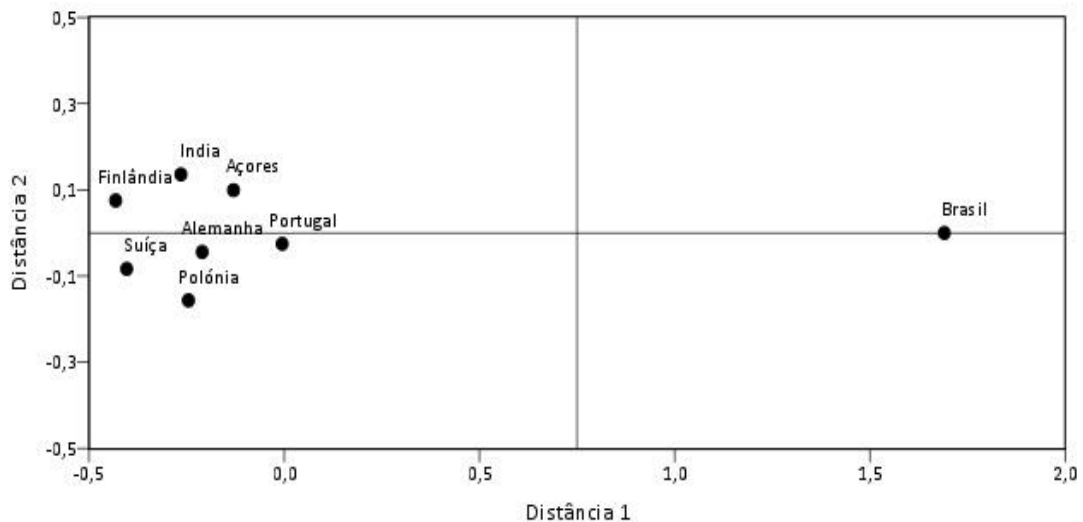


Fig. 11 – *Multidimensional Scaling* usando as distâncias genéticas linearizadas (F_{st} Slatkin), para o *locus* da SCA2, entre as diferentes populações.

CAPÍTULO 4: DISCUSSÃO

4.1. ANÁLISE DE FRAGMENTOS VS. SEQUENCIAÇÃO

As SCAs apresentam fenótipos neurológicos semelhantes, sendo por isso necessário para um diagnóstico exacto o recurso a testes moleculares, que detectam a mutação no gene causal da doença (Schelhaas *et al.*, 2000). Vários estudos têm determinado o número de repetições dos microsátélites nas doenças de expansão por análise de fragmentos (ver, entre outros, Bettencourt *et al.*, 2008; Freund *et al.*, 2009; Daoud *et al.*, 2011; Ross *et al.*, 2011). Durante este procedimento, todavia, é necessário ter em conta que a conversão dos pares de bases em unidades de repetição pode não ser linear e que a determinação do tamanho do fragmento por comparação apenas com o marcador padrão de tamanho molecular conhecido pode não ser exacta, sendo recomendado utilizar uma amostra de genótipo conhecido (Sequeiros *et al.*, 2010). Neste estudo, porém, não foi possível usar um marcador alélico (conjunto de alelos descritos para um *locus*), uma vez que, que os *loci* da SCA2, SCA6 e DRPLA nunca foram anteriormente genotipados nas instalações do CIRN, na Universidade dos Açores. Em relação ao *locus* da SCA3, embora já tivesse sido feita a genotipagem de várias amostras, estes alelos não foram no entanto confirmados por sequenciação, não podendo por isso ser usados como marcador alélico. Neste estudo, os diferentes resultados obtidos por AF e sequenciação, nos *loci* da SCA3 (± 1 CAG), SCA2 (+ 3 CAG), SCA6 (+ 3 CAG) e DRPLA (+ 2 CAG), vieram confirmar que a utilização de um marcador alélico é essencial para a determinação correcta do tamanho do alelo. No entanto, no *locus* da SCA3, apenas se verificaram diferenças no tamanho de alguns dos alelos. A diferença observada entre os resultados obtidos no *locus* da SCA3 e nos restantes *loci* poderá dever-se ao tamanho do fragmento no *locus* da SCA3 ser maior do que o tamanho do fragmento dos outros *loci*, e também ao sequenciador e, concretamente, Polímero (POP7) utilizados no SEEBMO, os quais permitem a obtenção de resultados mais sensíveis e precisos.

4.2. *LOCUS* DA SCA3

A distribuição alélica no *locus* da SCA3 confirmou a inexistência de alelos no intervalo expandido na amostra analisada, que poderá ser utilizada como população controlo, na segunda fase da investigação.

O perfil alélico na amostra estudada foi estatisticamente idêntico ao reportado anteriormente para a população açoriana (Bettencourt *et al.*, 2008), e estatisticamente

diferente do observado na população finlandesa (Juvonen *et al.*, 2005) e brasileira (Freund *et al.*, 2009). Os resultados da distância genética revelaram uma proximidade significativa entre a amostra estudada e a população açoriana (Bettencourt *et al.*, 2008), e uma distância significativa com as populações finlandesa (Juvonen *et al.*, 2005) e brasileira (Freund *et al.*, 2009). Porém, as diferenças significativas observadas entre a amostra estudada e a população finlandesa poderão dever-se, em parte, ao facto de terem sido analisados, na população finlandesa, cromossomas normais de doentes. Devido à escassez de dados comparáveis acerca da frequência de alelos normais no *locus* da SCA3 foram feitas comparações genéricas, a partir de gráficos das frequências alélicas de outras populações. Globalmente, as frequências alélicas observadas na amostra estudada foram semelhantes às observadas nas populações gaulesa (Majounie *et al.*, 2009), de Portugal Continental (Silveira *et al.*, 2002; Lima *et al.*, 2005) e brasileira (Silveira *et al.*, 2002). Por outro lado, a distribuição alélica parece ser diferente da observada nas populações japonesa e norte americana (Takano *et al.*, 1998; Lee *et al.*, 2011a). A semelhança entre o perfil alélico da amostra estudada e da população brasileira (Silveira *et al.*, 2002) poderá ser explicada com base na ligação histórica entre estas duas populações, pois, segundo Martins *et al.* (2007), existe uma baixa diversidade e uma relação filogenética próxima entre os haplótipos dos doentes de famílias portuguesas e brasileiras; ambas as séries de doentes brasileiros analisados (Silveira *et al.*, 2002; Martins *et al.*, 2007) são naturais da região Rio Grande do Sul. Segundo a classificação de Andrés *et al.* (2003) a média de tamanho dos alelos na amostra estudada foi elevada (média > 20 CAG), sendo concordante com os resultados observados por diversos autores (ver, entre outros, Lima *et al.*, 2005; Bettencourt *et al.*, 2008).

A variância observada na amostra estudada (28,86) foi superior à observada na população indiana (18,09; n = 198) (Andrés *et al.*, 2003) e parece ser semelhante à observada na população europeia (26,76; n = 641) (Andrés *et al.*, 2003), embora o tamanho da amostra analisado nesta população seja cerca de quatro vezes superior, o que poderá indicar que a amostra estudada apresenta uma maior variabilidade quando comparada com esta população. O valor de heterozigotia observado (H = 80,8%) neste estudo parece ser semelhante ao valor observado numa amostra representativa da população portuguesa (inclui indivíduos normais de todos os distritos portugueses) (H = 80%; n = 1,844) (Lima *et al.*, 2005) e numa amostra de indivíduos normais açorianos (H

= 78,9%; n = 256) (Bettencourt *et al.*, 2008). A sequenciação dos alelos revelou que o padrão interno das repetições CAG, (CAG)₂CAAAAG(CAG)CAA(CAG)₈, estava de acordo com anteriormente descrito por Kawaguchi *et al.* (1994).

Segundo vários autores a frequência de alelos normais de maior tamanho está correlacionada com a frequência relativa da doença (ver, entre outros, Takano *et al.*, 1998; Basu *et al.*, 2000). Todavia, os resultados observados na amostra em estudo parecem não estar em concordância com esta teoria, pois embora, a prevalência da MJD nos Açores seja de 1:2597 (Araújo, submetida), apenas 7,1% dos alelos observados possuem mais de 27 repetições CAG. Por outro lado, estes resultados parecem estar em concordância com os resultados reportados por Lima *et al.* (2005). Os resultados observados num estudo recente, feito na população açoriana (Bettencourt *et al.*, 2008), parecem estar também em conformidade com os estes resultados, dado que apenas foram observados ≈ 10% de alelos normais grandes. Noutros estudos, realizados nas populações do País de Gales (Majounie *et al.*, 2007) e do Brasil (Freund *et al.*, 2009), não foram também reportadas quaisquer associações. Freund *et al.* (2009) sugeriram, todavia, que um estudo realizado numa amostra maior, onde estivessem incluídas diversas regiões do Brasil com diferentes frequências da MJD, poderia levar a uma correlação negativa ou positiva entre a frequência dos alelos normais grandes e a frequência da doença. Segundo Lima *et al.* (2005) a ausência de alelos intermediários na população normal é importante para compreender o padrão mutacional no gene *ATXN3*. Num estudo recente, Martins *et al.* (2006) demonstraram a existência de um mecanismo mutacional *multistep*. Para além disso, o principal factor que influencia a instabilidade dos alelos parece ser o *background* do cromossoma e não o número de repetições CAG e interrupções (que ocorrem no início das repetições CAG) (Igarashi *et al.*, 1996; Martins *et al.*, 2008).

Neste estudo, a curva de distribuição alélica revelou uma assimetria positiva, em oposição à assimetria negativa (excesso de alelos normais pequenos) observada numa amostra portuguesa [- 0,185 ± 0,057 (EP)] (Lima *et al.*, 2005) e numa amostra açoriana [- 0,249 ± 0,152 (EP)] (Bettencourt *et al.*, 2008). A diferença observada, todavia, poderá ser devido à diferença no tamanho da amostra, uma vez que, o número da amostra em estudo foi inferior ao analisado nestas populações. Embora a distribuição alélica da amostra estudada indique um desvio à direita, 84% dos alelos normais tinham menos de 27 repetições, e destes 76,3% tinham um número de repetições inferior ou igual a

23 CAG. Este valor foi semelhante ao observado por Bettencourt *et al.* (2008) e superior ao observado por Lima *et al.* (2005).

4.3. Locus DA SCA2

O perfil alélico reportado para o *locus* da SCA2 foi estatisticamente idêntico ao observado noutras populações europeias, nomeadamente ao da população polaca, suíça, alemã e indiana (Choudhry *et al.*, 2001; Sobczak & Krzyzosiak, 2004; Gispert *et al.*, 2012); por outro lado, foi significativamente distinto do observado nas populações finlandesa (Juvonen *et al.*, 2005), de Portugal Continental (Ramos *et al.*, 2010) e brasileira (Freund *et al.*, 2009). As diferenças observadas entre a amostra estudada e as populações de Portugal Continental (Ramos *et al.*, 2010) e da Finlândia (Juvonen *et al.*, 2005) poderão dever-se ao facto de terem sido analisados nestas populações cromossomas normais de doentes. Os resultados da distância genética revelaram uma proximidade significativa entre a amostra em estudo e as populações europeias e indiana (Choudhry *et al.*, 2001; Sobczak & Krzyzosiak, 2004; Juvonen *et al.*, 2005; Bettencourt *et al.*, 2008; Ramos *et al.*, 2010; Gispert *et al.*, 2012), e uma distância genética significativa com a população brasileira (Freund *et al.*, 2009). De modo global, a distribuição alélica da amostra estudada parece ser similar à observada na população gaulesa (Majounie *et al.*, 2007), e distinta da observada nas populações cubana (Laffita-Mesa *et al.*, 2012), belga e holandesa (Van Damme *et al.*, 2011). Em todas estas populações, incluído na amostra estudada, o alelo 22 foi o mais frequente, variando entre um mínimo de 75,92% na população cubana (Laffita-Mesa *et al.*, 2012) e um máximo de 92,8% na população finlandesa (Juvonen *et al.*, 2005). Na amostra em estudo não foi reportado nenhum alelo 13 ou 26, estando estes resultados de acordo com Laffita-Mesa *et al.* (2012), os quais indicaram que o alelo 13 foi descrito somente na população cubana, e o alelo 26 nas populações cubana e checa (Bauer *et al.*, 2005). A média do número de repetições CAG observado na amostra em estudo foi designada “longa” de acordo com a classificação de Andrés *et al.* (2003). Os genótipos 22/22 (60,88%) e 22/23 (13,75%) foram os mais frequentes na amostra estudada, assim como, na população cubana (Laffita-Mesa *et al.*, 2012), sendo que ambas as populações mostraram uma distribuição genotípica com um desvio à direita.

A variância observada na amostra estudada aparenta ser menor do que a observada na população geral cubana (1,66 vs. 5,31) (Laffita-Mesa *et al.*, 2012). Esta

desigualdade poderá dever-se ao elevado grau de polimorfismo, do *locus* da SCA2, na população cubana, uma vez que, o tamanho da amostra estudada foi cerca de duas vezes a amostra cubana. A amostra em estudo ($H = 16,7\%$) possui uma menor diversidade genética do que a população de Portugal Continental ($H = 31\%$, $n = 91$) (Silveira *et al.*, 2002), e uma diversidade similar à observada nas populações indiana ($H = 14,3\%$, $n = 200$) (Andrés *et al.*, 2003) e brasileira ($H = 15\%$, $n = 54$) (Silveira *et al.*, 2002). A distribuição alélica observada na amostra estudada revelou uma assimetria positiva [$4,071 \pm 0,194$ (EP)], indicando uma tendência para alelos normais de maior tamanho. O desvio à direita na curva de distribuição alélica na amostra estudada parece ser superior ao reportado para a população geral cubana [$1,66 \pm 0,26$] (Laffita-Mesa *et al.*, 2012) e para os cromossomas normais de famílias de Portugal Continental com SCA2 [$3,364 \pm 0,239$ (EP)] (Ramos *et al.*, 2010), ao passo que os cromossomas normais de doentes na população indiana apresentam um desvio à esquerda (Basu *et al.*, 2000).

A sequenciação de alguns alelos com 22 repetições CAG revelou a existência de apenas duas configurações internas (8+4+8 e 13+8). Estas configurações ocorrem em todas as populações estudadas em diferentes rácios: 3:1, 5:1, 4,5:1, 3,5:1 e 6:1, na amostra estudada, polaca (Sobczak & Krzyzosiak, 2004), francesa (Imbert *et al.*, 1996), indiana (Choudhry *et al.*, 2001) e japonesa (Sanpei *et al.*, 1996), respectivamente. A diferença na frequência das duas variantes parece estar de acordo com Ramos *et al.* (2011), os quais sugeriram que o alelo $(CAG)_{13}CAA(CAG)_8$ foi originado a partir da perda da interrupção CAA a 5' no alelo $(CAG)_8CAA(CAG)_4CAA(CAG)_8$. Por outro lado, os alelos $(CAG)_{20}CAA(CAG)_8$ e $(CAG)_{24}CAA(CAG)_8$ observados na amostra em estudo foram descritos somente nas populações indiana (Choudhry *et al.*, 2001) e cubana (Laffita-Mesa *et al.*, 2012), respectivamente. Estes resultados poderão estar relacionados com facto de ter sido documentado, na província de Holguín (Cuba) e na população indiana, um efeito fundador, isto é, todos os alelos expandidos possuem um haplótipo comum (ver, entre outros, Choudhry *et al.*, 2001; Hernandez *et al.*, 1995).

Segundo vários autores, de entre os quais Cleary & Pearson (2003), os alelos de maior tamanho possuem uma maior probabilidade de expansão patogénica do que os alelos normais pequenos, e podem contribuir para o aparecimento de mutações *de novo* (Laffita-Mesa *et al.*, 2012). Este estudo revelou, na amostra estudada, a existência de uma tendência para os alelos normais grandes. Os resultados observados sugerem a

inexistência de correlação entre a frequência destes alelos e a ausência de casos reportados de SCA2 na região dos Açores. Por outro lado, Laffita-Mesa *et al.* (2012) sugerem que os alelos intermediários na população cubana possam ter evoluído dos alelos com 23-31 repetições, tendo em conta a elevada frequência (19%) destes alelos. Dada a presença de instabilidade inter-geracional, tais alelos podem dar origem a alelos de penetrância completa, nas sucessivas gerações. Resultados observados por Takano *et al.* (1998) revelaram igualmente que a elevada frequência destes alelos na população caucasiana, em comparação com a japonesa, está correlacionada com uma maior prevalência da doença nas populações caucasianas. Noutras populações, no entanto, parece não haver uma correlação entre a frequência de alelos normais grandes e a prevalência da doença (ver, entre outros, Basu *et al.*, 2000; Saleem *et al.*, 2000; Majounie *et al.*, 2007). Estes resultados parecem indicar que a frequência dos alelos normais de maior tamanho poderão não ser um bom indicador da frequência da doença nestas populações, pois, segundo vários autores, os alelos sem interrupções ou com apenas uma interrupção CAA na extremidade 3' das repetições têm uma maior predisposição para a instabilidade e expansão (Saleem *et al.*, 2000; Choudhry *et al.*, 2001; Sobczak & Krzyzosiak, 2004), dado que o primeiro passo do processo mutacional é a perda da interrupção CAA a 5'. No caso da SCA2, a frequência destes alelos será assim um melhor indicador da incidência da doença, em oposição, à frequência dos alelos normais de maior tamanho, contribuindo para explicar melhor as diferenças observadas entre diferentes populações (Saleem *et al.*, 2000; Sobczak & Krzyzosiak, 2004). Para podermos atestar a validade dos resultados reportados neste estudo seria preciso sequenciar todos os alelos normais grandes para determinar a configuração interna das repetições CAG, de modo a verificar se estes alelos têm uma propensão à expansão, em sucessivas gerações. Para além disso, é necessário também aumentar o tamanho da amostra estudada para aumentar o poder estatístico.

Na amostra estudada foi observado um alelo intermediário com 33 repetições CAG, com apenas uma interrupção CAA na extremidade 3'. Na maioria dos estudos realizados em populações controlo, não foram observados alelos com ≥ 31 repetições CAG (ver, entre outros, Freund *et al.*, 2009; Corrado *et al.* 2011; Lee *et al.*, 2011b; Van Damme *et al.*, 2011; Gispert *et al.*, 2012), com excepção dos estudos realizados nas populações de descendência caucasiana (Ross *et al.*, 2011), francesa (Daoud *et al.*, 2011) e alemã (Gispert *et al.*, 2012). Todos os alelos intermediários observados, tanto

na amostra estudada como noutras populações, foram encontrados em heterozigotia (ver, entre outros, Corrado *et al.*, 2011; Ross *et al.*, 2011).

4.3.1. ASSOCIAÇÃO DAS REPETIÇÕES CAG COM DOENÇAS NEURODEGENERATIVAS

SCA2

Os alelos intermediários com 32 e 33 repetições CAG possuem uma penetrância reduzida, tendo sido somente descritos em alguns doentes SCA2 com um início muito tardio dos sintomas (Futamura *et al.*, 1998; Santos *et al.*, 1999; Fernandez *et al.*, 2000). Para além disso, sabe-se ainda que a vasta maioria dos alelos patogénicos com mais de 30 CAG são ininterruptos, e portanto, instáveis (Imbert *et al.*, 1996; Pulst *et al.*, 1996; Sanpei *et al.*, 1996), ainda que Costanzi-Porrini *et al.* (2000) tenham reportado dois doentes que tinham um alelo 34 com uma interrupção CAA. Foram também observados alelos intermediários (32 e 33 CAG) com pelo menos uma interrupção CAA em doentes SCA2 com parkinsonismo (Hussey *et al.*, 2002; Silveira *et al.*, 2002; Payami *et al.*, 2003; Kim *et al.*, 2007). Tendo em conta os factos acima indicados, poderá dizer-se que o indivíduo que possui o alelo intermediário poderá possuir uma susceptibilidade para desenvolver a doença SCA2, porém, o facto de o alelo possuir uma interrupção CAA poderá indicar que este alelo é estável, não estando propenso a uma expansão, nas sucessivas gerações.

Outras doenças neurodegenerativas

Estudos recentes postularam que os alelos expandidos no gene *ATXN2* com interrupções podem produzir um fenótipo clínico mais heterogéneo assemelhando-se a outras doenças neurodegenerativas, incluído entre outras a PD, PSP e ALS (Furtado *et al.*, 2002; Kim *et al.*, 2007). Vários estudos observaram que os alelos com 24 a 33 CAG aumentam o risco de ALS em diversas populações, sendo que o tamanho das repetições associado a este aumento varia entre diferentes populações (ver, entre outros, Corrado *et al.*, 2011; Daoud *et al.*, 2011; Ross *et al.*, 2011; Van Damme *et al.*, 2011; Gispert *et al.*, 2012). Ross *et al.* (2011) observaram que os alelos com mais de 30 CAG estão associados a um aumento do risco de desenvolver outra doença neurodegenerativa, nomeadamente a PSP. Na amostra estudada foram observados alelos com > 24 CAG (n = 3/156, 4,8%), > 27 CAG (n = 2/156, 1,2%) e > 30 CAG (n = 1/156, 0,6%), tendo a sequenciação dos alelos com mais de 27 repetições revelado apenas uma interrupção

CAA. A frequência dos alelos (CAG)₂₄₋₃₃ observados na amostra estudada foi semelhante à observada por Daoud *et al.* (2011) (n = 24/471, 5,1%), enquanto a frequência dos alelos (CAG)₃₁₋₃₃ foi superior à observada por Ross *et al.* (2011) (n = 4/4877, 0,2%). Embora não seja possível tirar ilações sobre a correlação destes alelos com a frequência destas doenças na região dos Açores, os resultados observados neste estudo poderão sugerir uma possível associação dos alelos normais grandes e/ou intermediários como factor de risco para o desenvolvimento destas doenças neurodegenerativas. Todavia, para poder testar a validade desta hipótese é necessário fazer um estudo caso-controlo, e aumentar o número da amostra estudada, que iria ser usada como grupo controlo.

4.4. Locus DA SCA6

A distribuição alélica reportada para amostra em estudo foi estatisticamente diferente da observada nas populações finlandesa (Juvonen *et al.*, 2005) e brasileira (Freund *et al.*, 2009). Os resultados da distância genética revelaram uma distância significativa entre a amostra em estudo e as duas populações anteriormente referidas. As diferenças observadas entre a amostra estudada e a população finlandesa poderão dever-se, em parte, devido ao facto da análise das frequências alélicas nesta população ter sido feita em cromossomas normais de doentes. Globalmente, o perfil alélico observado na amostra estudada foi similar ao observado nas populações brasileira, de Portugal Continental (Silveira *et al.*, 2002) e gaulesa (Majounie *et al.*, 2007). Por outro lado, o perfil alélico parece ser diferente da observada nas populações japonesa (Takano *et al.*, 1998) e norte americana (Lee *et al.*, 2011a).

A variância observada na amostra em estudo (5,70) foi superior à observada nas populações europeia (3,65; n = 303), indiana (4,99; n = 200), africana (1,35; n = 300) e da Ásia oriental (4,03; n = 327) (Andrés *et al.*, 2003), ainda que o tamanho da amostra seja menor, o que poderá indicar uma maior variabilidade alélica na amostra estudada. A diversidade genética observada (H = 73,1%) na amostra em estudo parece ser semelhante à reportada nas populações brasileira (H = 68%; n = 60) (Silveira *et al.*, 2002), europeia (H = 73,9%; n = 303), da Ásia Oriental (H = 74,7%; n = 327) e indiana (H = 79,6%; n = 200) (Andrés *et al.*, 2003), e ainda superior à observada nas populações africana (H = 62,3%; n = 300) (Andrés *et al.*, 2003) e de Portugal Continental (H = 60%; n = 95) (Silveira *et al.*, 2002). De acordo com a classificação de Andrés *et al.* (2003) a

média do tamanho dos alelos na amostra em estudo pode ser considerada pequena (média < 15 CAG), sendo estes resultados concordantes com os resultados reportados por este autor.

Segundo diferentes autores existe uma correlação entre a prevalência da doença e a frequência de alelos normais grandes (> 13 CAG). Os resultados observados neste estudo, onde apenas 3,2% dos alelos observados possuíam mais de 13 repetições CAG, parecem estar em concordância com esta hipótese, uma vez que, não foram reportados casos de SCA6 até ao presente momento na região dos Açores. Para podermos testar a validade desta hipótese, todavia, seria necessário aumentar o tamanho da amostra em estudo. Resultados observados na população de Portugal Continental parecem estar também em concordância com esta hipótese (Silveira *et al.*, 2002; Vale *et al.*, 2010). Takano *et al.* (1998) observou que 20% dos alelos normais na população japonesa tinham mais de 13 repetições, enquanto na população caucasiana estes correspondiam apenas a 4%, sendo estes resultados consistentes com a prevalência desta ataxia nestas populações. Os resultados reportados em duas populações brasileiras (Silveira *et al.*, 2002; Freund *et al.*, 2009) parecem suportar a hipótese de Lopes-Cendes (1999), que sugeriu que as diferenças observadas na frequência das várias SCAs em diferentes populações brasileiras poderão dever-se à diferente origem étnica desses indivíduos. O mesmo se poderá verificar na população norte americana, pois embora tenha sido reportado que esta ataxia é relativamente frequente nos Estados Unidos (Moseley *et al.*, 1998), a frequência de alelos normais grandes observados por Lee *et al.* (2011a) foi muito baixa.

Contrariamente ao observado nos *loci* da SCA3 e SCA2 a distribuição alélica no *locus da SCA6* mostrou uma assimetria negativa. Estes resultados são similares aos observados nas populações caucasiana, japonesa e finlandesa (Takano *et al.*, 1998; Juvonen *et al.*, 2005), e distinto dos resultados obtidos para a população indiana (Basu *et al.*, 2000) em que se observou um desvio à direita.

4.5. Locus DA DRPLA

O perfil alélico observado na amostra em estudo foi estatisticamente diferente do reportado para a população finlandesa (Juvonen *et al.*, 2005). Os resultados da distância genética revelaram uma distância significativa entre as duas populações. Tal como previamente referido relativamente aos outros *loci*, as diferenças observadas

entre estas populações poderão dever-se, em parte, devido ao facto do estudo das frequências alélicas na população finlandesa ter sido feita em cromossomas normais de doentes. Globalmente, o perfil alélico observado na amostra em estudo aparenta ser semelhante ao observado nas populações gaulesa, caucasiana, de Portugal Continental e brasileira (Takano *et al.*, 1998; Silveira *et al.*, 2002; Majounie *et al.*, 2007), sendo aparentemente diferente do observado na população norte americana (Lee *et al.*, 2011a), em que o alelo 27 foi o mais frequente. A média de tamanho dos alelos normais na amostra estudada pode ser de acordo com a classificação de Andrés *et al.* (2003) considerada pequena (média < 15 repetições), sendo este resultado concordante com o reportado noutras populações (ver, entre outros, Silveira *et al.*, 2002; Andrés *et al.*, 2003).

A variância observada na amostra estudada (12,31) parece ser semelhante à observada nas populações europeia (11,29; n = 156) e indiana (12,91; n = 196), e por outro lado, aparenta ser superior à observada na população africana (7,26; n = 150) (Andrés *et al.*, 2003). Na amostra em estudo, a diversidade genética observada (H = 84,2%) parece ser similar à reportada em diversas populações, tais como a de Portugal Continental (H = 81%; n = 96), brasileira (H = 82; n = 65) (silveira *et al.*, 2002), africana (H = 84,3%; n = 150), europeia (H = 85,3%; n = 156) e indiana (H = 86,1%; n = 196) (Andrés *et al.*, 2003). A sequenciação dos alelos revelou que o padrão interno das repetições CAG estava de acordo com anteriormente descrito por Koide *et al.* (1994).

A frequência dos alelos normais grandes (> 17 CAG) observados na amostra em estudo (10,5%) foi inferior à observada nas populações japonesa (24%) (Takano *et al.*, 1998) e gaulesa (11,8%) (Majounie *et al.*, 2007) e semelhante à observada na população caucasiana (10%) (Takano *et al.*, 1998), tendo sido, contudo, superior à observada nas populações de Portugal Continental (6%) e brasileira (4%) (Silveira *et al.*, 2002). Parte da diferença observada entre estas populações poderá, no entanto, dever-se a três factores: (1) utilização de diferentes técnicas para a determinação do tamanho do alelo, sendo que a técnica utilizada nestes estudos possui uma menor resolução; (2) análise de alelos normais de doentes e não indivíduos saudáveis e (3) análise de uma amostra nestas populações com tamanho superior ao da amostra estudada.

A frequência de alelos normais grandes, na amostra estudada, no entanto, parece não estar associada com a frequência da doença, dado que não foi reportado nenhum caso de DRPLA nos Açores. Estes resultados parecem estar em concordância com a

hipótese sugerida por Silveira *et al.* (2002), e recentemente corroborada por Martins *et al.* (2006). Silveira *et al.* (2002), não observaram qualquer associação entre a frequência dos alelos normais grandes e a frequência da doença (3%) na população de Portugal Continental, tendo proposto que estes resultados poderiam estar relacionados com o haplótipo associado aos alelos normais grandes. Mais tarde, Martins e colaboradores (2006) corroboraram esta hipótese, ao observarem que todos os alelos expandidos, na população portuguesa, tinham o haplótipo A1-B1, e que este estava presente também nos alelos normais grandes. Embora, contrariamente ao observado nos cromossomas normais na população japonesa (Yanagisawa *et al.*, 1996), onde se regista a maior prevalência a nível mundial, este haplótipo específico (A1-B1) não foi o mais frequente (Martins *et al.*, 2006). Os resultados deste estudo parecem estar de acordo com Silveira *et al.* (2002), não suportando a teoria de que a frequência dos alelos normais grandes está associada com a frequência da doença (Takano *et al.*, 1998), todavia, seria necessário fazer um estudo haplotípico para poder atestar a validade desta suposição.

A distribuição alélica no *locus* da DRPLA na amostra estudada demonstrou uma assimetria negativa semelhante à observada nas populações caucasiana e finlandesa, e diferente da observada na população japonesa (Takano *et al.*, 1998; Juvonen *et al.*, 2005).

CAPÍTULO 5: CONCLUSÕES

No presente trabalho realizou-se a genotipagem dos *loci* *ATXN2*, *CACNA1A* e *ATN1* em indivíduos açorianos saudáveis, dispondo-se agora de uma amostra que irá ser utilizada como grupo controlo num estudo de potenciais genes modificadores da MJD na população açoriana. Para esta mesma amostra confirmou-se molecularmente a inexistência de alelos no intervalo expandido, no *locus* da SCA3, aspecto fundamental para a utilização da mesma na segunda fase da investigação.

Este trabalho veio confirmar que a utilização de um marcador alélico no diagnóstico por análise de fragmentos (ainda não disponível comercialmente para as SCAs) é a melhor estratégia para evitar a ocorrência de erros no número de repetições CAG, evitando assim consequências maiores, como por exemplo, quando estão em questão alelos normais de maior tamanho (sobretudo no caso das SCAs em que a fronteira entre a amplitude normal e patogénica é reduzida ou inexistente) e alelos intermediários. Todavia, quando não existe essa possibilidade, a determinação exacta do tamanho das repetições deverá ser confirmada por sequenciação. Para além disso, todos os alelos que estejam dentro de uma gama de tamanhos em que a presença ou ausência de interrupções possa ser relevante para o fenótipo ou instabilidade do alelo, devem ser sequenciados para determinar o padrão interno das repetições. Este trabalho permitiu ainda verificar a relevância do sequenciador e consequentemente do polímero utilizado na AF, sendo estes aspectos particularmente importantes, quando do diagnóstico e realização de testes preditivos.

Nos *loci* das SCA2 e DRPLA encontraram-se frequências relativamente elevadas dos alelos normais grandes. Não estão reportadas, todavia, as patologias associadas a estes *loci* na população estudada. Outros factores podem estar também envolvidos na dinâmica mutacional destes *loci*, tais como, o padrão interno das repetições CAG e o haplótipo associado a estes alelos, permanecem por estudar. Por outro lado na SCA6 a baixa frequência de alelos normais grandes é coerente com o facto de não estarem reportados casos de SCA6 nos Açores.

Tendo em conta que a prevalência dos diferentes subtipos de SCAs variam entre diferentes populações e que o único diagnóstico possível é o molecular, é importante ter uma estimativa da frequência destas doenças na população para poder adoptar uma estratégia adequada de planeamento de diagnóstico.

Perspectivas futuras:

Para além da utilização dos resultados obtidos no presente trabalho para o estudo de modificadores da MJD, este poderá ter continuidade em vários pontos. Assim, poder-se-á:

- Contribuir para a criação de um marcador alélico usando amostras de genótipo conhecido para as SCAs no laboratório do CIRN, para reduzir a probabilidade de ocorrência de erros na determinação no número de repetições, particularmente em doenças causadas por um pequeno número de repetições;
- Aumentar o tamanho da amostra para poder confirmar a ausência de correlação entre a elevada frequência de alelos normais grandes, no *locus ATXN2*, e ausência de casos reportados de SCA2 nos Açores. E determinar a configuração interna dos alelos normais grandes neste *locus* para conferir se estes têm uma propensão para a instabilidade e conseqüente expansão em sucessivas gerações;
- Fazer um estudo caso-controlo da ALS, dado que:
 - Foi reportado um alelo intermediário na amostra estudada, sendo que estes casos são pouco frequentes em populações controlo;
 - Existem casos reportados de ALS nos Açores;
 - Existe uma associação entre a frequência de alelos normais grandes, no *locus* da SCA2, e um aumento do risco de desenvolver esta doença;
- Realizar um estudo haplotípico, no *locus* da DRPLA, de modo a confirmar qual o haplótipo associado aos alelos normais grandes, e verificar a existência de uma correlação entre os haplótipos e a ausência de casos diagnosticados de DRPLA nos Açores.

CAPÍTULO 6: REFERÊNCIAS BIBLIOGRÁFICAS

- Andrés AM, Lao O, Soldevila M, *et al.* (2003). Dynamics of CAG repeat *loci* revealed by the analysis of their variability. *Human mutation*, **21**(1): 61–70.
- Araújo M (submetida). Geographic profiling of Machado-Joseph disease: a health geography pilot – study in the Azores. Tese de Mestrado, Universidade dos Açores.
- Basu P, Chattopadhyay B, Gangopadhaya PK, *et al.* (2000). Analysis of CAG repeats in SCA1, SCA2, SCA3, SCA6, SCA7 and DRPLA *loci* in spinocerebellar ataxia patients and distribution of CAG repeats at the SCA1, SCA2 and SCA6 *loci* in nine ethnic populations of eastern India. *Human genetics*, **106**(6): 597–604.
- Bauer PO, Zumrova A, Matoskaet V, *et al.* (2005). Absence of spinocerebellar ataxia type 3/Machado-Joseph disease within ataxic patients in the Czech population. *European Journal of Neurology*, **12**(11): 851–857.
- Bettencourt C, Silva-Fernandes A, Montiel R, *et al.* (2007). Triplet Repeats: Features, Dynamics and Evolutionary Mechanisms. In: Santos C & M Lima (Eds.). *Recent Advances in Molecular Biology and Evolution: Applications to Biological Anthropology*. pp. 83–114. Kerala Research Signpost, India.
- Bettencourt C, Fialho RN, Santos C, *et al.* (2008). Segregation distortion of wild-type alleles at the Machado-Joseph disease *locus*: a study in normal families from the Azores islands (Portugal). *Journal of Human Genetics*, **53**(4): 333–339.
- Bettencourt C, Santos C, Montiel R, *et al.* (2010). Increased transcript diversity: novel splicing variants of Machado-Joseph Disease gene (*ATXN3*). *Neurogenetics*, **11**(2): 193–202.
- Bettencourt C & Lima M (2011). Machado-Joseph Disease: From First Descriptions to New Perspectives. *Orphanet Journal of Rare Diseases*, **6**: 35. Doi:10.1186/1750-1172-6-35.
- Bettencourt C, Santos C, Coutinho P, *et al.* (2011). Parkinsonian Phenotype in Machado-Joseph Disease (MJD/SCA3): a Two-case Report. *BMC Neurology*, **11**(1): 131. Doi:10.1186/1471-2377-11-131.
- Brouwer JR, Willemsen R, Oostra B (2009). Microsatellite Repeat Instability and Neurological Disease. *BioEssays*, **31**: 71–83.
- Burnett BG & Pittman RN (2005). The polyglutamine neurodegenerative protein ataxin 3 regulates aggresome formation. *Proceedings of the National Academy of Sciences USA*, **102**(12): 4330–4335.

- Cancel G, Dürr A, Didierjean O, *et al.* (1997). Molecular and clinical correlations in spinocerebellar ataxia 2: a study of 32 families. *Human Molecular Genetics*, **6**: 709–715.
- Carvalho DR, La Rocque-Ferreira A, Rizzo IM, *et al.* (2008). Homozygosity enhances severity in spinocerebellar ataxia type 3. *Pediatric Neurology*, **38**(4): 296–299.
- Castel AL, Cleary JD, Pearson CE (2010). Repeat Instability as the Basis for Human Diseases and as a Potential Target for Therapy. *Nature Reviews Molecular Cell Biology*, **11**: 165–170.
- Chen Y, Huang R, Yang Y, *et al.* (2011). Ataxin-2 Intermediate-length Polyglutamine: a Possible Risk Factor for Chinese Patients with Amyotrophic Lateral Sclerosis. *Neurobiology of Aging*, **32**(10): 1925.e1–5.
- Chou AH, Yeh TH, Ouyang P, *et al.* (2008). Polyglutamine-expanded ataxin-3 causes cerebellar dysfunction of SCA3 transgenic mice by inducing transcriptional dysregulation. *Neurobiology of Disease*, **31**: 89–101.
- Choudhry S, Mukerji M, Srivastava K, *et al.* (2001). CAG repeat instability at SCA2 locus: anchoring CAA interruptions and linked single nucleotide polymorphisms. *Human Molecular Genetics*, **10**(21): 2437–2446.
- Chung MY, Ranum LP, Duvick LA, *et al.* (1993). Evidence for a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type I. *Nature Genetics*, **5**: 254–258.
- Cleary JD & Pearson CE (2003). The contribution of *cis*-elements to disease-associated repeat instability: clinical and experimental evidence. *Cytogenetic and Genome Research*, **100**: 25–55.
- Corrado L, Mazzini L, Oggioni GD, *et al.* (2011). ATXN2 CAG Repeat Expansions Are Interrupted in ALS Patients. *Human Genetics*, **130**(4): 575–580.
- Costa MC & Paulson HL (2011). Toward understanding Machado-Joseph disease. *Progress in Neurobiology*, **97**: 239–257.
- Costanzi-Porrini S, Tessarolo D, Abbruzzese C, *et al.* (2000). An interrupted 34-CAG repeat SCA2 allele in patients with sporadic spinocerebellar ataxia. *Neurology*, **54**: 491–493.
- Coutinho P (1992). Doença de Machado-Joseph: Tentativa de Definição, Tese de Doutorado, Instituto de Ciências Biomédicas Abel Salazar, Porto, 247pp.

- Craig K, Keers SM, Archibald K, *et al.* (2004). Molecular epidemiology of spinocerebellar ataxia type 6. *Annals of Neurology*, **55**: 752–755.
- Daoud H, Belzil V, Martins S, *et al.* (2011). Association of Long ATXN2 CAG Repeat Sizes with Increased Risk of Amyotrophic Lateral Sclerosis. *Archives of Neurology*, **68**(6): 739–742.
- Derek L & Bonini NM (2008). Polyglutamine genes interact to modulate the severity and progression of neurodegeneration in *Drosophila*. *PLoS Biology*, **6**(2): 266–274.
- DeStefano L, Cupples L, Maciel P, *et al.* (1996). A Familial Factor Independent of CAG Repeat Length Influences Age at Onset of Machado-Joseph Disease. *American Journal of Human Genetics*, **59**(1): 119–127.
- Elden AC, Hyung-jun K, Michael PH, *et al.* (2010). Ataxin-2 Intermediate-length Polyglutamine Expansions Are Associated with Increased Risk for ALS. *Nature*, **466** (7310): 1069–1075.
- Ellegren H (1991). DNA typing of museum birds. *Nature*, **354**: 113.
- Ellegren H (2004). Microsatellites: Simple Sequences with Complex Evolution. *Nature Reviews Genetics*, **5**(6): 435–445.
- Everett CM & Wood NW (2004). Trinucleotide repeats and neurodegeneration disease. *Brain*, **127**: 2385–2405.
- Excoffier L, Laval G, Schneider S (2005). Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**: 47–50.
- Fernandez M, McClain ME, Martinez RA, *et al.* (2000). Late-onset SCA2: 33 CAG repeats are sufficient to cause disease. *Neurology*, **55**: 569–572.
- Fletcher CF, Lutz CM, O’Sullivan TN, *et al.* (1996). Absence epilepsy in Tottering mutant mice is associated with calcium channel defects. *Cell*, **87**: 607–617.
- Freund AA, Scola RH, Teive HG, *et al.* (2009). Spinocerebellar ataxias: microsatellite and allele frequency in unaffected and affected individuals. *Arquivos de Neuropsiquiatria*, **67**(4): 1124–1132.
- Friedman MJ, Li SH, Li XJ (2006). Polyglutamine disease. *In*: Fry M & Usdin K (Eds.). *Human Nucleotide Expansion Disorders. Nucleic Acids and Molecular Biology*, Vol. 19. pp. 200-231. Springer-Verlag, Berlin, Heidelberg.
- Furtado S, Farrer M, Tsuboi Y, *et al.* (2002). SCA2 presenting as parkinsonism in an Alberta family: clinical, genetic, and PET findings. *Neurology*, **59**: 1625–1627.

- Futamura N, Matsumura R, Fujimoto Y, *et al.* (1998). CAG repeat expansions in patients with sporadic cerebellar ataxia. *Acta Neurologica Scandinavica*, **98**: 55–59.
- Gaspar C, Lopes-Cendes I, Hayes S, *et al.* (2001). Ancestral Origins of the Machado-Joseph Disease Mutation: a Worldwide Haplotype Study. *American Journal of Human Genetics*, **68**(2): 523–528.
- Gatchel JR & Zoghbi HY (2005). Diseases of Unstable Repeat Expansion: Mechanisms and Common Principles. *Nature Reviews Genetics*, **6**(10): 743–755.
- Geschwind DH, Perlman S, Figueroa CP, *et al.* (1997). The prevalence and wide clinical spectrum of the spinocerebellar ataxia type 2 trinucleotide repeat in patients with autosomal dominant cerebellar ataxia. *American Journal of Human Genetics*, **60**: 842–850.
- Ghosh S & Feany MB (2004). Comparison of pathways controlling toxicity in the eye and brain in *Drosophila* models of human neurodegenerative disease. *Human Molecular Genetics*, **13**: 2011–2018.
- Gispert S, Kurz A, Waibel S, *et al.* (2012). The modulation of Amyotrophic Lateral Sclerosis risk by Ataxin-2 intermediate polyglutamine expansions is a specific effect. *Neurobiology of Disease*, **45**(1): 356–361.
- Gomes-Pereira M & Monckton DG (2006). Chemical Modifiers of Unstable Expanded Simple Sequence Repeats: What Goes up, Could Come Down. *Mutation Research*, **598**(1-2): 15–34.
- Gomez CM, Thompson RM, Gammack JT, *et al.* (1997). Spinocerebellar ataxia type 6: gaze-evoked and vertical nystagmus, Purkinje cell degeneration, and variable age of onset. *Annals of Neurology*, **42**: 933–950.
- Guo SW & Thompson EA (1992). Performing the exact test of Hardy-Weinberg proportions for multiple alleles. *Biometrics*, **48**: 361–372.
- Hancock JM (1999). Microsatellites and other simple sequences: Genomic context and mutational mechanisms. In: Goldstein D & Schlötterer C (Eds.). *Microsatellites: Evolution and Applications*. Oxford University Press, New York.
- Harding AE (1984). *The Hereditary Ataxias and Related Disorders*. Edinburgh. Churchill Livingstone.
- Hayes S, Turecki G, Brisebois K, *et al.* (2000). CAG Repeat Length in RAI1 Is Associated with Age at Onset Variability in Spinocerebellar Ataxia Type 2 (SCA2). *Human Molecular Genetics*, **9**(12): 1753–17538.

- Hernandez A, Magarino C, Gispert S, *et al.* (1995). Genetic mapping of the spinocerebellar ataxia 2 (SCA2) locus on chromosome 12q23-q24.1. *Genomics*, **25**: 433–435.
- Hussey J, Lockhart PJ, Seltzer W, *et al.* (2002). Accurate determination of ataxin-2 polyglutamine expansion in patients with intermediate-range repeats. *Genetic Testing*, **6**: 217–220.
- Huynh DP, Del Bigio MR, Ho DH, *et al.* (1999). Expression of ataxin-2 in brains from normal individuals and patients with Alzheimer's disease and spinocerebellar ataxia 2. *Annals of Neurology*, **45**: 232–241.
- Huynh DP, Yang HT, Vakharia H, *et al.* (2003). Expansion of the polyQ repeat in ataxin-2 alters its Golgi localization, disrupts the Golgi complex and causes cell death. *Human Molecular Genetics*, **12**: 1485–1496.
- Ichikawa Y, Goto J, Hattori M, *et al.* (2001). The genomic structure and expression of *MJD*, the Machado-Joseph disease. *Archives of Neurology*, **46**: 413–422.
- Igarashi S, Takiyama Y, Cancel G, *et al.* (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. *Human Molecular Genetics*, **5**: 923–932.
- Igarashi S, Koide R, Shimohata T, *et al.* (1998). Suppression of aggregate formation and apoptosis by transglutaminase inhibitors in cells expressing truncated DRPLA protein with an expanded polyglutamine stretch. *Nature Genetics*, **18**(2): 111–117.
- Ikeuchi T, Takano H, Koide R, *et al.* (1997). Spinocerebellar ataxia type 6: CAG repeat expansion in $\alpha 1A$ voltage-dependent calcium channel gene and clinical variations in Japanese populations. *Annals of Neurology*, **42**: 879–884.
- Imbert G, Saudou F, Yvert G, *et al.* (1996). Cloning of the Gene for Spinocerebellar Ataxia 2 Reveals a Locus with High Sensitivity to Expanded CAG/glutamine Repeats. *Nature Genetics*, **14**(3): 285–291.
- Jardim LB, Silveira I, Pereira ML, *et al.* (2001). A survey of spinocerebellar ataxia in South Brazil – 66 new cases with Machado-Joseph disease, SCA7, SCA8, or unidentified disease-causing mutations. *Journal of Neurology*, **248**: 870–876.

- Jardim L, Silveira I, Pereira ML, *et al.* (2003). Searching for Modulating Effects of SCA2, SCA6 and DRPLA CAG Tracts on the Machado-Joseph Disease (SCA3) Phenotype. *Acta Neurologica Scandinavica*, **107**(3): 211–214.
- Jeffreys AJ, Turner M, Debenham P (1991). The efficiency of multilocus DNA fingerprint probes for individualization and establishment of family relationships, determined from extensive casework. *American Journal of Human Genetics*, **48**: 824–840.
- Jen JC, Yue Q, Karrim J, *et al.* (1998). Spinocerebellar ataxia type 6 with positional vertigo and acetazolamide responsive episodic ataxia. *Journal of Neurology, Neurosurgery, and Psychiatry*, **65**: 565–568.
- Jurka J & Pethiyagoda C (1995). Simple repetitive DNA sequences from primates: compilation and analysis. *Journal of Molecular Evolution*, **40**: 120–126.
- Juvonen V, Hietala M, Kairisto K, *et al.* (2005). The occurrence of dominant spinocerebellar ataxias among 251 Finnish ataxia patients and the role of predisposing large normal alleles in a genetically isolated population. *Acta Neurologica Scandinavica*, **111**(3): 154–162.
- Kanazawa I (1999). Molecular Pathology of Dentatorubral-pallidoluysian Atrophy. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **354** (1386): 1069–1074.
- Kaneko A, Narabayashi Y, Itokawa K, *et al.* (1997). A Case of Machado-Joseph Disease Presenting with Spastic Paraparesis. *Journal of Neurology, Neurosurgery & Psychiatry*, **62**(5): 542–543.
- Kawaguchi Y, Okamoto T, Taniwaki M, *et al.* (1994). CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. *Nature Genetics*, **8**: 211–228.
- Kim JM, Hong S, Kim GP, *et al.* (2007). Importance of low-range CAG expansion and CAA interruption in SCA2 Parkinsonism. *Archives of Neurology*, **64**: 1510–1518.
- Knight SP, Richardson MM, Osmand AP, *et al.* (1997). Expression and distribution of the dentatorubral-pallidoluysian atrophy gene product (atrophin-1/drplap) in neuronal and non-neuronal tissues. *Journal of the Neurological Sciences*, **146**(1): 19–26.
- Koh JY & Cotman CW (1992). Programmed cell death: its possible contribution to neurotoxicity mediated by calcium channel antagonists. *Brain Research*, **587**(2): 233–240.

- Koide R, Ikeuchi T, Onodera O, *et al.* (1994). Unstable Expansion of CAG Repeat in Hereditary Dentatorubral-pallidoluysian Atrophy (DRPLA). *Nature Genetics*, **6**: 9–13.
- Komeichi K, Sasaki H, Yabe I, *et al.* (2001). Twenty CAG repeats are sufficient to cause the SCA6 phenotype. *Journal of Medical Genetics*, **38**: e38.
- Komure O, Sano A, Nishino N, *et al.* (1995). DNA Analysis in Hereditary Dentatorubral-pallidoluysian Atrophy: Correlation Between CAG Repeat Length and Phenotypic Variation and the Molecular Basis. *Neurology*, **45**: 143–149.
- Kurohara K, Kuroda Y, Maruyama H, *et al.* (1997). Homozygosity for an allele carrying intermediate CAG repeats in the dentatorubral-pallidoluysian atrophy (DRPLA) gene results in spastic paraplegia. *Neurology*, **48**: 1087–1090.
- Laffita-Mesa M, Vela LC, Falco NS, *et al.* (2012). Unexpanded and Intermediate CAG Polymorphisms at the SCA2 Locus (*ATXN2*) in the Cuban Population: Evidence about the Origin of Expanded SCA2 Alleles. *European Journal of Human Genetics*, **20**: 41–49.
- Lander ES, Linton LM, Birren B, *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature*, **409**: 860–892.
- Lee T, Li YR, Chesi A, *et al.* (2011a). Evaluating the prevalence of polyglutamine repeat expansions in amyotrophic lateral sclerosis. *Neurology*, **76**(24): 2062–2065.
- Lee T, Li YR, Ingre C, *et al.* (2011b). Ataxin-2 Intermediate-length Polyglutamine Expansions in European ALS Patients. *Human Molecular Genetics*, **20**(9): 1697–1700.
- Leggo J, Dalton A, Morrison PJ, *et al.* (1997). Analysis of Spinocerebellar Ataxia Types 1, 2, 3 and 6, Dentatorubral-pallidoluysian Atrophy, and Friedreich's Ataxia Genes in Spinocerebellar Ataxia Patients in the UK. *Journal of Human Genetics*, **34**: 982–985.
- Leone M, Bottacchi E, D'Alessandro G, *et al.* (1995). Hereditary ataxias and paraplegias in Valle d'Aosta, Italy: A study of prevalence and disability. *Acta Neurologica Scandinavica*, **91**: 183–187.
- Li SH, McInnis MG, Margolis RL (1993). Novel triplet repeat containing genes in human brain: cloning, expression, and length polymorphisms. *Genomics*, **16**:572–579.

- Li F, Macfarlan T, Pittman RN, *et al.* (2002). Ataxin-3 is a histone-binding protein with two independent transcriptional corepressor activities. *Journal of Biological Chemistry*, **277**(47): 45004–45012.
- Lim J, Hao T, Shaw C, *et al.* (2006). A protein-protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration. *Cell*, **125**: 801–814.
- Lima M (1996). *Doença de Machado-Joseph nos Açores: Estudo epidemiológico, biodemográfico e genético*. Dissertação de Doutoramento, Universidade dos Açores, 299pp.
- Lima M, Mayer FM, Coutinho P, *et al.* (1998). Origins of a mutation: Population Genetics of Machado-Joseph disease in the Azores (Portugal). *Human Biology*, **70**(6): 1011–1023.
- Lima M, Costa MC, Montiel R, *et al.* (2005). Population genetics of wild-type CAG repeats in the Machado-Joseph disease gene in Portugal. *Human heredity*, **60**(3): 156–163.
- Lima M (2010). *Expecções ao Paradigma Mendeliano. A doença de Machado-Joseph como Caso de Estudo*. Síntese de um plano lição para obtenção do título de agregado, Universidade dos Açores, 37pp.
- Lopes-cendes I (1999). *Molecular Genetic Studies of Patients with Autosomal Dominant Spinocerebellares Ataxias*. PhD dissertation. McGill University, Montreal, 276pp.
- Lu CS, Wu Chou YH, Kuo PC, *et al.* (2004). The parkinsonian phenotype of spinocerebellar ataxia type 2. *Archives of Neurology*, **61**: 35–38.
- Maciel P, Gaspar C, DeStefano AL, *et al.* (1995). Correlation between CAG repeat length and clinical features in Machado-Joseph disease. *American Journal of Human Genetics*, **57**(1): 54–61.
- Maciel P, Costa MC, Ferro M, *et al.* (2001). Improvement in the molecular diagnosis of Machado-Joseph disease. *Archives of neurology*, **58**(11): 1821–1827.
- Majounie E, Wardle M, Muzaimia M, *et al.* (2007). Case control analysis of repeat expansion size in ataxia. *Neuroscience letters*, **429**: 28–32.
- Mariotti C, Gellera C, Grisoli M, *et al.* (2001). Pathogenic effect of an intermediate-size SCA-6 allele (CAG)(19) in a homozygous patient. *Neurology*, **57**: 1502–1504.
- Martins S, Matamá T, Guimarães L, *et al.* (2003). Portuguese Families with Dentatorubro-pallidoluysian Atrophy (DRPLA) Share a Common Haplotype of Asian Origin. *European Journal of Human Genetics*, **11**(10): 808–811.

- Martins S, Calafell F, Wong VC, *et al.* (2006). A multistep mutation mechanism drives the evolution of the CAG repeat at MJD/SCA3 locus. *European Journal of Human Genetics*, **14**(8): 932–940.
- Martins S, Calafell F, Gaspar C, *et al.* (2007). Asian Origin for the Worldwide-spread Mutational Event in Machado-Joseph Disease. *Archives of Neurology*, **64**(10): 1502–1508.
- Martins S, Coutinho P, Silveira I, *et al.* (2008). Cis-acting factors promoting the CAG intergenerational instability in Machado-Joseph disease. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, **147B**: 439–446.
- Maruyama H, Nakamura S, Matsuyama Z, *et al.* (1995). Molecular features of the CAG repeats and clinical manifestation of Machado-Joseph disease. *Human Molecular Genetics*, **4**(5): 807–812.
- Matilla-Dueñas A, Sánchez I, Corral-Juan M, *et al.* (2010). Cellular and molecular pathways triggering neurodegeneration in the spinocerebellar ataxias. *Cerebellum*, **9**(2): 148–166.
- Matos C, Macedo-Ribeiro S, Carvalho AL (2011). Polyglutamine Diseases: The Special Case of Ataxin-3 and Machado-Joseph Disease. *Progress in Neurobiology*, **95**: 26–48.
- Matsuyama Z, Kawakami H, Maruyama H, *et al.* (1997). Molecular features of the CAG repeats of spinocerebellar ataxia 6 (SCA6). *Human Molecular Genetics*, **6**(8): 1283–1287.
- Miller SA, Dykes DD, Polesky HF (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, **16**(3): 1215.
- Moseley ML, Benzow KA, Schut LJ, *et al.* (1998). Incidence of dominant spinocerebellar and Friedreich triplet repeats among 361 ataxia families. *Neurology*, **51**: 1666–1671.
- Nagafuchi S, Yanagisawa H, Sato K, *et al.* (1994). Dentatorubral and Pallidolusian Atrophy Expansion of an Unstable CAG Trinucleotide on Chromosome 12p. *Nature Genetics*, **6**: 14–18.
- Naito H & Oyanagi S (1982). Familial myoclonus epilepsy and choreatetosis: hereditary dentatorubral-pallidolusian atrophy. *Neurology*, **32**: 798–807.
- Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**: 583–590.

- Ohta T & Kimura M (1973). A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a genetic population. *Genetical Research*, **22**: 201–204.
- Okazawa H (2003). Polyglutamine diseases: a transcription disorder? *Cellular and Molecular Life Sciences*, **60**: 1427–1439.
- Orr HT & Zoghbi HY (2007). Trinucleotide Repeat Disorders. *Annual Review of Neuroscience*, **30**: 575–621.
- Paulson HL & Fischbeck KH (1996). Trinucleotide repeats in neurogenetic disorders. *Annual Review of Neuroscience*, **19**: 79–107.
- Paulson HL, Das SS, Crino PB, *et al.* (1997a). Machado-Joseph disease gene product is a cytoplasmic protein widely expressed in brain. *Annals of Neurology*, **41**(4): 453–462.
- Paulson HL, Perez MK, Trottier Y, *et al.* (1997b). Intranuclear inclusions of expanded polyglutamine protein in spinocerebellar ataxia type 3. *Neuron*, **19**(2): 333–344.
- Payami H, Nutt J, Ganther S, *et al.* (2003). SCA2 may present as levodopa-responsive parkinsonism. *Movement Disorders*, **18**: 425–429.
- Pearson CE & Sinden RR (1998). Trinucleotide repeat DNA structures: dynamic mutations from dynamic DNA. *Current Opinion Structural Biology*, **8**(3): 321–330.
- Pearson CE, Eichler EE, Lorenzetti D, *et al.* (1998). Interruptions in the triplet repeats of SCA1 and FRAXA reduce the propensity and complexity of slipped strand DNA (S-DNA) formation. *Biochemistry*, **37**: 2701–2708.
- Pearson CE, Edamura NK, Cleary JD (2005). Repeat instability: mechanisms of dynamic mutations. *Nature Reviews Genetics*, **6**: 729–742.
- Pietrobon D (2002). Calcium channels and channelopathies of the central nervous system. *Molecular Neurobiology*, **25**: 31–50.
- Pulst SM, Nechiporuk A, Nechiporuk T, *et al.* (1996). Moderate Expansion of a Normally Biallelic Trinucleotide Repeat in Spinocerebellar Ataxia Type 2. *Nature Genetics*, **14**: 269–275.
- Pulst SM, Santos N, Wang D, *et al.* (2005). Spinocerebellar ataxia type 2: polyQ repeat variation in the CACNA1A calcium channel modifies age of onset. *Brain*, **128**: 2297–2303.

- Ramos EM, Martins S, Alonso I, *et al.* (2010). Common Origin of Pure and Interrupted Repeat Expansions in Spinocerebellar Ataxia Type 2 (SCA2). *American Journal of Medical Genetics Part B*, **153B**: 524–531.
- Restituto S, Tompson RM, Eliet J, *et al.* (2000). The polyglutamine expansion in spinocerebellar ataxia type 6 causes a beta subunit-specific enhanced activation of P/Q-type calcium channels in *Xenopus oocytes*. *Journal of Neuroscience*, **20**: 6394–6403.
- Rettig J, Sheng ZH, Kim DK, *et al.* (1996). Isoform-specific interaction of the alpha1A subunits of brain Ca²⁺ channels with the presynaptic proteins syntaxin and SNAP-25. *Proceedings of the National Academy of Sciences USA*, **93**(14): 7363–7368.
- Richards RI & Sutherland GR (1992). Dynamic mutations: a new class of mutations causing human disease. *Cell*, **70**: 709–712.
- Ross O, Rutherford NJ, Baker M, *et al.* (2011). Ataxin-2 Repeat-length Variation and Neurodegeneration. *Human Molecular Genetics*, **20**(16): 3207–3212.
- Sahba S, Nechiporuk A, Figueroa KP, *et al.* (1998). Genomic structure of the human gene for spinocerebellar ataxia type 2 (SCA2) on chromosome 12q24.1. *Genomics*, **47**: 359–364.
- Sakai T & Kawakami H (1996). Machado-Joseph disease: A proposal of spastic paraplegic subtype. *Neurology*, **46**(3): 846–847.
- Saleem Q, Choudhry S, Mukerji M, *et al.* (2000). Molecular Analysis of Autosomal Dominant Hereditary Ataxias in the Indian Population: High Frequency of SCA2 and Evidence for a Common Founder Mutation. *Human Genetics*, **106**(2): 179–187.
- Sanpei K, Takano H, Igarashi S, *et al.* (1996). Identification of the Spinocerebellar Ataxia Type 2 Gene Using a Direct Identification of Repeat Expansion and Cloning Technique, DIRECT. *Nature Genetics*, **14**: 277–284.
- Santos N, Aguiar J, Fernandez J, *et al.* (1999). Molecular diagnosis of a sample of the Cuban population with spinocerebellar ataxia type 2. *Biotechnologia Aplicada*, **16**: 219–221.
- Satterfield TF & Leo Pallanck J (2006). Ataxin-2 and its *Drosophila* homolog, ATX2, physically assemble with polyribosomes. *Human Molecular Genetics*, **15**(16): 2523–2532.

- Schelhaas HJ, Ippel PF, Beemer FA, *et al.* (2000). Similarities and differences in the phenotype, genotype, and pathogenesis of different spinocerebellar ataxias. *European Journal of Neurology*, **7**: 309–314.
- Schöls L, Gispert S, Vorgerd M (1997). Spinocerebellar Ataxia Type 2 Genotype and Phenotype in German Kindreds. *Archives of Neurology*, **54**: 1073–1080.
- Schöls L, Szymanski S, Peters S, *et al.* (2000). Genetic background of apparently idiopathic sporadic cerebellar ataxia. *Human Genetics*, **107**: 132–137.
- Schöls L, Bauer P, Schmidt T, *et al.* (2004). Review Autosomal Dominant Cerebellar Ataxias: Clinical Features, Genetics, and Pathogenesis. *The Lancet*, **3**: 291–304.
- Schlötterer C, Amos W, Tautz D (1991). Conservation of polymorphic simple sequence *loci* in cetacean species. *Nature*, **354**: 63–65.
- Schlötterer C (2000). Evolutionary dynamics of microsatellite DNA. *Chromosoma*, **109**(6): 365–371.
- Sequeiros J (1989). Análise genética da variação fenotípica na doença de Machado-Joseph. Tese de Doutorado. Instituto de Ciências Biomédicas de Abel Salazar, Universidade do Porto, 236pp.
- Sequeiros J, Coutinho P (1993). Epidemiology and clinical aspects of Machado-Joseph disease. *Advances in neurology*, **61**: 139–153.
- Sequeiros J, Martindale J, Seneca S, *et al.* (2010). EMQN Best Practice Guidelines for Molecular Genetic Testing of SCAs. *European Journal of Human Genetics*, **18**(11): 1173–1176.
- Sequeiros J, Martins S, Silveira I (2012). Epidemiology and Population Genetics of Degenerative Ataxias. *In*: Vinken PJ & Bruyn GW (Eds.). *Ataxic Disorders. Handbook of Clinical Neurology*, Vol. 103 (3rd series), pp. 227–251. Elsevier.
- Silva MC, Coutinho P, Pinheiro CD, *et al.* (1997). Hereditary ataxias and spastic paraplegias: methodological aspects of a prevalence study in Portugal. *Journal of Clinical Epidemiology*, **50**: 1377–1384.
- Silveira I, Coutinho P, Maciel P, *et al.* (1998). Analysis of SCA1, DRPLA, MJD, SCA2, and SCA6 CAG Repeats in 48 Portuguese Ataxia Families. *American Journal of Medical Genetics*, **81**: 134–138.
- Silveira I, Miranda C, Guimarães BS, *et al.* (2002). Trinucleotide Repeats in 202 Families with Ataxia - A Small Expanded (CAG)*n* Allele at the SCA17 *locus*. *Archives of neurology*, **59**: 623–629.

- Sobczak K & Krzyzosiak WJ (2004). Patterns of CAG repeat interruptions in SCA1 and SCA2 genes in relation to repeat instability. *Human mutation*, **24**(3): 236–247.
- Soong BW & Paulson HL (2007). Spinocerebellar ataxias: an update. *Current Opinion in Neurology*, **20**(4): 438–446.
- SPSS Inc. SPSS (2009) for windows-release 18.0.0. Chicago, 1993–2007.
- Stephens K, Tait JF, Jacobs DS, *et al.* (2001). Molecular Genetic Testing. Laboratory Test Handbook. *In*: Jacobs DS, DeMott WR, Oxley DK (Eds.). pp 701–718. Cleveland, Lexi-Comp Inc.
- Stevanin G, Dürr A, Brice A (2000). Clinical and molecular advances in autosomal-dominant cerebellar ataxias: from genotype to phenotype and physiopathology. *European Journal of Human Genetics*, **8**: 4–18.
- Strachan T & Read AP (1999). Human Molecular Genetics, 2nd edition. Wiley-Liss, New York.
- Subramanian S, Madgula VM, George R, *et al.* (2003). Triplet repeats in human genome: distribution and their association with genes and other genomic regions. *Bioinformatics*, **19**: 549–552.
- Takahashi H, Ohama E, Naito H, *et al.* (1988). Hereditary dentato-rubral-pallidoluysian atrophy: clinical and pathologic variants in a family. *Neurology*, **38**: 1065-1070.
- Takano H, Cancel G, Ikeuchi T, *et al.* (1998). Close associations between prevalences of dominantly inherited spinocerebellar ataxias with CAG-repeat expansions and frequencies of large normal CAG alleles in Japanese and Caucasian populations. *American journal of human genetics*, **63**(4): 1060–1066.
- Takiyama Y, Nishizawa M, Tanaka H, *et al.* (1993). The gene for Machado-Joseph disease maps to human chromosome 14q. *Nature Genetics*, **4**(3): 300–304.
- Takiyama Y, Igarashi S, Rogaeva EA, *et al.* (1995). Evidence for inter-generational instability in the CAG repeat in the *MJD1* gene and for conserved haplotypes at flanking markers amongst Japanese and Caucasian subjects with Machado-Joseph disease. *Human Molecular Genetics*, **4**(7): 1137–1146.
- Tan NC, Zhou Y, Tan AS, *et al.* (2004). Spinocerebellar ataxia type 2 with focal epilepsy - an unusual association. *Annals, Academy of Medicine, Singapore*, **33**: 103–106.
- Trott A, Jardim LB, Ludwig HT, *et al.* (2006). Spinocerebellar ataxias in 114 Brazilian families: clinical and molecular findings. *Clinical Genetics*, **70**: 173–176.

- Tsuji S (1997). Molecular Genetics of Triplet Repeats: Unstable Expansion of Triplet Repeats as a New Mechanism for Neurodegenerative Diseases. *Internal Medicine*, **36**(1): 3–8.
- Tuite PJ, Rogaeva EA, George-Hyslop PHST *et al.* (1995). Dopa-responsive parkinsonism phenotype of Machado-Joseph disease: confirmation of 14q CAG expansion. *Annals of Neurology*, **38**: 684–687.
- Uchihara T, Fujigasaki H, Koyano S, *et al.* (2001). Non-expanded polyglutamine proteins in intranuclear inclusions of hereditary ataxias - triple-labeling immune fluorescence study. *Acta Neuropathologica*, **102**: 149–152.
- Uversky VN (2009). Intrinsic Disorder in Proteins Associated with Neurodegenerative Diseases. In: Ovádi J & Orosz F (Eds.). *Protein Folding and Misfolding: Neurodegenerative Diseases*. Focus on Structural Biology, Vol. 7. pp. 21–75. Springer.
- Vale J, Bugalho P, Silveira I, *et al.* (2010). Autosomal Dominant Cerebellar Ataxia: Frequency Analysis and Clinical Characterization of 45 Families from Portugal. *European Journal of Neurology*, **17**(1): 124–128.
- Van Damme P, Veldink JH, van Blitterswijk M, *et al.* (2011). Expanded ATXN2 CAG Repeat Size in ALS Identifies Genetic Overlap Between ALS and SCA2. *Neurology*, **76**(24): 2066–2072.
- van de Warrenburg BP, Sinke RJ, Verschuuren-Bemelmans CC, *et al.* (2002). Spinocerebellar ataxias in the Netherlands: Prevalence and age at onset variance analysis. *Neurology*, **58**: 702–708.
- van de Warrenburg BP, Hendriks H, Dürr A, *et al.* (2005). Age at onset variance analysis in spinocerebellar ataxias: a study in a Dutch-French cohort. *Annals of Neurology*, **57**(4): 505–512.
- Velázquez-Pérez L, Cruz GS, Falcón NS, *et al.* (2009). Molecular Epidemiology of Spinocerebellar Ataxias in Cuba: Insights into SCA2 Founder Effect in Holguin. *Neuroscience Letters*, **454**(2): 157–60.
- Velázquez-Pérez L, Rodríguez-Labrada R, Cesar García-Rodríguez J, *et al.* (2011). A Comprehensive Review of Spinocerebellar Ataxia Type 2 in Cuba. *Cerebellum*, **10** (2): 184–198.
- Walker D & DeWaard M (1998). Subunit interaction sites in voltage-dependent Ca²⁺ channels: role in channel function. *Trends in Neurosciences*, **21**: 148–54.

- Weber JL & Wong C (1993). Mutation of Human Short Tandem Repeats. *Human Molecular Genetics*, **2**(8): 1123–1128.
- Zhang S, Xu L, Lee J, *et al.* (2002). *Drosophila* atrophin homolog functions as a transcriptional corepressor in multiple developmental processes. *Cell*, **108**: 45–56.
- Zhuchenko O, Bailey J, Bonnen P, *et al.* (1997). Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the α_{1A} -voltage-dependent calcium channel. *Nature Genetics*, **15**: 62–69.
- Zoghbi HY & Orr HT (2000). Glutamine Repeats and Neurodegeneration. *Annual Review of Neuroscience*, **23**: 217–247.
- Yamada M, Tan CF, Inenaga C, *et al.* (2004). Sharing of polyglutamine localization by the neuronal nucleus and cytoplasm in CAG-repeat diseases. *Neuropathology and Applied Neurobiology*, **30**(6): 665–675.
- Yamada M, Sato T, Tsuji S, *et al.* (2008). CAG Repeat Disorder Models and Human Neuropathology: Similarities and Differences. *Acta Neuropathologica*, **115**(1): 71–86.
- Yanagisawa H, Fujii K, Nagafuchi S, *et al.* (1996). A unique origin and multistep process for the generation of expanded DRPLA triplet repeats. *Human Molecular Genetics*, **5**: 373–379.
- Yazawa I, Nukina N, Hashida H, *et al.* (1995). Abnormal gene product identified in hereditary dentatorubral-pallidoluysian atrophy (DRPLA) brain. *Nature Genetics*, **10**: 99–103.
- Yu Z, Zhu Y, Chen-Plotkin AS, *et al.* (2011). PolyQ Repeat Expansions in *ATXN2* Associated with ALS Are CAA Interrupted Repeats. *PLoS One*, **6**(3): e17951.