

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

**AVALIAÇÃO DE FATORES GENÉTICOS COMO  
POTENCIAIS MODIFICADORES DA  
IDADE DE INÍCIO DA DOENÇA DE HUNTINGTON**

Aline Francielle Damo Souza

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

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Maria Luiza Saraiva Pereira

Porto Alegre, Maio de 2013

## **INSTITUIÇÕES E FONTES FINANCIADORAS**

Este trabalho foi realizado no Laboratório de Identificação Genética (IN.GENE) localizado no Centro de Pesquisa Experimental do Hospital de Clínicas de Porto Alegre em parceria com o Serviço de Genética Médica do mesmo hospital, contando com apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), do Fundo de Incentivo à Pesquisa e Eventos – Hospital de Clínicas de Porto Alegre e da Fundação de Amparo à Pesquisa do Rio Grande do Sul (FAPERGS). Durante o período de realização do trabalho a aluna contou com uma bolsa de estudos concedida pela Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES).

*“Tudo quanto te vier à mão para fazer, faze-o conforme as tuas forças,  
porque para onde todos nós um dia iremos, não há obra nem projeto,  
nem conhecimento, nem sabedoria alguma” - Adaptado de Eclesiastes 9:10*

## AGRADECIMENTOS

À minha orientadora, Prof<sup>a</sup>. Dr<sup>a</sup>. Maria Luiza Saraiva Pereira, pela oportunidade de fazer parte do grupo seletivo de pesquisadores do LIG. Obrigada por toda a paciência e por todo o suporte pessoal e científico durante este tempo de mestrado. Os “puxões de orelha” certamente contribuíram para o meu crescimento pessoal e científico.

Aos estatísticos do Hospital de Clínicas de Porto Alegre, especialmente à estatística Vanessa Bielefeldt Torman, pela assistência e paciência nas muitas consultorias realizadas.

A esta universidade pela oportunidade de ensino, pelo incentivo à educação e principalmente à pesquisa, por todo o suporte de infraestrutura e por toda a contribuição para minha formação pessoal, intelectual, científica e profissional durante o tempo do mestrado. Aos órgãos de fomento pelo financiamento desta pesquisa: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundo de Incentivo à Pesquisa e Eventos (FIPE) - Hospital de Clínicas de Porto Alegre (HCPA) e Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES).

Aos colegas do LIG: Marina, Márcia, Fernanda, Hugo, Gabriel, Tailise, Édina, Eduarda, Júlia obrigada pela acolhida e por todas as contribuições. Certamente elas e vocês foram essenciais para que este trabalho pudesse ser realizado.

A Deus, meu amigo mais fiel, pelo dom da Vida, pelo amor incondicional, por se importar com cada detalhe da minha vida. Se hoje cheguei até aqui foi porque a Tua mão me sustentou! Aos meus pais, Maura e Luiz, por terem me gerado, me amado, zelado pela minha vida durante todos os meus dias. Mãe, obrigada por todo o amor, carinho, dedicação, suporte, por ter estado sempre presente em todos os momentos, especialmente naqueles em que eu não tinha estrutura para suportar tudo sozinha. À minha avó, Adelinda (*in memoriam*), agradeço por todo o cuidado e carinho que sempre teve e por cada ensinamento que hoje e sempre carregarei comigo. Dindo e Dinda, vocês são meus segundos pai e mãe. Obrigada por todo o cuidado que tiveram e tem conosco. Agradeço a todos os meus familiares que sempre me apoiaram e torceram pra que eu pudesse chegar até aqui e à segunda família que Deus me deu por todo o amor e suporte.

## SUMÁRIO

PRINCIPAIS ABREVIATURAS.....	6
RESUMO.....	7
ABSTRACT.....	9
1. INTRODUÇÃO.....	9
1.1 Doença de Huntington .....	9
1.1.1 Manifestações Clínicas .....	10
1.1.2 Prevalência da DH .....	12
1.1.3 Identificação do Gene .....	13
1.1.4 Proteína huntingtina .....	14
1.2 Modificadores genéticos .....	16
1.2.1 Estudos de <i>Genome Wide Association</i> .....	16
1.2.2 Genes modificadores na DH.....	18
1.2.2.1 <i>GRM1 (metabotropic glutamate receptor 1)</i> .....	19
1.2.2.2 <i>SGK1 (serum/glucocorticoid regulated kinase 1)</i> .....	23
1.2.2.3 <i>TGM2 (C polypeptide, protein-glutamine-gamma-glutamyltransferase)</i> .....	24
1.2.2.4 <i>HAP1 (huntingtin associated protein 1)</i> .....	28
1.2.2.5 Ataxias espinocerebelares.....	29
2. JUSTIFICATIVA .....	33
3. OBJETIVOS .....	34
3.1 Objetivo Geral .....	34
3.2 Objetivos Específicos .....	34
4. RESULTADOS .....	35
4.1 Artigo 1 .....	36
4.2 Artigo 2.....	62
5. DISCUSSÃO .....	87
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	95

## PRINCIPAIS ABREVIATURAS

A	adenina
ATXN1	ataxina 1 ( <i>ataxin 1</i> )
ATXN2	ataxina 2 ( <i>ataxin 2</i> )
ATXN3	ataxina 3 ( <i>ataxin 3</i> )
ATXN7	ataxina 7 ( <i>ataxin 7</i> )
C	citossina
CACNA1A	subunidade 1 $\alpha$ do canal de cálcio voltagem dependente do tipo P/Q ( <i>calcium channel, voltage dependent, P/Q type, <math>\alpha</math>1A subunit</i> )
DH	doença de Huntington
DNA	ácido desoxirribonucleico ( <i>desoxyribonucleic acid</i> )
DRPLA	atrofia dentatorubral-palidoluisiana ( <i>Dentatorubral-pallidoluisian atrophy</i> )
EHW	equilíbrio de Hardy-Weinberg
G	guanina
GMR1	receptor metabotrópico de glutamato 1 ( <i>metabotropic glutamate receptor 1</i> )
GWAS	estudos de associação do genoma ( <i>genome-wide association studies</i> )
HAP1	proteína associada à huntingtina ( <i>huntingtin associated protein 1</i> )
HTT	gene da huntingtina ( <i>HTT gene</i> )
PPP2R2B	subunidade regulatória B da proteína fosfatase 2, $\beta$ ( <i>Protein phosphatase 2, regulatory subunit B, <math>\beta</math></i> )
SBMA	atrofia muscular espinobulbar ( <i>spinal-bulbar muscular atrophy</i> )
SCA	ataxia espinocerebelar ( <i>spinocerebellar ataxia</i> )
SGK1	cinase regulada por soro e glicocorticoide ( <i>serum/glucocorticoid regulated kinase 1</i> )
SNP	polimorfismo de base única ( <i>single nucleotide polymorphism</i> )
TBP	proteína ligadora de TATA-box ( <i>TATA-box binding protein</i> )
TGM2	transglutaminase 2 ( <i>C polypeptide, protein-glutamine-gamma-glutamyltransferase</i> )

## RESUMO

A doença de Huntington (DH) é o distúrbio mais comum entre as doenças neurodegenerativas de padrão de herança autossômica dominante com idade de início na vida adulta. A DH é causada por uma expansão das repetições trinucleotídicas CAG no gene *HTT* e caracteriza-se por movimentos involuntários, distúrbios emocionais severos e déficit cognitivo. A idade de início dos sintomas é parcialmente determinada pelo número de repetições CAG, sendo o restante influenciado por fatores genéticos e ambientais.

As variantes polimórficas rs1331639 e rs6923492 no gene *GRM1*, rs1763505 e rs1075427 no gene *SGK*, rs7270898 e rs7275079 no gene *TGM2*, rs4523977 no gene *HAP1*, número de repetições CAG no alelo normal no gene *HTT* e nos alelos normais dos genes relacionados às ataxias espinocerebelares SCA1, SCA2, SCA3, SCA6, SCA7, SCA12, SCA17 e DRPLA foram escolhidas como genes candidatos a modificador, avaliando a associação com a idade de início dos sintomas em pacientes com DH.

Na população em estudo, o número de repetições CAG no alelo mutante do gene *HTT* explicou ~65% da variação da idade de início. Os SNPs rs6923492 (*GRM1*) e rs1075427 (*SGK1*) não estavam em EHW na população testada ( $\chi^2=7,586$ ;  $p=0,023$  e  $\chi^2=7,239$ ;  $p=0,027$ , respectivamente) e por isso foram excluídos das análises posteriores. Entre as variantes polimórficas restantes, associação moderada foi encontrada com a variação da idade de início para o genótipo AA (M441) do rs4523977 no gene *HAP1* ( $p=0,044$ ) e para o número de repetições CAG presentes no alelo longo do gene *ATXN1* ( $p=0,045$ ). Este último fator explicou 4,5% do restante da variação da idade de início não atribuída às repetições CAG de *HTT*. Os alelos com 24 repetições do gene *ATXN3*, 9 repetições do gene *ATXN7* e 11 repetições do gene *PPP2R2B* foram mais frequentemente encontradas no grupo de pacientes quando comparado ao grupo controle ( $p=0,003$ ;  $p<0,0001$  e  $p<0,0001$ , respectivamente). O alelo com 8 repetições do gene *PPP2R2B* foi encontrado com uma frequência de 43% exclusivamente no grupo de pacientes. Estes alelos podem estar relacionados com a patogênese da DH.

Os resultados aqui encontrados são de grande relevância para uma melhor elucidação de fatores genéticos atuantes na modulação da idade de início na DH. Tais resultados devem ser mais aprofundados e necessitam de outros estudos que corroborem os achados deste trabalho.

## ABSTRACT

Huntington's disease (HD) is the most common autosomal dominantly inherited neurodegenerative disorder with age at onset (AO) in adult life. HD is characterized by motor dysfunction, cognitive deficits, and psychiatric disturbances. An expansion of CAG trinucleotide repeats in the first exon of *HTT* gene is the cause of HD. Mean AO is around 40 years and is mainly determined by CAG repeats length of the mutant *HTT* allele. The remaining variation observed in HD patient is probably due to genetic and/or environmental factors.

Polymorphic variants rs1331639 and rs6923492 in *GRM1* gene, rs1763505 and rs1075427 in *SGK* gene, rs7270898 and rs7275079 in *TGM2* gene, rs4523977 in *HAPI* gene, CAG repeat length in normal allele of *HTT* gene and normal alleles of *ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7*, *PPP2R2B*, *TBP*, and *DRPLA* gene were chosen as candidates to modifiers and a potential association to AO was analyzed in patients with HD.

In our population, CAG repeat length in mutant *HTT* allele explained ~65% of variation in AO. A modest association with AO variation in genotype AA (M441) of rs4523977 in *HAPI* gene ( $p=0.044$ ), and of CAG repeat length in long allele of *ATXN1* gene ( $p=0.045$ ). The latter explained an additional 4.5% of the AO besides influence of CAG repeats length of mutant allele in *HTT* gene. Alleles with 24 repeats in *ATXN3* gene, 9 repeats in *ATXN7* gene, and 11 repeats in *PPP2R2B* genes were significantly more frequent in case group when compared to control group ( $p=0.003$ ;  $p<0.0001$  and  $p<0.0001$ , respectively). Allele with 8 repeats in *PPP2R2B* gene was found in a frequency of 43% in case group only. These alleles could be related with HD pathogenesis.

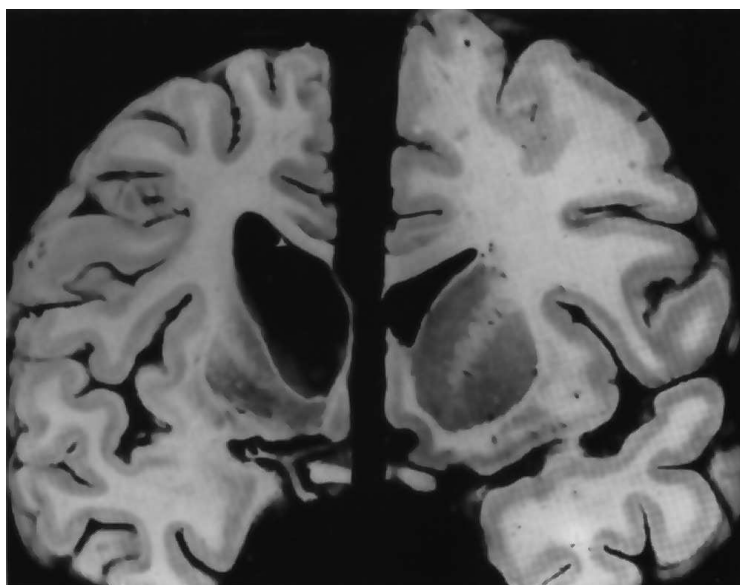
In summary, results presented here are relevant for a better understanding of genetic factors that act in AO modulation of HD. Such results might be further explored and additional studies are required to corroborate data presented here.



# 1. INTRODUÇÃO

## 1.1 Doença de Huntington

A doença de Huntington (DH) (OMIM 143100) é a doença neurodegenerativa de herança autossômica dominante mais comum com início das manifestações clínicas na vida adulta. A idade média de início é de 40 anos, sendo que já foram descritos casos precoces, com idade de início inferior a 20 anos, e tardios, acima dos 60 anos. Um padrão seletivo de atrofia cerebral é observado nos pacientes com a doença, havendo perda neuronal e gliose na região neostriatal. Nesta região, que compreende os núcleos putâmen e caudado, a atrofia se estabelece precocemente, sendo que o núcleo caudado é o preferencialmente afetado. No entanto, com a progressão da doença, a seletividade é perdida e todo o cérebro é acometido pela atrofia, resultando em uma perda de massa cerebral que pode chegar a 200 ou 300 gramas (Figura 1) (Hayden, 2001). Até o momento não existe tratamento clínico capaz de conter a progressão da doença. Este fato reflete na morte dos doentes após aproximadamente 18 anos da idade de início da manifestação dos sintomas motores e cognitivos.



**Figura 1.** Secção coronal cerebral. Atrofia estriatal severa em um indivíduo com DH (esquerda) em comparação com um indivíduo normal (direita). (Fonte: Huntington Disease in The metabolic and Molecular Basis of Inherited Disease. 8th edition. 2001).

### 1.1.1 Manifestações Clínicas

A DH é caracterizada clinicamente por um quadro progressivo de anormalidades motoras, especialmente associadas com o sistema extrapiramidal, alterações cognitivas e disfunções psiquiátricas (Hayden, 2001). As anormalidades motoras na DH podem ser segregadas em anormalidades motoras voluntárias e involuntárias. Entre os movimentos involuntários que marcam o início da doença encontram-se: coreia, hipercinesia, inquietação, movimento anormal dos olhos, hiperreflexia, movimentos excessivos e inadequados com os dedos das mãos e dos pés durante estresse emocional. Coreia, do grego “choreia”, significa dançar, e refere-se aos movimentos rápidos, irregulares e bruscos dos membros, tronco e rosto desencadeados pela doença. A amplitude da coreia e sua gravidade podem variar muito de indivíduo para outro. Os movimentos coreicos são

bastante frequentes no início da doença, e geralmente se agravam com a progressão da doença. Em fases mais avançadas da doença, os movimentos tendem a ser reduzidos, manifestando-se por hipocinesia, rigidez e distonia acentuados. Os prejuízos nos movimentos voluntários são da mesma forma progressivos. A mobilidade vai sendo reduzida à medida que a doença avança e as quedas são bastante frequentes. Os déficits no controle motor fino podem ser primeiramente notados como uma deficiência leve de escrita. Ao final da doença o indivíduo torna-se incapaz de movimentar-se sozinho e possui um controle volitivo mínimo sobre os movimentos dos membros. Disartria e disfagia refletem tanto o prejuízo do controle, como a piora dos movimentos involuntários da boca, língua e lábios. A disfunção progride para uma total incapacidade de tolerar a nutrição oral (Roos, 2010).

Em relação aos sintomas cognitivos, todos os pacientes com DH apresentam um declínio global na capacidade cognitiva, manifestando, inicialmente, um padrão de lentidão no pensamento, alterações na personalidade, mudanças afetivas e falha na habilidade de adquirir novos conhecimentos. Alguns estudos identificaram a presença de alterações cognitivas sutis, no entanto bem definidas, antes do início das manifestações motoras (Tabrizi *et al*, 2009; Rupp *et al*, 2010). As mudanças iniciais frequentemente envolvem perda da flexibilidade mental e falha na execução de funções como planejamento e organização de atividades sequenciais.

Embora distúrbios psiquiátricos sejam tão característicos quanto os distúrbios motores e cognitivos, eles aparecem com menos frequência e não estão necessariamente relacionados à gravidade das anormalidades motoras, mudanças cognitivas ou mesmo ao tamanho da mutação genética causadora da DH (Hayden, 2001). Os sintomas mais

frequentemente relatados são: mudanças emocionais e de humor, ansiedade, irritabilidade, apatia, sintomas obsessivos e compulsivos e psicose. O risco de suicídio é bastante aumentado nesses indivíduos, especialmente diante da incerteza do resultado do teste molecular para diagnóstico da DH, com o início dos sintomas e quando a dependência para tarefas diárias aumenta (van Duijn *et al*, 2007).

### **1.1.2 Prevalência da DH**

Pesquisas realizadas em diferentes partes do mundo, incluindo o noroeste europeu, o Reino Unido, a Escandinávia, a América do Norte, a Austrália, o Japão e a África do Sul, chegaram a um consenso geral sobre a frequência da DH em populações euro-descendentes. A prevalência estimada foi de 3 a 7 casos a cada 100.000 indivíduos.

A frequência da DH entre os habitantes da costa oeste do lago Maracaibo, no Estado de Zulia, Venezuela, é a maior em todo o mundo, com o acometimento de 14.000 indivíduos pela DH (*National Institute of Neurological Disorders and Stroke – NIH*, 2013), enquanto que em outras regiões da Venezuela, a prevalência estimada é de um caso a cada 200.000 habitantes (Paradisi *et al*, 2008).

Estudos genealógicos indicam que uma migração inicial do noroeste europeu foi o fator primário responsável pela dispersão do gene da doença ao redor do mundo (Hayden, 2001).

### 1.1.3 Identificação do Gene

A primeira descrição mais precisa e cuidadosa a respeito desta doença foi realizada pelo médico norte-americano George Huntington em 1872 (“*On Chorea*” publicado no *Medical and Surgical Reporter of Philadelphia* - volume 26, nº. 15, de abril de 1872). Desde então, muitos avanços foram descritos na área. O mapeamento do gene associado à DH foi realizado em 1983 (Gusella *et al*, 1983). Entretanto sua identificação só foi realizada dez anos depois, em 1993 (MacDonald *et al*, 1993).

O gene associado à DH, conhecido como *IT15* ou *HTT* (gene da huntingtina), está localizado no braço curto do cromossomo 4, *locus* 4p16.3, compreende 200 Kb e está dividido em 67 éxons. O primeiro éxon apresenta uma sequência repetitiva dos nucleotídeos CAG, a qual codifica uma sequência de poliglutaminas (poli-Q) na proteína. Essas repetições são polimórficas, sendo que, em indivíduos normais, o número de repetições CAG varia de 6 a 35. A DH é causada por 36 ou mais repetições trinucleotídicas; no entanto, as manifestações clínicas características ocorrem em indivíduos com número de repetições acima de 40 (c.52CAG(>40)) (MacDonald *et al*, 1993). A extensão de 36 a 39 repetições (c.52CAG(36\_39)) leva a uma penetrância incompleta da doença ou à manifestação bastante tardia dos sintomas. Repetições CAG entre 27 e 35 (c.52CAG(27\_35)) são instáveis, sendo chamados de alelos intermediários, e estes alelos estão propensos a modificações durante a reprodução (MacDonald *et al*, 1993).

O tamanho das repetições CAG expandidas se constitui o fator mais crítico na determinação da idade de início da manifestação dos sintomas característicos da DH. Essa correlação se estabelece de maneira inversa, ou seja, quanto maior o número de repetições CAG, mais precocemente o indivíduo manifesta os sintomas da doença. O auge da idade

de início varia de 35 a 50 anos, sendo que 10% dos indivíduos doentes desenvolvem os sintomas antes dos 20 anos, sendo essa forma conhecida como DH juvenil. Nesses indivíduos, o número de repetições CAG ultrapassa as 60 (MacDonald *et al*, 1993).

#### **1.1.4 Proteína huntingtina**

O gene *HTT* codifica uma proteína de 348 kDa denominada huntingtina (htt), a qual é largamente expressa nos diferentes tecidos humanos, com maior expressão no tecido cerebral. A htt apresenta uma extensão característica de poliglutaminas (poli-Q) na região N-terminal, a qual é codificada a partir das repetições trinucleotídicas CAG presentes no gene *HTT*. Esta sequência de poli-Q inicia no 18<sup>o</sup> aminoácido e, em indivíduos normais, apresenta de 6 a 35 resíduos de glutamina.

A presença da expansão trinucleotídica resulta em uma proteína mutante com um trecho anormalmente alongado de glutamina (MacDonald *et al*, 1993). Atribui-se a essa htt mutante a causa de disfunção, neurodegeneração e das características clínicas primárias em pacientes com a doença por meio de um ganho patológico de funções tóxicas (Landles and Bates, 2004). A determinação do mecanismo preciso de toxicidade mediada por poli-Q é motivo de grandes investigações e existem evidências de que interações aberrantes proteína-proteína estariam presentes na patogênese da doença.

Outras doenças neurodegenerativas progressivas de manifestação tardia também são causadas por uma expansão das repetições trinucleotídicas CAG na região codificante do gene associado à doença. Entre elas, podemos citar a atrofia muscular espinobulbar (SBMA), a atrofia dentatorubral-palidoluisiana (DRPLA) e as ataxias espinocerebelares 1,

2, 3, 6, 7 e 17. Essas doenças de poliglutaminas apresentam como característica comum a presença de uma proteína mutante com um trecho alongado de poliglutaminas que, por um ganho de função tóxica, leva à disfunção neuronal que geralmente ocorre na vida adulta. O fato de os indivíduos já nascerem com a mutação e expressá-la desde o nascimento implica na existência de mecanismos compensatórios que, até certo momento da vida, conseguem neutralizar o efeito tóxico da proteína. Tais mecanismos ainda não foram bem elucidados (Takahashi *et al*, 2010).

As doenças de poliglutaminas apresentam um biomarcador patológico comum, a presença de agregados proteicos no citoplasma e/ou no núcleo das células neuronais. Por muito tempo suspeitou-se que as inclusões proteicas formadas a partir de proteínas com cauda de poliglutamina expandida fossem a causa patogênica da doença (Davies *et al*, 1998). No entanto, trabalhos mais recentes vêm demonstrando uma falta de associação entre a formação de agregados proteicos com a morte neuronal e, pelo contrário, que este seria um mecanismo protetor da forma tóxica da proteína (Miller *et al*, 2010; Arrasate and Finkbeiner, 2012).

Já foi demonstrado que a forma difusa da proteína htt, bem como o tamanho da expansão poli-Q, podem prever a morte neuronal (Arrasate *et al*, 2004). Além disso, a formação de corpos de inclusão reduz os níveis da htt difusa, reduzindo o risco de morte dos neurônios estriatais, aumentando sua sobrevivência, enquanto que neurônios que não formaram agregados de htt permaneceram com o risco de morte da mesma forma elevado (Arrasate *et al*, 2004; Miller *et al*, 2010). Neste caso, a formação de agregados proteicos seria um mecanismo protetor de resposta contra a toxicidade proteica.

## 1.2 Modificadores genéticos

Apesar do número de repetições trinucleotídicas CAG ser o principal fator determinante da idade de início dos sintomas clínicos da DH, este número explica apenas 42 a 73% da variabilidade na idade de início, sendo sugerido anteriormente que o restante deve ser influenciado por fatores genéticos e ambientais (Wexler *et al*, 2004). Indivíduos da mesma família e, mesmo indivíduos com número idêntico de repetições CAG, apresentam variabilidade significativa na idade de início. Em um grupo de 370 indivíduos com DH, sendo 165 gêmeos, foi sugerida a ocorrência de modificadores genéticos que, independentemente do número de repetições CAG, contribuiriam em 11 a 19% para a variação de idade de início (Rosenblatt *et al*, 2001).

### 1.2.1 Estudos de *Genome Wide Association*

Os estudos de *Genome Wide Association* (GWAS) realizam uma busca pelo genoma de variações de sequência de DNA investigando uma possível associação entre genótipo e fenótipo em centenas de milhares a milhões de SNPs. O objetivo é encontrar genes envolvidos na determinação de traços humanos específicos, tal como altura, bem como em doenças genéticas humanas.

Alguns estudos GWAS vêm sendo desenvolvidos, especialmente em nível familiar, a fim de encontrar regiões no genoma que contenham fatores genéticos familiares responsáveis pela parte não CAG da variação da idade de início da DH. Em um desses estudos, os cromossomos 4 e 6 foram identificados como prováveis regiões ligadas à DH, além de evidência marginal de ligação para os cromossomos 1, 2, 5 e 18 (Li *et al*, 2003).



As regiões de maiores valores de LOD (*logarithm of odds*) foram 6p, 6q e 4p (2,29; 2,28; 1,93, respectivamente). Nestas regiões encontram-se genes candidatos a modificador, como o *CtBP* (*Corepressor c-terminal-binding protein*) (*locus* 4p16.3), bem como outros genes previamente descritos como modificadores da DH, como o *UCHL1* (4p14) (Naze *et al*, 2002) e o *GRIK2* (6q16.3) (Chattopadhyay *et al*, 2003).

Mais recentemente, procurando replicar um estudo anterior, as regiões 6q23-24 e 18q22 foram identificadas como sugestivamente ligadas à variação da idade de início na DH (Li *et al*, 2006). Ainda nesse estudo, análises epistáticas indicaram que a região 6q23-24 não interage com as outras cinco regiões avaliadas, podendo, assim, conter genes que participam de vias independentes na DH. Esta região contém 128 genes conhecidos, sendo dois deles de particular interesse: o *SGK* (*serum and glucocorticoid regulated kinase*) e o *GRM1* (*metabotropic glutamate receptor 1*), os quais aparecem superexpressos no cérebro de pacientes com DH e apresentam funções neuroprotetoras. Outro gene presente nesta região, descrito previamente como modificador, é o *GRIK2*. Na região 18q22, 36 genes são conhecidos, entre eles o *NEDD4L* (*neural precursor cell expressed, developmentally down-regulated 4-like*), o qual parece interagir com a *htt* e ter papel importante na patogênese da doença (Li *et al*, 2006).

Outro dois novos *loci* no cromossomo 2 (2p25 e 2q35) potencialmente ligados à DH foram descobertos em outro trabalho, além dos *loci* 5p14 e 5q32 (Gayan *et al*, 2008). Além disso, a provável ligação do *locus* 6q22, previamente descrito por outros trabalhos, foi confirmada. Os indivíduos incluídos neste estudo são provenientes da Venezuela e fazem parte do maior grupo de famílias com a doença já encontrado no mundo. Tal fato

possibilita uma análise mais precisa dos fatores genéticos familiares que determinam parte da variação da idade de início da DH (Gayan *et al*, 2008).

A identificação de modificadores genéticos na DH pode representar um grande avanço na identificação dos mecanismos capazes de retardar a idade de início da manifestação da doença.

### **1.2.2 Genes modificadores na DH**

A análise genotípica de genes candidatos é uma alternativa apropriada para identificar modificadores genéticos. Polimorfismos nestes genes não exercem efeito em indivíduos sem a doença, no entanto, podem modificar o curso da doença em indivíduos doentes.

Alguns trabalhos foram publicados com a identificação de genes modificadores da idade de início da DH. O efeito do polimorfismo S18Y, localizado no gene *UCHL1* (*ubiquitin carboxy-terminal hydrolase L1*), sobre a variação da idade de início foi avaliado por dois grupos distintos. O primeiro trabalho encontrou correlação entre a presença do SNP S18Y e a manifestação mais tardia da doença, determinando 13% da variação da idade de início (Naze *et al*, 2002). Um segundo estudo confirmou a associação de S18Y com a modulação da idade de início em pacientes com DH (1,7% da variação da idade de início). No entanto, foi observado que indivíduos homocigotos para o alelo Y apresentavam os sintomas oito anos mais cedo em relação aos outros genótipos (Metzger *et al*, 2006). Pacientes euro-descendentes homocigotos para o genótipo M441 (substituição de uma treonina por uma metionina) no gene *HAPI* (*huntingtin associated protein 1*)

manifestaram a doença oito anos mais tarde em relação aos pacientes homocigotos para o genótipo T441. Ensaio funcionais demonstraram que HAP1-M441 humana, além de interagir com a htt mutante mais fortemente que a HAP1-T441 humana, reduz os produtos de degradação de htt solúveis e protege da toxicidade mediada pela mesma (Metzger *et al*, 2008). Outro trabalho publicado mais recentemente demonstrou que entre os genes relacionados à autofagia - *Atg* (*autophagy-related*) – o polimorfismo V471A (troca de alanina por valina) em *Atg7* está associado com a antecipação da doença em quatro anos (Metzger *et al*, 2010).

Outros genes participantes de vias relacionadas à htt e/ou que foram identificados por *GWAS* em regiões potencialmente ligadas à DH serão abordados como candidatos a modificadores genéticos do fenótipo da doença. São eles: *GRM1*, *SGK1* e *TGM2*. Além disso, será feita a validação do polimorfismo M441 em *HAP1* como modificador da DH na população em estudo. Foram incluídos neste estudo genes relacionados às ataxias espinocerebelares (SCAs) de repetição CAG: *ATXN1* (SCA1), *ATXN2* (SCA2), *ATXN3* (SCA3), *CACNA1A* (SCA6), *ATXN7* (SCA7), *PPP2R2B* (SCA12), *TBP* (SCA17) e *DRPLA* (DRPLA), a fim de avaliarmos uma possível influência desses genes causadores de doenças neurodegenerativas na patogênese da DH.

### **1.2.2.1 *GRM1* (metabotropic glutamate receptor 1)**

Glutamato é o principal neurotransmissor excitatório no sistema nervoso de vertebrados e está envolvido em uma variedade de processos neuronais, como diferenciação, sobrevivência ou morte neuronal, proliferação, desenvolvimento das células neuronais e gliais e plasticidade na transmissão sináptica em cérebros adultos e em

desenvolvimento (Nakanishi and Masu, 1994). No entanto, por uma ação exacerbada deste neurotransmissor, conhecida como excitotoxicidade do glutamato, uma série de respostas de dano neuronal pós-isquemia ou hipoglicemia, epilepsia, tolerância e dependência a drogas, ansiedade, depressão, bem como processos neurodegenerativos agudos e crônicos estariam sendo desencadeadas (Meldrum, 2000).

Em relação ao envolvimento do glutamato em processos neurodegenerativos, três possíveis mecanismos são considerados: i) glutamato exógeno ou compostos relacionados ingeridos pela dieta, estariam agindo nos receptores de glutamato causando dano neuronal; ii) glutamato endógeno, liberado dos neurônios, poderia contribuir para o processo de neurodegeneração aguda em resposta à isquemia ou ao dano traumático cerebrais; iii) a ativação dos receptores de glutamato contribuiria para a morte celular em doenças neurodegenerativas crônicas, como a DH (Meldrum, 2000). Esta última hipótese é a atualmente postulada para os processos neurodegenerativos crônicos.

Os receptores de glutamato são divididos em duas classes: os ionotrópicos, que compreendem canais iônicos que permitem a entrada seletiva de cátions e mediam a neurotransmissão excitatória rápida, e os metabotrópicos (mGluRs), que são proteínas transmembrana, membros da família C de receptores acoplados à proteína G intracelular (GPCR) (Nakanishi and Masu, 1994).

Existem, pelo menos, três diferentes tipos de receptores de glutamato ionotrópicos, classificados em dois subgrupos: NMDA (receptor N-metil-D-aspartato) com as subunidades NR1, NR2A-NR2D, NR3A, NR3B e não-NMDA, que compreende os receptores AMPA (ácido  $\alpha$ -amino-3-hidróxi-5-metil-4-isoxazolepropionato), com as

subunidades GluR1, GluR2, GluR3 e GluR4, e os receptores cainato, que incluem as subunidades GluR5, GluR6, GluR7, KA1 e KA2 (Nakanishi and Masu, 1994).

Oito mGluRs distintos foram descritos em mamíferos e encontram-se divididos em três subgrupos com base na homologia de sequência, nas vias preferenciais de transdução de sinal e na sua farmacologia (Pin and Duvoisin, 1995). Os receptores do grupo I, que inclui os receptores mGluR1 e mGluR5, estão predominantemente acoplados à ativação da fosfolipase C via proteína  $G_q$  e são seletivamente ativados por 3,5-dihidroxifenilglicina (3,5-DHPG), cujo efeito é a produção de inositol trifosfato e subsequente mobilização de cálcio intracelular e ativação da proteína cinase C (PKC) (Ferraguti *et al*, 2008). O grupo II (mGluR2 e mGluR3) e o grupo III (mGluR4, mGluR6, mGluR7 e mGluR8) estão acoplados a proteínas  $G_i$  e são ativados, respectivamente, por 2R-4R ácido aminopiperidindicarboxílico e 2-amino-4-fosfonobutirato (Ferraguti *et al*, 2008). Estes grupos regulam negativamente a formação de cAMP (Pin and Duvoisin, 1995).

A ação excitotóxica do glutamato através de excessiva ativação dos seus receptores pode ser a causa da grande perda celular neuronal observada na DH, especialmente na região dos neurônios estriatais. Vários subtipos de mGluRs estão diferencialmente expressos na população de neurônios estriatais, predominando os mGluRs do grupo I (mGluR1 e mGluR5) (Kerner *et al*, 1997; Testa *et al*, 1998). É possível que essa distribuição diferencial possa resultar em diferentes respostas eletrofisiológicas cerebrais e explique a vulnerabilidade à excitotoxicidade célula-específica descrita na região do estriado no período inicial da manifestação da doença em pacientes com DH.

Os receptores mGluR do grupo I interagem com a proteína htt mutada e a optineurina (OPTN), uma proteína que interage com a htt, atua em antagonismo ao

agonista do receptor mGluR1a, de forma a bloquear a excitabilidade do receptor. Além disso, coprecipitação da htt normal foi evidenciada com o receptor mGluR1a, sendo que, quando OPTN é superexpressa, a coprecipitação da htt normal com mGluR1a aumenta em quase três vezes. No entanto, a expressão da htt normal e htt mutante sozinhas não teve efeito na formação de inositol fosfato (InsP) mediada pelo acoplamento do receptor mGluR1a com proteína Gq. Por outro lado, quando há coexpressão de OPTN, apenas a htt mutante leva ao aumento da ligação OPTN-mGluR1 e consequente aumento do antagonismo de OPTN sobre o receptor. Assim, não apenas OPTN, mas também htt mutante, pode antagonizar a sinalização de mGluR na DH (Anborgh *et al*, 2005).

Uma vez que os sintomas da DH só começam a aparecer em geral na vida adulta, é possível que tanto vias neuronais protetoras estejam ativadas ou mesmo que os sinais de morte celular estejam atenuados na fase pré-sintomática da doença. A estimulação com um agonista de mGluR1/5 evidenciou a formação atenuada de InsP, a partir da ativação do receptor mGluR1/5, até os 9 meses de idade de ratos pré-sintomáticos homocigotos *knock-in* para a htt mutante (*Hdh<sub>Q111</sub>/Hdh<sub>Q111</sub>*) quando comparados com ratos normais (*Hdh<sub>Q20</sub>/Hdh<sub>Q20</sub>*) (Ribeiro *et al*, 2010). No entanto, essa diferença na formação de InsP em ambos os modelos animais não foi observada a partir dos 11 meses de idade (Ribeiro *et al*, 2010). Estes dados sugerem que a sinalização do Grupo I de mGluRs pode ser diferencialmente modulada na presença da htt mutada e que esta modulação pode variar de acordo com a idade.

Os receptores mGluR1 estão envolvidos na regulação da excitabilidade neuronal, na plasticidade sináptica, na seleção sináptica e na liberação de neurotransmissores, mecanismos, estes, importantes para o desenvolvimento cerebral, aprendizado e

neuroproteção (Ferraguti *et al*, 2008). Desta forma, os receptores mGluR1 poderiam ser fatores-chave na determinação da idade de início em pacientes com DH.

O gene *GRM1* está localizado no *locus* 6q24 em humanos e codifica o receptor metabotrópico de glutamato 1 (mGluR1). Este gene se estende por 410 kb, dividido em 10 éxons.

O SNP rs1331639 está localizado no íntron 2 e o SNP rs6923492, presente no éxon 8 (c2977 > T), resulta em uma troca do aminoácido prolina por uma serina no resíduo 993 (P993S) (Downey *et al*, 2009).

#### **1.2.2.2 *SGK1 (serum/glucocorticoid regulated kinase 1)***

O gene *SGK1* está localizado no cromossomo 6, no *locus* 6p23, e codifica uma proteína serina/treonina cinase bastante dinâmica que desempenha um importante papel na resposta ao estresse celular. Entre muitas outras funções, esta cinase ativa determinados canais de sódio, potássio e cloro, sugerindo um envolvimento na regulação de processos de sobrevivência celular, excitabilidade neuronal e apoptose (Lang *et al*, 2009).

Algumas vias de transdução de sinal já foram descritas na fisiopatologia da DH. Entre elas está uma via neuroprotetora que envolve diretamente a fosforilação da proteína htt no resíduo de serina 421 (S421) pela serina/treonina cinase AKT. AKT é ativada por IGF-1 nos neurônios estriatais e a ativação dessa via inibe a morte celular induzida pela htt mutante e resulta na formação reduzida de inclusões intranucleares características da DH. É possível que a fosforilação da htt mutante altere a sua interação com outras proteínas, reduzindo sua toxicidade e a formação dos agregados proteicos (Humbert *et al*, 2002).

Além da AKT, IGF-1 leva à ativação de outra cinase que pode ter papel importante na promoção da sobrevivência celular, a proteína SGK. Nas células estriatais, SGK é ativada em resposta a IGF-1 e, assim como a AKT, promove a fosforilação da htt no resíduo S421 *in vitro* e *in vivo*. Essa fosforilação leva ao aumento da sobrevivência das células neuronais do estriado que expressam htt mutante. Tanto em amostras cerebrais como em modelo celular de DH, SGK apresentou níveis proteicos aumentados em três diferentes regiões do cérebro: cerebelo, córtex e estriado (Rangone *et al*, 2004).

Além de ser ativada pós-transcricionalmente em resposta a IGF-1, SGK também é regulada em nível transcricional em resposta a sinais extracelulares. Estes sinais incluem dano ao DNA, estímulos metabólicos, hormonais e estresse (Leong *et al*, 2003). A via p38/MAPK é a responsável pelo aumento da expressão de SGK em resposta a diferentes estresses, como sorbitol, aquecimento, radiação UV e peróxido de hidrogênio (Bell *et al*, 2000). Em modelo celular de DH, com o uso de inibidores desta via, os níveis de expressão de SGK permaneceram constantes, mostrando que na DH o aumento desta proteína depende da via de sinalização p38/MAPK (Rangone *et al*, 2004).

O gene *SGK1* possui comprimento de 418 kb, organizado em 14 éxons. Os SNPs escolhidos para análise neste gene são rs1763505 (íntron 1) e rs1075427 (íntron 7).

### **1.2.2.3 *TGM2 (C polypeptide, protein-glutamine-gamma-glutamyltransferase)***

O gene *TGM2* está localizado no cromossomo 20, no locus 20q12 e codifica a enzima *transglutaminase tipo 2* (tecidual) – TGM2. Esta proteína se caracteriza por sua versatilidade multifuncional dentro da família das transglutaminases, sendo a mais



largamente expressa em tecidos de mamíferos, envolvida em uma variedade de funções bioquímicas e presente em diferentes sublocalizações celulares (Fesus and Piacentini, 2002). Entre suas principais atividades estão: i) modificações pós-traducionais de proteínas envolvendo ligações cruzadas  $\text{Ca}^{+2}$ -dependentes entre resíduos de glutamina e de lisina e a incorporação de aminas primárias a resíduos de glutamina; ii) atividades de proteína isomerase dissulfeto e de proteína cinase; iii) mediação da transdução de sinal transmembrana e das interações entre proteínas de superfície celular e a matriz extracelular. Em baixos níveis de  $\text{Ca}^{+2}$ , TGM2 atua como proteína-G, ligando-se e hidrolisando GTP com afinidade e ação catalítica similar a proteínas-G clássicas (Fesus and Piacentini, 2002).

A atividade de ligação cruzada  $\text{Ca}^{+2}$ -dependente de TGM2 entre resíduos de glutamina (Q) de uma proteína/peptídeo substrato e lisina (K) de uma proteína/peptídeo cosubstrato é relacionada com a formação de agregados proteicos nas principais doenças neurodegenerativas, entre elas a DH. A formação de agregados na presença de TGM2 foi evidenciada *in vitro* com construtos de htt contendo domínios  $Q_n$ . Em geral, a formação de agregados foi mais proeminente quanto maior o número de glutaminas e com as formas mais truncadas da proteína htt (Karpuj *et al*, 1999). Neste mesmo estudo foi observada a superexpressão desta proteína em extratos cerebrais de córtex e de cerebelo de pacientes com DH quando comparados a amostras controle, sendo detectada maior atividade da mesma nos extratos corticais nucleares se comparados aos extratos citosólicos (Karpuj *et al*, 1999).

Uma melhor elucidação das implicações dos achados acima citados foi obtida com a demonstração do envolvimento de TGM2 na disfunção energética e transcricional

observada em doenças neurodegenerativas que envolvem proteínas poli-Q expandidas. A inativação *in vitro* de duas proteínas relacionadas ao metabolismo energético da célula, a gliceraldeído-3-fosfato-desidrogenase (GAPDH) e o complexo  $\alpha$ -cetoglutarato desidrogenase (KGDHC), foi observada quando as mesmas estavam envolvidas em ligações cruzadas com uma proteína heteróloga apresentando poli-Q expandido. Este fato pode estar interligado com a ruptura da homeostase energética cerebral observada nessas doenças (Cooper *et al*, 1997). Subsequentemente, a histona H1 foi identificada como substrato para TGM2, sugerindo a participação da mesma no remodelamento da cromatina e na regulação da expressão gênica (Cooper *et al*, 2000).

Ficou evidenciado que TGM2 poliamina a cauda N-terminal da histona H3 (Zainelli *et al*, 2005). A poliaminação da cauda N-terminal de H3 aumenta a carga positiva da mesma e a sua propensão a interagir mais fortemente com o DNA, o qual é carregado negativamente. A TGM2 participaria, dessa maneira, da alteração da estrutura da cromatina atuando como repressora da transcrição de genes alvo (Zainelli *et al*, 2005). Assim, uma superexpressão acompanhada com hiperatividade de TGM2 poderia desregular a transcrição gênica. Corroborando esses achados, a inibição química de TGM2 em neurônios derivados de modelo animal de DH expressando a proteína htt mutante completa resultou na normalização dos níveis de expressão (em pelo menos 25%) de 196 genes dos 461 analisados (Zainelli *et al*, 2005). Utilizando o ensaio de imunoprecipitação da cromatina (CHIP) foi evidenciado que TGM2 ocupava extensivamente as regiões promotora/intensificadora de dois genes essenciais na produção energética: *PGC-1 $\alpha$*  (*peroxisome proliferator-activated receptor- $\gamma$  coactivator 1- $\alpha$* ) e *COX-1* (*cytochrome c oxidase subunit I*), de modo a reprimir a transcrição dos mesmos (Zainelli *et al*, 2005). A

redução da expressão de PGC-1 $\alpha$  e de genes relacionados contribui para as disfunções mitocondriais e metabólicas características da DH.

Estudos com inibição química de TGM2 com cistamina – inibidor de TGM2 – e geneticamente por meio de *knockout* de *TGM2*, mostraram melhora do processo neurodegenerativo em modelo animal murino de DH, levando-se em consideração os sintomas apresentados, com o retardo da idade de início das disfunções motoras e o prolongamento do tempo de sobrevivência dos mesmos (Mastroberardino *et al*, 2002; Bailey and Johnson, 2006).

Embora a vasta maioria das proteínas humanas seja polimórfica, devido, principalmente, a polimorfismos de nucleotídeo único (SNP), é de se esperar que TGM2 possua um alto grau de conservação genômica com limitadas variações no gene, uma vez que esta desempenha inúmeras funções importantes no organismo. De fato, ficou evidenciado, após análise sistêmica do gene, que este possui baixa frequência de variações genéticas na sequência codificante da proteína (Kiraly *et al*, 2011).

Poucos dados de frequência alélica estão disponíveis para os SNPs neste gene e os valores de MAF (*minimum allele frequency*) são bastante baixos, o que dificulta a identificação das variantes na população de estudo. Os SNPs rs7270898 e rs7275079 (íntron 2) estão localizados em uma região de possível regulação da transcrição gênica, podendo influenciar a expressão do gene *TGM2* em pacientes com DH, sendo relativamente polimórficos na população.

#### 1.2.2.4 *HAPI (huntingtin associated protein 1)*

O gene *HAPI* está localizado no cromossomo 17, *locus* 17q21.2-q21.3, apresenta 12,1 kb de extensão e se divide em 11 éxons, codificando a proteína de mesmo nome. *HAPI* foi a primeira proteína descrita a interagir com a proteína htt. Essa interação é mais forte com a htt mutada do que com a htt normal. A expressão de *HAPI*, ao contrário da proteína htt, é restrita ao cérebro, sendo seletivamente expressa em neurônios (Li *et al*, 1995). *HAPI* estaria envolvida com tráfego axonal, uma vez que interage com proteínas de transporte dependentes de microtúbulos, com tráfego intracelular e endocitose, baseado na sua associação com organelas endossomais, com o receptor 1,4,5-inositol trifosfato tipo I e com o receptor de androgênios, além de estar envolvida com o transporte de BDNF (*Brain-Derived Neurotrophic Factor*) (Engelender *et al*, 1997; Tang *et al*, 2003; Gauthier *et al*, 2004; McGuire *et al*, 2006; Takeshita *et al*, 2006). Tais funções corroboram sua interação direta com a proteína htt, fazendo de *HAPI* um forte candidato a modificador genético do fenótipo da DH.

Novos polimorfismos foram identificados no gene *HAPI* e investigados quanto a uma possível associação com a determinação da idade de início em 900 indivíduos com DH provenientes da Europa (Metzger *et al*, 2008). O polimorfismo rs4523977 (g.5157502G>A) foi identificado como modificador da idade de início nos pacientes incluídos no estudo. A mudança polimórfica do resíduo treonina por metionina na posição 441 (T441M) foi responsável por um atraso na manifestação dos sintomas em 8 anos para os indivíduos homozigotos para o genótipo M441. Através de ensaios funcionais, evidenciou-se uma associação forte entre a proteína *HAPI*-M441 e a proteína htt mutante,

com a redução dos produtos de degradação de htt solúveis e a proteção contra a toxicidade mediada pela mesma (Metzger *et al*, 2008).

Dois anos após a identificação de *HAPI* como modificador da idade de início da DH, outro grupo reproduziu o mesmo estudo, com a inclusão de alguns genes previamente identificados como modificadores, em uma população de 419 pacientes, menos da metade utilizada no primeiro estudo. A mesma associação, embora modesta, foi encontrada para o alelo de risco A. A diferença na média da idade de início para o genótipo M441 homocigoto em relação aos demais genótipos foi bastante sutil, mas mostrou uma tendência de retardo no início da manifestação da DH para este grupo (Taherzadeh-Fard *et al*, 2010).

O SNP rs4523977 será incluído nesse estudo visando a validação da ação protetora deste polimorfismo na população em questão.

#### **1.2.2.5 Ataxias espinocerebelares**

As ataxias espinocerebelares (SCAs) são um grupo de manifestações patológicas neurodegenerativas raras, de herança autossômica dominante, que apresentam como característica comum a degeneração cerebelar progressiva, com conseqüente prejuízo no equilíbrio, na marcha, nos movimentos coordenados dos membros, bem como na fala.

À medida que as várias SCAs foram sendo identificadas, um número correspondente à ordem cronológica das descobertas foi sendo atribuído a cada uma delas (SCA1, SCA2, e assim sucessivamente), sendo a SCA36 a mais recentemente descrita (até 20 de dezembro de 2012) (Kobayashi *et al*, 2011). O número desta última forma de SCA

não corresponde diretamente ao total de SCAs descritas até o momento, uma vez que dentro desta classificação foram incluídas outras síndromes relacionadas que não ataxias espinocerebelares propriamente ditas. Desconsiderando essas síndromes relacionadas, há um total de 30 diferentes *loci* associados às SCAs, sendo que para 21 delas já foram identificadas as variantes genéticas associadas (Hersheson *et al*, 2012).

O grupo mais prevalente dentro das ataxias espinocerebelares inclui as SCAs causadas por expansão das repetições trinucleotídicas CAG na região codificante dos genes, as quais codificam o aminoácido glutamina. As doenças de “poliglutaminas”, como são conhecidas as SCAs pertencentes a este grupo, incluem a SCA1, SCA2, SCA3, SCA6, SCA7, SCA17 e DRPLA (Tabela 1). Esse grupo de SCAs apresenta em comum a produção de uma proteína mutante com um trecho anormalmente alongado de poliglutaminas, à qual se atribui as causas das manifestações clínicas observadas em indivíduos com a doença. A SCA12 também é causada por uma expansão das repetições CAG, porém estas se localizam na região 5' não-codificante do gene *PPP2R2B* (Holmes *et al*, 1999) (Tabela 1). O mecanismo de patofisiologia da SCA12 ainda é desconhecido. Assim como na DH, as SCAs se manifestam quando o número de repetições ultrapassa o limiar considerado normal e a idade de início da manifestação dos sintomas também é determinada por uma correlação inversa ao número de repetições CAG. O número de repetições tende a aumentar ao longo das gerações, resultando em uma idade de início mais precoce e em um fenótipo mais grave de uma geração para outra; este evento é chamado de antecipação genética (Hersheson *et al*, 2012).

**Tabela 1. Ataxias espinocerebelares dominantes de repetição CAG**

Nome	OMIM	Locus	Gene	Proteína	Tipo de mutação	% SCAs
SCA1	164400	6p22.3	<i>ATXN1</i>	Ataxina 1	Expansão CAG entre 39 e 81 repetições puras	6–27%
SCA2	183090	12q24.12	<i>ATXN2</i>	Ataxina 2	Expansão CAG mais de 32 repetições puras	13–18%
SCA3	109150	14q32.12	<i>ATXN3</i>	Ataxina 3	Expansão CAG entre 54 e 86	20–50%
SCA6	183086	19p13.2	<i>CACNA1A</i>	Subunidade $\alpha$ 1A do canal de cálcio voltagem dependente tipo P/Q	Expansão CAG entre 21 e 33	13–15%
SCA7	164500	3p14.1	<i>ATXN7</i>	Ataxina 7	Expansão CAG entre 36 e 450	3–5%
SCA12	604326	5q32	<i>PPP2R2B</i>	Subunidade regulatória B da proteína fosfatase 2 $\beta$	Repetição CAG em 5' UTR maior que 51 repetições	Rara
SCA17	607136	6q27	<i>TBP</i>	Proteína ligante de TATA-box	Expansão CAG/CAA entre 43 e 66	Rara
DRPLA	125370	12p13.31	<i>DRPLA</i>	Atrofina 1	Expansão CAG entre 48 e 93	Rara

Ao contrário das mutações estáticas que são estavelmente transmitidas para a prole, as mutações de repetições nucleotídicas são dinâmicas, com um processo de mutação continuada nos tecidos e ao longo de gerações. Trechos longos de repetições são mais suscetíveis à mutação do que trechos curtos. O mecanismo molecular que rege a instabilidade das mutações dinâmicas ainda não está totalmente esclarecido. A instabilidade do DNA pode ser causada pela própria expansão ou em conjunto com *cis*-elementos gene-específicos que se localizam próximos à expansão e por *trans*-elementos que atuam na replicação e reparo do DNA (Pearson *et al.*, 2005).

Já foram descritas associações entre alelos normais longos em *cis*-elementos gene-específicos, e a modulação do fenótipo de doenças neurodegenerativas. É o caso da associação entre alelos longos de *ATXN2* (> 29 repetições CAG) e o risco aumentado para

ELA (Esclerose Lateral Amiotrófica) (Daoud *et al*, 2011). Da mesma forma, a presença de alelos normais longos no gene *CACNA1A* mostrou-se associada com a manifestação precoce da ataxia espinocerebelar do tipo 2 (SCA2) (Pulst *et al*, 2005).

Assim, repetições CAG longas em alelos normais de genes relacionados a doenças de poliglutaminas podem influenciar na variação da idade de início em pacientes com DH. Os genes relacionados às ataxias espinocerebelares SCA1 (*ATXN1*), SCA2 (*ATXN2*), SCA3 (*ATXN3*), SCA6 (*CACNA1A*), SCA7 (*ATXN7*), SCA12 (*PPP2R2B*), SCA17 (*TBP*) e DRPLA (*DRPLA*) foram selecionados como candidatos a modificadores da idade de início da DH.



## 2. JUSTIFICATIVA

A DH é a forma mais comum das doenças neurodegenerativas hereditárias. Esta se manifesta em geral na vida adulta, sendo a idade de início dos sintomas parcialmente determinada pelo número das repetições CAG expandidas no gene relacionado à doença. O restante da variabilidade da idade de início deve estar relacionado a fatores genéticos e ambientais. A compreensão da ação desses fatores genéticos na modulação da idade de início da DH pode ser bastante elucidativa quanto aos mecanismos envolvidos na patogênese da doença, bem como pode ser o ponto de partida para a geração de intervenções terapêuticas eficientes na contenção da doença.

Alguns genes modificadores da idade de início já foram descritos na DH. A análise de variantes polimórficas em genes candidatos a modificadores é uma estratégia útil na determinação de fatores genéticos moduladores do fenótipo da doença. Os genes *GRM1*, *SGK1*, *ATXN1*, *PPP2R2B* e *TBP* estão em regiões, ou em regiões próximas, identificadas por estudos de *GWAS* como associadas à DH e podem estar diretamente envolvidos na determinação da idade de início de indivíduos com DH. O gene *TGM2* parece ter função bastante importante na patogênese da DH, por sua interação com a proteína htt e por seu envolvimento na disfunção energética e transcricional observada em doenças neurodegenerativas que envolvem proteínas poli-Q expandidas. O gene *HAP1* já foi descrito como modificador da DH e a sua validação em subgrupos diferentes da população poderá reforçar sua importância na fisiopatologia da doença. Os genes relacionados às SCAs (*ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7*, *PPP2R2B*, *TBP* e *DRPLA*) podem atuar como elementos moduladores do fenótipo de doenças neurodegenerativas através de alelos normais longos em *cis*-elementos gene-específicos.

### 3. OBJETIVOS

#### 3.1 Objetivo Geral

Identificar a ação de alguns fatores genéticos na modulação da idade de início da DH;

#### 3.2 Objetivos Específicos

3.2.1 Determinar as frequências alélicas e genotípicas de variantes polimórficas nos genes, *GRM1* (rs1331639 e rs6923492), *SGK1* (rs1763505 e rs1075427), *TGM2* (rs7270898 e rs7275079) e *HAP1* (rs4523977) em pacientes e em indivíduos controles;

3.2.2 Avaliar a presença de associação entre as variantes genéticas acima citadas e a variabilidade da idade de início dos sintomas clínicos em indivíduos com DH;

3.2.3 Correlacionar o número de repetições CAG no alelo normal do gene *HTT* com a idade de início dos sintomas em indivíduos com DH;

3.2.4 Determinar o número de repetições CAG em ambos os alelos dos genes *ATXN1* (SCA1), *ATXN2* (SCA2), *ATXN3* (SCA3), *CACNA1A* (SCA6), *ATXN7* (SCA7), *PPP2R2B* (SCA12), *TBP* (SCA17) e *DRPLA* (DRPLA) em casos e controles, comparar as frequências alélicas encontradas e correlacionar o tamanho dos alelos com a variação da idade de início em indivíduos com DH.

## 4. RESULTADOS

Os resultados deste trabalho serão apresentados na forma de dois artigos científicos.

#### 4.1 Artículo 1

### **SEARCHING FOR GENETIC MODIFIERS OF AGE AT ONSET IN HUNTINGTON'S DISEASE**

Este manuscrito será submetido para a revista *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*.

## SEARCHING FOR GENETIC MODIFIERS OF AGE AT ONSET IN HUNTINGTON'S DISEASE

Running head title: Genetic modifiers of Huntington's disease age at onset

Aline Francielle Damo Souza BSc<sup>1,2</sup>; Tailise Conte Gheno MSc<sup>1</sup>; Carlos Rieder MD, PhD<sup>3</sup>; Laura Bannach Jardim MD PhD<sup>1,4</sup>; Maria Luiza Saraiva-Pereira PhD<sup>1,2,5</sup>

<sup>1</sup>Laboratório de Identificação Genética– Centro de Pesquisa Experimental e Serviço de Genética Médica, Hospital de Clínicas de Porto Alegre (HCPA); <sup>2</sup>Programa de Pós-Graduação em Biologia Celular e Molecular – Universidade Federal do Rio Grande do Sul (UFRGS); <sup>3</sup>Serviço de Neurologia – HCPA; <sup>4</sup>Departamento de Medicina Interna – UFRGS; <sup>5</sup>Departamento de Bioquímica – UFRGS.

### **Corresponding author:**

Maria Luiza Saraiva-Pereira, PhD  
Medical Genetics Service  
Hospital de Clínicas de Porto Alegre  
Rua Ramiro Barcelos, 2350  
CEP 90035-903 – Porto Alegre – Brazil  
Tel: +55 51 3359-8011  
Fax: +55 51 3359-8010  
e-mail: mlpereira@hcpa.ufrgs.br

**ABSTRACT**

Huntington's disease is the most common inherited neurodegenerative disorder generally starting at adulthood. Mean age at onset (AO) is around 40 years and is mainly determined by CAG repeats length of the mutant *HTT* allele. The remaining variation observed in HD patient is due to genetic and environmental factors. Some genetic modifiers of AO have already been described and are useful to understand the HD pathogenesis. *GRM1*, *SGK1*, *TGM2* and *HAPI* genes were assessed in this study in a cohort of 91 Brazilian unrelated HD patients. Association of *HAPI* and AO were confirmed as previously reported with a modest but significant correlation ( $p=0.044$ ). Lack of association of other tested variables may be due to small size sample. Therefore, multicenter studies should be performed with those genes or other good candidates to HD modifiers.

**Key words:** Huntington's disease, genetic modifiers, age at onset, single nucleotide polymorphism (SNP)

## INTRODUCTION

Huntington's disease (HD) (OMIM 143100) is the most common neurodegenerative disorder that is autosomal dominantly inherited with clinical age at onset (AO) in adult life. HD is characterized by motor dysfunction, cognitive deficits, and psychiatric disturbances that starts commonly between 35–50 years of age (Hayden, 2001). HD associated mutation is an expansion of CAG trinucleotide repeats in the first exon of *HTT* gene that encodes for huntingtin (htt) protein (MacDonald *et al*, 1993). The mutant protein with an abnormally long polyglutamine stretch is responsible for the dysfunction, neurodegeneration, and the primary clinical characteristics shown by HD patients. CAG repeat length is known to be the main factor associated to AO. However, it is responsible for only 42-73% of its variation that leads to the occurrence of environmental and genetic modifiers (Wexler *et al*, 2004). Based on studies in siblings, familial genetic factors might account for 11 to 19% of AO variation (Rosenblatt *et al*, 2001). Among these genetic factors, other genes can be able of modifying the course of disease.

Genome wide association studies (GWAS) have been performed and genomic regions have been associated to AO variation in HD (Gayan *et al*, 2008; Li *et al*, 2003; Li *et al*, 2006). A great number of genes were identified in these regions and *UCHL1* gene located at *locus* 4p14 (Naze *et al*, 2002) and *GRIK2* gene located at *locus* 6q16.3 (Chattopadhyay *et al*, 2003) have been previously identified as potential genetic modifiers. Polymorphisms in these genes can modulate mechanisms responsible for delaying or anticipating the onset of the disorder.

We have chosen four genes among a wide variety of candidates to assess their potential association to AO in this study. *GRM1* (*metabotropic glutamate receptor 1*) and *SGK*

(*serum and glucocorticoid regulated kinase*) are located in regions that have been associated with AO modulation by GWAS (Li *et al*, 2006) and are involved in neuroprotection; *TGM2* (*C* polypeptide, *protein-glutamine-gamma-glutamyltransferase*) plays a central role in HD pathogenesis, regulating gene transcription and formation of htt aggregates; and *HAPI* (*huntingtin associated protein 1*) has been previously described as a modifier in HD (Metzger *et al*, 2008), and it was included in this work to validate the data in our population.



## SUBJECTS AND METHODS

Ninety-one unrelated Brazilian HD patients were previously diagnosed by molecular analysis. Age at onset (AO) of each patient was defined as the date when patient or a close relative noticed the first symptom. Either the patient or a close relative have signed an informed consent to participate in this study. We have also included 100 unrelated subjects as controls.

DNA was isolated from peripheral blood by standard methods (Miller *et al*, 1988). CAG repeat length was determined by PCR (Polymerase Chain Reaction) using primers HuntF 5'-GACCCTGGAAAAGCTGATGAAGGC-3' and HuntAR 5'-TGGCGGCTGTTGCTGCTGCTG-3'. The latter was labeled in 5' end with 6-FAM (6-carboxifluorescein). PCR reaction was performed in a total of 25  $\mu$ L, containing 40 ng of genomic DNA, 4 pmol of each primer, 2 mM dNTPs, 1.5 mM MgCl<sub>2</sub>, 1X *GeneAmp*<sup>®</sup> *Gold Buffer*, 10% (v/v) DMSO (dimethyl sulfoxide), and 1.25 U of *AmpliTaq Gold DNA polymerase*. Amplification was as follows: denaturation step at 95°C for 10 minutes, followed by 5 cycles at 94°C for 30 seconds and 72°C for 3 minutes, 5 cycles at 94°C for 30 seconds and 70°C for 3 minutes, 5 cycles at 94°C for 30 seconds, 68°C for 1 minute and 72°C for 2 minutes and 20 cycles at 94°C for 30 seconds, 60°C for 1 minute and 72°C for 3 minutes and a final extension at 72°C for 20 minutes. Following amplification, PCR product was submitted to capillary electrophoresis in an *ABI3130xl* Genetic Analyzer (Applied Biosystems) using a mixture of 0.5  $\mu$ L of each sample, 9.3  $\mu$ L of Hi-Di<sup>™</sup> Formamide (Applied Biosystems) and 0.2  $\mu$ L of GeneScan<sup>™</sup> 500 LIZ<sup>™</sup> size standard. Samples were denatured at 95°C for 5 minutes and left on ice for 5 minutes before being injected during 22 seconds at 3 kV and left running during 40 minutes at 60°C in a 50 cm

long capillary containing Polymer POP-7 (Applied Biosystems). Data was then analyzed by GeneMapperID<sup>®</sup> software 4.0 (Applied Biosystems).

Allelic discrimination of SNPs was performed by qualitative real time PCR using the TaqMan<sup>®</sup> PCR Assay (Applied Biosystems), which is based on fluorescent minor groove binding probes. Assays used in this study are listed in Table SI and were ordered from Applied Biosystems. PCR reaction was performed in a final volume of 8  $\mu$ L containing 4 ng of DNA, 0.2  $\mu$ L of specific TaqMan assay and 4  $\mu$ L of 2X PCR Genotyping Master Mix (Applied Biosystems). Amplification included an initial step at 50°C for 2 minutes (activation of AmpErase UNG function), AmpliTaq<sup>®</sup> Gold activation at 95°C for 10 minutes, followed by 40 cycles of denaturation at 95°C for 15 seconds and annealing extension at 60°C for 1 minute. Allelic discrimination step was performed at 60°C for 1 minute. PCR products were analyzed by the Sequence Detection System Software version 1.2.1 (Applied Biosystems) in the ABI Prism 7500 through allelic discrimination plot.

The majority of SNPs were selected using SNP Tagger tool of HAPLOVIEW software, available from The International HapMap Consortium, 2009. rs4523977 (*HAPI*) and rs6923492 (*GRMI*) were included based on previously published data (Metzger *et al*, 2008); (Downey *et al*, 2009).

Statistical analyses were performed using SPSS Statistics 18 (Statistical Package for Social Sciences, version 18). Allelic and genotypic frequencies were estimated by gene count and compared by Pearson Chi-square ( $\chi^2$ ). Genotypic frequencies were tested to deviation of Hardy-Weinberg equilibrium, using Chi-square test with significance level of 5%.

Correlation between AO and CAG repeat number in mutated allele was estimated using simple linear regression, with logarithmic transformation of AO to a better fit on

determination coefficient ( $R^2$ ). The influence of polymorphism on AO was assessed by multiple linear regressions including each SNP individually.

## RESULTS

Our cohort included patients with CAG repeat size in the mutated allele ranging from 39 to 79 repeats with a mean repeat number of 47 ( $47 \pm 7.6$  repeats), while normal allele CAG repeats ranged from 10 to 30 repeats with a mean of 18.3 ( $18.3 \pm 3.4$  repeats). Mean AO was 37.3 years ( $37.3 \pm 12.4$  years) ranging from 6 to 72 years. In our case, length of CAG repeat was responsible for 65.3% ( $R^2=0.653$ ;  $p<0.0001$ ) of variation in AO (Figure S1). The regression model was well adjusted after a logarithmic transformation of dependent variable AO. According to this function, an increase in one CAG unit is responsible for a decrease of 0.95 years at AO.

Allelic and genotypic frequencies distribution in cases and controls are presented in table I. rs6923492 allelic and genotypic frequencies were statistically different ( $\chi^2=5.079$ ;  $p=0.024$  and  $\chi^2=14.079$ ;  $p=0.001$ , respectively). Genotypic distribution was tested to Hardy-Weinberg Equilibrium (HWE) and both rs1075427 (*SGK1*) and rs6923492 (*GRM1*) were not in HWE in both cases and controls ( $\chi^2=7.239$ ;  $p=0.027$  and  $\chi^2=7.586$ ;  $p=0.023$ ). Both SNPs were then excluded from further analysis.

**Table I.** Allelic and genotypic SNP frequencies distribution

Allele/Genotype	HD patients n = 91	Normal subjects n = 100	$\chi^2$	$p^a$
<b>rs6923492 (GMR1)</b>				
C	102 (0.56)	89 (0.44)		
T	80 (0.44)	111 (0.56)	5.079	0.031*
C/C	32 (0.35)	13 (0.13)		
C/T	38 (0.42)	63 (0.63)		
T/T	21 (0.23)	24 (0.24)	14.079	0.001*
<b>rs1331639 (GRM1)</b>				
C	111 (0.61)	126 (0.63)		
T	71 (0.39)	74 (0.37)	0.164	0.752
C/C	35 (0.38)	36 (0.36)		
C/T	41 (0.45)	54 (0.54)		
T/T	15 (0.16)	10 (0.10)	2.359	0.317
<b>rs1763505 (SGK1)</b>				
C	123 (0.68)	135 (0.68)		
T	59 (0.32)	65 (0.33)	0.000	1.000
C/C	38 (0.42)	46 (0.46)		
C/T	47 (0.52)	43 (0.43)		
T/T	6 (0.06)	11 (0.11)	1.956	0.396
<b>rs1075427 (SGK1)</b>				
C	104 (0.57)	112 (0.56)		
T	78 (0.43)	88 (0.44)	0.051	0.837
C/C	36 (0.40)	33 (0.33)		
C/T	32 (0.35)	46 (0.46)		
T/T	23 (0.25)	21 (0.21)	2.315	0.314
<b>rs7270898 (TGM2)</b>				
A	90 (0.49)	105 (0.53)		
G	92 (0.51)	95 (0.47)	0.355	0.609
A/A	21 (0.23)	25 (0.25)		
A/G	48 (0.53)	55 (0.55)		
G/G	22 (0.24)	20 (0.20)	0.513	0.807

<b>rs7275079 (TGM2)</b>					
<b>C</b>	91 (0.50)	95 (0.47)			
<b>T</b>	91 (0.50)	105 (0.53)	0.238	0.682	
<b>C/C</b>	22 (0.24)	20 (0.20)			
<b>C/T</b>	47 (0.52)	55 (0.55)			
<b>T/T</b>	22 (0.24)	25 (0.25)	0.510	0.792	
<b>rs4523977 (HAP1)</b>					
<b>A</b>	30 (0.16)	41 (0.20)			
<b>G</b>	152 (0.84)	159 (0.80)	1.016	0.357	
<b>A/A</b>	4 (0.05)	3 (0.03)			
<b>A/G</b>	22 (0.24)	35 (0.35)			
<b>G/G</b>	65 (0.71)	62 (0.62)	2.797	0.268	

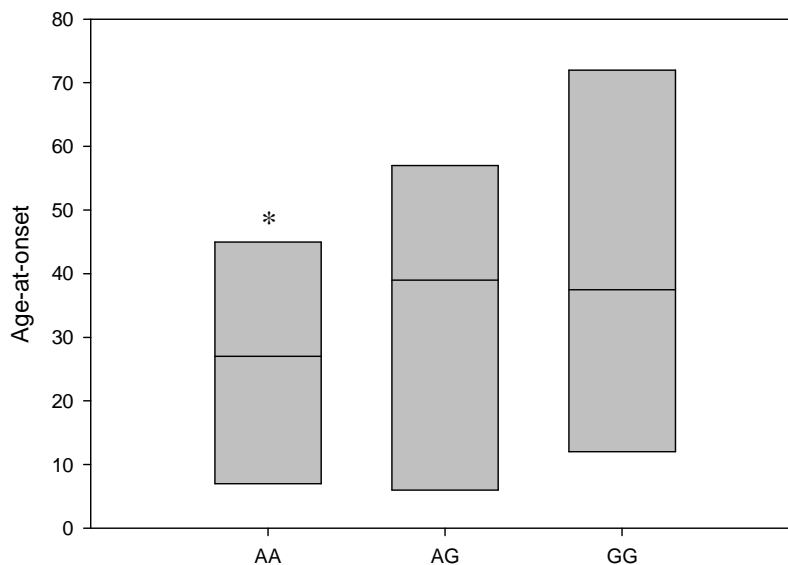
Frequencies are shown within parentheses. Fisher's exact test ( $\chi^2$ ) was used to compare allelic and genotypic frequencies in cases and controls. <sup>a</sup> P values < 0.05 were considered significant.

Regression model was used to assess the influence of rs1331639, rs1763505, rs7270898, rs7275079, and rs4523977 in variation of AO together with length of CAG tract. Each SNP genotype was assessed one by one. These results are shown in table II. We have confirmed correlation of rs4523977 AA genotype and AO ( $p=0.044$ ) that was previously published (Metzger *et al*, 2008). Mean AO of AA genotype was 27 years ( $27 \pm 16.3$  years) and this is statistically significant from other genotypes (Figure 1). Lack of association with late onset might be a consequence of small homozygous patients number ( $n=4$ ). Examining the variation in model correlation ( $\Delta R^2$ ) we see that all models increased the  $R^2$  value, even though not statistically significantly, especially rs4523977 ( $\Delta R^2=0.016$ ).

**Table II.** Single and multiple linear regression models

<b>Model</b>	<b>R<sup>2</sup></b>	<b>ΔR<sup>2</sup></b>	<b>p<sup>a</sup></b>	<b>Power<sup>b</sup></b>
CAG x AO	0.653	-	<0.0001*	1.000
CAG x AO + rs7270898	0.654	0.001	0.886	0.068
AA			0.660	0.072
AG			0.664	0.072
GG			0.886	0.068
CAG x AO + rs7275079	0.654	0.001	0.886	0.068
CC			0.658	0.072
CT			0.927	0.051
TT			0.886	0.068
CAG x AO + rs4523977	0.669	0.016	0.130	0.418
AA			0.044*	0.524
AG			0.771	0.060
GG			0.130	0.418
CAG x AO + rs1763505	0.657	0.004	0.664	0.115
CC			0.675	0.070
CT			0.452	0.116
TT			0.664	0.115
CAG x AO + rs1331639	0.654	0.001	0.949	0.058
CC			0.753	0.061
CT			0.785	0.058
TT			0.949	0.058

<sup>a</sup> P values < 0.05 were considered significant. <sup>b</sup> Statistical power observed for tested sample.



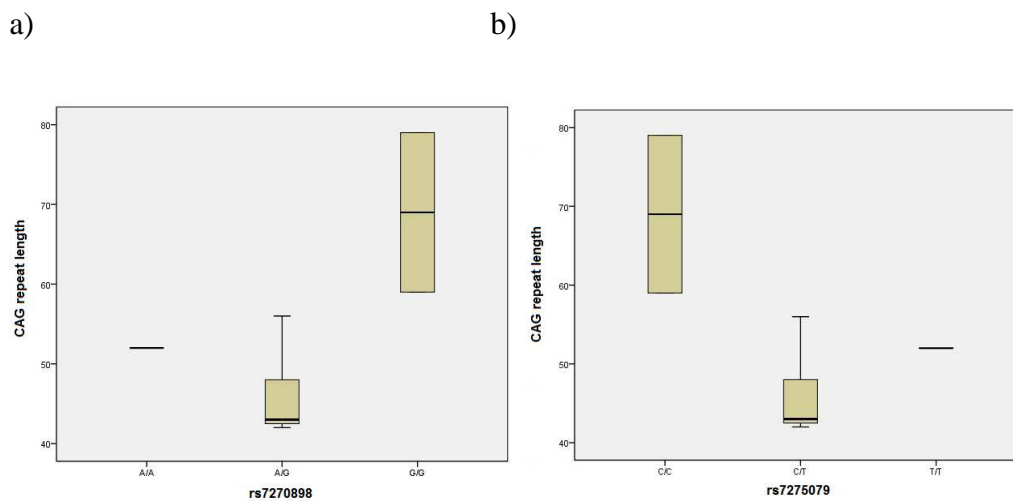
**Figure 1.** AO distribution in rs4523977 AA, AG, and GG genotypes. Mean AO (black line in boxes) were AA=27 ( $27 \pm 16.3$  years) (ranging from 7 to 45 years); AG=39 ( $39 \pm 13.3$  years) (ranging from 6 to 57 years) and GG=37.5 ( $37.5 \pm 11.7$  yaers) (ranging from 12 to 72 years).

Patients were divided into three groups according to expected AO variation: (i) patients with AO more than one SD (standard deviation) below average (-1SD), named early-onset group (n=10); (ii) patients with AO between -1 and +1 SD from average, named average AO (n=74); (iii) patients with AO one SD above the average (+1SD), named late-onset group (n=7). Genotypic frequencies of each SNP in early-onset, average, and late-onset group were compared by Chi-square test with Monte Carlo correction. No statistical significant difference was found between group frequencies (rs7270898 -  $p=0.079$ ; rs7275079 -  $p=0.056$ ; rs4523977 -  $p=0.394$ ; rs1763505 -  $p=0.743$ , and rs1331639 -  $p=0.821$ ).



Distribution of CAG repeats in early-onset group shows no difference to all genotypes of each tested SNP in both groups. A tendency of association was observed in rs7270898 CC and rs7275079 GG genotypes ( $p=0.074$ ) with large CAG repeat length (Figure 2). Only the genotypes CC ( $n=3$ ) and CT ( $n=7$ ) were found to rs1763505 ( $p=0.491$ ). The Mann-Whitney test was not significant to CAG length distribution in both genotypes ( $U=7.500$ ;  $p=0.491$ ). Results of additional SNPs are shown in figure S2.

Similar analysis was done in the late-onset group. All patients were heterozygous in rs7270898 and rs7275079 (AG and CT, respectively) and showed the same CAG repeat length average  $47.86 \pm 13.67$  repeats. Only the genotypes AG ( $n=2$ ) and GG ( $n=5$ ) and none AA (M411) were found to rs4523977. This explains the lack of association of M441 with late-onset disease. Only genotypes CC ( $n=4$ ) and CT ( $n=3$ ) were detected in the analysis of rs1763505. CAG length distribution was the same in both genotypes in each SNP ( $U=4.500$ ;  $p=0.845$  and  $U=5.500$ ;  $p=0.858$ , respectively). To rs1331639 (CC ( $n=2$ ), CT ( $n=3$ ) and TT ( $n=2$ )) we used Kruskal-Wallis test ( $p=0.978$ ). Distribution of each genotype is condensed in figure S3.



**Figure 2.** CAG repeat length distribution of mutant alleles among genotypes of rs7270898 (a) and rs7275079 (b) in early-onset group. There is a tendency of longer CAG repeat length to be associated with GG and CC genotypes ( $69 \pm 14.14$  repeats), while the heterozygous genotypes (AG - rs7270898 and CT - rs7275079) showed lower average ( $46 \pm 5.26$  repeats).  $p=0.074$  for both rs7270898 and rs7275079 - Kruskal-Wallis test. (AA n=1; AG n=7; GG n=2; TT n=1; CT n=7; CC n=2.).

## DISCUSSION

This is the first study in a Brazilian population to assess the influence of genetic modifiers in modulate AO. Our results confirm association of *HAP1* M441 genotype to variation in HD AO ( $p=0.044$ ), even though we were not able to confirm the delay caused by this genotype reported by the original study (Metzger *et al*, 2008). A tight binding of HAP1-M441 to mutant htt when compared to T441 has been reported previously and this protein change does not alter its interaction of wild-type htt and with other interacting proteins (Metzger *et al*, 2008). HAP1-M441 colocalizes with htt aggregates and seems to be responsible for stabilizing htt aggregates as well as reducing the amount of soluble degraded htt products and apoptosis reduction. Results presented here confirm that *HAP1* is a powerful genetic modifier of HD and further studies to understand the molecular mechanism of HD responsible for the delay of AO must be performed.

We have not found association of tested SNPs in *SGK1* and *GRM1* genes and AO. Therefore additional SNPs remain to be studied in these genes in order to provide a more extensive view of a possible influence of this region on AO.

We have observed a tendency of genotypes GG and CC in both rs7270898 and rs7275079, respectively, to longer CAG repeat length ( $p=0.074$ ) in the early-onset group. However, we were not able to confirm *TGM2* as genetic modifier of HD, according to the evaluated SNPs. The transglutaminase cross linking activity  $\text{Ca}^{+2}$ -dependent between the carboxamide (terminal amide) of a glutamine (Q) residue in a protein substrate with the  $\epsilon$  amino group of a lysyl (K) residue of another protein is responsible for this function (Cooper *et al*, 2002). In addition, TGM2 seems to act as a repressor of target genes by changing chromatin conformation. It is critical for essential genes to energetic production,

as *PGC-1 $\alpha$*  and *COX-1* that contributes to mitochondrial and metabolic dysfunction characteristics of HD (Zainelli *et al*, 2005). However, this protein is highly conserved across species and this gene is low polymorphic (Kiraly *et al*, 2011).

*TGM2* is a good modifier candidate regarding its involvement in the formation of protein aggregates in several neurodegenerative diseases including HD. We have observed a tendency of genotypes GG and CC in both rs7270898 and rs7275079 to longer CAG repeat length ( $p=0.074$ ) in the early-onset group. However, we were not able to confirm *TGM2* as genetic modifier of HD, according to the evaluated SNPs. The transglutaminase cross linking activity  $\text{Ca}^{+2}$ -dependent between the carboxamide (terminal amide) of a glutamine (Q) residue in a protein substrate with the  $\epsilon$  amino group of a lysyl (K) residue of another protein is responsible for this function (Cooper *et al*, 2002). In addition, *TGM2* seems to act as a repressor of target genes by changing chromatin conformation. It is critical for essential genes to energetic production, as *PGC-1 $\alpha$*  and *COX-1* that contributes to mitochondrial and metabolic dysfunction characteristics of HD (Zainelli *et al*, 2005). However, this protein is highly conserved across species and this gene is low polymorphic (Kiraly *et al*, 2011). This fact can be difficult identification of *TGM2* as a genetic modifier by genotypic analysis.

Lack of association of other tested variables may be due to some of the following issues. As we have included specific SNPs, we cannot rule out that additional polymorphic regions in these genes could be responsible for a modifier effect in HD patients. These SNPs were chosen based on GWAS performed in a population with different ethnic background; therefore we might be studying genomic regions that are not so relevant for our mixed population. The other issue is related to the limited number of patients included

in these analyzes which can result in lack of power to detect genetic factor with low penetrance. We can then conclude that further studies should be performed including additional SPNs in these genes or extending the analysis to other regions of the genome in a larger population with similar ethnic background to address possible genetic modifiers of HD.

## **ACKNOWLEDGMENTS**

This research was supported by the following Brazilian funding agencies: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Rio Grande do Sul (FAPERGS), and Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES).

**Conflict of interest:** We declare that there are no conflicts of interests.

## REFERENCES

- Chattopadhyay, B, Ghosh, S, Gangopadhyay, PK, Das, SK, Roy, T, Sinha, KK, Jha, DK, Mukherjee, SC, Chakraborty, A, Singhal, BS, *et al.* 2003. Modulation of age at onset in Huntington's disease and spinocerebellar ataxia type 2 patients originated from eastern India. *Neurosci Lett* 345(2):93-6.
- Cooper, AJ, Jeitner, TM, Gentile, V, Blass, JP. 2002. Cross linking of polyglutamine domains catalyzed by tissue transglutaminase is greatly favored with pathological-length repeats: does transglutaminase activity play a role in (CAG)(n)/Q(n)-expansion diseases? *Neurochem Int* 40(1):53-67.
- Downey, PM, Petro, R, Simon, JS, Devlin, D, Lozza, G, Veltri, A, Beltramo, M, Bertorelli, R, Reggiani, A. 2009. Identification of single nucleotide polymorphisms of the human metabotropic glutamate receptor 1 gene and pharmacological characterization of a P993S variant. *Biochem Pharmacol* 77(7):1246-53.
- Gayán, J, Brocklebank, D, Andresen, JM, Alkorta-Aranburu, G, Zameel Cader, M, Roberts, SA, Cherny, SS, Wexler, NS, Cardon, LR, Housman, DE. 2008. Genomewide linkage scan reveals novel loci modifying age of onset of Huntington's disease in the Venezuelan HD kindreds. *Genet Epidemiol* 32(5):445-53.
- Hayden, MR, Kremer, B. 2001. Huntington Disease. In: Charles R. Scriver ALB, William S. Sly, David Valle, editor. *Metabolic and Molecular Basis of Inherited Disease*. 8th edition ed. p 5677-5701.
- Kiraly, R, Barta, E, Fesus, L. 2011. Polymorphism of transglutaminase 2: unusually low frequency of genomic variants with deficient functions. *Amino Acids*.
- Li, JL, Hayden, MR, Almqvist, EW, Brinkman, RR, Durr, A, Dode, C, Morrison, PJ, Suchowersky, O, Ross, CA, Margolis, RL, *et al.* 2003. A genome scan for modifiers of age at onset in Huntington disease: The HD MAPS study. *Am J Hum Genet* 73(3):682-7.
- Li, JL, Hayden, MR, Warby, SC, Durr, A, Morrison, PJ, Nance, M, Ross, CA, Margolis, RL, Rosenblatt, A, Squitieri, F, *et al.* 2006. Genome-wide significance for a modifier of age at neurological onset in Huntington's disease at 6q23-24: the HD MAPS study. *BMC Med Genet* 7:71.
- MacDonald, M, Ambrose, CM, Duyao, MP, Myers, RH, Lin, C, Srinidhi, L, Barnes, G, Taylor, SA, James, M, Groot, N *et al.* 1993. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. *Cell* 72(6):971-83.
- Metzger, S, Rong, J, Nguyen, HP, Cape, A, Tomiuk, J, Soehn, AS, Propping, P, Freudenberg-Hua, Y, Freudenberg, J, Tong, L, *et al.* 2008. Huntingtin-associated protein-1 is a modifier of the age-at-onset of Huntington's disease. *Hum Mol Genet* 17(8):1137-46.

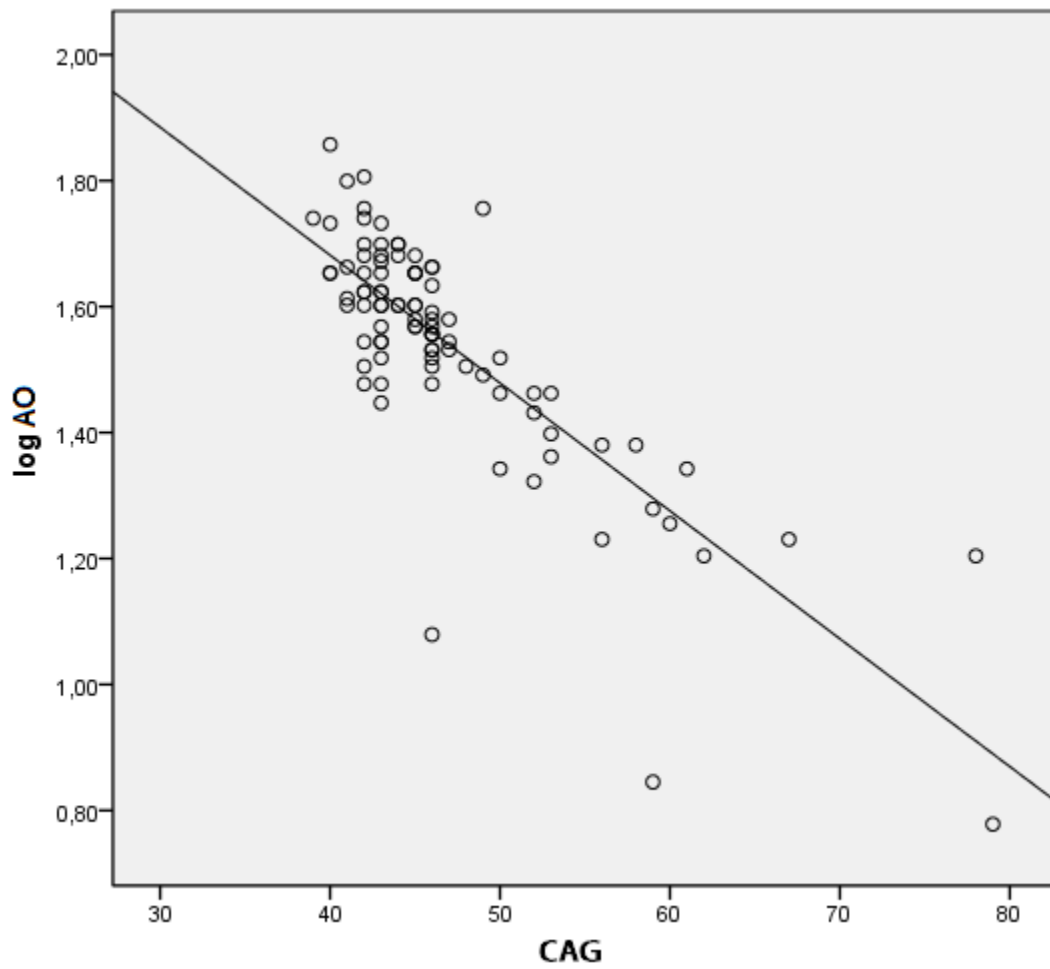
- Miller, SA, Dykes, DD, Polesky, HF. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16(3):1215.
- Naze, P, Vuillaume, I, Destee, A, Pasquier, F, Sablonniere, B. 2002. Mutation analysis and association studies of the ubiquitin carboxy-terminal hydrolase L1 gene in Huntington's disease. *Neurosci Lett* 328(1):1-4.
- Rosenblatt, A, Brinkman, RR, Liang, KY, Almqvist, EW, Margolis, RL, Huang, CY, Sherr, M, Franz, ML, Abbott, MH, Hayden, MR, *et al.* 2001. Familial influence on age of onset among siblings with Huntington disease. *Am J Med Genet* 105(5):399-403.
- Wexler, NS, Lorimer, J, Porter, J, Gomez, F, Moskowitz, C, Shackell, E, Marder, K, Penchaszadeh, G, Roberts, SA, Gayan, J, *et al.* 2004. Venezuelan kindreds reveal that genetic and environmental factors modulate Huntington's disease age of onset. *Proc Natl Acad Sci U S A* 101(10):3498-503.
- Zainelli, GM, Dudek, NL, Ross, CA, Kim, SY, Muma, NA. 2005. Mutant huntingtin protein: a substrate for transglutaminase 1, 2, and 3. *J Neuropathol Exp Neurol* 64(1):58-65.



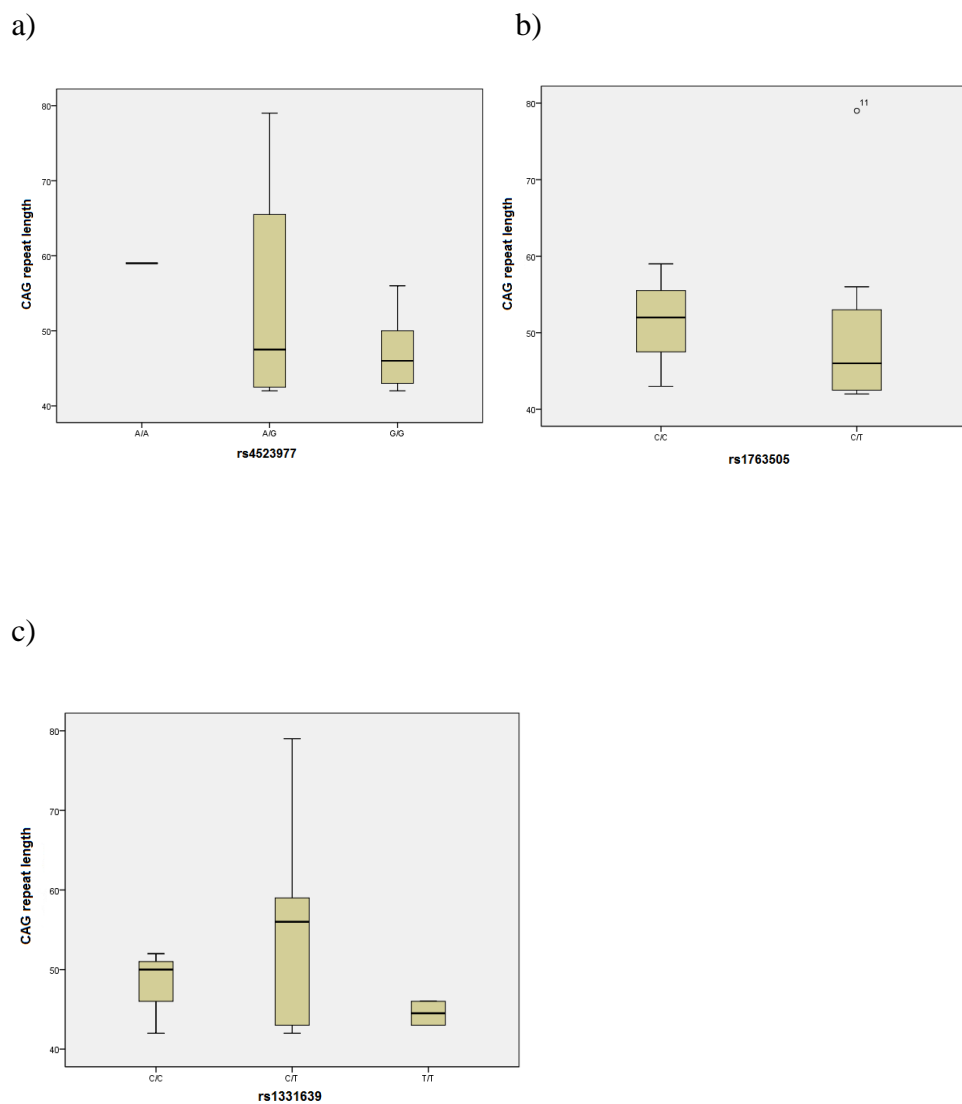
**SUPPLEMENTARY MATERIAL**

**Table S1.** SNPs included in this study and probes sequences

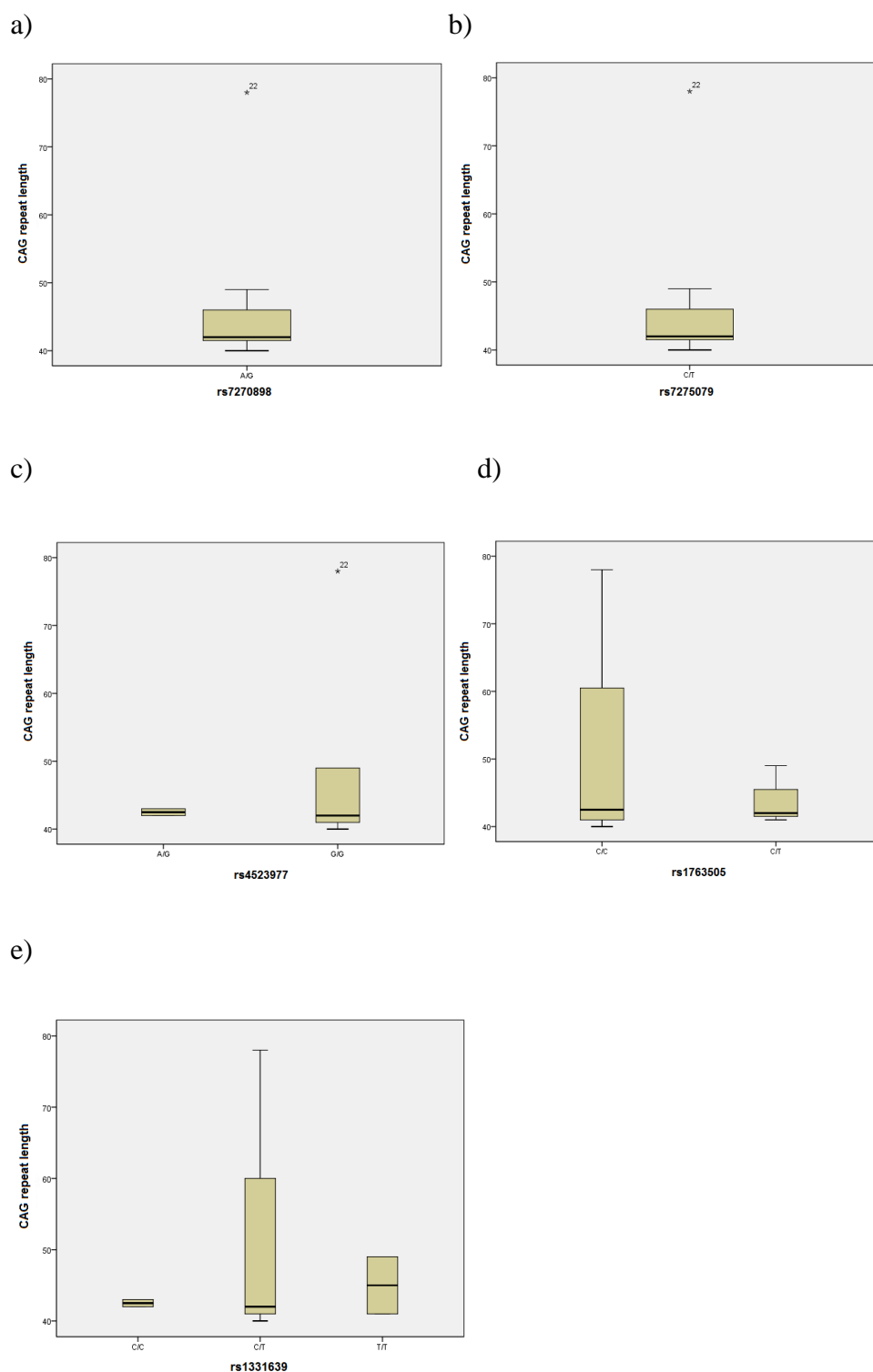
<b>Gene</b>	<b>Chromosome location</b>	<b>dbSNP ID</b>	<b>Genomic location</b>	<b>Chromosome position</b>	<b>Gene location</b>	<b>Probes sequences (VIC/FAM)</b>
<i>GRM1</i>	6q24.3	rs1331639	g.50601326C>T	chr6:146431869-146431869	Intron 2	TATTTAAGGCATTTATTTTGTATTT[C/T]CTCATAGGCTGTTTTGTATGGTTAT
<i>GRM1</i>	6q24.3	rs6923492	g.50924781T>C	chr6:146755324-146755324	Exon 8	CGCGGCGACCACTCCGCCTCTGCCG[C/T]CCCACCTGACCGCAGAGGAGACCCC
<i>HAPI</i>	17q21.2	rs4523977	g.5157502G>A	chr17:39883350-39883350	Exon 9	CAAGAACCACCCAAACCTAAGTGTT[C/T]TGAGGCCAGGCCAGCACCCCCTGCA
<i>SGK1</i>	6q23.2	rs1763505	g.38750575T>C	chr6:134581118-134581118	Intron 1	GTAAATCAGCAAACATGAGCAATAA[C/T]GATTTATATTTGCATGGCATTCTAT
<i>SGK1</i>	6q23.2	rs1075427	g.38673661C>T	chr6:134504204-134504204	Intron 7	ATCGTGCTTTTTTCTAGGTCAATAAG[C/T]TTATCCTGCAGGAATGAGAAGAGGA
<i>TGM2</i>	20q11.23	rs7275079	g.6989337C>T	chr20:36793245-36793245	Intron 2	CAAGAACCACCCAAACCTAAGTGTT[C/T]TGAGGCCAGGCCAGCACCCCCTGCA
<i>TGM2</i>	20q11.23	rs7270898	g.6987920G>A	chr20:36791828-36791828	Intron 2	AAGTGCCTTGTATGTGAATTATCTC[A/G]ACTTAAATCCCAAAACAACCTCCGCT



**Figure S1.** Simple linear regression graph between the independent variable CAG repeat length (CAG) and the dependent variable AO after a logarithmic transformation (logAO). ( $R^2=0.653$ ;  $p<0.0001$ ).



**Figure S2.** Box spot graphs of CAG repeat length distribution in the mutant allele among genotypes of a) rs4523977 (AA=1, AG=4, GG=5;  $p=0.444$ ), b) rs1763505 (CC=3, CT=7;  $p=0.491$ ) and c) rs1331639 (CC=3, CT=5, TT=2;  $p=0.625$ ) in early-onset group. Kruskal-Wallis and Mann-Whitney tests with statistical significance to  $p<0.05$ .



**Figure S3.** Box spot graphs of CAG repeat length distribution among genotypes in assessed SNPs in late-onset group. There was not statistical difference in distribution of CAG repeat length among genotypes to each SNP. rs7270898,

AG=7 ( $47.8 \pm 13.6$ ); rs7275079, CT=7 ( $47.8 \pm 13.6$ ); rs4523977, AG=2 ( $42.5 \pm 0.7$ ) e GG=5 ( $50.0 \pm 16.0$ ),  $p=0.845$ ; rs1763505, CC=4 ( $50.7 \pm 18.2$ ) and CT=3 ( $44.0 \pm 4.3$ ),  $p=0.858$ ; rs1331639, CC=2 ( $42.5 \pm 0.7$ ), CT=3 ( $53.3 \pm 21.3$ ), TT=2 ( $45.0 \pm 5.6$ ),  $p=0.978$ . (SNP, sample to each genotype (average  $\pm$  SD),  $p$  value). Kruskal-Wallis and Mann-Whitney tests with statistical significance to  $p<0.05$ .

## 4.2 Artigo 2

### **HUNTINGTON'S DISEASE: POLY-Q REPEAT VARIATION IN *ATXN1* AS A POTENTIAL MODIFIER OF AGE AT ONSET**

Este manuscrito será submetido para a revista *Movement Disorders*.

## HUNTINGTON'S DISEASE: POLY-Q REPEAT VARIATION IN *ATXN1* AS A POTENTIAL MODIFIER OF AGE AT ONSET

Aline Francielle Damo Souza BSc<sup>1,2</sup>; Tailise Conte Gheno MSc<sup>1</sup>; Gabriel Vasata Furtado MSc<sup>1</sup>; Carlos Rieder MD<sup>3</sup>; Laura Bannach Jardim PhD<sup>1,4</sup>; Maria Luiza Saraiva-Pereira PhD<sup>1,2,5</sup>

Genetics Identification Laboratory<sup>1</sup> – Experimental Research Center and Medical Genetics Service, Hospital de Clínicas de Porto Alegre (HCPA); Cellular and Molecular Biology<sup>2</sup> – Universidade Federal do Rio Grande do Sul (UFRGS); Neurology Service – HCPA<sup>3</sup>; Internal Medicine Department – UFRGS<sup>4</sup>; Department of Biochemistry – UFRGS<sup>5</sup>.

### **Corresponding author:**

Maria Luiza Saraiva-Pereira, PhD  
Medical Genetics Service  
Hospital de Clínicas de Porto Alegre  
Rua Ramiro Barcelos, 2350  
CEP 90035-903 – Porto Alegre – Brazil  
Tel: +55 51 3359-8011  
Fax: +55 51 3359-8010  
e-mail: mlpereira@hcpa.ufrgs.br

**Acknowledgments:** This research was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de amparo à pesquisa do Rio Grande do Sul (FAPERGS) and Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES).

**Conflict of interest:** We declare that there are no conflicts of interests.

Title: HUNTINGTON'S DISEASE: POLY-Q REPEAT VARIATION IN *ATXN1* AS A POTENTIAL MODIFIER OF AGE AT ONSET (93 characters)

Running title: *ATXN1* as a potential HD age at onset modifier

Number of words in the manuscript: 3.663

Number of figures: 3 figures

Number of tables: 4 tables



## ABSTRACT

**Background:** A number of human genetic neurodegenerative diseases are caused by an expansion of CAG repeat tract in disease causative gene region. A common feature among all these trinucleotide repeat diseases is strong inverse correlation between CAG repeat length and age at disease onset. Despite of it, individuals with the same number of CAG repeats show high variation in disease onset pointing to the occurrence of other factors that modulate age at disease onset. **Purpose:** We hypothesized that long normal alleles in other genes associated to neurodegenerative diseases caused by CAG expansion can be associated to premature disease onset in HD. **Methods:** We assessed here influence of normal *HTT* allele and long normal alleles in *ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7*, *PPP2R2B*, *TBP* and *DRPLA* genes in 89 unrelated Brazilian patients with Huntington's disease. **Results:** CAG repeat length in mutant allele was responsible for ~58% of age at disease onset variance. Of the genetic factors tested, only normal CAG repeats in *ATXN1* long allele showed a modest but significant correlation with AO ( $p=0.045$ ) explaining 4.5% of the residual variation in age at disease onset. **Conclusion:** Further studies are required to confirm *ATXN1* gene as genetic modifier of age at Huntington's disease onset.

**Keywords:** Huntington's disease, *HTT* gene, CAG repeats, long normal alleles, trinucleotide repeat diseases, age at onset.

## INTRODUCTION

Huntington's disease (HD) (OMIM 143100) is a neurodegenerative disorder caused by an expansion of CAG trinucleotide repeat above 36 units in the first exon of *HTT* gene which codes for a mutant huntingtin (htt) protein<sup>1</sup>. HD is one of nine neurodegenerative diseases caused by an expansion of CAG repeat located in gene coding region. In addition to HD, these disorders include spinal and bulbar muscular atrophy (SBMA), spinocerebellar ataxias (SCAs) 1, 2, 3, 6, 7, 17 and dentatorubral-pallidoluysian atrophy (DRPLA) and are known as polyglutamine diseases. Spinocerebellar ataxia type 12 is caused by expanded CAG repeats that, unlike polyglutamine diseases, are located in 5' non-coding of *PPP2R2B* gene (Table 1). The molecular mechanisms involved in these disease developments are under extensive investigation.

**Table 1.** Human diseases caused by CAG expanded repetitions

<b>Disease</b>	<b>MIM</b>	<b>Locus</b>	<b>Gene</b>	<b>Protein</b>	<b>Type of mutation</b>
HD	143100	4p16.3	<i>HTT</i>	Huntingtin	CAG expansion more than 36 repeats
SBMA	313200	Xq12	<i>AR</i>	Androgen receptor	CAG expansion ranging from 38 to 62 repeats
SCA1	164400	6p22.3	<i>ATXN1</i>	Ataxin 1	CAG expansion ranging from 39 to 81 pure repeats
SCA2	183090	12q24.12	<i>ATXN2</i>	Ataxin 2	CAG expansion more than 32 pure repeats
SCA3	109150	14q32.12	<i>ATXN3</i>	Ataxin 3	CAG expansion ranging from 54 to 86 repeats
SCA6	183086	19p13.2	<i>CACNA1A</i>	Calcium channel, voltage dependent P/Q type, $\alpha 1A$ subunit	CAG expansion ranging from 21 to 33 repeats
SCA7	164500	3p14.1	<i>ATXN7</i>	Ataxin 7	CAG expansion ranging from 36 to 450 repeats
SCA12	604326	5q32	<i>PPP2R2B</i>	Protein phosphatase 2, regulatory subunit B, $\beta$	CAG expansion in 5'UTR more than 51 repeats
SCA17	607136	6q27	<i>TBP</i>	TATA-box binding protein	CAG/CAA expansion ranging from 43 to 66 repeats
DRPLA	125370	12p13.31	<i>DRPLA</i>	Atrophin 1	CAG expansion ranging from 48 to 93 repeats

A common feature among all these polyglutamine diseases is strong inverse correlation between CAG repeat length and age at disease onset (AO). In HD, the CAG repeat length explains 42-73% of AO variation, suggesting that there can be *cis*- and *trans*-acting genetic factors, non-allelic genetic modifiers, stochastic and environmental factors influencing age of onset (AO) in addition to the pathogenic allele itself <sup>2</sup>.

Another shared characteristic among CAG repeat diseases is a genetic phenomenon called genetic anticipation, in which disease severity increases and/or AO decreases from one generation to next. Unlike static mutations, which are retained in somatic tissues and stably transmitted to offspring, the repeat mutation is dynamic, with a continued mutation process across generations. Longer stretches are more likely to undergo expansion than shorter ones. The molecular mechanism that leads to instability of dynamic mutations is not totally clear up to date. DNA instability could be caused by expansion itself or in assembly with gene-specific *cis*-elements located near the expansion and *trans*-acting DNA proteins that act on DNA replication and repair <sup>2</sup>.

Associations between long normal alleles in gene-specific *cis*-elements and phenotypic modulation of neurodegenerative diseases have already been described. Association between long normal alleles of *ATXN2* (<29 CAG repeats) and increased risk to ALS (Amyotrophic Lateral Sclerosis) has been shown previously <sup>3</sup>. Similarly, presence of long normal CAG repeats in *CACNA1A* gene (SCA6) was associated to earlier AO than expected in spinocerebellar ataxia type 2 (SCA2) <sup>4</sup>.

Therefore, the number of CAG repeats in the normal *HTT* allele and long normal CAG repeats in genes responsible for CAG expanded diseases can affect variation of AO observed in patients with Huntington's disease. Genes associated to spinocerebellar ataxias

such as *ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7*, *PPP2R2B*, *TBP*, and *DRPLA* genes were chosen as *cis*-acting candidates to genetic modifiers of AO in HD patients.

## SUBJECTS AND METHODS

In the present study, 89 unrelated Brazilian HD patients were included. These patients were assessed clinically and HD was confirmed by molecular diagnosis. AO of each patient was defined as the date when the patient, or a close relative, noticed the first symptom, either neurologic or motor manifestation. Patients have signed an informed consent term to participate in the research. In addition to HD patients, 360 unrelated normal subjects were included as controls.

DNA isolation was performed from a peripheral blood sample collected in EDTA (*ethylene diamine tetracetic acid*) by standard procedure<sup>5</sup>.

Mutant and normal *HTT* allele, *ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7*, *PPP2R2B*, *TBP* and *DRPLA* polymorphic CAG region were analyzed by PCR (Polymerase Chain Reaction) followed by capillary electrophoresis in *ABI3130xl* Genetic Analyzer (Applied Biosystems) and results were processed by GeneMapperID<sup>®</sup> software 4.0 (Applied Biosystems). A list of primers used in this study is available in supplementary materials Table S1.

PCR to evaluate CAG repeats in the *HTT* gene was performed to a final volume of 25  $\mu$ L, containing 40 ng of genomic DNA, 4 pmol of primers HuntF and HuntAR, 2 mM of dNTPs, 1,5mM of MgCl<sub>2</sub>, 1X *Gold Buffer GeneAmp*<sup>®</sup>, 10% (v/v) DMSO (dimethyl sulfoxide), and 1,5U/ $\mu$ L of *AmpliTaq Gold DNA polimerase*. Amplification was performed as follows: initial denaturation step and polymerase activation at 95°C for 10 minutes, followed by 5 cycles at 94°C for 30 seconds and 72°C for 3 minutes, 5 cycles at 94°C for 30 seconds and 70°C for 3 minutes, 5 cycles at 94°C for 30 seconds, 68°C for 1 minute and

72°C for 2 minutes and 20 cycles at 94°C for 30 seconds, 60°C for 1 minute and 72°C for 3 minutes and a final extension at 72°C for 20 minutes.

Amplification of *ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7* genes was performed as multiplex PCR in a final volume of 25 µL, containing 50 ng of genomic DNA, 20 pmol of primers Rep1 and Rep2, 10 pmol of primers SCA2A and SCA2B, 20 pmol of primers MJD52 and MJD25, 10 pmol of primers S-5-F1 and S-5-R1, 20 pmol of primers SCA7F and SCA7R, 2 mM of dNTPs, 1,5 mM MgCl<sub>2</sub>, 1X *Gold Buffer GeneAmp*<sup>®</sup>, 10% (v/v) DMSO (dimethyl sulfoxide), and 1,25U/µL of *AmpliTaq Gold DNA polimerase*. Amplification was performed as follows: initial denaturation step and polymerase activation at 95°C for 10 minutes, followed by 5 cycles at 94°C for 30 seconds and 72°C for 3 minutes, 5 cycles at 94°C for 30 seconds and 70°C for 3 minutes, 5 cycles at 94°C for 30 seconds, 68°C for 3 minutes, 25 cycles at 94°C for 30 seconds, 55°C for 1 minute e 68°C for 3 minutes and a final extension at 60°C for 60 minutes.

Amplification of *PPP2R2B*, *TBP*, and *DRPLA* genes was performed as multiplex PCR in a final volume of 25 µL, containing 100 ng of genomic DNA, 10 pmol of primers SCA12F and SCA12R, 20 pmol of primers SCA17F and SCA17R, 20 pmol of primers DRPLAF and DRPLAR, 2 mM dNTPs, 1.5 mM of MgCl<sub>2</sub>, 1X *Gold Buffer GeneAmp*<sup>®</sup>, 10% (v/v) DMSO (dimethyl sulfoxide), and 1.5 U/µL of *AmpliTaq Gold DNA polimerase*. Amplification was performed as follows: with a first step of denature and Taq DNA polymerase activation at 95°C by 10 minutes, followed by 5 cycles a 94°C for 30 seconds and 72°C for 3 minutes, 5 cycles a 94°C for 30 seconds and 70°C for 3 minutes, 5 cycles at 94°C for 30 seconds, 68°C for 1 minute, 72°C for 2 minutes and 20 cycles at 94°C for 30 seconds, 60°C for 1 minute and 72°C for 3 minutes and a final extension at 60°C for 60 minutes.

An aliquot of 0.5  $\mu$ L of each sample was mixed to 9.3  $\mu$ L of Hi-Di™ Formamide (Applied Biosystems) and 0.2  $\mu$ L of GeneScan™ 500 LIZ™ Size Standard and this mixture was submitted to capillary electrophoresis after denaturation. Samples were electroinjected during 22 seconds at 3 kV in the Genetic Analyzer and electrophoresis were performed at 15 kV during 40 minutes at 60°C in a 50 cm long capillary containing the matrix™ Polymer (Applied Biosystems).

Statistical analyses were performed using SPSS Statistics 18 (Statistical Package for Social Sciences, version 18). CAG tract length was determined for all nine loci – *HTT*, *ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7*, *PPP2R2B*, *TBP*, and *DRPLA* loci. Allele distribution was compared between cases and controls by Mann-Whitney test, which compares distribution around mean value of alleles in each group. Mean value was determined considering both individuals alleles indistinctly in cases and in controls. Allele frequency to each gene was assessed dividing samples of cases and controls in two groups: short allele and long allele according to mean value. Short group consisted of all alleles shorter than mean value. Long group contained alleles longer than median value one unit, excluding median value. Frequencies of alleles in each group were compared to cases and controls by Chi-square test.

To differentiate each measured CAG tract, we refer to them as the short and the long allele of a particular gene (*HTT*, *ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7*, *PPP2R2B*, *TBP* or *DRPLA*). A correlation between AO and length of mutant CAG repeat was estimated using simple linear regression model with logarithmic transformation of AO for a well fit in determination coefficient ( $R^2$ ). Influence of normal CAG repeat length in *HTT*, *ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7*, *PPP2R2B*, *TBP* and *DRPLA* genes in AO was assessed by multiple linear regressions. ANOVA test was used to compare average AO in

patients carrying no long allele, one long allele, and two long alleles. Size of normal CAG repeats of *HTT*, *ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7*, *PPP2R2B*, *TBP*, and *DRPLA* (DRPLA) genes were joined to the CAG x AO regression model to assess the influence of genes on AO in HD. We assessed the influence of long allele regarding the presence of short allele, sum of both alleles, and the importance of each allele individually. The regression models tested, the determination value ( $R^2$ ) as well as variation in determination coefficient ( $\Delta R^2$ ), and P-values are presented in the table 3.



## RESULTS

### Allele distribution in cases and controls

Mean size of mutant allele in HD patients was 46.6 ( $46.6 \pm 6.5$  repeats), ranging from 39 to 78 CAG repeats while mean size of normal allele was 18.3 ( $18.3 \pm 3.4$  repeats) CAG repeats, ranging from 10 to 30. AO ranged from 7 to 72 years of age.

In *ATXN1* gene, normal alleles of HD patients ranged from 18 to 36 CAG repeats and allele with 30 repeats was the most frequent in both cases and controls groups (37%). Short alleles ranged from 18 to 32 and long alleles from 21 to 36.

In *ATXN2* gene, normal alleles ranged from 17 to 32 CAG repeats being the allele with 22 repeats the most frequent in both cases and controls (80%). Short alleles ranged from 17 to 24 and long alleles from 22 to 32.

In *ATXN3* gene, normal alleles ranged from 14 to 42 CAG repeats. Normal allele with 23 repeats was the most frequently identified in controls (23%) and normal allele with 14 repeats was the most frequently found in cases (22%). Short alleles ranged from 14 to 29 and long alleles from 14 to 42.

Normal alleles in the *CACNA1A* gene ranged from 4 to 15 CAG repeats being the allele with 11 repeats the most frequent (42%). Short alleles ranged from 4 to 15 and long alleles from 7 to 15 CAG repeats.

Normal alleles ranged from 7 to 15 CAG repeats in *ATXN7* gene, being the allele with 10 repeats the most frequently observed in both cases and controls (54%). Short alleles ranged from 7 to 12 and long alleles from 9 to 15 CAG repeats.

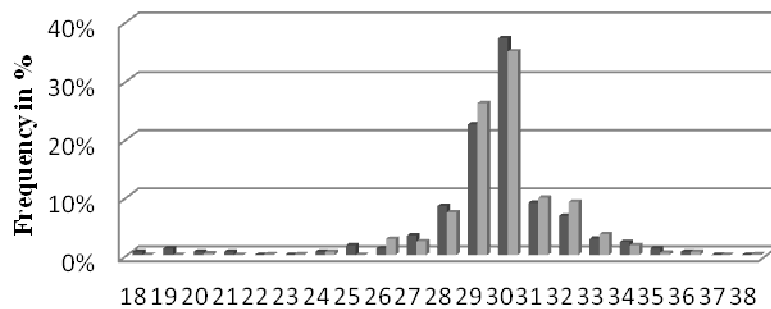
In *PPP2R2B* gene, normal alleles ranged from 6 to 17 CAG repeats being the allele with 10 the most frequent in controls (51% of controls) and allele 8 the most frequent in cases (43% of cases). Short alleles ranged from 6 to 14 and long alleles from 8 to 17. In *TBP*, normal alleles ranged from 25 to 38 CAG repeats being allele 35 the most frequent in controls (29% of controls) and allele 36 the most frequent in cases (29% in cases). Short alleles ranged from 25 to 36 and long alleles from 27 to 38. And in *DRPLA*, normal alleles ranged from 6 to 20 CAG repeats being allele 14 the most frequent in cases and controls (31% in cases). Short alleles ranged from 6 to 18 and long alleles from 8 to 20.

*ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7*, *PPP2R2B*, *TBP* and *DRPLA* allele distribution, as well as their frequencies in cases and in controls are shown in figure 1. Differences in allele distribution in cases and controls were significant to *ATXN2* ( $p=0.023$ ) gene, *ATXN7* gene ( $p=0.04$ ), and to *TBP* gene ( $p<0.001$ ). Frequency of short and long alleles in cases and in controls was assessed by Chi-square test. Mean dates, short and long allele frequencies,  $\chi^2$  value and P-value are presented in table 2. Statistically significant difference was observed between allele distribution in cases and controls in both *ATXN2* and *ATXN3* genes ( $\chi^2=5.053$ ;  $p=0.026$  and  $\chi^2=5.780$ ;  $p=0.017$ , respectively). Long allele was found more frequently in cases when compared to controls in these genes. Alleles with 24 CAG repeats in *ATXN3* gene, 9 CAG repeats in *ATXN7* gene, and 11 CAG repeats in *PPP2R2B* gene were visibly more frequent in the case group than in the control group. Fisher's exact test was used to compare frequencies, and allele with 24 CAG repeats in the *ATXN3* gene was significantly more frequent in cases ( $p=0.003$ ). Similar feature was also observed in *ATXN7* gene, where allele with 9 CAG repeats was more frequently found in cases ( $p<0.0001$ ), and in alleles with 11 CAG repeats in *PPP2R2B* gene ( $p<0.0001$ ). Alleles with 7, 8, and 12 CAG repeats in *PPP2R2B* gene were found exclusively in the

case group with frequencies of 0.04, 0.43, and 0.06, respectively. Analyzing the AO of individuals with these alleles, the average AO of allele with 24 CAG repeats of *ATXN3* was 39.8 years, of allele 9 of *ATXN7* gene was 38.8 years, of alleles 7, 8, 11, and 12 CAG repeats in *PPP2R2B* gene the averages were 38.2 years, 37.6 years, 38 years, and 36.6 years, respectively. The overall mean of AG in HD patients was 37.4 years. These results indicate that these alleles might participate in modulate HD.

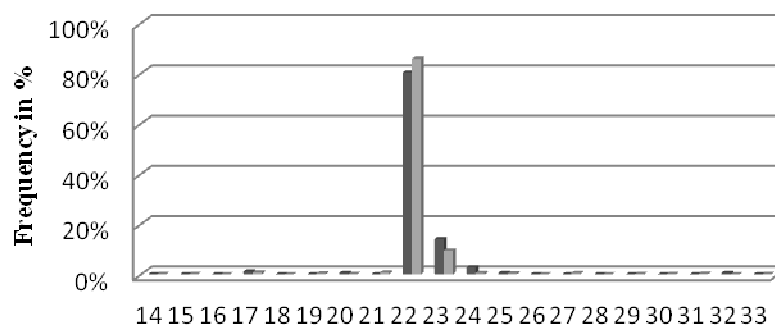
a)

### *ATXN1* alleles

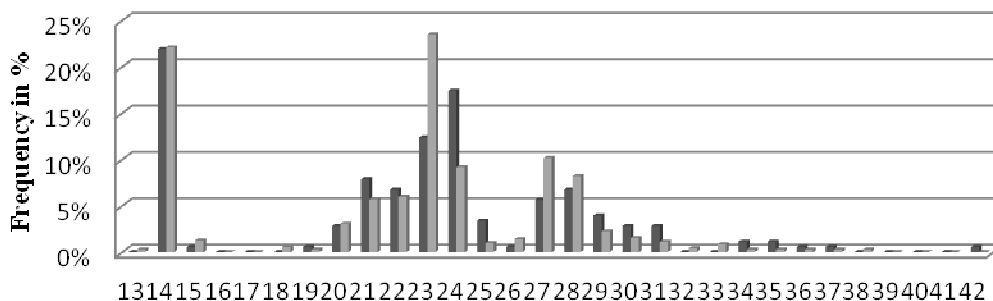


b)

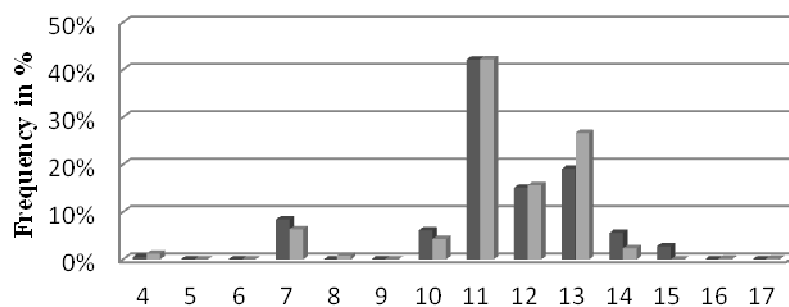
### *ATXN2* alleles



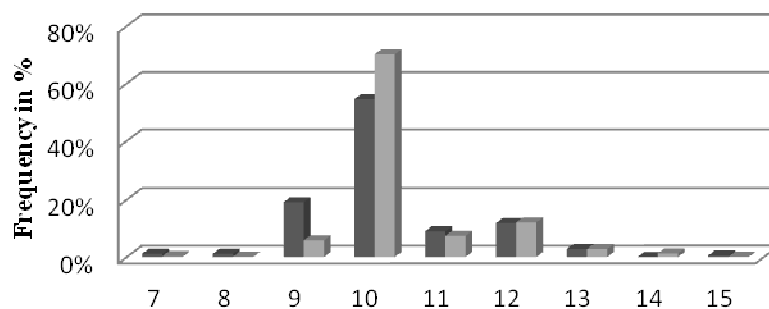
c)

**ATXN3 alleles**

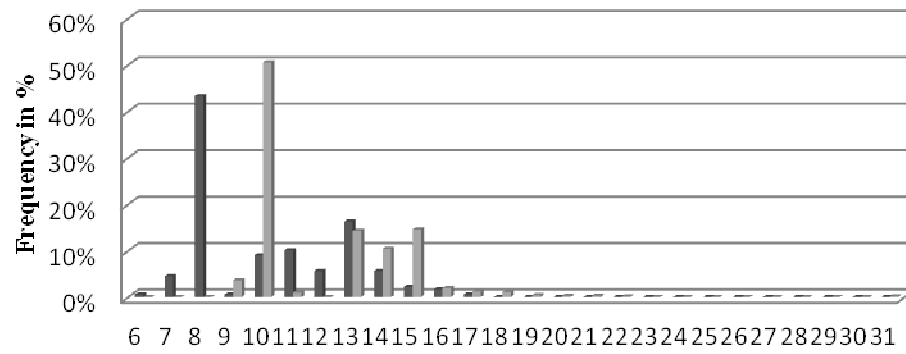
d)

**CACNA1A alleles**

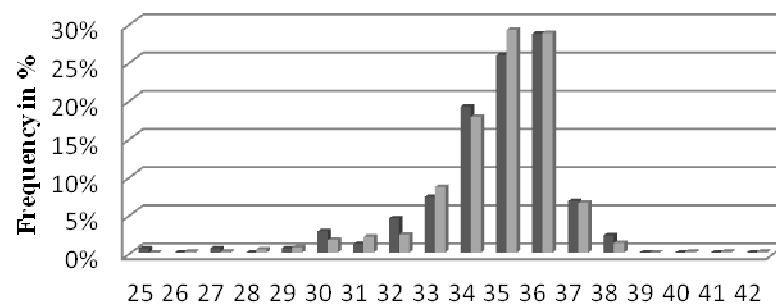
e)

**ATXN7 alleles**

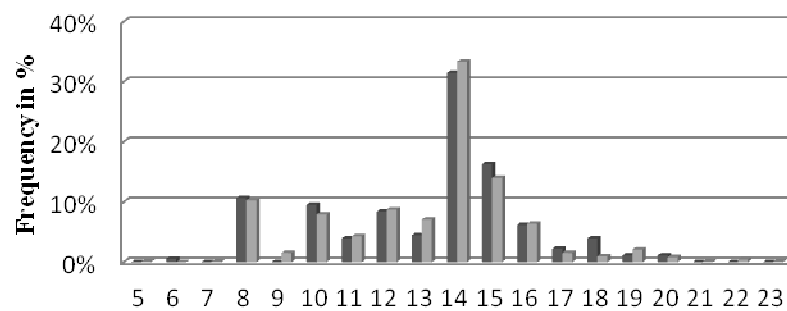
f)

**PPP2R2B alleles**

g)

**TBP alleles**

h)

**DRPLA alleles**

**Figure 1.** Allele distribution and frequency of (a) *ATXN1*, (b) *ATXN2*, (c) *ATXN3*, (d) *CACNA1A*, (e) *ATXN7*, (f) *PPP2R2B*, (g) *TBP* and (h) *DRPLA*. Black bars indicate patients with HD and gray bars normal individuals.

**Table 2.** Comparison between frequencies of short and long alleles in cases and in controls

Gene	Mean	Short Allele		Long Allele		$\chi^2$	<i>p</i> value <sup>a</sup>
		Cases	Controls	Cases	Controls		
<i>ATXN1</i>	30	138 (0.77)	537 (0.75)	40 (0.23)	183 (0.25)	0.663	0.440
<i>ATXN2</i>	22	146 (0.82)	636 (0.88)	32 (0.18)	84 (0.12)	5.053	0.026*
<i>ATXN3</i>	23	94 (0.53)	451 (0.63)	84 (0.47)	269 (0.37)	5.780	0.017*
<i>CACNA1A</i>	11	102 (0.57)	396 (0.55)	76 (0.43)	324 (0.45)	0.307	0.614
<i>ATXN7</i>	10	135 (0.76)	550 (0.76)	43 (0.24)	170 (0.24)	0.024	0.922
<i>PPP2R2B</i>	10	103 (0.58)	390 (0.54)	75 (0.42)	330 (0.46)	0.788	0.401
<i>TBP</i>	35	111 (0.62)	454 (0.63)	67 (0.38)	266 (0.37)	0.030	0.931
<i>DRPLA</i>	14	123 (0.69)	529 (0.73)	55 (0.31)	191 (0.27)	1.371	0.260

Frequencies are shown within parenthesis. <sup>a</sup>Chi-square test with statistical significance to  $p < 0.05$ .

### **Influence of CAG repeat length in mutant and normal *HTT* gene and in SCAs related genes**

CAG repeat number showed an inverse correlation with AO and explained 57.7% ( $R^2=0.577$ ;  $p < 0.0001$ ) of AO variability in our sample population (S1 figure). The equation for the regression model was:  $\log AO = a + b (CAG)$ , where  $a = 2.450$  and  $b = -0.019$ . According to this function, an increase in one unit in CAG repeats will decrease the AO in 0.95 years.

Our results indicate that normal allele in *HTT* gene is not associated to AO in this population ( $R^2=0.579$ ;  $p=0.526$ ). When long allele of *ATXN1* gene was assessed in presence and added with short allele, p-value indicated a possible correlation of CAG

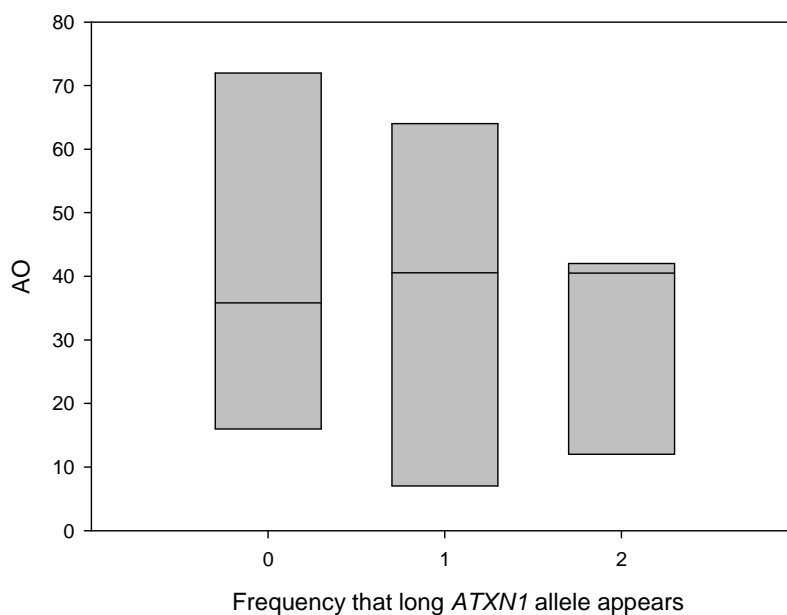
repeat length in *ATXN1* gene and AO ( $p=0.156$  and  $p=0.107$ ). When alleles were analyzed individually, a modest but significant association was observed between long allele alone and variation in AO ( $R^2=0.596$ ;  $p=0.045$ ) with an additional AO variance of 4.5%. Figure 2 shows AO averages according to the absence or presence of long allele in *ATXN1* gene. Presence of one or two long alleles in *ATXN1* seems to have a correlation with late AO when compared to the absence of long allele ( $p=0.203$ ). We observed that the presence of long allele tends to reduce the low value in the AO range (Figure 2). When analyzed the presence of short allele, the long allele of *CACNA1A* and of *ATXN7* genes showed a correlation tendency with AO, but this increase was not statistically significant ( $R^2=0.593$ ;  $p=0.083$  and  $R^2=0.591$ ;  $p=0.105$ ).

**Table 3.** Single and multiple linear regression models

Regression Model	Presence of both alleles			Alleles analyzed individually		
	R <sup>2</sup>	ΔR <sup>2</sup>	p <sup>a</sup>	R <sup>2</sup>	ΔR <sup>2</sup>	p <sup>a</sup>
CAG x AO	0.577	-	<0,0001*			
CAG x AO + <i>ATXN1</i>						
Short allele	0.590	0.013	0.833	0.580	0.003	0.413
Long allele			0.156	0.596	0.019	0.045*
Sum of both alleles	0.589	0.012	0.107			
CAG x AO + <i>ATXN2</i>						
Short allele	0.579	0.002	0.508	0.579	0.002	0.503
Long allele			0.938	0.577	0.000	0.712
Sum of both alleles	0.579	0.002	0.516			
CAG x AO + <i>ATXN3</i>						
Short allele	0.584	0.007	0.641	0.577	0.000	0.744
Long allele			0.227	0.580	0.003	0.413
Sum of both alleles	0.579	0.002	0.490			
CAG x AO + <i>CACNA1A</i>						
Short allele	0.593	0.016	0.510	0.579	0.002	0.484
Long allele			0.083	0.578	0.001	0.605
Sum of both alleles	0.577	0.000	0.814			
CAG x AO + <i>ATXN7</i>						
Short allele	0.591	0.014	0.700	0.578	0.001	0.626
Long allele			0.105	0.577	0.000	0.804
Sum of both alleles	0.578	0.001	0.679			
CAG x AO + <i>PPP2R2B</i>						
Short allele	0.577	0.000	0.719	0.577	0.000	0.717
Long allele			0.985	0.577	0.000	0.870
Sum of both alleles	0.577	0.000	0.777			
CAG x AO + <i>TBP</i>						
Short allele	0.583	0.006	0.292	0.582	0.005	0.310
Long allele			0.528	0.577	0.000	0.933
Sum of both alleles	0.579	0.002	0.555			
CAG x AO + <i>DRPLA</i>						
Short allele	0.577	0.000	0.994	0.577	0.000	0.979
Long allele			0.630	0.578	0.001	0.568
Sum of both alleles	0.577	0.000	0.747			

<sup>a</sup>Statistical significant level established in  $p < 0,05$ .





**Figure 2.** Distribution of AO according to frequency with which long CAG repeats appears in *ATXN1*: absence (0), one time (1) and two times (2). We observe that presence of long allele tend (1 and 2) to associate with later average AO (black line in boxes) compared with the absence of long CAG repeats (0). Averages AO: (0)=35.82; (1)=40.54; (2)=40.50 (ANOVA test:  $F=1.626$ ;  $p=0.203$ ).

## DISCUSSION

Unstable trinucleotide repeats are related to neurodegenerative diseases either as a cause of disease or as a modulation factor. The molecular mechanism of dynamic mutations is not completely clear, however the occurrence of *cis*-acting elements and *trans*-acting factors seems to be a suitable explanation for anticipation. *Cis*-acting elements are directly related to expanded repeat, such as copy number and repeat sequence. *Trans*-acting factors can contribute to instability of repeats by interacting with them.

HD and SCA1 are among nine neurodegenerative diseases caused by CAG repeat expansions located in gene coding region. There is no evidence of interaction between ataxin 1 and other poly-Q protein to date, but some evidences pointing to its potential participation in modulate other neurodegenerative diseases have been reported. Allele with 31 CAG repeats in *ATXN1* gene was found to be associated with HD and myotonic dystrophy type 1 (MD) in patients from Yugoslav, suggesting a *trans*-acting effect of this allele or its gene product on repeat instability in HD and MD <sup>6, 7</sup>, although another published work has not found the same association in German patients <sup>8</sup>. These controversial results might be due to differences in ethnic population and/or sample size. *ATXN1* normal alleles of patients with SCA1 were found to be more unstable than alleles of identical size from healthy people <sup>9, 10</sup>. *ATXN1* is located in chromosome 6p22-p23, a region previously identified as associated with AO modulation in HD, showing a potential influence of *ATXN1* in the AO variation observed in HD patients <sup>11</sup>.

In the present study we found a modest, but significant, correlation between the long normal allele of *ATXN1* gene and variation in AO of HD patients ( $p=0.045$ ). This correlation explained an additional AO variance of 4.5%. When we analyzed AO averages

regarding presence of one, two, or none long normal *ATXN1* allele, a tendency in increasing AO average with presence of one or two long alleles ( $p=0.203$ ) was observed. On the other hand, the presence of two long alleles tends to maintain AO under ~40 years, thus indicating a possible influence in reducing AO. However, further studies are required to validate these results.

The influence of normal *HTT* allele in modulate HD is yet controversial<sup>12</sup>. Although some studies have found association of AO variation with normal CAG repeats in *HTT* gene, this association was not found to our HD patients ( $p=0.526$ ). We have found some alleles that appear differentially in cases and in controls. Among them we highlight allele with 24 CAG repeats in *ATXN3* gene. This allele was the second most prevalent in HD patients when analyzing *ATXN3* gene (17%) and could be considered a long allele. It may have an important participation in HD disease.

Understanding modifiers factors of a disease is relevant for better comprehension of pathogenesis and it might facilitate the identification of environmental effects on phenotype. In summary, our results corroborate the hypothesis that CAG polymorphic repeats related to other neurodegenerative diseases can modulate HD phenotype. We have reported, for the first time, the influence of long alleles in *ATXN1* gene on AO variation in a group of HD patients.

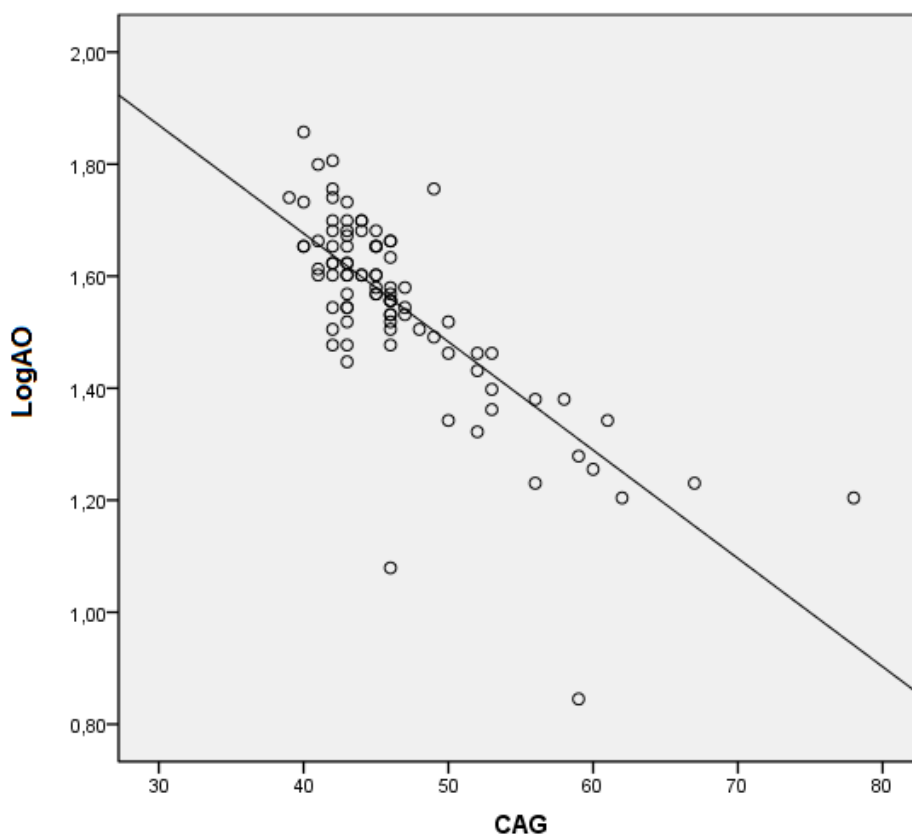
## REFERENCES

1. MacDonald, M, Ambrose, CM, Duyao, MP *et al.* A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. *Cell* 1993;72(6):971-983.
2. Pearson, C, Nichol, EK and Cleary, JD. Repeat instability: mechanisms of dynamic mutations. *Nat Rev Genet* 2005;6(10):729-742.
3. Daoud, H, Belzil, V, Martins, S *et al.* Association of long ATXN2 CAG repeat sizes with increased risk of amyotrophic lateral sclerosis. *Arch Neurol* 2011;68(6):739-742.
4. Pulst, SM, Santos, N, Wang, D *et al.* Spinocerebellar ataxia type 2: polyQ repeat variation in the CACNA1A calcium channel modifies age of onset. *Brain* 2005;128(Pt 10):2297-2303.
5. Miller, S, Dykes, DD and Polesky, HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988;16(3):1215.
6. Savic, D, Topisirovic, I, Keckarevic, M *et al.* Is the 31 CAG repeat allele of the spinocerebellar ataxia 1 (SCA1) gene locus non-specifically associated with trinucleotide expansion diseases? *Psychiatr Genet* 2001;11(4):201-205.
7. Keckarevic, D, Culjkovic, B, Savic, D *et al.* The status of SCA1, MJD/SCA3, FRDA, DRPLA and MD triplet containing genes in patients with Huntington disease and healthy controls. *J Neurogenet* 2000;14(4):257-263.
8. Hellenbroich, Y, Kaulich, M, Opitz, S, Schwinger, E and Zuhlke, C. No association of the SCA1 (CAG)<sup>31</sup> allele with Huntington's disease, myotonic dystrophy type 1 and spinocerebellar ataxia type 3. *Psychiatr Genet* 2004;14(2):61-63.
9. Koefoed, P, Hasholt, L, Fenger, K *et al.* Mitotic and meiotic instability of the CAG trinucleotide repeat in spinocerebellar ataxia type 1. *Hum Genet* 1998;103(5):564-569.
10. Richards, RI. Dynamic mutations: a decade of unstable expanded repeats in human genetic disease. *Hum Mol Genet* 2001;10(20):2187-2194.
11. Li, JL, Hayden, MR, Almqvist, EW *et al.* A genome scan for modifiers of age at onset in Huntington disease: The HD MAPS study. *Am J Hum Genet* 2003;73(3):682-687.
12. Aziz, N, van Roon-Mom, WM and Roos, RA. CAG repeat size in the normal HTT allele and age of onset in Huntington's disease. *Mov Disord* 2011;26(13):2450-2451; author reply 2451.

## SUPPLEMENTARY MATERIAL

**Table S1.** List of primers used in this study

Gene	Primer sequences 5'-3'	References
<i>HTT</i>	HuntF: GACCCTGGAAAAGCTGATGAAGGC HuntAR: TGGCGGCTGTTGCTGCTGCTG	
<i>ATXN1</i>	Rep1: AACTGGAAATGTGGACGTAC Rep2: CAACATGGGCAGTCTGAG	Orr <i>et al</i> , 1993
<i>ATXN2</i>	SCA2-A: GGGCCCCTCACCATGTCG SCA2-B: CGGGCTTGCGGACATTGG	Pulst <i>et al</i> , 1996
<i>ATXN3</i>	MJD52: CCAGTGACTACTTTGATTTCG MJD25: TGGCCTTTCACATGGATGTGAA	Kawaguchi <i>et al</i> , 1994
<i>CACNA1A</i>	S-5-F1: CACGTGTCCTATTCCCCTGTGATCC S-5-R1: TGGGTACCTCCGAGGGCCGCTGGTG	Zhuchenko <i>et al</i> , 1997
<i>ATXN7</i>	SCA7F: GAGCGGAAAGAATGTCCGAG SCA7R: CACGACTGTCCCAGCATCACTT	David <i>et al</i> , 1997
<i>PPP2R2B</i>	SCA12F: CCTCGCCTTTAATGCACCAGCC SCA12R: GCGCCAGCGCACTCACCCCTC	
<i>TBP</i>	SCA17F: CCCACAGCCTATTCAGAACACC SCA17R: GGGACGTTGACTGCTGAACGG	
<i>DRPLA</i>	DRPLAF: CCACCCACCAGTCTCAACACATC DRPLAR: GGAGGGAGACATGGCGTAAGG	



**Figure S1.** Simple linear regression graph between the independent variable CAG repeats length in mutant allele (CAG) and dependent variable AO after logarithmic transformation (LogAO). ( $R^2=0.577$ ;  $p<0.0001$ ).

## 5. DISCUSSÃO

O presente trabalho teve como principal objetivo a identificação de novos genes modificadores da idade de início da DH e a validação de *HAPI* como gene modificador para esta população. Para tal, 91 indivíduos com DH originários de seis diferentes estados do Brasil e da Argentina foram genotipados. A distribuição foi de 59 pacientes do Rio grande do Sul, 11 de São Paulo, 10 do Rio de Janeiro, 4 de Santa Catarina, 3 do Rio Grande do Norte e 1 da Bahia e 3 pacientes de Buenos Aires. Além desses pacientes, 100 indivíduos controle foram incluídos no estudo de SNPs e 360 para o estudo com genes associados a repetições CAG. O genótipo encontrado para cada variante polimórfica em cada indivíduo foi associado à idade de início do mesmo, considerando sempre o número de repetições CAG no gene causador da doença, buscando evidências de que tais variantes estão envolvidas na determinação do fenótipo idade de início.

A idade média de início dos sintomas encontrada nesta população foi de 37,3 anos ( $37,3 \pm 12,4$  anos). As médias do número de repetições CAG nos alelos do gene *HTT* foram 18,3 repetições CAG ( $18,3 \pm 3,4$  repetições CAG) para o alelo normal e 47 repetições CAG ( $47 \pm 7,6$  repetições CAG) para o alelo expandido. Para esta população, o número de repetições CAG explicou 65,3% ( $R^2=0,653$ ) da variabilidade da idade de início, mostrando uma forte influência na determinação da AO ( $p<0,0001$ ).

Para o primeiro grupo de resultados, quando os SNPs foram avaliados como um todo, não foi encontrada correlação estatisticamente significativa com a variação da idade de início, no entanto, na segregação por genótipo, o genótipo AA do rs4523977 mostrou-se correlacionado com a idade de início ( $p=0,044$ ), sendo a média de idade para este genótipo

de 27 anos ( $27 \pm 16,3$  anos). Este resultado corrobora os achados de Meztger sobre a associação do genótipo AA com a variação da idade de início, no entanto, não pudemos confirmar o atraso na manifestação dos sintomas causados por este genótipo, como reportado no estudo original, o qual foi realizado em mais de 900 indivíduos europeus (Metzger *et al*, 2008). Considerando a amostra analisada (91 pacientes, sendo apenas 4 homocigotos para M441) pelo fato de havermos encontrado associação, concluímos que *HAP1* é um excelente candidato a modificador da DH e deve ser melhor explorado para a compreensão do mecanismo molecular da DH e como este genótipo polimórfico pode atuar na modulação da idade inicial de manifestação da doença. Com base nos polimorfismos analisados, não fomos capazes de confirmar *SGK1* e *GRM1* como modificadores genéticos da idade de início, apesar de estes estarem localizados em regiões previamente descritas como associadas à DH por *GWAS*. Testamos apenas um SNP em cada um desses genes, uma vez que os outros dois escolhidos não estavam em EHW. Sugerimos que outros polimorfismos sejam testados nestes genes. *TGM2* é um bom candidato a modificador pelo seu envolvimento na formação de agregados proteicos nas principais doenças neurodegenerativas, incluindo a DH (Cooper *et al*, 2002), além de poder atuar como repressor de genes alvo mudando a conformação da cromatina. Tal fato é crítico para genes chave na produção energética, como *PGC-1 $\alpha$*  e *COX-1*, o que contribuiria para as disfunções mitocondriais e metabólicas características da DH (Zainelli *et al*, 2005). Entretanto, pelo fato de esta proteína ser altamente conservada entre as espécies, este gene é pouco polimórfico (Kiraly *et al*, 2011). Para rs7270898 e rs7275079, observamos uma tendência de associação entre os genótipos GG e CC com um número maior de repetições CAG no grupo precoce ( $p=0,074$ ). Estes genótipos podem estar associados a uma maior instabilidade das repetições CAG, levando à antecipação genética, com uma idade de início



mais precoce e à maior gravidade da doença. Por fim, o gene *TGM2* não foi confirmado como modificador genético da DH nesse estudo.

Para o segundo grupo de resultados, a análise de repetições CAG, dentro da faixa de normalidade, em outros genes que estão relacionados a outras doenças neurodegenerativas ou mesmo no alelo normal do gene *HTT* é bastante interessante com vistas à compreensão da ação de fatores genéticos em *cis*-elementos na modulação de uma doença. A influência do alelo normal de *HTT* na modulação da DH é bastante controversa (Aziz *et al*, 2011). Não foi encontrada associação deste com a variação da idade de início para esta população ( $R^2=0,655$ ;  $p=0,57$ ). Com a genotipagem de casos e controles para os genes relacionados às SCAs, observamos que os pacientes possuem uma tendência a apresentar alelos mais longos do que os controles. Este fato é bastante visível para *ATXN3* em que a partir do alelo 29, a frequência é maior nos casos, sendo o alelo 24 o mais frequente nos controles (figura 1 do artigo 2). Este é um indicativo de que existe um mecanismo compartilhado que leva à instabilidade em sequências repetitivas, que não apenas do gene causador da doença, ou que os genes contendo repetições polimórficas interajam entre si de alguma maneira ainda desconhecida. Apesar de termos encontrado diferenças de distribuição dos alelos para os genes *ATXN2* ( $p=0,023$ ), *ATXN7* ( $p=0,04$ ) e para *TBP* ( $p<0,001$ ), e a diferença de frequência dos alelos para os genes *ATXN2* e *ATXN3* ( $\chi^2=5,053$ ;  $p=0,026$  e  $\chi^2=5,780$ ;  $p=0,017$ , respectivamente) em casos e controles, foi o alelo longo de *ATXN1* que mostrou correlação moderada com a variação da idade de início. A presença deste alelo explicou 4,5% da variação da idade de início para esta população ( $p=0,045$ ). Esta foi a primeira vez que *ATXN1* foi apontado como modificador da DH. O alelo com 31 repetições no gene *ATXN1* foi associado com DH e distrofia miotônica (DM) em pacientes da Iugoslávia, sugerindo um efeito do elemento-*cis* alelo 31 ou do produto

gênico na instabilidade das repetições na DH e na DM (Keckarevic *et al*, 2000; Savic *et al*, 2001). Este resultado não foi confirmado por trabalho publicado com pacientes da Alemanha (Hellenbroich *et al*, 2004). Diferenças étnicas da população e o tamanho da amostra podem ser responsáveis pelos dados controversos encontrados. Os alelos normais de *ATXN1* de pacientes com SCA1 foram descritos como sendo mais instáveis do que alelos de tamanho idêntico em pessoas normais (Koefoed *et al*, 1998; Richards, 2001). *ATXN1* está localizado no cromossomo 6p22-p23, região esta previamente descrita por estudos de *GWAS* associada à modulação da idade de início na DH (Li *et al*, 2003).

Alguns alelos foram encontrados diferencialmente em casos e controles. Entre eles, destacamos o alelo 24 de *ATXN3*. Este alelo foi o segundo mais prevalente nos pacientes quando analisamos *ATXN3* (17% dos casos). Ele está entre os alelos normais longos e pode ter participação importante na patogênese da DH.

Em resumo, nossos resultados dão suporte à ideia de que repetições CAG polimórficas relacionadas a outras doenças neurodegenerativas podem modular o fenótipo da DH. Nesse trabalho foi demonstrada a influência de alelos no gene *ATXN1* com a idade de início da DH. Estudos *in vitro*, utilizando proteínas recombinantes, imunoprecipitação, espectrometria de massas, modelo celular de DH ou mesmo programas de predição de interações proteicas, podem ser válidas na identificação da real interação entre huntingtina e ataxina 1 como fatores de ação *trans*, participando na modulação genética e fenotípica da DH.

Discrepâncias entre os dados observados podem ter explicações diversas, entre elas as diferenças de origem genética ou ambiental das diferentes populações estudadas ou a falta de poder estatístico para detectar fatores genéticos de baixa penetrância.

Um gene é um modificador se a alteração de sua estrutura ou expressão acarreta uma alteração da manifestação do fenótipo associado com a mutação primária causadora da doença, neste caso a expansão CAG no gene *HTT*. A identificação de modificadores genéticos em humanos está atualmente limitada à natural ocorrência de variantes genéticas que são observadas na população como um todo. Uma vez identificada a variante genética capaz de alterar o fenótipo do indivíduo doente, este deve ser incorporado em desenhos de ensaios clínicos envolvendo tal fenótipo. As intervenções que estes ensaios clínicos são idealizados para testar são candidatos a modificadores, químicos ou outros, da patogênese da DH e a inclusão da genotipagem para importantes variantes genéticas, que também modifiquem a patogênese da DH, aumentará o poder do ensaio clínico na detecção de um efeito da intervenção aplicada, uma vez que controlará o efeito genético de fundo do paciente testado. Para um modificador genético ser utilizado como uma intervenção terapêutica por si só, é necessário definir, pelo menos em parte, o mecanismo de ação da variante genética que gera o efeito modificador para que se possa saber como utilizar as vias em que este modificador está inserido ou as proteínas envolvidas como alvo para desenvolvimento de abordagens terapêuticas. Isto irá requerer a utilização de modelos animais, avaliando a extrapolação das vias/ proteínas encontradas para o ser humano.

A principal esperança atualmente é o aperfeiçoamento da tecnologia de indução de células tronco pluripotentes (iPS - *induced pluripotent stem*) que promete fornecer células pluripotentes e produtos diferenciados a partir do material biológico de cada paciente (Camnasio *et al*, 2012). Esta estratégia representa um grande avanço tanto para definir o efeito de modificadores genéticos validados em células expressando o fenótipo da DH e também por permitir a realização de triagens genéticas a partir de células humanas, como por exemplo, estudos de RNA de interferência, de superexpressão, a fim de identificar

modificadores que não foram encontrados em estudos de *GWAS* por falta de variação funcional inerente à população em geral.

Desde a descoberta do gene relacionado à patogênese da DH, muitos avanços já foram alcançados nesta área e a descoberta e caracterização de modificadores genéticos é uma delas. Com a descoberta e o aperfeiçoamento de novas ferramentas na área da ciência experimental estes, que até então são apenas dados científicos, poderão tornar-se uma medida paliativa ou mesmo definitiva para doenças tão devastadoras como a DH.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

- Anborgh, PH, Godin, C, Pampillo, M, Dhami, GK, Dale, LB, Cregan, SP, Truant, R and Ferguson, SS (2005). "Inhibition of metabotropic glutamate receptor signaling by the huntingtin-binding protein optineurin." *J Biol Chem* 280(41): 34840-34848.
- Arrasate, M and Finkbeiner, S (2012). "Protein aggregates in Huntington's disease." *Exp Neurol* 238(1): 1-11.
- Arrasate, M, Mitra, S, Schweitzer, ES, Segal, MR and Finkbeiner, S (2004). "Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death." *Nature* 431(7010): 805-810.
- Aziz, N, van Roon-Mom, WM and Roos, RA (2011). "CAG repeat size in the normal HTT allele and age of onset in Huntington's disease." *Mov Disord* 26(13): 2450-2451; author reply 2451.
- Bailey, CD and Johnson, GV (2006). "The protective effects of cystamine in the R6/2 Huntington's disease mouse involve mechanisms other than the inhibition of tissue transglutaminase." *Neurobiol Aging* 27(6): 871-879.
- Bell, LM, Leong, ML, Kim, B, Wang, E, Park, J, Hemmings, BA and Firestone, GL (2000). "Hyperosmotic stress stimulates promoter activity and regulates cellular utilization of the serum- and glucocorticoid-inducible protein kinase (Sgk) by a p38 MAPK-dependent pathway." *J Biol Chem* 275(33): 25262-25272.
- Camnasio, S, Carri, AD, Lombardo, A, Grad, I, Mariotti, C, Castucci, A, Rozell, B, Riso, PL, Castiglioni, V, Zuccato, C, Rochon, C, Takashima, Y, Diaferia, G, Biunno, I, Gellera, C, Jaconi, M, Smith, A, Hovatta, O, Naldini, L, Di Donato, S, Feki, A and Cattaneo, E (2012). "The first reported generation of several induced pluripotent stem cell lines from homozygous and heterozygous Huntington's disease patients demonstrates mutation related enhanced lysosomal activity." *Neurobiol Dis* 46(1): 41-51.
- Chattopadhyay, B, Ghosh, S, Gangopadhyay, PK, Das, SK, Roy, T, Sinha, KK, Jha, DK, Mukherjee, SC, Chakraborty, A, Singhal, BS, Bhattacharya, AK and Bhattacharyya, NP (2003). "Modulation of age at onset in Huntington's disease and spinocerebellar ataxia type 2 patients originated from eastern India." *Neurosci Lett* 345(2): 93-96.
- Cooper, AJ, Jeitner, TM, Gentile, V and Blass, JP (2002). "Cross linking of polyglutamine domains catalyzed by tissue transglutaminase is greatly favored with

pathological-length repeats: does transglutaminase activity play a role in (CAG)(n)/Q(n)-expansion diseases?" *Neurochem Int* 40(1): 53-67.

Cooper, AJ, Wang, J, Pasternack, R, Fuchsbauer, HL, Sheu, RK and Blass, JP (2000). "Lysine-rich histone (H1) is a lysyl substrate of tissue transglutaminase: possible involvement of transglutaminase in the formation of nuclear aggregates in (CAG)(n)/Q(n) expansion diseases." *Dev Neurosci* 22(5-6): 404-417.

Cooper, AJ, Sheu, KR, Burke, JR, Onodera, O, Strittmatter, WJ, Roses, AD and Blass, JP (1997). "Transglutaminase-catalyzed inactivation of glyceraldehyde 3-phosphate dehydrogenase and alpha-ketoglutarate dehydrogenase complex by polyglutamine domains of pathological length." *Proc Natl Acad Sci U S A* 94(23): 12604-12609.

Daoud, H, Belzil, V, Martins, S, Sabbagh, M, Provencher, P, Lacomblez, L, Meininger, V, Camu, W, Dupre, N, Dion, PA and Rouleau, GA (2011). "Association of long ATXN2 CAG repeat sizes with increased risk of amyotrophic lateral sclerosis." *Arch Neurol* 68(6): 739-742.

Davies, SW, Beardsall, K, Turmaine, M, DiFiglia, M, Aronin, N and Bates, GP (1998). "Are neuronal intranuclear inclusions the common neuropathology of triplet-repeat disorders with polyglutamine-repeat expansions?" *Lancet* 351(9096): 131-133.

Downey, PM, Petro, R, Simon, JS, Devlin, D, Lozza, G, Veltri, A, Beltramo, M, Bertorelli, R and Reggiani, A (2009). "Identification of single nucleotide polymorphisms of the human metabotropic glutamate receptor 1 gene and pharmacological characterization of a P993S variant." *Biochem Pharmacol* 77(7): 1246-1253.

Engelender, S, Sharp, AH, Colomer, V, Tokito, MK, Lanahan, A, Worley, P, Holzbaur, EL and Ross, CA (1997). "Huntingtin-associated protein 1 (HAP1) interacts with the p150Glued subunit of dynactin." *Hum Mol Genet* 6(13): 2205-2212.

Ferraguti, F, Crepaldi, L and Nicoletti, F (2008). "Metabotropic glutamate 1 receptor: current concepts and perspectives." *Pharmacol Rev* 60(4): 536-581.

Fesus, L and Piacentini, M (2002). "Transglutaminase 2: an enigmatic enzyme with diverse functions." *Trends Biochem Sci* 27(10): 534-539.

Gauthier, LR, Charrin, BC, Borrell-Pages, M, Dompierre, JP, Rangone, H, Cordelieres, FP, De Mey, J, MacDonald, ME, Lessmann, V, Humbert, S and Saudou, F (2004). "Huntingtin controls neurotrophic support and survival of neurons by enhancing BDNF vesicular transport along microtubules." *Cell* 118(1): 127-138.

Gayan, J, Brocklebank, D, Andresen, JM, Alkorta-Aranburu, G, Zameel Cader, M, Roberts, SA, Cherny, SS, Wexler, NS, Cardon, LR and Housman, DE (2008).

- "Genomewide linkage scan reveals novel loci modifying age of onset of Huntington's disease in the Venezuelan HD kindreds." *Genet Epidemiol* 32(5): 445-453.
- Gusella, JF, Wexler, NS, Conneally, PM, Naylor, SL, Anderson, MA, Tanzi, RE, Watkins, PC, Ottina, K, Wallace, MR, Sakaguchi, AY and et al. (1983). "A polymorphic DNA marker genetically linked to Huntington's disease." *Nature* 306(5940): 234-238.
- Hayden, M, Kremer, B. 'Huntington Disease'. Metabolic and Molecular Basis of Inherited Disease. 8th edition. 2001. 5677-5701
- Hellenbroich, Y, Kaulich, M, Opitz, S, Schwinger, E and Zuhlke, C (2004). "No association of the SCA1 (CAG)<sub>31</sub> allele with Huntington's disease, myotonic dystrophy type 1 and spinocerebellar ataxia type 3." *Psychiatr Genet* 14(2): 61-63.
- Hersheson, J, Haworth, A and Houlden, H (2012). "The inherited ataxias: genetic heterogeneity, mutation databases, and future directions in research and clinical diagnostics." *Hum Mutat* 33(9): 1324-1332.
- Holmes, SE, O'Hearn, EE, McInnis, MG, Gorelick-Feldman, DA, Kleiderlein, JJ, Callahan, C, Kwak, NG, Ingersoll-Ashworth, RG, Sherr, M, Sumner, AJ, Sharp, AH, Ananth, U, Seltzer, WK, Boss, MA, Vieria-Saecker, AM, Epplen, JT, Riess, O, Ross, CA and Margolis, RL (1999). "Expansion of a novel CAG trinucleotide repeat in the 5' region of PPP2R2B is associated with SCA12." *Nat Genet* 23(4): 391-392.
- Humbert, S, Bryson, EA, Cordelieres, FP, Connors, NC, Datta, SR, Finkbeiner, S, Greenberg, ME and Saudou, F (2002). "The IGF-1/Akt pathway is neuroprotective in Huntington's disease and involves Huntingtin phosphorylation by Akt." *Dev Cell* 2(6): 831-837.
- Karpuj, MV, Garren, H, Slunt, H, Price, DL, Gusella, J, Becher, MW and Steinman, L (1999). "Transglutaminase aggregates huntingtin into nonamyloidogenic polymers, and its enzymatic activity increases in Huntington's disease brain nuclei." *Proc Natl Acad Sci U S A* 96(13): 7388-7393.
- Keckarevic, D, Culjkovic, B, Savic, D, Stojkovic, O, Kostic, V, Vukosavic, S and Romac, S (2000). "The status of SCA1, MJD/SCA3, FRDA, DRPLA and MD triplet containing genes in patients with Huntington disease and healthy controls." *J Neurogenet* 14(4): 257-263.
- Kerner, JA, Standaert, DG, Penney, JB, Jr., Young, AB and Landwehrmeyer, GB (1997). "Expression of group one metabotropic glutamate receptor subunit mRNAs

in neurochemically identified neurons in the rat neostriatum, neocortex, and hippocampus." *Brain Res Mol Brain Res* 48(2): 259-269.

- Kiraly, R, Barta, E and Fesus, L (2011). "Polymorphism of transglutaminase 2: unusually low frequency of genomic variants with deficient functions." *Amino Acids*.
- Kobayashi, H, Abe, K, Matsuura, T, Ikeda, Y, Hitomi, T, Akechi, Y, Habu, T, Liu, W, Okuda, H and Koizumi, A (2011). "Expansion of intronic GGCCTG hexanucleotide repeat in NOP56 causes SCA36, a type of spinocerebellar ataxia accompanied by motor neuron involvement." *Am J Hum Genet* 89(1): 121-130.
- Koefoed, P, Hasholt, L, Fenger, K, Nielsen, JE, Eiberg, H, Buschard, K and Sorensen, SA (1998). "Mitotic and meiotic instability of the CAG trinucleotide repeat in spinocerebellar ataxia type 1." *Hum Genet* 103(5): 564-569.
- Landles, C and Bates, GP (2004). "Huntingtin and the molecular pathogenesis of Huntington's disease. Fourth in molecular medicine review series." *EMBO Rep* 5(10): 958-963.
- Lang, F, Artunc, F and Vallon, V (2009). "The physiological impact of the serum and glucocorticoid-inducible kinase SGK1." *Curr Opin Nephrol Hypertens* 18(5): 439-448.
- Leong, ML, Maiyar, AC, Kim, B, O'Keeffe, BA and Firestone, GL (2003). "Expression of the serum- and glucocorticoid-inducible protein kinase, Sgk, is a cell survival response to multiple types of environmental stress stimuli in mammary epithelial cells." *J Biol Chem* 278(8): 5871-5882.
- Li, JL, Hayden, MR, Almqvist, EW, Brinkman, RR, Durr, A, Dode, C, Morrison, PJ, Suchowersky, O, Ross, CA, Margolis, RL, Rosenblatt, A, Gomez-Tortosa, E, Cabrero, DM, Novelletto, A, Frontali, M, Nance, M, Trent, RJ, McCusker, E, Jones, R, Paulsen, JS, Harrison, M, Zanko, A, Abramson, RK, Russ, AL, Knowlton, B, Djousse, L, Mysore, JS, Tariot, S, Gusella, MF, Wheeler, VC, Atwood, LD, Cupples, LA, Saint-Hilaire, M, Cha, JH, Hersch, SM, Koroshetz, WJ, Gusella, JF, MacDonald, ME and Myers, RH (2003). "A genome scan for modifiers of age at onset in Huntington disease: The HD MAPS study." *Am J Hum Genet* 73(3): 682-687.
- Li, JL, Hayden, MR, Warby, SC, Durr, A, Morrison, PJ, Nance, M, Ross, CA, Margolis, RL, Rosenblatt, A, Squitieri, F, Frati, L, Gomez-Tortosa, E, Garcia, CA, Suchowersky, O, Klimek, ML, Trent, RJ, McCusker, E, Novelletto, A, Frontali, M, Paulsen, JS, Jones, R, Ashizawa, T, Lazzarini, A, Wheeler, VC, Prakash, R, Xu, G, Djousse, L, Mysore, JS, Gillis, T, Hakky, M, Cupples, LA, Saint-Hilaire, MH, Cha, JH, Hersch, SM, Penney, JB, Harrison, MB, Perlman, SL, Zanko, A, Abramson, RK, Lechich, AJ, Duckett, A, Marder, K, Conneally, PM, Gusella, JF, MacDonald,



- ME and Myers, RH (2006). "Genome-wide significance for a modifier of age at neurological onset in Huntington's disease at 6q23-24: the HD MAPS study." *BMC Med Genet* 7: 71.
- Li, XJ, Li, SH, Sharp, AH, Nucifora, FC, Jr., Schilling, G, Lanahan, A, Worley, P, Snyder, SH and Ross, CA (1995). "A huntingtin-associated protein enriched in brain with implications for pathology." *Nature* 378(6555): 398-402.
- MacDonald, M, Ambrose, CM, Duyao, MP, Myers, RH, Lin, C, Srinidhi, L, Barnes, G, Taylor, SA, James, M, Groot, N, MacFarlane, H, Jenkins, B, Anderson, MA, Wexler, NS, Gusella, JF, Bates, GP, Baxendale, S, Hummerich, H, Kirby, S, North, M, Youngman, S, Mott, R, Zehetner, G, Sedlacek, Z, Poustka, A, Frischauf, A, Lehrach, H, Buckler, AJ, Church, D, Doucette-Stamm, L, O'Donovan, MC, Riba-Ramirez, L, Shah, M, Stanton, VP, Strobel, SA, Draths, KM, Wales, JL, Dervan, P, Housman, DE, Altherr, M, Shiang, R, Thompson, L, Fielder, T, Wasmuth, JJ, Tagle, D, Valdes, J, Elmer, L, Allard, M, Castilla, L, Swaroop, M, Blanchard, K, Collins, FS, Snell, R, Holloway, T, Gillespie, K, Datson, N, Shaw, D, Harper, PS (1993). "A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group." *Cell* 72(6): 971-983.
- Mastroberardino, PG, Iannicola, C, Nardacci, R, Bernassola, F, De Laurenzi, V, Melino, G, Moreno, S, Pavone, F, Oliverio, S, Fesus, L and Piacentini, M (2002). "'Tissue' transglutaminase ablation reduces neuronal death and prolongs survival in a mouse model of Huntington's disease." *Cell Death Differ* 9(9): 873-880.
- McGuire, JR, Rong, J, Li, SH and Li, XJ (2006). "Interaction of Huntingtin-associated protein-1 with kinesin light chain: implications in intracellular trafficking in neurons." *J Biol Chem* 281(6): 3552-3559.
- Meldrum, BS (2000). "Glutamate as a neurotransmitter in the brain: review of physiology and pathology." *J Nutr* 130(4S Suppl): 1007S-1015S.
- Metzger, S, Saukko, M, Van Che, H, Tong, L, Puder, Y, Riess, O and Nguyen, HP (2010). "Age at onset in Huntington's disease is modified by the autophagy pathway: implication of the V471A polymorphism in Atg7." *Hum Genet* 128(4): 453-459.
- Metzger, S, Rong, J, Nguyen, HP, Cape, A, Tomiuk, J, Soehn, AS, Propping, P, Freudenberg-Hua, Y, Freudenberg, J, Tong, L, Li, SH, Li, XJ and Riess, O (2008). "Huntingtin-associated protein-1 is a modifier of the age-at-onset of Huntington's disease." *Hum Mol Genet* 17(8): 1137-1146.
- Metzger, S, Bauer, P, Tomiuk, J, Laccione, F, Didonato, S, Gellera, C, Soliveri, P, Lange, HW, Weirich-Schwaiger, H, Wenning, GK, Melegh, B, Havasi, V, Baliko, L, Wiczorek, S, Arning, L, Zaremba, J, Sulek, A, Hoffman-Zacharska, D, Basak,

- AN, Ersoy, N, Zidovska, J, Kebrdlova, V, Pandolfo, M, Ribai, P, Kadasi, L, Kvasnicova, M, Weber, BH, Kreuz, F, Dose, M, Stuhmann, M and Riess, O (2006). "The S18Y polymorphism in the UCHL1 gene is a genetic modifier in Huntington's disease." *Neurogenetics* 7(1): 27-30.
- Miller, J, Arrasate, M, Shaby, BA, Mitra, S, Masliah, E and Finkbeiner, S (2010). "Quantitative relationships between huntingtin levels, polyglutamine length, inclusion body formation, and neuronal death provide novel insight into huntington's disease molecular pathogenesis." *J Neurosci* 30(31): 10541-10550.
- Nakanishi, S and Masu, M (1994). "Molecular diversity and functions of glutamate receptors." *Annu Rev Biophys Biomol Struct* 23: 319-348.
- Naze, P, Vuillaume, I, Destee, A, Pasquier, F and Sablonniere, B (2002). "Mutation analysis and association studies of the ubiquitin carboxy-terminal hydrolase L1 gene in Huntington's disease." *Neurosci Lett* 328(1): 1-4.
- Paradisi, I, Hernandez, A and Arias, S (2008). "Huntington disease mutation in Venezuela: age of onset, haplotype analyses and geographic aggregation." *J Hum Genet* 53(2): 127-135.
- Pearson, CE, Nichol Edamura, K and Cleary, JD (2005). "Repeat instability: mechanisms of dynamic mutations." *Nat Rev Genet* 6(10): 729-742.
- Pin, JP and Duvoisin, R (1995). "The metabotropic glutamate receptors: structure and functions." *Neuropharmacology* 34(1): 1-26.
- Pulst, SM, Santos, N, Wang, D, Yang, H, Huynh, D, Velazquez, L and Figueroa, KP (2005). "Spinocerebellar ataxia type 2: polyQ repeat variation in the CACNA1A calcium channel modifies age of onset." *Brain* 128(Pt 10): 2297-2303.
- Rangone, H, Poizat, G, Troncoso, J, Ross, CA, MacDonald, ME, Saudou, F and Humbert, S (2004). "The serum- and glucocorticoid-induced kinase SGK inhibits mutant huntingtin-induced toxicity by phosphorylating serine 421 of huntingtin." *Eur J Neurosci* 19(2): 273-279.
- Ribeiro, FM, Paquet, M, Ferreira, LT, Cregan, T, Swan, P, Cregan, SP and Ferguson, SS (2010). "Metabotropic glutamate receptor-mediated cell signaling pathways are altered in a mouse model of Huntington's disease." *J Neurosci* 30(1): 316-324.
- Richards, RI (2001). "Dynamic mutations: a decade of unstable expanded repeats in human genetic disease." *Hum Mol Genet* 10(20): 2187-2194.
- Roos, RA (2010). "Huntington's disease: a clinical review." *Orphanet J Rare Dis* 5(1): 40.

- Rosenblatt, A, Brinkman, RR, Liang, KY, Almqvist, EW, Margolis, RL, Huang, CY, Sherr, M, Franz, ML, Abbott, MH, Hayden, MR and Ross, CA (2001). "Familial influence on age of onset among siblings with Huntington disease." *Am J Med Genet* 105(5): 399-403.
- Rupp, J, Blekher, T, Jackson, J, Beristain, X, Marshall, J, Hui, S, Wojcieszek, J and Foroud, T (2010). "Progression in prediagnostic Huntington disease." *J Neurol Neurosurg Psychiatry* 81(4): 379-384.
- Savic, D, Topisirovic, I, Keckarevic, M, Keckarevic, D, Major, T, Culjkovic, B, Stojkovic, O, Rakocevic-Stojanovic, V, Mladenovic, J, Todorovic, S, Apostolski, S and Romac, S (2001). "Is the 31 CAG repeat allele of the spinocerebellar ataxia 1 (SCA1) gene locus non-specifically associated with trinucleotide expansion diseases?" *Psychiatr Genet* 11(4): 201-205.
- Tabrizi, SJ, Langbehn, DR, Leavitt, BR, Roos, RA, Durr, A, Craufurd, D, Kennard, C, Hicks, SL, Fox, NC, Scahill, RI, Borowsky, B, Tobin, AJ, Rosas, HD, Johnson, H, Reilmann, R, Landwehrmeyer, B and Stout, JC (2009). "Biological and clinical manifestations of Huntington's disease in the longitudinal TRACK-HD study: cross-sectional analysis of baseline data." *Lancet Neurol* 8(9): 791-801.
- Taherzadeh-Fard, E, Saft, C, Wieczorek, S, Epplen, JT and Arning, L (2010). "Age at onset in Huntington's disease: replication study on the associations of ADORA2A, HAP1 and OGG1." *Neurogenetics* 11(4): 435-439.
- Takahashi, T, Katada, S and Onodera, O (2010). "Polyglutamine diseases: where does toxicity come from? what is toxicity? where are we going?" *J Mol Cell Biol* 2(4): 180-191.
- Takeshita, Y, Fujinaga, R, Zhao, C, Yanai, A and Shinoda, K (2006). "Huntingtin-associated protein 1 (HAP1) interacts with androgen receptor (AR) and suppresses SBMA-mutant-AR-induced apoptosis." *Hum Mol Genet* 15(15): 2298-2312.
- Tang, TS, Tu, H, Chan, EY, Maximov, A, Wang, Z, Wellington, CL, Hayden, MR and Bezprozvanny, I (2003). "Huntingtin and huntingtin-associated protein 1 influence neuronal calcium signaling mediated by inositol-(1,4,5) triphosphate receptor type 1." *Neuron* 39(2): 227-239.
- Testa, CM, Friberg, IK, Weiss, SW and Standaert, DG (1998). "Immunohistochemical localization of metabotropic glutamate receptors mGluR1a and mGluR2/3 in the rat basal ganglia." *J Comp Neurol* 390(1): 5-19.
- van Duijn, E, Kingma, EM and van der Mast, RC (2007). "Psychopathology in verified Huntington's disease gene carriers." *J Neuropsychiatry Clin Neurosci* 19(4): 441-448.

- Wexler, NS, Lorimer, J, Porter, J, Gomez, F, Moskowitz, C, Shackell, E, Marder, K, Penchaszadeh, G, Roberts, SA, Gayan, J, Brocklebank, D, Cherny, SS, Cardon, LR, Gray, J, Dlouhy, SR, Wiktorski, S, Hodes, ME, Conneally, PM, Penney, JB, Gusella, J, Cha, JH, Irizarry, M, Rosas, D, Hersch, S, Hollingsworth, Z, MacDonald, M, Young, AB, Andresen, JM, Housman, DE, De Young, MM, Bonilla, E, Stillings, T, Negrette, A, Snodgrass, SR, Martinez-Jaurrieta, MD, Ramos-Arroyo, MA, Bickham, J, Ramos, JS, Marshall, F, Shoulson, I, Rey, GJ, Feigin, A, Arnheim, N, Acevedo-Cruz, A, Acosta, L, Alvir, J, Fischbeck, K, Thompson, LM, Young, A, Dure, L, O'Brien, CJ, Paulsen, J, Brickman, A, Krch, D, Peery, S, Hogarth, P, Higgins, DS, Jr. and Landwehrmeyer, B (2004). "Venezuelan kindreds reveal that genetic and environmental factors modulate Huntington's disease age of onset." *Proc Natl Acad Sci U S A* 101(10): 3498-3503.
- Zainelli, GM, Dudek, NL, Ross, CA, Kim, SY and Muma, NA (2005). "Mutant huntingtin protein: a substrate for transglutaminase 1, 2, and 3." *J Neuropathol Exp Neurol* 64(1): 58-65.

## CURRICULUM VITAE RESUMIDO

Souza, AFD

## 1. DADOS PESSOAIS

Nome: Aline Francielle Damo Souza

Local e data de nascimento: Passo Fundo, Rio Grande do Sul, Brasil, 23/02/1987

Endereço profissional: Laboratório de Identificação Genética – Centro de Pesquisa Experimental – Hospital de Clínicas de Porto Alegre, Rua Ramiro Barcelos, 2350, CEP: 90035-903 – Porto Alegre – RS

Telefone profissional: (51) 3359.7689

e-mail: alinefds2@hotmail.com

2. FORMAÇÃO: Graduação em Biomedicina (Universidade Federal do Rio Grande do Sul, de 2007 a 2010);
3. ESTÁGIOS: Iniciação científica no Laboratório de Biologia Molecular Endócrina e Tumoral no projeto de pesquisa: Efeitos do silenciamento do receptor de androgênios e seus coreguladores FHL2 e N-CoR por siRNAs na expressão gênica e resposta proliferativa de células de carcinoma prostático. Integrantes: Aline Francielle Damo Souza; Responsável: Dr<sup>a</sup>. Gisele Branchini e Orientadora: Prof. Dr<sup>a</sup>. Ilma Simoni Brum da Silva. Financiadores: Conselho Nacional de Desenvolvimento Científico e Tecnológico - Auxílio financeiro. De janeiro de 2009 a dezembro de 2010. Participação na realização de cultura de células de tecido prostático, extração de RNA, RT-PCR, PCR convencional e em tempo real. Apresentação de pôsteres no Salão de Iniciação Científica (2009 e 2010), Semana Científica do HCPA (2009, 2010 e 2011), no *International Symposium on Cell Signaling*, Florianópolis (2010).

#### 4. ARTIGOS COMPLETOS PUBLICADOS:

- 1- Santin, AP ; Souza, AFD; Brum, IS; Furlanetto, TW. Validation of Reference Genes for Normalizing Gene Expression in Real-Time Quantitative Reverse Transcription PCR in Human Thyroid Cells in Primary Culture Treated with Progesterone and Estradiol. *Molecular Biotechnology* **JCR**, v. 51, p. 1-5, 2012.
- 2- Souza, AFD; Brum, IS; Neto, BS; Berger, M; Branchini, G. Reference gene for primary culture of prostate cancer cells. *Molecular Biology Reports* **JCR**, v. 39, p. 23269617, 2012.

#### 5. RESUMOS E TRABALHOS APRESENTADOS EM CONGRESSOS:

- 1- Souza, AFD; Gheno TC; Rieder C; Pedroso JL; Barsottini O; Salarini D; Pereira E; Vargas FR; Godeiro C; Toralles MBP; Jardim LB; Saraiva-Pereira ML. Avaliação de polimorfismos no gene da Transglutaminase 2 (*TGM2*) e sua relação com a variação da idade de início em pacientes com Doença de Huntington. Em: XXIV Congresso Brasileiro de Genética Médica, 2012, Porto Alegre. Anais On Line, 2012.
- 2- Souza, AFD; Martiny, PB; Brum, IS; Neto, BS; Berger, M; Branchini, G. Reference gene for primary culture of prostate cancer cells. Em: 15th International Congress on Hormonal Steroids and Hormones & Cancer, 2012, Kanazawa. 15th International Congress on Hormonal Steroids and Hormones & Cancer. Kanazawa, 2012. v. 15. p. 142-142.
- 3- Souza, AFD; Branchini, G; Pizzolato, LS; Neto, BS; Berger, M; Brum, IS. Searching for a housekeeping gene for gene expression studies in primary culture of prostate cancer. Em: International Symposium on Cell Signaling, 2010, Florianópolis. Abstract Book, 2010. p. 9-112.