

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE BIOCÊNCIAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

**Estudo de haplótipos em famílias com Ataxia Espinocerebelar tipo 10  
(SCA10): evidências de um efeito fundador da mutação**

**Giovana Bavia Bampi**

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

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

Porto Alegre, março de 2015

## **INSTITUIÇÕES E FONTES FINANCIADORAS**

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

## **AGRADECIMENTOS**

À minha orientadora, Profa. Maria Luiza Saraiva Pereira, pela oportunidade, convivência e ensinamentos, os quais foram essenciais para meu crescimento científico e profissional.

Às minhas queridas professoras e “madrinhas científicas”, Marion Schiengold e Kátia Kvitko, pelos ensinamentos, aprendizados, conselhos e, acima de tudo, por terem contribuído muito no momento de preparação para entrar no Programa de Pós-Graduação.

Ao Prof. Aldo Mellender de Araújo, por ter sempre se mostrado disposto a me ajudar no estudo de Evolução Biológica, essencial para minha aprovação no exame de admissão do Mestrado.

À Profa. Sídia Callegari Jacques, pelos assessoramentos em bioestatística, sempre disposta ajudar e sanar qualquer dúvida.

À Profa. Maria Cátira Bortoloni e à Dra. Tábita Hünemeier, pela fundamental contribuição e conhecimento na realização deste trabalho.

Ao colega e amigo, Rafael Bisso Machado, pela convivência, alegria, divertimento e descontração, quando eu achava que tudo estava perdido.

Ao Elmo, secretário do PPGBM, quem sempre se mostrou muito atencioso, solícito e competente sempre que precisei.

Aos colegas do Laboratório de Identificação Genética, pela amizade, convivência e intercâmbio de conhecimento.

Aos meus pais e meu irmão, pelo amor incondicional, carinho e confiança.

Aos meus amigos, pela torcida e compreensão nos momentos de ausência.

## SUMÁRIO:

LISTA DE ABREVIATURAS.....	5
LISTA DE FIGURAS.....	6
RESUMO.....	7
ABSTRACT.....	8
1. INTRODUÇÃO.....	9
1.1. Ataxias espinocerebelares.....	9
1.2. Ataxia espinocerebelar tipo 10.....	10
1.2.1. Manifestações clínicas.....	11
1.2.1.1. Sinais cerebelares e sintomas.....	11
1.2.1.2. Convulsões.....	12
1.2.1.3. Sintomas extracerebelares.....	12
1.2.2. Aspectos moleculares.....	13
1.2.3. Aspectos genéticos .....	15
1.2.4. A proteína e aspectos fisiopatológicos.....	17
1.2.5. Diagnóstico.....	18
1.2.6. Ancestralidade.....	19
1.3. Justificativa.....	20
2. OBJETIVOS.....	21
2.1. Objetivo geral.....	21
2.2. Objetivos específicos.....	21
3. RESULTADOS.....	22
<i>Haplotype study in SCA10 families provides further evidence for a common     ancestral origin of the mutation</i> .....	23
4. DISCUSSÃO.....	51
5. CONCLUSÕES.....	54
6. REFERÊNCIAS.....	55

## LISTA DE ABREVIATURAS

A	adenina
AIMs	marcadores informativos de ancestralidade ( <i>Ancestry Informative Markers</i> )
bp	pares de base ( <i>base pairs</i> )
C	citosina
cM	centiMorgans
DNA	ácido desoxirribonucleico ( <i>desoxyribonucleic acid</i> )
DRPLA <i>Atrophy</i> )	atrofia dentato-rubro-palido-lusiana ( <i>Dentatorubral Pallido Luysian</i>
G	guanina
hnRNP	ribonucleoproteínas nucleares heterogêneas ( <i>heterogeneous nuclear ribonucleoprotein</i> )
Kb	kilobase
μMolar	micromolar
mRNA	ácido ribonucleico mensageiro ( <i>messenger ribonucleic acid</i> )
mtDNA	DNA mitocondrial
OMIM	<i>Online Mendelian Inheritance in Man</i>
PCR	reação em cadeia da polimerase ( <i>Polymerase Chain Reaction</i> )
pmol	picomol
RNA	ácido ribonucleico ( <i>ribonucleic acid</i> )
RP-PCR	<i>Repeat-Primed PCR</i>
SCA	ataxia espinocerebelar ( <i>Spinocerebellar Ataxia</i> )
SNP	polimorfismo de nucleotídeos individuais ( <i>Single Nucleotide Polymorphism</i> )
STR	polimorfismo de repetições nucleotídicas ( <i>Short Tandem Repeats</i> )
T	timina

## LISTA DE FIGURAS

	Página
<b>Figura 1.</b> Representação esquemática do gene <i>ATXN10</i>	14

## RESUMO

A ataxia espinocerebelar tipo 10 (SCA10) é uma doença neurodegenerativa rara de herança autossômica dominante caracterizada por atrofia cerebelar com alterações da marcha e, em alguns casos, convulsões. A SCA10 é causada por expansões de repetições pentanucleotídicas ATTCT no íntron 9 do gene *ATXN10*, o qual se localiza no *locus* 22q13. Alelos normais apresentam entre 10 a 29 repetições e o alelo patogênico apresenta entre 800 a 4.500 repetições. Até o momento, casos de SCA10 foram descritos apenas em pacientes miscigenados de países do continente americano como México, Brasil, Argentina, Venezuela, Colômbia, Estados Unidos e, mais recentemente, Peru. A origem ameríndia auto declarada pelos pacientes com SCA10 e a ausência de casos em países europeus e asiáticos indicam a hipótese de ocorrência de um efeito fundador da mutação nas populações nativas americanas. O objetivo deste trabalho foi investigar a hipótese de origem ancestral comum da mutação no gene *ATXN10*. As amostras analisadas foram proveniente de 16 famílias brasileiras e de 21 famílias peruanas com SCA10. Além do grupo de pacientes, um grupo controle composto por 48 indivíduos saudáveis da população indígena Quechua do Peru foi também incluída na análise assim como 51 controles brasileiros de um estudo anterior. Os resultados obtidos mostraram que o haplótipo 19CGGC14 associado ao alelo da expansão está presente em 46,8% das famílias de brasileiros e 62,8% das famílias de peruanos. As frequências de ambos os grupos não é estatisticamente diferente dos controles Quechua (57,3%), sendo diferente dos controles brasileiros (11,8%) ( $p < 0,001$ ). Entretanto, origem etnogeográfica da mutação ainda é desconhecida. O haplótipo comum mínimo foi expandido incluindo outros dois marcadores polimórficos, os quais integram dois haplótipos com alta prevalência em populações nativo americanas com o intuito de obter uma aproximação da origem da região cromossômica onde a mutação está inserida. Dois haplótipos mais frequentes 19-13-CGGC-14-10 e 19-15-CGGC-14-10 foram identificados nos controles indígenas Quechua, com frequências relativas de 14,3% e 13,3% respectivamente. O segundo haplótipo mais frequente em Quechuas, 19-15-CGGC-14-10, é encontrado em 50,0% das famílias brasileiras e em 64,7% das famílias Peruanas com SCA10. Esses achados corroboram a hipótese de origem ameríndia da mutação.

## ABSTRACT

Spinocerebellar ataxia type 10 is a rare autosomal dominant neurodegenerative disorder characterized by progressive cerebellar ataxia and epilepsy in some cases. The disease is caused by a pentanucleotide ATTCT expansion in intron 9 of the *ATXN10* gene, which is located at locus 22q13.3. Normal alleles range from 10 to 29 repeats while mutant alleles range from 800 to 4,500 repeats. SCA10 has only been described so far in admixed patients from American countries such as Mexico, Brazil, Argentina, Venezuela, Colombia, United States and more recently Peru. The self-declared Amerindian ancestry by patients and the absence of SCA10 in European and Asian countries leads to the hypothesis of a mutation founder effect in the Native American populations. The aim of this study was to investigate the hypothesis of a common ancestral origin of *ATXN10* mutation. Samples analyzed were from 16 Brazilian families, 21 Peruvian families with SCA10. In addition to patient samples, 48 healthy individuals of Indigenous Quechua from Peru were also included in the laboratorial analyses along with 51 Brazilian controls from a previous study. Our data has shown that 19CGGC14-shared haplotype was found in 46.8% of Brazilian and in 62.8% of Peruvian families. Frequencies from both groups are not statistically different from Quechua controls (57.3%), but they are statistically different from Brazilian controls (11.8%) ( $p < 0.001$ ). However, the mutation ethno-geographical origin remains unclear. The minimal common haplotype was expanded by including two additional polymorphic markers that are found at high prevalence in two haplotypes in Native American populations aiming to shed light on the chromosome region ancestry where the mutation arose. Two frequent haplotypes, 19-13-CGGC-14-10 and 19-15-CGGC-14-10 were identified in Indigenous Quechua controls, with relative frequencies of 14.3% and 13.3% respectively. The second most frequent haplotype in Quechuas, 19-15-CGGC-14-10, is found in 50.0% of Brazilian and in 64.7% of Peruvian families with SCA10. These findings corroborate the hypothesis of a Native American ancestry of the mutation.



## 1. INTRODUÇÃO

### 1.1. Ataxias espinocerebelares

As ataxias espinocerebelares (*Spinocerebellar ataxia* - SCAs) constituem um grupo de doenças neurológicas com grande heterogeneidade clínica e genética apresentando um padrão de herança autossômica dominante. Os indivíduos afetados apresentam irregularidades no ajuste fino da postura e movimentos causadas pela perda progressiva dos neurônios do cerebelo e suas conexões.

O processo neurodegenerativo não se limita apenas ao cerebelo, podendo expandir o quadro clínico. Células da retina, do nervo óptico, da medula óssea, dos gânglios basais, do córtex, da coluna espinhal e do sistema nervoso periférico também podem ser afetadas. Os primeiros sinais clínicos são resultantes da atrofia cerebelar, os quais se expandem a estruturas extra-cerebelares, de forma característica a cada subtipo de ataxia. Os sinais clínicos observados, tais como a idade de início dos sintomas, a velocidade de progressão da doença e a gravidade do quadro clínico dependem, portanto, do tipo de ataxia espinocerebelar.

As causas primárias de várias doenças neurodegenerativas – como algumas SCAs – são mutações dinâmicas. O mecanismo proposto para as mutações dinâmicas envolve a expansão de uma sequência nucleotídica durante a transmissão vertical de geração para geração alcançando um tamanho patológico específico para cada doença. Ao contrário, indivíduos que não manifestam a doença permanecem com o número de repetições dentro do limite da variação normal.

Até o momento, mais de trinta *loci* distintos e mais de vinte e cinco genes foram associados às formas mendelianas de SCAs. Sete delas – SCA1, SCA2, SCA3, SCA6, SCA7, SCA12, SCA17 – e a DRPLA (atrofia dentato-rubro-palido-lusiana) estão associadas a alterações de expansões trinucleotídicas CAGs (Bird, 2012). A SCA8 está associada a repetições trinucleotídicas CTG. Expansões pentanucleotídicas e hexanucleotídicas estão associadas à SCA10 e à SCA36, respectivamente (Bird, 2012). Mutações de ponto e deleções também podem estar associadas a outras formas de SCAs: SCA5, SCA13, SCA14, SCA16 e SCA27. Todavia, nas demais SCAs, os genes e,

portanto, as mutações, ainda não foram identificados nem caracterizados. Os mecanismos que acarretam as expansões patológicas dos alelos ainda não estão bem esclarecidos.

## 1.2. Ataxia espinocerebelar tipo 10 (SCA10)

O primeiro relato de uma família portadora de uma forma distinta de SCA foi descrito em 1998, sendo confirmada como SCA mais tardiamente (Grewal et al., 1998). No ano seguinte, a localização do gene de SCA10 foi mapeada em um estudo envolvendo duas famílias grandes mexicanas, as quais apresentavam ataxia puramente cerebelar acompanhadas de convulsões (Matsuura et al., 1999; Zu et al., 1999). Subsequentemente, mais quatro famílias não aparentadas foram descritas no México, revelando um fenótipo ainda mais diverso (Rasmussen et al., 2001). Em 2004, a confirmação de cinco famílias brasileiras com SCA10 foi o primeiro relato da presença de SCA10 sem origem mexicana (Teive et al., 2004). Ao contrário das famílias mexicanas, todos os 28 indivíduos dessas famílias brasileiras apresentavam ataxia cerebelar sem convulsões, demonstrando a outra extremidade do espectro dos fenótipos de SCA10. A SCA10 é uma das ataxias mais frequentes no México e no Brasil (Rasmussen et al., 2000; Teive et al., 2004).

Nos últimos 5 anos, mais relatos de pacientes com SCA10 foram publicados em outros países da América do Sul. Em 2007, uma família argentina, com ancestralidade espanhola e ameríndia, foi descrita com 6 membros afetados (Gatto et al., 2007). Em 2009, a SCA10 foi descrita em 3 gerações de uma família venezuelana com 5 indivíduos afetados (Gallardo et al., 2009). O quadro clínico dessas famílias de ambas nacionalidades é semelhante às famílias mexicanas, com ataxia cerebelar acompanhadas de convulsões. Em 2012, o primeiro caso de coexistência de duas ataxias em um único indivíduo foi descrito: um paciente mexicano de 54 anos descendente da tribo indígena Yaqui com SCA2 e SCA10 (Kapur et al., 2012). Em seguida, o caso de uma paciente colombiana manifestando concomitantemente doença de Huntington e SCA10 também foi publicado (Roxburgh et al., 2013). Até o momento, o único caso de paciente com SCA10 de origem estadunidense descrito nos Estados Unidos menciona que a avó materna do paciente seria de ascendência indígena pertencente à tribo Sioux (Bushara et al., 2013). O único caso de ocorrência da doença fora da América foi publicado por pesquisadores na Itália. No entanto, a paciente é uma imigrante peruana, com ascendência auto declarada puramente

indígena, proveniente da região de Paijan ao noroeste da costa do Peru (Leonardi et al., 2014). E mais recentemente, uma família boliviana com 3 afetados sendo 2 deles com características fenotípicas de SCA2 e SCA10 (Baizabal-Carvallo & Jankovic, 2012; Baizabal-Carvallo et al., 2015) e outra família proveniente da Guatemala foram descritas (Trikamji et al., 2015). Até o momento, a SCA10 foi descrita exclusivamente em pacientes miscigenados do continente Americano (Teive et al., 2013). Os dados presentes na literatura sugerem que a doença teve origem em populações nativas americanas e tenha se espalhado para o resto do continente americano.

### 1.2.1. Manifestações clínicas

Todos os pacientes com SCA10 apresentam disfunções cerebelares progressivas. A função do cerebelo está relacionada ao ajuste, coordenação e refinamento dos complexos movimentos motores, resultando em um padrão de movimentos suaves e integrados. A fluidez e a acurácia dos movimentos são perdidas devido a disfunções cerebelares. A doença apresenta uma expressividade clínica variável com início do aparecimento dos sintomas clínicos variando entre os 12 e 80 anos de idade.

#### 1.2.1.1. Sinais cerebelares e sintomas

Aos sinais cerebelares incluem-se ataxia da marcha e membros, disartria e distúrbios oculares. O primeiro sinal clínico geralmente se manifesta como marcha e postura desequilibradas associado a um grau variável de ataxia nos membros, configurando um modo de andar característico, com os pés afastados, para compensar a falta de equilíbrio, o que lembra o caminhar de um indivíduo em estado de embriaguez. A ataxia dos membros é caracterizada por movimentos irregulares e descoordenados, os quais não se justificam por fraqueza muscular ou perda da sensibilidade. A falta de coordenação dos braços e pernas pode afetar a capacidade de realizar tarefas diárias simples, como comer e escrever (Grewal et al., 1998; Rasmussen et al., 2001; Grewal et al., 2002).

No exame neurológico, o paciente demonstra ataxia de marcha, disdiadococinesia, tremor de intenção e dismetria. A ataxia é seguida pela disartria; dificuldades na

articulação das palavras e fala arrastada, ou explosiva com hesitações frequentes e pausas inapropriadas (Matsuura et al., 2000; Zu et al., 2001; Teive et al., 2004).

Alterações nos movimentos oculares também podem ser observados (discinesia), comumente apresentada como intrusões sacádicas hipométricas durante o acompanhamento ocular, que pode progredir para uma vibração ocular conjugada com breves oscilações dos olhos durante a tentativa de fixação ou movimento (dismetria ocular). Nistagmo pode ser observado em alguns pacientes com SCA10. Essas irregularidades no movimento dos olhos podem aparecer nos estágios iniciais da doença.

Estudos de neuroimagem com ressonância magnética ou tomografia computadorizada revelaram uma atrofia cerebelar específica envolvendo o vermis e os dois hemisférios, enquanto outras estruturas do cérebro, como córtex cerebral e o tronco encefálico, parecem ser minimamente afetadas (Lin & Ashizawa, 2005).

#### 1.2.1.2. Convulsões

As crises convulsivas estão associadas à SCA10 de forma variável. A epilepsia se manifesta, geralmente, através de convulsões motoras generalizadas e/ou convulsões parciais complexas poucos anos após o início da ataxia cerebelar. Drogas antiepiléticas, como carbamazepina, fenitoína e ácido valpróico são eficientes na maioria dos casos (Rasmussen et al., 2001; Grewal et al., 2002).

O eletroencefalograma interictal é anormal em muitos pacientes com SCA10 com epilepsia e em alguns que não apresentam convulsões. Os achados mais comuns são disfunção cortical difusa com atividades lentas e desorganizadas. Irritabilidade focal cortical ou atividades lentas também podem ser observadas em alguns indivíduos. Muito provavelmente as atividades epiléticas refletem os efeitos patogênicos fora do cerebelo, particularmente no córtex cerebral, causados pela mutação que desencadeiam a SCA10.

A prevalência das convulsões em pacientes mexicanos é de 72,2%. Considerando a mesma família, a prevalência variou entre 25% a 80% (Matsuura et al., 2000; Rasmussen et al., 2001). Em um estudo publicado recentemente, 13 de 20 pacientes brasileiros (65%) apresentaram crises convulsivas atingindo uma prevalência próxima àquela dos mexicanos (Castilhos et al., 2013). Por outro lado, em famílias brasileiras diagnosticadas no estado do

Paraná, apenas 3 de 80 (3,75%) indivíduos acometidos apresentaram convulsões (Teive et al., 2010).

### 1.2.1.3. Sintomas extracerebelares

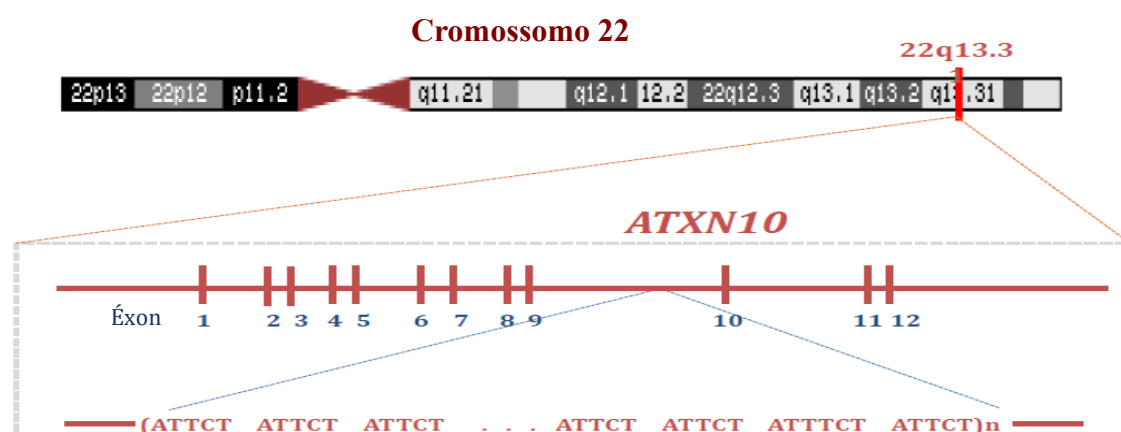
Sinais extracerebelares, tais como transtorno do humor, disfunção cognitiva, sinais piramidais leves e neuropatia periférica, também são observados em pacientes mexicanos (Matsuura et al., 2000; Grewal et al., 2002). Um grau variável de sinais piramidais foi descrito, incluindo hiperreflexia, espasmos nas pernas e sinal de Babinski. Os indivíduos acometidos reclamaram de uma perda sensitiva suave nas extremidades inferiores e estudos de condutividade nervosa confirmaram a presença de polineuropatia (Grewal et al., 2002). Alguns pacientes apresentam baixo quociente de inteligência (QI). Características depressivas, agressivas e/ou irritabilidade foram detectadas pela avaliação psiquiátrica baseada no Inventário Multifásico Minnesota da Personalidade (MMPI- *Minnesota Multiphasic Personality Inventory*) (Lin & Ashizawa, 2005; Alonso et al., 2006). Relatos de tendências violentas em alguns pacientes levaram à especulação de que irritação e agressividade possam ser manifestações subclínicas de mania e psicose em pacientes com SCA10 (Wexler & Fogel, 2011). Apenas uma família configurou um quadro de alterações hepática, cardíaca e hematológicas ainda não bem esclarecido se causado em decorrência da SCA10 ou de uma outra doença concomitante (Rasmussen et al., 2001).

Os pacientes brasileiros, por outro lado, demonstraram apenas o fenótipo cerebelar e, ocasionalmente, sinais piramidais leves (Teive et al., 2004). Três pacientes brasileiras da mesma família, durante o último mês de gestação, manifestaram sintomas no início agudo da doença muito possivelmente desencadeado por fatores hormonais (Teive et al., 2011). Distonia e parkinsonismo foi relatado em famílias argentinas (Gatto et al., 2007). Mais recentemente, foi publicado o primeiro caso de uma família proveniente da Guatemala com pai e quatro filhos afetados, sendo dois deles apresentando esquizofrenia, psicose e depressão, além do sintoma clássico, como marcha atáxica (Trikamji et al., 2015).

### 1.2.2. Aspectos moleculares

O *locus* de SCA10 foi mapeado na região 22q13.1 por análise de ligação, de forma independente, em duas grandes famílias mexicanas (Matsuura et al., 1999; Zu et al., 1999)

conforme a figura 1. A hipótese até então formulada era de que a mutação responsável pela SCA10 seria uma expansão de repetições nucleotídicas instáveis e a identificação foi realizada através de clonagem posicional. Logo após, um polimorfismo de repetições pentanucleotídeos em tandem (ATTCT)<sub>n</sub> no íntron 9 do gene *E46L* foi caracterizado como sendo o responsável pela SCA10 (Matsuura et al., 2000). Atualmente este gene é denominado *ATXN10* e abrange 172,8kb de DNA genômico com uma sequência aberta de leitura (ORF – *open reading frame*) de 1.428pb (Matsuura et al., 2000).



**Figura 1: Representação esquemática do gene *ATXN10*.** A mutação que causa a doença SCA10 é uma mutação dinâmica de repetições pentanucleotídicas ATTCT e está localizada no íntron 9 do gene *ATXN10*. (Adaptado de Teive et al., 2004)

O número de repetições é polimórfico na espécie humana, mas, quando o número dessas repetições está além da normalidade, pode desencadear na neurodegeneração. Indivíduos normais apresentam de 10 a 29 repetições ATTCT, sendo 82% deles heterozigotos compostos dentro deste espectro de variação (Matsuura et al., 2000). Alelos expandidos variam entre 800 (4kb) a 4500 repetições (22,5kb). A mutação causadora de SCA10 é, portanto, a maior expansão de microssatélites conhecida no genoma humano até o momento (Matsuura & Ashizawa, 2012).

Em famílias brasileiras com SCA10 de ancestralidade portuguesa e ameríndia, os alelos expandidos apresentavam 400 e 760 repetições ATTCT. Porém, a presença ou não dos sintomas é variável entre os indivíduos dentro da mesma família. Esses achados sugerem que o limiar patogênico das repetições possa ser inferior ao limiar considerado até o momento (Alonso et al., 2006). Em uma grande família brasileira de sete indivíduos com aproximadamente 850 repetições em um alelo, apenas um indivíduo manifestou a doença aos quatro anos de idade; os demais são assintomáticos até a data da publicação (Raskin et

al., 2007). O grande intervalo entre as repetições normais e as patogênicas possibilita a ocorrência de fenótipo pré-mutacional sugerindo casos de penetrância reduzida da mutação. Estudos adicionais são necessários para determinar em qual faixa entre 33 a 850 repetições ATTCT está associada à penetrância reduzida (Raskin et al., 2007). Em um indivíduo afetado com idade de início precoce (14 anos) foi identificado um alelo com 280 repetições, sendo que o mesmo alelo foi também encontrado na mãe assintomática.

### 1.2.3. Aspectos genéticos

A SCA10 é uma doença com padrão de herança autossômico dominante. As repetições pentanucleotídicas (ATTCT) expandidas no gene *ATXN10* são a causa primária de SCA10. Todas as expansões de repetições ATTCT segregam com o fenótipo de SCA10 em famílias afetadas e estão ausentes em mais de 1000 cromossomos normais (Ashizawa et al., 2004). Além disso, o tamanho da expansão repetida apresenta uma correlação inversa com a idade de início das manifestações clínicas. Por fim, o gene *ATXN10* é amplamente expresso no cérebro; a diminuição da sua expressão afeta preferencialmente o sistema nervoso central, como o cerebelo e o córtex cerebral (Ashizawa et al., 2004).

A antecipação é um fenômeno clínico associado, geralmente, à idade de início precoce dos sintomas e/ou a um fenótipo mais grave nas gerações sucessivas. Antecipação foi descrita numa grande família com SCA10 (Zu et al., 1999). Esse fenômeno reflete a natureza dinâmica de algumas mutações, as quais podem mudar de uma geração para outra, aumentando durante o processo de replicação (Lin et al., 2005). O coeficiente de correlação inversa entre o tamanho da expansão e a idade de início da doença foi 0,34, sugerindo que o tamanho da repetição ATTCT pode explicar apenas um terço da variação da idade de início. Outros fatores, portanto, devem influenciar a idade de início dos sintomas. Essa é uma característica encontrada num número considerável de doenças neurológicas, como Síndrome do X frágil, doença de Huntington, distrofia miotônica tipo I e outras SCAs (Ashizawa et al., 2004).

As repetições ATTCT expandidas demonstram instabilidade dependendo do tamanho da repetição quando da transmissão de geração para geração. O padrão de instabilidade é dependente da origem de herança: paterna ou materna. A expansão em SCA10 exhibe uma tendência à transmissão paterna, ocorrendo mais frequentemente durante a gametogênese masculina. Por outro lado, a transmissão materna é principalmente

acompanhada por pequena ou nenhuma mudança no número de repetições. Entretanto, algumas contrações no número de repetições de uma geração para outra associado à antecipação também foi observado na transmissão paterna, ressaltando uma complexa correlação genótipo-fenótipo ao que se refere o tamanho das repetições e a idade de início dos primeiros sintomas (Matsuura et al., 2004). Também foi evidenciada a presença de mosaicismos tanto intra- quanto inter-geracional, quando analisadas linhagens somáticas de diferentes tecidos bem como em linhagens germinativas, indicando instabilidade mitótica e meiótica (Matsuura et al., 2004). Essa propriedade dinâmica da expansão das repetições nucleotídicas contribui para variadas manifestações clínicas da doença.

Os padrões de expressão fenotípica e instabilidade são variáveis tanto dentro quanto entre famílias. Um dos mecanismos propostos para esses fenômenos é de que a configuração dos alelos pode variar (Matsuura et al., 2004). Neste estudo, duas famílias foram descritas com fenótipos distintos ao que se refere à ocorrência de convulsões e correlação do tamanho das repetições com a idade de início da doença. Todos os pacientes de uma mesma família mostraram repetições contínuas “puras”, ao passo que numa segunda família as repetições revelaram um padrão complexo, com duas interrupções repetitivas, ATTTCT e ATATTCT. Esse padrão foi semelhante entre os indivíduos afetados da família ao longo de três gerações.

Esses achados questionam a ideia consagrada de que a doença é causada por expansões ininterruptas, sugerindo que o grau de pureza das repetições pode surgir como um fator modificador da doença (Matsuura et al., 2006). O mesmo pode ser verificado em outras doenças neurológicas tais como SCA1, SCA2, ataxia de Friedrich, síndrome do X-frágil, distrofia miotônica tipo I, nas quais as sequências ininterruptas parecem estabilizar as unidades repetitivas e a perda das interrupções está associada à instabilidade e a expansões repetitivas (Matsuura et al. 2000; Matsuura et al., 2006; Braida et al., 2010). Ainda que outros estudos sejam necessários para confirmar essa hipótese, ela surge como uma possível explicação para a variação na frequência de convulsões entre famílias mexicanas e brasileiras.

Nenhuma correlação entre o tamanho dos alelos expandidos e o fenótipo de convulsões foi demonstrado (Teive et al., 2004). Também, nenhuma correlação entre o tamanho das expansões e a gravidade da doença foi encontrada, sugerindo que fatores família-dependentes podem influenciar nesse aspecto (Matsuura & Ashizawa, 2012).



Expansões nas repetições ATTCT do tipo *de novo* ainda não foram identificadas (Teive et al., 2011).

#### 1.2.4. A proteína e aspectos fisiopatológicos

O produto normal do gene *ATXN10* é a proteína ataxina 10, a qual é composta por 475 aminoácidos (März et al., 2004). Análises da sequência de aminoácidos da ataxina 10 demonstram que a proteína não possui domínio transmembrânico, nem localização nuclear, ou outro tipo de peptídeo sinal, *clusters* ou padrões incomuns de aminoácidos ionizados (März et al., 2004). Análises citológicas mostraram que a ataxina 10 está localizada predominantemente no citoplasma do cérebro de ratos, de camundongos e seres humanos. Ela é uma proteína globular que tende a formar um complexo homotrimérico ponta a ponta e contém duas repetições *armadillo* na extremidade carbóxi-terminal (aminoácidos 287-433). Proteínas associadas a membranas, tal como a  $\beta$ -catenina, também apresentam essa configuração demonstrando que a proteína ataxina 10 pertence à família de proteínas com repetições *armadillo* (März et al., 2004).

Essa proteína é fortemente expressa no cérebro, no coração e nos músculos, mas sua expressão é também detectada em uma ampla variedade de tecidos (White et al., 2010). O real efeito patogênico da mutação ainda não está bem esclarecido. Porém, estudos em modelo animal nocaute indicam atividade vital especialmente no cerebelo, pois sua deficiência resulta em morte celular (März et al., 2004) e a superexpressão resulta na diferenciação e extensão dos axônios e dendritos das células PC12 (Waragai et al., 2006). A ataxina 10 no cérebro interage com a subunidade  $\beta 2$  da proteína G e com a transferase de  $\beta$ -N-acetilglucosamina ligada a oxigênio, o qual reforça a hipótese desta proteína estar relacionada com o desenvolvimento neuronal (März et al., 2004; Waragai et al., 2006).

Os eventos de expansões de repetições ATTCT ocorrem no íntron 9. Os estudos *in vitro* em linfoblastos e fibroblastos derivados de pacientes com SCA10 sugerem que as repetições ATTCT não interferem na transcrição e nem no processamento pós-transcricional do alelo mutante de *ATXN10*. A repetição expandida ATTCT é transcrita em AUUCU expandida no mRNA mutante não processado e o íntron 9, que contém as expansões AUUCU é corretamente removido por mecanismo de *splicing* (White et al., 2010).

Conseqüentemente, os níveis de mRNA processados permanecem inalterados. Entretanto, camundongos *knock-out* para *ATXN10* morreram em estágios embrionários. Os heterozigotos Sca10-nulo não manifestam modificações fenotípicas (Wakamiya et al., 2006). As repetições nucleotídicas não interferem na transcrição dos genes vizinhos *PPARA/FBLN1*. Além disso, uma translocação no íntron 2 do gene *ATXN10*, responsável pela interrupção do gene e pela haploinsuficiência da ataxina 10, não causa SCA10 em humanos (Keren et al., 2010). Dessa forma, perda ou ganho de função, haploinsuficiência e interferência em genes vizinhos muito provavelmente não interferem na função fisiológica da ataxina 10 e, portanto, não parecem ser os fatores responsáveis pela SCA10 (Wakamiya et al., 2006).

Alternativamente, a ideia de que o mecanismo causador da SCA10 é mediado por um possível efeito tóxico do ganho de função do RNA com a expansão das repetições AUUCUU tem sido considerada. White e colaboradores mostraram que o RNA com a repetição expandida AUUCUU no transcrito mutante do gene *ATXN10* interage com a proteína nuclear hnRNP K, levando à apoptose via translocação massiva da proteína quinase C (PKC) para a mitocôndria, sugerindo que a SCA10 é causada por esse mecanismo. A interação do RNA anômalo com outras proteínas leva à formação de agregados (White et al., 2010).

Para confirmar a toxicidade mediada pelo ganho de função do RNA, um novo modelo de camundongo transgênico foi desenvolvido, no qual as repetições expandidas de pentanucleotídeos foram reconstruídas na região 3' não traduzida (3' UTR) para garantir que as repetições fossem transcritas, mas não traduzidas. Os camundongos foram submetidos a testes para avaliar mudanças neurológicas, histopatológica e celulares. Aos 6 meses, eles apresentaram marcha irregular e um aumento na suscetibilidade de convulsões, semelhante ao fenótipo clínico de SCA10. O córtex cerebral, o hipocampo e o núcleo da ponte também apresentaram perda neuronal. Mudanças moleculares e celulares semelhantes às previamente encontradas no modelo celular de SCA10 foram evidenciadas no cérebro desses animais (White et al., 2012).

#### 1.2.5. Diagnóstico

Devido à grande semelhança do quadro sintomatológico entre várias SCAs e à grande variabilidade fenotípica observada dentro das famílias, um diagnóstico confiável não pode ser baseado tão somente em aspectos clínicos. Ferramentas moleculares se tornam úteis para a confirmação do diagnóstico, bem como para a testagem preditiva. O número elevado de repetições pentanucleotídicas presente no alelo mutante do gene *ATXN10* não permite o uso da PCR como opção diagnóstica. A análise por *Southern Blot* é a metodologia mais indicada. Entretanto, essa metodologia é trabalhosa e a obtenção do resultado é demorada, o que dificulta o seu uso em laboratórios de diagnóstico (Matsuura & Ashizawa, 2002). Nesse caso, a metodologia de RP-PCR (*Repeat Primed PCR*) surge como uma alternativa para identificação de alelos mutantes. Entretanto, esta metodologia apresenta limitações, como a impossibilidade de quantificação do número de repetições que restringe a definição do número de repetições nucleotídicas presentes na amostra analisada.

#### 1.2.6. Ancestralidade

Uma vez elucidada as bases genéticas da doença, muitas famílias de grupos étnicos distintos com ataxia espinocerebelar autossômica dominante foram testadas (Matsuura et al., 2000). Pacientes com ataxia espinocerebelar autossômica dominante sem mutações identificadas para outras SCAs e DRPLA, incluindo indivíduos norteamericanos eurodescendentes, francocanadenses, italianos, franceses, espanhóis, portugueses, poloneses ou populações asiáticas como japoneses, chineses, foram testados para a expansão no gene *ATXN10* e os resultados desses estudos foram negativos (Fujigazaki et al., 2002; Matsuura et al., 2002; Sasaki et al., 2003; Brusco et al., 2004; Jiang et al., 2005; Seixas et al., 2005; Wang et al., 2010; Vale et al., 2010; Sulek-Piatkowska et al., 2010).

As histórias genealógicas e as características fenotípicas dessas famílias com SCA10 sugerem ancestralidade ameríndia nesses pacientes. Uma origem ancestral comum da expansão das repetições ATTCT em pacientes com SCA10 nas Américas foi determinada como sendo possivelmente proveniente da população ameríndia, a qual se espalhou para as populações miscigenadas do México, do Brasil e, provavelmente, para outros países da América Latina devido a um efeito fundador (Ashizawa et al., 2001). Entretanto, existem poucas evidências moleculares que corroboram essa hipótese.

### 1.3. Justificativa

Os dados disponíveis até o momento indicam uma origem ancestral comum da mutação, conforme trabalhos publicados anteriormente (Almeida et al., 2009). Um trabalho recente do nosso grupo identificou um haplótipo comum em 19 dos 20 pacientes estudados (12 famílias) com SCA10 pertencentes à coorte da Rede Neurogenética (Gheno, 2012). Esse haplótipo é o mesmo identificado previamente em 3 famílias do Rio Grande do Sul e em 1 das 3 famílias mexicanas incluídas em um estudo anterior (Almeida et al., 2009). A frequência do mesmo é comum em indivíduos controles mas difere significativamente quando comparada aos pacientes (Gheno, 2012).

Esses achados corroboram a hipótese de uma origem ancestral comum para SCA10 nas Américas. Provavelmente a mutação tenha se fixado na população americana pela existência de um efeito fundador sendo descrita apenas em populações ameríndias (Almeida et al., 2009). A determinação de haplótipos através de marcadores polimórficos em regiões próximas ao gene *ATXN10* é importante para a confirmação da origem ancestral ameríndia da mutação.

## 2. OBJETIVOS

### 2.1. Objetivo geral

O objetivo desse trabalho foi investigar a origem ancestral da mutação no gene *ATXN10*.

### 2.2. Objetivos específicos

- a) confirmar a ocorrência de um haplótipo ancestral comum em pacientes com SCA10 de diferentes origens geográficas;
- b) identificar marcadores próximos ao gene *ATXN10* que sejam informativos de ancestralidade ameríndia;
- c) inferir os haplótipos nos pacientes incluídos neste estudo;
- d) comparar a distribuição desses haplótipos entre grupos de pacientes;
- e) correlacionar e identificar a origem dos haplótipos associados ao alelo mutante;
- f) avaliar a relação filogenética entre os haplótipos.

### **3. RESULTADOS**

Os resultados deste trabalho serão apresentados a seguir sob a forma de manuscrito, o qual será submetido à publicação no periódico *European Journal of Neurology*.

## **Haplotype study in SCA10 families provides further evidence for a common ancestral origin of the mutation**

Giovana B Bampi, MSc<sup>1,2,3</sup>; Rafael Bisso-Machado, PhD<sup>1</sup>; Tábita Hünemeier, PhD<sup>1</sup>; Tailise C Gheno, MSc<sup>1,3</sup>; Gabriel V Furtado, MSc<sup>1,3</sup>; Diego Veliz-Otani<sup>4</sup>; Mario Cornejo-Olivas, MD<sup>4</sup>, Pillar Mazzeti, MD, PhD<sup>4</sup>; Maria Cátira Bortolini, PhD<sup>1</sup>; Laura B Jardim, MD, PhD<sup>1,2,3,5</sup>; Maria Luiza Saraiva-Pereira, PhD<sup>1,2,3,6</sup>

<sup>1</sup>Department of Genetics – Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

<sup>2</sup>Medical Genetics Service – Hospital de Clínicas de Porto Alegre, Brazil

<sup>3</sup>Laboratory of Genetics Identification – Research Center – Hospital de Clínicas de Porto Alegre, Brazil

<sup>4</sup>Neurogenetics Research Center – Instituto de Ciencias Neurológicas, Peru

<sup>5</sup>Department of Internal Medicine – Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

<sup>6</sup>Department of Biochemistry – Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

**Running title:** Haplotype analysis and ancestral origin of the mutation in SCA10

### **Corresponding author:**

Maria Luiza Saraiva-Pereira, PhD

Hospital de Clínicas de Porto Alegre – Serviço de Genética Médica

Ramiro Barcelos 2350 – CEP 90035-903

Porto Alegre, Rio Grande do Sul, Brazil

e-mail: mlpereira@hcpa.ufrgs.br

Phone: 55 51 3359 8011

Fax: 55 51 3359 8010

**Total number of words in the text (excluding references, tables, and figures): 2434**

## ABSTRACT

**Background** Spinocerebellar ataxia type 10 (SCA10) is an autosomal dominant neurodegenerative disorder characterized by progressive cerebellar ataxia and epilepsy. The disease is caused by a pentanucleotide ATTCT expansion in intron 9 of the *ATXN10* gene on chromosome 22q13.3. So far SCA10 has shown a geographical distribution throughout America since it has only been described in admixed patients self-reporting Amerindian ancestry from different countries. Currently available data suggests that SCA10 mutation probably might have arisen in Amerindian population and has spread out to the rest of the continent during the peopling of the Americas. The real ancestral origin of SCA10 mutations remains under speculation.

**Methods** Samples of SCA10 patients from two Latin American countries were analyzed, being 16 families from Brazil (29 patients) and 21 families from Peru (27 patients) as well as 49 healthy individuals from Indigenous Quechua population and 51 healthy Brazilian individuals. Four polymorphic markers spanning a region of 5.2 cM harboring the ATTCT expansion were used to define the haplotypes, which were genotyped by different approaches.

**Results** Our data has shown that 19CGGC14-shared haplotype was found in 46.8% of Brazilian and in 62.8% of Peruvian families. Frequencies from both groups are not statistically different from Quechua controls (57.3%), but they are statistically different from Brazilian controls (11.8%) ( $p < 0.001$ ). The second most frequent haplotype in Quechuas, 19-15-CGGC-14-10, is found in 50.0% of Brazilian and in 64.7% of Peruvian families with SCA10.

**Conclusions** These findings bring valuable evidence that ATTCT expansion may have arisen in a Native American chromosome.

**KEYWORDS:** SCA10, *ATXN10*, haplotypes, ancestral origin, Amerindians, founder effect



## INTRODUCTION

Spinocerebellar ataxia type 10 (SCA10, OMIM#603516) is a late onset autosomal dominant neurodegenerative disorder characterized by progressive cerebellar ataxia and epilepsy [1,2]. Extra-cerebellar symptoms can be also observed such as psychiatric disorder, cognitive disability, polyneuropathy, and pyramidal signs [1-5]. Apparently, clinical manifestation may vary according to patients' ethnical background [3-5]. The disease is caused by an expansion of the pentanucleotide repeat ATTCT in intron 9 of the *ATXN10* gene, which is located at chromosome 22q13.3 [1,2]. Normal alleles can vary from 9 to 32 repeats, while alleles from affected subjects range from 800 up to 4,500 repeats [6-9]. Reduced penetrance has been found for intermediate size alleles of 280 to 850 repeats [10,11].

So far SCA10 has shown a geographical distribution throughout the Americas since it has only been described in patients from Mexico, Brazil, Argentina, Venezuela, Colombia, United States and more recently Peru [12-20]. Interestingly, the majority of patients has self-declared a family history of Amerindian ancestry, and the literature specifically reports two affected subjects descending from the Yaqui, in Mexico, and from the Sioux, also from North America [17,18]. Previous haplotype studies in a limited number of Brazilian and Mexican families have shown a shared common haplotype [21].

Hence, currently available data suggests that SCA10 mutation might have arisen in Native American populations and it is present nowadays in urban populations due to admixture with Europeans and Africans who arrived in the Americas in the XVI and XVII centuries [22-25]. Therefore, the real ancestral origin of SCA10 mutations remains under speculation. Herein, we present the most extensive haplotype study in patients from two different geographical regions bringing molecular evidence that sustains the hypothesis of Amerindian ancestral founder mutation in SCA10 patients.

## MATERIALS AND METHODS

### *SCA10 families and control population*

This study was carried out in 16 Brazilian and 21 Peruvian unrelated families including 29 and 27 individuals, respectively. Brazilian families are from 5 different Brazilian States: 7 from Rio Grande do Sul, 3 from Santa Catarina, 3 from São Paulo, 1 from Goiás, and 2 from Rio Grande do Norte. Peruvians families are also from different regions of Peru, mainly from Pacific coast and highlands regions. Samples from Peru and from Brazilian States but Rio Grande do Sul were available through Rede Neurogenética. Inclusion criteria were suggestive clinical signs and symptoms of ataxia with a confirmed molecular diagnosis of SCA10 (an expanded ATTCT allele). Samples from 49 unrelated Quechua individuals and 51 unrelated Brazilians individuals were included as controls [21]. All participants provided signed informed consent. This study is part of a more comprehensive project, which was approved by the Ethics Committee of Clinical Hospital of Porto Alegre (project #07-259).

### *ATXN10 Mutation analysis*

Peripheral blood samples (5 mL) were collected from patients and healthy individuals. Genomic DNA was isolated from peripheral blood leukocytes using salting-out method [26]. Expanded alleles were detected by repeat-primed polymerase chain reaction (RP-PCR) using a *locus* specific primer, a repeat specific primer, and a tailed primer (Table S1). Amplification was carried out under conditions previously described [27].

### *Genotyping of polymorphisms flanking the ATXN10 gene*

All subjects were genotyped for four highly polymorphic short tandem repeats (STRs) and four single nucleotide polymorphisms (SNPs). Initially, haplotypes were constructed with four SNPs (rs5764850 - C/A, rs72556348 - G/A, rs72556349 - G/A, rs72556350 - C/T) and two dinucleotides STRs (D22S1140 and D22S1153). In order to gain insight about the ancestral origin of the haplotype two tetranucleotides STRs (D22S532 and GATA030P)

were added to the original set of markers, spanning a region of 5.2 cM harboring the ATTCT repeats. The entire multiloci combination is shown in Figure 1. Amplification of the polymorphic regions was performed with primer sequences listed in Table S1. Three different genotyping methods were used. First, rs5764850 was determined by qualitative assay by real-time PCR using a validated TaqMan<sup>®</sup> PCR Assay (Applied Biosystems). PCR reaction was performed in a final volume of 8 $\mu$ L containing 4ng of DNA, 0.2 $\mu$ L of specific TaqMan assay, and 4 $\mu$ L of 2x PCR Genotyping Master Mix (Applied Biosystems) (Assay ID C\_30474239\_39). Amplification included an initial step at 50°C for 2 min (activation of AmpErase UNG function), 95°C for 10 min (AmpliTaq<sup>®</sup>Gold activation), followed by 40 cycles of denaturation at 95°C for 15 seconds and annealing/extension at 60°C for 1 min. Allelic discrimination step was performed at 60°C for 1 minute. Amplicons were analyzed by Sequence Detection System software version 1.2.1 (Applied Biosystems) in the ABI Prism 7500 Sequence Detector System (Applied Biosystems) through allelic discrimination plot.

The remaining SNPs were genotyped by DNA sequencing. The three SNPs were amplified in a single fragment of 483 bp. A 25 $\mu$ L reaction was prepared containing 100ng of genomic DNA, 5.0pmol of each primers ATXN10SNPR and ATXN10SNPF1; 200 $\mu$ M dNTPs (Amersham-Pharmacia), 10mM Tris-HCl (pH 8.3), 50mM KCl, 0.75mM MgCl<sub>2</sub>, and 1.5U of Taq DNA Polymerase (Invitrogen). Amplification was performed in a thermal cycler as follows: initial denaturation at 94°C for 10min, followed by 30 cycles at 94°C for 30 sec, 63°C for 1 min and 72°C for 30 sec, and final extension at 72°C for 10 min. Amplification was verified by resolving an aliquot using a 1.5%(w/v) electrophoresis gel. Amplicons were submitted to 2.5U Exonuclease I and 0.25U Shrimp Alkaline Phosphatase purification before sequencing reaction, which was performed using the BigDye<sup>®</sup> Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems), according to manufacturer's instructions. Amplicons were resolved through capillary electrophoresis on an ABI3130xl Genetic Analyzer (Applied Biosystems).

Finally, the identification of four STRs markers were performed simultaneously by multiplex PCR reaction carried out in a final volume of 25 $\mu$ L, containing 20ng of genomic DNA, 2.5 pmol of each primer, 200 $\mu$ M dNTPs (Amersham-Pharmacia), 15mM Tris-HCl

(pH 8.3), 75 mM KCl (PCR Buffer II – Applied Biosystems); 0.75 mM MgCl<sub>2</sub> and 1U of AmpliTaq<sup>®</sup>Gold DNA polymerase (Applied Biosystems). One primer of each pair was labeled on its 5' end with different fluorescent dye in order to discriminate each fragment. Multiplex PCR amplification was performed in a thermal cycler, followed by 30 cycles as: initial denaturation at 95°C for 10 min and 56°C for 30 seconds for annealing, and final extension at 72°C for 5 min. After amplification, an aliquot of PCR products was mixed with formamide (HiDi<sup>™</sup> Formamide, Applied Biosystems) and GeneScan<sup>™</sup> 500-LIZ (Applied Biosystems), and capillary electrophoresis was performed in an ABI 3130xl Genetic Analyzer (Applied Biosystems) using a 36cm x 50µm capillary containing Performance Optimized Polymer-7 (POP-7, Applied Biosystems). Samples were injected during 22s at 3.0kV and eletrophoresed at 15kV for 40 min at 60°C. Amplicon lengths were estimated using GeneScan<sup>™</sup> 500-LIZ molecular weight through GeneMapper<sup>®</sup> 3.2 software (Applied Biosystems). Samples with length determined by DNA sequencing were used as internal controls in the amplification reaction.

#### *Statistical analysis*

Allele frequencies were determined by direct counting. Haplotypes were inferred using PHASE v. 2.1.1 [28]. Statistical analysis was performed using WINPEPI program. Chi-square test was used to compare alleles frequencies between controls and patients. The null hypothesis was rejected at  $p < 0.05$ . Phylogenetic analysis was conducted employing Network Software 4.6.1.3 to assign the phylogenetic relationship between haplotypes [29].

## RESULTS

A total of 29 Brazilian (16 families) and 27 Peruvian (21 families) SCA10 patients, and 49 healthy Quechua controls were genotyped using eight ATXN10 gene-flanking markers. Polymorphic alleles markers were arbitrarily named according to the number of repetitive tracts. For those families with more than one affected member, the genotype of the index case was chosen to perform statistical analysis in order to avoid counting the same haplotype twice. Taking into account the intragenic SNPs, CGGC haplotype was conserved in all SCA10 families, being the most common among Brazilians (78.2%), Peruvians (90.5%), and Quechuas (94.9%).

Allelic frequencies distribution of genetic markers among all samples included in this study are shown in Table S2. D22S532 is the only marker in which allelic frequencies showed statically difference between each patients groups and Quechuas controls. Initially, haplotypes were constructed with informative SNPs and STRs D22S1140 and D22S1153, spanning a region of ~ 1.7Mb, flanking the repeats on both sides. Results showed 16 and 13 different haplotypes defined within Brazilian and Peruvian families, respectively. Considering those, 2 Brazilian haplotypes were linked to SCA10 mutation while 5 different Peruvians haplotypes were SCA10-linked (Table 1). Amerindian Quechua control population showed 14 different haplotypes. The most frequent SCA10-linked shared haplotype was 19CGGC14 accounting for 53.0% Brazilian and 59.5% Peruvian families. This haplotype is also present in Amerindian Quechuas with a very similar frequency (57.0%) as shown in Table S3. No statistically significant differences were found between patients and controls. Considering all markers as an extended haplotype, the most frequent one in both patients groups was 19-15-CGGC-14-10.

## DISCUSSION

SCA10 is a rare disease with unknown prevalence and it is believed to exist anywhere also but in America [30-36]. The unique occurrence of SCA10 ever observed outside America was reported in a patient born in the Paijan region of Peru, who declared pure Amerindian ancestry [18]. The vast majority of affected individuals had self-declared as having Amerindian ancestry.

In this study, common haplotype 19CGGC14 was demonstrated to be present in a large cohort of SCA10 families from Brazil and Peru. Frequencies regarding this 19CGGC14 SCA10-linked do not differ among groups (Table 1). This haplotype has been previously denominated as the “SCA10 haplotype” once it was carried by 2 families from the Southernmost state of Brazil and in one Mexican family, and in 19 out of 20 patients from a previous study of our group [37], signaling that the mutation has a common ancestral origin [20].

The Quechua from Peru were chosen for the present study due to their elevated contribution to the genetic pool of the Peruvian population. All Peruvian SCA10 cases in our cohort show geographic distribution overlapping with Quechua speaking-regions. Interestingly, our data has shown that 19CGGC14 shared haplotype is the most frequent among Quechua controls (57.3%), which does not differ statistically from the frequency in Brazilian (46.8%) and Peruvian patients (62.8%), but are statistically different from Brazilian controls (11.8%) ( $p < 0.001$ ) as shown in Table 2 [20].

Still, this data has to be interpreted carefully when extrapolating to affirm the exclusive Amerindian origin of SCA10 mutation, as previous study has shown this multiloci combination is quite frequent in normal chromosomes of Brazilian population as well as in Mexican and Portuguese [20]. Moreover, due to the great admixture of American population, whether the causative mutation was originated in Native American chromosomes remains obscure.

In order to address this question we then extended the haplotype by including two more STRs, D22S532 and GATA030P, since considering this inclusion there are two forming

haplotypes, h2 (13-10) and h3 (15-10) carried by almost 40% of the Amerindians [38]. The two most frequent extended haplotypes in controls are 19-13-CGGC-14-10 (14.2%) followed by 13-15-CGGC-14-10 (13.2%) (Table S4). The 13-15-CGGC-14-10 is carried by 50.0% of Brazilians and 64.7% Peruvians families as shown in Table 3. The causative mutation may not necessarily have occurred in the most prevalent haplotype. Frequencies of referring haplotypes were compared among native populations around the world and among 19CGGC14 SCA10-linked in both patients groups and are shown in Figure 2. Interestingly, h3 is shown to be in *cis* with 19CGGC14 and the mutant allele, and is present in 7 out of 14 Brazilian (50.0%) and in 11 out of 17 (64.7%) Peruvian families. These findings are in agreement to our previous hypothesis of a common ancestry of the mutation and further reinforce that the studied block of markers flanking ATTCT mutation is probably in an Amerindian portion of the chromosome. Furthermore, a founder effect might be contributing to the distribution of SCA10 throughout American countries [20].

Data from different fields of knowledge, including Genetics, give support to the Asian origin of nowadays Native Americans. They would have migrated from somewhere near southwest Siberia, passing through Bering Strait and finally reaching the American continent through North America. However, there is no consensus on arrival data of these migrants, and on the number of main migration waves as well. Nevertheless, Bering played a central role, as the migrating population stayed there for at least 5,000 years. Sufficient time in which numerous autochthonous mutations found only in Native Americans may have arisen [39, 40]. Moreover, it is probable that the origin of SCA10 mutation may have occurred early in their entry into the Americas, considering the wide distribution of the disease in this continent. The southward model of human dispersal pointed out to a decreasing genetic diversity as distance from Bering Strait increases [38].

Also, there is a higher level of genetic diversity and a lower level of population structure in Western South America when compared to Eastern South America. This may explain a relatively higher diversity of Peruvian haplotypes associated to the SCA10 mutation when compared to Brazilian families, since the latter have shown only two different haplotypes 19CGGC14 and 19CGGC18 associated to SCA10 mutation. The 19CGGC18 is different

from the most common haplotype by a four repeat unit in the D22S1140 marker as a result of a relatively common event in STRs represented by gain or loss of repeat units.

CGGC haplotype-lineage networks were constructed in order to better understand the phylogenetic relationship among haplotypes associated to SCA10. Networks considering CGGC with 2 STRs and CGGG with 4 STRs have shown similar pattern among groups, as shown in Figure 3 and 4, respectively. This is an additional evidence of Amerindian ancestry. A bimodal-like pattern would be seen in networks of patients if they had a different *ATXN10* gene origin, which is a pattern characteristic of historically isolated populations. Besides, both networks present long branches in its extremities that characterize population isolation-by-distance, which would be unlikely to be expected in admixed populations.

We should take into consideration that haplotypes h2 and h3 are found in other populations around the World despite their high prevalence in Native American tribes. Unfortunately, there are no exclusive genetic markers of Native American along the chromosome region assessed in the present studied that could give undoubtedly information about the SCA10 ethnic origin. Considering this limitation, additional studies should focus in evaluating the individual genetic ancestry of family members carrying the most common haplotype by employing different approaches to further assess the founder effect for SCA10 in America.



## **CONCLUSION**

The present study is the most extensive ever performed addressing the SCA10 ancestral origin and including a large number of SCA10 families from two different geographical regions in South America. Up to date the Amerindian ancestry of the mutation was hypothesized based on theoretical evidence only. We provide here additional molecular evidence that supports the Amerindian common ancestry for all patients.

## **ACKNOWLEDGEMENTS**

The authors are grateful to the patients and their families for providing biological material for this study. We would also like to thank Vanessa Paixão-Cortês, Nelson Jurandi Rosa Fagundes, and Virginia Ramallo for their contribution in this study.

## **FUNDINGS**

This research was supported by Brazilian Funding Agencies: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Fundo de Incentivo a Pesquisa e Eventos do HCPA (FIPE-HCPA).

## **CONFLICTS OF INTERESTS**

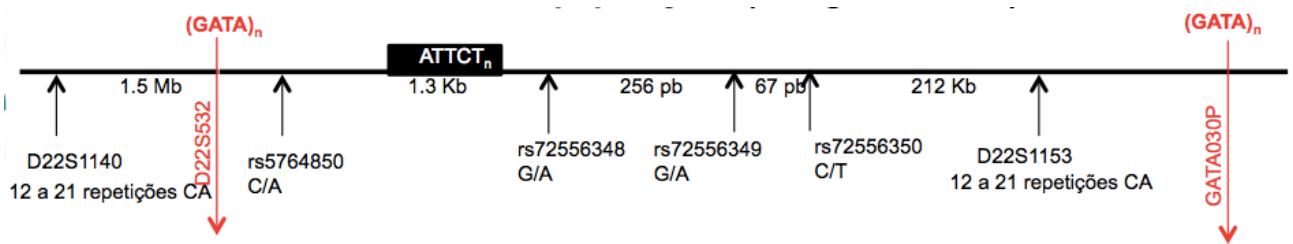
All authors declare themselves to have no conflict of interest in this research work.

## REFERENCES

1. Teive HA, Arruda WO, Raskin S, Munhoz RP, Zavala JA, Werneck LC, Ashizawa. Spinocerebellar ataxia type 10 – A review. *Parkinsonism Relat Disord* 2011, **17**:655-661.
2. Ashizawa T. Spinocerebellar ataxia type 10. *Handb Clin Neurol* 2012, **103**:507-519.
3. Teive HA, Roa BB, Raskin S, Fang P, Arruda WO, Neto YC, Gao R, Werneck LC, Ashizawa T. Clinical phenotype of Brazilian families with spinocerebellar ataxia 10. *Neurology* 2004, **63**:1509-1512.
4. Rasmussen A, Matsuura T, Ruano L, Yescas P, Ochoa A, Ashizawa T, Alonso E. Clinical and genetic analysis of four Mexican families with spinocerebellar ataxia 10. *Annu Neurol* 2001, **50**: 234-239.
5. Teive HA, Munhoz RP, Raskin S, Fang P, Arruda WO, de Paola L, Werneck LC, Ashizawa T. Spinocerebellar ataxia type 10: Frequency of epilepsy in a large sample of Brazilian patients. *Mov Disord* 2010, **25**:2875-2878.
6. Matsuura T, Achari M, Khajavi M, Bachinski LL, Zoghbi HY, Ashizawa T. Mapping of the gene for a novel spinocerebellar ataxia with pure cerebellar signs and epilepsy. *Annual Neurol* 1999, **45**:407-411.
7. Matsuura T, Yamagata T, Burgess DL, Rasmussen A, Grewal RP, Watase K, *et al*. Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. *Nat Genet* 2000, **26**:191-194.
8. Grewal RP, Tayag E, Figueroa KP, Zu L, Durazo A, Nunez C, Plust SM. Clinical and genetic analysis of a distinct autosomal dominant spinocerebellar ataxia. *Neurology* 1998, **51**:1423-1426.
9. Grewal RP, Achari M, Matsuura T, Durazo A, Tayag E, Zu L, Pulst SM, Ashizawa T. Clinical features and ATTCT repeat expansion in spinocerebellar ataxia type 10. *Ach Neurol* 2002, **59**:1285-1290.
10. Alonso I, Jardim LB, Artigas O, Saraiva-Pereira ML, Matsuura, Ashizawa T, Sequeiros J, Silveira I. Reduced penetrance of intermediate size alleles in spinocerebellar ataxia type 10. *Neurology* 2006, **66**:1602-1604.
11. Raskin S, Ashizawa T, Teive HA, Arruda WO, Fang P, Gao R, White MC, Werneck LC, Roa B. Reduced penetrance in a Brazilian family with spinocerebellar ataxia type 10. *Arch Neurol* 2007, **64**:591-594.
12. Matsuura T, Ranum LPW, Volpini V, Pandolfo M, Sasaki H, Tashiro K, Watase K, Zoghbi HY, Ashizawa T. Spinocerebellar ataxia type 10 is rare in populations other than Mexicans. *Neurology* 2002, **58**:983-984.
13. Rasmussen A, Yescas P, Matsuura T, *et al*. Molecular diagnosis of spinocerebellar ataxias in Mexican population. *Am J Hum Genet* 2000, **67** Suppl 2:A1902.
14. Gatto EM, Gao R, White MC, Uribe Roca MC, Etcheverry JL, Persi G, Poderoso JJ, Ashizawa T. Ethnic origin and extrapyramidal signs in an Argentinian spinocerebellar ataxia type 10 family. *Neurology* 2007, **69**:216-218.
15. Gallardo M, Soto A. Clinical characterization of a Venezuelan family with spinocerebellar ataxia type 10. *Mov Disord* 2009, **24**:S12.
16. Roxburgh HR, Smith OC, Lim JG, Bachman DF, Byrd E, Bird TD. The unique co-occurrence of spinocerebellar ataxia type 10 (SCA10) and Huntington disease. *J Neurol Sci* 2013, **324**:176-178

17. Kapur SS, Goldman JG. Two in one: report of a patient with spinocerebellar ataxia type 2 and 10. *Arch Neurol* 2012, **69**(9):1200-1203.
18. Bushara K, Bower M, Liu J, McFarland KN, Landrian I, Hutter D, Teive HA, Rasmussen A, Mulligan CJ, Ashizawa T. Expansion of the spinocerebellar ataxia type 10 (SCA10) repeat in a patient with Sioux Native American Ancestry. *PLoS One* 2013, **8**(11):e81342.
19. Leonardi L, Marcotulli C, McFarland NK, Tessa A, DiFabio R, Santorelli MF, Pierelli F, Ashizawa T, Casali C. Spinocerebellar ataxia type 10 in Peru: the missing link in the Amerindian origin of the disease. *J Neurol* 2014, **57**:8-11.
20. Teive HA, Ashzawa T. Spinocerebellar ataxia type 10: from Amerindians to Latin Americans. *Curr Neurol Neurosci Rep* 2013, **13**:393.
21. Almeida T, Alonso I, Martins S, Ramos EM, Azevedo L, Ohno K, Amorin A, Saraiva-Pereira ML, Jardim LB, Matsuura T, Sequeiros J, Silveira I. Ancestral origin of the ATTCT repeat expansion in spinocerebellar ataxia type 10 (SCA10). *PLoS One* 2009, **4**(2):e4553.
22. Avena S, Via M, Ziv E, Pérez-Stable JE, Gignoux RC, Dejean C *et al.* Heterogeneity in genetic admixture across different regions of Argentina. *PLoS One*, 2012 **7**(4):e34695.
23. Price AL, Patterson N, Yu F, Cox DR, Waliszewska A, *et al.* A genomewide admixture map in Latino populations. *Am J Hum Genet* 2007, **80**:1024-1036.
24. Lins TC, Vieira RG, Abreu BS, Grattapaglia D, Pereira RW. Genetic composition of Brazilian population samples based on a set of twenty-eight ancestry informative SNPs. *Am J Hum Biol* 2010, **22**:187-192.
25. Seldin MF, Tian C, Shigeta R, Scherbarth HR, Silva G, *et al.* Argentine population genetic structure: a large variance in Amerindian contribution. *Am J Phys Antropol* 2007, **132**:455-462.
26. Miller AS, Dykes DD, Plesky HF. A sample salting-out procedure of extracting DNA from nucleated cells. *Nucleic Acids Res* 1998, **16**:1215.
27. Matsuura T, Ashizawa T. Polymerase chain reaction amplification of expanded ATTCT repeat in spinocerebellar ataxia type 10. *Ann Neurol* 2002, **51**:271-2.
28. Stephans M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 2001, **68**(4):978-989.
29. Network 4.5.1.6. Available at: <http://www.fluxus-engineering.com/>.
30. Brusco A, Gellera C, Cognoli C, Saluto A, Castucci A, Michielotto C, Fetoni V, Mariotti C, Migoni N, Di Donato S, Taroni F. Molecular genetics of hereditary spinocerebellar ataxia: mutation analysis of spinocerebellar ataxia genes and CAG/CTG repeat expansion detection in 225 Italian families. *Arch Neurol* 2004, **61**:727-733.
31. Fujigasaki H, Tardieu S, Camuzat A, Stevanin G, LeGuern E, Matsuura T, Ashizawa T, Durr A, Brice A. Spinocerebellar ataxia type 10 in the French population. *Ann Neurol* 2002, **51**:408-409.
32. Sulek-Piatkowska A, Zdzienicka E, Raczyńska-Rakowicz M, Krysa W, Rajkiewicz M, *et al.* The occurrence of spinocerebellar ataxias caused by dynamic mutations in Polish patients. *Neurol Neurochir Pol* 2010, **44**: 238–245.
33. Vale J, Bugalho P, Silveira I, Sequeiros J, Guimaraes J, *et al.* Autosomal dominant cerebellar ataxia: frequency analysis and clinical characterization of 45 families from Portugal. *Eur J Neurol* 2010, **17**:124–128.

34. Jiang H, Tang BS, Xu B, Zhao GH, Shen L, *et al.* Frequency analysis of autosomal dominant spinocerebellar ataxias in mainland Chinese patients and clinical and molecular characterization of spinocerebellar ataxia type 6. *Chin Med J (Engl)* 2005, **118**: 837–843.
35. Xie QY, Liang XL, Li XH. [Molecular genetics and its clinical application in the diagnosis of spinocerebellar ataxias]. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 2005, **22**: 71–73.
36. Wang J, Shen L, Lei L, Xu Q, Zhou J, *et al.* Spinocerebellar ataxias in mainland China: an updated genetic analysis among a large cohort of familial and sporadic cases. *Zhong Nan Da Xue Xue Bao Yi Xue Ban* 2011, **36**: 482–489.
37. Gheno TC, Furtado GV, Saute JAM, Donis KC, Fontanari AMV, Emmel VE, Pedroso JL, Barsottini O, GodeiroJunior C, van der Linden H, Pereira ET, Marques Jr W, Castilhos RM, Alonso I, Sequeiros J, Cornejo-Olivas M, Mazzetti-Soler P, Torman VL, Jardim LB, Saraiva-Pereira ML, on behalf of Rede Neurogenetica. Spinocerebellar ataxia type 10 (SCA10): ancestral haplotype and progression rate in Peru and Brazil. Submitted for publication.
38. Wang S, Jr. Lewis MC, Jakobsson M, Ramachandran S, Ray N, Bedoya G, Rojas W *et al.* Genetic Variation and Population Structure in Native Americans. *PloS Genetics* 2007, **3**(11): e185.
39. Schoroeder Kb, Schurr TG, Long JC, Rosenberg NA, Crawford MH, Tarskaia LA, Osipova LP, Zhadanov SI, Smith DG. A private allele ubiquitous in the Americas. *Biol Lett* 2007, **3**:218-223.
40. Fagundes NJ, Kanitz R, Eckert R, Valls AC, Bogo MR, Salzano FM, Smith DG, Silva WA Jr, Zago MA, Ribeiro-dos-Santos SE, Petzl-Erler ML, Bonatto SL Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas *Am J Hum Genet* 2008, **82**:583-592.



**Figure 1. Scheme of the chromosome 22 region including the *ATXN10* gene harboring the SCA10 mutation.** The ATTCT repeat and its flanking region is represented, showing the location of four intragenic SNPs (rs5764850, rs72556348, rs72556349, and rs72556350) and four microsatellite markers (D22S1140, D22S532, D22S1153, and GATA030P) used for haplotype reconstruction.

**Table 1. Frequencies of haplotypes associated to SCA10**

Haplotypes	Brazilian families SCA10-linked alleles	Peruvians families SCA10-linked alleles
	(n=16)	(n=21)
19 CGGC 14	0.875 (14) <sup>a</sup>	0.808 (17) <sup>a</sup>
19 CGGC 18	0.125 (2)	0.048 (1)
21 CGGC 17	0.000	0.048 (1)
21 CGGC 20	0.000	0.048 (1)
18 CGGC 17	0.000	0.048 (1)

<sup>a</sup>Frequencies were compared using the Qui-square test ( $X^2 = 3.003$ ;  $p = 0.930$ )

**Table 2. Haplotype 19CGGC14 vs. others haplotypes among groups**

Haplotypes	Brazilian patients (16)	Peruvians patients (21)	Quechuas controls (49)	Brazilian controls (51) [21]
	n (%)	n (%)	n (%)	n (%)
19CGGC14	15 (46.8)	26 (62.8)	56 (57.3)	12 (11.8)*
others	17 (53.2)	16 (27.2)	42 (42.7)	90 (88.2)
Total	32 (100.0)	42 (100.0)	98 (100.0)	102 (100.0)

\* This group is statistically different from the others ( $\chi^2=55.0$ ;  $p<0.001$ ; sidak-adjusted pairwise comparison).



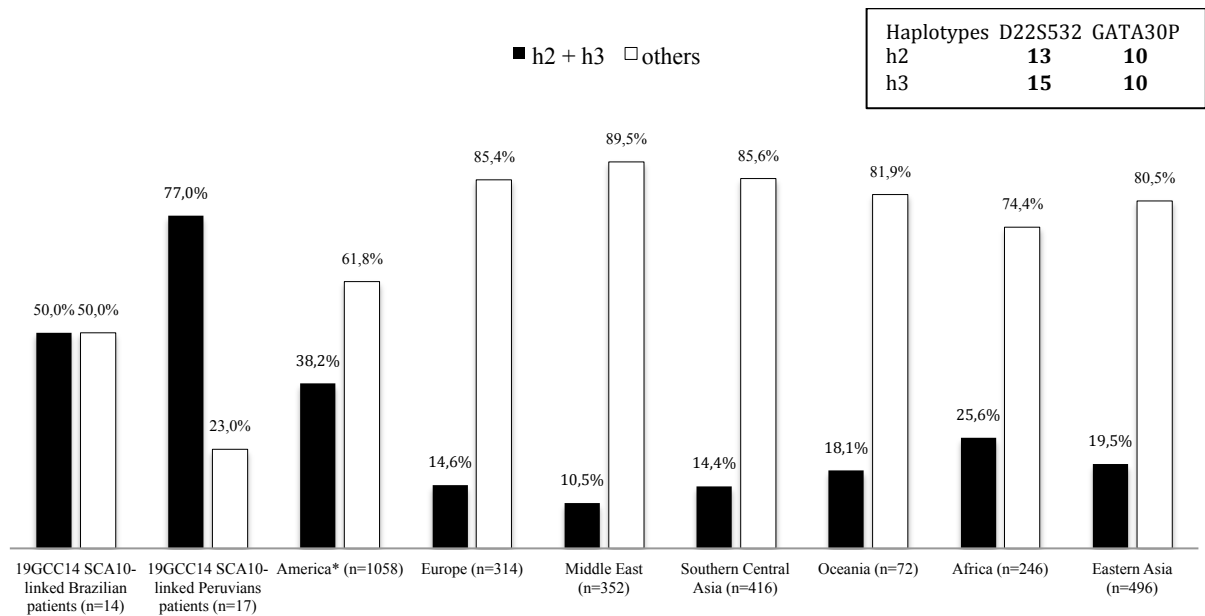
**Table 3. Frequencies of the extended haplotypes associated to SCA10**



Haplotypes	19CGGC14 SCA10-linked in Brazilians <sup>a</sup> (n=14)	19CGGC14 SCA10-linked in Peruvians <sup>a</sup> (n=17)	Quechuas 19CGGC14-carriers <sup>a</sup> (n=58)
19 <b>15</b> CGGC 14 <b>10</b>	0.500 (7)	0.646 (11)	0.224 (13)
19 <b>15</b> CGGC 14 <b>8</b>	0.357 (5)	0.000	0.017 (1)
19 <b>15</b> CGGC 14 <b>9</b>	0.143 (2)	0.059 (1)	0.052 (3)
19 <b>15</b> CGGC 14 <b>7</b>	0.000	0.059 (1)	0.000
19 <b>13</b> CGGC 14 <b>10</b>	0.000	0.118 (2)	0.241 (14)
19 <b>14</b> CGGC 14 <b>9</b>	0.000	0.059 (1)	0.017 (1)
19 <b>16</b> CGGC 14 <b>5</b>	0.000	0.059 (1)	0.000
Others <sup>b</sup>	0.000	0.000	0.449 (26)

<sup>a</sup> Frequencies were compared by Fisher's exact test with Bonferroni correction for multiple comparisons; Brazilians vs. Peruvians ( $p=0.037$ ); Brazilians vs. Quechuas ( $p<0.001$ ); Peruvians vs. Quechuas ( $p<0.001$ ).

<sup>b</sup> Remaining alleles are distributed among 26 others 19CGGC14-carrier haplotypes (Table S4)



**Figure 2. Distribution of haplotype frequencies of ancestral informative markers among ethno-geographical groups and 19CGGC14 SCA10-linked among patients.** Corresponding haplotype names to their alleles are shown in upper black box. Black bar represent the combined alleles of D22S532 and GATA30P markers that are carried by 38.2% of Amerindians according to the study of Wang et al, 2007. Other haplotypes (white bars) are distributed in low frequency among 46 different haplotypes. Interestingly, both highly prevalent haplotypes in America (h2 and h3) are also presented in half of Brazilian and 77.0% of Peruvian SCA10 patients.

**Table S1. Primers used for markers genotyping and expanded allele detection**

<b>Markers</b>	<b>Primer name</b>	<b>Primers sequence (5'-3')</b>	<b>Genotyping method</b>
D22S1140	60541F	GCCTCAGTGTCCCCATCTG	Multiplex PCR
	60541R	VIC-TCTCTGGTGTACGCGCAC	
D22S532	D22S532F	AGGAGGCAGAGGTTACAGTA	Multiplex PCR
	D22S532R	PET-GTGGTCTGAGAAGATACTGA	
rs72556348	ATXN10SNPF1	TCTTTTGGAGATGAAGTCTCTCTATGTTG	Sequencing
	ATXN10SNPR	GCAGATAACGTGATCATCACGGG	
rs72556349	ATXN10SNPF1	TCTTTTGGAGATGAAGTCTCTCTATGTTG	Sequencing
	ATXN10SNPR	GCAGATAACGTGATCATCACGGG	
rs72556350	ATXN10SNPF1	TCTTTTGGAGATGAAGTCTCTCTATGTTG	Sequencing
	ATXN10SNPR	GCAGATAACGTGATCATCACGGG	
D22S1153	6553F	TGAGCCATGATCGCACCAC	Multiplex PCR
	6553R	NED-GCCTTCCTGCCAAGTTTGG	
GATA030P	GATA030PF	GCGATCATATCTGCGATTCA	Multiplex PCR
	GATA030PR	PET-GTTTCTTCACTCCCAGCTTCTTTGCTT	
<b>Mutation</b>	<b>Primer name</b>	<b>Primers sequence (5'-3')</b>	<b>Genotyping method</b>
ATTCT expansion	SCA10TPF	FAM-CAGATGGCAGAAATGATAAACTCAA	Repeat-primed PCR <sup>27</sup>
	SCA10TPR1	TACGCATCCCAGTTTGAGACGAGAATAGAATAGA ATAGAATAGAAT	
	TPR2 (tail)	TACGACATCCCAGTTTGAGCG	

<sup>27</sup>Matsuura & Ashizawa, 2002.

**Table S2. Frequency of alleles at marker loci D22S1140, D22S532, D22S1153, GATA030P and SNPs in ATTCT-mutation-bearing and control Quechuas chromosome.**

Alleles	SCA10 patients		BF vs PF	Quechua controls	BF vs QC	PF vs QC
	Brazilians (n= 16)	Peruvians (n=21)	<i>P</i> value <sup>a</sup>	(n=49)	<i>P</i> value <sup>a</sup>	<i>P</i> value <sup>a</sup>
<b>D22S1140</b>						
12	0.094	0.047		0.041		
15	0.031	0.000		0.000		
16	0.000	0.000		0.010		
17	0.031	0.000		0.000		
18	0.031	0.024	0.640	0.020	0.016	0.203
19	0.656	0.810		0.888		
20	0.063	0.024		0.031		
21	0.094	0.095		0.010		
<b>D22S532</b>						
7	0.000	0.024		0.000		
12	0.000	0.000		0.010		
13	0.219	0.214		0.459		
14	0.063	0.143	0.776	0.204	0.001*	0.004*
15	0.688	0.571		0.327		
16	0.030	0.048		0.000		
<b>rs5764850</b>						
C	0.844	0.905		0.940		
A	0.156	0.095		0.060		
C/C	0.688	0.810	0.488	0.880	0.137	0.487
C/A	0.313	0.190		0.120		
A/A	0.000	0.000		0.000		
<b>rs72556348</b>						
G	0.875	0.970		1.000		
A	0.125	0.030		0.000		
G/G	0.750	0.952	0.159	1.000	0.003*	0.300
G/A	0.250	0.048		0.000		
A/A	0.000	0.000		0.000		
<b>rs72556349</b>						
G	0.938	0.976		0.990		
A	0.063	0.024		0.010		
G/G	0.875	0.952	0.575	0.980	0.150	0.512
G/A	0.125	0.048		0.020		
A/A	0.000	0.000		0.000		
<b>rs72556350</b>						
C	1.000	1.000		1.000		
T	0.000	0.000		0.000		
C/C	1.000	1.000	-	1.000	-	-
C/T	0.000	0.000		0.000		
A/A	0.000	0.000		0.000		
<b>D22S1153</b>						
14	0.688	0.691		0.716		
15	0.031	0.000	0.627	0.000	0.660	0.936
16	0.000	0.000		0.010		

17	0.031	0.095		0.061		
18	0.094	0.119		0.112		
19	0.031	0.000		0.020		
20	0.125	0.095		0.071		
21	0.000	0.000		0.010		
<b>GATA030P</b>						
5	0.218	0.119		0.092		
7	0.000	0.024		0.041		
8	0.250	0.119		0.153		
9	0.125	0.095	0.316	0.163	0.114	0.127
10	0.344	0.500		0.531		
11	0.063	0.143		0.020		

<sup>a</sup> *p*-values were obtained by Fisher's exact test with Bonferroni for multiple comparisons

\*statistically significant values (*p*<0.05)

BF: Brazilian families

PF: Peruvian families

QC: Quechua controls

**Table S3. Haplotypes frequencies distribution among SCA10 patients and controls.**

Haplotypes	16 Brazilian families (32 alleles)	21 Peruvians families (42 alleles)	49 Quechuas controls (98 alleles)	51 Brazilian controls <sup>[21]</sup> (102 alleles)
CGGC	0.782	0.905	0.949	0.706
AGGC	0.094	0.000	0.041	0.197
CAGC	0.031	0.071	0.000	0.000
AAAC	0.031	0.000	0.000	0.078
CAAC	0.031	0.000	0.000	0.019
AGAC	0.000	0.000	0.010	0.000
AAGC	0.031	0.024	0.000	0.000
19CGGC14	0.468	0.628	0.573	0.118
19CGGC18	0.063	0.071	0.113	0.000
19CGGC20	0.063	0.071	0.061	0.000
12AGGC14	0.031	0.024	0.000	0.020
21AAGC19	0.031	0.000	0.000	0.000
21AGGC14	0.031	0.000	0.000	0.000
20AGGC20	0.031	0.000	0.000	0.000
17AAAC20	0.031	0.000	0.000	0.020
15CGGC14	0.031	0.000	0.000	0.020
20CGGC15	0.031	0.000	0.000	0.000
21CAGC14	0.031	0.000	0.000	0.000
18CGGC14	0.031	0.000	0.020	0.020
12CGGC18	0.031	0.024	0.000	0.059
19CAAC14	0.031	0.000	0.000	0.000
19CGGC17	0.031	0.024	0.061	0.039
12CGGC14	0.031	0.000	0.041	0.020
21CGGC17	0.000	0.024	0.000	0.000
21CGGC18	0.000	0.024	0.000	0.059
19AGGC14	0.000	0.024	0.041	0.020
19CGGC19	0.000	0.000	0.020	0.000
21CGGC20	0.000	0.024	0.000	0.000
18CGGC17	0.000	0.024	0.000	0.000
20AAAC14	0.000	0.024	0.000	0.020
16CGGC14	0.000	0.000	0.010	0.000
19CGGC21	0.000	0.000	0.010	0.020
19AGAC14	0.000	0.000	0.010	0.000
20CGGC14	0.000	0.000	0.020	0.000

20CGGC20	0.000	0.000	0.010	0.000
21CGGC16	0.000	0.000	0.010	0.000
21AGGC17	0.000	0.024	0.000	0.000
12AAAC20	0.000	0.000	0.000	0.020
12CGGC17	0.000	0.000	0.000	0.039
15CGGC20	0.000	0.000	0.000	0.020
16AGGC14	0.000	0.000	0.000	0.020
16AGGC15	0.000	0.000	0.000	0.020
16CGGC15	0.000	0.000	0.000	0.020
17AGGC14	0.000	0.000	0.000	0.039
17AGGC15	0.000	0.000	0.000	0.020
17AGGC20	0.000	0.000	0.000	0.020
17CAAC18	0.000	0.000	0.000	0.020
17CGGC14	0.000	0.000	0.000	0.039
17CGGC15	0.000	0.000	0.000	0.039
17CGGC17	0.000	0.000	0.000	0.020
17CGGC18	0.000	0.000	0.000	0.039
18AAAC18	0.000	0.000	0.000	0.020
18AGGC14	0.000	0.000	0.000	0.020
20AGGC19	0.000	0.000	0.000	0.000
20CGGC18	0.000	0.000	0.000	0.020
20CGGC21	0.000	0.000	0.000	0.000
21CGGC14	0.000	0.000	0.000	0.078
21CGGC21	0.000	0.000	0.000	0.020
21AAAC14	0.000	0.000	0.000	0.020
21AGGC19	0.000	0.000	0.000	0.020
22CGGC14	0.000	0.000	0.000	0.020

---

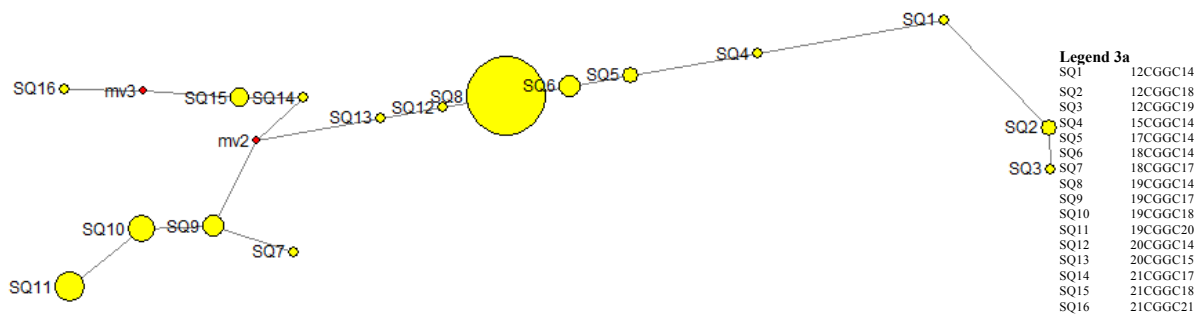
**Table S4. List of extended haplotypes and their frequencies in Indigenous Quechua controls**

#	Haplotypes	Absolute	Relative
		frequency	frequency
n=98			
1	12 <b>13</b> CGGC 20 <b>7</b>	1.000	0.010
2	12 <b>15</b> CGGC 14 <b>10</b>	3.000	0.031
3	16 <b>13</b> CGGC 14 <b>10</b>	1.000	0.010
4	18 <b>15</b> CGGC 14 <b>8</b>	1.000	0.010
5	18 <b>15</b> CGGC 14 <b>10</b>	1.000	0.010
6	19 <b>12</b> CGGC 19 <b>10</b>	1.000	0.010
7	19 <b>13</b> CGGC 14 <b>5</b>	7.000	0.071
8	19 <b>13</b> CGGC 14 <b>8</b>	3.000	0.031
9	19 <b>13</b> CGGC 14 <b>9</b>	7.000	0.071
10	19 <b>13</b> CGGC 14 <b>10</b>	14.000	0.143
11	19 <b>13</b> CGGC 14 <b>11</b>	1.000	0.010
12	19 <b>13</b> CGGC 17 <b>8</b>	1.000	0.010
13	19 <b>13</b> CGGC 18 <b>10</b>	7.000	0.071
14	19 <b>13</b> CGGC 19 <b>10</b>	1.000	0.010
15	19 <b>14</b> CGGC 14 <b>7</b>	1.000	0.010
16	19 <b>14</b> CGGC 14 <b>8</b>	3.000	0.031
17	19 <b>14</b> CGGC 14 <b>9</b>	1.000	0.010
18	19 <b>14</b> CGGC 14 <b>10</b>	2.000	0.020
19	19 <b>14</b> CGGC 17 <b>8</b>	3.000	0.031
20	19 <b>14</b> CGGC 17 <b>10</b>	1.000	0.010
21	19 <b>14</b> CGGC 18 <b>8</b>	3.000	0.031
22	19 <b>14</b> CGGC 20 <b>9</b>	2.000	0.020
23	19 <b>14</b> CGGC 20 <b>10</b>	1.000	0.010
24	19 <b>14</b> AGAC 14 <b>7</b>	1.000	0.010
25	19 <b>15</b> CGGC 14 <b>5</b>	2.000	0.020
26	19 <b>15</b> CGGC 14 <b>8</b>	1.000	0.010
27	19 <b>15</b> CGGC 14 <b>9</b>	3.000	0.031
28	19 <b>15</b> CGGC 14 <b>10</b>	13.000	0.133
29	19 <b>15</b> CGGC 20 <b>9</b>	1.000	0.010
30	19 <b>15</b> CGGC 20 <b>10</b>	1.000	0.010
31	19 <b>15</b> CGGC 21 <b>10</b>	1.000	0.010
32	19 <b>15</b> AGGC 14 <b>10</b>	3.000	0.031
33	19 <b>15</b> AGGC 14 <b>11</b>	1.000	0.010
34	20 <b>13</b> CGGC 17 <b>10</b>	1.000	0.010
35	20 <b>14</b> CGGC 20 <b>9</b>	1.000	0.010
36	20 <b>15</b> CGGC 14 <b>7</b>	1.000	0.010
37	21 <b>13</b> CGGC 16 <b>9</b>	1.000	0.010
38	19 <b>14</b> CGGC 18 <b>10</b>	1.000	0.010

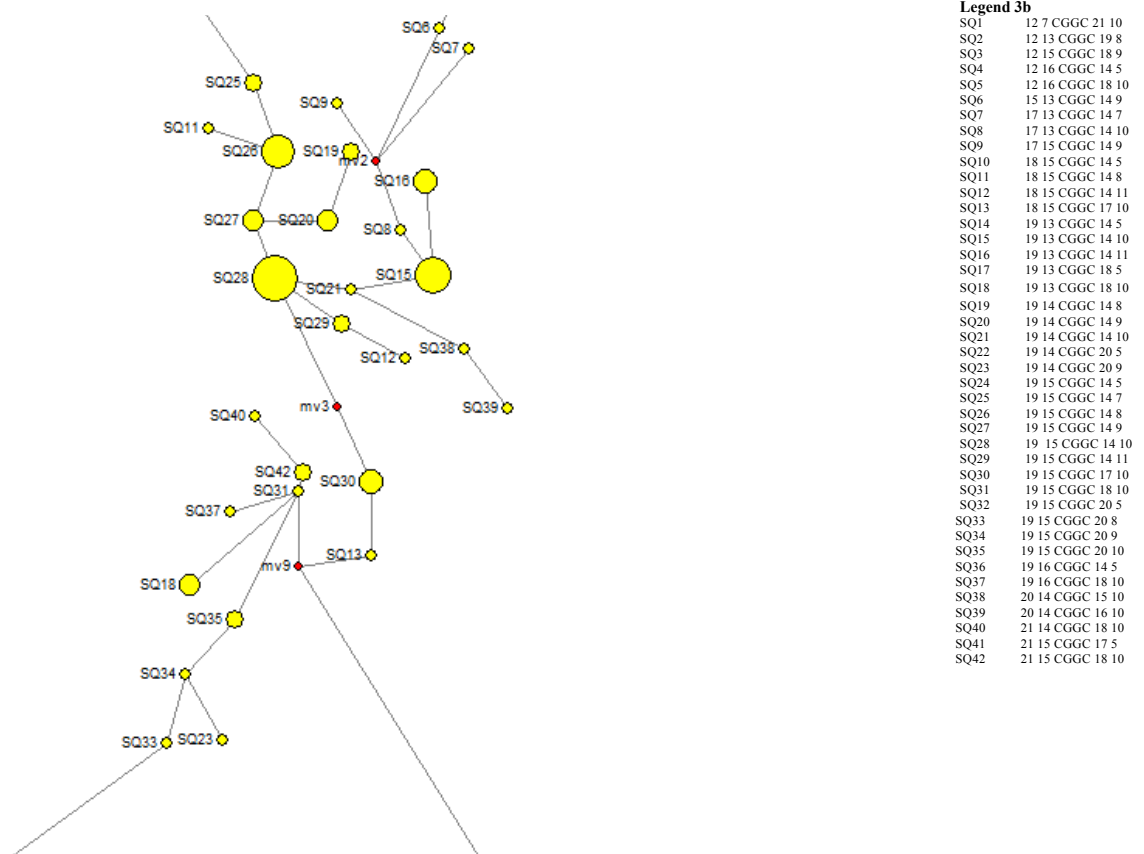
All haplotypes were inferred by PHASE<sup>[28]</sup>. Additional informative markers D22S532 and GATA030 are shown in **bold**. Two most frequent haplotypes are boxed.



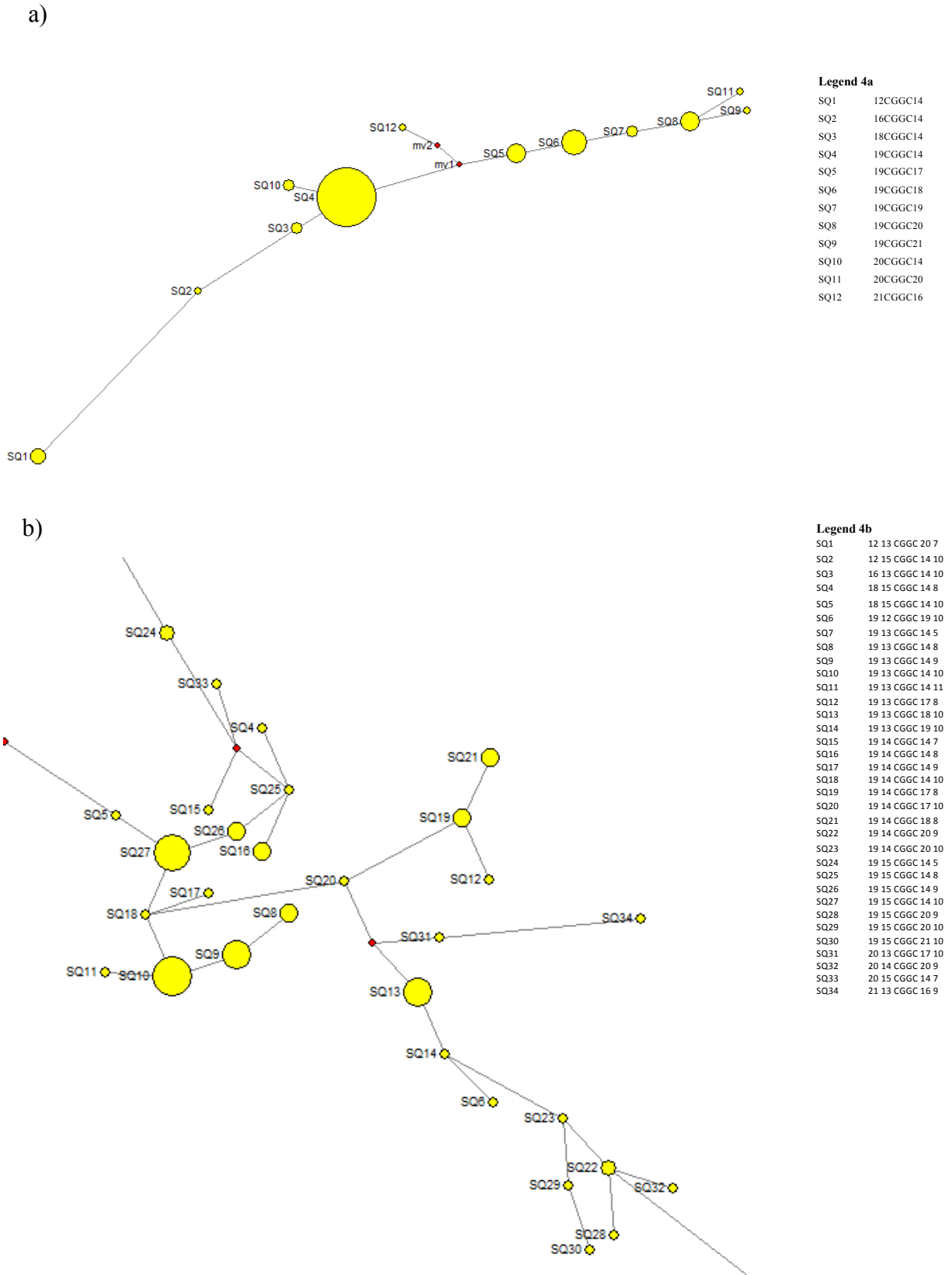
a)



b)



**Figure 3. Networks of CGGC haplotype-lineage in patients.** a) Phylogenetic relationship including two STRs (D22S1140 and D22S1153); b) Phylogenetic relationship including four STRs (D22S1140, D22S1153, D22S532, and GATA030P).



**Figure 4. Networks of CGGC haplotype-lineage in the Quechua population.** a) Phylogenetic relationship including two STRs (D22S1140 and D22S1153); b) Phylogenetic relationship including four STRs (D22S1140, D22S1153, D22S532, and GATA030P).

#### 4. DISCUSSÃO

A SCA10 é uma doença rara e sua prevalência é desconhecida. Em geral, a prevalência das SCAs, assim como outras doenças genéticas, é bastante variável e pode estar associada com a composição étnica das populações. Uma característica relevante dessa doença é que todos os pacientes acometidos com SCA10 são provenientes de países do continente americano. Além disso, frequentemente, a maioria dos pacientes com SCA10 relatam ter uma ancestralidade indígena. Essas evidências, juntamente com a inexistência da doença em outras populações europeias e asiáticas, corrobora a hipótese de um efeito fundador para a mutação associada à doença (Teive et al., 2013). Não há na literatura estudos em populações africanas para as quais a SCA10 tenha sido testada.

No presente estudo novas famílias foram identificadas como portadoras do alelo expandido, sendo essas provenientes de diferentes regiões do Brasil e do Peru. Buscando investigar a origem ancestral da mutação, primeiramente, uma análise para a identificação de haplótipos foi realizada utilizando marcadores intragênicos (SNPs) e extragênicos (STRs) já publicados na literatura. Esse haplótipo foi definido por Almeida *et al.* (2009) em todas as três famílias do estado do Rio Grande do Sul e em uma das três famílias Mexicanas estudadas e é denominado na literatura como o “haplótipo SCA10” (Bushara et al., 2013).

Um estudo anterior do nosso grupo, verificou que esse mesmo haplótipo está presente em 19 dos 20 pacientes com alelos expandidos no gene *ATXN10* (Ghenó, 2012). A presença consistente de um haplótipo único, 19CGGC14, sugere uma ancestralidade comum a todos os pacientes com SCA10. No presente estudo esse mesmo haplótipo está presente em 46,8% dos pacientes brasileiros e 62,8% dos pacientes peruanos. Além disso, ele também está numa frequência semelhante de 57,3% na população indígena controle, não havendo diferença estatisticamente significativa entre essa e pacientes de ambos países. No entanto, apesar dessa ser mais uma evidência relevante, é difícil saber, com base apenas nesses marcadores, se realmente a mutação tem origem nativa americana uma vez que esse haplótipo também é o mais frequente em indivíduos normais de outras populações amostradas por Almeida *et al.* (2009), porém há diferença estatisticamente significativa entre as populações amostradas e os controles brasileiros. Não há marcadores próximos à

expansão que sejam exclusivos da população indígena nativa americana, dificultando assim a elucidação dessa questão.

Visando ter uma aproximação da origem do bloco compartilhado, neste estudo expandimos o haplótipo incluindo mais dois marcadores informativos de ancestralidade, D22S532 e GATA030P, abrangendo uma região de 5.4 cM. As combinações alélicas desses marcadores apresentam dois haplótipos, os quais conferem uma prevalência de quase 40% em populações nativas americanas, sendo esses mais frequentes na América em comparação a outras populações ao redor do mundo (Wang et al., 2007). O haplótipo estendido foi investigado em 16 famílias brasileiras, 21 famílias peruanas e, como controles, 49 indivíduos saudáveis da mesma população indígena Quechua do Peru.

A partir das análises dos haplótipos, a existência do haplótipo comum CGGC foi confirmada em todos os pacientes com SCA10. Interessantemente, a presença do haplótipo 19CGGC14, em 7 das 14 famílias brasileiras (50,0%) e 11 das 17 (64,7%) famílias peruanas; está em *cis* com o haplótipo que vem a ser o segundo mais frequente (15-10) na população indígena controle (13,6%), configurando o haplótipo estendido 19-15-CGGC-14-10. Não necessariamente o evento de expansão pode ter ocorrido no alelo mais frequente. Esses achados estão de acordo com a nossa hipótese de origem ancestral comum da mutação e apontam para uma aproximação à origem nativa americana do bloco que contém a mutação patogênica.

A ideia de que os nativos americanos vieram da Ásia não é nova. Mas existe controvérsia sobre a data da chegada desses migrantes e o número de ondas migratórias. Os estudos com marcadores genéticos, uniparentais mtDNA e cromossomo Y, mostram que a Beríngia desempenhou um papel central no modelo de povoamento das Américas. Sendo assim, a mesma não serviu apenas de corredor de passagem, mas foi também uma região onde várias mutações específicas foram acumuladas, resultando em haplogrupos encontrados somente nos nativos americanos e que os distinguem das populações asiáticas que lhes deram origem (Bonatto e Salzano, 1997). Por conseguinte, as populações ancestrais dos paleoíndios ali permaneceram por um longo período, provavelmente ~5.000 anos, tempo suficiente para que essas mutações autóctones pudessem ocorrer. É provável que a mutação associada à SCA10 tenha ocorrido em populações ancestrais dos nativos americanos e tenha se espalhado para o restante do continente de acordo com o modelo de povoamento da América pelo *Homo sapiens*. Esse efeito fundador da mutação em

populações nativas americanas estaria, portanto, contribuindo para o aparecimento da doença no continente. Entretanto, estudos adicionais são necessários para comprovar essa hipótese.

É importante salientar que os marcadores que compõem o haplótipo estendido não são exclusivos de populações nativas americanas. Por isso, para complementar o resultado da análise de haplótipos que nos indicaram para a ancestralidade nativa americana da região cromossômica em questão, nós propomos o emprego futuro de uma outra metodologia. Essa seria a genotipagem de SNPs, abrangendo todo genoma, a qual nos daria uma estimativa aproximada da ancestralidade genética de pacientes portadores do haplótipo compartilhado. Assim, os resultados da análise haplotípica somados à informação da ancestralidade individual obtida utilizando uma ampla gama de marcadores informativos tornaria possível a inferência da origem do bloco com grande probabilidade de acerto.

Por fim, este trabalho consiste no mais extenso estudo sobre a origem ancestral da mutação de SCA10 até o momento. Nós identificamos um haplótipo estendido em famílias de duas origens etnogeográficas distintas que compartilham um haplótipo também frequente em população indígena Quechua. Esse achado é mais uma evidência que corrobora a hipótese de que a mutação associada à SCA10 possa ter ocorrido num cromossomo de origem nativa americana.

## 5. CONCLUSÕES

Considerando apenas os SNPs intragênicos, o haplótipo CGGC se mostrou consistente e compartilhado entre todos os pacientes incluídos no estudo. Considerando o haplótipo denominado “haplótipo SCA10”, 19CGGC14, esse se mostrou o mais frequente nas três populações amostradas: pacientes brasileiros (46,8%), pacientes peruanos (62,8%) e Quechuas (57,3%), não havendo diferença estatisticamente significativa entre os grupos.

Após a determinação do “haplótipo SCA10” comumente encontrado na amostra de pacientes, dois marcadores adicionais foram estudados. O haplótipo 19-15-CGGC-14-10 associado à mutação apresentou uma frequência de 50,0% e 64,7% em pacientes brasileiros e peruanos, respectivamente. A comparação das frequências anteriormente citadas com a população controle indígena Quechua apontou para a ancestralidade nativa americana da mutação uma vez que o haplótipo estendido associado à SCA10, 19-15-CGGC-14-10, é o segundo mais prevalente entre alelos controles (13,2%) seguido do haplótipo 19-13-CGGC-14-10 (14,2%).

## 6. REFERÊNCIAS:

- Almeida T, Alonso I, Martins S, Ramos EM, Azevedo L, Ohno K, Amorin A, Saraiva-Pereira ML, Jardim LB, Matsuura T, Sequeiros J, Silveira I (2009) Ancestral origin of the ATTCT repeat expansion in spinocerebellar ataxia type 10 (SCA10). *PLoS One* 4:e4553.
- Alonso I, Jardim LB, Artigas O, Saraiva-Pereira ML, Matsuura, Ashizawa T, Sequeiros J, Silveira I (2006) Reduced penetrance of intermediate size alleles in spinocerebellar ataxia type 10. *Neurology* 66:1602-4.
- Baizabal-Carvalho JF, Jankovic J (2012) Spinocerebellar ataxia type 2 and 10: more than a coincidental association? *Arch Neurol* 69(11):1524-5.
- Baizabal-Carvalho JF, Xia G, Botros P, Laguna J, Ashizawa T, Jankovic J (2015) Bolivian Kindred with combined spinocerebellar ataxia types 2 and 10. *Acta Neurol Scand* DOI: 10.1111/ane.12371.
- Ashizawa T (2012) Spinocerebellar ataxia type 10. *Handb Clin Neurol* 103:507-19
- Bird, TD Hereditary Ataxia Overview 1998 Oct 28 [Updated 2012 May 31]. In: Pagon RA, Bird TD, Dolan CR, et al., editors. *GeneReviews*<sup>TM</sup> [Internet]. Seattle (WA): University of Washington, Seattle; 1993-. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK1138/>.
- Bonato SL, Salzano (1997) A single and early migration for the peopling of the Americas supported by mitochondrial DNA sequence data. *Proc Natl Acad Sci USA* 94:1866-1871.
- Braida C, Stefanatos RK, Adam B, Mahajan N, Smeets HJ, Niel F, Goizet C, Arveiler B, Koenig M, Lagier-Tourenne C, Mandel JL, Faber CG, de Die-Smulders CE, Spaans F and Monckton DG (2010) Variant CCG and GGC repeats within the CTG expansion dramatically modify mutational dynamics and likely contribute toward unusual symptoms in some myotonic dystrophy type 1 patients. *Hum Mol Genet* 19:1399-412.
- Brusco A, Gellera C, Cognoli C, Saluto A, Castucci A, Michielotto C, Fetoni V, Mariotti C, Migoni N, Di Donato S, Taroni F (2004) Molecular genetics of hereditary spinocerebellar ataxia: mutation analysis of spinocerebellar ataxia genes and CAG/CTG repeat expansion detection in 225 Italian families. *Arch Neurol* 61:727-33.
- Bushara K, Bower M, Liu J, McFarland KN, Landrian I, Hutter D, Teive HA, Rasmussen A, Mulligan CJ, Ashizawa T (2013) Expansion of the spinocerebellar ataxia type 10 (SCA10) repeat in a patient with Sioux Native American Ancestry. *PLoS One* 8(11):e81342.
- Cagnoli C, Michielotto C, Matsuura T et al (2004) Detection of large pathogenic expansions in FRDA1, SCA10, and SCA12 genes using a simple fluorescent repeat primed PCR assay. *J Mol Diagn* 6:96-100.
- Castilhos RM, Furatado GV, Gheno TC, Schaeffer P, Russo A, Barsottini O, Pedroso JL, Salarini DZ, Vargas FR, Lima MAFD, Godeiro C, Silva LCS, Toralles MBP, Santos S, Jr HL, Wanderley HY, Frassinetti P, Medeiros PFV, Pereira ET, Ribeiro E, Saraiva-Pereira ML, Jardim LB (2013) Spinocerebellar ataxias in Brazil – frequencies and modulating effects of related genes. *The Cerebellum* [online].
- Fujigasaki H, Tardieu S, Camuzat A, Stevanin G, LeGuern E, Matsuura T, Ashizawa T, Durr A, Brice A (2002) Spinocerebellar ataxia type 10 in the French population. *Ann Neurol* 51:408-9.

- Gallardo M, Soto A (2009) Clinical characterization of a Venezuelan family with spinocerebellar ataxia type 10. *Mov Disord* 24:S12.
- Gatto EM, Gao R, White MC, Uribe Roca MC, Etcheverry JL, Persi G, Poderoso JJ, Ashizawa T (2007) Ethnic origin and extrapyramidal signs in an Argentinian spinocerebellar ataxia type 10 family. *Neurology* 69:216-8.
- Gheno, TC. *Ataxia Espinocerebelar tipo 10: identificação de casos e determinação de haplótipos no gene ATXN10*. 2012. 45f. Dissertação (Mestrado em Genética e Biologia Molecular) – Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2012.
- Grewal RP, Tayag E, Figueroa KP, Zu L, Durazo A, Nunez C, Plust SM (1998) Clinical and genetic analysis of a distinct autosomal dominant spinocerebellar ataxia. *Neurology* 51:1423-6.
- Grewal RP, Achari M, Matsuura T, Durazo A, Tayag E, Zu L, Pulst SM, Ashizawa T (2002) Clinical features and ATTCT repeat expansion in spinocerebellar ataxia type 10. *Ach Neurol* 59:1285-90.
- Harding AE (1984): *The hereditary ataxias and related disorders*. Churchill Livingstone, New York.
- Jiang H, Tang BS, Xu B, Zhao GH, Shen L, Tang JG, Li QH, Xia K (2005) Frequency analysis of autosomal dominant spinocerebellar ataxias in mainland Chinese patients and clinical and molecular characterization of spinocerebellar ataxia type 6. *Chin Med J (Engl)* 118:837-43.
- Kapur SS, Goldman JG (2012) Two in one: report of a patient with spinocerebellar ataxia type 2 and 10. *Arch Neurol* 69(9):1200-3.
- Keren B, Jacqueline A, Depienne C, Leite P, Durr A, Carpentier W, Benyahia B, Ponsot G, Soubrier F, Brice A, Héron D (2010) Evidence against haploinsufficiency of human ataxin 10 as a cause of spinocerebellar ataxia type 10. *Neurogenetics* 11:273-4
- Leonardi L, Marcotulli C, McFarland NK, Tessa A, DiFabio R, Santorelli MF, Pierelli F, Ashizawa T, Casali C (2014) Spinocerebellar ataxia type 10 in Peru: the missing link in the Amerindian origin of the disease. *J Neurol* 57:8-11.
- Lin X, Ashizawa T (2005) Recent progress in spinocerebellar ataxia. *The Cerebellum* 4:37-42.
- Marz P, Probst A, Lang S, Schwager M, Rose-John S, Otten U and Ozbek S (2004) Ataxin-10, the spinocerebellar ataxia type 10 neurodegenerative disorder protein is essential for survival of cerebellar neurons. *J Bio Chem* 279:33542-50.
- Matsuura T, Achari M, Khajavi M, Bachinski LL, Zoghbi HY, Ashizawa T (1999) Mapping of the gene for a novel spinocerebellar ataxia with pure cerebellar signs and epilepsy. *Annual Neurol* 45:407-11.
- Matsuura T, Ashizawa T (2002) Polymerase chain reaction amplification of expanded ATTCT repeat in spinocerebellar ataxia type 10. *Ann Neurol* 51:271-2.
- Matsuura T, Fang P, Lin X, Khajavi M, Tsuji K, Rasmussen A, Grewal RP, Achari M, Alonso ME, Pulst SM, Zoghbi HY, Nelson DL, Roa BB, Ashizawa T (2004) Somatic and germline instability of the ATTCT repeat in spinocerebellar ataxia type 10. *Am J Hum Genet* 74:1216-24.
- Matsuura T, Fang P, Pearson CE, Jayakar P, Ashizawa T, Roa BB, Nelson DL (2006) Interruptions in the expanded ATTCT repeat of spinocerebellar ataxia type 10: repeat purity as a disease modifier? *Am J Hum Genet* 78:125-9.



- Matsuura T, Ranum LPW, Volpini V, Pandolfo M, Sasaki H, Tashiro K, Watase K, Zoghbi HY, Ashizawa T (2002) Spinocerebellar ataxia type 10 is rare in populations other than Mexicans. *Neurology* 58:983-4.
- Matsuura T, Yamagata T, Burgess DL, Rasmussen A, Grewal RP, Watase K, Khajavi M, McCall AE, Davis CF, Zu L, Achari M, Pulst SM, Alonso E, Noebels JL, Nelson DL, Zoghbi HY, Ashizawa T (2000) Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. *Nat Genet* 26:191-4.
- Matsuura T, Ashizawa T. Spinocerebellar Ataxia Type 10. 2002 Apr 23 In: Pagon RA, Adam MP, Bird TD, et al., editors. *GeneReviews™* [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2013. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK1175/>.
- Raskin S, Ashizawa T, Teive HA, Arruda WO, Fang P, Gao R, White MC, Werneck LC, Roa B (2007) Reduced penetrance in a Brazilian family with spinocerebellar ataxia type 10 *Arch Neurol* 64:591-4.
- Rasmussen A, Matsuura T, Ruano L, Yescas P, Ochoa A, Ashizawa T, Alonso E (2001) Clinical and genetic analysis of four Mexican families with spinocerebellar ataxia 10. *Annu Neurol* 50: 234-9.
- Rasmussen A, Yescas P, Matsuura T *et al.* (2000) Molecular diagnosis of spinocerebellar ataxias in Mexican population. *Am J Hum Genet* 67 Suppl 2:A1902
- Roxburgh HR, Smith OC, Lim JG, Bachman DF, Byrd E, Bird TD (2013) The unique co-occurrence of spinocerebellar ataxia type 10 (SCA10) and Huntington disease. *J Neurol Sci* 324:176-178.
- Sasaki H, Yabe I, Tashiro K (2003) The hereditary spinocerebellar ataxias in Japan. *Cytogenet Genome Res* 100:198-205.
- Seixas AI, Maurer MH, Lin M, Callahan C, Ahuja A, Matsuura T, Ross CA, Hisama Fm, Silveira I, Margolis RL (2005) FXTAS, SCA10 and SCA17 in American patients with movement disorders. *Am J Med Genet* 136:87-9.
- Stephans M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68(4):978-989.
- Sulek-Piatkowska A, Zdzienicka E, Raczynska-Rakowicz M, Krysa W, Rajkiewicz M, et al. (2010) The occurrence of spinocerebellar ataxias caused by dynamic mutations in Polish patients. *Neurol Neurochir Pol* 44: 238–245.
- Teive HA, Roa BB, Raskin S, Fang P, Arruda WO, Neto YC, Gao R, Werneck LC, Ashizawa T (2004) Clinical phenotype of Brazilian families with spinocerebellar ataxia 10. *Neurology* 63:1509-12.
- Teive HA, Arruda WO, Raskin S, Munhoz RP, Zavala JA, Werneck LC, Ashizawa T (2011) Spinocerebellar ataxia type 10 – A review. *Parkinsonism Relat Disord* 17:655-61.
- Teive HA, Munhoz RP, Raskin S, Fang P, Arruda WO, de Paola L, Werneck LC, Ashizawa T (2010) Spinocerebellar ataxia type 10: Frequency of epilepsy in a large sample of Brazilian patients. *Mov Disord* 25:2875-8.
- Teive HA, Ashzawa T (2013) Spinocerebellar ataxia type 10: from Amerindians to Latin Americans. *Curr Neurol Neurosci Rep* 13:393.
- Trikamji B, Singh P, Mishra S (2015) Spinocerebellar ataxia-10 with paranoid schizophrenia. *Ann Indian Acad Neurol* 18(1):93-95
- Wakamiya M, Matsuura T, Liu Y, Schuster GC, Gao R, Xu W, Sarkar PS, Lin X, Ashizawa T (2006) The role of ataxin 10 in spinocerebellar ataxia type 10 pathogenesis. *Neurology* 67:607-13.

- Wang JL, Wu YQ, Lei LF, Shen L, Jiang H, Zhou YF, Yi JP, Zhou J, Yan XX, Pan Q, Xia K, Tang BS (2010) Polynucleotide repeat expansion of nine spinocerebellar ataxia subtypes and dentatorubral-pallidoluysian atrophy in healthy Chinese Han population. *Zhonghua Yi Xue Chuan Xue Za Zhi* 27:501-5.
- Waragai M, Nagamitsu S, Xu W, Li YJ, Lin X, Ashizawa T (2006) Ataxin 10 induces neuritogenesis via interaction with G-protein beta2 unit. *J Neurosci Res* 83:1170-8.
- Wexler E, Fogel BL (2011) New-onset psychosis in a patient with spinocerebellar ataxia type 10. *Am J Psychiatry* 168:1339-40.
- White MC, Gao R, Xu W, Mandal SM, Lim JG, Hazra TK, Wakamiya M, Edwards SF, Raskin S, Teive HA, Zaghbi HY, Sarkar PS, Ashzawa T (2010) Inactivation of hnRPP K by expanded intronic AUUCU repeat induces apoptosis via translocation of PKCdelta to mitochondria in spinocerebellar ataxia 10. *PLoS Genet* 6:e1000984.
- White MC, Xia G, Gao R, Wakamiya M, Sarkar SP, McFarland K, Ashizawa T (2012) Transgenic mice with SCA10 pentanucleotide repeats show motor phenotype and susceptibility to seizure: a toxic gain-of-function model. *J Neuro Sci* 90:706-14.
- Vale J, Bugalho P, Silveira I, Sequeiros J, Guimarães J, Coutinho P (2010) Autosomal dominant cerebellar ataxia: frequency analysis and clinical characterization of 45 families from Portugal. *Eur J Neurol* 17:124-8.
- Zu L, Figueroa KP, Grewal R, Pulst SM (1999) Mapping of a new autosomal dominant spinocerebellar ataxia to chromosome 22. *Am J Hum Genet* 64:594-9.