


Haplotype Study in SCA10 Families Provides Further Evidence for a Common Ancestral Origin of the Mutation

Giovana B. Bampi^{1,2,3} · Rafael Bisso-Machado⁴ · Tábita Hünemeier⁵ · Tailise C. Gheno^{1,2,3} · Gabriel V. Furtado^{1,2,3} · Diego Veliz-Otani⁶ · Mario Cornejo-Olivas⁶ · Pillar Mazzeti⁶ · Maria Cátira Bortolini¹ · Laura B. Jardim^{1,2,3,7} · Maria Luiza Saraiva-Pereira^{1,2,3,8}  · on behalf of Rede Neurogenetica

Received: 26 September 2016 / Accepted: 29 August 2017 / Published online: 13 September 2017
© Springer Science+Business Media, LLC 2017

Abstract 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. SCA10 has shown a geographical distribution throughout America with a likely degree of Amerindian ancestry from different countries so far. Currently available data suggest that SCA10 mutation might have spread out early during the peopling of the Americas. However, the ancestral origin of SCA10 mutation remains under speculation.

Electronic supplementary material The online version of this article (doi:10.1007/s12017-017-8464-8) contains supplementary material, which is available to authorized users.

✉ Maria Luiza Saraiva-Pereira
mlpereira@hcpa.edu.br

- ¹ Department of Genetics, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil
- ² Medical Genetics Service, Hospital de Clínicas de Porto Alegre, Ramiro Barcelos 2350, Porto Alegre, Rio Grande do Sul 90035-903, Brazil
- ³ Laboratory of Genetics Identification – Research Center, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil
- ⁴ Centro Universitario de Tacuarembó, Universidad de la República, Tacuarembó, Uruguay
- ⁵ Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo, Brazil
- ⁶ Neurogenetics Research Center, Instituto Nacional de Ciencias Neurológicas, Lima, Peru
- ⁷ Department of Internal Medicine, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil
- ⁸ Department of Biochemistry, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

Samples of SCA10 patients from two Latin American countries were analysed, 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 harbouring the ATTCT expansion were used to define the haplotypes, which were genotyped by different approaches. Our data have shown that 19-CGGC-14 shared haplotype was found in 47% of Brazilian and in 63% of Peruvian families. Frequencies from both groups are not statistically different from Quechua controls (57%), but they are statistically different from Brazilian controls (12%) ($p < 0.001$). The most frequent expanded haplotype in Quechuas, 19-15-CGGC-14-10, is found in 50% of Brazilian and in 65% of Peruvian patients with SCA10. 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 cerebellar ataxia with variably associated extracerebellar findings that most characteristically include epilepsy, but also psychiatric disorder, cognitive disability, polyneuropathy, and pyramidal signs (Teive et al. 2004, 2010, 2011; Ashizawa 2012). Apparently, clinical spectrum may vary according to patients' ethnical background (Teive et al. 2004, 2010; Rasmussen et al. 2001). 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 (Teive et al. 2011; Ashizawa 2012). Normal alleles can vary from 9 to 32 repeats, while alleles from affected subjects range from 800 up to 4500 repeats (Matsuura et al. 1999, 2000; Grewal et al. 1998, 2002). Reduced penetrance has been found for intermediate size alleles of 280–850 repeats (Alonso et al. 2006; Raskin et al. 2007).

SCA10 has shown geographical distribution throughout the Americas. Until very recently, it has only been described in patients from Mexico, Brazil, Argentina, Venezuela, Colombia, USA, and Peru (Matsuura et al. 2002; Rasmussen et al. 2001; Gatto et al. 2007; Gallardo and Soto 2009; Roxburgh et al. 2013; Kapur and Goldman 2012; Bushara et al. 2013; Leonardi et al. 2014; Teive and Ashizawa 2013). Interestingly, except cases involving a Japanese and a Chinese Han families, patients have a recognized family history of Amerindian ancestry, and literature data report two affected subjects descending from the Yaqui, in Mexico, and from the Sioux, also from North America (Kapur and Goldman 2012; Bushara et al. 2013; Wang et al. 2015; Naito et al. 2017). Previous haplotype studies in a limited number of Brazilian and Mexican families have shown a shared common haplotype (Almeida et al. 2009).

The recent description of two families from Asia carrying SCA10 suggests that the original SCA10 mutation might have occurred very early, before the divergence of Proto-Amerindians from ancestral Asians (Wang et al. 2015; Naito et al. 2017). SCA10 is nowadays present in admixed American populations due to admixture with Europeans and Africans who arrived in the Americas in the sixteenth and seventeenth centuries (Avena et al. 2012; Price et al. 2007; Lins et al. 2010; Seldin et al. 2007). However, the ancestral origin of SCA10 mutations remains unclear, since this possible unique origin has to be confirmed with more extensive analyses. SCA10 is a rare disease belonging to a group of genetically heterogeneous neurodegenerative disorders. More than 40 loci have been associated with spinocerebellar ataxia with overlapping clinical symptoms so far. It is well known that founder effects play an important role in the prevalence of genetic disorders that are restricted to certain geographical areas. Therefore, elucidating aspects related to the natural history of the disease will definitely assist clinicians and improve awareness of this disease, providing a more comprehensive genetic counselling that will reflect in a better life quality for patients. Herein, we present a wider study of patients from two different geographical regions bringing molecular evidence to test the hypothesis of the presence of an 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 five 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 healthy Amerindian “Quechua” individuals from Southern Peru and 51 unrelated Brazilians individuals were included as controls (Almeida et al. 2009). 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 leucocytes using salting-out method (Miller et al. 1998). 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 (Matsuura et al. 2002).

Genotyping of Polymorphisms Flanking the ATXN10 Gene

All subjects were genotyped for four highly polymorphic short tandem repeats (STRs) and four intragenic single nucleotide polymorphisms (SNPs). Initially, haplotypes were constructed with four SNPs (rs764850—C/A, rs72556348—G/A, rs72556349—G/A, rs72556350—C/T) and two dinucleotide STRs (D22S1140 and D22S1153), as published previously (Almeida et al. 2009; Gheno et al. 2017). In order to gain deeper insight about the ancestral origin of the haplotype, two additional tetranucleotide loci (STRs: D22S532 and GATA030P) were added to the original set of markers, spanning a region of 5.2 cM harbouring the ATTCT repeats. These markers were selected based on two criteria: (1) physical distance from SCA10

locus and (2) the presence of alleles with higher frequency in Native Americans than in other populations. The additional STRs data set used for selection was published elsewhere (Wang et al. 2007). The genomic position of markers and mutation location are indicated in Fig. 1.

Amplification of the polymorphic regions is 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[®] PCR Assay (Applied Biosystems). PCR was performed in a final volume of 8 μ L containing 4 ng of DNA, 0.2 μ L of specific TaqMan assay, and 4 μ L of 2 \times PCR Genotyping Master Mix (Applied Biosystems) (Assay ID C_30474239_39). Amplification included an initial step at 50 $^{\circ}$ C for 2 min (activation of AmpErase UNG function), 95 $^{\circ}$ C for 10 min (AmpliTaq[®] Gold activation), followed by 40 cycles of denaturation at 95 $^{\circ}$ C for 15 s, and annealing/extension at 60 $^{\circ}$ C for 1 min. Allelic discrimination step was performed at 60 $^{\circ}$ C for 1 min. Amplicons were analysed 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 μ L reaction was prepared containing 100 ng of genomic DNA, 5.0 pmol of each primers ATXN10SNPR and ATXN10SNPF1, 200 μ M dNTPs (Amersham-Pharmacia), 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 0.75 mM MgCl₂, and 1.5 U of Taq DNA Polymerase (Invitrogen). Amplification was performed in a thermal cycler as follows: initial denaturation at 94 $^{\circ}$ C for 10 min, followed by 30 cycles at 94 $^{\circ}$ C for 30 s, 63 $^{\circ}$ C for 1 min, and 72 $^{\circ}$ C for 30 s, and final extension at 72 $^{\circ}$ C for 10 min. Amplification was verified by resolving an aliquot

using a 1.5% (w/v) electrophoresis gel. Amplicons were submitted to 2.5 U Exonuclease I and 0.25 U Shrimp Alkaline Phosphatase purification before sequencing reaction, which was performed using the BigDye[®] 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 was performed simultaneously by multiplex PCR carried out in a final volume of 25 μ L, containing 20 ng of genomic DNA, 2.5 pmol of each primer, 200 μ M dNTPs (Amersham-Pharmacia), 15 mM Tris-HCl (pH 8.3), 75 mM KCl (PCR Buffer II—Applied Biosystems), 0.75 mM MgCl₂, and 1 U of AmpliTaq[®] Gold DNA Polymerase (Applied Biosystems). One primer of each pair was labelled 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 $^{\circ}$ C for 10 min and 56 $^{\circ}$ C for 30 s for annealing, and final extension at 72 $^{\circ}$ C for 5 min. After amplification, an aliquot of PCR products was mixed with formamide (HiDi[™] Formamide, Applied Biosystems) and GeneScan[™] 500-LIZ (Applied Biosystems), and capillary electrophoresis was performed in an ABI 3130xl Genetic Analyzer (Applied Biosystems) using a 36 cm \times 50 μ m capillary containing Performance Optimized Polymer-7 (POP-7, Applied Biosystems). Samples were injected during 22 s at 3.0 kV and electrophoresed at 15 kV for 40 min at 60 $^{\circ}$ C. Amplicon lengths were estimated using GeneScan[™] 500-LIZ molecular weight through GeneMapper[®] 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 (Stephens et al. 2001). Statistical analysis was performed using WINPEPI program. Chi-square test was used to compare allele frequencies between patients and controls. 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.

Results

A total of 29 Brazilian (16 families) and 27 Peruvian (21 families) SCA10 patients as well as 49 Amerindian Quechua controls were genotyped using eight *ATXN10* gene-flanking markers (Fig. 1). Alleles were arbitrarily named

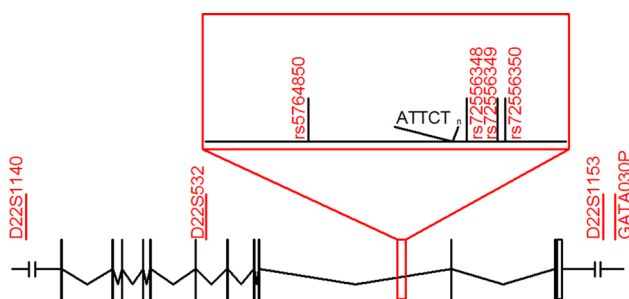


Fig. 1 Diagram of chromosome 22 region including *ATXN10* gene harbouring the SCA10 mutation. The ATTCT repeat and flanking regions include exons (vertical lines) and intergenic/introns (horizontal line), showing location of four intragenic SNPs and four microsatellite markers used for full haplotype reconstruction. Distance from ATTCT expansion is as follows: -1.5 Mb for D22S1140; -68 Kb for D22S532; -1198 bp for rs5764850; $+48$ bp for rs72556348; $+303$ bp for rs72556349; $+370$ bp for rs72556350; $+21.3$ Kb for D22S1153; and $+78.5$ Kb for GATA030P

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%), as expected.

Allelic frequency distribution of genetic markers among all samples included in this study is shown in Table S2. D22S532 is the only marker in which allelic frequencies showed statically difference between each patient groups and Quechua controls. Initially, haplotypes were constructed with informative SNPs and STRs D22S1140 and D22S1153, spanning a region of ~1.7 Mb, flanking the pathogenic expansion 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 controls showed 14 different haplotypes. The most frequent SCA10-linked shared haplotype, 19-CGGC-14, was present in 47% of Brazilian and in 63% of Peruvian families (Table 2). This haplotype is also present in Amerindian Quechua controls with a very similar frequency (57%) as shown in Table S3. No statistically significant differences were found between patients and Amerindian Quechua controls.

Considering all markers as an extended haplotype, the most frequent one in both patient groups was 19-15-CGGC-14-10.

Frequencies of more common haplotype (H2 and H3), considering markers D22S532 and GATA30P only, observed in the present study are shown comparatively to other global populations (Fig. 2). Black bars represent the sum of frequencies of both haplotypes (H2 + H3), and they were shown to be frequent in Amerindians (38.2%) (America data in Fig. 2), as defined previously by Wang et al. (2007). These haplotypes (H2 and H3) correspond to haplotype #10 and #28, respectively, in our study (Table S4). White bars represent the additional 46 different haplotypes. Interestingly, both highly prevalent haplotypes

in Amerindians (H2 and H3) were also found at high frequency in mutant alleles of both groups of SCA10 patients: Brazilians SCA10 (50.0%) and Peruvians (77.0%) (Fig. 2).

A median-joining network considering haplotypes constructed based on CGGC plus 2 dinucleotide STRs (D22S1140 and D22S1153) is shown in Fig. 3. This network included SCA10 patients, Quechua controls and Brazilian controls. It is relevant to mention that this network was build up with data from this study and from data published before, but always including the same 6 markers defined above. This network has shown similar pattern among groups, displaying a unimodal-like pattern (Fig. 3).

Discussion

SCA10 is a rare disease with unknown prevalence that has been mostly described in Latin American populations (Brusco et al. 2004; Fujigasaki et al. 2002; Sulek-Piatkowska et al. 2010; Vale et al. 2010; Wang et al. 2011). To date SCA10 has been reported outside America in just three cases: in a patient born in the Northern coast of Peru, suggesting a significant Amerindian ancestry¹⁹, in a Japanese family (Wang et al. 2015), and, more recently, in a Chinese Han family (Naito et al. 2017).

In the present study, common haplotype 19-CGGC-14 was demonstrated to be present in a large cohort of SCA10 families from both Brazil and Peru. Frequencies of this 19-CGGC-14 SCA10-linked do not differ among groups (Table 1). This haplotype has been originally referred as the “SCA10 haplotype”, considering that it was carried by 2 families from the Southernmost state of Brazil and in one Mexican family as well as in 19 out of 20 patients from a previous study of our group (Gheno et al. 2017), signalling that the mutation has a probable common origin (Almeida et al. 2009).

Amerindian Quechua controls from Peru were chosen for the present study due to their high contribution to the genetic pool of the contemporary mixed Peruvian population. All Peruvian SCA10 cases in our cohort show geographical distribution overlapping with traditional Quechua speaking regions. Interestingly, our data have shown that

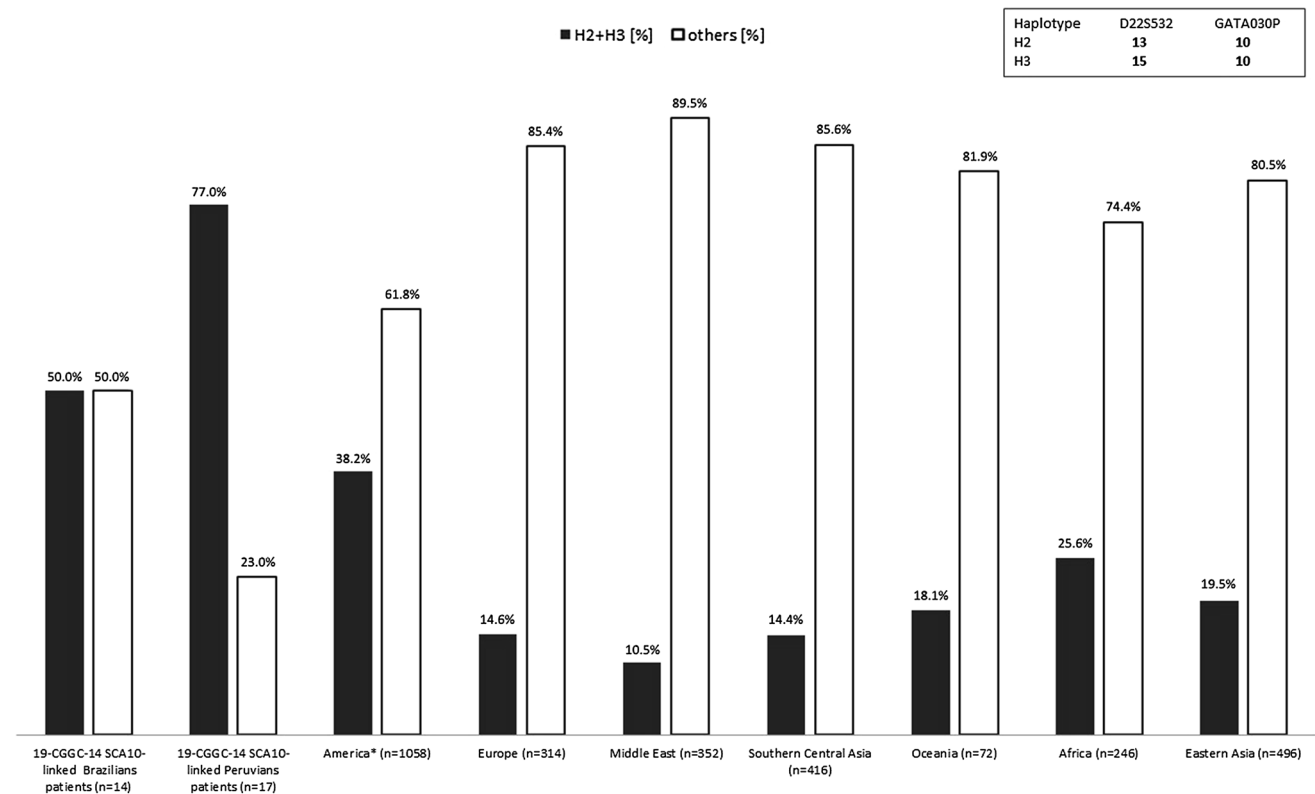
Table 1 Frequencies of haplotypes associated with SCA10

Haplotypes	Brazilian families SCA10-linked alleles (n = 16)	Peruvians families SCA10-linked alleles (n = 21)
19-CGGC-14	0.875 (14) ^a	0.808 (17) ^a
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)

^aFrequencies were compared using the Chi-square test ($X^2 = 3.003$; $p = 0.930$)

Table 2 Haplotype 19-CGGC-14 versus others haplotypes among groups

Haplotypes	Brazilian patients (16) <i>n</i> (%)	Peruvian patients (21) <i>n</i> (%)	Quechua controls (49) <i>n</i> (%)	Brazilian controls (51) ^a <i>n</i> (%)
19-CGGC-14	15 (46.8)	26 (62.8)	56 (57.3)	12 (11.8) ^b
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)

^aWang et al. (2015)^bThis group is statistically different from the others ($X^2 = 55.0$; $p < 0.001$; Sidak-adjusted pairwise comparison)**Fig. 2** Distribution of haplotype frequencies of ancestral informative markers among ethno-geographical groups and 19-CGGC-14 SCA10-linked among patients. Names of corresponding haplotype and their alleles are shown in the box in the upper right part of the figure. Black bars represent the combined alleles of D22S532 and GATA30P

markers, while white bars represent all other haplotypes. Data from Brazilian and Peruvian SCA10 patients were generated by this current study. Data from the additional seven different regions were from healthy individuals from Wang et al. (2007)

19-CGGC-14 shared haplotype is the most frequent among Amerindian Quechua controls (57%), which do not differ statistically from those observed for Brazilian (47%) and Peruvian patients (63%). However, these numbers are statistically different from Brazilian controls (12%) ($p < 0.001$) as shown in Table 2.

It is well known that America has a history of extensive admixture between three continental populations: Africans, Europeans, and Native Americans. Native American ancestry in Brazil is higher in Northwest (Amazonia), while African and European are more widespread, being

the most admixed country in the continent (Ruiz-Linares et al. 2014). The Brazilian population heterogeneity may explain statistical difference between Brazilian patients and controls. Peruvian population has a substantial Native American ancestry widespread, very low African ancestry, and European ancestry in Northern/Central areas (Ruiz-Linares et al. 2014). Therefore, under Amerindian ancestral origin of SCA10 mutation, statistical difference was not expected between Peruvians patients and Quechua controls, although samples were not tested for individual admixture proportion.

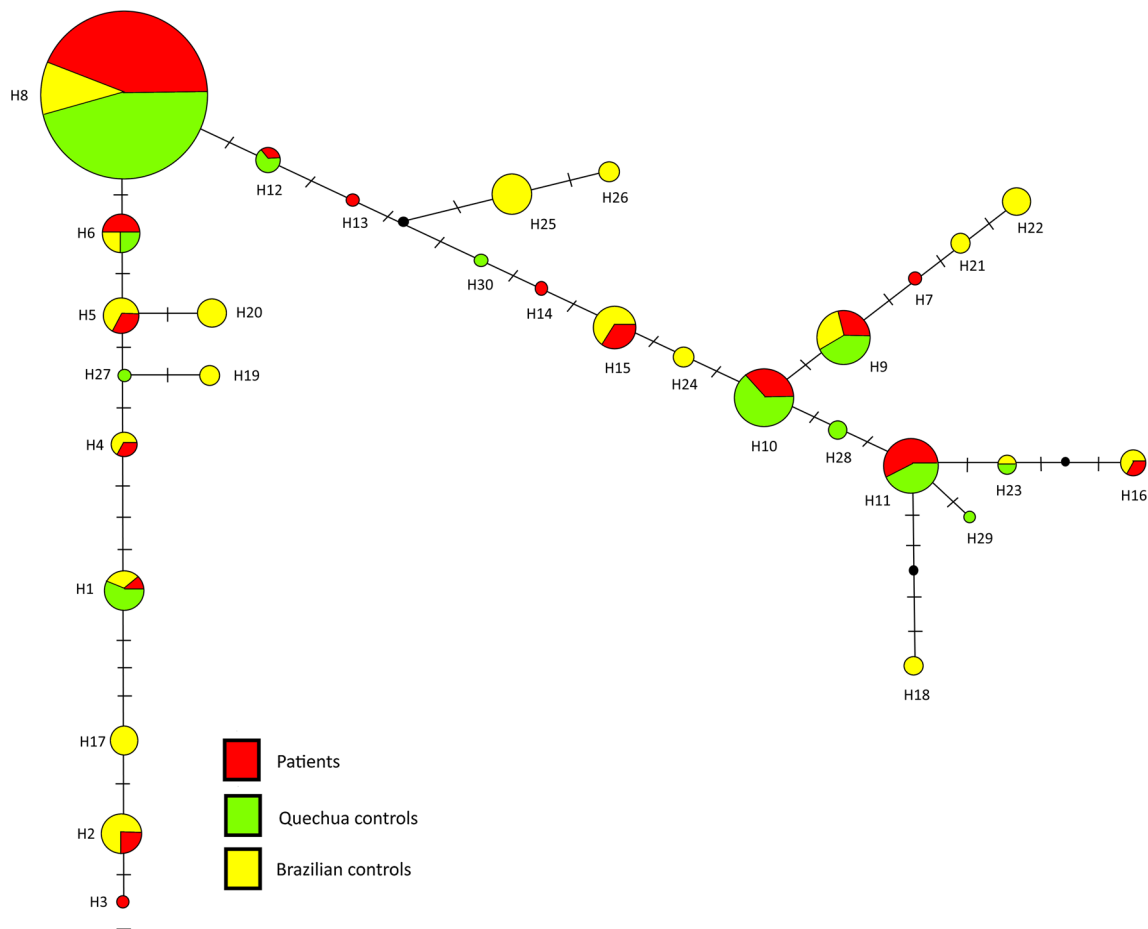


Fig. 3 Network of CGGC haplotype lineage in SCA10 patients, Quechua controls and Brazilian controls. Phylogenetic relationship including two STRs (D22S1140 and D22S1153). H1: 12-CGGC-14; H2: 12-CGGC-18; H3: 12-CGGC-19; H4: 15-CGGC-14; H5: 17-CGGC-14; H6: 18-CGGC-14; H7: 18-CGGC-17; H8: 19-CGGC-14; H9: 18-CGGC-17; H10: 19-CGGC-18; H11: 19-CGGC-20; H12: 20-CGGC-14; H13: 20-CGGC-15; H14: 21-CGGC-17; H15:

21-CGGC-18; H16: 21-CGGC-21; H17:12-CGGC-17; H18: 15-CGGC-20; H19: 16-CGGC-15; H20: 17-CGGC-15; H21: 17-CGGC-17; H22: 17-CGGC-18; H23: 19-CGGC-21; H24: 20-CGGC-18; H25: 21-CGGC-14; H26: 22-CGGC-14; H27: 16-CGGC-14; H28: 19-CGGC-19; H29: 20-CGGC-20; H21: 21-CGGC-16

Nevertheless, these data have to be interpreted carefully when declaring the exclusive Amerindian origin of SCA10 mutation. This multiloci combination is quite frequent in normal chromosomes of Brazilian populations as well as in Mexicans and Portuguese, as previously shown (Almeida et al. 2009).

In order to address this query, we then extended the haplotype by including two additional STRs, D22S532 and GATA030P. Considering the inclusion of these two loci, two haplotypes, H2 (13–10) and H3 (15–10), are carried out by almost 40% of the Amerindians (Wang et al. 2007), unlike what happens to populations from other continents, as shown in Fig. 1. Hence, we assumed that both could be considered a good indicator of Amerindian ancestry. The two most frequent extended haplotypes in controls are 19-13-CGGC-14-10 (14%) followed by 19-15-CGGC-14-10 (13%) (Table S4). Haplotype 19-15-CGGC-14-10 is

carried by 50% of Brazilians and 65% of Peruvians 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 19-CGGC-14 SCA10-linked in both patient groups are shown in Fig. 1. Interestingly, H3 is shown to be in *cis* with 19-CGGC-14 and the mutant allele, being present in 7 out of 14 Brazilians (50%) and in 11 out of 17 (65%) Peruvians. These findings are in agreement to our previous hypothesis of a common ancestry of the mutation and reinforce further that the studied chromosome block of markers flanking ATTCT mutation has a probable Amerindian origin. Furthermore, a founder effect might be contributing to the distribution of SCA10 throughout American countries (Almeida et al. 2009).

Table 3 Frequencies of the extended haplotypes associated with SCA10

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

^aFrequencies were compared by Fisher's exact test with Bonferroni correction for multiple comparisons; Brazilians versus Peruvians ($p = 0.037$); Brazilians versus Quechuas ($p < 0.001$); Peruvians versus Quechuas ($p < 0.001$)

^bRemaining alleles are distributed among 26 others 19-CGGC-14-carrier haplotypes (Table S4)

Data from different fields of knowledge, including genetics, support an Asian origin of nowadays Native Americans. Migrants from Asia settled Beringia, an empty land mass of about 1 million km² exposed due to the last glacial maximum in the end of Pleistocene (Bortolini et al. 2003). They remained there for at least 5000 years; enough time in which numerous autochthonous mutations found only in Native Americans may have arisen (Bortolini et al. 2003, 2014; Fagundes et al. 2008; González-José et al. 2008). After the disappearance of Beringia and the emergence of Bering Strait, this particular population entered and colonized the America (González-José et al. 2008; Bortolini et al. 2014). Then, it is possible to speculate that the origin of SCA10 mutation may have occurred in this Beringian population, considering the wide distribution of the disease in the American continent.

Further, there is a higher level of genetic diversity and a lower level of population structure in Western South America (Andean area) when compared to Eastern South America (Amazonia, Chaco, and Brazilian Central Plateau; Fagundes et al. 2008; Tarazona-Santos et al. 2001). This may explain a relatively higher diversity of Peruvian haplotypes associated with the SCA10 mutation when compared to Brazilian families, since the latter have shown two different haplotypes 19-CGGC-14 and 19-CGGC-18 associated with SCA10 mutation. The 19-CGGC-18 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 improve the understanding of the phylogenetic

relationship among haplotypes associated with SCA10. Similar pattern was observed among groups, demonstrating a unimodal-like pattern (Fig. 3). This type of pattern is expected to be observed in networks of patients with same mutation origin, while different mutation origins generate a bimodal-like pattern.

We should take into consideration that haplotypes H2 and H3 are also found in other populations throughout the world despite their high prevalence in Native Americans. Unfortunately, there are no exclusive genetic markers of Native American along the chromosome region assessed in this study that could give undoubtedly information about the SCA10 geographical 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.

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

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

Funding 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).

Compliance with Ethical Standards

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

References

- Almeida, T., Alonso, I., Martins, S., Ramos, E. M., Azevedo, L., Ohno, K., et al. (2009). Ancestral origin of the ATTCT repeat expansion in spinocerebellar ataxia type 10 (SCA10). *PLoS ONE*, *4*(2), e4553.
- Alonso, I., Jardim, L. B., Artigalás, O., Saraiva-Pereira, M. L., Matsuura, T., Ashizawa, T., et al. (2006). Reduced penetrance of intermediate size alleles in spinocerebellar ataxia type 10. *Neurology*, *66*(10), 1602–1604.
- Ashizawa, T. (2012). Spinocerebellar ataxia type 10. *Handbook of Clinical Neurology*, *103*, 507–519.
- Avena, S., Via, M., Ziv, E., Pérez-Stable, E. J., Gignoux, C. R., Dejean, C., et al. (2012). Heterogeneity in genetic admixture across different regions of Argentina. *PLoS ONE*, *7*(4), e34695.
- Bortolini, M. C., González-José, R., Bonatto, S. L., & Santos, F. R. (2014). Reconciling pre-Columbian settlement hypotheses requires integrative, multi-disciplinary, and model-bound approaches. *Proceedings of the National Academy of Science of USA*, *111*(2), 213–214.
- Bortolini, M. C., Salzano, F. M., Thomas, M. G., Stuart, S., Nasanen, S. P., Bau, C. H., et al. (2003). Y-chromosome evidence for differing ancient demographic histories in the Americas. *American Journal of Human Genetics*, *73*(3), 524–539.
- Brusco, A., Gellera, C., Cagnoli, C., Saluto, A., Castucci, A., Michielotto, C., et al. (2004). Molecular genetics of hereditary spinocerebellar ataxia: Mutation analysis of spinocerebellar ataxia genes and CAG/CTG repeat expansion detection in 225 Italian families. *Archives of Neurology*, *61*(5), 727–733.
- Bushara, K., Bower, M., Liu, J., McFarland, K. N., Landrian, I., Hutter, D., et al. (2013). Expansion of the spinocerebellar ataxia type 10 (SCA10) repeat in a patient with Sioux Native American Ancestry. *PLoS ONE*, *8*(11), e81342.
- Fagundes, N. J., Kanitz, R., Eckert, R., Valls, A. C., Bogo, M. R., Salzano, F. M., et al. (2008). Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas. *American Journal of Human Genetics*, *82*(3), 583–592.
- Fujigasaki, H., Tardieu, S., Camuzat, A., Stevanin, G., LeGuern, E., Matsuura, T., et al. (2002). Spinocerebellar ataxia type 10 in the French population. *Annals of Neurology*, *51*(3), 408–409.
- Gallardo, M., & Soto, A. (2009). Clinical characterization of a Venezuelan family with spinocerebellar ataxia type 10. *Movement Disorder*, *24*, S12.
- Gatto, E. M., Gao, R., White, M. C., Uribe Roca, M. C., Etcheverry, J. L., Persi, G., et al. (2007). Ethnic origin and extrapyramidal signs in an Argentinian spinocerebellar ataxia type 10 family. *Neurology*, *69*(2), 216–218.
- Gheno, T. C., Furtado, G. V., Saute, J. A. M., Donis, K. C., Fontanari, A. M. V., Emmel, V. E., et al. (2017). Spinocerebellar ataxia type 10: Common haplotype and disease progression rate in Peru and Brazil. *European Journal of Neurology*, *24*(7), 892–e36.
- González-José, R., Bortolini, M. C., Santos, F. R., & Bonatto, S. L. (2008). The peopling of America: Craniofacial shape variation on a continental scale and its interpretation from an interdisciplinary view. *American Journal of Physical Anthropology*, *137*(2), 175–187.
- Grewal, R. P., Achari, M., Matsuura, T., Durazo, A., Tayag, E., Zu, L., et al. (2002). Clinical features and ATTCT repeat expansion in spinocerebellar ataxia type 10. *Archives of Neurology*, *59*(8), 1285–1290.
- Grewal, R. P., Tayag, E., Figueroa, K. P., Zu, L., Durazo, A., Nunez, C., et al. (1998). Clinical and genetic analysis of a distinct autosomal dominant spinocerebellar ataxia. *Neurology*, *51*(5), 1423–1426.
- Kapur, S. S., & Goldman, J. G. (2012). Two in one: Report of a patient with spinocerebellar ataxia type 2 and 10. *Archives of Neurology*, *69*(9), 1200–1203.
- Leonardi, L., Marcotulli, C., McFarland, K. N., Tessa, A., DiFabio, R., Santorelli, F. M., et al. (2014). Spinocerebellar ataxia type 10 in Peru: The missing link in the Amerindian origin of the disease. *Journal of Neurology*, *261*(9), 8–11.
- Lins, T. C., Vieira, R. G., Abreu, B. S., Grattapaglia, D., & Pereira, R. W. (2010). Genetic composition of Brazilian population samples based on a set of twenty-eight ancestry informative SNPs. *American Journal of Human Biology*, *22*(2), 187–192.
- Matsuura, T., Achari, M., Khajavi, M., Bachinski, L. L., Zoghbi, H. Y., & Ashizawa, T. (1999). Mapping of the gene for a novel spinocerebellar ataxia with pure cerebellar signs and epilepsy. *Annals of Neurology*, *45*(3), 407–411.
- Matsuura, T., Ranum, L. P., Volpini, V., Pandolfo, M., Sasaki, H., Tashiro, K., et al. (2002). Spinocerebellar ataxia type 10 is rare in populations other than Mexicans. *Neurology*, *58*(6), 983–984.
- Matsuura, T., Yamagata, T., Burgess, D. L., Rasmussen, A., Grewal, R. P., Watase, K., et al. (2000). Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. *Nature Genetics*, *26*(2), 191–194.
- Miller, A. S., Dykes, D. D., & Plesky, H. F. (1998). A sample salting-out procedure of extracting DNA from nucleated cells. *Nucleic Acids Research*, *16*(3), 1215.
- Naito, H., Takahashi, T., Kamada, M., Morino, H., Yoshino, H., Hattori, N., et al. (2017). First report of a Japanese family with spinocerebellar ataxia type 10: The second report from Asia after a report from China. *PLoS ONE*, *12*(5), e0177955.
- Network 4.5.1.6. <http://www.fluxus-engineering.com/>. Accessed June 2, 2015.
- Price, A. L., Patterson, N., Yu, F., Cox, D. R., Waliszewska, A., McDonald, G. J., et al. (2007). A genome-wide admixture map in Latino populations. *American Journal of Human Genetics*, *80*(6), 1024–1036.
- Raskin, S., Ashizawa, T., Teive, H. A., Arruda, W. O., Fang, P., Gao, R., et al. (2007). Reduced penetrance in a Brazilian family with spinocerebellar ataxia type 10. *Archives of Neurology*, *64*(4), 591–594.
- Rasmussen, A., Matsuura, T., Ruano, L., Yescas, P., Ochoa, A., Ashizawa, T., et al. (2001). Clinical and genetic analysis of four Mexican families with spinocerebellar ataxia 10. *Annals of Neurology*, *50*(2), 234–239.
- Roxburgh, H. R., Smith, O. C., Lim, J. G., Bachman, D. F., Byrd, E., & Bird, T. D. (2013). The unique co-occurrence of spinocerebellar ataxia type 10 (SCA10) and Huntington disease. *Journal of the Neurological Science*, *324*(1–2), 176–178.
- Ruiz-Linares, A., Adhikari, K., Acuña-Alonzo, V., Quinto-Sanchez, M., Jaramillo, C., Arias, W., et al. (2014). Admixture in Latin America: Geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals. *PLoS Genetics*, *10*(9), e1004572.

- Seldin, M. F., Tian, C., Shigeta, R., Scherbarth, H. R., Silva, G., Belmont, J. W., et al. (2007). Argentine population genetic structure: A large variance in Amerindian contribution. *American Journal of Physical Anthropology*, 132(3), 455–462.
- Stephens, M., Smith, N. J., & Donnelly, P. (2001). A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, 68(4), 978–989.
- Sulek-Piatkowska, A., Zdzienicka, E., Raczyńska-Rakowicz, M., Krysa, W., Rajkiewicz, M., Szirkowiec, W., et al. (2010). The occurrence of spinocerebellar ataxias caused by dynamic mutations in Polish patients. *Neurologia i Neurochirurgia Polska*, 44(3), 238–245.
- Tarazona-Santos, E., Carvalho-Silva, D. R., Pettener, D., Luiselli, D., De Stefano, G. F., Labarga, C. M., et al. (2001). Genetic differentiation in South Amerindians is related to environmental and cultural diversity: Evidence from the Y chromosome. *American Journal of Human Genetics*, 68(6), 1485–1496.
- Teive, H. A., & Ashizawa, T. (2013). Spinocerebellar ataxia type 10: From Amerindians to Latin Americans. *Current Neurology and Neuroscience Reports*, 13(11), 393.
- Teive, H. A., Munhoz, R. P., Arruda, W. O., Raskin, S., Werneck, L. C., & Ashizawa, T. (2011). Spinocerebellar ataxia type 10—A review. *Parkinsonism Related Disorders*, 17(9), 655–661.
- Teive, H. A., Munhoz, R. P., Raskin, S., Arruda, W. O., de Paola, L., Werneck, L. C., et al. (2010). Spinocerebellar ataxia type 10: Frequency of epilepsy in a large sample of Brazilian patients. *Movement Disorder*, 25(16), 2875–2878.
- Teive, H. A., Roa, B. B., Raskin, S., Fang, P., Arruda, W. O., Neto, Y. C., et al. (2004). Clinical phenotype of Brazilian families with spinocerebellar ataxia 10. *Neurology*, 63(8), 1509–1512.
- 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. *European Journal of Neurology*, 17(1), 124–128.
- Wang, S., Lewis, C. M., Jakobsson, M., Ramachandran, S., Ray, N., Bedoya, G., et al. (2007). Genetic variation and population structure in Native Americans. *PLoS Genetics*, 3(11), e185.
- Wang, K., McFarland, K. N., Liu, J., Zeng, D., Landrian, I., Xia, G., et al. (2015). Spinocerebellar ataxia type 10 in Chinese Han. *Neurology Genetics*, 1(3), e26.
- Wang, J., Shen, L., Lei, L., Xu, Q., Zhou, J., Liu, Y., et al. (2011). 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*, 36(6), 482–489.