The Cerebellum (2019) 18:841–848
https://doi.org/10.1007/s12311-019-01057-x

ORIGINAL PAPER

ATXN10 Microsatellite Distribution in a Peruvian Amerindian Population

Diego Véliz-Otani[1] · Miguel Inca-Martinez[1,2] · Giovana B. Bampi[3,4,5] · Olimpio Ortega[1] · Laura B. Jardim[3,4,5,6] · Maria Luiza Saraiva-Pereira[3,4,5,7] · Pilar Mazzetti[1] · Mario Cornejo-Olivas[1,8]

Published online: 24 July 2019

Abstract

Spinocerebellar ataxia type 10 (SCA10) is a repeat expansion disease occurring mostly in Latin America, suggesting that the mutation spread with the peopling of the Americas, or that Amerindian populations, have a higher *ATXN10* mutability. High frequency of large normal alleles is associated with prevalence and relative frequency of other repeat expansion diseases. To test whether the allele distribution of the SCA10-causing *ATXN10* microsatellite in an Amerindian Peruvian population differs from that of other populations. The *ATXN10* allele distribution in a Quechua Peruvian population from Puno, Peru, is similar to that of Finland. Mean allele size and mode were also similar to those of Mexico, Japan, and white Europeans. *ATXN10* allele distribution in a healthy Amerindian population from Peru does not differ from that of other populations.

Keywords Amerindian · Quechua · *ATXN10* · Spinocerebellar ataxia type 10 · Large normal allele

Introduction

Spinocerebellar ataxia type 10 is a neurodegenerative disease with autosomal dominant inheritance caused by the abnormal expansion of an $(ATTCT)_n$ microsatellite at the *ATXN10* gene (locus 22q13.31) [1, 2]. Normal alleles range between 9 and 32 repeats [2–4], and mutant alleles between 280 and 4500 repeats [2]. Incomplete penetrance is observed between 280 and 850 repeats [4–6]. The frequency of incomplete penetrance alleles is unknown since $(ATTCT)_n$ genotyping is performed mostly by RP-PCR, an assay that is cost-effective but uncapable of repeat-sizing [7].

SCA10 cases have been identified almost exclusively in individuals with Native American ancestry [8–17] and are probably absent in European lineages [18]. Only two SCA10 cases without Native American ancestry have been reported, a Chinese Han [19] and a Japanese [20]. Twenty-seven self-declared mestizo SCA10 cases from 21 Peruvian families have been described, making it the most common SCA in Peru reported so far [16].

The limited geographic distribution of SCA10 cases is likely explained by one of the two following hypotheses: (i) The SCA10 expansion has occurred only once or (ii) the mutation has occurred more than once but in individuals carrying a predisposing genotype. Both hypotheses assume that either an ancient mutation or predisposing alleles spread from Beringia to the Americas during the last glaciation period. Subsequent back migration to Asian continent might explain the report of 2 SCA10 families from China and Japan [19, 20].

Frequency of large normal alleles has been proposed to be associated with prevalence of other repeat expansion diseases

✉ Diego Véliz-Otani
diegovelizotani@gmail.com

[1] Neurogenetics Research Center, Instituto Nacional de Ciencias Neurológicas, 1271 Ancash St., Barrios Altos, Lima, Lima, Peru

[2] Lerner Research Institute, Genomic Medicine, Cleveland Clinic Foundation, Cleveland, OH, USA

[3] Post-Graduation Program of Genetics and Molecular Biology, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

[4] Medical Genetics Service, Hospital de Clínicas de Porto Alegre, Ramiro Barcelos 2350, Porto Alegre, Rio Grande do Sul 90035-903, Brazil

[5] Laboratory of Genetics Identification – Research Center, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil

[6] Department of Internal Medicine, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

[7] Department of Biochemistry, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

[8] Center for Global Health, Universidad Peruana Cayetano Heredia, Lima, Peru

(Table 1) [21–23]. Some spinocerebellar ataxias, such as MJD/SCA3, SCA6, and DRPLA, show a relatively high prevalence in Japanese population, where large normal alleles are found in high frequency. Conversely, Caucasian populations show a comparatively low prevalence for these diseases and a low frequency of large normal alleles [21]. Furthermore, SCA1 and SCA2 show a low prevalence in Japanese compared with Caucasian population, and a low frequency of large normal alleles.

Extensive research on Huntington Disease (HD) has identified a plausible mechanism linking large normal alleles frequency with disease prevalence. Populations with higher prevalence of HD (caused by the expansion of a $(CAG)_n$ microsatellite on the *HTT* gene) show larger alleles on average [22, 23]. Specifically, Western European and White Hispanic populations show the highest HD prevalence worldwide as well as the largest mean *HTT* allele sizes on healthy population. Moreover, these populations also exhibit the greatest frequency of large normal alleles and intermediate alleles in control population, as compared with populations with low HD prevalence, such as Black South Africans, Chinese, Japanese, and Finnish populations [22]. Furthermore, there is a high correlation between frequency of large normal alleles and disease prevalence. In addition, instability of the *HTT* microsatellite increases with allele size, even within non-pathogenic range [24]. Altogether, these observations suggest that the higher mutation rate of longer alleles increases the disease risk in the overall population and explains the correlation between frequency of large normal alleles and repeat-expansion-disease prevalence [23].

The *ATXN10* allele distribution has not been studied in Native American populations. Most of the SCA10 cases have some Amerindian ancestry; however, genetic factors that might increase the *ATXN10* mutation rate have not been explored, with most studies focusing on testing the unique-origin hypothesis based on haplotype studies [17, 19]. A previous study analyzed the haplotype frequency in a sample of 49 Peruvian control subjects of self-declared Quechua ancestry [16]. However, the *ATXN10* allele distribution and frequency of large normal alleles (≥ 17 pentanucleotide repeats [4]) in Native American populations remain unstudied.

The Peruvian population results from the admixture between diverse Native American ethnic groups and Europeans, mostly Spaniards. The admixture between the Native American groups began before the arrival of the Spaniards; therefore, most current Andean and Coastal ethnic groups are closely related [25]. The current Peruvian population shows ~ 80% average Native American continental ancestry, but there exists great variability within and between departments (i.e., geopolitical divisions) [25, 26]. The largest Native American ethnic group in Peru, according to reported self-identity, is the Quechua, which makes up 22% of the population over 12 years old [27]. Puno, the department from where the sample donors of this study originate, has over 90% Amerindian ancestry [25].

In this study, we analyzed whether there is evidence of an association between SCA10 frequency in Peruvian population and frequency of large normal alleles in a healthy Native American population. For that purpose, we genotyped the *ATXN10* microsatellite in samples from 49 self-identified Quechua individuals from Puno, Peru, collected for a previous study on the genetics of Parkinson Disease [28]. We also reanalyzed previously obtained [16] genotype data of 4 SNP and 4 STR markers to distinguish locus-specific from demographic phenomena. This is the first study analyzing the *ATXN10* allele distribution in a Native American community.

Table 1 Frequency of large normal alleles

	Puno	Finland	*MJE
Normal alleles	95 (96.9%)	471 (93.8%)	561 (92.9%)
Large normal alleles	3 (3.1%)	31 (6.2%)	43 (7.1%)

*Pooled Mexican, Japanese, and European samples from ref. [2]

Patients and Methods

Subjects

DNA samples were obtained from the Neurogenetics Research Center's DNA bank. All 49 DNA samples were isolated from peripheral blood mononuclear cells from self-identified Quechua individuals, whose previous 2 generations had lived in the Puno region in Peru. These individuals were recruited for a previous study on the genetics of Parkinson disease in Latin America; all of them were healthy at the time of collection and had no family history of neurological disorders. Analysis of 29 ancestry informative markers revealed that this sample had over 90% Amerindian ancestry [29]. These are the same samples that were used for a SCA10 haplotype study [16]. The sample donors gave informed consent for further use of their DNA samples in other genetic studies related to neurological disorders. The study was approved by the Institutional Review Board at *Instituto Nacional de Ciencias Neurológicas* (INCN). Statistical power analysis using the Minsage software showed that the used sample size reached 97.4% probability to detect a large normal allele (≥ 17 repeats, population frequency 7.1% [4]).

Methods

Genotyping

ATXN10 microsatellite was PCR-amplified using a previously described protocol [13]. PCR products were measured by

capillary electrophoresis using Hi-Di formamide buffer to provide denaturing conditions at HCPA using an ABI310XL instrument. Genotype data of 4 STRs and 4 SNPs obtained for a previous study were included [16].

Statistical and Genetic Analyses

Frequency of large normal alleles in Peruvian population was compared with available data from Finnish [22], Mexican, Japanese and Caucasian populations [2] by a proportion test. Normality was assessed by the Shapiro-Wilk test. Allele distributions from Peruvian and Finnish population were compared by a Kolmogorov-Smirnov test modified to work with discrete distributions implemented in the R package dgof v1.2 [30]. We did not include other population in this analysis because tabulated data was unavailable. All statistical analyses were performed using R v3.3 [31]. Hardy-Weinberg equilibrium (HWE) was tested using genepop [32]. F_{IS} statistic was estimated by the Weir and Cockerham method [33]. We used the tabix program [34] to search the 1000 Genomes Project [35] database for single nucleotide variants (SNVs) or small insertions/deletions (indels) in the *ATXN10* gene (Chr 22:45671798–45845307, version GRCh38.p12). We checked whether any variant was located in the primer-binding regions (forward: (5′)45795261–45795281(3′), reverse: (3′)45795438–45795457(5′)).

Results

Alleles showed a non-normal (Shapiro-Wilk test p value = 3.177×10^{-6}) unimodal distribution ranging between 11 and 17 repeats, with mean allele size 14 ± 1.13. The most frequent allele was 14 (48.0%, Fig. 1), which is in agreement with most studied populations, except for one European population whose most frequent allele was 13, followed by the 14 repeats allele (Table 2).

Three (3.1%) large normal alleles were found, all 17-repeat long. No alleles over 17 repeats were found. Frequency of large normal alleles was 6.2% (31/502) in Finnish population [36], and 7.1% (43/604) in a combined sample of Mexican, Japanese, and White Europeans, called "MJE sample" in the remaining of this text [2] (Table 1). Frequency of large normal alleles was similar in Puno (Quechua Amerindian population from this study), Finland [36], and MJE (proportion test p value = 0.303).

We then used a modified Kolmogorov-Smirnov test to compare the allele distribution in the Quechua sample with that of Finland [36], and found no statistically significant differences ($p = 0.333$). The empirical cumulative distributions of both samples are shown in Fig. 2.

The $(ATTCT)_n$ microsatellite was not in Hardy-Weinberg equilibrium in the Quechua sample ($p = 0.0063 \pm 0.0021$ [S.E]). The observed heterozygosity (Ho = 0.49) was lower than the expected heterozygosity (He = 0.706) and the F_{IS} coefficient was very high (0.308). The low heterozygosity observed in the Quechua sample disagrees with the observed values in Cyprus, MJE, and Venezuela [38] (Table 2).

In order to distinguish between demographic and *ATXN10*-specific phenomena, we also analyzed 4 STRs and 1 SNP markers flanking the $(ATTCT)_n$ microsatellite that had been previously genotyped to study the *ATXN10* haplotypes in this Quechua sample [16]. The original dataset consisted of 4 STRs and 4 SNPs, but 3 SNPs were excluded from this study because they were either monomorphic or had a private allele in our sample. One STR marker (GATA030P) showed a deviation from HWE significant at the nominal value ($p = 0.017$, Table 3). None of the flanking markers showed statistical significant deviation from HWE after Bonferroni correction for multiple testing ($p > 0.05/6$). However, the $(ATTCT)_n$ repeat deviation from HWE remained statistically significant

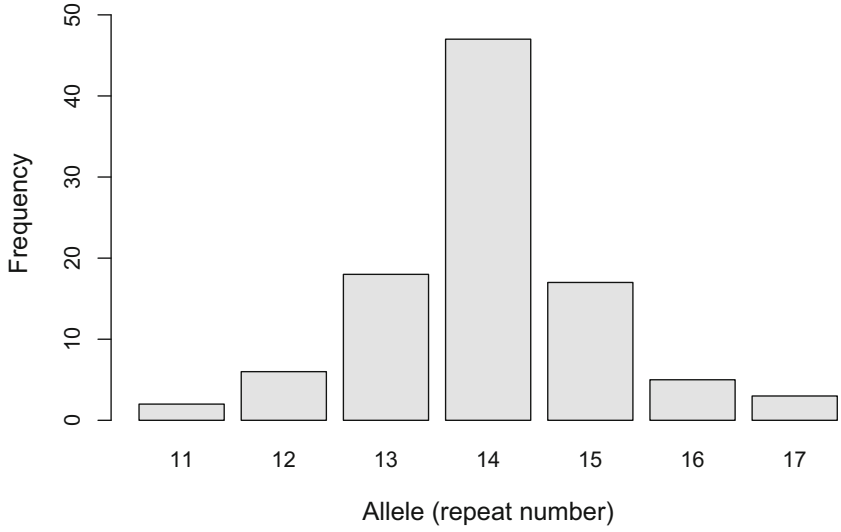

Fig. 1 Allelic distribution of the *ATXN10* pentanucleotide in the Quechua Population

Table 2 Allele distribution summary statistics

Location	N†	HWE‡	Mode (Frequency %)	Mean	Range	Heterozygosity	Reference
Puno	98	No	14 (48%)	14	11–17	0.49	This study
Finland	502	ND*	14 (38.6%)	14.07	10–20	ND	[36]
Cyprus	116	ND	14 (37%)	ND	10–20	0.81	[37]
Japan	100	Yes	14(27%)	ND	10–20	0.82	[2]
Europe	250	Yes	13 (30%)	ND	11–20		
Mexico	254	Yes	14 (29%)	ND	11–22		

† Number of chromosomes, ‡ Hardy-Weinberg equilibrium, *not determined

(Table 3). The F_{IS} statistic of the 5 (4 STRs, 1 SNP) remaining flanking markers ranged from −0.168 to 0.130, and the weighted F_{IS} across loci was −0.021. The Rho_{IS} statistic of the 4 flanking STRs ranged between −0.219 and 0.407, and the weighted Rho_{IS} across loci was of 0.009. DNA variants in the primer-binding sites may give rise to null alleles and decrease the apparent heterozygosity [39]. Therefore, we searched the 1000 genomes database for SNVs or indels near the $(ATTCT)_n$ microsatellite but found none.

Discussion

SCA10 is the most frequent ataxia in Peru known so far, and we wondered whether this is caused by either a founder effect of an ancestral mutant allele or whether is caused by high frequency of predisposing alleles in Native populations from America. Large normal alleles have been proposed to act as predisposing alleles for several repeat expansion disorders [21–23]. We studied the normal *ATXN10* allele distribution in an exemplary Native American ethnic group, the Quechua population from Puno, Peru. ATTCT length distribution among normal Quechua population was unimodal with little asymmetry, and similar to those of Caucasian, Mexican, and Japanese populations. Altogether, our results suggest that differences in allele distribution do not explain differences in SCA10 relative frequency across populations.

Haplotype studies suggested that SCA10 has a unique ancestral origin. The occurrence of a single ancient mutation in an Amerindian or proto-Amerindian population prior to the split between Asians and Amerindian lineages would explain the limited geographic distribution of SCA10 cases [13, 16, 40]. A 4-SNP haplotype (CGGC) is found in all expansion-bearing chromosomes from case reports from China, Japan, Venezuela, and the Native American Sioux Tribe [17, 19, 20, 38]. Furthermore, this haplotype shows a high frequency in mutation-carrying chromosomes of Brazilian- and Mexican-affected families [16, 40]. Nevertheless, association between specific haplotypes and expanded alleles does not necessarily point to a unique ancestral mutation's origin. For example, large normal alleles of the *HTT* gene are more frequent on specific haplotypes significantly associated with HD [23]. This phenomenon raises the question of whether chromosomes bearing these haplotypes also bear *cis*-acting elements that increase HD risk, or if these haplotypes are merely associated with HD through linkage disequilibrium with expansion-prone, large normal alleles [22, 23, 41].

Therefore, we analyzed the *ATXN10* allele distribution in a Quechua sample and compared it with that of other populations in order to assess whether frequency of large normal alleles in a Native Peruvian population is higher than in other

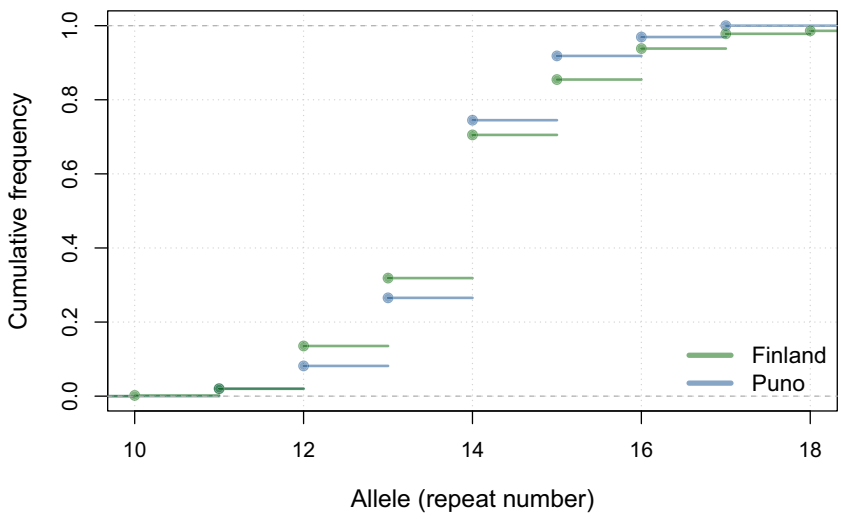

Fig. 2 Empirical cumulative distribution of allele size in Puno and Finland

Table 3 Summary statistics

Locus	[a]Distance	[b]Ho	[c]He	[d]HWE (*p* value)	F_{IS}	Rho_{IS}
D22S1140	− 1.5 Mb	0.184	0.211	0.193	0.130	− 0.019
D22S532	− 68 Kb	0.694	0.645	0.495	− 0.077	0.060
rs5764850	− 1198 bp	0.102	0.098	1.000	− 0.044	NA
(ATTCT)n	[e]NA	0.490	0.706	0.007*	0.308	0.316
rs72556348	48 bp	0.000	0.000	[f]ND	ND	NA
rs72556349	303 bp	0.020	0.020	ND	ND	NA
rs72556350	370 bp	0.000	0.000	ND	ND	NA
D22S1153	21.3 Kb	0.551	0.473	0.715	− 0.168	− 0.219
GATA030P	78.5 Kb	0.612	0.673	0.017	0.092	0.407
Weighted (with ATTCT)				0.025	0.062	0.044
Weighted (without ATTCT)				0.200	− 0.021	0.009

[a] Distance from each marker to the $(ATTCT)_N$ microsatellite

[b] Observed heterozygosity

[c] Expected heterozygosity

[d] Hardy-Weinberg equilibrium

[e] Not applicable

[f] Not determined

[*] Significant after Bonferroni correction: $p < (0.05/6)$

ethnic groups, which would provide a plausible mechanism explaining the high relative frequency of SCA10 in Peruvian population.

Allele distribution in the Quechua sample is similar to that of other populations. All populations analyzed so far, including the Quechua one, have shown a nearly symmetrical non-normal unimodal distribution (Fig. 1), which may suggest a negligible or little upward-biased mutation process [42, 43]. Furthermore, the 14-repeats allele is the most common in every population studied, except for the White Europeans, where it was the second most frequent, close to the 13-repeat allele, which was the most frequent [2]. The allele range observed in this study [11–17] was narrower than the observed in other populations (Table 2), probably due to the small sample size used in this study. Frequency of large normal alleles in Puno was similar to that of MJE [2] and Finnish [36] samples (Table 1). Moreover, a modified Kolmogorov-Smirnov test showed no difference between the allele distributions of Puno and Finland (Fig. 2), in spite of the great genetic distance among Amerindian and European populations [44]. The similar allele distributions in all studied populations coupled with its nearly symmetrical distribution suggest that the evolution of the ATTCT microsatellite is constrained by its mutational process [45]; however, we are unable to formally test this hypothesis with the current data.

Genotypic frequency of Puno differed from other locations in spite of the similar allele frequencies. The Quechua sample showed low heterozygosity (0.49) compared with Cyprus (0.81), Venezuela (0.82), and MJE (Table 2). The *ATXN10* microsatellite was not under Hardy-Weinberg equilibrium (HWE) in Puno, in contrast with MJE (Table 2). The HWE in MJE is noteworthy, because mixing samples with different allele frequency give rise to the Wahlund effect, consisting of an apparent decreased heterozygosity and hence departure from HWE [46]. Thus, the HWE in MJE further stresses out that the *ATXN10* microsatellite shows similar allele frequencies in otherwise genetically disparate populations.

Deviation from HWE in the Quechua sample is unlikely caused solely by null alleles. The most recognized source of null alleles is the presence of DNA variants in the primer-binding sites [39]. Therefore, we searched the 1000 Genomes Project data for SNVs and small indels in the primer-binding sites but found none. Other possible cause of null alleles is differential amplification of alleles of different size [39]. This phenomenon would cause different allele frequencies between populations, but allele distribution in the Quechua sample was similar to that of other populations (Fig. 2 and Table 2). Altogether, our results suggest that the low heterozygosity is unlikely caused by null alleles.

Low heterozygosity is unlikely caused by population inbreeding or structure. The district, from where the samples were obtained, has a population of just over 6500 individuals, according to the 2017 National Census [27]. Nevertheless, the weighted F_{IS} and Rho_{IS} across 4 flanking STRs and 1 SNP were close to zero (− 0.021 and 0.009, respectively, Table 3), suggesting that there is little inbreeding in this population. Similarly, if low heterozygosity at the ATTCT microsatellite was caused by population structure (Wahlund effect), it would cause a high F_{IS} value in the flanking markers; however, this was not observed [46].

The ATTCT microsatellite shows replication-mediated instability through a length-dependent DNA unwinding activity. The *ATXN10* expanded alleles show paternal transmission-dependent intergenerational instability, intra-tissue mosaicism, and inter-tissue repeat-length variability [47]. By contrast, the normal alleles show little instability. $(ATTCT)_{27}$ and $(ATTCT)_{48}$ show size-instability, but $(ATTCT)_8$ and $(ATTCT)_{13}$ do not in a yeast model. Furthermore, $(ATTCT)_{27}$ and $(ATTCT)_{48}$ can replace the c-myc replicator DNA unwinding element, but $(ATTCT)_8$ and $(ATTCT)_{13}$ cannot [48]. *ATXN10* instability seems to be mediated by a strand-switching mechanism during replication, where the nascent Okazaki fragment from the lagging strand functions as template for the synthesis of the leading nascent strand [49]. Under this model, the size of each expansion event should be a multiple of the Okazaki fragment size, thus explaining the massive ATTCT tract size observed in SCA10. This instability mechanism would function in expanded alleles only, in agreement with the unique-origin hypothesis. Under this scenario, the ancestral mutation would have arisen by gene conversion eased by the abundance of ALU and LINE elements in the intron 9 of *ATXN10* [40].

Altogether, our data suggest that allele distribution of the *ATXN10* allele in the Peruvian sample does not differ from that of other populations. Thus, we found no evidence to support that high frequency of SCA10 cases in Peru is caused by a high frequency of mutation-prone alleles in the most prominent Native American ancestral component in Peruvian population. These features might contribute to better understand the origin of the SCA10 mutation with significant contribution on genetic counseling in specific populations. Our results help understand the epidemiology of SCA10 in Peru and possibly other Latin American populations.

Future studies should focus on the evolution of the ATTCT microsatellite at the *ATXN10* gene. Founder effects and population divergence due to genetic drift after long periods of population splits should cause different allele frequencies between populations [46]; however, this was not the case. The similar allele frequencies in all the studied populations so far suggest that the ATTCT microsatellite evolution is constrained by its mutation mechanisms. Moreover, the low heterozygosity specific to the ATTCT microsatellite, and not observed in the flanking markers, suggests that this region may be under selection in the Quechua population [50]. Inferring the selection mechanism possibly acting on this locus requires more studies.

One caveat of this study was the relatively small sample size, which prevented us from comparing $(ATTCT)_n$ repeat size allele frequency among *ATXN10* haplotypes. We suggest that a well-powered joint analysis of *ATXN10* haplotypes and ATTCT repeat-length distribution would help further understand the limited geographical distribution of SCA10 cases.

Conclusion

ATXN10 allele distribution and frequency of large normal alleles in a Quechua community from Puno are similar to those of Finland, Mexico, Japan, and other European countries, implying that frequency of large normal alleles does not explain differences in SCA10 frequency across populations. Thus, our results favor the alternate hypothesis that SCA10 has a unique ancestral origin. Studying the repeat size allele distribution of each *ATXN10* haplotype would further help pinpoint the hypothesis that best explains the geographical distribution of SCA10.

Acknowledgments We are grateful to Victoria Marca and Karina Milla-Neyra for the administrative and logistic support as well as lab assistance.

Compliance with Ethical Standards

The sample donors gave informed consent for further use of their DNA samples in other genetic studies related to neurological disorders. The study was approved by the Institutional Review Board at *Instituto Nacional de Ciencias Neurológicas* (INCN).

Conflict of Interest The authors declare that they have no competing interests.

References

1. Zu L, Figueroa KP, Grewal R, Pulst S-M. Mapping of a new autosomal dominant spinocerebellar ataxia to chromosome 22. Am J Hum Genet. 1999;64:594–9.
2. 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;26:191–4.
3. Wang J, Wu Y, Lei L, Shen L, Jiang H, Zhou Y, et al. Polynucleotide repeat expansion of nine spinocerebellar ataxia subtypes and dentatorubral-pallidoluysian atrophy in healthy Chinese Han population. 2010.
4. Matsuura T, Fang P, Pearson CE, Jayakar P, Ashizawa T, Roa BB, et al. Interruptions in the expanded ATTCT repeat of spinocerebellar ataxia type 10: repeat purity as a disease modifier? Am J Hum Genet. 2006;78:125–9.
5. Alonso I, Jardim LB, Artigalas O, Saraiva-Pereira ML, Matsuura T, Ashizawa T, et al. Reduced penetrance of intermediate size alleles in spinocerebellar ataxia type 10. Neurology. 2006;66:1602–4.
6. Raskin S, Ashizawa T, Teive HA, Arruda WO, Fang P, Gao R, et al. Reduced penetrance in a Brazilian family with spinocerebellar ataxia type 10. Arch Neurol. 2007;64:591–4.
7. Matsuura T, Ashizawa T. Polymerase chain reaction amplification of expanded ATTCT repeat in spinocerebellar ataxia type 10. Ann Neurol. 2002;51:271–2.
8. Teive HAG, Munhoz RP, Raskin S, Arruda WO, de Paola L, Werneck LC, et al. Spinocerebellar ataxia type 10: frequency of epilepsy in a large sample of Brazilian patients. Mov Disord. 2010;25:2875–8.
9. Grewal RP, Achari M, Matsuura T, Durazo A, Tayag E, Zu L, et al. Clinical features and ATTCT repeat expansion in spinocerebellar ataxia type 10. Arch Neurol. 2002;59:1285–90.

10. Rasmussen A, Matsuura T, Ruano L, Yescas P, Ochoa A, Ashizawa T, et al. Clinical and genetic analysis of 4 Mexican families with spinocerebellar ataxia type 10. Ann Neurol. 2001;50:234–9.
11. Gatto EM, Gao R, White MC, Roca MCU, Etcheverry JL, Persi G, et al. Ethnic origin and extrapyramidal signs in an Argentinean spinocerebellar ataxia type 10 family. Neurology. 2007;69:216–8.
12. de Castilhos RM, Furtado GV, Gheno TC, Schaeffer P, Russo A, Barsottini O, et al. Spinocerebellar ataxias in Brazil—frequencies and modulating effects of related genes. Cerebellum. 2014;13:17–28.
13. Gheno TC, Furtado GV, Saute JAM, Donis KC, Fontanari AMV, Emmel VE, et al. Spinocerebellar ataxia type 10: common haplotype and disease progression rate in Peru and Brazil. Eur J Neurol. 2017;24:892–e36.
14. Abstracts of The Movement Disorder Society's Thirteenth International Congress of Parkinson's Disease and Movement Disorders. Mov Disord. 2009;24:S1–S653.
15. Roxburgh RH, Smith CO, 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;324:176–8.
16. Bampi GB, Bisso-Machado R, Hünemeier T, Gheno TC, Furtado GV, Veliz-Otani D, et al. Haplotype study in SCA10 families provides further evidence for a common ancestral origin of the mutation. NeuroMolecular Med. 2017;19:501–9.
17. Bushara K, Bower M, Liu J, McFarland KN, Landrian I, Hutter D, et al. Expansion of the spinocerebellar ataxia type 10 (SCA10) repeat in a patient with Sioux Native American ancestry. PLoS One. 2013;8:e81342.
18. Matsuura T, Ranum LPW, Volpini V, Pandolfo M, Sasaki H, Tashiro K, et al. Spinocerebellar ataxia type 10 is rare in populations other than Mexicans. Neurology. 2002;58:983–3.
19. Wang K, McFarland KN, Liu J, Zeng D, Landrian I, Xia G, et al. Spinocerebellar ataxia type 10 in Chinese Han. Neurol Genet [Internet]. 2015 [cited 2018 Jan 10];1. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4809459/
20. Naito H, Takahashi T, Kamada M, Morino H, Yoshino H, Hattori N, et al. First report of a Japanese family with spinocerebellar ataxia type 10: the second report from Asia after a report from China. PLoS One. 2017;12:e0177955.
21. Takano H, Cancel G, Ikeuchi T, Lorenzetti D, Mawad R, Stevanin G, et al. Close associations between prevalences of dominantly inherited spinocerebellar ataxias with CAG-repeat expansions and frequencies of large normal CAG alleles in Japanese and Caucasian populations. Am J Hum Genet. 1998;63:1060–6.
22. Squitieri F, Andrew SE, Goldberg YP, Kremer B, Spence N, Zelsler J, et al. DNA haplotype analysis of Huntington disease reveals clues to the origins and mechanisms of CAG expansion and reasons for geographic variations of prevalence. Hum Mol Genet. 1994;3:2103–14.
23. Kay C, Collins JA, Wright GEB, Baine F, Miedzybrodzka Z, Aminkeng F, et al. The molecular epidemiology of Huntington disease is related to intermediate allele frequency and haplotype in the general population. Am J Med Genet B Neuropsychiatr Genet. 2018;177:346–57.
24. Semaka A, Kay C, Doty C, Collins JA, Bijlsma EK, Richards F, et al. CAG size-specific risk estimates for intermediate allele repeat instability in Huntington disease. J Med Genet. 2013;50:696–703.
25. Harris DN, Song W, Shetty AC, Levano KS, Cáceres O, Padilla C, et al. Evolutionary genomic dynamics of Peruvians before, during, and after the Inca Empire. PNAS. 2018;115:E6526–35.
26. Sandoval JR, Salazar-Granara A, Acosta O, Castillo-Herrera W, Fujita R, Pena SD, et al. Tracing the genomic ancestry of Peruvians reveals a major legacy of pre-Columbian ancestors. J Hum Genet. 2013;58:627–34.
27. Instituto Naional de Estadística e Informática. Resultados Definitivos de los Censos Nacionales 2017. Puno [Internet]. Lima, Perú: Instituto Nacional de Estadística e Informática; 2018 Oct. Available from: https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1563/.
28. Cornejo-Olivas M, Marca V, Dorschner MO, Inca Martinez M, Medina A, Shetty AC, et al. Target sequencing analysis of Parkinson's disease genes in a healthy Amerindian population from Puno-Peru. San Diego, California, USA; 2014 [cited 2018 Jul 2]. Available from: http://www.ashg.org/2014meeting/abstracts/fulltext/f140122321.htm.
29. Cornejo-Olivas M, Torres L, Velit-Salazar MR, Inca-Martinez M, Mazzetti P, Cosentino C, et al. Variable frequency of LRRK2 variants in the Latin American research consortium on the genetics of Parkinson's disease (LARGE-PD), a case of ancestry. NPJ Parkinsons Dis. 2017;3:19.
30. Arnold TB, Emerson JW, worldwide RCT and contributors. dgof: discrete goodness-of-fit tests [internet]. 2013 [cited 2018 Aug 2]. Available from: https://CRAN.R-project.org/package=dgof.
31. R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013. 2014.
32. Raymond M, Rousset F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered. 1995;86:248–9.
33. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. Evolution. 1984;38:1358–70.
34. Li H. Tabix: fast retrieval of sequence features from generic TAB-delimited files. Bioinformatics. 2011;27:718–9.
35. The 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012;491:56–65.
36. Juvonen V, Hietala M, Kairisto V, Savontaus M-L. The occurrence of dominant spinocerebellar ataxias among 251 Finnish ataxia patients and the role of predisposing large normal alleles in a genetically isolated population. Acta Neurol Scand. 2005;111:154–62.
37. Votsi C, Zamba-Papanicolaou E, Georghiou A, Kyriakides T, Papacostas S, Kleopa KA, et al. Investigation of SCA10 in the Cypriot population: further exclusion of SCA dynamic repeat mutations. J Neurol Sci. 2012;323:154–7.
38. Paradisi I, Ikonomu V, Arias S. Spinocerebellar ataxias in Venezuela: genetic epidemiology and their most likely ethnic descent. J Hum Genet. 2016;61:215–22.
39. Dakin EE, Avise JC. Microsatellite null alleles in parentage analysis. Heredity. 2004;93:504–9.
40. Almeida T, Alonso I, Martins S, Ramos EM, Azevedo L, Ohno K, et al. Ancestral origin of the ATTCT repeat expansion in spinocerebellar ataxia type 10 (SCA10). PLoS One. 2009;4:e4553.
41. Warby SC, Montpetit A, Hayden AR, Carroll JB, Butland SL, Visscher H, et al. CAG expansion in the Huntington disease gene is associated with a specific and targetable predisposing haplogroup. Am J Hum Genet. 2009;84:351–66.
42. Falush D, Almqvist EW, Brinkmann RR, Iwasa Y, Hayden MR. Measurement of mutational flow implies both a high new-mutation rate for Huntington disease and substantial underascertainment of late-onset cases. Am J Hum Genet. 2001;68:373–85.
43. Falush D. Haplotype background, repeat length evolution, and Huntington's disease. Am J Hum Genet. 2009;85:939–42.
44. Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, et al. Worldwide human relationships inferred from genome-wide patterns of variation. Science. 2008;319:1100–4.
45. Falush D, Iwasa Y. Size-dependent mutability and microsatellite constraints. Mol Biol Evol. 1999;16:960–6.
46. Templeton AR. Population genetics and microevolutionary theory. 1st ed. Hoboken, N.J: Wiley-Liss; 2006.
47. Matsuura T, Fang P, Lin X, Khajavi M, Tsuji K, Rasmussen A, et al. Somatic and germline instability of the ATTCT repeat in

spinocerebellar ataxia type 10. Am J Hum Genet. 2004;74:1216–24.

48. Liu G, Bissler JJ, Sinden RR, Leffak M. Unstable spinocerebellar ataxia type 10 (ATTCT)·(AGAAT) repeats are associated with aberrant replication at the ATX10 locus and replication origin-dependent expansion at an ectopic site in human cells. Mol Cell Biol. 2007;27:7828–38.
49. Cherng N, Shishkin AA, Schlager LI, Tuck RH, Sloan L, Matera R, et al. Expansions, contractions, and fragility of the spinocerebellar ataxia type 10 pentanucleotide repeat in yeast. Proc Natl Acad Sci U S A. 2011;108:2843–8.
50. Walsh B, Lynch M. Evolution and selection of quantitative traits: I. Foundations. [Internet]. 1st ed. Unpublished manuscript [cited 2017 Mar 4]. Available from: nitro.biosci.arizona.edu/zbook/NewVolume_2/newvol2.html.