Short Report

Analysis of CAG repeats in five SCA loci in Mexican population: epidemiological evidence of a SCA7 founder effect


Spinocerebellar ataxias (SCA) are a heterogeneous group of neurodegenerative disorders. CAG (cytosine-adenine-guanine) trinucleotide repeat expansions in the causative genes have been identified as the cause of different SCA. In this study, we simultaneously genotyped SCA1, SCA2, SCA3, SCA6, and SCA7 applying a fluorescent multiplex polymerase chain reaction assay. We analyzed 10 families with SCA (64 patients) from five different communities of Veracruz, a Mexican southeastern state, and identified 55 patients for SCA7 and 9 for SCA2, but none for SCA1, SCA3, or SCA6. To our knowledge, this sample represents one of the largest series of SCA7 cases reported worldwide. Genotyping of 300 healthy individuals from Mexican population and compiled data from different ethnicities showed discordant results concerning the hypothesis that SCA disease alleles arise by expansion of large normal alleles.

Conflict of interest

The authors declare that they have no conflicts of interest.
Autosomal dominant spinocerebellar ataxias (SCA) include several neurodegenerative disorders characterized by progressive cerebellar dysfunction (1, 2). Over 30 distinct SCA types have been described (2) and at least 8 of these are caused by a cytosine-adenine-guanine (CAG) trinucleotide repeat expansion in the coding region of the respective genes, resulting in an expanded glutamine repeat (2, 3).

The general prevalence of SCA is 5–7/100,000 inhabitants (4, 5); however, their incidence may vary widely within different ethnicities or geographical regions. Studies reported worldwide show that SCA3 is the most common ataxia, followed by SCA2 and SCA6. Studies conducted in Latin American populations reveal that SCA3 is the most frequent SCA in Brazil (6–8), whereas SCA2 exhibits the highest incidence in Cuba (5, 9) and Mexico (10).

With the aim of developing the study of SCA in Mexico, we set up a method based on a fluorescent multiplex polymerase chain reaction (PCR) to test simultaneously for SCA1, SCA2, SCA3, SCA6, and SCA7 in a single assay (11). We diagnosed 64 patients with SCA from Veracruz State and determined the distribution of CAG repeat alleles at these five SCA loci in Mexican general population.

Materials and methods

Subjects

A case-finding pilot study was conducted for >2 years (2010–2011) by the Rehabilitation and Special Education Center of Veracruz (CREEVER-DIF) and the National Institute of Rehabilitation (INR) to identify individuals affected with late-onset familial cerebellar ataxia in five different communities of Veracruz State (Fig. 1). A number of resources were used for the study, including (i) a medical staff composed of neurologists, geneticists, molecular biologists and physical therapists transferred from the INR to Veracruz to perform clinical evaluation of SCA patients and to collect blood samples from SCA patients and their relatives in their communities themselves; (ii) patients identified in the pilot study were contacted for prospective evaluation; and (iii) all neurologists practicing in the studied region’s private sector were contacted requesting referrals of patients with SCA to our mobile medical unit. Clinical evolution of patients is described in Supporting Information Appendix S1. Genotyping of patients with SCA and control individuals, as well as as statistical analysis of the data are described in Appendix S1.

Results

Genetic and epidemiological analysis of patients with SCA

We estimated the occurrence of SCA1, SCA2, SCA3, SCA6, and SCA7 in five different communities of Veracruz State (Fig. 1) that were previously identified as having high prevalence of SCA by a case-finding pilot study. A total of 10 families with SCA were collected, including 64 affected individuals and 75 non-affected relatives. We identified eight families with SCA7, comprising 55 affected individuals and 17 asymptomatic individuals, and two families with SCA2, comprising nine affected individuals (Fig. 2a,b). No family with mutation in the SCA1, SCA3, or SCA6 gene was identified.

Distribution of patients with SCA and SCA subtype prevalence rate in the different communities are shown in Table 1. Our data indicate that SCA7 is the most common SCA, with general prevalence of 10.63/100,000 inhabitants, exceedingly higher than worldwide prevalence (<1/100,000) (12, 13). Distribution of SCA7 expanded alleles is shown in Fig. 3a.

Clinical features of patients with SCA

Patients with SCA7 have a mean age at onset of 32.74 ± 17.87 years, similar to that reported for other populations (14, 15). Both SCA7 and SCA2 patients exhibited the typical symptomatology of their respective disease (16, 17). It is noteworthy that the SCA2 patient bearing a short CAG track (32 repeats) (Fig. 2b, family #1) exhibited a mild phenotype with decrease in saccadic eye movement as the sole symptom, whereas his offspring displayed a more severe phenotype.

Distribution of CAG repeats at SCA1, SCA2, SCA3, SCA6, and SCA7 loci in healthy Mexican individuals

Frequencies of the trinucleotide repeat alleles in different SCA genes were analyzed in 300 healthy subjects.
Analysis of CAG repeats in five SCA loci in Mexican population

Fig. 1. Geographic origin of Mexican patients with spinocerebellar ataxia (SCA). Patients with SCA were recruited from five different communities of Veracruz State. These communities encompass a small region of 1200 km².

Fig. 2. Pedigrees of (a) SCA7 and (b) SCA2 families. Numbers below the symbols indicate repeat number at the SCA locus. Filled symbols indicate clinically affected individuals and symbols with a central point mark indicate asymptomatic carriers.

from Mexican general population and in 70 healthy individuals from Tlaltetela (Fig. 3b). Intriguingly, a healthy subject carrying an allele in the pathological range (39 repeats) was found; however, further DNA sequencing revealed the presence of two CAT interruptions within the CAG repeat configuration of the 39 repeat alleles (data not shown), which might explain the lack of symptomatology. Allelic frequency of SCA1, SCA2, SCA3, and SCA6 loci showed no significant deviations from the Hardy and Weinberg
Fig. 3. Allelic distribution of CAG (cytosine-adenine-guanine) repeats of various SCA loci in Mexican population. (a) Distribution of expanded CAG repeat alleles in 55 patients with SCA7 and 17 asymptomatic relatives. Expanded alleles ranged from 34 to 72 CAG repeats. (b) Distribution of CAG repeat alleles at various SCA loci in 300 healthy subjects from Mexican general population and 70 healthy subjects from Tlaltetela population. Genotyping was performed as described in Appendix S1. X axis represents CAG repeat alleles.

<table>
<thead>
<tr>
<th>Community</th>
<th>Inhabitants</th>
<th>SCA2 patients</th>
<th>Prevalence ratea</th>
<th>SCA7 patients</th>
<th>Prevalence ratea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xalapa City</td>
<td>457,928</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0.66</td>
</tr>
<tr>
<td>Tlaltetela</td>
<td>4528</td>
<td>6</td>
<td>132.51</td>
<td>37</td>
<td>817.14</td>
</tr>
<tr>
<td>Tuzamapan</td>
<td>6824</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>161.2</td>
</tr>
<tr>
<td>Cosautlán</td>
<td>4429</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>90.31</td>
</tr>
<tr>
<td>Xico</td>
<td>35,188</td>
<td>3</td>
<td>8.53</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>508,897</td>
<td>9</td>
<td>1.74</td>
<td>55</td>
<td>10.63</td>
</tr>
</tbody>
</table>

Table 1. Distribution of patients with spinocerebellar ataxia (SCA) and prevalence rates in the Veracruz communities

a[(Number of cases) (100,000 inhabitants)]/total population of inhabitants for each community (2010–2012).

The relationship between the presence of large normal alleles and the relative frequency of the different SCA was explored in Mexican population and compared with other ethnicities (Table 2). Interestingly, despite having a relatively high percentage of large normal alleles, no SCA1 case has been found in either Mexican or Brazilian populations (18); contrariwise, a correlation between the rate of large normal alleles and SCA1 frequency is observed in Australians and Caucasians (19, 20). Regarding SCA2, the occurrence of large normal alleles appears to correlate with the high frequency of this SCA in Mexico and Cuba (10, 21). On the other hand, strong correlation is apparent between the frequency of large normal alleles and the occurrence of SCA3 in Spanish, Indian, and Japanese populations, but not in Brazilian (19, 22–25). Regarding SCA6, the low frequency of large normal alleles appears to correlate with the absence of this SCA in Mexicans. Australian and Spanish populations...
Table 2. Frequency of large normal alleles and prevalence of five different spinocerebellar ataxias in different ethnicities

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>SCA1</th>
<th>SCA2</th>
<th>SCA3</th>
<th>SCA6</th>
<th>SCA7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Large ANs (%)</td>
<td>Frequency of cases</td>
<td>Large ANs (%)</td>
<td>Frequency of cases</td>
<td>Large ANs (%)</td>
</tr>
<tr>
<td>Mexican</td>
<td>27.5 (n = 600)</td>
<td>0</td>
<td>6.33 (n = 600)</td>
<td>45.4</td>
<td>2.38 (n = 600)</td>
</tr>
<tr>
<td>Cuban</td>
<td>nd</td>
<td>0</td>
<td>8 (n = 2695)</td>
<td>86.79</td>
<td>nd</td>
</tr>
<tr>
<td>Spaniard</td>
<td>19.66 (n = 294)</td>
<td>5.6</td>
<td>3.3 (n = 294)</td>
<td>15.3</td>
<td>7.22 (n = 294)</td>
</tr>
<tr>
<td>Indian</td>
<td>12 (n = 270)</td>
<td>32</td>
<td>4 (n = 270)</td>
<td>22.9</td>
<td>7 (n = 270)</td>
</tr>
<tr>
<td>Japanese</td>
<td>4 (n = 176)</td>
<td>3</td>
<td>1 (n = 359)</td>
<td>5</td>
<td>21 (n = 275)</td>
</tr>
<tr>
<td>Caucasian</td>
<td>16 (n = 574)</td>
<td>15</td>
<td>3 (n = 355)</td>
<td>14</td>
<td>9 (n = 641)</td>
</tr>
<tr>
<td>Australian</td>
<td>38 (n = 1295)</td>
<td>30</td>
<td>2 (n = 889)</td>
<td>15</td>
<td>5 (n = 1135)</td>
</tr>
<tr>
<td>Brazilian</td>
<td>14.24 (n = 308)</td>
<td>0</td>
<td>2.79 (n = 308)</td>
<td>5.22</td>
<td>1.75 (n = 308)</td>
</tr>
</tbody>
</table>

nd, not documented.

aNumber of chromosomes analyzed for each population.

bNumber of cases with an ataxia type specific × 100/total ataxia cases.
Magaña et al.

exhibited high percentages of SCA6 large normal alleles; however, while the former displays high frequency of SCA6 (20), the latter has low frequency of SCA6 cases (22). SCA7 analysis was precluded because of scarce genetic studies reported for this SCA.

Discussion

We describe CAG repeat analysis at SCA1, SCA2, SCA3, SCA6, and SCA7 loci in a Mexican population sample of 64 patients and 75 non-affected relatives belonging to 10 families with SCA recruited from five communities of Veracruz State, 300 unaffected subjects from Mexican general population, and 70 healthy individuals from Tlaltetela community. We identified 55 patients with SCA7 and 9 with SCA2 representing, to our knowledge, one of the largest series of SCA7 worldwide. The clinical features found in both SCA7 and SCA2 patients were consistent with their well-recognized symptomatology. We conclude that clinical examination concomitantly with multiplex PCR-based testing provides an accurate diagnosis for differentiating among SCA.

In other populations, SCA7 is one of the rarest forms of SCA (12, 13). In fact, our data revealing a relatively high prevalence of SCA7 (85.94%) are in sharp contrast with a previous study of Mexican population depicting SCA7 with low occurrence (7.4%) and that SCA2, SCA10 and SCA3 are the most common subtypes in Mexico (10). Moreover, these authors reported 18.5% of patients with non-identified SCA, whereas this study had no undiagnosed patients. The reason for these differences is not clear; however, it might be because the former study analyzed a sample representative of the whole Mexico territory, whereas this study was focused on the central Veracruz State region. The communities analyzed, founded by Spanish landowners and populated by Native Americans during the XVIII century, were settled in a region surrounded by a chain of mountain and hills, which has undermined their communication with other populations and limited human migration. Thus, it is likely that an ancestral mutation spreads over Veracruz communities because of their geographical isolation as well as cultural causes, such as consanguineous marriages. We believe that this geographic cluster of patients with SCA7 may be explained by a founder effect, as previously reported for Swedish and Finnish populations (14). Stevanin et al. (26) revealed multiple origins of SCA7 in 41 families from different ethnic backgrounds. The authors described four major haplotypes for the centromeric SCA7 region distributed in four distinct geographical areas, indicating regional founder effects. Further haplotype analysis of SCA7 in Veracruz communities is warranted to determine the mutation’s origin.

The association between prevalence of different SCA and frequency of their respective large normal alleles was found in some populations (19, 21, 23). We noted a certain correlation between the relatively high prevalence of SCA2 and SCA3 and the high frequencies of their respective large normal alleles in Mexican population. However, the data compiled from different ethnic populations revealed certain inconsistencies regarding this hypothesis, such as high percentage of large normal alleles with low prevalence or absence of a particular SCA (e.g. SCA6 in Spanish population) (20–22). Discrepancies across populations could be due to variability in the number of individuals analyzed. Interestingly, intermediate-length CAG repeat expansions in the ATXN2 gene (>28 repeats) were associated with an increased risk for amyotrophic lateral sclerosis (ALS) (27, 28), and raised the possibility that the relatively high proportion of SCA2 large normal alleles present in Mexico would be reflected in ALS frequency. Although some clinical studies on Mexican ALS patients have been published (29, 30), further epidemiological studies are required to define ALS prevalence in Mexico.

In summary, we genotyped 55 patients with SCA7 and 9 patients with SCA2, and provided, to our knowledge, the first report on CAG repeat allele distribution at five different SCA loci in normal Mexican population. The relatively high frequency of SCA7 in Veracruz might be the result of a founder premutation.

Supporting Information

The following Supporting information is available for this article:
Appendix S1. Supplemental materials and methods.
Additional Supporting information may be found in the online version of this article.

Acknowledgements

This work was supported by ICyTDF, grant number PICSA12-162, to J.J.-M. This article is dedicated to the patients and the members of SCA-affected families. We thank Lic. Adrian Morales, Dr. Hugo Azamar, and Lic. Juan Antonio Nemi for their facilities and for help with the logistics of sample collection.

References

Analysis of CAG repeats in five SCA loci in Mexican population