Reviewer's report

Title: High frequency of Machado-Joseph disease identified in Southeastern Chinese kindreds with spinocerebellar ataxia

Version: 1 Date: 10 September 2009

Reviewer: alfredo brusco

Reviewer's report:

The paper by Gan et al., describes the screening of SCA3 in Southeastern Chinese patients with cerebellar ataxia. SCA3 is found at higher prevalence in this region, than previously reported.

The paper makes an interesting point about SCA3 in China, however considering the state of the art of Spinocerebellar ataxia (SCA) genetics, and the fact that the poliglutamine expanded genes were described more than 10 years ago, I would have expected the screening of the most known genes (SCA1, SCA2, and maybe SCA6 and SCA7) together with SCA3. This would add a more complete picture of poly-Gln SCAs in Southeastern China. The results here reported are interesting, although SCA3 is already known to be a major SCA gene in China.

Furthermore some important technical problems need to be addressed.

Major points:

1) The calculation of glutamines (CAGs) in the ATXN3 gene is fundamental for the aim of this work. The sequence in figure 4A shows a 10 repeat allele, but it is reported as a 14 CAGs. This error needs to be checked, and potentially may change percentages. The expanded allele in Fig.4B is again referred as 81, but I can only count 76-77 repeats. Actually the end of the sequence does not allow a precise estimate of the repeats, likely due to somatic mosaicism, and would need cloning.

2) Introduction: The number of CAGs in ATXN3 for normal, mutable normal, reduced penetrance and full penetrance alleles are well summarized in Geneclinics. I would refer to this site (and the references therein) as an updated review of the literature. See: http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=gene&part=sca3

Normal alleles should be defined as #27 and #31. Large normal allele are >31 and <44.

Allele distribution found in SCA patients would be helpful.

Further points:

1) abstract: line 4: the sentence: "The efficiency...reported previously" should be changed to better explain the concept that a technical problem may explain why SCA3 expansions were overlooked by others.
2) line 9: “138 unrelated probands”

3) Throughout the text: please approximate percentages to no more than 3 figures (e.g., 72.46% becomes 72.5%).

4) The conclusion of the authors that large normal alleles may explain the high frequency of SCA3 expansions is plausible, but does not explain why they found a higher frequency than other authors.

   The technical artefact, due to a more efficient PCR, a different population selected, or more stringent clinical criteria are to take into account.

6) Introduction and text. It is difficult to sustain that other groups have found lower SCA3 expansions because of a mistake/technical problems. This may be suggested in discussion, or sustained by proofs.

7) Introduction page 4 line 3. “This related frequency”: is it “relative frequency”? Check throughout the text.

8) Page 5 line 3. indicate V/cm. Line 5: “the transport ratio of the DNA marker”; what is it?

   Line 6: the equation is not explained; what is Y, a, X, and b. Is it really necessary to specify all this?

   Line 9. some normal alleles: “some” is not acceptable. In figure 4 and results it seems one.

9) Table 1. Is the first column indicating the number of CAGs? I don’t understand the need of an entire table. Number of subjects (not reported) are so small that it would be helpful to discuss the category 28-31 CAGs, instead of the single CAGs.

10) Table 2. Erase chi square column and leave the P. Indicate p<0.001, for instance when below that value.

11) It is not necessary to indicate chi square and P values. It is clear that here the differences are due to different ethnic background. Please approximate numbers.

Page numbering is missing. Figure numbers are microscopic.

   Figure 2: indicate number of alleles on the Y axis.

   Figure 5 is not useful and very difficult to see.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**
I declare that I have no competing interests