Reviewer's report

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

Version: 2 Date: 13 November 2009

Reviewer: Conceição Bettencourt

Reviewer's report:

I recognize the effort of the authors to improve their manuscript. However, there are still some issues that should be further worked.

- Major essential revisions

1. In the Discussion and Conclusion sections, the authors should be more careful in the way they present their conclusions:

   a) As I said previously, the results of the present study raise the hypothesis that large normal alleles may constitute a reservoir from which the expanded alleles may be emerging. The authors stated that “the high frequency of MJD may be attributed to the high frequency of large ANs in the present study”. Since the authors did not find evidence of de novo mutations, they should be more cautious and say, for example, “The high frequency of large ANs associated with such a high relative frequency of MJD suggests that in this population large ANs may constitute a reservoir from which the expanded alleles may be emerging”.

   b) It is difficult to believe that lower relative frequencies of MJD found by other groups working with SCA patients were only due to technical problems in PCR, related to the use of Taq polymerase, especially when groups studying Portuguese and Brazilian patients found higher frequencies. Furthermore, caution is needed to avoid false negatives in a molecular diagnosis, and any case of homozigosity for normal alleles should be confirmed by other methods. Even using La Taq polymerase, 6 cases of homozigosity were identified by the authors and, as they recognize, these still need to be confirmed. Despite the higher efficiency of La Taq polymerase in amplifying the expanded alleles, and consequently in diminishing the number of samples that would require further analysis, the authors should not question the results of all other groups that used Taq polymerase just based on this. The present results simply allow the authors to conclude that the use of La Taq has advantages over Taq polymerase in this type of analysis, by amplifying the expanded alleles easily without the need of further PCR optimization or the application of other techniques such as Southern blot. Therefore, the text in the Discussion and Conclusion sections should be revised accordingly. For example, instead of stating “We have reported the characteristics of CAG repeats in ATXN3 based on a relatively large normal Chinese population and the highest relative frequency of MJD observed in Asian SCA patients. This may be attributed to the high frequency of large normal alleles and a more efficient PCR using LA Taq polymerase.”, the authors could be more
prudent saying something like “We have reported the characteristics of CAG repeats in ATXN3 based on a relatively large normal Chinese population and the highest relative frequency of MJD observed in Asian SCA patients. The results support the hypothesis that in this population large ANs may constitute a reservoir from which the expanded alleles may be emerging. Furthermore, LA Taq polymerase was proven to be more efficient than Taq polymerase in the amplification of the expanded alleles, facilitating and improving the molecular diagnosis.”

2. Although the authors added some text to the Abstract and Background sections to clarify the work hypothesis, the main goals of this study are still not obvious when reading those sections.

- Minor essential revisions
3. The English should be revised.

4. In the Abstract section, the authors wrote “…denaturing polyacrylamide gel electrophoresis and DNA sequencing in 150 normal individuals and 138 unrelated probands…” which is not completely correct, since they did not sequence all the samples. The authors should change the sentence in order to clarify this issue.

5. In the Molecular analysis item of Methods, where it is written “The DL2000 … were used as DNA size markers” it should be “The DL2000 … was used as DNA size marker”. Furthermore, in the legend of Figure 5 it is written “…M: D2000 marker”, but shouldn’t it be “…M: DL2000 marker”?

6. In the Statistical analysis of Methods, the first sentence should be changed, for example, to “All the statistical analyses were performed using SPSS software version 11.0 (SPSS...)”.

7. In the Analysis of CAG repeats in normal individuals of Results, where it is written “The difference in the frequency of large ANs between present study and other studies…” it should be “The difference in the frequency of large ANs between the present study and other studies...”.

8. In the Discussion section, what do the authors intend to say when they wrote “proble”? Would it be “probable”?

- Discretionary revisions
9. In the Results section, the authors present the mean, median, variance and skewness of the normal alleles of normal individuals. However, this kind of characterization is missing for normal and expanded alleles of MJD patients and this could be important to better visualize their distribution.

**Level of interest:** An article of importance in its field

**Quality of written English:** Needs some language corrections before being
published

**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