Author’s response to reviews

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

Authors:

Shi-Rui Gan Dr. (ganshirui1981@163.com)
Jian-Jun Wu Dr. (jungliw@gmail.com)
Ning Wang Prof. (nwang63@yahoo.com.cn)
Sheng-Sheng Shi Dr. (shishengsheng@gmail.com)
Sheng-Tong Wen Dr. (wst3344@126.com)
Shen-Xing Murong Prof. (zhiyingwucn@yahoo.com.cn)
Chuan-Zhen Lu Prof. (luchuanzhen@yahoo.com.cn)
Zhi-Ying Wu Prof. (zhiyingwu67@yahoo.com)

Version: 2 Date: 22 October 2009

Author’s response to reviews: see over
A point-by-point response to the concerns

In response to the reviewer Alfredo Brusco:

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.

The initial region of CAG repeats of ATXN3 gene has the sequence of (CAG)2 CAA AAG CAG CAA. We included the AAG and two CAA variant triplets in determining the number of CAG repeats just as suggested by Seneca et al (Eur J Hum Genet, 2008, 16:913-920) and as most other articles involving in calculation of number of CAG repeats in ATXN3 gene, though some articles (Schöls, et al. Hum Mol Genet, 1995, 4: 1001-1005) excluded variants of AAG and CAA.

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

Thanks for your concern. We have revised the number range in “intermediate alleles” and in “Abnormal allele with full penetrance” just as suggested in the site (http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=gene&part=sca3) in the background, and adopt the site as reference 5.

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.

Thanks again. We have changed the sentence to “The efficiency of amplification for CAG repeats is affected by the GC content of the amplified sequence when Taq polymerase was used in the polymerase chain reaction (PCR), however, the Taq polymerase has been used for PCR in nearly all studies reported previously”.

2) line 9: “138 unrelated probands”

We have corrected it.

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

We have corrected them throughout the text.

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.

The reason for a higher frequency found in the present study than other authors is due to a more efficient PCR using LA Taq polymerase and we have added this content in the Conclusion of Abstract and the last paragraph of Discussion.

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.

Base on your suggestion, we have added some contents in fourth paragraph from the Discussion section to show that using Taq polymerase in PCR was probable to lead to false negative in molecular testing of SCA3.

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

We have corrected all of them 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.

“the transport ratio of the DNA marker” means the distance of all fragments of the DNA marker migrated in the polyacrylamide gel.

“Y” indicated the size of fragments of PCR products, “X” indicated the transport ratio of fragments of PCR products, “a” and “b” indicated coefficient and exponent that all generated by the statistical package, respectively. For readers’ comprehension of the equation, we think it is necessary to specify all this, and thank you for your reminding.

We have added the specific number of the normal alleles sequenced in “Molecular analysis” from the Method section.

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.

Yes, the first column indicated the number of CAGs, and we have changed “copy number of CAG repeats” to “number of CAG repeats” for better understanding.

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

We have revised them just as your suggestion.

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.

We have revised as your suggestion. The number of alleles has been indicated on Figure 2. We have changed the Figure 5 into new one with color.
In response to the reviewer Conceição Bettencourt:

Major Compulsory Revisions

1. Although reading the paper we can infer the questions pointed by the authors, the work hypothesis should be clearly stated in the Abstract and in the Background sections.

Thanks for your advice, we have added some contents in the Abstract and Background sections based on your suggestion.

2. In the Methods section, the authors should specify the number of the sequenced samples. Furthermore, statistical analysis should be further detailed.

We have added the specific number of the sequenced samples in “Molecular analysis” and some statistical details in “Statistical analysis” from the Method section.

3. The results of the present study raise the hypothesis that the large normal alleles may constitute a reservoir from which the expanded allele may be emerging and the authors affirmed that “the high frequency of MJD may be attributed to the high frequency of large ANs in the present study”. Is there any evidence of de novo mutations that could support this statement?

We regret to say that there are no de novo mutations in present study. However, we will carry out the pedigree analysis in further study.

4. The results from the comparison between LA Taq and Taq polymerase are only mentioned in the Discussion section. They should come first in the Results section and the corresponding methodological aspects should be explained in Methods. Were the PCR reaction conditions the same? How many samples were used in the comparison? What was the estimated error for amplifications with Taq polymerase?

We have added some contents in “Molecular analysis” from the Method section, in “Analysis of MJD expanded alleles” from the Result section and in Discussion section based on your suggestion. The PCR reaction conditions were the same and the same 12 MJD patients were used in the comparison. Lower amplifying efficiency of Taq polymerase may due to the enzyme’s lack of proofreading activity as reported by Arezi et al (Anal Biochem, 2003, 321:226-235).

5. In the Discussion section, the authors wrote: “we suppose that Asian origin of MJD and founder effect may also contribute to the high frequency of MJD in the present study”. Could the frequency of the large normal alleles be also affected by a founder effect? What other factors could be influencing the distribution of wild-type alleles?

We presume that the frequency of the large normal alleles could be also affected by a founder effect. In addition, different population histories may also have had an effect on allele distributions (Andres et al, Hum Mutat. 2003, 21:61-70), and the distributions of the various numbers of CAG repeats are likely to be in a dynamic state depending on the mutation frequencies of the CAG repeats of the corresponding genes (Takano et al, Am J Hum Genet 1998, 63:1060-1066).
Minor Essential Revisions:

6. In the Background section, the authors state “The characteristics of CAG repeats in a large normal population including Acadian, African American, Caucasian, Greenland Inuit and Thai individuals were analyzed in 1996 [14].” Without disregarding the importance of the referenced work, there are several and more recent studies in this field, why the restriction to this reference from 1996?

Since the characteristics of CAG repeats of normal alleles was not the major point in present study, we delete the sentences of “The characteristics of CAG repeats in a large normal population including Acadian, African American, Caucasian, Greenland Inuit and Thai individuals were analyzed in 1996 [14]. However, we have not yet found any research on the characteristics of CAG repeats in a large Chinese mainland population though Jiang et al [15] analyzed the range of CAG repeats in 110 normal subjects of Han population in Northeastern China” when we were in revision of the Background section. But anyway, we would like to thank you for your concern.

7. In the end of the “Analysis of CAG repeats in normal individuals” from the Results section, the authors wrote: “In accordance with the previous report, … (>27 repeats) were defined as large ANs [21]. The frequencies of large ANs in the present study were calculated and compared to those of Japanese … and Thai [14].” In those two sentences, the authors are explaining methodological aspects, thus they should come in the Methods section, and in the Results the authors should only refer to table 1.

We have transferred the first sentence from the Method section to “Statistical analysis”. The second sentence has been changed to “The difference in the frequency of large ANs between present study and other studies involving in Japanese [9], Indian [10], Czech [11] populations and a combined population comprised of Acadian, Black, Caucasian, Inuit and Thai [21] is shown in table 1.” and still put in “Analysis of CAG repeats in normal individuals” on the Results section.

8. In the Results section, where it is “estimated to contain 14 CAG repeats is confirmed by sequencing” it should be “estimated to contain 14 CAG repeats was confirmed by sequencing”

We have corrected it.

9. In the Discussion section, where it is “which was also the most common allele found in the study of Limprasert et al [14] and Takano et al [21].” It should be “which was also the most common allele found in the studies of Limprasert et al [14] and Takano et al [21]”

We have corrected it.