Reviewer’s report

Title: A genomic copy number variant analysis implicates the MBD5 and HNRNPU genes in Chinese children with infantile spasms and expands the clinical spectrum of 2q23.1 deletion

Version: 1 Date: 5 March 2014

Reviewer: Weimin Bi

Reviewer’s report:

- Major Compulsory Revisions

1. The conclusion that the molecular mechanism leading to pathogenic CNVs in IS appears conserved among different ethnic backgrounds needs to be revised. For the four pathogenic CNVs, three of them actually haven’t been reported in individuals with primary diagnosis of IS as stated in the manuscript. However, it appears that the genes involved in IS and/or seizure, LIS1, HNRNPU and MBD5, are also the causative genes in Chinese IS patients. I suggest to say “Our study also supports that the molecular mechanism of infantile spasm appears conserved among different ethnic backgrounds”.

2. The duplication in Xp22.31 in S134 contains other genes in addition to the two genes, STS and VCX3A, mentioned in the manuscript. Duplication of Xp22.31 has been reported in multiple papers (Am J Med Genet A. 2012 Feb;158A(2):461-4; Eur J Med Genet. 2010 Mar-Apr;53(2):93-9; Hum Mol Genet. 2011 May 15;20(10):1975-88). There is a debate on whether this duplication is a cause of intellectual disability and/or developmental delay or a benign variant. In addition, the reference of Palka-Bayard-de-Volo et al 2012, is not an appropriate one for Xp22.3 deletion.

3. The case S67 has a duplication at 1q21.1. This duplication is flanked by large segmental duplications and it is a recurrent duplication that has been previously reported. The size of this duplication was smaller in Brunetti et al [23], which is due to different probe coverage in this region for the arrays used in these studies.

4. The phenotype of absent hallux in patient S162 was not previously reported in 2q23.1 deletion patients. However, MBD5 mutations/ deletions have been shown to cause craniofacial abnormalities and hand/foot abnormalities in almost all the patients. It is necessary to add this information in the discussion.

5. In the section of Methods, the sentence of “A DLR score above 0.20 was set as the cutoff criteria for false CNVs” is unclear. Do you mean you use a DLR score above 0.20 to indicate poor quality of array data.

6. The manuscript needs to be carefully checked throughout the text in order to accurately present the results. Examples are as below:
5a. In Figure 1C and Figure 2C, the brain MRI images seem to be switched between S100 who has a 17p13.3 deletion and lissencephaly and S37 who has 1q44 deletion and delayed myelination.

5b. In “Results”, section of “Genome-wide copy number variation analysis”, the CNVs were placed into 3 classes, not 4 classes.

5c. In “Results”, the proband S134 has a deletion in17p12 including PMP22. But the father carries the same duplication. Should the father carry a deletion instead of duplication as stated in the manuscript?

5d. In Figure 1 legend, the distal breakpoint is within intron 1 of LIS1, not intron 2.

- Minor Essential Revisions

1. Both 17p13.3 and 17p13 are used for the deletion in individual S100. 17p13.3 needs to be consistently used throughout the manuscript.

2. Both IS and ISS are used for Infantile spasms. IS should be used all the time.

3. In abstract, section of “results”, line 8 from bottom, “for” should be used instead of “form”.

4. In “Methods”, what is the temperature used for primer extension?

5. The section of “CNVs of unknown variants of clinical significance” should be “CNVs of unknown clinical significance”. In this section, “YAP1, a protein important” should be “YAP1, a protein important”.

6. In “Discussion”, the reference “Vijav et al” should be “Tiwari et al.” and this paper actually identified CNVs in 13 trio, not 11 IS cases.

7. Figure 2 legend, the gene name of HNRNPU was misspelled.

8. In “Author’s information”, “Children’s Fudan University” should be “Children’s Hospital of Fudan University”.

- Discretionary Revisions

1. Please add page number.

2. The Xq22.1 duplication in the male patient S34 is inherited from the mother. Checking for this duplication in maternal male relatives may help determine the clinical significance of this duplication.

3. Suppl. Table S1, there is one column on “Dev at Last Review”. It is necessary to indicate the patients’ ages for the last reviews.

4. In section of “Materials and Methods”, DNA from patients and controls was digested. It is unclear which controls were used for array CGH. It is better to provide website or citation on how to access the published database for the CNVs from ~1000 Han Chinese individuals.

5. Figures 1D and 2D is better to placed right behind Figures 1A and 2A, respectively, to be consistent with the order in Figure 3. The texts for the colored bars need to be consistent between Figure 1D and figure 2D.
**Level of interest:** An article whose findings are important to those with closely related research interests

**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