Reviewer's report

Title: Genotyping and Drug resistance patterns of M. tuberculosis strains in Pakistan

Version: 2 Date: 6 April 2008

Reviewer: Stefan Niemann

Reviewer's report:

Although revised, the paper of Tanveer et al. does not follow a clear strain classification scheme as already suggested in the first review. This renders analysis and interpretation of the data difficult and in parts useless. Several passages of the paper still lack clarity, flow and are partially not fully correct and difficult to understand.

Details are outlined below:

Major Compulsory Revisions:

1. The authors still stick with a comparison of spoligotypes with Spoldb4 and classified their strains in previously known types e.g. CAS1 or orphan types such as Pak 1-12. This can be done, but it should be clearly mentioned that both belong to the same phylogenetic lineage namely CAS. To prove this, the authors are encouraged to analyze the strains for presence or absence of lineage specific deletions e.g. RD750 in the case of CAS. The authors should organize their results part from bigger classification to smaller groups. First classify all strains in major phylogenetic lineages CAS, Haarlem, Dehli and name this clearly in Fig. 1. Then classify e.g. CAS strains in subtypes that can be shared e.g. CAS1 or orphan Pak1. The term genogroup should be replaced by a more appropriate term such as genotype of phylogenetic lineage.

2. Although not an expert in statistics, this reviewer is scared about the statistical analyses. First of all, for comparisons, strains should be grouped in valid phylogenetic lineages that should be defined by valid markers. In my opinion it makes no sense to compare CAS1 and Beijing (which are well defined) with other ill defined groups shared types and orphan types which represent a mixture of strains from different phylogenetic lineages. If such an analysis is desired, multilogistic regression procedures might be necessary. This should be checked by a statistician.

3. It is also necessary that the strain grouping in the different tables should be clearly visible to the reader. Again, the major clades should be included in Fig. 1. Each clade can be further distinguished in finer groupings.

4. The author need to discuss their extremely high MDR rate or skip the data. In the introduction, they mention a recent study reporting a rate of 2% MDR in
untreated patients only. If they can not provide more detailed patient information, at least thorough interpretation of the DR data should be presented in the discussion.

• Discretionary Revisions

• Minor Essential Revisions

1 Abstract. The abstract should be revised according to the revisions suggested above.

2 Page 5, lines 117-120. I don’t understand that sentence. Please clarify

3 Page 6, line 123 - 125. Change to CAS. What means Bejing and Beijign like?

4 Page 6, line 131 - 132. Add correct citation.

5 Page 6, line 131 – 132. You are not really working on molecular epidemiology. This would need markers with higher resolution. Chan to population structure. I would also not insist to much on drug resistance. This would require to present more patient data, at least the classification in untreated and previously treated in order to allow a useful interpretation of the data.

6 Page 6, line 136 – 138. Other deletions are much more useful for the purpose of the study e.g. RD 750 for CAS classification.

7 Results should be changed as suggested above. First classify all strains in major phylogenetic lineages CAS, Haarlem, Dehli and name this clearly in Fig. 1. Then classify e.g. CAS strains in subtypes that can be shared e.g. CAS1 or orphan Pak1. The term genogroup should be replaced by a more appropriate term such as genotype of phylogenetic lineage.

8 Deletion analysis. Well, it is already well known that TBD1 is deleted in all M. tuberculosis sensu stricto strains except EAI.

9 Pulmonary versus extra pulmonary. If presented, this analysis should be revised and done with valid grouping. This should also be checked by a statistician.

10 Drug resistance. See above.

11 Discussion. In general - change according to changes in results section.

12 Discussion. Please revise according to genotype classification. CAS strains (including CAS1, 2, Pak 1-2 etc) are dominant of which CAS1 is the most common strain type. Follow this line for the other genotypes. Please have in mind that for Beijing you have only the major phylogenetic lineage since you have no further discriminator power with spoligotyping.

13 Discussion. Pulmonary versus extra pulmonary. Change according to new analysis.

14 Discussion. Please provide a thorough discussion of your high MDR rate before discussion different resistance rates among genotypes.
Level of interest: An article of importance in its field

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

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:

I declare that I have no competing interests