Reviewer’s report

Title: A low clustering rate with predominance of Haarlem strains in patients without known risk factors for multi-drug resistant tuberculosis in North-Eastern Lima, Peru

Version: 1 Date: 27 February 2013

Reviewer: M. Cristina Gutierrez

Reviewer’s report:

A low clustering rate with predominance of Haarlem strains in patients without known risk factors for multi-drug resistant tuberculosis in North-Eastern Lima, Peru

BMC Infectious Diseases

This study focuses on a selected population of adults (>18 year-old) with new smear-positive pulmonary TB and without known risk factors for MDR-TB, living in a geographical area with high exposure to TB. The aim is to assess the burden of TB due to recent transmission, to identify predominant circulating TB lineages and to investigate the drug resistance patterns associated with these strains. These are important questions to be solved for understanding the TB epidemiologic patterns in this community and improving its control. The authors present original data in a properly structured and well written manuscript. The study has efficiently determined the population structure of the Mycobacterium tuberculosis isolated from this group of TB patients. However, this study design fails to surely assess the burden of TB due to recent transmission. The accuracy of resistance testing to rifampin remains to be confirmed.

Major Compulsory Revisions

1. My main concern is that the studied sample is not sufficient to infer the rate of recent transmission of TB within this population in North-Eastern Lima. In its present format the study could be leading to misinterpret that the rate of recent TB transmission for this population is low, in spite of living in an area of high exposure to and transmission of TB. The authors should suppress this subject from the objectives and conclusions, carefully review the comments in the discussion, and focus the manuscript on the population structure and resistance of MTB population. Specifically, a seven-month period is too short for a molecular epidemiologic study of TB transmission in the general population; the sample is not exhaustive, representing only 16% of TB cases notified in San Juan de Lurigancho, the studied area of North-Eastern Lima (which notifies more than 2130 TB cases per year); the sample is biased, only MTB strains from a selected group of TB patients within the general population are compared, whereas transmission could have occurred between the selected population and the rest of TB patients in the area - the suspects to have a MDRTB- <18 year-old
patients are excluded, whereas TB in youngest people is an indicator of active transmission. The need of future studies to accurately estimate the rate of recent TB transmission could be mention in the summary, together the identification of risk factors.

2. The authors indicate that they used 40ug/ml of rifampin to test on Middlebrook 7H10 medium the susceptibility of the MTB strains. However the standard final concentration to test this drug on 7H10 is 1.0 ug/ml. Which was the real concentration that they tested? If 40ug/ml, the interpretation of the results for this drug and their conclusions are not valid, and the rate of rifampin resistance and MDR could be underestimated.

3. Methods, Fingerprint analysis. Please review the first phrase, it is unclear the process followed to compare the patterns as only the spoligopatterns can be compared to SpolDB4.0. In the same line, it is unclear along the description of the results whether the authors determined the spoligotyping family for the strains using only MIRU-VNTRplus (as it is shown in Figure 2) or they also checked in SpolDB4.0 the strains that did not match the spoligotypes of the reference strains of MIRU-VNTRplus. The process to assign strains to one family should be described in detail in Methods.

4. It is unclear whether MIRU-VNTR typing was performed with all the 199 strains or only with strains clustered by spoligotyping, this point should be clearly indicated in the text. Additionally, the manuscript provides scarce information on MIRU-VNTR results, which should have been more exploited. E.g., were all MIRU-VNTR profiles congruent with the family classification by spoligotyping? A situation that makes the analysis more robust and can be verified using MIRU-VNTRplus.

5. Percents in the text need revision. For instances, in Results page 6, 50 patients is not 24.8% of 314. I also count 44 H-resistant strains in Table 1, this corresponds to 22.1 % instead of 14.4% indicated in the abstract.

6. It is indicated that LAM strains were less likely to be H resistant than the other lineages (0.14 vs. 0.19 in the abstract; 0.84 RR, 95% CI 0.19-2.56 in the results). This data seems to have not statistical significance. Table 3 shows lower RR for family U strains (0.43)? P values should be added to Table 3, and the lack (or not) of significance mentioned in the text.

7. The manuscript is in some instances ambiguous regarding the covered population. Along the entire text, the authors should make clear that suspected MDRTB patients were excluded, and conclusions concern only TB patients without risk factors for MDRTB. Idem for the structure of the MTB population and the predominant circulating families, which cannot be extrapolated to the entire population, particularly because without additional data it cannot be excluded that MDRTB cases could be due to a predominant strain.

Minor Essential Revisions
1. Methods, Culture and drug-susceptibility testing, please indicate MTB in full.

2. Methods, MIRU-VNTR. “MIRU-VNTR is a PCR-based typing method ... to be polymorphic in MTB”. Please review the phrase, in the present format it refers to only 15-locus MIRU-VNTR typing, whereas the method is independent of the number of loci tested.

3. Methods, Fingerprint analysis. Please note that double alleles at 2 or more MIRU-VNTR loci can be considered not only mixed infections but also possible cross-contaminations. Additionally, in the phrase “Identical spoligotypes and MIRU-VNTR patterns ... clustered a cluster” something is missing, “clustered in a cluster”?

4. Table 1, polyresistant (non-MDR) H+E is indicated twice.

5. Table 2, there seems to be an error in the RR (95% CI) for either Sex or Residence as they are identical.

6. Legend to Figure 2 indicates that the dendrogram includes spoligotyping and MIRU15 patterns for the 199 isolates, however the tree has been constructed only based on spoligotyping data, and the MIRU15 patterns are not shown.

7. Reference n. 24, the title is missing in Bibliography section.

Discretionary Revisions

1. Table 2: Characteristics and MDR status. I would recommend to classify the population into 2 groups, 30 year-old or less, and more than 30 year-old, and reanalyze according to MDR status by chi square test as it looks to be a tendency to MDR (+) status among younger people.

2. Figure 2 adds little information to the already provided in Figure 1 that is descriptive enough. It could be more useful a figure focusing only on the orphan spoligotypes because specific ST patterns can be easily checked in SpolDB4.0.

**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