Reviewer’s report

Title: Modified genome comparison method: a new approach for identification of specific targets in molecular diagnostic tests using Mycobacterium tuberculosis complex as an example

Version: 0 Date: 03 Aug 2018

Reviewer: Peter Keller

Reviewer's report:

The authors describe an in silico procedure for the selection of target sequences for a molecular diagnostic test for the detection of Mycobacterium tuberculosis complex. They evaluated their novel molecular assay with a dilution series and 230 clinical specimens.

The methods and results are very well described and can be understood very easily.

Major comments:

1) Missing control for the number of M. tuberculosis complex complete genomes that contain the 5KST-PCR target sequence: In the results section (lines 53-60, page 9), the authors describe that the target sequence - which is a part of Rv2821c, a non-essential hypothetical protein - has been found in 209 strains of M. tuberculosis deposited with the NCBI genome database. Since Rv2821c is a non-essential gene and nothing is known about its conservation in the worldwide M. tuberculosis complex, I recommend the following two analyses: a) BLASTN search of an essential target sequence (e.g. 16S rDNA [rrs] or rpoB) against the identical NCBI genome set that has been used by the authors for their analysis. Ideally, the number of genomes with occurrence of the 5KST-PCR target (i.e. 209 genomes) is identical with the number of genomes with the rrs or rpoB target sequence (i.e. it should also be 209 genomes). b) BLASTN search against a genome collection of bioinformatically well characterized world-wide collected M. tuberculosis complex strains, e.g. the 5310 M. tuberculosis complex genomes that have been included in the following study: Manson AL et al. Genomic analysis of globally diverse Mycobacterium tuberculosis strains provides insights into the emergence and spread of multidrug resistance. Nat Genet. 2017 Mar;49(3):395-402. doi: 10.1038/ng.3767. DOI: 10.1038/ng.3767. There, I would expect that some strains will be lacking Rv2821c. It would be interesting to know, how many strains/ lineages do not contain a copy of Rv2821c/ 5KST-target.
2) How do the authors assess, whether the PCR assay is inhibited for an individual clinical sample? This control is lacking in the methods and in the results section. The authors should discuss how PCR inhibition can be seen in their assay.

Minor comments:

3) In low-incidence countries, M. tuberculosis infections are often paucibacillary. I.e. there are only a few bacteria per µl of respiratory material. The authors reduce the theoretically achievable analytical sensitivity of their assay by taking only 1 µl DNA extract (page 7, line 35) for the argument of a lower PCR inhibition rate. They use a relatively crude, low-tech method to extract the DNA, which is reasonable in resource limited settings. In order to allow the reader to understand the complete workflow, the text should also explain the steps prior to DNA extraction to illustrate whether there are additional concentrating steps. The authors should state which volume of primary clinical material has been used. They should state whether the material has been decontaminated prior to DNA extraction and whether there was a centrifugation step prior to DNA extraction. Commercial systems for the detection of M. tuberculosis complex often use higher amounts of DNA extract (e.g. 25 µl in case of Roche TaqMan MTB assay; see http://tbevidence.org/documents/rescentre/sop/MTB%20TaqMan%20PI.pdf). There is recent work by Chakravorty S. et al. (DOI: 10.1128/mBio.00812-17), that argues, that sensitivity of the commercial Cepheid GeneXpert MTB assay could be increased by a work-over of the assay (MTB/RIF vs. MTB/RIF Ultra version of the test) using higher amounts of sample input and DNA extract. The importance of the specimen input volume and DNA extract volume should be discussed. In the results section, the rate of PCR inhibition should be indicated for the clinical specimen set.

4) Introduction: In enumeration of PCR targets for the detection of M. tuberculosis complex (page 4, lines 12-12); important targets that are used in commercial assays are missing: 16S rDNA gene (rrs) and rpoB should be added.
Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

Yes

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

No

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?
4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?
5. Do you have any other financial competing interests?
6. Do you have any non-financial competing interests in relation to this paper?
If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal.