Genes
Reviewer
To whom it may concern

Geneva, 21st August 2019

Re: Response to reviewer 2 on “Viral Diagnostics Using Next-Generation Sequencing”

Dear Reviewer,

We would like to warmly thank you for your time and insightful comments in improving our manuscript on “Viral metagenomics in the clinical realm: lessons learned from a Swiss-wide ring trial”. We reply below to all your comments.

Line 51: next generation sequencing (NGS) was already defined in line 47.
We updated the text accordingly, thank you.

Line 62: Change 'SIB Swiss Institute of Bioinformatics' to ‘Swiss Institute of Bioinformatics(SIB)’. The abbreviation is then used on line 63.
We updated the text accordingly, thank you.

Table 3: Does pipeline ‘I’ have a method for RNA extraction? The box appears to be missing on the table.
Pipeline “I” used easyMAG for both DNA and RNA extraction. We clarified this in Table 3 by adding a missing cell border.

Lines 247 to 256: I understand the decision to reduce read numbers proportional to the read length, from the point of view of genome coverage; however, if this has a detrimental impact on longer-read pipelines insofar as the actual number of reads per target is lower. If virus detection and identification is purely based on coverage data, then there is no issue. If mapped read numbers plays a role, could this cause a bias against the longer-read pipelines? Presumably this is accounted for in the depth measurement (mapped reads * read length / genome size). In preparing for this study did you run analyses (in any pipeline) using identical numbers of reads regardless of read length? If so, were the outcomes different to those reported for the normalized read numbers?

We thank you for this very interesting comment. No preparatory analyses were performed before the ring trial implementation and thus we do not have numbers on using identical numbers of reads for different read lengths. When preparing the datasets, we wanted to normalize in some way the number of reads to be fair between the participating laboratories. Indeed, having more reads is certainly an advantage for virus detection and identification, but that comes at a cost, resulting in a necessary trade-off to enable routine implementation in a clinical setting (we updated the text accordingly in lines 252-254).
However, in order to ensure fairness among participants, we decided to normalize the number of reads per virus to the read length, as having more reads is certainly an advantage for virus detection and identification, but comes at a cost, resulting in a necessary trade-off to enable routine implementation in a clinical setting.

As you mention, we also came to the conclusion that using coverage and average depth (measured as number of mapped reads * read length / genome size) as performance metrics would not ultimately penalize longer-reads pipelines. Also, as now shown in Table 3 that we updated, read length in our ring trial ranged from 100 to 150bp only. We had planned to also analyze artificial minION long read data (cf. Table S2 and Dataset S2), but unfortunately none of the participants actually chose to test those datasets.

**Line 330: What does the (1) at the end of the line refer to?**
The (1) refers to the formula number according to Genes Authors guidelines. We had however forgotten to cite it and now refer to it in the legend of Figure 2.

**Line 338: it is not evident why pipelines A and B were based on the same FASTQ dataset, as the sample processing methods differ (lack of RNA extraction and post extraction enrichment in pipeline B)?**
Indeed, the wording in the text is a mistake, A and B submitted different FASTQ datasets, although originating from the same sequencing center. Thank you again for the careful reading, we updated the text accordingly (lines 353-354):

- Indeed, for the 7 pipelines, there were actually 5 sets of submitted FASTQ datasets from 4 sequencing centers, as pipelines E, F and J were based on a common FASTQ dataset.

**Additionally, does 'For quantitect whole genome amplification of cDNA molecules' (Figure 3 pipeline A) refer to the QuantiTect whole transcriptome kit?**
Indeed, the wording was not precise; we have updated Table 3 accordingly, thank you for noting this.

**The results section includes comment on the impact of sample processing (including potential points and levels of contamination) and bioinformatics pipelines (false positive and negative calling), however there doesn't seem to be much mention of these factors in the discussion. Do the authors draw any overall conclusions regarding best practices for sample preparation/sequencing and/or details of bioinformatics pipelines? Or is the intent not to direct the practices of individual testing centers directly, but to define output quality guidelines regardless of the methods used to achieve report outcomes.**

We very much thank the reviewer for this comment. When setting up the working group and implementing this trial, our idea was to view “NGS practices harmonization” as a means to achieve comparable high quality results, irrespective of the methodology chosen by each laboratory. Indeed, clinical laboratories need to integrate NGS as part of their existing practice, and may therefore have constrains e.g. on the choice of the nucleic acids extraction processes already in place for other existing analyses also requiring this step. We clarified this at the beginning of the Discussion as follows (lines 530-540):

- “The SIB working group on “NGS Microbes Typing and Characterization” and the ring trial that we implemented aimed at harmonizing NGS practices in clinical viral metagenomics. Harmonizing
means here for different clinical laboratories to achieve comparable results of high quality, irrespective of the methodology chosen at the experimental or bioinformatics levels, which may result in part from internal constrains as well (e.g. existing laboratory processes for nucleic acid extraction to be re-used as much as possible for metagenomics as well). Our results have highlighted various aspects that can impact overall performance at the experimental, databases and bioinformatics levels, offering insights to each participating laboratory into processes where they may further improve their workflows. In order for others to also benefit from this study and test their own workflows, we have published with this manuscript the database and all the datasets that were generated (Datasets S1 and S2)."

We remain at your disposal should you have any further questions or comments.
Thank you very much and kind regards,

Dr. Aitana Lebrand
Senior Project Manager
Clinical Bioinformatics