the WINNOWER

Pre-publication review of: Reagent and laboratory contamination can critically impact sequence-based microbiome analyses

JENNA LANG

ABSTRACT

This is a pre-publication review of the manuscript entitled: "Reagent and laboratory contamination can critically impact sequence-based microbiome analyses" by Salter *et. al*, 2014 http://www.biomedcentral.com/1741-7007/12/87

1. Is the question posed by the authors well defined?
Yes.
2. Are the methods appropriate and well described?
For the most part.

My main (minor) concern is that the input DNA was characterized as a pre-DNA extraction cell count. I assume that the different DNA extraction kits have different recovery efficiencies, and it would be nice to know whether that played any role in the kit-to-kit comparisons. But more importantly, it is extraordinarily rare for researchers to estimate this number before working with an environmental sample. Typically, one would measure DNA concentration post- DNA extraction. In order to increase the *value and impact* of this work, an attempt to provide guidelines that relate to standard practices would be nice. For example, I am not likely to proceed with 16S PCR if my input DNA is below the level of detection of my Qbit. That's ~ 0.2ng/uL. Which of your serial dilutions does this correspond to? I could calculate that based on the genome size of *S. bongori*, but I have no idea how much *S. bongori* DNA was actually extracted from your dilutions. Of course, for most of my samples, even if I have an input DNA concentration of 100ng/uL, I don't typically know how much of that is bacterial DNA. So, are you suggesting that everyone incorporate a bacterial qPCR step before every 16S PCR reaction? If so, then that (while onerous) would be more useful advice than suggesting that everyone make sure they are putting in at least 10^4 cells (as per Box 1).

A couple of other minor points:

1.   1) The authors suggest that replicates should be carried out using different reagents and batches. It might also be worthwhile to carry out serial dilutions of at least a few of the samples that you process with each kit/batch of reagents.

the WINNOWER

PRE-PUBLICATION REVIEW OF: REAGENT AND LABORATORY CONTAMINATION CAN CRITICALLY
IMPACT SEQUENCE-BASED MICROBIOME ANALYSES : **BIOLOGICAL SCIENCES**

2. 2) While a heavy-handed complete removal of every taxon that appears in the kit control is not likely to be appropriate in every case, some bioinformatic solution might be developed. Nick suggested a few ideas in the comments on the preprint, which I think would be worth including in a revision, even if as a loose bit of "thinking out loud." You might also take a look at the approach implemented here: https://peerj.com/articles/447/#supp-2

3. Are the data sound?
Yes.

4. Do the title and abstract accurately convey what has been found?
Yes.

5. Is the writing acceptable?
The British spelling of certain wourds makes the reading a bit awkward for me, as does the (ironic?) choice not to use the Oxford comma. But these are purely stylistic concerns.

6. Does the manuscript adhere to the relevant standards for reporting and data deposition?
Not yet. There is no mention of data deposition in the manuscript! I'm hopeful that all of the sequence data generated for this study will be made freely available, as it could be very useful to the community, e.g., to someone who wanted to build a reference database of common kit contaminants, or design a bioinformatic tool to flag potential contaminants.

7. Are limitations of the work clearly stated?
Sure.

8. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?
Yes.

9. Are the discussion and conclusions well balanced and adequately supported by the data?
Yes.

10. If additional data are required, are the requested data needed to support
a) the main point of the paper
b) only part of the conclusions (please specify) a point that is not essential to the main conclusion(s) of the paper (please specify)
See #2 for additional data requested (none essential.)

-Jenna M Lang