Response to Reviewers:

Thank you for the constructive criticisms. In response to these reviews, we have undertaken a near complete re-writing and re-organization of the manuscript. Where possible, we have responded to individual suggestions, but in many cases entire swaths of comments are addressed and revised.

General

This article by Gloor, Links, and Reid discusses the handling of microbiome sequencing data using compositional analysis. The authors provide two examples of data handling workflow (bacterial vaginosis data with two different treatment methods, Hsiao et al 2013 data). The authors attempt to show that the conclusions drawn by Hsiao et al. 2013 are likely false due the way they handled and interpreted their sequencing data.

There appears to be two purposes to the paper. One, a step-by-step walkthrough of how to analyze sequencing data, and two, an examination of published data and how using incorrect techniques can lead to flawed conclusions. These are both valuable messages, but these concepts need to be clearly defined in-text, and if included in the same manuscript, must be better melded together in the paper.

Lastly, the article requires significant proofing for grammar, readability, stylistic/formatting errors, and internal referencing errors. The figures are not of sufficient quality for publication and require revision for readability and formatting. The article contains sub-headings (most are bolded, but not all), but the authors need to clearly define sections for methods, results, and discussion. Overall, the paper does not flow very well and is a bit hard to read, and can be improved by improving cohesiveness and brevity in writing.

This work has the potential to be particularly valuable to the scientific community, and I would strongly encourage the authors to revise and resubmit.

These individual comments are largely addressed, but were done so en masse.

LN28-32: Break up into multiple sentences

LN48: Microbiology was still a very active discipline before human microbiome studies

LN57-61: Awkward phrasing.

L65: hereby is unnecessary (conciseness)

LN73: Filler (such is the nature of field), generalization.

LN77-80: Broad statement on the sample collection, can the author’s cite a study or review related to optimization of collection practices?

(DONE line 97)

LN84: is normally collected

LN86: or whatever, informal.

LN92: Conciseness (It should be recognized that)

LN94: there s/b the

LN97: semicolon should be used

LN100: if s/b of

LN111: appear italics

LN113: italics: negative correlation bias, did they authors intend to include a glossary? Unsure as to why certain phrases are in italics throughout document.

(this has been expanded considerably lines 116-159)

LN115: and others – vague, define everything or conclude the list of analysis.

LN116: no fourth analysis goal listed above.

LN126 and 129: centred, centered

LN137: conciseness: it is here that we have to remember

LN138: italics

LN143: with very 0 counts

LN144: , s/b .

LN156: a worked example?

LN158-159: behind the scenes, colloquial (Fixed line 274)

LN161: Define placebo as a placebo probiotic supplement or reverse order of two treatments (confusing to read in that order) . Add (BV) after bacterial vaginosis (not defined for abbreviation)

LN170: delete workhorse (Done line 286)

LN173: grammar, properly made and interpreted they

LN178:filler, it should be noted that

LN227: See , UniFrac

LN261: indent?

LN314-317: awkward

LN321: Sample processing is not really discussed, only data handling after sequencing (No, there are many other papers on this and we cited Walker 2015 as a good example case for caveats)

LN361-362: references are not numbered (fixed)

LN470: filed s/b field

LN481-482: much encouraging?

OTU is never defined (operational taxonomic unit) (Changed to taxa throughout)

Table has no title.

Figures:

Figures need properly labeled axes. Figures should be self-contained with legends so that they are understood without referring to the text.Figure 1, panels should be labeled (there is no Figure 1C LN133)

Figure 1 has been redone

Figure 2 has no title. Labels on Figure 2 are obscured and not legible. Characters are cut off on figures due to improper sizing (fig 3).

Figure 2 now has headings and the scale has been changed. We apologize for the overlapping labels, but this cannot be changed in the biplot coding since it is meant to show the exact relationships between datapoints. This is the reason that color was added and the code to generate the figure supplied in the supplement.

Figure 5 and 6 are not cited in-text and it is not well explained what is being examined from within the dataset.

Figure 6 has been dropped, and Figure 5 is now referenced. The data and analysis should be explained within text so that the reader can understand the comparison without having read Hsiao et al. 2013.

A better introduction to the paper is provided.

Reviewer: 2

We have now rewritten this in more of a review format with a list of software that can be accessed and two examples (and code) that the reader can use.

General

I am confused about what type of paper you are writing. Seems to be just a report from a workshop. Some conclusions seem very amateurish-style, e.g. " For a step by step guide a book outlining how to use the compositions R package by Van den Boogaart,and Tolosana-Delgado (2013) is particularly helpful, For others wishing to understand bioinformatics and data analysis of sequencing data in general terms, hopefully this paper is helpful, and encouragement will be given for members of their research group to enroll in specialized courses"·

Concluding that one possible value of this paper is that people will be willing to take a course in bioinfo or recommending a book is basically saying that there is little value in what is presented.

I hope you can turn this manuscript into a serious review with a meaningful final message for the reader.