Dear Dr. Ionescu,

Thank you for our quick evaluation of our revised manuscript. We have made both corrections you requested (see below).

We also made several other minor corrections after performing a final scan of our manuscript. These changes are noted below and are all very minor with the exception of a typo where in one case the tool Piphillin was incorrectly referred to as PAPRICA.

We did add the conflict of interest statement in our preprint in the last version, which you can see before the reference section.

Neither of us are PCI recommenders so we do not believe any disclosure is needed along those lines (at least one of us, GMD, would be very eager to become a recommender though!).

We should also mention that we did not receive an email from PCI Genomics and instead simply noticed the update on the website – just mentioning this in case it is a problem for others at the moment too.

-   Gavin Douglas and Morgan Langille

Minor corrections (requested by recommender):

1) We have added "Second," and removed "also" on lines 83-84.

2) We have removed "(hereafter 16S)" on line 1164.

Additional minor corrections:

-   Box 1 – changed "Many changes in the highly conserved region of the 16S rRNA gene" to be "Many changes in the highly conserved region**s** of the 16S rRNA gene"

-   Line 338 – Changed "The V4-V5 region was recently shown" to "**For example,** the V4-V5 region was recently shown".

-   Line 413 – Changed "In contrast, we found that open-reference operational taxonomic unit clustering resulted in a high rate of spurious ~~operational taxonomic units~~ compared to these methods" to "In contrast, we found that open-

reference operational taxonomic unit clustering resulted in a high rate of spurious ==identifications== compared to these methods"

- Line 452 – Changed: "when building phylogenetic trees representing a single species or ~~genera~~." to "…when building phylogenetic trees representing a single species or ==genus==."

- In last sentence of Figure 1 legend, changed: "…produces the best results when sequencing read depth is high and~~/or~~ community complexity is low" to be "produces the best results when sequencing read depth is high ==and== community complexity is low"

- Line 756-757, changed: "…numerous compositional approaches have been developed. For instance, several compositional correlation approaches ==~~have been developed~~…" to be "numerous compositional approaches have been developed. For instance, several compositional correlation approaches ==are now available==".

- In Figure 2 legend: Changed: "==~~The~~== illustrated example" to be "==This== illustrated example".

- Line 995, changed: "curated alignments of representative protein sequence" to be "curated alignments of representative protein sequence==s=="

- In Box 2, changed: "Because taxa and functional data types are qualitatively different from each other ==and== the choice of how to compare the two is based on somewhat arbitrary decisions on how to categorize them" to be "Because taxa and functional data types are qualitatively different from each other==,== the choice of how to compare the two is based on somewhat arbitrary decisions on how to categorize them".

- Line 1353, changed: "It is noteworthy that ==~~PAPRICA~~==, which represents a much simpler approach" to be "It is noteworthy that ==Piphillin==, which represents a much simpler approach"