

January 19, 2021

Dear Dr. Veronique van den Berghe,

Thank you and editorial board member Dr. Segata for handling our manuscript (“Compilation of longitudinal microbiota data and hospitalome from hematopoietic cell transplantation patients”) and encouraging a revision. We also thank the two anonymous referees for their constructive comments. We are pleased that both referees are positive and found our compiled data potentially useful for the field. We hereby present our revised manuscript, where we have addressed your and the referees’ comments. We include a detailed, point-to-point response to each comment below.

Editor’s Comments:

Comment 1: *“Dear Professor Xavier,*

First of all, our best wishes for the new year and please accept our apologies for the delay in sending you a decision on your manuscript. Your manuscript entitled "Compilation of longitudinal microbiota data and hospitalome from hematopoietic cell transplantation patients" has now been seen by the referees, whose comments are appended below. As you will see, the referees were supportive, but they raised important points and made suggestions for modifications, which they felt must be addressed before this work would be acceptable for publication at Scientific Data.

In line with Referee #1's comments, we request that you make all SRR accessions public or at least provide the SRA metadata. Also, could project level identifiers be added to the tblASVsamples file?

Based on the recommendation from the handling Editorial Board member, Nicola Segata, we therefore invite you to revise and resubmit your manuscript, taking into account the points raised.”

Author Response: Thank you and Dr. Segata for handling our manuscript and providing us feedbacks from referees during the pandemic. A delayed review process is totally understandable. Below we have addressed all their comments and made corresponding modifications in our manuscript.

In response to your specific concern, we have made all SRR accessions public and provided project-level identifiers (column “BioProject”) in the tblASVsamples file (<https://figshare.com/articles/dataset/samples/12016983>). We described this new column in the section “Data Records” of our manuscript (Line 259).

Comment 2: *“If you have not already done so, at this time, we ask you to use our metadata tool (<https://scientificdata.metadata-creator.com>) to submit key characteristics about the data described in SDATA-20-01150. Submitting your metadata at this point will result in a quicker publication process should your manuscript be accepted. Our in-house curation team uses the information submitted to finalise the machine-readable metadata for all accepted Data Descriptor manuscripts.”*

Author Response: We have submitted our metadata using the online tool.

Comment 3: *“At the same time, we ask that you ensure your manuscript complies with our format requirements explained in full in our Submission Guidelines:*

<https://www.nature.com/sdata/publish/submission-guidelines>

** Please ensure that your references, code and data citations conform fully to the Nature style. See the examples at the links below:*

<https://www.nature.com/sdata/publish/submission-guidelines#refs>>
and https://www.nature.com/sdata/publish/submission-guidelines#data_citations>

For example, with the current details reference 36 should become:
Liao, C. samples. figshare <https://doi.org/10.6084/m9.figshare.12016983.v6> (2020).

You may want to adapt the title and co-authors of your figshare files, so that the references become more informative as well.”

Author Response: We have checked the submission guideline thoroughly and ensured that our manuscript, including the reference style, conforms to the general guidelines.

Regarding data citations, we have put together all our datasets within a single figshare collection (<https://doi.org/10.6084/m9.figshare.c.5271128.v1>) and generated the following informative citation format including title and author

Liao, C. & Xavier, J. B. Compilation of longitudinal microbiota data and hospitalome from hematopoietic cell transplantation patients. *figshare* <https://doi.org/10.6084/m9.figshare.c.5271128.v1> (2021).

This reference (Lines: 608-610) has been used in the manuscript when referring to our data.

Comment 4: *“* Please upload all figures and tables as separate files. Tables should preferably be in Excel format. Line-art figures should be provided in a suitable vector-graphic format, ideally by saving them directly in the EPS or PDF formats from the program used to create them. Please note that some of the coloured boxes in Figure 5 are difficult to distinguish.”*

Author Response: We have uploaded all figures (no table was provided) as separate PDF files in vector-graphic format. For Figure 5, we created different patterns for boxes with similar colors to distinguish among them.

Comment 5: *“* For formatting reasons, please change the following section titles "Contributions" and "Acknowledgement" into "Author contributions" and "Acknowledgements" respectively. Please also replace "subjects" by "participants" at line 100.”*

Author Response: We have implemented these changes (Line 106, 515, 526).