

## 2nd round of revisions

Dear authors,

Thank you for submitting the first round of revisions. Both reviewers found you answered to most of their initial concerns. However, they noticed some issues remained that they would like to be addressed before the manuscript being accepted for publication.

Please address all these points and submit both a revised version of your manuscript and response letter.

I look forward to receiving your revised manuscript.

Yours sincerely

Thomas POLLET

*We thank the reviewers for their positive answers and for the new round of revisions. We hope that the new version of this article will be acceptable for publication.*

**Reviewed by anonymous reviewer, 04 Jan 2023 15:21**

I found that the authors have answered my initial concerns mostly sufficiently.

*We are very thankful again for the time and thoroughness you have dedicated to this review.*

There are two larger issues and some minor issues that I would like to be still addressed.

First, there is still a small lack of clarity with interactions. I note that the authors are careful of not mixing their observed associations as interactions, which is commendable, but this might be worth stating explicitly. For example, in paragraph from rows 127 to 148 as it will be one of the main motives in discussion where this is raised multiple times.

*We have included some elements in the introduction to state explicitly that empirical studies in the field enable to detect patterns of associations between pathogens and or microbiota. This is an important pre-requisite to define hypotheses about potential interactions and their underlying mechanisms.*

Second, the lack of broader theoretical framework causes some problems in relation to the Introduction and Discussion. There are a number of lenses to look at the assembly process for three different communities (e.g., stochastic/adaptive, extrinsic/intrinsic, environmental/?, transmission/interactions, colonization/maintenance, systemic/direct, local/global in introduction only!). I would go once through these passages and consider whether there is still room to simplify the conceptual jungle.

*We have included a paragraph in the introduction presenting the processes and the environmental/host factors that influence gut microbiota assembly.*

*Then for the helminths and the microparasites, we have detailed how interactions with gut microbiota could occur at a local and systemic scale.*

*We have homogenized some terms throughout the manuscript, e.g. environmental/extrinsic; host/intrinsic.*

*We hope that these changes facilitate the reading of the manuscript.*

**Minor issues:**

- Row 84: Is not the "variation in microbial fitness" rather a consequence of the variation of microbes differential reproductive success rather than a driver of said variation? If you want to contrast for "stochastic processes", then I would think something along the lines of "adaptive differences in microbes". *We agree with this proposal. We have changed the text in consequence.*

- Row 85: Should be "host". *This has been changed.*

- Row 84-90: I find everytime these classifications of different factors difficult. I would not consider parasite infection as "environmental factor". In previous sentence you have said that it is an extrinsic feature, which makes sense. *We agree and we have changed the text in consequence.*

- Row 102: or rather "...bacteriota interactions may thus be positive or negative (Loke & Lim, 2015), with..."? *We meant interactions between helminths and gut bacteriota. We have reformulated the text in consequence.*

- Row 118-126: The references should be used a bit more accurate. I would like to see a reference to systemic impact and for the claim that studies have focused on viruses, not on bacteria. *We have included several references to illustrate these assertions.*

- Row 139: This claim should have a reference, maybe to a classic study. *It was quite hard to choose, we have therefore included two references.*

- Row 515: I think that it might overstretching your results to say that you found that individuals were clustered within distinct enterotypes. The discussion that you are presenting considers this to be the case, but you had only one sample per individual, so technically, you rather have just a snap-shot in time. I think the first important step here would be to figure out whether distinct enterotypes are longitudinally constant or whether intraindividual variation trumps here interindividual variation. *We thank the reviewer for this comment. It is something that could be interesting to study. We have added a sentence to point this issue.*

- Row 537: It is strange to say that geographic location shapes the variation in gut bacteriota. I would expect that it is abiotic or biotic factors related to the location that does the shaping as you write. *We agree with the reviewer and we have changed the text to better explain what we previously meant with 'geographic location'.*

- Row 576: I do not understand this sentence. Do you mean to say that you did not find it? *Yes, we meant that we did not find strong evidence of three-way relationships between the gut microbiota, the gastro-intestinal helminth and pathogenic bacteria communities. We have changed the text accordingly.*
- Row 646: I do not follow the second idea fully. Not considering interactions, I would think that gut microbiota is more shaped by host immune function and/or host diet, whereas gut helminths are more limited by invasion success – I think this would be the thinking along the “everything is everywhere” paradigm of microbial ecology. Then again, this maybe goes beyond the authors’ focus as there is no general framework on how to combine transmission/external factors to within-host interactions in shaping the gut community. *We do not fully agree about what the reviewer indicates as the main processes influencing the development of the gut microbiota and helminths’ infestations. The host immune system is important in shaping host sensibility to helminths and the outcomes of helminths infestations (from elimination and expulsion to establishment, and more or less limitations of helminths development). Immunoheterogeneity among host individuals, populations and species is also known to be a main driver of helminth epidemiology. In addition, host diet may influence the exposure to helminths (e.g. soil or food transmitted helminths, ingestion of infested intermediate hosts). Therefore, we have not taken this comment into account. However, we may not have fully understood what the reviewer wanted to point out, so we are open to further discussions about this issue.*

### **Reviewed by J. Anders, 13 January 2023**

Bouilloud et al. have put in a significant amount of time and effort into revising this manuscript and it really shows. The overall story is much easier to understand, the writing is much clearer, and the ideas presented are more fully developed. This new iteration was enjoyable to read.

The authors have addressed most of my previous concerns, though I'm still concerned about the use of an older taxonomic classifier. But as long as an appropriate conformational check to show it's not affecting the results is provided it should be ok. Please see my comment below for more details. The only other concern I have is the lack of references provided.

There are quite a few sentences where ideas, information, and previous studies are talked about without a proper reference provided. The rest of my comments are mostly minor and whose purpose is to help improve the manuscript just a little bit more.

*We thank the reviewer for his positive comments that were very constructive.*

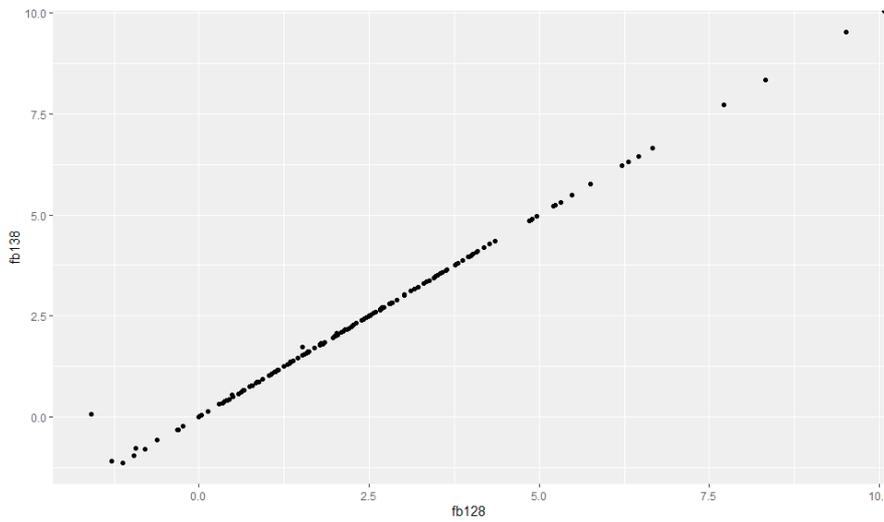
### Major comments

Lines 222 – 224: I agree that it is not always possible to use the most up to date version of such resources as they are updated routinely. However, SILVA SSU 132 was released in December 2017, presumably before you conducted your bioinformatics analyses in 2018. Also, version 138 was released in December 2019 (effectively 2020), more than 2 years before this paper was first submitted to PCI, providing plenty of time to incorporate it into your study. I agree that proper reporting of your methods is important and helps to interpret your results, but using such outdated information can cause misleading or inaccurate results. Especially since there have been significant increase in the number of bacterial species described and cataloged, as well as taxonomic classifications being reworked. It is the same reason why the greengenes reference database is not typically used anymore as it has not been updated in nearly 10 years. SILVA SSU 128 is only a couple years newer than greengenes. I understand that redoing all your analyses following OTU clustering is a major task that you are hesitant to do. I would be as well. Using a more updated reference database ultimately may not change the results of your study, especially your diversity analyses unless it is able to provide taxonomic classification for a larger number of OTUs that remained unclassified at the family with SILVA 128. What I am most concerned about is that it could change your Firmicutes / Bacteroidetes ratio estimates and lead to very different results for your differential abundance analyses (DESEQ2). I highly recommend you at least provide some form of confirmation that this is not impacting your results. For example, showing that the Firmicutes / Bacteroidetes ratio is similar for both the 128 and 138 releases.

*We thank the reviewer for his understanding.*

*As recommended, we have compared the estimations of the Firmicutes / Bacteroidetes ratio following the SILVA 128 and SILVA 138 databases. We have also compared the assignments given by these databases for the most abundant families present in our dataset (90% of the number of reads).*

***F/B ratio** : We found a significant correlation between the estimates obtained from the two datasets ( $r=0.99$ ;  $p$ -value  $<10^{-5}$ ). The difference that could have influenced the F/B ratio and thus the interpretation of our results is the inclusion of the phylum Tenericutes within the phylum Firmicutes in SILVA 138. However, this former phylum Tenericutes was only weakly represented in our dataset with only 2 individuals that had a number of reads  $>500$ . Therefore, changes between SILVA 128 and SILVA 138 had almost no impact on the Firmicutes/Bacteroidetes ratio.*



*Correlation between the F/B ratio estimated using the SILVA 128 and SILVA 138 databases*

*Assignations of the most abundant OTUs: The databases SILVA 128 and 138 provided very similar results at the family taxonomic rank, when considering the most abundant families in our dataset (>90% of all the reads). The main change concerned the Ruminococcaceae family that is now divided in two distinct families, the Oscillospiraceae and the Ruminococcaceae.*

**SILVA 128**

Cluster	Prevalence	TotalAbundance	RelativeAbundance	Phylum	Family
Cluster_42	1.00000000	304252	0.110285861	Firmicutes	Lachnospiraceae
Cluster_96	0.97515528	104315	0.037812305	Firmicutes	Ruminococcaceae
Cluster_117	0.89440994	38920	0.014107798	Firmicutes	Clostridiales2
Cluster_27	0.33540373	46489	0.016851424	Firmicutes	Christensenellaceae
Cluster_101	0.49689441	19241	0.006974515	Firmicutes	Christensenellaceae
Cluster_3	0.72049689	339882	0.123201093	Epsilonbacteraeota	Helicobacteraceae
Cluster_5	0.99378882	511689	0.185478030	Proteobacteria	Desulfobibrionaceae
Cluster_75	0.05590062	26106	0.009462954	Firmicutes	Erysipelotrichaceae
Cluster_14	0.90062112	590058	0.213885379	Firmicutes	Lactobacillaceae
Cluster_22	0.65217391	69014	0.025016330	Spirochaetes	Spirochaetaceae
Cluster_10	0.96273292	322802	0.117009901	Actinobacteria	Eggerthellaceae
Cluster_28	0.73291925	37151	0.013466567	Bacteroidetes	Bacteroidales1
Cluster_26	0.95652174	215421	0.078086226	Bacteroidetes	Muribaculaceae

**SILVA 138**

Cluster	Prevalence	TotalAbundance	RelativeAbundance	Phylum	Family
Cluster_42	1.00000000	304252	0.110285861	Firmicutes	Lachnospiraceae
Cluster_96	0.85093168	25054	0.009081623	Firmicutes	Ruminococcaceae
Cluster_213	0.73291925	53780	0.019494280	Firmicutes	Oscillospiraceae
Cluster_27	0.21739130	42573	0.015431944	Firmicutes	Clostridia_1
Cluster_101	0.49689441	19241	0.006974515	Firmicutes	Christensenellaceae
Cluster_3	0.72049689	339882	0.123201093	Campylobacterota	Helicobacteraceae
Cluster_5	0.99378882	511689	0.185478030	Desulfobacterota	Desulfobibrionaceae
Cluster_75	0.05590062	26106	0.009462954	Firmicutes	Erysipelotrichaceae
Cluster_14	0.90062112	590058	0.213885379	Firmicutes	Lactobacillaceae
Cluster_22	0.65217391	69014	0.025016330	Spirochaetota	Spirochaetaceae
Cluster_10	0.96273292	322802	0.117009901	Actinobacteriota	Eggerthellaceae
Cluster_28	0.49689441	29853	0.010821174	Bacteroidota	Bacteroidales_1
Cluster_26	0.95652174	215421	0.078086226	Bacteroidota	Muribaculaceae

*This outcome (no change in the main patterns and messages of our study) confirms the expectation that using a more updated reference database ultimately may not change the results of the study, especially for the diversity analyses. We are particularly indebted to the reviewer for having pointed out both this issue and a way to ensure the robustness of the work's messages. Furthermore, the public access to the « Review document » (allowed by the journal) will be of instrumental help for enlightening readers that would have similar concerns about the database version. This is why we have provided above the details of the analyses suggested by the reviewer. »*

### Minor comments

General comment on the use of commas: Generally speaking, in scientific writing it is better to avoid the use of commas for stylistic purposes and instead use them only for grammatical reasons. The overuse of commas for stylistic purposes can lead to the tendency of creating unnecessarily long and complex sentences that requires the reader to read multiple times to really understand it. It makes the reader tired, and is more difficult and time consuming to read through the manuscript. Please check their use throughout the manuscript as it was common throughout. Just to point out a few, there should be no comma after “fitness” (Line 77), “bacteriota” (Line 82), or “weight loss” (Line 105). *We have read the text carefully to take this comment into account.*

Also, please check the proper use of the singular / plural form of nouns.

Lines 102-104: Similar to the use of transition words at the beginning of a paragraph, “may thus” implies cause and consequence. However, what you are referring to when using “thus” is information provided within the previous paragraph. I recommend rephrasing this as something along the lines of, “Helminth – gut bacteriota interactions may be positive or negative with potentially both local and systemic physiological changes affecting host health”. *We have modified the text in consequence.*

Line 106-108: References? *We have added a reference to justify this assertion.*

Lines 116 to 124: Each of these sentences requires references. *We have added several references to justify these sentences.*

Lines 119 – 120: No “...” in the parentheses. The same on line 123. *We have deleted ‘...’*

Lines 131 – 132: References? *We have added two references to justify this assertion.*

Line 154: Do you mean “health status of the host”? In this particular sentence it is unclear if this is referring to the “health status” of the host or the gut microbiota. *Yes, we meant health status of the host. This has been clarified in the text.*

Line 196: “lower segments” not “segment” *This has been changed.*

Line 198: How was the lumen removed? Washed off with a solution or scrapped off with a tool? *We first used sterile tweezers to remove the lumen. Then we washed the piece of colon with ultrapure water.*

Line 219-220: I agree with your argument outlined in your response letter that OTUs have received a bad reputation recently, but that they are not obsolete and that sticking with OTUs for this manuscript is appropriate. That being said, the way this is currently phrased, it sounds like the purpose of this method for OTU clustering is to produce something similar to ASVs which begs the question, “why not just use ASVs?”. Also, there is no clustering of ASVs like OTUs, so they are inherently different which makes this statement a bit confusing. Do you mean it is of similar quality or reliability?

*We rephrased this sentence and removed the reference to ASVs, as we agree it was confusing. We didn't include more details in the revised manuscript but here are some answers to your question.*

*Since the community analyses are done at the phylum or family levels, we think that ASVs and OTUs are of similar quality. The advantage of OTUs is that the ASV approach enables to consider sequencing errors only, while the OTU approach also enables to consider PCR errors. The OTU sequences obtained therefore might contain less errors than the ASV sequences.*

*To our knowledge, there is only one paper that compared these two approaches (Bernard et al., 2022; DOI: 10.1093/bib/bbab318) and showed the benefits of using FROGS vs DADA2 pipeline.*

Line 234 - 235: “Number of reads per OTU” not “OTUs number of reads”? *This has been changed.*

Line 269: Referencing the references in another paper is not appropriate. Please provide the references you are referring to here. *We have added the original references in the text.*

Line 422-431: this should be one paragraph. *The three previous paragraphs have been merged.*

Line 587: The same as my comment regarding the reference on line 269. Referencing the references in another paper is not appropriate. Please provide the actually references you are referring to here. *We have added an original reference in the text.*

Line 596: This is a bit misleading as you are reporting a correlation between *H. mixtum* and *Lactobacillaceae* abundance whereas Reynolds et al. 2014 showed a direct relationship through experimental inoculation. I think it is really interesting that you found a similar correlation with *H. mixtum*, but I recommend rephrasing the last part of this sentence, “as observed in our study” to better reflect the limitations of your own data. *We have modified the text to dampen our statement. (“as suggested by the correlation detected in our study”)*

Lines 606 - 607: “of this later”? This may have been accidentally left in from a previous version of this sentence. *We have deleted this part of the sentence.*

Lines 607 – 610: This is more appropriate for the results section. I recommend deleting this entire sentence and making the sentence on line 605 – 607 the start of paragraph beginning on line 611. *We have deleted this sentence.*