

SOIL Discuss., referee comment RC1
<https://doi.org/10.5194/soil-2021-148-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.



Comment on soil-2021-148

Anonymous Referee #1

Referee comment on "Lower functional redundancy in "narrow" than "broad" functions in global soil metagenomics" by Huaihai Chen et al., SOIL Discuss., <https://doi.org/10.5194/soil-2021-148-RC1>, 2022

General Comments:

In this work, the authors present a synthetic analysis of publically available metagenome data to explore the concept of microbial functional redundancy. They divide annotated metabolic functions into five "broad" functions, mostly related to central carbon metabolism and five "narrow" functions related to nitrogen, phosphorus, and micronutrient metabolism. They then construct pairwise similarity comparisons between all soils in terms of taxonomy and functional gene diversity for each functional gene grouping of interest. They find that there is a greater positive correlation for "narrow" functions than "broad" functions, indicating taxonomic similarity correlates with functional gene similarity. They then go on to explore the taxonomic diversity within functional groupings and identify major phyla in which designated functional genes are more abundant. Overall, I found the study to be well designed, executed, and explained. While the results are not surprising, I think it is crucial to test assumptions with available large datasets and agree with the authors that this type of work is important in "generalizing patterns of microbial characteristics regulating biogeochemical cycling".

Specific Comments:

I think there may be utility in better defining the concept of functional redundancy used for this work early on in the introduction. One might be interested in strict redundancy (i.e., exact same set of functions) versus "partial redundancy" (i.e., share some set of functions but nevertheless differ in terms of additional functions or environmental requirements). An example of partial redundancy would be two ammonium oxidizers that

have different temperature preferences. I can see that a similar definition to that of Louca et al., 2018 ("with respect to a given function" - see glossary) is what is meant here, but I think it is worth re-stating clearly in the beginning of the manuscript. It is also worth mentioning somewhere in the introduction that redundancy may result from non-metabolic processes and thus not be captured in the current approach centered around metabolic redundancy (e.g., habitat preferences, particle attachment strategies, predation - see Louca et al., 2018). This inevitability that results from projecting from low to high dimensional space is mentioned in line 367, but could be mentioned sooner. See Garland et al., 2018 introduction (<https://doi.org/10.1038/s41396-018-0158-1>) for my exposure to this concept and the above example as well as Louca et al., 2018 (already cited many places).

I think some of this contradiction likely arises from the definition of "redundancy" as well and our limitations in measuring the factors controlling niche space (line 72-73).

Is there any prior precedent for the threshold values chosen for categorization in the network analysis (lines 254-255, Figure 5)? This reference should be in the methods section near the line specifying the chosen values.

Technical Corrections:

[line 54] – Change to "does not rapidly recover" (subject/verb agreement)

[line 60] – Change to "shifts" (subject/verb agreement)

[line 106] – Change to "pairwise" (repeated typo throughout manuscript)

[line 333] - Should state "based on certain taxonomic and functional databases"

[line 335] – I believe this should be "may contain potential bias"

Captions for Figure 3 and 4 need to be switched.