Review of: "Cross-kingdom metagenomic profiling of Lake Hillier reveals pigment-rich polyextremophiles and wide-ranging metabolic adaptations"

Horia Leonard Banciu[1]

1 Babes-Bolyai University of Cluj-Napoca

Potential competing interests: The author(s) declared that no potential competing interests exist.

The manuscript bioRxiv preprint doi: https://doi.org/10.1101/2022.02.17.480683 entitled “Cross-kingdom metagenomic profiling of Lake Hillier reveals pigment-rich polyextremophiles and wide-ranging metabolic adaptations” by Sierra M. et al., focuses on revealing the microbiome diversity and its possible involvement in driving the peculiar pink coloration of Lake Hillier (SW Australia). To achieve this aim, the authors mainly employed NGS as both amplicon sequencing and WGS approaches on environmental samples covering shore , water and aquatic sediments.

I have tried to provide a critical review of the manuscript in an attempt to aid in the improvement of the manuscript. My numerous major and minor comments are given on the manuscript files and scanned as PDF. However, I am here expressing only the major concerns as a list of statement. These statements, however, are also written in the manuscript filles annotated bty me, hopefully, in a readable handwriting format.

Title and Background

- In my view the title is misleading as some part of the results provide information on issues not related to lake coloration or metabolic “adaptation”, I meant that the (although interesting) BGC exploration is not well linked to the general aim of research defined in this manuscript nor with the title;
- In the same idea as above, the culturing approach is weakly supporting the title and aim of the research;
- In my opinion, the backround/ abstract could be much better as it currently provide little specific information, so, generally lacks focus; many parts are commonplace and should, instead, focus on the uniqueness of Lake Hillier (the pink color could not be the distinctive feature alone as this color is largely noted in many coastal hypersaline crystallizer ponds or some sea-derived shallow hypersaline lakes); Lake Hillier must be much better advertised in the Introduction part;

Materials and Methods

- No chemistry information is provided in the results and especially those indicating salinity and oxygen (or redox potential) in the assayed water/sediment samples; as an environmental microbiology-oriented manuscript, the metadata are essential. In the same line, I was not able to see if there any basic chemical assessment of pigments have been performed say, by simple spectrophotometric measurements of total extracted pigments. The absorption spectra of solvent-extracted pigments would have provided a wealth of information including the identification of bacteriochlorophylls (typical to PSB), bacterioruberin (Halobacteria Class) or salinixanthin (Salinibacter);
- I have some concerns on using 10% salinity cultivation media instead of 25-28% as the Lake Hillier setting; similar

concerns are expressed for the sample preparation as the samples have been reported to be diluted many fold prior to filtration – during this preparatory step (I roughly calculated that dilution was from 28% to 5-6% salinity which makes a high difference for halophiles), many osmotically adapted cells will lyse releasing their DNA prior to the biomass retainment of membrane filters;

- I am concerned whether TMD is validly published or accepted in the scientific community as the reference reporting this database appears as preprint only since 2019;
- I am concerned on overall low number of samples (two) scrutinized for WGS-inferred microbiome and the resulting relatively low number of MAGs (21); usually one may see these numbers as weakly supporting the scientific output;
- A detailed Table showing the precise number of samples, their origin and brief description together with their subsequent use (either for amplicon or WGS) is needed; ideally environmental data (such as salinity and ORP) should be provided;
- It is not clear how the Illumina short-reads were clustered to provide either OTUs or ASVs as these approaches are standard for amplicon-sequencing-based taxonomic approaches;

Results and Discussions

- As already stated above, the reason of analysing BGC within the environmental context (as suggested by the title) is not clearly explained; however, I found no further convincing discussion either;
- There is a recurrent and highly arguable use of the “species” term throughout the manuscript; I stress that the amplicon sequencing has low resolution below Family level, whereas metagenomics (with a higher resolution below Family level) has revealed only a handful of MAGs; at some point, the authors refer to over 13000 species, while this number varies a lot throughout the manuscript – a thorough reading by a microbial taxonomist is highly advisable;
- There is an inconsistent reference to fungi and algae; moreover, there is an inconsistent reference to Dunaliella as I found no reported evidence of Chlorophyta-related SSU RNA gene sequences;
- There is no clear discussion on the funglar diversity and their putative ecological roles;
- I am skeptical on naming viruses as the fourth domain of life but this could be supported by a valid reference;
- I am afraid that there are many more inconsistencies throughout the manuscript that add to some problematic English. I strongly recommend that the manuscript should be carefully read and proofread by native English authors;
- I have not made many comments on the Discussion as I noted that many parts of the results should be re-worded or more carefully re-assessed; in the end, I rehearse the idea of gaps between the former aim of research (microbiome diversity and inferred metabolic capabilities in the environmental context) and exploring BGC, perhaps reconsidering the title, aim and concept of the manuscript to include BGC would be advisable, otherwise I would suggest more focus on ecological roles.