**Biogeosciences
Discussions**

Interactive
Comment

Interactive comment on "Arctic Gypsum Endoliths: a biogeochemical characterization of a viable and active microbial community" *by* L. A. Ziolkowski et al.

Anonymous Referee #1

Received and published: 2 April 2013

This manuscript presents microbial diversity, lipid and C isotope analyses of gypsum endoliths from a cold desert region in the Arctic. The authors describe a moderate diversity of prokaryotic and eukaryotic organisms, which is nevertheless more diverse than that found in other gypsum endolithic communities of even more extreme environments, such as the Antarctic valleys or the Atacama desert, and which is active, as can be deduced from the incorporation of 14C isotopes dating back to the 1960s. Overall, the information presented in this manuscript contributes to enrich knowledge about these peculiar ecosystems.

However, I have some important concerns that need to be addressed before this

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

manuscript can be published:

- My major concern relates to the lack of very basic details about the analysis of pyrosequences and the determination of the microbial community composition. There is no mention whatsoever to the number of sequences produced, what was the chemistry used (Titanium, I guess?), etc. Furthermore, it is unclear if the authors carried out metagenomic sequencing of total DNA, as could be inferred from lines 23-24 page 2276 ("Total DNA was concentrated from multiple extractions ... and sent for pyrosequencing analyses..."), or if they sequenced SSU rDNA amplicons, as can be deduced from the description of primers used (lines 24 page 2276-line 2 page 2277) and lines 25-26 page 2279 ("Pyrosequencing of the 16S, 18S, and 23S plastid rRNA genes of..."). And if they pyrosequenced amplicons, did they mix them? How? How many libraries per run or line? When the authors say that only sequences of groups constituting more than 1% of the reads, how many are these? Later, "representative sequences" of what, how many and what was their proportion from the total? All the basic details are absent. The methodology to treat the pyrosequences is also far from clear. Pyrosequencing provides massif amounts of reads that cannot be easily analyzed by manual BLAST, without some kind of clustering process earlier, which is not described. There are also very important concerns regarding sequence analyses. Pyrosequencing introduces many errors. How did the authors control for quality? How many sequences were eliminated? Also, depending on the clustering method used to group the sequences in operational taxonomic units (if this was at all done – because it is not described, though standard procedure), the diversity can be highly overestimated. Provide the total number of sequences, describe exactly how you treated them and clustered them (what cut-off values were used) and provide a final number of OTUs retained as true per phylum (e.g. in a table).

- This is very important: Sequences must be deposited in a public database and made accessible to the scientific community. The authors should submit their sequences to the Sequence Read Archive (NCBI) and provide an accession number. This is a basic

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

ethical rule and a general requirement for publication.

- Phylogenetic analyses. In the methods section, the authors say that they select "representative" sequences and aligned them to generate a neighbor-joining phylogenetic tree. However, the figures show "maximum likelihood" trees. Please, precise. Also, indicate whether the ambiguously aligned positions were removed. Making phylogenetic trees with relatively few positions (pyrosequences) is dangerous, especially if sequences from disparate phyla are included. The tree in Figure 5 is an example. Alphaproteobacteria are not monophyletic. Cyanobacteria and chloroplasts do not cluster together. Likely, if chloroplast sequences were removed, a better resolution could be achieved for bacteria.

- The benefit of using 23S rRNA gene markers to describe the diversity of algae is unclear, since in most cases, the environmental diversity of plastid markers is not yet linked to the diversity of the nuclear 28S rRNA genes. And even if it was the case, the database for 28S rRNA genes is much poorer than that of the most widely used 18S rRNA gene. I do not understand why the authors use fungal 18S rRNA primers and 23S rRNA plastid primers instead of using universal eukaryotic 18S rRNA primers, which would have led to the amplification of both, fungi and all kinds of algae, plus other protists. Indeed, here, if there are other protists associated to the endoliths, we are missing them, because they have not been targeted.

- The authors use "Algae" as if algae constituted a taxon: it is not. Algae group a variety of photosynthetic eukaryotes belonging to very different phyla (green algae, red algae, diatoms, euglenas, etc). Therefore, the authors should use the corresponding taxon name. Also, vernacular names should not be written starting with a capital letter, only Latin names should.

Interactive comment on Biogeosciences Discuss., 10, 2269, 2013.

Interactive
Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper