Biogeosciences Discuss., 12, C10231–C10234, 2016
www.biogeosciences-discuss.net/12/C10231/2016/
© Author(s) 2016. This work is distributed under
the Creative Commons Attribute 3.0 License.

Biogeosciences
Discussions
Open Access
EGU

Interactive comment on "The keystone species of Precambrian deep bedrock biosphere belong to *Burkholderiales* and *Clostridiales*" *by* L. Purkamo et al.

L. Purkamo et al.

lotta.purkamo@gmail.com

Received and published: 17 March 2016

The authors thank the referee for the constructive remarks and criticism that will be helpful in improving this manuscript. Both Referee #1 and Referee #2 address the reliability of metagenomic prediction with PICRUSt for the quantitative analysis of certain pathways. As we mention in the manuscript, PICRUSt analysis may only be used as a prediction of the properties of the metagenome (P. 18113, l.17-18). The high values of weighted NSTI (nearest sequenced taxon index) also denote the uncertainty of the prediction. This is why we do not infer any major conclusions in the discussion part, although we believe that assessing the predicted metagenome we can have an

overview of the top-level functionality of the metagenome of this environment. We can tone down the results part of this analysis, as Referee #2 suggests. In addition, we consider omitting the results of the HUMANn analysis (p.18114, l.1-8 and p.18121, l. 20→p.18122, l. 4 and l. 27→p.18123, l. 2), where we have used these predicted metagenomes as starting material. This analysis depends ultimately to the partial 16S rRNA sequences, so stating that certain pathways are complete may not be feasible. Specific comments from Referee #2: We consider changing the title for more descriptive "Microbial co-occurence patterns in Precambrian deep bedrock biosphere" and we can state in the abstract and again at the discussion part that the samples come from groundwater (p.18104, l. 4, p.18124, l. 4). We could also add groundwater as one of the keywords.

P6_L16: Why was DNase treatment not applied to all the RNA extracts? We used DNase treatment only the RNA extracts that were verified to contain residual DNA in order to avoid the loss of material, as the amount of retained RNA was low already in the beginning, and during the DNase treatment we would have lost more of the precious material.

P8_L11: What is the coverage of the archaeal 16S rRNA gene diversity with the primer pair used for the nested PCR? Different primer pairs are known to be biased against certain archaeal groups. See for example Teske et al. (2008). Primers used in the archaeal nested PCR are reported to have frequent mismatch rate to some archaeal groups, such as SAGMEGs (Hadesarchaea)(Teske et al. 2008, Wang & Qian 2009). This might lead to underestimation of these groups in the final results. However, we did find a large amount of SAGMEG sequences in one depth of the bedrock, further reinforcing the results from previous study (Purkamo et al. 2013). Using the Ribosomal Database Project's probe match platform, we detected that archaeal primers used in the first step of the nested PCR matched 39% of the archaeal 16S rRNA genes in the database, and the primers used in the second step 31%. Nevertheless, phylum Thermococci, in which newly described Hadesarchaea (SAGMEGs) belong to, is cov-

ered much better as the overall archaea, as the coverage range of the used primers is from 85-91%. We can add this primer match data as a supplementary table to the manuscript if necessary.

P8_L20: What were the PCR conditions for the 16S rRNA genes? Quantitative PCR conditions for bacterial 16S rRNA genes are described in the manuscript in P7_L26→. Archaeal 16S rRNA -targeted qPCR conditions are previously described in Bomberg et al. (2015): "The qPCR reactions were performed in 10μL reaction volumes using the KAPA 2 × Sybr FAST qPCR-kit on a LightCycler480 qPCR machine (Roche Applied Science, Germany) on white 96-well plates (Roche Applied Science, Germany) sealed with transparent adhesive seals (4titude, UK). Each reaction contained 2.5 μM of relevant forward and reverse primer and 1μL DNA extract. Each reaction was run in triplicate and no-template control reactions were used to determine background fluorescence in the reactions. The qPCR conditions consisted of an initial denaturation at 95 °C for 10 min followed by 45 amplification cycles of 15s at 95 °C, 30s at 55 °C and 30s at 72 °C with a quantification measurement at the end of each elongation. A final extension step of three minutes at 72 °C was performed prior to a melting curve analysis. This consisted of a denaturation step for 10 s at 95 °C followed by an annealing step at 65 °C for one minute prior to a gradual temperature rise to 95 °C at a rate of 0.11 °C s-1 during which the fluorescence was continuously measured.". Amplicon library construction for sequencing is described in p.18111, lines 2-19.

Results P15_L5: I see in figure 4 that there are some phototrophs in the RNA-based analysis from the 2260 m sample. How do the authors explain this? In this sample, 4,5% of the sequences affiliated with Streptophyta, most likely representing plant material. These are probable contaminants that could be originating from the drilling operations, where some peat material was used in the drilling mud as a stabilizer in the deepest parts of the drill hole.

P15_L17: There exists a recently published paper on a SAGMEG genome. These archaea have been renamed Hadesarchaea (Baker, Nature microbiology, 2016). Please

change conclusions in the discussion part based on findings in this paper as well. We will change the discussion part concerning SAGMEGs, as this new paper sheds light to the carbon metabolism of these archaea.

Discussion P.20_L16: (. . .) these organisms can prevail 'in the same environment' or 'same environments'. Should be "in the same environment".

P22_L3: the authors don't offer an explanation as to why the SAGMEG dominate the DNA-based analysis but methanogens dominate the RNA- based analysis at 967m. Is the number of reads obtained for each analysis significantly different? Please address this issue. The number of reads in these samples was highly similar, so this is probably not the reason for this result. However, as different microbes have different amount of ribosomes, depending on the activity of the microbe and the need for production of proteins, this could explain the difference between total and active archaeal community structure at 967 m. There are reports from deep subsurface environments with similar results (Posiva report by Bomberg & Itävaara 2012 http://www.posiva.fi/files/3096/WR_2012-27.pdf, Miettinen et al. 2015 doi: 10.3389/fmicb.2015.01203). In addition, it was demonstrated in recent study that only certain microbes in the total population are active but their numbers might be so low that they are not even detectable via DNA-based analysis (Rajala et al. 2015). These observations can be included to the manuscript.

P22_L16: this observation could be due to low coverage by the primer pair used for the nested PCR, and by the fact that 2 PCRs were carried out for the archaeal diversity analysis. We agree, and this notion can be implemented into the manuscript.

Interactive comment on Biogeosciences Discuss., 12, 18103, 2015.