Biogeosciences Discuss., 12, C6537–C6547, 2015
www.biogeosciences-discuss.net/12/C6537/2015/

Biogeosciences
Discussions
Open Access

Interactive comment on "Estimation of microbial metabolism and co-occurrence patterns in fracture groundwaters of deep crystalline bedrock at Olkiluoto, Finland" *by* M. Bomberg et al.

Anonymous Referee #2

Received and published: 15 October 2015

1 Overview

This paper presents the 16S rDNA data from 12 distinct sampling locations within the Fennoscandian Shield fracture system. The authors have generated a very complete dataset worthy of publication however, in my opinion, the authors need to correct some of their analytical methods as well as refocus their discussion before it is fit for publication.

2 Major Comments

From the data presented, their main results include (1) the correlation of rare taxa with geophysical and geochemical settings $\rightarrow$ discussion on rare biosphere and (2) predicted metabolisms derived from PICRUSt. My main concerns with the analysis are as follows:

1. The description of the approach taken to perform the correlation analysis and subsequent significance are not clearly stated. For example it is not stated whether or not the data was normalized in any manner prior to analysis (this is also omitted when describing their implementation of CCA on the data). More importantly, the authors do not seem to correct for the high false discovery rate when testing multiple hypotheses. If I understood correctly, the authors tested the pairwise correlation of 732 genera (651 bacterial and 81 archael). Under this scenario there would be $(731 * 732)/2 = 267,546$ comparisons made. Recall that the p-value is defined as the probability of obtaining a result equal to or more extreme than what was actually observed, assuming that the null hypothesis (there is no correlation between the pair in question) is true. Thus, a p-value threshold of 0.01 (as denoted in the methods section, I noticed figure 4 uses $p < 0.001$) would, in expectation, yield $2,675$ $(= 0.01 * 267,546)$ tested hypotheses that would appear significant just by random chance. *Note, that in lines 2 and 3 of page 13831, the authors report the number of genera (101) that were found to be significantly correlated to other genera not the number of significant correlations identified.* To correct for the false discovery rate, the authors need to use some sort of correction for multiple hypothesis testing such as the q-value (Storey, John D. "The positive false discovery rate: a Bayesian interpretation and the q-value." Annals of statistics (2003): 2013-2035.) or the Bonferroni. Until an appropriate test for significance is conducted, the results and discussion based on the correlation analyses, in my opinion, are not yet in publishable form.

2. Although the authors note (in lines 22 to 27 of page 13822 and elsewhere) that the method of functional inference from 16S data using PICRUSt has limitations, almost the entire discussion of the submitted paper is based on the predicted functional metabolisms identified through this method. My main concern is that they are too focused on reporting the PICRUSt results rather than how the metabolisms (and 16S data) may relate to the larger context of the paper, environment type, and field. This is especially concerning when the PICRUSt results may contradict other observations that the authors report based on the 16S data such the statement about the low abundance of sulfur metabolisms (PICRUSt) on page 13825 lines 10-12 and the results of their CCA (as wells as their correlations that need correcting – see above – in lines 14 to 17 page 13833) that show communities that correspond to increased sulfur and sulfate concentrations. What is changing between these sites if not the relative abundance of sulfate/sulfur reducers? What other thoughts might you have on this? Another anomaly that concerns me is the presence of methanogens and absence of the Wood-Ljungdahl (acetyl-CoA) pathway in archaea – a pathway that is a feature of methanogens. In my opinion, there is too much dependence on reporting these results that may not be entirely accurate and are subject to interpretation (two of my major concerns are listed above). The discussion and presentation of PICRUSt results can be enhanced by performing a deeper investigation and interpretation of the results – i.e. have the same trends (as far as predicted metabolisms) been observed before in the Fennoscandian Shield? How do they compare to metagenomes from the area or other subsurface sites? Other questions that seem important but appear to be largely ignored: Is there a difference in sulfur metabolism in the sites OL-KR5,6,9,13,23 than the rest as, taxonomically, they seem to correspond to increase in S and sulfate concentrations. What are the taxa? Are these taxa the same in the sites? Also the discussion of rare versus core is interesting but not well developed. Some things I would like to see more of are: Do the predicted metabolisms vary within the core set vs the rare?

What does that observation potentially say about the theory of rare biosphere?

3 Minor Comments:

3.1 Abstract

- Lines 6-10 are confusing and need clarification. For example, I think you are referring to 95 and 99% of the alpha diversity but it is unclear
- Significance needs to be reexamined as described in **Major Comment (1)**
- "It may consist of remnants of microbial communities prevailing in earlier conditions on Earth" is a bit misleading in the context of the rest of the paper. In my opinion, as written and throughout the paper, the discussion on the rare biosphere is rather ambiguous and subject to misinterpretation by the reader and need clarification. For example, the rare biosphere itself is not necessarily ancient (as stated in line 21 of page 13821) – it is a feature that is present in the microbial community observed today. However, the introduction of these taxa into the Fennoscandian Shield may have happened a long time ago and, over time, the taxa have persisted in the environment at low abundance. This is a feature not just of the Fennoscandian site but something that relates to all microbial communities. Also, it is important to note that in Sogin et al. (2008), the aforementioned mechanism is not the only avenue by which the rare biosphere may appear. Sogin et al. (2008) note: *"The large number of highly diverse, low-abundance OTUs constitutes a 'rare biosphere' that is largely unexplored. Some of its members might serve as keystone species within complex consortia; others might simply be the products of historical ecological change with the potential to become dominant in response to shifts in environmental conditions (e.g., when local or global*

change favors their growth). Because we know so little about the global distribution of members of the rare biosphere, it is not yet possible to know whether they represent specific biogeographical distributions of bacterial taxa, functional selection by particular marine environments, or cosmopolitan distribution of all microbial taxa (the 'everything is everywhere' hypothesis)"

3.2 Introduction

- Lines 4 and 5 of page 13821, I think it would be important to note that a core microbiome in the South African subsurface has been reported *(Magnabosco, Cara, Memory Tekere, Maggie CY Lau, Borja Linage, Olukayode Kuloyo, Mariana Erasmus, Errol Cason et al. "Comparisons of the composition and biogeographic distribution of the bacterial communities occupying South African thermal springs with those inhabiting deep subsurface fracture water." Frontiers in microbiology 5 (2014))*. This may also serve as an interesting paper for further comparison as the same region (V6) of the 16S was sequenced
- Line 21 of page 13821 again, as written it is slightly misleading. See section Abstract, last comment

3.3 Methods

- 2.4: I think the use cell counts and qPCR were a nice addition to the paper. I'm curious as to why primers were not V6 for the qPCR as it was the primer used in amplification
- 2.6: What quality filtering method was used? When talking about the rare biosphere it is important to not that the QC step can greatly influence the number of taxa and size of the rare biosphere. See *(Huse, Susan M., David Mark Welch,*

Hilary G. Morrison, and Mitchell L. Sogin. "Ironing out the wrinkles in the rare biosphere through improved OTU clustering." Environmental microbiology 12, no. 7 (2010): 1889-1898.) and *(Eren, A. Murat, Joseph H. Vineis, Hilary G. Morrison, and Mitchell L. Sogin. "A filtering method to generate high quality short reads using Illumina paired-end technology." (2013): e66643.)* for more information. Typically, is suggested to use a 100% overlap when working with sequences form the V6 region (Eren et al. 2013).

- 2.7: Concerns are listed in **Major Comment (1)**. Please also specify if any normalization was performed and correct for multiple hypothesis testing

3.4 Results

- 3.2: Sequences statistics
  - Lines 14-20 on page 13828 (Chao and ACE) are difficult to follow, please clarify
  - A similarity index between samples would be helpful. It is somewhat illustrated in figure 3 but it would be interesting to get a sense of how many taxa are shared vs how many taxa are present. A visualization of tables 3 and 4 would achieve a similar objective
  - Can add a citation about the difficulty in assigning taxonomy to short sequences (lines 5-6 page 13829). GAST (Huse, Susan M., Les Dethlefsen, Julie A. Huber, D. Mark Welch, David A. Relman, and Mitchell L. Sogin. "Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing." PLoS Genet 4, no. 11 (2008): e1000255.) is a tool that has been used minimize this problem and may serve as a useful citation
- 3.3: Core communities

- The abstract is misleading, there is in fact a core community that makes up an extremely high proportion of the dataset. Please address
- Several of the genera are "TaxaX, Other" which is not very informative and may inflate the number of "shared". What about the number of shared OTUs vs not shared? This is where a similarity index or a visualization of the taxonomic data may be useful

- 3.4: Impact of
  - Again see **Major Comment (1)**
  - It would be helpful to have a better sense of the taxonomic composition and environmental parameters of the samples before getting into the discussion of the CCA. Sections describing these 2 data types would be helpful earlier in the results section
  - As in 3.3: Core Communities, you report a lot of Actinobacteria/Other, Gammaproteobacteria/Other, etc which I, as a reader interpret as unclassified species within the aforementioned phylum or class. This encompasses quite a lot of diversity. It would be helpful for you to describe "Other" category more completely and state exactly what it means. Points to think about: How might this larger group of Gammaproteobacteria/Other (i.e. how many OTUs fit this classification) influence your interpretations when compared to Gammaproteobacteria/Shewanella

- 3.5: Co-occurrence network
  - see **Major Comment (1)**

- 3.6: Predicted metabolic
  - cite PICRUSt line 13 page 13831

  - See **Major Comment (2)**
  - Line 20 page 13831: has PICRUSt revealed distinguishable features between environmental datasets before? Are the 16S datasets rather similar → similar metabolic predictions? Again this is where some sort of similarity index would be helpful for the larger interpretation.
  - The features listed (membrane transport, carbohydrate metabolism, glycolysis) are not necessarily distinguishing features of environments. Maybe there is a difference between predicted metabolisms of the core community and rare?
  - Table 9 is hard to read, I think it can be placed in the supplement

3.5 Discussion

- Lines 1-6 on page 13833 was difficult to follow
- Line 10 on page 13833 "genera, respectively seem low". Have you considered the number of genera within the GreenGenes reference dataset? This may deflate the number of observed genera.
- Tables 5-8 could be in the supplement
- Which sulfur and sulfate reducers are present? Who are the archaea? (lines 13-17 page 13833)
- 4.1: Energy Metabolism
  - line 14 and 25 page 13834: It is important to note that mixotrophy (ability to shift between autotrophy and heterotrophy) has also been suggested to be an important option for low-energy, subsurface systems. See:

1. Moser, Duane P., Thomas M. Gihring, Fred J. Brockman, James K. Fredrickson, David L. Balkwill, Michael E. Dollhopf, Barbara Sherwood Lollar et al. “Desulfotomaculum and Methanobacterium spp. dominate a 4-to 5-kilometer-deep fault.” Applied and Environmental Microbiology 71, no. 12 (2005): 8773-8783.
2. Magnabosco, Cara, Kathleen Ryan, Maggie CY Lau, Olukayode Kuloyo, Barbara Sherwood Lollar, Thomas L. Kieft, Esta van Heerden, and Tullis C. Onstott.“A metagenomic window into carbon metabolism at 3 km depth in Precambrian continental crust.” The ISME journal (2015).
3. Osburn, Magdalena R., Douglas E. LaRowe, Lily M. Momper, and Jan P. Amend. “Chemolithotrophy in the continental deep subsurface: Sanford Underground Research Facility (SURF), USA.” Frontiers in microbiology 5 (2014). –a potential for both chemolithotrophy and heterotrophy (not necessarily within the same organisms) in the Sanford Underground Research Facility is reported here

– Line 4 page 13835, my concern about PICRUSt, methanogens, and the Wood-Ljungdahl pathway are summarized in (Marjor Comment 2)
– I think it would be helpful for the discussion to include more about how their results compare to Purkamo et al. (2015) and other subsurface sites. For example, a study by Itävaara et al sampled the Fennoscandian Shield at various depths saw a marked change in community composition with depth. It would be interesting to see how this study compares since they are from what seem to be a similar locality *(Itävaara, Merja, Mari Nyyssönen, Anu Kapanen, Aura Nousiainen, Lasse Ahonen, and Ilmo Kukkonen. “Characterization of bacterial diversity to a depth of 1500 m in the Outokumpu deep borehole, Fennoscandian Shield.” FEMS microbiology ecology 77, no. 2 (2011): 295-309.)*
– Another point of comparison are the whole genome metagenomes from Out-

okumpu by Nyyssönen et al 2014. Are the same metabolisms predicted by PICRUSt as identified in the metagenomes? *(Nyyssönen, Mari, Jenni Hultman, Lasse Ahonen, Ilmo Kukkonen, Lars Paulin, Pia Laine, Merja Itävaara, and Petri Auvinen. "Taxonomically and functionally diverse microbial communities in deep crystalline rocks of the Fennoscandian shield." The ISME journal 8, no. 1 (2014): 126-138.)*
– And there are several other useful and interesting citations for comparison on page 13832 lines 16-18 as well as those listed in the first and second comments of this section and

1. Dong, Yiran, Charu Gupta Kumar, Nicholas Chia, PanâĂŘJun Kim, Philip A. Miller, Nathan D. Price, Isaac KO Cann et al.“Halomonas sulfidaerisâĂŘdominated microbial community inhabits a 1.8 kmâĂŘdeep subsurface Cambrian Sandstone reservoir.” Environmental microbiology 16, no. 6 (2014): 1695-1708.
2. Fukuda, Akari, Hiroki Hagiwara, Toyoho Ishimura, Mariko Kouduka, Seiichiro Ioka, Yuki Amano, Urumu Tsunogai, Yohey Suzuki, and Takashi Mizuno. “Geomicrobiological properties of ultra-deep granitic groundwater from the Mizunami Underground Research Laboratory (MIU), central Japan.” Microbial ecology 60, no. 1 (2010): 214-225.

• 4.2-4.5

– See **Major Comment (2)** and comments on **Section 4.1: Energy Metabolism**

4 Tables and Figures

• Table 3: This table is hard to follow (especially when trying to make comparisons between depths as a whole), can you present the data another way?

- Tables 5-9 should be in the supplement. Also, tables 5-8 will change once correcting for the false discovery rate.
- Figure 2: Can you reorder the legend colors by depth? This can also be in the supplement
- Figure 3: The text is a bit small and hard to read. Perhaps you can reduce the number geochemical parameters on the triplot to those that you deem most interesting. Discussion on normalization and how bacteria and archaea entries were input into the community matrix prior to CCA is also needed.
- Figure 4: Again, I must stress that the authors need to correct how they define a "significant correlation". I suspect the authors use the $p < 0.001$ cutoff in the figure instead of the $p < 0.01$ cutoff described in the methods section because there were too many "significant" edges to visualize.

Interactive comment on Biogeosciences Discuss., 12, 13819, 2015.