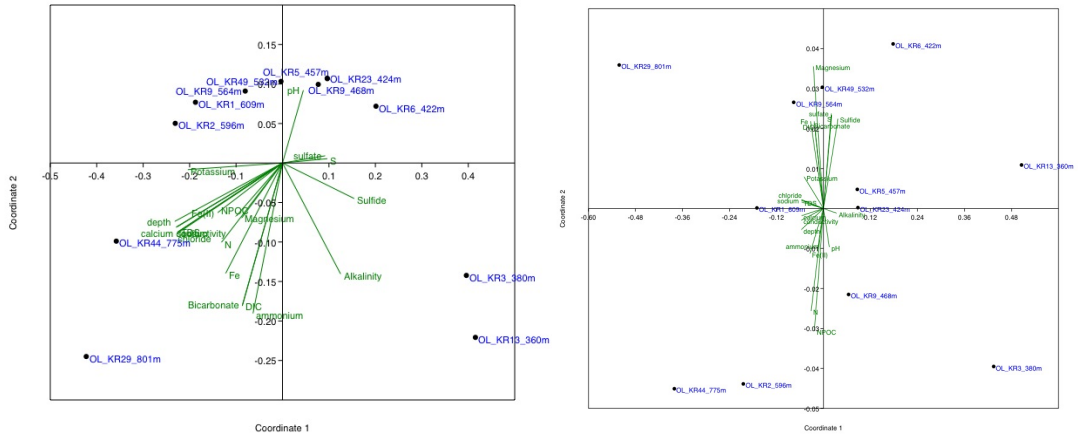
Reviewer 2

1. The data was tested again. Normality tests (Shapiro-Wilkins, Anderson-Darling) were run on the chemical and physical parameters as well as on the taxonomical data obtained from the samples. Only parameters and taxa for which the null hypothesis could be rejected with both the S-W and A-D tests when p(normal) and p(Montecarlo)< 0.05 were included in the subsequent test. One-way ANOVA (Levine's test for homogeneity of variance from means and from medians) and Kruskal-Wallis test indicated significant differences between samples (p=0). Correlation between physicochemical parameters and bacterial and archaeal taxa was tested with the Mann-Whithney pairwise test and the p-values were corrected using the Bonferroni method. As the reviewer well predicted we ended up with no significant correlations. After discussions with several knowledgeable statisticians we have come to the conclusion that we cannot apply Pearson correlations to out data matrix. Thus we are forced to remove these analyses from the paper.

Instead of these tests we have performed a non-metric multidimensional scaling test on the archaeal and bacterial communities vs. the environmental parameters. The archaeal data is presented in the left plot and the bacterial in the right plot.



2. PICRUSt provides an estimation based on the data present in the PICRUSt database. Only well-characterized and whole-genome sequenced microbial species are present. For example, the ANME-2D representative Candidatus Methanoperedens nitroreducens is not yet included. However, representatives of most methanogenic clusters are present. At least the M. nitroreducens has been shown to have the genes for the Wood-Ljungdal pathway, but when looking at the C fixation pathways of other methanogens in KEGG, they don't appear to have the whole WL-patway. Most seem to have the carbon monoxide dehydrogenase and the acetyl-CoA decarboxylase/synthase, but these enzymes are also involved in the methanogenesis from CO2 and acetate. In fact, it appears that KEGG shows the bacterial form of the WL pathway, where the CO2 is refuced to formate, and the archaeal form where CO2 is reduced to CO and a methyl group that goes into the methanogenesis, in the methanogenesis pathway. The discussion is based on the results obtained from our PICRUSt analyses, but we will alter the discussion

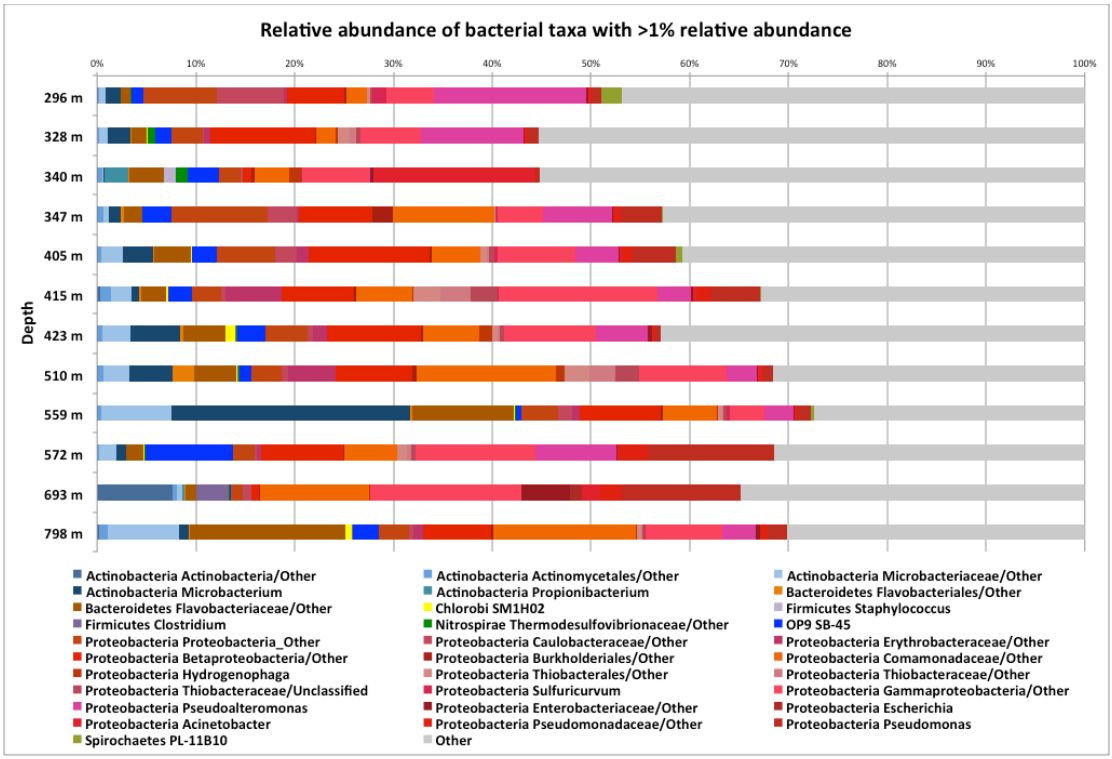
to concider that the archaeal and bacterial WL carbon fixation pathways are different. We will also specifically look for the presence of carbon monoxide dehydrogenase and the acetyl-CoA decarboxylase/synthase in the archaeal community, even though the whole WL-pathway may not be present. The discussion about the correspondence of the microbial groups to environmental factors, such as sulphate and sulfur, will be revised based on the new analyses. The correspondence analyses will be omitted, since they can't after correction of the p-values be considered significant anymore. We will compare our results to other published data fro the Fennoscandian shield. However, there are no published metagenomes from the Olkiluoto site, only from Outokumpu (Nyyssönen et al., 2014). We will compare our estimations to metagenomes from the Swedish part of the Fennoscandian shield (Wu et al., 2015), but as far as we have studied different Fennoscandian Shield sites so far they are all interestingly very different from each other.

Since a lot of space will be freed when the correspondence tables are removed we can show more pathway maps. For example, the sulfur metabolism (as brought up by the reviewer). The taxa in all the sites are presented in the supplementary tables. We can look for the roles of specific taxa in different metabolic pathways and we can compare the core vs the rare biosphere metabolisms. Here I think it might be more important to focus on the core groups as they represent the biggest part of the community.

3. Minor comments
3.1 Abstract – this will be revised. We mean 95 and 99% of the total number of sequences obtained from the bacterial and archaeal communities, respectively. The significance statement will be revised. The discussion about the rare biosphere will be revised as suggested.
3.2 Introduction – these are good points and will be included in the introduction
3.3 Methods – **2.4:** we have used the 'qPCR-primers' previously in our work and used for sequencing on the 454 platform. Based on these older results the primers have been deemed quite specific and to detect bacteria and archaea broadly. In addition, the archaeal qPCR was a bit tricky and we tested several different primer pairs in order to find the one that worked most reliably. The sequencing for this work was not done by us, but by the Census of Deep Life collaboration and in the CoDL the method has been standardized for the v6 region. We also wanted a slightly longer fragment that that produced by the v6 primers. The v6 primers used also consisted of a mix of primers, and we wanted to use only one primer/direction. **2.6:** The fastq files were combined in mothur using default parameters and the resulting fasta files were screened with QIIME allowing for no errors in barcodes (primers were removed by mothur). **2.7:** these concerns are addressed above.
3.4 Results – **3.2:** Chao1 and ACE results will be clarified. Similarity indices for the samples have been calculated and will be presented. The tables 3 and 4 have been visualized as presented below. Maybe one of these would work?

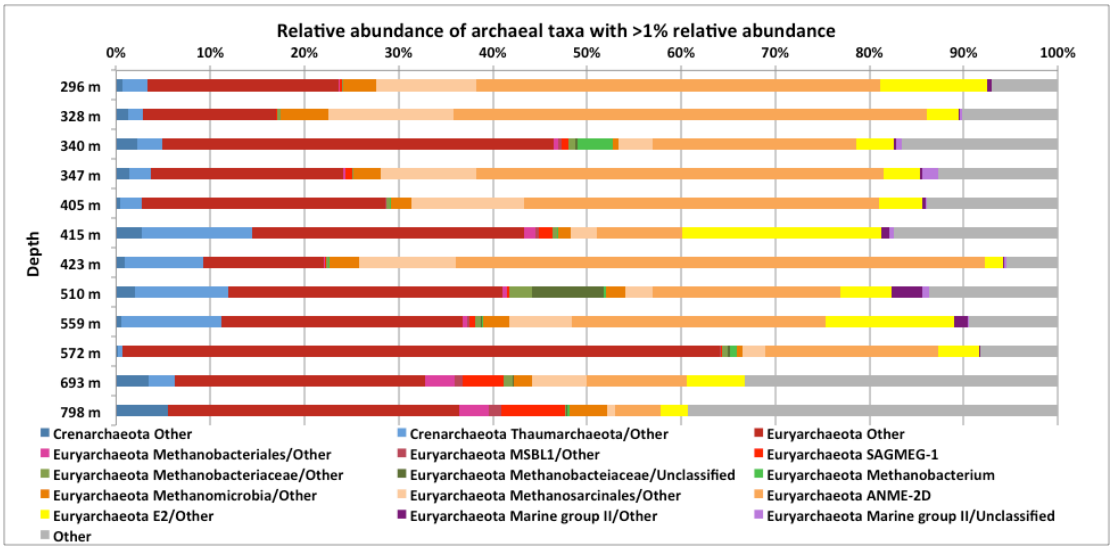
Relative abundance of the most common bacterial taxa:

		296 m	328 m	340 m	347 m	405 m	415 m	423 m	510 m	559 m	572 m	693 m	798 m
Actinobacteria	Actinobacteria/Other	0.04%	0.02%	0.09%	0.05%	0.09%	0.34%	0.09%	0.10%	0.09%	0.04%	7.63%	0.18%
	Actinomycetales/Other	0.17%	0.14%	0.45%	0.63%	0.28%	1.09%	0.43%	0.52%	0.35%	0.22%	0.41%	0.85%
	Microbacteriaceae/Other	0.63%	0.87%	0.10%	0.50%	2.27%	2.03%	2.88%	2.69%	7.07%	1.68%	0.61%	7.27%
	Microbacterium	1.56%	2.32%	0.10%	1.17%	3.08%	0.79%	5.01%	4.35%	24.23%	1.01%	0.02%	0.97%
	Propionibacterium	0.00%	0.00%	2.44%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.21%	0.00%
Bacteroidetes	Flavobacteriales/Other	0.05%	0.08%	0.09%	0.39%	0.05%	0.17%	0.31%	2.18%	0.21%	0.03%	0.06%	0.07%
	Flavobacteriaceae/Other	0.90%	1.58%	3.46%	1.81%	3.68%	2.58%	4.30%	4.24%	10.28%	1.70%	1.10%	15.76%
Chlorobi	SM1H02	0.04%	0.11%		0.01%	0.12%	0.17%	1.00%	0.16%	0.08%	0.07%		0.70%
Firmicutes	Staphylococcus	0.01%	0.01%	1.24%	0.01%	0.01%	0.01%	0.01%	0.00%	0.00%	0.01%	0.03%	0.01%
	Clostridium	0.00%	0.00%	0.01%	0.00%	0.02%	0.02%	0.02%	0.07%	0.01%	0.01%	3.27%	0.04%
Nitrospirae	Thermodesulfovibrionaceae/Other	0.08%	0.73%	1.23%	0.03%	0.04%	0.03%	0.12%	0.09%	0.06%	0.14%	0.10%	0.02%
OP9	SB-45	1.23%	1.62%	3.12%	2.92%	2.52%	2.41%	2.93%	1.25%	0.62%	8.91%	0.07%	2.66%
Proteobacteria	Proteobacteria_Other	7.43%	3.18%	2.21%	9.70%	5.81%	2.94%	4.24%	3.00%	3.70%	2.10%	1.19%	3.01%
	Caulobacteraceae/Other	6.72%	0.17%	0.23%	3.00%	2.22%	0.38%	0.51%	0.70%	1.34%	0.20%	0.96%	0.49%
	Erythrobacteraceae/Other	0.41%	0.58%	0.03%	0.21%	1.27%	5.73%	1.46%	4.75%	0.78%	0.45%	0.00%	0.99%
	Betaproteobacteria/Other	5.71%	10.64%	0.80%	7.38%	12.25%	7.35%	9.48%	7.76%	8.24%	8.39%	0.69%	6.87%
	Burkholderiales/Other	0.28%	0.09%	0.30%	2.11%	0.20%	0.22%	0.19%	0.50%	0.19%	0.05%	0.17%	0.18%
	Comamonadaceae/Other	2.08%	2.03%	3.58%	10.29%	4.89%	5.65%	5.69%	14.04%	5.44%	5.40%	11.06%	14.48%
	Hydrogenophaga	0.03%	0.14%	1.21%	0.06%	0.03%	0.12%	1.33%	0.95%	0.15%	0.01%	0.08%	0.11%
	Thiobacterales/Other	0.24%	1.26%	0.00%	0.04%	0.62%	2.78%	0.59%	2.46%	0.44%	0.95%	0.00%	0.38%
	Thiobacteraceae/Other	0.06%	0.59%	0.00%	0.03%	0.19%	3.05%	0.19%	2.65%	0.13%	0.48%		0.14%
	Thiobacteraceae/Unclassified	0.19%	0.45%	0.00%	0.04%	0.52%	2.75%	0.41%	2.38%	0.35%	0.36%		0.28%
	Sulfuricurvum	1.44%	0.02%	0.07%	0.16%	0.41%	0.04%	0.05%	0.04%	0.29%	0.01%	0.02%	0.03%
	Gammaproteobacteria/Other	4.84%	6.12%	6.78%	4.56%	7.82%	16.08%	9.28%	8.79%	3.49%	12.17%	15.27%	7.84%
	Pseudoalteromonas	15.34%	10.42%	0.13%	7.08%	4.45%	3.44%	5.20%	3.07%	2.95%	8.22%	0.00%	3.40%
	Enterobacteriaceae/Other	0.22%	0.15%	0.31%	0.10%	0.09%	0.18%	0.49%	0.15%	0.06%	0.12%	4.98%	0.34%
	Escherichia	0.00%		0.02%		0.00%	0.00%	0.00%	0.00%			1.15%	0.00%
	Acinetobacter	0.13%	0.06%	16.27%	0.06%	0.18%	0.37%	0.10%	0.28%	0.10%	0.04%	1.89%	0.07%
	Pseudomonadaceae/Other	0.23%	0.28%	0.10%	0.69%	1.15%	1.42%	0.22%	0.26%	0.31%	2.97%	2.17%	0.76%
	Pseudomonas	1.00%	0.99%	0.49%	4.14%	4.31%	5.03%	0.53%	0.96%	1.24%	12.80%	12.06%	1.92%
Spirochaetes	PL-11B10	2.03%	0.04%	0.00%	0.09%	0.63%	0.05%	0.07%	0.05%	0.36%	0.02%		0.05%
Other		46.89%	55.30%	55.14%	42.73%	40.79%	32.78%	42.89%	31.54%	27.41%	31.45%	34.81%	30.11%



Relative abundance of the most common archaeal taxa:

		296 m	328 m	340 m	347 m	405 m	415 m	423 m	510 m	559 m	572 m	693 m	798 m
Crenarchaeota	Other	0.63%	1.26%	2.30%	1.35%	0.40%	2.70%	0.89%	1.97%	0.59%	0.23%	3.46%	5.53%
	Thaumarchaeota/Other	2.64%	1.54%	2.56%	2.31%	2.31%	11.78%	8.31%	9.92%	10.59%	0.48%	2.79%	0.02%
Euryarchaeota	Other	20.35%	14.12%	41.59%	20.43%	25.85%	28.86%	12.89%	29.18%	25.60%	63.34%	26.61%	30.92%
	Methanobacteriales/Other	0.14%	0.08%	0.45%	0.21%	0.06%	1.16%	0.06%	0.39%	0.53%	0.10%	3.12%	3.08%
	MSBL1/Other	0.05%	0.02%	0.42%	0.06%	0.01%	0.43%	0.02%	0.06%	0.16%	0.07%	0.80%	1.38%
	SAGMEG-1	0.17%	0.12%	0.67%	0.67%	0.09%	1.44%	0.15%	0.29%	0.69%	0.13%	4.38%	6.71%
	Methanobacteriaceae/Other	0.11%	0.26%	0.76%	0.13%	0.31%	0.40%	0.17%	2.38%	0.57%	0.55%	0.90%	0.19%
	Methanobacteiaceae/Unclassified	0.02%	0.01%	0.30%	0.04%	0.04%	0.01%	0.05%	7.64%	0.08%	0.27%	0.20%	0.06%
	Methanobacterium	0.05%	0.07%	3.65%	0.05%	0.06%	0.13%	0.06%	0.24%	0.18%	0.70%	0.02%	0.24%
	Methanomicrobia/Other	3.41%	5.07%	0.66%	2.87%	2.25%	1.41%	3.13%	2.02%	2.81%	0.68%	1.91%	3.97%
	Methanosarcinales/Other	10.72%	13.29%	3.59%	10.14%	11.93%	2.67%	10.37%	2.87%	6.54%	2.33%	5.75%	0.88%
	ANME-2D	42.90%	50.29%	21.65%	43.21%	37.68%	9.09%	56.11%	19.99%	27.00%	18.45%	10.62%	4.88%
	E2/Other	11.37%	3.34%	3.95%	3.83%	4.59%	21.23%	1.94%	5.40%	13.61%	4.36%	6.16%	2.84%
	Marine group II/Other	0.38%	0.15%	0.25%	0.35%	0.38%	0.75%	0.15%	3.31%	1.53%	0.05%		
	Marine group II/Unclassified	0.02%	0.22%	0.69%	1.60%	0.14%	0.49%	0.23%	0.71%	0.05%	0.07%		0.01%
Other		7.05%	10.15%	16.52%	12.74%	13.90%	17.45%	5.45%	13.64%	9.46%	8.19%	33.28%	39.28%



The proposed reference to Huse et al. (2008) will be added and discussed.

3.3: In the abstract, we mean that a core community has not been identified before now. This will be addressed both here and in the abstract. The TaxaX, Other is a problem. However, these 'Others' is what the GG database gives for many of the more unknown groups. Similarity indices based on OTUs have been calculated and will be presented in the paper. I will attempt to visualize the shared and not shared OTUs in a heatmap representation of the data. Challenging with the number of OTUs obtained.

3.4: These comments will be addressed and the text edited as suggested. A closer look into the 'Other' will be taken.

3.5: The co-occurrence network will be recalculated and a new figure presented.

3.6: Reference will be added. PICRUSt is rather new and has not been used for many environmental studies yet. However, we tried in on a bog community where Acidobacteria were prevailing and the results were different from the ones we got here. Staley et al found that the PICRUSt performed quite well on riverine microbiomes when 16S rRNA gene data was compared with

metagenomic data, but they did call for caution in the interpretation of the PICRUSt results.
- Staley C, Gould TJ, Wang P, Phillips J, Cotner JB, Sadowsky MJ. Core functional traits of bacterial communities in the Upper Mississippi River show limited variation in response to land cover. *Frontiers in Microbiology*. 2014;5:414. doi:10.3389/fmicb.2014.00414.

We have employed PICRUSt to another deep bedrock dataset and will compare the results.

The listed features are those that were the most common. The presentation could be altered to show only the important metabolic cycles and leave out the membrane transport etc. We will test the core vs rare metabolisms. The table 9 can be removed to the supplements.

3.5 – Discussion
L1-6 on P13833 will be revised
L10 – we used the GG reference for the taxonomic assignments, so this is certainly a possibility. The GG was used, because PICRUSt is not compatible with other reference databases.
Tables 5-8 will be removed.
L13-17 – will be revised according to the new analyses
4.5 – mixotrophy will be included in the discussion and references added
L4, P13835 – again, PICRUSt is only a prediction that is based on the KEGG pathways. In the KEGG pathway maps on different methanogens the full WL is not shown. On the other hand, in KEGG the WL pathway is presented as it is in the bacteria, where CO2 is reduced to formate while in the archaea (methanogens) the CO2 is reduced to CO and to a methyl group bound to tetrahydropterin (in Berg et al., 2010).
-We will include the Outokumpu references in the discussion. The Outokumpu borehole was sampled to a depth of 2.5km while our Olkiluoto data is more focused on depths 300-800 m, which does not give the same depth perspective as in Outokumpu. We have also focused on specific fracture zones in Olkiluoto, while the Outokumpu samples are obtained from the water column spanning the whole borehole. A new paper is recently published where Outokumpu fracture zones have been investigated by 454 amplicon sequencing, and this will be the best paper for comparison with our results. It wasnot yet available when this paper was submitted.

Outokumpu is very different from Olkiluoto and only one of the metagenomes from Outokumpu is from a depth close to those examined in out paper. Nevertheless, we will compare the results and include them in the discussion.

The suggested papers have not been included in the discussion because despite the deep subsurface environments, the sites are still very different. However, we will strive to include these papers.

4.2-4.5: the text will be revised based on reviewer's comments.

Table3-4; presented as figures above.
Figure 2 – the legends will be reordered and the figure placed in supplements
Figure 3 – the figure will be revised and presented as NMDS plots for bacteria and archaea separately
Figure 4: an new network figure will be presented.