We would like to thank Reviewer #1 for her/his time, effort, and valuable comments. We have prepared a response taking into account all the points raised, as described below. We show the reviewer's comments in bold, while our responses are formatted as standard text. Line numbers refer to the original manuscript.

The manuscript by Krauze et al. examined microbial communities found within recently deglaciated cryosols in Antarctica in order to couple dominant pedogenic processes with microbial community structure. This is an exciting topic within the scope of Biogeosciences and the authors have identified an interesting model system for their study. The manuscript was well written with clear language and was presented nicely as well. Unfortunately, serious flaws within the experimental design and methodology employed for assessing microbial communities have effectively prevented the authors from being able to draw any meaningful, scientifically robust conclusions regarding the microbiota within their system. The first major problem is the experimental design. There was no experimental replication at any site and because the entire study consisted of four cores and four depths were examined for each core, there were only 16 samples in total. This would not be a problem in a simple system, but this is a complex system with many variables to account for. As soon as a simple factor is taken into account, depth for instance, the effective sample size decreases. Comparisons involving only the top layers, bottom layers or single cores are restricted to four observations. This number is important because regression typically requires five or more observations in order to be able to calculate significance values and a statistical evaluation of regression would allow statements such as: With depth, pH increased (paraphrased from lines 177-178), "The Shannon index showed a decreasing trend in diversity with depth" (line 197), "With depth, the relative abundances of Gemmatimonadetes and Actinobacteria increased, while the relative abundance of Bacteroidetes and Verrucomicrobia decreased" (lines 205-206), "the microbial communities became less diverse and more similar with increasing depth across all investigated soil profiles" (lines 258-259). Without statistics backing up these statements, they constitute opinions, not evidence. I will concede that it is typical to report a handful of opinions and observations, particularly when there is a visual indication of a relationship despite the lack of statistical support. However a close examination of this manuscript reveals that apparently no statements were evaluated using any sort of statistical test. This is not typical or acceptable. As a reader of this manuscript, I don't know what has been evaluated statistically and therefore I don't know what constitutes a scientific statement on the microbial community.

We agree that additional biological replicates would increase the significance of our observations. With technical triplicates of every sample for the microbiological methods, we tried to cover spatial heterogeneity of the investigated soils, which worked out well (see. Fig. 4). Anyway, to account for the low number of replications any conclusion drawn from our statistics was formulated in a careful fashion in our manuscript. In order not to be limited by the quantity of replicates, we did not compare or made statements regarding single sites or depth increments.

The statements mentioned in your comment are related to changes we observed across the whole data set (e.g. depth-dependent changes in pH or microbial diversity) taking into account 16 independent samples. In our opinion, this number of replicates allows for the statements we made. The same should be true for the NMDS, which highlights environmental factors shaping the microbial communities (Fig. 4).

The second problem I identified involved specific methodological choices in the microbial community processing / analysis. I find it difficult to accept that 100% of the microbial community in any cryosol could be classified. This is likely a direct result of the OTU picking method, which involved mapping sequence data to a database. Results like this indicate that the entire community is not being reported. Instead the current analysis compares proportions of classified data, which may or may not represent the community accurately. To know for sure, it is necessary to know how much of the microbial community is "missing" in the analysis because it did not map to the Silva database. It is often helpful to track the fate of raw data, which requires some additional accounting. The authors should list each dataset in a supplemental table. This table should include, for each individual dataset: SRA object ID, sample ID, number of reads obtained prior to any QC (raw data), number of reads that successfully merged, number of reads that remained after trimmomatic, reads that remained after chimera removal and finally, reads that were classified into OTUs which is the final number of reads that were used for analysis.

Thank you for your suggestion. A table with the relevant information in the supplement was added (Tab. S4) and referred to in the beginning of "3.2 Characterization and quantification of the microbial communities" in the Results section.

Originally, completely unassigned OTUs have been identified, but were very low in their abundance and summarized with other low abundant phyla under the term "Others" in Figure 3. Our data was reanalysed using the amplicon sequence variant (ASV) approach (see the last comment). Similar to the OTU analysis, this approach resulted in low abundances of completely unassigned sequences (minimum: 0.01% in KGI A 10 – 20 cm; maximum: 0.28% in KGI A 0 – 1 cm). Low abundances of unassigned reads in Antarctic soils are common in recent literature (0.28 % of the total data set of Meier et al., 2019; 1.1 % of the total data set in Kim et al., 2019). The very low abundance of unassigned sequences does not mean that an identification to genus/species level for the remaining sequences was possible, though. Many of the OTUs/ASVs shown in Fig. 5, which represent the most abundant reads in the data set, were just classified to the order or family level.

Negative PCR controls should always be included in such a table. Actually, I was surprised and somewhat concerned that there was no mention of negative controls. The method of Meier, 2019 used a total of 30 PCR cycles, which is quite high but understandable if the DNA yields are low, which apparently was the case for some samples (lines 151-152). Given the combination of high number of PCR cycles and low biomass, negative controls are absolutely necessary. Table 2 indicates that the deeper soils have much less template for PCR and the authors also concluded that the deeper soils are more similar to one another (line 259). If the deeper samples resemble negative controls due to having a small amount of template, it should come as no surprise that they resemble one another. This very real possibility needs to be ruled out.

A negative control was of course analysed, but not mentioned in the manuscript. The extraction included a negative control, which was afterwards handled like a sample: The amount of DNA was measured by using a Qbit; the extract was used as a template in the PCR; and the resulting PCR product sent for sequencing.

Relevant information was added to "2.4 Nucleic acids extraction" in the Material and Methods section:

"In addition, a negative control without any template but the material and chemicals of the extraction kit was included." (L. 152)

We made minor changes to “2.5 Illumina HiSeq-Sequencing” in the Material and Methods section:

“Total genomic extracts of each sample as well as an extraction negative control and a positive control (*Escherichia coli*) were sequenced using tagged 515F (5’-GTGCCAGCMGCCGCGGTAA-3’) and 806R (5’-GGACTACHVGGGTWTCTAAT-3’) primers after Caporaso et al. (2010).” (L. 155ff)

Additionally, we now not only refer to the sequencing data set in the ENB, but provide an ASV table (Tab. S5) in the supplement, including all analysed replicates, and both positive and negative controls. We refer to this table in “3.2 Characterization and quantification of the microbial communities” in the Results section. We hope the increased transparency is helpful for the evaluation of our sequencing data.

Looking at the data, the samples with very low biomass (and therefore very low DNA concentrations) do not resemble the community found in the negative control. The most abundant sequences in the negative control were classified as Cutibacterium, Corynebacterium, Staphylococcus, Methylophilaceae, Undibacterium, Pseudomonas and Streptococcus, which had relative abundances >5 % in the negative control. None of these taxa made up more >0.6 % of the total community after merging the triplicates of a sample. Some taxa found in the negative control occur as possible contaminants in some samples, but in very low relative abundances.

We included the negative control in a NMDS. It is located quite far from the cluster of analysed samples, including the deeper ones, and therefore differs significantly in its community composition. With this, we are confident that the deeper samples do not reflect a contamination during extraction and the following steps, but represent the actual community.

In addition, it would be quite helpful to include information on DNA extracted per gram of soil for each sample (and replicate) and which samples were extracted three times. The methods just say “samples with low yields”.

We are happy to include a table (Tab. S3) showing the amount of soil used for DNA extraction for every replicate in the supplementary material. “Samples with low yields” were highlighted. This table is referred to in “2.4 Nucleic acids extraction” in the Material and Methods section.

As a side note, I don’t understand the point of pooling, since the method of Meier 2019 used a fixed volume of genomic DNA.

As described in Meier et al. (2019), a fixed volume of DNA extract was used in the PCR, and a fixed amount of DNA of every PCR product was used for sequencing. The mention of pooling was solely referred to the pretreatment of the “samples with low yields” before molecular biological work.

Every replicate of those samples was extracted in triplicates, and the resulting extracts of a sample merged afterwards. Vacuum centrifugation reduced the volume of this pool to 50 µl. This step was necessary for the following molecular biological work (PCR, qPCR) to work.

We hope to clarify this by rewording the relevant part in “2.4 Nucleic acid extraction” in the Material and Methods section as follows:

“Sample replicates with very low DNA yields (Tab. S3) were extracted in triplicates. These extraction triplicates of a sample replicate were merged and after reducing their volume to 50 µl by vacuum centrifugation ready for the following molecular biological work.” (L. 151f)

Finally, the microbial resesarch community is largely moving beyond OTUbased analysis in favor of amplicon sequence variant (ASV) analysis. The authors may consider switching over to this high-resolution approach that in some cases can provide deep insights into microbial community structure: (https://www.bioconductor.org/packages/devel/bioc/vignettes/dada2/inst/doc/dada2-intro.html).

We agree. As recommended, we switched to the ASV analysis and reanalysed the sequencing data using this more modern approach. Additionally, the taxonomical classification was done using the latest release of the SILVA database (138.1)

To account for the new method, we have rewritten 2.6 Bioinfortmatics and statistical analysis in the Material and Methods section as follows:

“Raw sequencing data obtained by Illumina HiSeq (2 x 300 bp) were checked for quality with FastQC (Andrews et al., 2010). The data was demultiplexed by using the *make.contigs* function in Mothur (version 1.39.5; pdiff = 2, bdiff = 1, and default setting for others; Schloss et al., 2009). According to the resulting report files, a filtering step was implemented to get fastq sequence identifiers for sequences with a minimum overlap of >25 bases, maximum mismatches of <5 bases and no ambigious bases. Next, these sequences were extracted with the *filterbyname.sh* function from BBTools (Bushnell et al., 2014) from the raw paired-end fastq file. With QIIME1, sequence orientation was checked and corrected by using the *extract_barcodes.py* function and the primers were removed using the *awk* command (Caporaso et al., 2010). DADA2 was used for filtering, dereplication, chimera check, sequence merge, and amplicon sequence variants (ASV) calling (Callahan et al., 2016). The output of DADA2 was taxonomically classified by using QIIME2 (Bolyen et al., 2019) and USEARCH (Edgar, 2010) with SILVA138 (Quast et al., 2013). Resulting data were visualized using R and PAST4 (Hammer et al., 2001).” (L. 160 – 163)

All figures based on the sequencing data as well as the calculation of the diversity indices were redone based on the new ASV table. Excluding some changes in their names, the relative abundances of the presented phyla (see Fig. 3) were not affected substantially by switching over to the ASV analysis.

176 **Table S3: Amount of soil used for DNA extraction and resulting DNA concentrations.**

sample	soil used for extraction [mg]	DNA concentration [ng/µl]
KGI_A_0_1_a	557	7.13
KGI_A_0_1_b	550	8.2
KGI_A_0_1_c	530	6.87
KGI_A_1_10_a	510	0.121
KGI_A_1_10_b	564	0.213
KGI_A_1_10_c	585	0.103
KGI_A_10_20_a	1560	n. d.
KGI_A_10_20_b	1585	n. d.
KGI_A_10_20_c	1631	n. d.
KGI_A_20_40_a	1510	0.4
KGI_A_20_40_b	1614	0.511
KGI_A_20_40_c	1574	0.376
KGI_B_0_1_a	505	13.3
KGI_B_0_1_b	494	10.1
KGI_B_0_1_c	541	9.13
KGI_B_1_10_a	590	4.61
KGI_B_1_10_b	590	4.83
KGI_B_1_10_c	549	4.11
KGI_B_10_20_a	592	0.09
KGI_B_10_20_b	538	0.04
KGI_B_10_20_c	529	0.08
KGI_B_20_80_a	1679	0.07
KGI_B_20_80_b	1665	0.09
KGI_B_20_80_c	1639	0.04
KGI_C_0_1_a	485	78.4
KGI_C_0_1_b	500	69.6
KGI_C_0_1_c	550	61.2
KGI_C_1_10_a	548	18.5
KGI_C_1_10_b	544	16.5
KGI_C_1_10_c	553	12.4
KGI_C_10_20_a	530	0.8
KGI_C_10_20_b	570	1.17
KGI_C_10_20_c	565	1.13
KGI_C_20_40_a	550	0.273
KGI_C_20_40_b	570	0.428
KGI_C_20_40_c	562	0.101
KGI_D_0_3_a	552	49.6
KGI_D_0_3_b	505	49.4
KGI_D_0_3_c	585	22.6
KGI_D_3_15_a	560	5.46
KGI_D_3_15_b	514	3.87
KGI_D_3_15_c	512	4.97
KGI_D_15_27_a	562	1.35
KGI_D_15_27_b	561	1.27
KGI_D_15_27_c	586	1.37

KGI_D_27_60_a	599	0.322
KGI_D_27_60_b	550	0.337
KGI_D_27_60_c	548	0.501

177
178

179 **Table S4: Number of sequencing reads after each processing step.**

sample	input	filtered	denoised	merged	Non-chimera	0.01% cutoff
	number of reads					
KGI_A_0_1_a	823555	760184	744842	728869	722825	706168
KGI_A_0_1_b	416652	386897	378089	368092	365389	359124
KGI_A_0_1_c	923881	848435	824280	799372	786254	772324
KGI_A_1_10_a	473836	436670	427179	416676	413400	410050
KGI_A_1_10_b	234950	218204	212241	204825	203788	201502
KGI_A_1_10_c	352840	330275	323230	316343	314148	311452
KGI_A_10-20_a	36923	33865	29718	26130	25958	25922
KGI_A_10-20_b	18966	17275	14209	12268	12259	12252
KGI_A_10-20_c	25875	23729	19871	17313	17308	17300
KGI_A_20_40_a	905510	839764	828694	814373	808267	790148
KGI_A_20_40_b	846168	790425	779458	764252	758482	737591
KGI_A_20_40_c	893543	830439	821643	808578	801514	778446
KGI_B_0_1_a	802491	738909	726606	711994	705038	683957
KGI_B_0_1_b	767042	707072	685728	660207	650385	634947
KGI_B_0_1_c	776140	717207	691876	662511	644483	630913
KGI_B_1-10_a	870220	801229	775510	742168	726536	700009
KGI_B_1-10_b	763113	708362	698110	683584	679585	655095
KGI_B_1-10_c	984339	904070	867753	821445	790478	763807
KGI_B_10_20_a	902645	835484	829409	818332	809608	796113
KGI_B_10_20_b	50853	47043	41987	37904	37729	37613
KGI_B_10_20_c	45144	41598	36756	32746	32618	32557
KGI_B_20_80_a	797077	739614	734477	727607	713996	706767
KGI_B_20_80_b	666167	617867	613646	605338	594001	587915
KGI_B_20_80_c	59859	55236	50726	47120	46730	46649
KGI_C_0_1_a	748549	692351	673800	648601	635472	605915
KGI_C_0_1_b	896152	828214	784842	728938	687678	661043
KGI_C_0_1_c	780292	724237	701902	672607	650189	619970
KGI_C_1_10_a	1072129	979738	941569	893583	866371	830430
KGI_C_1_10_b	796453	734715	714128	686413	675197	651021
KGI_C_1_10_c	945868	879402	852158	818052	799307	765149
KGI_C_10_20_a	914463	847130	837834	821943	814309	788037
KGI_C_10_20_b	724047	663724	657595	646501	642849	625122
KGI_C_10_20_c	379777	348578	342785	334296	331763	325740
KGI_C_20_40_a	805325	743983	737314	727505	720592	705527
KGI_C_20_40_b	883724	818626	808175	789997	782377	762057
KGI_C_20_40_c	651443	603352	595920	583664	578353	564023
KGI_D_0_3_a	779880	715854	698512	675201	664504	626027
KGI_D_0_3_b	787246	723529	685098	643359	617163	586001
KGI_D_0_3_c	898392	826444	770393	701924	643856	619137

KGI_D_3_15_a	815335	754410	739659	718535	709893	677501
KGI_D_3_15_b	950935	871121	850341	820666	806893	766920
KGI_D_3_15_c	930601	858951	806053	732072	650658	627047
KGI_D_15_27_a	913147	851119	839448	822669	812036	784157
KGI_D_15_27_b	1074041	989909	980950	964946	953467	917368
KGI_D_15_27_c	869976	799507	791364	777160	770119	749250
KGI_D_27_60_a	740686	687100	681491	671687	666209	645648
KGI_D_27_60_b	768476	714698	707202	695134	689840	671215
KGI_D_27_60_c	852927	790621	782334	769492	763600	737767
Negative control	11045	9999	9955	9899	9899	9819
Positive control	651120	590140	589778	585268	585268	585250

References

Andrews, S.: FastQC: a quality control tool for high throughput sequence data, 2010.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., ... & Bai, Y.: Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature biotechnology, 37(8), 852-857, 2019.

Bushnell, B.: BBTools software package, 2014.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P.: DADA2: high-resolution sample inference from Illumina amplicon data. Nature methods, 13(7), 581-583, 2016.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Widmann, J., Yatsunenko, T., Zaneveld, J., and Knight, R.: QIIME allows analysis of high-throughput community sequencing data, Nature methods, 7, 335–336, 10.1038/nmeth.f.303, 2010.

Edgar, R.: Usearch, 2010

Kim, M., Lim, H. S., Hyun, C. U., Cho, A., Noh, H. J., Hong, S. G., and Kim, O. S.: Local-scale variation of soil bacterial communities in ice-free regions of maritime Antarctica, Soil Biology and Biochemistry, 133, 165-173, 2019.

Meier, L. A., Krauze, P., Prater, I., Horn, F., Schaefer, C. E. G. R., Scholten, T., Wagner, D., Mueller, C. W., and Kühn, P.: Pedogenic and microbial interrelation in initial soils under semiarid climate on James Ross Island, Antarctic Peninsula region, Biogeosciences, 16, 2481–2499, 10.5194/bg-16-2481-2019, 2019.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F. O.: The SILVA ribosomal RNA gene database project: improved data processing and web-based tools, Nucleic acids research, 41, D590-596, 10.1093/nar/gks1219, 2013.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J. and Weber, C. F.: Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and environmental microbiology, 75(23), 7537-7541, 2009.