Relation of aggregate stability and microbial diversity in an incubated sandy soil

Frederick Büks[1], Philip Rebensburg[2], Peter Lentzsch[2] and M. Kaupenjohann[1]

[1] Chair of Soil Science, Department of Ecology, Technische Universität Berlin.
[2] Leibniz Center for Agricultural Landscape Research (ZALF e.V.), Müncheberg, Germany

Correspondence to: F. Büks (frederick.bueks@tu-berlin.de)

Final response to #Referee1

Dear Referee1,
thank you very much for your important suggestions for the improvement of this work. In the following I present how to include them in a revised article.

1) The choice of microorganism sources: soil born - air-born. Does not it by definition biases things towards greater aggregation in soil-born case, because they will for sure have fungi? If the authors are truly after biofilms they should choose a more biofilm oriented set of sources.

Our objective was to investigate the influence of truly different microbial populations on aggregate stability. *Delmont et al. (2014)* recently found that the development of microbial communities is controlled by physical-chemical properties of soils rather than start population: e.g. an initial population taken from a forest soil was given on a sterile grassland soil and there developed like the original grassland population. Therefore we were forced to be careful and chose two inoculates that definitely have no potential to converge their microbial abundances. The possibility to suppress fungi by antibiotics to see the pure bacterial effect, was discarded because of their potential effect on soil DOC and other microorganisms. Therefore, fungi are part of this biofilm. This information will be included in the revised paper.

2) It is not clear where the air-born microorganisms come from. Line 195 paragraph talks about sterile air supply and line 215 paragraph states that exposure to unsterile air was done after the incubation?

“The incubation of both variants took place in columns with sterile air supply to get e.g. similar evaporation rates and hinder permanent infection. On the other hand, during sampling the air-born variant was exposed to room air to force infection, whereas the soil-born variant remained under sterile handling. After each sampling, both variants were reconnected to sterile air supply.” We are going to explicate that in the revised article.

3) Description of statistical methods is not clear. There are two factors here – two soil treatments and several sampling times. Why this is not analyzed as a two factor experiment with repeated measures? There were only three replications analyzed - how the tests for normality and equal variance could be conducted with so few data points. What is meant by "variant"?
AND
If I understand correctly, there were only three true replications of the studied systems? Given very high variability of soil aggregation data, no wonder that no statistically significant differences were observed. But the observed tendencies appear to be consistent with the authors’ hypothesis. In such cases it is strongly recommended to conduct post-hoc power analysis to address the sufficiency of the replications and the size of the differences that could be statistically detected given the observed variability and the numbers of replications used. I would strongly recommend the authors to conduct

such analysis.
We restructured our statistic analysis in matters of your suggestions:
“The statistic analysis of microbial populations and SOC release comprised calculation of mean values, standard deviations and analysis of variance. After application of the Shapiro-Wilk test *(Shapiro and Wilk, 1965)* samples were assumed to be normally distributed. T-test for unequal variances (Welch test) was used to test for significant differences of class, domain and total DNA concentrations between both incubated variants (SP_{soil} and SP_{air}) *(Ruxton, 2006)*. The Bonferroni correction was applied *(Bland and Altmann, 1995)*. Total bacterial populations were assumed to be similar, if the difference is less than 1.2 ng DNA/ mg dry soil in the variants. SOC releases of SP_{soil}, SP_{air}, SP_{pure} and A_{pure} were analyzed using one way ANOVA *(Christensen, 1996)*.”
The threshold of 1.2 ng DNA/ mg dry soil results in a narrowing of the “final period” from day 49-76 to day 51-76. This will be included to the manuscript.
Power analyses is mostly used to estimate the needed sample size, to detect an assumed effect size with a given type 2 error rate. We didn't do a post-hoc power analysis for the following reasons: In case of the taxonomic comparison between the variants SP_{soil} and SP_{air} significant differences in beta-proteobacteria, acidobacteria and fungi are given – therefore both variants are different and a type II error analysis is not necessary. The comparison of aggregate stability on the other hand showed no significant difference between variants. However, we forewent to apply a post-hoc power analysis in this case, since it only explains what is already known – the power of three parallels is low. Therefore we will add the following to the end of the discussion: “Our results give a first insight to the relation of microbial community composition, SOC release and aggregate stability. A more quantitative analysis would require more replicate samples, probably inclusion of soils from different land use and different microbial communities. However, this was beyond the scope of the present study.”

4) Results section can be shortened and a lot of things in Discussion should be moved to the Results. As of now the Results contain a lot of verbal descriptions of how numbers go up and down on the figures and this is not helpful. I would suggest to focus on bringing to reader’s attention the key trends and points of interest instead (those are present in the Discussion and should be moved to the Results – e.g., the material in l. 405 paragraph).
Thank you very much. I will do so after consulting the editors to best match their structural requirements.

5) Should show on the figures and in Table 3 when the differences are statistically significant and when they are not.
I will include the level of significance to tables and figures.

6) Ll 384-385 – unclear, please rewrite.
“Our results do not support this hypothesis, as strongly diverging microbial populations did not cause significantly different aggregate stabilities.”

References

(Bland and Altmann, 1995) Bland, J. Martin, and Douglas G. Altman. "Multiple significance tests: the Bonferroni method." Bmj 310.6973 (1995): 170.

(Christensen, 1996) Christensen, Ronald. "One-way ANOVA." Plane Answers to Complex Questions. Springer New York, 1996. 79-93.

(Delmont et al., 2014) Delmont, Tom O., et al. "Microbial community development and

unseen diversity recovery in inoculated sterile soil." Biology and Fertility of Soils 50.7 (2014): 1069-1076.

(Ruxton, 2006) Ruxton, Graeme D. "The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test." Behavioral Ecology 17.4 (2006): 688-690.

(Shapiro and Wilk, 1965) Shapiro, Samuel Sanford, and Martin B. Wilk. "An analysis of variance test for normality (complete samples)." Biometrika 52.3/4 (1965): 591-611.

Best regards,
Frederick Büks