From: ismecomms@springernature.com
Subject: Decision on ISMECOMMS-23-00126A-T
Date: July 18, 2023 at 4:34 AM
To: isabelle.laforest.lapointe@gmail.com
Cc: hauke.smidt@wur.nl, kirsten.kuesel@uni-jena.de

Manuscript Number: ISMECOMMS-23-00126A-T
Title: "Dynamics of moss-microbial community assembly and functional capacity in boreal forests"

Dear Professor Laforest-Lapointe,

Thank you for submitting your manuscript to ISME Communications.

As you will see from the reports included below, your manuscript has been reviewed by our editorial board and external reviewers. Upon careful consideration, your manuscript has been rejected.

The journal receives many more manuscripts than we can publish, and as a result we are forced to reject many submissions that are of merit but we feel are a little short of the excellence for which we are striving at ISME Communications.
More specifically, reviewers consistently expressed concerns regarding:
-the lack of a clear motivation,
-inadequate sample size in the natural site vs. mined site analysis (minimum sample size of 2, which is not sufficient to account for the spatial variation)
- lack of justification for the bioinformatic approach used which is not very common in environmental metagenomics

We appreciate your interest in ISME Communications and hope that you will consider us for your exciting work in the future.

Sincerely,
Hauke Smidt
Editor in Chief
ISME Communications

Kirsten Küsel
Editor
ISME Communications

Comments from reviewer(s): Referee #1 (Comments to the Author):

The manuscript by Ishak et al is a metagenomic study on the drivers of microbial communities and their functions in four species of moss occurring in boreal forests. The justification for this work is that many of these mosses have been neglected as most moss-associated microbiome work is focused on Sphagnum spp. Furthermore, the use of metagenomics in a field where most microbial ecology research have used only amplicon sequencing presents an advance to the field and some of the results are very interesting.

However, I have a few concerns regarding the robustness of some of the results presented here.

The authors address three main aims: 1) to characterize microbial communities and their function, 2) study moss species and moss section as drivers of moss associated microbial communities and 3) determine the impact of mining/post mining sites and moss associated microbes.

Aims 1 and 2 is addressed by studying 6 samples collected in September 2021, while aim no. 3 (mine vs natural sites) is studied with 10 samples collected in June. When these 10 samples are broken down to what was collected, the effective sample size for some conditions is effectively 2, which is not enough for robust statistical analysis (see fig. 3d as example). The manuscript will be better without this part. Unfortunately, a large part of the discussion is also dedicated to this analysis and I personally believe these results are not robust.

Further comments

The introduction section provides a lot of interesting background information in paragraphs that were sometimes very long and tedious to read. For example, paragraph 2 of the introduction starts with a sentence on new insights on possible drivers of moss-associated microbial assemblages, but a few sentences later lists very specific functions such as quorum sensing, pathogen antagonism and biofilm formation is mentioned – hence very different subjects which lack of a clear topic making it hard to read. The same goes for the paragraph starting at L69 where biotic and abiotic drivers are mentioned. The paragraph is very long and the topics in this paragraph are quite different from each other. The green and brown sections are mentioned here for the first time in the introduction, but it really needs more attention since it is a big part of the story (and some of the most interesting results).

Overall, the introduction could be better organised with single clear themes for each paragraph.

The term "functional portrait" sounds strange – how about functional profiles or functional perspective (which you also use in the paper sounds much better).

Methods

Methods

L145 – I had to go to the supplementary information to find out how many samples of the different mosses were collected. A bit more detail is required here in the main text. I refer to the text: "the green sections of a total of 10 individuals of Polytrichum juniperinum and Tortella tortuosa were sampled from a natural and a mine site (Table S2)."
So this was 10 in total? In essence, in the natural area 6 samples were collected from P. juniperum (n=3) and T. tortuosa (n=3), and at the mine site the same species but two individuals of each species which means n= 2 x 2. More clarity is needed in the main text, because it lends a lot of context and credibility to the results. Two samples in a metagenomic study in a single treatment may be too little to confidently capture the variation and error in the models later used. Furthermore, the design is quite unbalanced, since the number of samples collected in June and September differs substantially.

L165 -> 10 and 165 metagenomic libraries were generated, but I don't gather from this sentence what the 10 and 165 refers to.

Bioinformatic methods

The authors made use of a read-based approach to assign taxonomy to their sequencing data. This was followed by producing reference genomic data using ChocoPhlan and the UniRef database. To my knowledge, these methods are more commonly used in human microbiome metagenomics studies and not so common in soil and other environmental metagenomics. However, I think they are fine as long as its use is justified over that of conventional assembly, annotation and mapping. One or two sentences at the beginning of this section is required to justify this approach. My first thought is "why no assemblies?" and many readers will ask the same.

Was the default Kraken database used and does it have adequate coverage for environmental fungi?

For the taxonomy analysis, I gather that in some cases samples were rarefied to as low as 260 000 reads (Table S3). If a minimum of 10 million reads (L169) were obtained, why were these rarefied to such a low number? Did you lose that many when filtering host DNA?

L205-206: community composition and assembly used interchangeably? Which is it, because these terms have different meaings?

L225 – if fungi made up 1% of all taxa, is it worth mentioning the Saccharomycetes at all?

L240 – 250 – It would be great to know if these KOs were from cyanobacteria (or what other taxon),– even at phylum level

Were the N-fixing genes the only difference found and was it only nifU and none of the other nif genes? If the taxonomic context could be added, the authors could comment on the possible loss of cyanobacteria associated with brown colouring and loss of N fixing genes?

L274 – 284 – These results are interesting, but I am sceptical of the few samples taken at the mine sites.

L297 – Yes, there were some KOs that were different, but do these really constitute whole "pathways"?

L338- without knowing what they are this statement does not mean much. Many (or most KOs) are associated with non-microbial functions are merely misannotated.

L420 - A single nif gene was differentially abundant between the green and brown sections. Could the authors comment whether one of many nif genes is enough to "demonstrate nitrogen metabolic genes' abundances differ between moss green and brown sections". How about the other nif genes required for N-fixation?

L421 – "in mines" sounds like samples come from inside a mine.
L422 – except for the heat shock protein found, there was very little evidence of this.

Mine site analysis – even though the glm/lm used in the analysis accounts for unbiased designs, I have my doubts whether this result is robust given that n=2 (or n=4, combining the two species). The difference in data points in fig 3 b and d also makes this clear.

Discussion

The discussion section is interesting, and the authors provide interesting insights into the results. However, much emphasis is placed on the natural sites vs. mined sites analysis, of which I believe the results are not robust due to the very few samples used in this analysis.

Again, if all fungi made less than 1% of the total community, is it worth mentioning?

Paragraph startign at L329 is very short – I understand that the hypothesis around Cyanobacteria in brown sections is very important, thus the authors could elaborate more beyond the possibility of relic DNA being present.

Figures and tables
Fig 1a – The map is very blurry? Could it be improved?
Fig 2 – Looking at fig 2 I am struck again that the sampling and two main questions in this study is quite unbalanced. Fig 2a+c (moss species and sections) is very interesting and impressive, but in comparison Fig2b+d on the right looks sparse with very few samples.
Fig 3 – Same as above. I find the results interesting, but the number of samples in Fig 3 b+d is an issue, especially looking at T. tortuosa in mined sites

Fig 5 – this fig is quite hard to read with all different size scales for every subplot. Maybe some of these can be reorganized and combined to create a clearer impression of the results. The colour legend in the middle is also a bit confusing.

Referee #2 (Comments to the Author):

Review report for "Dynamics of moss-microbial community assembly and functional capacity in boreal forests" by Ishak et al.

In the manuscript (MS), the authors aimed to identify microbial community composition as well as their functions associated with mosses – both in green and brown parts of the mosses. Differences in community composition as well as microbial diversity between the moss sections were found (with e.g. brown moss sections having a higher diversity). Further, the authors collected moss samples near a mine and found different communities and functions in those samples compared to samples collected in a natural forest site.

This MS deals with an interesting study system, and the authors use novel techniques to address open questions. However, I have some comments for the authors to consider:

Title:
The title is misleading as you do not assess "dynamics" of microbial communities on mosses. This would imply a temporal change / assessment. You do collect mosses at 2 time points but these are analysed separately and not linked.

Abstract:
L. 19: why do you mention leaf litter? There are so many factors that control microbial communities on mosses, why do you mention this one? You have not tested this, and you could mention any factor
L. 24: What is the link between the 5 and the 2 species? Why do you separate those? Did you not use the same Polytrichum species for the mine-part of the study?
L. 26: taxa such as?
L. 28-33: this bit could be improved to get a bit more interpretation of your data here, and speculate which factors drive the differences, why etc.

Introduction:
In general, the introduction is a bit too long, and one sentence is added after the other without drawing any conclusions or synthesize. Which are the main drivers, what are the mechanisms and what is still unknown?
L. 39: lichens are not plants
L. 44 and following: Why do you bring temperature into the text here? You did not look at temperature. Your introduction is already very long, and I would suggest you focus on the factors you have actually assessed here in your study
L. 61 and following: are these "functions" anything special for these bacteria? Is this not common amongst bacteria? Consider deleting or highlight some specific functions.
L. 70 and throughout: always good to write in which way (positive or negative) factors affect processes
L. 73: how and why does litter do that?
L. 74: how is moss phylogeny a driver of bacterial community structure? What is it about the phylogeny that does it? Which traits? Try to get at mechanisms (here and throughout)
L. 76: what is the problem that only 2 species have been studied? Spell it out
L. 85: this is not a good enough motivation for the study
L. 92: but we still would not know if they are actually expressing those genes, right?
L. 98: this is not a good enough motivation for the study- what do we get if we do this?
L. 102: what type of mining?
L. 104: what is a positive impact on soil?
L. 108 and following: how can you distinguish between if the characteristics you are interested in are from the site/pollution or from the moss species that has been selected for to grow in polluted sites. Is this not a confounding factor in your study? Only certain moss species can grow in certain sites (conditions). So, is it the site that drives community structure or the moss that is selected by the site...?
L. 114: This is indeed interesting. I think you should consider having this angle as a starting point and main topic of your paper. Why is a function not affected by metal pollution? Who can grow there, and what genes do they have? Also, as you hypothesize that the pollution will be the driver of community structure, it makes sense to mention this (more) and first?
L. 125: in no place in your MS do you mention what type of mine this is. And is it still in use? We need much more information on the site (mine)! This will affect your entire MS – why this mine, what is the main pollutant and what has been found before etc.?
L. 130: where does the biofilm come from? This has not been properly introduced

M&M:
Much more details are needed to be able to understand what and why you have done what you have done. E.g. why 2 sampling times and then you did different analyses on the different samples. What was the idea behind this? Also, many functions have a somewhat seasonal dynamic, with often lower nitrogen fixation later in the year.
Why did you choose these moss species? Also, you have 3 species of Polytrichum – how different do you expect them to be in terms of microbial community structure – you write earlier that phylogeny is an important driver, yet, you picked some very closely related species. I would expect traits to be very similar (chemical, morphological) that will in the end determine which microbes are there and how active they are?
L. 142: what do you mean with dynamics?
L. 171: I think there is one out for Pleurozium? https://academic.oup.com/g3journal/article/9/9/2791/6026388

L. 182 and following: could you write out which analyses have been done on which dataset to achieve which outcome?

Results:
L. 254: again, the location determines who can grow there – but what about the location drives community structure? It is not “location” or “mine” vs. “natural”. It will likely be a continues factor such as nutrient availability, light etc.

Discussion:
I would suggest that you structure your discussion according to your hypotheses-that way it may also become a bit more condensed.
L. 289: What type of land use types did you investigate?
L. 290: which other factors have you measured that could affect it? If you assessed those 3 (moss species, “site”, moss section), then they will be a significant part of the explanation. If you would have measured other factors, they may have contributed to explaining the pattern. It all depends on the factors you chose...

Referee #3 (Comments to the Author):

This is a strong paper with a particularly well-written discussion section. I have two major comments and some line comments. I appreciate that the authors plan to make their data and code freely available.

1. As far as I can tell, in your statistical models, moss species is treated as a fixed effect with 4 levels (one for each species). This model structure does not reflects the fact that three of your species are within the same genus. These species are not therefore fully "independent". In the past, papers on microbial assemblages in mosses have run analyses on the genus level or accounted for phylogenetic relatedness among species in their model structures. This point is illustrated by your results and discussion (L397-405). I understand the reasoning, but I think it undermines your claims about the role of moss species and would encourage either consideration of one of the approaches above or including more of an explanation for why this analysis structure was selected and how the results should be interpreted due to that structure.

2. There is a need in the literature to include less-studied species and I appreciate that the authors do so here. However, I think the selection of these species and the impact that may play on their results needs to be better addressed in the introduction, methods, and discussion. Why were these species selected? How do these species differ from other boreal mosses in ways that may be relevant to their microbiome? These species tend to have much lower documented rates of nitrogen fixation compared to other mosses--how might that effect the conclusions that can be drawn from the data presented here?

Line comments
39 unnecessary comma after lichens
42 Both cited sources are specifically about N2 fixation, not C sequestration, and I don’t think “facilitated” correctly describes the role that microbes play is moss-associated N2 fixation. You could consider looking at Cornelissen 2007 and Lindo 2013 for studies about the connection between moss-associated N2 fixation and C sequestration
42-44 I’m not sure if this is the correct interpretation of the Lin et al. reference, since they found that P limitation has a stronger influence over organic matter decomposition. Also other studies have found that mosses contribute more than 50% of new nitrogen to these ecosystems—usually the majority depending on N deposition rates.
47 a reference is needed for the impact of temperature increases on moss-associated nitrogen fixation
88-89 Again, I’m not sure that these are the correct references.
101 Recommend moving this paragraph up to the sections discussing climate change and moss community structures.
128 Hypothesis is unclear: are you predicting that green vs brown will be 2nd most important in microbial community assembly with species identity 3rd? Or that green vs brown and species identity will be equally informative?

141 June should be capitalized (September as well)
184 lmer is not a part of base R; if mixed models were used, you should specify the package (likely lme4 or nlme). Since mixed models were used to evaluate the claims in your hypotheses, the fixed and random effects structure should be included in the main text, not the supplementary material.

Recommended citations-For N2 fixation measurements specifically taken in different parts of moss shoots
Solheim et al., Associations between Arctic cyanobacteria and mosses (Symbiosis 37, 2004)
Leppanen et al., Nitrogen fixation and methanotrophy in forest mosses along a N deposition gradient (Environmental and Experimental Botany 90, 2013)
Gavazov et al., Isotopic analysis of cyanobacterial nitrogen fixation associated with subarctic lichen and bryophyte species (Plant and Soil 333, 2010)

**