

Reply to the comments on the manuscript acp-2020-1065:

Comment: The dataset is interesting and original. This was analyzed through statistics, looking for trends between the different size fractions and between the different variables investigated, which is rather consistent with the underlying objective to identify a signature of specific PBAP. However, the absolute values themselves are neither discussed nor positioned respect to literature. It would be interesting to have a paragraph for discussing these, independently from trends.

Response: Absolute values obtained from the chemical and microbiological species were discussed at 3.2 section. In addition, these results were compared with those found in the literature. It was observed that concentration values were slightly low in comparison with those reported our previous study (Helin et al., 2017). However, it could be explained considering the sampling period used in this research, corresponding to autumn-winter with some episodes of snow. In addition, limitations diverted from the technique used for the determination of total DNA on the determination of low DNA concentrations were considered.

Comment: The title is probably too general and somehow inappropriate, as this is more about the interrelationships between variables than about the characterization itself. Also, is there any evidence that the particles looked for are indeed exclusively primary? And biological?

Response: We have followed the suggestion of the reviewer and modified the title of the manuscript. The new title is "Chemical and microbiological components of bioaerosols in the boreal forest and their interrelationship with environmental variables".

Comment: Figure 1 and/or Tables S8-S11 could include a line with the sum of all fractions, which would thus correspond to the total aerosols load.

Response: Figure 1 and Tables S8-S11 were modified as requested.

Comment: I have a major concern with qPCR data, and these are fundamental in this study. First, there is not even a mention of the genes targeted.

Response: Thank you for this comment. We have now added information that the qPCR is targeting at bacterial 16S and fungal 18S ribosomal DNA.

Comment: Second, and most importantly, the results: values around 1-10 genes (supposedly 16S and 18S rRNA)/m³ of air are reported, indicating the presence of 10 cells/m³ at the very most, which is absolutely not consistent when the literature reports orders of magnitude higher values around 10⁵-10⁶ copies/m³ in much more remote contexts (see notably (Dommergue et al., 2019; Šantl-Temkiv et al., 2017; Tignat-Perrier et al., 2019, 2020). Even the same authors reported incomparable values in previous publication (Helin et al., 2017), so either the data themselves are not valid as they largely underestimate the actual situation, or it could be that the unit used is wrong, or again that there was a mistake in the conversion to air volumes. It would also be interesting to have indicated somewhere the cycle thresholds used for quantifications.

Response: Thank you for raising this concern. The units are lower than in our previous publication (Helin et al. Atmos. Chem. Phys., 17, 13089, 2017) and in other publications because the samples were collected in this study in September-November, and not during growing season. We also showed in our previous publication (Helin et al. Atmos. Chem. Phys., 17, 13089, 2017) that the PBAPs have seasonal variation and that in the winter samples collected in February and March the gene copy numbers were lower than in samples collected during other months. In Helin et al. 2017 publication the extracted DNA amounts were also higher than in our current manuscript. In addition, we would like to point out that in the study described in Helin et al (2017) publication, the whole filter was used for the DNA extraction, but in the present study a half of the filter was used for the extraction of free amino acids and sugars, and another half for the extraction

of DNA, and in this way we used the same sample for both chemical and microbiological analysis. Due to low amounts of template, we could unfortunately not redo the qPCR with higher amount of sample.

In this work, the observed gene copy numbers varied between 0.1-200 16S gene copies/m³ in case of bacteria, 0.05-400 18S gene copies/m³ in case of fungi, and 0.05-6 16S gene copies/m³ in case of *Pseudomonas*. However, the obtained gene copy numbers varied highly depending on the filter size and sampling time, and thus the averaged values shown in Figures 2 and 3 seem low, since they represent average gene copy numbers across all the studied filter sizes in all the studied time points. Similar time-scale variation and size segregation were observed both in the present study and in the Helin et al. 2017 publication, but this was not emphasized here due to the different focus of the manuscript

Comment: This work is basically a repeat (improved?) of that published in 2017 by Helin et al, with different approaches and added with new variables like saccharides. There are at several occasions (auto)-plagiarism of this reference in the experimental section (maybe acceptable there (?)).

Response: The techniques used for the determination of chemical compounds, microbiological species and total DNA were based on those used in our previous research published in 2017 by Helin et al. with modifications. In this study amino acids and sugars could be successfully separated in a single run by hydrophilic interaction liquid chromatography (HILIC) under 23 minutes. In addition, our present system allowed the determination of amino acids, sugars and microbes from the same collected filter sample. Further potential auto-plagiarism was deleted from experimental section.

The later reference is barely cited in the results and discussion section. However the present work would probably benefit to be positioned in context, with the findings discussed respect to previous ones.

Response: Our previous research has been positioned in context in the introduction section. In our previous study amino acids, bacteria and fungi were determined in aerosol samples collected at SMEAR II station to establish seasonal variations and size distributions. Additionally, the effect of few local meteorological factors and potential emission sources was also evaluated. Even though the observations of concentrations and distribution of different PBAPs are accumulating, there is still lack of a comprehensive understanding of the processes behind the different observations and on detailed chemical characterisation of the particles. In this new study, chemical compounds (amino acids and saccharides), microbial species (bacteria, fungi and *Pseudomonas*) and total DNA concentrations were determined and different statistical tools were used to clarify the relationship between particle size, environmental and meteorological conditions and the composition of PBAPs. In addition, potential chemical signals from microbes in aerosol particles. Finally, the potential connections between gas phase VOCs and the microbiological composition of the aerosol particles, bacterial, fungal or *Pseudomonas* gene copy numbers.

The choice of targeting in particular *Pseudomonas* among the humongous biodiversity that exists in the air must be justified. This is probably not obvious for everyone. . . Also, it might be useful to specify that *Pseudomonas* is a genus of bacteria at least in the introduction, this might not be obvious for every readers of ACP and it is presented as a distinct category.

Response: Introduction section has been modified according to reviewer suggestions. It now includes the explanation for the selection of *pseudomonas*.

Comment: Unless I missed something, Table S1 and Figure 1 and Tables S8-S11 are the same data. However there are many inconsistencies, for instance the max values indicated for DNA, *Pseudomonas* and AA appear different from those in the figure. Can you check for any error and make the appropriate corrections.

Response: The data in Tables and Figures have been checked, revised and updated.

Comment: There is no mention of the results concerning control filters used for correcting chemical data: can you provide some information on what was found, if any contaminant was detected, and how the correction was done? Were there any such controls for microbiology (in addition of negative qPCR controls)?

Response: Thank you for this comment, we did analyze also control filters, in addition to the negative controls in qPCR. We have now added a description of the control filters to the method section: "In addition to the actual samples, we extracted DNA from eight blank filters, determined their DNA concentrations and used the extracts as templates in qPCR. All the blank filters were below detection limit in both DNA concentration assay with Qubit as well as in all of the three qPCR assays."

Specific comments:

Comment: L21 and throughout the manuscript: Specify which gene when mentioning gene copy numbers as it has no sense without this information.

Response: Specific genes used in the determination of the gene copy numbers have been included in the text.

Comment: In Figure 1, the labels PM 2.5 and PM 10 are misleading as these are actually not PM 2.5 or PM 10 in the sense $PM < 2.5$ or < 10 , but rather PM1-2.5 and PM2.5-10.

Response: Figure 1 and tables have been modified and revised.

Comment: L34: ". . .the influence of microbes. . .": The term "influence" suggests active intervention, is this what is meant? or does this rather refer to the contribution to the pool of chemical compounds? This should be clarified by modifying "influence" if appropriate.

Response: The sentence has been clarified as requested.

Comment: L58: What is meant by "the role of amino acids in the atmosphere"? "impact" might be more appropriate?

Response: The sentence was modified according to the suggestions of the reviewer.

Comment: L77: "Viruses can be frequently found in the airborne. . ." state? (word missing)

Response: The sentence has been revised.

Comment: Section 2.3: the latin names of trees must be italicized.

Response: Latin names have been now written with italics.

Comment: L 193-195: italicize latin organisms' names.

Response: Latin organisms' names have been now written with italics.

Comment: L233: Pearson correlations were used. Was/how the normality of data verified?

Response: As stated in the text, standardized Skewness and Kurtosis tests were used for the evaluation of the normality of data distribution. Additional logarithmic transformation of the data was needed to ensure the normal distribution of the input variables. This point has now been stressed in the text.

Comment: Section 3.6: (link between microbiology and VOC): aerosols for microbiological analyses were collected at 23m, above the canopy, while VOCs were screened by PRT-MS inside the canopy at 8.4 m above ground. Why this discrepancy? And how could this had influenced the data? It is known that above-canopy

and below-canopy air can be decoupled and can have different signatures (Gabey et al., 2010; Jocher et al., 2020).

Response: Sampling inlet for the impactor used for the collection of the samples was placed 5 m over the ground. Clarification was added to the section 2.3. 8.4 m above ground was the best option to collect the PRT-MS data of VOCs without any big differences.