Biogeosciences Discuss., 4, S164–S168, 2007
www.biogeosciences-discuss.net/4/S164/2007/
© Author(s) 2007. This work is licensed
under a Creative Commons License.

Biogeosciences Discussions

BGD
4, S164–S168, 2007

Interactive Comment

Interactive comment on "Molecular genetics and diversity of primary biogenic aerosol particles in urban, rural, and high-alpine air" *by* V. Després et al.

Anonymous Referee #3

Received and published: 23 March 2007

General comments: Overall, this manuscript presents a thoughtful contribution to the literature exploring biogenic particles in the atmosphere. The work appears to have been carefully done, and the discussion provides a thorough overview of the results and their implications for future studies. I provide the following suggestions for strengthening the manuscript.

Specific comments: 1. The title is limiting, particularly in that it neglects to note the important technical contribution of the work–specifically in your findings of substantial contamination of filters and the challenge this poses to data analysis and interpretation. Furthermore, the reference to 'molecular genetics' is misleading and may dis-

EGU

suade readers who are otherwise interested in the primary findings of this paper. Key contributions of the paper are in the techniques presented and in the general overview of what organisms were present and their very approximate abundances. I encourage the authors to revise their title to emphasize these points.

2. Error estimates need to be associated with the diversity index data! *Variation in the Shannon-Wiener:

Var = ([SUM(pi)(lnpi)2 - SUM((pi)(lnpi))2]/N) - ((S - 1)/2N2)

You should check this out in an ecological statistics text to make sure that I have included all the parentheses in the correct place! This information should be included for each index. ALSO..this gives you the information that you need to contrast locations for significant differences in diversity using a t-test. Without the statistics, the data are difficult to assess.

Another issue associated with the diversity index data, you apparently combined ALL data from all sampling dates for each location. Is this correct? This should be clarified in the methods and/or the results. Are there any biological reasons why this does/does not make sense? Is there a reason that you did not calculate diversity data for each individual sampling data? Do you have any information on the variation in diversity of your samples over time? It seems that you are missing what may be some important pieces of information in your data. Also, to what degree is overall diversity a function of perhaps only one or two samples? For example, with the rural samples, did you happen to sample one day when the fields were being plowed, generating a high diversity, and all other dates had quite low diversity? I would encourage you to provide more detailed information for the reader. Though there are some significant constraints to going too far with the diversity analyses (especially for the sequence data, where your samples are quite limited), it is worthwhile to provide more a more comprehensive overview of what your results show regarding diversity over space (locations) AND time (different samples within the same location. Your ability to provide specific information

on the variability among samples at different times from the same location could be very helpful for future research work, specifically in providing guidance into the need for repeated samples. Can I sample just once? To what degree do samples in the same location taken at 5 different times provide totally different information? Though this information may be fairly well established for culturable microbes in the atmosphere, there is little in the way of data that explores this issue for DNA.

3. I recommend a number of revisions to the tables and figures to enhance their value to the readers. a. For all tables and figures, the use of M (Munich), H (Hohenpeissenberg), and Z (Zugspitze) should be changed to U (urban), R (rural), and A (alpine). This is both because it is much easier for the reader to associate meaning with the coding factor, and also it avoids confusion with headings in another table in which H refers to the diversity index. b. In Table 1, delimit urban, rural, and alpine samples by a line or some other notification. This would enhance the potential for the readers to look at this table and readily pull out meaningful information. c. Given the large number of blanks for the TSP samples, I was not convinced that these needed to be included in Table 1. In particular, I would recommend considering including perhaps only an average value for the DNA mass and DNA concentration data for each location, with a range of sampling dates. If you do not intend to provide more thorough analyses or discussion of these data, than the detail presented in the table is not warranted. d. In Table 4, I would encourage you to make it easier for your reader by making the column labels more explicit. Rather than H (seq), etc., why not have the column label:

Diversity (H) Evenness Diversity Evenness Species Richness (sequence) (sequence) (T-RFLP) (T-RFLP) (T-RFLP)

Of course, the figure label would still clarify the Shannon-Weaver, etc.

e. For Table 5, the measurements are not clear. The label notes: “Percentage of different bacterial groups found in PM2.5 samples from ...”. Percentage of what? Neither columns nor rows total 100, so this is not a percentage of the group. What are the scal-

BGD

4, S164–S168, 2007

Interactive Comment

EGU

ing factors? Is this the percentage of samples? This needs to be clarified. Also, you may want to consider presenting this table as 2 stacked figures–one for the sequence data and one for the T-RFLP data. This may be more accessible to the reader.

f. Why have you calculated the relationship between DNA and PM mass for only the urban samples? Did you perform similar analyses for the rural and alpine sites? Is this a generalizable relationship, or is this specific to the location? Also, should define particle mass in the figure legend.

g. Figure 2 was quite difficult to read. I assume that you mean to have this published in color to permit the reader to track the different T-RF peaks? However, even assuming this is in color, it is not a particular easy figure to read. I would encourage you to consider presenting these data in a table. This way the reader can compare numbers much more readily among sampling locations. At the very least, the y-axis needs to be labeled clearly for the reader. But a table may be more appropriate for these data.

Technical comments: p. 350, l. 25: I would recommend rephrasing: Over 80% of the 53 bacterial sequences could be matched to T-RF peaks, though 60% of T-RF peaks did not correspond with any of the bacterial sequences.

p. 351, l. 11: omit 's' at the end of clouds

p. 352, l. 2: change to : 'of plant and animal fragments.'

p. 352, l. 14: 'loss' rather than losses

p. 352, l. 15: ...organisms can be identified TO...

p. 353, l. 10: ...which HAVE not been addressed in earlier publications.

p. 354, l. 17: what is a dry lawn? Does this have some commonly accepted interpretation?

p. 358, l. 12-14: How many replicates were used in this analysis? This seems to be a very important question, and one worthy of detailed data presentation. It is unfortunate

that you have information from only one DNA extract. Did you really do this for only one band ('The brightness of the band was reduced about 50% compared to the vector control')? Further detail and clarification here would help.

p. 360, l. 7-9: This is sentence is awkwardly written. Rephrase.

p. 360, l. 21-27: While the analysis of the corresponding number of haploid human genomes inhaled, why not also provide a brief view of the number of genomes of OTHER, smaller-genomed organisms that may be expected to be inhaled in that equilvalent DNA quantity? In fact, your data suggest that we aren't inhaling many human genomes, but rather Proteobacterial, Actinobacterial, etc. genomes. How many genome equivalents of these are we inhaling each day?

p. 361, l. 12-19: nicely written!

p. 367, l. 21: what does 'are statistically not well founded' mean? Do you just mean that sample sizes were too small to provide an accurate estimate of abundance? State this more clearly.

p. 368, l. 20: change to '...was not very efficient at extracting DNA from fungal spores...'

Interactive comment on Biogeosciences Discuss., 4, 349, 2007.