Biogeosciences Discuss., 4, S1902–S1916, 2007
www.biogeosciences-discuss.net/4/S1902/2007/

Biogeosciences Discussions

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

V. Després et al.

Received and published: 13 November 2007

We thank C. Morris, B. Moffett and an anonymous referee (Referee #3) for constructive comments and suggestions, which are highly appreciated and have been taken into account in the revised manuscript. Detailed responses are given below, focusing first on a couple of general aspects (manuscript title and statistical analysis) and then on specific comments from individual referees.

General Aspects

1) Manuscript Title

C. Morris, comment 1: The title needs a minor change. -Molecular genetics- is the

study of the molecular structure and function of genes. However, this paper addresses only the molecular diversity of aerosols. Hence -genetics- should be removed from the title.

Referee #3, comment 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 dissuade readers who are otherwise interested in the primary finding 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 approximate abundance. I encourage the authors to revise their title and emphasis these points.

U.Pöschl (co-author): Interactive comment on -Molecular genetics and diversity of primary biogenic aerosol particles in urban, rural, and high-alpine air- by V. Després et al., Biogeosciences Discuss., 4, S178–S180, 2007. www.biogeosciences-discuss.net/4/S178/2007/

R. Conrad (co-author): -The terms used in the title of a scientific paper should be correct and precise. On the other hand, it is often difficult to find the right term for a generalizing feature, which is acceptable across different disciplines. Thus, the term -molecular analysis- is frequently used among microbial ecologists to depict an analysis which was based on studying biomolecules instead of life organisms. However, this term is admittedly sloppy and potentially misleading in an interdisciplinary journal. The term -genetic analysis- is also misleading, since among biologists it is usually reserved for applying methods out of a genetic tool box rather than just sequencing or fingerprinting individual genes. Because of these complications, I suggest to change the title using a more specific term by replacing -molecular analysis- with -ribosomal gene analysis-. This would avoid some of the earlier confusion. Still it is a general term, since it comprises fingerprinting of the ribosomal genes, sequence analysis of the ribosomal genes and phylogenetic analysis of the ribosomal genes.-

BGD

4, S1902–S1916, 2007

Interactive Comment

Response: To satisfy all comments and requests, we intend to revise the title as follows: -Characterization of primary biogenic aerosol particles in urban, rural, and high-alpine air by DNA sequence and restriction fragment analysis of ribosomal RNA genes-

2) Statistical Analysis

C. Morris: The authors do not attempt to make any statistical comparisons about the diversity among different sites. There are some anecdotal remarks about observed differences between sites, but otherwise the effort to sample at different locations is not well valorized. Could the authors improve on this part of the work with existing data (i.e. other data that might not have been presented here)? Otherwise, it is not clear why such an effort was made to sample at different locations. C. Morris, comment 11: The experimental design does not allow any statistical comparisons to be made among the different sampling locations. Is there any way that the data can be better exploited so as to test hypotheses (via statistical analyses) about differences in locations?

Referee #3, comment 2b: Also this gives you the information that you need to contrast locations for significant differences in diversity using a t-test. Without the statistics, that data are difficult to assess

Response: We fully agree that further statistical analyses will be needed to characterize and corroborate the differences between different sampling locations. We are planning to perform such analyses in future studies building on the methods and findings of this study and on larger numbers of samples collected specifically for this purpose. Within the present exploratory study, however, we see no need and no gain in further statistical analyses, because there are no additional data available yet to support robust statistics and definitive conclusions about local differences and temporal variations. Nevertheless, we think that the observed and reported variations between the investigated urban, rural and high alpine samples (taken from earlier measurement campaigns with different aims) are instructive and relevant for orientation and future investigations.

BGD
4, S1902–S1916, 2007

Specific Comments

3) Individual comments by C. Morris:

Morris: The authors have presented data indicating that -clean- filters might harbour significant amounts of DNA. However, their arguments could be improved if they demonstrated that this DNA can be amplified. If the background DNA cannot be recognized by primers used in a study of aerosol biological diversity, then does this contamination really introduce a serious error and the need for decontamination procedures? Morris, comment 9: This section on blank and background samples provides information about the presence of DNA on filters before they are used for sampling and that which might accumulate during handling after sampling. Data is given about the quantity of DNA. Is data available about the identity of this DNA? It would be useful to know what organisms it corresponds to. Was there any attempt to amplify this DNA? If it could not be amplified, then it could be argued that it might not necessarily be an important source of error for diversity studies.

Response: The DNA extracted from blank and background samples could be amplified (PCR) with a universal bacterial primer pair, cloned but not sequenced. Thus, the bacteria present on these blanks could not be further specified. This issue will be addressed in the revised manuscript.

Morris: The authors present data about the taxonomic identity of the DNA sequences detected here. I understand that it would be unusual for the authors not to include this data in a table: it serves as a point of reference. However, it is somewhat redundant with existing information: the authors note that the taxonomic groups detected are rather typical of aerobiological studies. Remarks should be added that some of this DNA likely corresponds to non viable organisms as no attempt was made by the authors to use selective techniques to isolate DNA in so-called -viable- cells: Response: A remark concerning the analysis of viable as well as non viable organisms will be added in the revised manuscript.

BGD

4, S1902–S1916, 2007

Morris, comment 2: This paragraph is about anomalies in characterization of air samples due to contaminants and chemical modifications. Hence, shouldn't -On the other hand- be changed to -Furthermore- as the information that follows is not a contradiction to what preceded but rather an additional illustration? Response: It will be changed as suggested

Morris, comment 3: Remove the word -biological-. When is an organism not -biological-?

Response: As confirmed by English and other language dictionaries, the term organism is not restricted to biology. See for example: http://www.merriam-webster.com/dictionary/organism 1: a complex structure of interdependent and subordinate elements whose relations and properties are largely determined by their function in the whole 2: an individual constituted to carry on the activities of life by means of organs separate in function but mutually dependent: a living being

Morris, comment 4: On pg 354 the authors indicate that they did not decontaminate the TSP filters. It should be explained why these filters were not decontaminated whereas the others were.

Response: We had no possibility to decontaminate the TSP filters, because these had been collected by other researchers for other purposes. An additional remark will clarify this in the revised manuscript.

Morris, comment 5: -Aliquot- is generally used to describe a volume. Here the word is used to describe a section cut from the filter. It would be best to state: -PM2.5 filter pieces-

Response: We are not aware that the word aliquot would be restricted to volumes. Nevertheless, we will follow the suggestion to avoid any further confusion.

Morris, comment 6: The description of how the data were normalized is not clear

Response: The data was normalized as described in Lüdemann et al., 2000, except

BGD

4, S1902–S1916, 2007

Interactive Comment

that the peak height was used instead of the peak area. This information will be added in the revised manuscript.

Morris, comment 7: Concerning statistical analysis and data related to identity of sequences: did all sequences belong to known groups? Were there unidentifiable sequences? This is not clearly presented. Response: All sequences did belong to known groups and there were no unidentifiable sequences. This information is given in 3.2.1.1.

Morris, comment 8: Atmosphere aerosol samples: Information is presented about the adequacy of the DNA extraction kit. Were there many measures of efficiency whereby decontaminated filters were seeded with a known amount of DNA followed by extraction? This would be useful.

Response: The reported tests were aimed at finding out if it is possible at all to extract DNA from different filter types. In this exploratory study we have not yet quantitatively determined the efficiency of the applied DNA extraction kit. Quantitative seeding and recovery experiments will be performed in follow-up studies.

Morris, comment 9: listed and answered above.

Morris, comment 10: Sequence and phylogeny: In this section the authors present the identities of the sequences detected based on blasting in the NCBI data base and on phylogenetic analyses. It must be understood that all these identities are the best HYPOTHETICAL identities based on similarities with known organisms. In this light, it is VERY misleading to bring up names of members of some of these large groups that have a high -sales value- such as Bacillus anthracis. There is no data to support that such organisms were in samples. Why not suggest the name of any other random Bacillus species or other spore-former?

Response: We understand the referee s concern, but we think that our formulations are not misleading but adequate for a paper that is targeted at a wide audience of geoscientists, most of whom are not expert microbiologists. We named Bacillus anthracis

BGD

4, S1902–S1916, 2007

as a prominent example of species which fall into the groups of bacteria detected in our samples (Despres et al., BGD, 2007, p. 364): -The Firmicutes include common soil bacteria like the endospore-forming bacilli, but also prominent pathogens like Bacillus anthracis or Bacillus thuringiensis, which is agriculturally used as an insecticide.- To us the naming of a well-known example is not a matter of -sales value- but a matter of orientation for non-experts. Moreover, the presence of endemic pathogens in air has been shown in a different publication (Brodie et al., 2007) and is thus not -far fetched-.

Morris, comment 11: listed and answered above.

Morris, comment 12: The authors state -These findings are consistent with the results of bacterial sequence analyses in air particulate matter at urban, and rural locations- They should add -and with most previous studies using culture-based methods- (and cite the appropriate references). Response: We will add the suggested statement and references in the revised manuscript.

Morris, comment 13: This result about the absence of fungal DNA is very surprising!!

Response: As stated in the original manuscript (section 3.2.2) we think that -the applied soil DNA extraction kit was not efficient or the amplification was inhibited- in the reported investigations. Ongoing follow-up experiments indeed indicate high abundance and diversity of fungal DNA in urban and rural air, which will be reported in follow-up studies.

Morris, comment 14: Animal sequences. In this section the authors again make a misleading remark related to the identity of their sequences. The Alveolata: Apicomplexa, otherwise known as the Sporozoa, are common parasites of insects and vertebrates. Hence there is a large list of possible examples that the authors could cite to illustrate this type of organism. I do not think that it is appropriate to single out a human pathogen when there is no specific data. In the long run this could lead to dangerous consequences for this field of research and does not favour the comprehension of microbiology by the general public.

BGD
4, S1902–S1916, 2007

EGU

Response: In principle, we maintain the perspective already outlined in response to comment 10. Nevertheless, we will follow the suggestion to give more general examples and revise the statement as follows: -These spore forming unicellular eukaryotes are known as common parasites of insects and vertebrates.-

Morris, comment 15: change -were- to -we- and eliminate the comma after -calculated-. This will read better. Morris, comment 16: change -have been -to -were- and add -of the- in front of -National Center- Morris, comment 17: References: There are many inconsistencies in the format used to list the references, especially concerning the use of capital letters in titles. Response: It will be changed as suggested

4) Individual comments by B. Moffett:

Moffett, comment 1: It appears only a single PCR was carried out for each analysis and that 35 cycles were used. I feel bias would be reduced if several replica reactions were pooled prior to analysis and if the number of cycles were reduced, particularly if inferences about abundances are to be made.

Response: We thank the referee for this suggestion, and we agree that PCR replica will be important for further investigations and the determination of abundances. We did not perform replica in this first exploratory study, because the amount of DNA extract was very limited, but we are planning to include replica in follow-up studies.

Moffett, comment 2: I found this a bit confusing. It is not clear to me how many clones were taken for each PCR reaction and how these relate to table 3. Were 7 products obtained for a single PCR? This needs clarification.

Response: After the cloning step a PCR was performed using a single clone as a template. Depending on the success of the cloning step up to seven clones (resulting from one cloning step) were analyzed in seven separate PCRs. A remark clarifying this issue will be added in the revised manuscript.

Moffett, comment 3: The tentative tying together of the T-RFLP data and the sequence

BGD

4, S1902–S1916, 2007

data is of interest and I wonder if greater confidence would be obtained by labeling both primers with different fluorophores. If fragments from both ends correspond to a particular sequence this would be a more rigorous test of a match. Response: Thank you very much for this interesting suggestion. We will test this approach in future studies.

Moffett, comment 4: In discussing the differences between different studies this is more likely to be a combination of both aerosol variability and technical points as aerosols are thought to be highly variable in time and space and this would be exacerbated by variation in techniques.

Response: We will clarify in the revised manuscript that both, aerosol variability and technical points, may lead to differences between the results of different studies.

Moffett, comment 5: I feel a bit more explanation or at least a reference regarding the sequence data being -biased by the cloning procedure- is required.

Response: A reference addressing this aspect will be added in the revised manuscript. (v. Wintzingerode, F; Göbel, U.B.; Stackebrandt, E. 1997. Determination of microbiological diversity in environmental samples: pitfalls of PCR-based rRNA analysis. FEMS Microbiology Reviews 21, 213-229)

Moffett, comment 6: insert -fragment- between restriction and length Response: It will be changed as suggested

Moffett, comment 7: Table 1 complete blanks with ND if appropriate Response: n.d. for -not determined- will be added in the revised manuscript where appropriate

Moffett, comment 8: Table 5 complete unknown row with 0 if appropriate Response: In the revised manuscript, Table 5 will be replaced by Figure 4.

5) Individual comments by anonymous Referee #3:

Referee #3, Comm. 1: answered above (general aspect 1)

BGD

4, S1902–S1916, 2007

Interactive Comment

Referee #3, Comm. 2a): Error estimates need to be associated with the diversity index data! *Variation in the Shannon-Weaver: 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.

Response: In this exploratory study, we have calculated the Shannon-Weaver indices just as a rough indicator of the observed diversity, not as a robust parameter for further statistical analysis. As outlined above, the data set is too small for further statistical analysis. To calculate robust error estimates we would have to analyze multiple independent samplings, calculate their indices and from the calculated indices then the mean and the errors. The data required for this approach, however, are not available.

Referee #3, Comm. 2c): 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 results. Are there any biological reasons why this does/ does not make sense? Is there any 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.

Response: For each location all data were combined for the calculation of Shannon Weaver Indices. This will be clarified in the revised manuscript (caption of Table 4). We agree that more detailed investigations of individual samples would be interesting, and we are planning to perform such investigations in future studies. As outlined above, however, we think that more data than available in this exploratory study will be required for robust statistics and definitive conclusions about temporal variations.

Referee #3, Comm. 2d): Also, to what degree is overall diversity a function of perhaps one or two samples? For example, with the rural samples, did you happen to sample one day when the fields were plowed, generating a high diversity, and all other dates

BGD

4, S1902–S1916, 2007

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 with 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 fore future research work, specifically in providing guidance into the need foe repeated samples. Can I sample just once? To what degree do samples in the same location taken at 5 different time 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.

Response: We thank the reviewer for these very thoughtful comments. Our aims and ideas in general go into the same direction. However, we think that this kind of analyses would go beyond the limits of the presented set of aerosol samples and measurement data. In addition, we do not have more information about the environmental conditions of sampling (e.g. plowing in short distance etc). As outlined above, we think that more data than available in this exploratory study will be required for robust statistics and definitive conclusions about spatial and temporal variations (PCR replica, extraction of different parts of the same filter, larger number of samples, etc.). We are planning to address such aspects in follow-up studies.

Referee #3, Comm. 3a): I recommend a number of revisions to the tables and figures to enhance their value to the readers. 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. Response: Changes will be made as suggested, using -HA- for high-alpine.

BGD

4, S1902–S1916, 2007

Referee #3, Comm. 3b): In Table 1, delimit urban, rural and high-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 Response: Lines will be included as suggested.

Referee #3, Comm. 3c): Given the large number of blanks for the TSP samples, I was not convinced that these needed to be included in Table 1 Response: TSP filter information will be removed from the table due to their large number of blanks.

Referee #3, Comm. 3d): 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

Response: The provided information is the basic set of data typically required for characterization of atmospheric aerosol samples, which we consider potentially relevant for interested readers. Thus, we intend to keep the mass and concentration data in Table 1 and to complement these by overview information about the number of sequences and T-RFs observed per sample. In addition, we will re-arrange Table 3 in the revised manuscript for improved readability.

Referee #3, Comm. 3e): 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 a column label: Diversity (H) Eveness Diversity Eveness Species Richness (Sequence)(Sequence) (T-RFLP)(T-RFLP)(T-RFLP). Of course the figure label would still clarify the Shannon-Weaver, etc. Response: The readability of the Table will be improved in the revised manuscript.

Referee #3, Comm. 3f): For Table 5, the measurements are not clear. The label notes: -Percentages 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.

BGD

4, S1902–S1916, 2007

Interactive Comment

What are the scaling 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. Response: Thanks for this helpful suggestions. In the revised manuscript, Table 5 is replaced by a new figure (Figure 4), which should be easier to read.

Referee #3, Comm. 3g): Why have you calculated the relationship between DNA and PM mass only for 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, you should define particle mass in the figure legend.

Response: Similar analyses were performed but did not exhibit such a pronounced correlation, which is likely due to the more constant environmental conditions at the urban sampling location. We consider the observed correlation specific for the location. Further investigations will be needed to find out more about general relationships between DNA and PM mass in the atmosphere. The definition of PM will be added.

Referee #3, Comm. 3h): 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 be more appropriate for these data. Response: We will improve the labeling of x- and y-axis to make the figure better readable, but we think that the figure is more instructive than a table would be.

Referee #3, tech Comm. 1): 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 Response: We rephrased the sentence to make our remark more clear.

Referee #3, tech Comm. 2): p. 351, l. 11: Omit -s- at end of clouds Referee #3, tech

BGD

4, S1902–S1916, 2007

Interactive Comment

EGU

Comm. 3): p. 352, l. 2: Change to -of plant and animals fragments- Referee #3, tech Comm. 4): p. 352, l. 174: -loss- rather than losses Referee #3, tech Comm. 5): p. 352, l. 15: organisms can be identified TO Referee #3, tech Comm. 6): p. 353, l. 10: Which -Have- not been addressed in earlier publications.

Response: Manuscript will be revised as suggested.

Referee #3, tech Comm. 7): p. 354, l. 17: What is a dry lawn? Does this have some commonly accepted interpretation?

Response: There is no -accepted interpretation-. The instrument was situated next to a lawn in southern Bavaria. Thus, -dry lawn- is just a term to describe the countryside in which the samples were taken.

Referee #3, tech Comm. 8): 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.

Response The inhibition test was performed once at a single DNA extract for the PM2.5 samples. However, we did also test one sampled coarse particle DNA polypropylene filter and one sampled cellulose nitrate filter and did find that in the cellulose nitrate filter the PCR reaction was completely inhibited. We will add this additional information in the revised manuscript as we agree that inhibition is an important aspect for future systematic DNA analyses of aerosol samples. This is why we developed and tested a method for detecting and characterizing inhibition, which had not been addressed in earlier studies. We think, however, that more data than available in this exploratory study will be required for robust statistics and definitive conclusions. Therefore, we are planning to address such aspects in follow-up studies.

Referee #3, tech Comm. 9): p. 360, l. 7-9: This sentence is awkwardly written.

BGD

4, S1902–S1916, 2007

Rephrase. Response: Will be rephrased in the revised manuscript.

Referee #3, tech Comm. 10): 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 equivalent 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?

Response: As suggested by the referee we included the calculations for a proteobacterial genome in the revised manuscript. Under the assumption of an average DNA concentration of ˜7 ng m-3, average DNA amounts of ˜4 fg per haploid bacterial genome (E.coli) and adult human breathing rates between 5 and 120 L min-1 (sleep vs. sports), an adult person living in a city can be expected to inhale every day about 0.05-1.2 μg DNA, corresponding to 107-108 haploid bacterial genomes (E.coli). This according information about the proteobacterium E.coli will be added in the document.

Referee #3, tech Comm. 11): 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.

Response: We do indeed refer here to the small number of obtained sequences which is too small for any robust statistical analysis. We will add a remark in the revised manuscript.

Referee #3, tech Comm. 12): p. 368, l.20: change to -was not very efficient at extracting DNA from fungal spores- Response: Will be changed as suggested.

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