Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2020-324-AC2, 2021

Biogeosciences
Discussions

Interactive comment on "Determining the hierarchical order by which the variables of sampling season, dust outbreaks occurrence, and sampling location, can shape the airborne bacterial communities in the Mediterranean basin" *by* Riccardo Rosselli et al.

Riccardo Rosselli et al.

squart@unipd.it

Received and published: 12 March 2021

Anonymous Referee #3 Received and published: 7 February 2021

General Remarks: The manuscript describes a study on the parameters affecting airborne microbial community composition, e.g., season, dust intrusion, geographic proximity to the dust source. These are important questions in the study of aerobiology, especially in the Mediterranean basin that is prone to increasing frequency of Saharan

dust intrusions. The study presents a surprising result, according to which the time of the sampling is the most significant factor affecting the airborne microbial community composition. Although seasonal differences have been demonstrated in previous studies, at various locations, I have no knowledge of any that have resulted in such overwhelming differences between two sampling campaigns at the same location, under similar atmospheric conditions. This does not come to doubt the validity of the results; however, extra-caution should be taken to ensure that no confounding variables are responsible for this result. A possible cause for this result might stem from batch effects, e.g., DNA extraction, amplification and sequencing conducted by two different people, on two different occasions might be sufficient in producing such differences. Therefore, the authors are urged to specify whether actions were taken to prevent any batch effects.

ANSWER: The possible batch effect issue does not appear to apply in this case as the processing has been done by the same single operator in all sampling times. A chart with the data details showing the processing uniformity of the throughput is provided in response to a further comment below and is available as Supplementary Table S2.

Other general suggestions: 1. The term "seasonality" can be used if a cyclic change over seasons is shown, the difference between May and September of a single year is better referred to as "temporal".

ANSWER: We find the comment very appropriate; as a matter of fact we have changed also the title of this manuscript by substituiting the word 'season' with 'period' (The new title is: Determining the hierarchical order by which the variables of sampling period, dust outbreaks occurrence, and sampling location, can shape the airborne bacterial communities in the Mediterranean basin). Moreover the term has been corrected throughout the manuscript in whichever occasion it had been used to refer to this single-year campaign and not to truly recurring phenomena. The use of the words season or seasonal in the paper is therefore limited to the descriptive context, while in any instance in which we infer/suggest interpret something from the observed data, we

are now using the terms time, temporal or period.

2. Only a single sample per location per month represents the ambient conditions, therefore it is hard to compare dusty to clear conditions. In the absence of several control samples per site, per month, one cannot appreciate the natural variation of the airborne community. With the current study design, the samples representing the same month or the same site cannot provide information on dusty vs. clear days. Possible comparisons can only be made between sites and between sampling periods (September / May). Clear to dusty conditions can only be compared across the entire dataset. However, the great variance observed between May and September probably masks the role of dust in changing the atmospheric bacterial community.

ANSWER: One aspect that needs to be considered here is that an atmospheric sampling is not to be regarded with the same conceptual metrics that would apply if one were to study liquid environments as e.g. a seashore or solid ones, as a farm plot, for which cases 'one' sample could correspond to the 50 ml filling a falcon tube dipped in water, or a gram of soil scooped from the ground. In skypost air filtering the operation is carried out continuously for days and one sample, in our case, accumulates the content of 56160 liters of continuously changing atmosphere, which takes into account the variations that occur during all those hours, inclusive of the day/night shifts. One sample is therefore not a 'point' but a built-in averaged replication protocol for the chosen window of events. Moreover, since it is already known from literature that, as microbial community composition is concerned, even in the absence of dust outbreaks, the ambient state of the atmosphere is not stable either, our goal was not to compare an hypothetical status quo with an altered one. The meaning of the 'control' here was to catch 1) the first possible timeframe after the stopping of a northbound dusty wind outbreak (it occurred in May) or 2) the latest possible timeframe of a situation before the onset of a dust-carrying change of wind regime (it occurred in September). Thence, in the latter event the control is not intended as a situation of calm that could represent a period of unknown length, but rather the time-zero sample of the dust event itself.

While for the former case in May, the control is symmetrically designed as the quiet after the storm.

Specific remarks:

L. 32: "concerted succession..." – The use of the term "succession" implies bacterial growth and selection. Please rephrase throughout the manuscript.

ANSWER: The term was corrected.

L. 54: particle size can well exceed 10 um. Dust storms often carry larger particles (Ryder et al., 2018).

ANSWER: The size class limit has been corrected and the reference added.

L. 61: Please provide a specific website address, the home page of WHO is insufficient.

ANSWER: The correct references (instead of a website) have been placed (Prospero et al. 2002, Schepansky et al. 2007) L. 73-76: Should rephrase: according to the cited paper these genes are not specific to atmospheric bacteria, it is suggested that their presence might enable bacterial survival in the atmosphere.

ANSWER: Correction made

L. 88: Change "until" to: "up to", or: "reaching".

ANSWER: Correction made (up to).

L. 149-151: this is not so clear. What is the filtering step? What do the two filters represent? A single sampling event? Two consecutive sampling days?

ANSWER: The sentence was still referring to the sampling (filtering air for 24 h). The two filters are parallel replicates treated independently up to the DNA extraction. The two consecutive sampling days was incorrectly written to mean the two periods of 12 h each in which the sampling was divided during the dust outbreaks. The text was edited to clarify the procedure as follows: "The experimental design involved: 2 sampling sites

at the opposite corners of the Sardinia island (Sassari vs. Cagliari), 2 sampling periods (May vs. September) 2 meteorological conditions (absence vs. presence of a dust outbreak). In each of these situations, two replicate samples were taken and processed independently throughout the DNA extraction step to be pooled before sequencing. Samples were collected on Teflon filters (Sartorius Stedim Biotech) by using a Skypost Tecora apparatus (compliant to the European legislation 96/62/gmeCE) processing 39 liters of air per minute. To constitute 'a sample' a continuous 24h-long air intake through the filters was performed. In the case of the dust outbreaks the 24h sampling was further divided in two periods, by considering independently the first 12 hours and the second 12 hours. The number of resulting samples was therefore 12; namely the module of three conditions: (a) Control; (b) Dust h 1-12, (c) Dust h 12-24; multiplied by 2 sampling periods and by the 2 sampling places, resulting in 3 x 2 x 2= 12 samples. As technical note, the scope and meaning of 'controls' here was that of samples that could allow to individuate the shift between one condition and its adjacent one. In our cases, catching the sudden change of wind regime by sampling immediately before or after a dust outbreak. Therefore, the controls were thus defined as single 24h time points flanking the key dust events."

The methods section should clearly indicate how many samples were collected, what was the duration of each sampling event, their dates, etc. It is advised to add a table that sums all the sampling data. DNA extraction and sequencing:

ANSWER: besides the above editing of the text we have included in the Supplementary material as indicated, a table showing the distribution of the reads output throughout the sampling which clarifies the homogeneity of the protocol outcome and addresses the queries on the possible batch variability ruling out such possibility. (Table S2. Details on the sequencing output quality and evenness of distribution across samples)

The choice of relatively large amplicons (_1400 bp) along with 93 cycles in a paired-end sequencing is surprising. Although there are various platforms that allow pre-processing of nonoverlapping sequences, this might introduce more errors to the

alignment step and to the taxonomic classification. Could the very high number of counts per sample (min. 1109571) arise from some sequencing pre-processing error? Were all samples sequenced on a single lane? Were they separated to different lanes according to their dates? Were the amplification and/or sequencing conducted in batches?

ANSWER: The sequencing strategy chosen at the time was using the Nextera XT DNA protocol via a whole amplification of the 16S rRNA and a shotgun sequencing with 93bp x 2 paired-end reads. All the DNA samples have been therefore sequenced in the same flow lane to avoid biases due to different sequencing batches. The high number of reads obtained from each sample is due to the high efficiency protocol for very low amounts of DNA.

L. 182: "OUT table" should be "OTU table"

ANSWER: Correction made

Figure 3: Please increase the font and provide clear titles and a color legend within the figure, and not as part of the caption. In general, pie-charts tend to be less clear than bar-charts, since the human eye estimates height differences better than area differences. It is possible to create a dendrogram including all samples, better emphasizing the clustering of samples according to the time of sampling. Also, does each pie-chart represent a single sample or a group of samples? It is not clear how many samples were used for this study.

ANSWER: Fig. 3 was restructured graphically by enlarging the charts, the fonts, the scale and by adding the color coded taxonomy in the figure itself. As now clarified by the above answers, in the figure each chart represents a single 24h sample (for the dust event samples, the results of the first and and second 12h periods and have been merged in a single pie to compare each case with the same filtering duration). The information has been added in the text and legend.

Section 3.2: The authors describe in detail the differences observed between the samples. These samples represent an array of conditions: dust events vs. clear days; September vs. May; Cagliari vs. Sassari. A clearer representation might be achieved if this section was divided and each comparison was described separately.

ANSWER: Our description was following the microbiology clades as leading topics and for each main phylum or class we comment the differences in their occurrence in relation to the three variables of time, meteorology and geography, by the order which explains the partitioning of each phylum. We have checked the rearrangement of the section as suggested, but in that case, there would be a six-fold multiplication of the description of each taxonomical group (e.g. for Actinobacteria we would need to describe their statuses starting over each time in the six different paragraphs for May, September, Cagliari, Sassari, Dust and Control) and in some cases no relevant facts apply for many phyla. The result appeared to convey a more dispersed view when compared to the presentation of patterns by-microbiology in which we underlined only the variables involved in main differences and we could also group the description of taxa that showed common behavior for different variables. We also find that having redrawn and anticipated Fig.3 and repositioned the former Table 4 (Now table1, allows to follow the text of section 3.2 in a much smoother way.

L. 315-321: The authors assume that taxa that will show the least variability between two samples of the same dust event, belong to a local core microbiome. This assumption should be better explained, and answer the following questions: 1. Since dust events introduce new taxa, they are expected to "dilute" the core microbiome, resulting in a variation between dusty/clear conditions. Why not look for the bacteria that decrease in abundance when dust events occur?

ANSWER: The method indeed lacked clarity and an important premise. Addressing first the comment on why not look at bacteria that decrease in abundance being diluted (passively) by mass immigration of others. The issue has to do with the distinction between actual population dynamics (ecologically-ruled) and mathematical effects of

sampling from a ballot box of objects, all of which compete for the constrained 100% format of results (probability-ruled). The problem arises when comparing communities at different time points (and even more with DNA-based methods with fixed total DNA amount processed), in which results are based on percent values (relative abundance of taxa). In that case the multiplication of any, determines obviously a reduction of the relative abundance of others when those do not grow at equal or higher rate. Therefore one given group could have been increasing, but its share in the sum could appear as if it had instead decreased if a different group has increased faster. This consideration, that applies inevitably for all metagenomics studies, should be kept in mind for all kinds of interpretations about increases and decreases which could be either real or apparent (when driven by a stronger change of a different group). As consequence, comparing different sampling points through time is linked to this inevitable constraint: the compositional nature of the dataset binds all relative frequencies to each other. Therefore, since the sum of them is bound to give always 100% the decrease of a given species could be either apparent (driven by the increase of another), or real (due to its actual negative population dynamics). The problem is that the two causes can not be uncoupled by just comparing their frequencies at the two sampling times. Moreover, even an actual increase of a given species could be masked by the parallel increase of another at a faster pace (or by its net immigration into the scene). This is the reason by which we consider with caution the possibility of looking at decreasing taxa as indicative of their actual biological activities or fate.

As regards the first part of the query, we revised the text to address the comment as follows:

“Besides the comparisons that included all OTUs to put in evidence community variations, in parallel we exploited an additional opportunity to detect possible dust-specific taxa. The rationale was to seek differential enrichment within the dust storm, by dissecting the process, during its progression, splitting its onset from its fully established stage. To this aim we collected separately the filters of the first 12h of the event, and

replaced them with new ones that collected air during the second lapse (hours 12 to 24). Thus the availability of two timeframes, both within the dust event, allowed to verify which OTUs would be incrementally enriched along with the progression of the stormy condition. This allowed to better refine the bacterial deposition dynamics during the outbreaks. From the visual and physical points of view, an increase in the inflow of air particulate was observed for the 12-24 h period, confirming the differential level of deposition occurring in the maturity stage of the meteorological phenomenon. This within-outbreaks set up was essentially aiming at individuating taxa that would display high variation in relation to dust events in comparison to those who would not. The latter were considered to represent the common core of bacteria that were constantly present in samples, irrespective of the changing meteorological events. To apply this distinction, the criterion was to set a cutoff value with respect to the percent of variation occurring between the first 12 h of the collection time and the second half of it. The choice of this threshold was considered critical and, in order to ensure robust conclusions, we deemed necessary to require a considerable consistency of variation. Pointing at this objective, only the taxa which displayed a mean variation higher than $\frac{1}{2}$ of the corresponding standard deviation were taken into account. The resulting level of variation in the two sampling stations is reported in Tab. 2 (Formerly Tab. 1) and the corresponding number of orders is displayed in Tab. 3 (Formerly Tab. 2). The Sassari (North-facing) collection site was the one that in both seasons resulted to feature the highest number of significantly changing taxa. The identities of these are shown in Supplementary Fig.S7 (May event), and Supplementary Fig. S8 (September event). In the graphs, the first 12h lapse is plotted above the baseline and the second (12-24 h) is on the specular position below."

We also added a clearer legend to the Table (former Table 1) as follows:

" Extent of OTUs change across cell harvesting time during the same dust outbreak sampling. The percent variation (either increase or decrease) of a given OTUs abundance between the values found in the community obtained by the first 12h sampling

and the ones resulting from the following 12h lapse was computed. The average, minimum and maximum percent variation between counts are shown. Only taxa displaying a difference in percentages higher than half of their standard deviation were selected for the present comparison."

2. Comparing two samples of 12 hrs each of a single dust event seems arbitrary. If PM10 mean concentrations throughout the examined 24 hrs remained stable, more dust-related taxa would show low variance, and be considered as "core microbiome". How do the authors assure that this is not the case here?

ANSWER: PM10 data by collecting services are made available in delayed fashion, but in our sampling we could decide the splitting not on an arbitrary pre-assumption but as, during preliminary trials in 2013 and in the February 2014 campaign (Published in Ref. Rosselli et al., 2015), we could witness in real time that an increase in the inflow of air particulate was occurring during the second (12-24 h) timeframe.

There is some confusion regarding which taxa were considered as significant for this analysis – L. 316 states that "The latter were considered to represent the common core: : :", referring to the taxa that showed lower variation. Following, on L. 320, it is stated that: "Only taxa which displayed a mean variation higher than 1/2 of the corresponding standard deviation were considered." In the following tables' captions it is stated that the numbers represent taxa that exceeded the threshold. Please clarify what is the aim of this analysis – to find a core microbiome or to find the "immigrant" bacterial community.

ANSWER: As mentioned above in relation to the compositional dataset constraints, the pursuit was to put in evidence differences that would be least affected by the apparent indirect mathematical effect that is common to all these studies. The common core is not the ideal word here as we would better define those as the bulk of taxa that display a behavior which is only indirectly affected by that of the others (i.e. ecologically unaffected and only mathematically affected in their relative frequencies resulting by

the behavior of others). The core microbiome here is to be seen as a complement background within which the immigrants are impinging. These aspects have been further detailed in the text of the section.

Tables 1 and 2: On table 1 the rows' order represents the two locations, alternately; on Table 2 the rows' order represents first Sassari and then Cagliari. Uniformity between tables is advised.

ANSWER: The former Table 1 (now Table 2) was reordered to achieve uniformity

Table 3: This is the first clear indication of the number of samples taken at each site, under the selected conditions.

ANSWER: We agree, having now anticipated the samples outline (3x2x2=12) in the text at above described in the Materials and methods should help to avoid confusion.

Figure 4: When the data is expected to vary along several parameters, it is advised to examine other PC axes, e.g., PC3 and PC4. This way, the effect of the most influencing variable (in this case – the time of the sampling) is less expressed, and gives way to see other parameters, possibly. Instead of showing the same PCA triplicated, the authors can attempt to make a more condense view of PC1 vs. PC2, using shapes and colors etc., and add figures displaying other PC axes, if they indeed show some correspondence to the other examined variables.

ANSWER: We have inspected further axes, however the explained variability added by PC3, PC4 etc. is of marginal increment (PC2 is already as low as 14%) and the representations were not adding clarity. However, the clearest complement to the data shown in Fig.4 is in our opinion Fig.5 (Formerly Fig. 6) which draws on the same PCA data and extracts its information by the discriminant analysis.

Figure 5: There's some redundancy in figures and tables: there is no new information arising from the PCoA in Figure 5, that wasn't provided by the PCA in Figure 4. Table 3 provides in detail the diversity indices of the samples, and Figure 5 A and B display it

as a boxplot. There's no need for both, the table can be moved to the supplementary information.

ANSWER: We agree and Fig. 5 has now been placed in the Supplementary material

L. 416: Using a Bonferroni correction for multiple comparisons is unnecessarily stringent. It resulted in only 6 taxa that were significantly differentially abundant when comparing the two sampling periods. A result of only six significant genera is probably of a low ecological relevance. It is more advised to apply Benjamini-Hochberg correction to achieve better statistical power, and a greater collection of significant taxa. It is often the case with very low p-values, that they represent rare taxa that are found in very few samples. I advise the authors to inspect the six significant taxa, and make sure this is not the case here. Moreover, other statistical tools, better fitting microbiome data analysis, are available for differential abundance tests, e.g., ANCOM, ANCOM-BC, MaAsLin2, etc. These methods should be preferred over ANOVA tests for compositional datasets.

ANSWER: We are grateful for the advice and for the uncommon acknowledgment of having been statistically even too severe. We have inspected the loosening of stringency effect, resulting in a more generalized array of taxa but supported by less robust p values, whose ecology is however less clearly linked to a nexus to the situation and whose presence appears more stochastic. We have also checked that for the six taxa that stood the Bonferroni test, are in all cases either consistently present in all samples or recur with frequencies higher than 1% in them. We therefore would maintain the high stringency output as the one we feel more confident to prudently describe.

Table 5: This table is somewhat overcrowded and demands an intense reading to draw conclusions from. It presents pair-wise distances between individual samples, and not between groups of samples (May vs. September, CA vs. SS, etc.), which provides no statistical significance to the observations arising from it. It also seems somewhat redundant, there is no added value in this table over the dendrogram, PCA and PCoA

already presented.

ANSWER: It is true and we have experimented different ways to reorganize that, including a heatmap-style arrangement with all the entries ending in a single matrix rather than three separate table sections, which however ended up more complex as it needs a tridimensional representation of the same/different season, dust/control and Cagliari/Sassari location. But actually the key message that this color table is meant to convey is just the fact that the low similarity pairwise comparisons (conditionally formatted as red-yellow shaded cells) are almost all observed in the upper table. i.e. the between-seasons comparisons, while the green ones (higher similarities) are all distributed in the two same-period comparisons. We further stressed this aspect to direct the readers' attention to the way to get the bottom line information from this table. The message is reinforced by the other figures as mentioned but this one shows the Bray-Curtis values and allows each pair to be inspected, which is an information that the PCA, PCoA and dendrogram are not providing with such detail. The table serves also a source of the dissimilarity matrix values for the Bray-Curtis based multivariate analyses, that a different reviewer has asked to provide.

L. 473: the declared goal of the study is indeed significant to the understanding the processes affecting the airborne microbial community in the Mediterranean; however, the design of the study suffers from too few samples. There are too few controls, and only two sampling campaigns, representing two seasons. Seasonality cannot be determined without repetitive sampling during the same season over several years. The presented differences can be referred to as temporal.

ANSWER: While as regards the number of control samples we have answered above which was the meaning of them (control as time zero of the dust itself), we fully agree on the fact that we can not interpret as seasonal (recurrent) patterns what we observe in a year and that we should stick to the term temporal to comment these observation and we have modified the terminology throughout the manuscript, including the change of its title.

L. 484: This remark is very true.

ANSWER: We appreciate the positive comment, which was indeed also already addressing the previous (L.473) point.

L. 488: This finding is very different from what was suggested by others before, e.g., Gat et al. (2017); Lang-Yona et al. (2020); Bowers et al. (2013); Caliz et al. (2018). Seasonal variations were shown in previous studies, yet they were not as extreme as shown in this study. Sampling location and aerosol back-trajectories were usually more significant in determining the airborne community composition. Did the authors make sure that no batch effects or other confounding variables stand behind the results presented here?

ANSWER: The cited studies indeed find also variations as acknowledged. In our case, as part of the prior comments'answers, we can not exclude that the single year that we analyzed could represent as particularly variable one and that repeating the same comparisons throughout different years would lead to lower variability as that caught by other authors (some of which in their own single-year analyses). As mentioned above, we can instead rule out the batch effect or confounding issues due to the sampling as it was repeated by the same operator and as the raw sequencing outputs (added supplementary table) do not show evidences of inter-sample variability in terms of technical throughput.

L. 527-532: The referral to PM10 concentrations is significant, especially when considering the changes in community composition during dust events. According to Figure S1, the dust events on May and September have tripled the atmospheric PM10 concentrations, compared with clear days. This significant change in PM10 is expected to result in a significant change in the airborne bacterial community (see Mazar et al., 2016). Yet, this change seems minor according to Fig. 3, 4 and 5. How do the authors explain this result?

ANSWER: Partly by the fact that PM10 is a physical measure of particulate size class

and does not straightly equate with a content of airborne biota. Mostly because the bulk of PM10 over industrialized or inhabited territories includes combustion particulate (which is by its source devoid of microbial cells or intact DNA) and also by the fact that in dust outbreaks a vast majority of the airborne fine material is not loaded with microbes as it comes from airlifted particles from desert zones in which a strong selection is exerted against surface life by unshielded UV radiation exposure, dry conditions and absence or scarcity of primary productivity. Results (as in Mazar et al. 2016) could also be dependent on the distance of the sampling location from the departing site of the airborne material and by the population density of the land crossed before discharge or in the surrounding of the sampling outpost.

L. 551-554: As stated before, sometimes the lowest p-values are given by the rarest taxa. Please make sure this is not the case for these taxa.

ANSWER: The issue has been addressed above and we refer to the prior comment

L. 576-593: Due to the low number of control samples, it is impossible to draw conclusions on dusty vs. clean conditions of the same season or the same location.

ANSWER: The text was rephrased to account for what we relied above to the same issue. Since the airborne community composition does change daily even during periods that do not feature the dust carrying episodes, there is not a stable condition that could be considered as the durable control community. Even the evening and morning opposite breeze regimes that occur daily in coastal locations impart modifications. Therefore taking a series of 'controls' intended as samples in different days before or after a dust storm, would have consisted in just as many different samples. The idea was instead to catch the shift that corresponds to the sudden change of wind regime immediately before or after a dust outbreak. Therefore, as explained. The controls are bound to be single time points that flank the two key events. The concept has been better outlined in the revised manuscript.

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2020-324, 2020.