

Dear Dr., Cédric Hubas,

Please find attached a revised version of the manuscript entitled “Clumpy coexistence in phytoplankton: the role of functional similarity in community assembly” by Caio Graco-Roza et al. Thank you for considering our manuscript for resubmission in PCI Ecology. We would also like to thank the reviewers for the very thorough review. We have now implemented all the minor suggestions by the reviewer or explained otherwise the reasons not to incorporate their suggestions. Below we present a point-by-point response to reviewers. We thank you and the reviewers for your time, and hope you find this version suitable for recommendations.

Yours sincerely,

Caio Graco-Roza

#### **Review of Graco-Roza et al. - Comments to the Authors**

I would like to thank the authors for the revised version of their manuscript entitled “**Clumpy coexistence in phytoplankton: the role of functional similarity in community assembly**”. In its current form, I found this paper interesting and the authors have made an important work in order to clarify their study, especially the ‘Methods’ and ‘Results’ sections. They have addressed all my previous comments and I think that the paper will be a valid contribution to the literature on this topic, providing quantitative information about the mechanisms that promote species coexistence and explain community structure. There are a few (minor) points that I would like the authors to address and/or discuss in the final version of the manuscript.

#### **Abstract**

**[1]** The authors could clarify the meaning of “clumps” in a clearer way, *i.e.*, “*a group of coexistent species in a community*” to help non-specialist readers to better understand.

**Answer :** We modified the sentence for clarity.

**[2]** line 40: “*In sum*” might be removed from the last sentence.

**Answer : Done.**

### **Introduction**

[3] line 51: “*exogenous spatio-temporal mechanisms*”.

**Answer : Done.**

[4] lines 82-88. I found this paragraph very interesting and relevant. The authors might clarify that the challenge that they highlight here – if I am correct – is mainly related to a statistical/numerical viewpoint.

**Answer : Done.**

### **Methods**

[5] line 161: “*based on the location of steep slopes the river elevation profile*”. Please, correct this sentence.

**Answer : Done.**

[6] lines 161-162: “*located close to the sampling sites*” might be replaced by “located in the upper and lower courses of the river”.

**Answer : Done.**

[7] line 190: “*100 individuals of the dominant species*”. Please clarify how which basis were determined the dominant species.

**Answer :** We corrected the sentence for clarity. Here, we meant that we counted 100 individuals of the most abundant in the sample.

[8] line 200: “*was measured, and the presence of aerotopes, mucilage, flagella, and siliceous exoskeletal structures*”. A part of the sentence is missing.

**Answer :** We corrected the sentence.

[9] lines 186-208. The authors might refer to Table S1 in this section. Please, correct “*Mophology based functional group*” in Table S1 and use the same abbreviates than in the main text (e.g., “vol” vs “V”) and the same units.

**Answer :** Thank you for noting this. We have now fixed the Table S1.

[10] lines 186-208. Please clarify in this paragraph if the classification into MBFG you performed here to group species was based on a numerical approach or only on threshold values to assess species in each group.

**Answer :** My apologies if I did not understand this question very well. We assigned species to MBFG following the classification key provided by Kruk et al 2010 (Figure 2 in their paper). This could be done only following the tresholds in the key. However, we used the functional traits in other analysis so we measured them for all species irrespectively to any treshold provided by Kruk et al. 2010.

[11] I really appreciate the sketch diagram that links statistical analyses and the three hypotheses. I would suggest to add this figure directly in the main paper as it will greatly help the reader to follow the flowchart of the analyses and the link with the results. The first steps of the work, e.g., calculation of species biovolume, data transformation, etc... might be added to the figure to be as exhaustive as possible.

**Answer :** We improved the figure and now included it in the main text (Figure 2).

[12] lines 210-214. This paragraph might be replaced at the end of the ‘Methods’ section.

**Answer :** Done

[13] lines 219-221. Please, specify a), b) and c) for each matrix/data frame.

**Answer : Done**

[14] line 238 vs line 264. Could the authors clarify why two different log-transformation approaches were applied and why a  $\log_{10}(x+1)$  transformation was not performed for the individual volume of species?

**Answer :** We used  $\log_{10}(x+1)$  because it is a standard procedure in limnologic studies. Regarding the individual volume of species, there is no need to sum one because  $\text{volume} > 0$  for all the species. The choice for the  $\log_2$  followed recommendations of Fort et al 2010. We have included this reference also in the sentence describing the log for clarity.

[15] line 314: “*and one by diving the actual  $F_{Dist}$* ” might be replaced by “*and one by dividing the actual  $F_{Dist}$* ”, if I am correct.

**Answer : Done**

## **Results**

[16] I suggest the authors (and this applies to the other sections) to replace “temporal” by “seasonal” (e.g., lines 358, 381) throughout the paper, as the study was performed on two seasons in a single year. For Figures 3 (line 401) and 4 (line 425), it might be helpful to specify: “**Figure 3.** Seasonal distribution of phytoplankton...” and “**Figure 4.** Spatial distribution of phytoplankton...”

**Answer : Done.**

[17] For Figures 3 and 4. I suggest the authors to sort the species code from the highest to the lowest biovolume, if possible; e.g., Figure 3. A3: 028 -> 012 -> 027 -> 026 -> 140.

**Answer :** We decided to order the species by their codes to ease the comparison between plots, e.g., spp. 140 is always the first of the clump I. If the reader do not find species 140 in the first position, it means species 140 was not observed in the given season/stretch.

[18] lines 446-448. For consistency,  $R^2$  values might be mentioned following the same format than the one used in Tables 3 and 4.

**Answer :** Done.

[19] Table 4 : “*Dependent variable: log<sub>10</sub> Biovolume*”.

**Answer :** Done

## **Discussion**

[20] The ‘Discussion’ section is very interesting and well-written. I found that the link between the results, previous studies and theory is clearly highlighted by the authors. If of interest I suggest the authors to consider the paper “**Size differences predict niche and relative fitness differences between phytoplankton species but not their coexistence**” by Gallego et al. (2019; 10.1038/s41396-018-03307) that is quite in line with the conclusions made in this study: “*size is more than a key trait controlling physiological and population-level aspects of phytoplankton, it is also relevant for community-level phenomena such as niche and fitness differences which influence coexistence and biodiversity*”.

**Answer :** Thank you for this suggestion. We included both of the references now in the discussion.

[21] I suggest the authors to add a few lines (or a short paragraph) about potential limitations/assumptions related to their dataset, especially the length of the time-series (could the conclusions be related - only - to the sampling frame or could the conclusions be generalised?) and the consideration of the two contrasted seasons (wet vs dry). It has been shown that annual phytoplankton succession is strongly influenced by the interaction between the niche of species and environmental conditions, at the community level “**Annual phytoplankton succession results from nicheenvironment interaction**” by Caracciolo et al. (2021; 10.1093/plankt/fbaa060). Did the strong -

and erratic – seasonality of phytoplankton species, the phenological signature of each species, influence the analysis and ensuing conclusions?

**Answer:** This is a very interesting question. There are indeed some limitations inherent to our data. These are now summarised in a short paragraph also with suggestion on future directions.