

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 valuable comments. We have now considered all the comments and suggestions made by the reviewers and made corresponding changes to the manuscript. We agreed with most of the points raised by the reviewers and the changes performed, or otherwise explained our reasons. We hope that this revised version of the manuscript meets the standards required by PCI Ecology to be published.

We present below the changes made point-by-point. Should you have any further comments or suggestions, we are willing to consider those as well, of course.

Yours sincerely,

Caio Graco-Roza

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

*The manuscript by Graco-Roza and colleagues, entitled “**Clumpy coexistence in phytoplankton: the role of functional similarity in community assembly**”, investigated both spatial and temporal distributional patterns of a phytoplankton community located in a tropical region of Brazil, the Piabanha river, under the hypothesis of Emergent Neutrality. To do this, they raised two main questions: 1) differences in species biovolume are linked to their functional differences at the clump level and 2) species with higher distinctiveness will show higher relative biovolume within clump; and use a 1-year time series to confront theoretical expectations with field observations. From their analysis, they identified two stable main clumps in body size with species belonging to three MBFGs, and found that among-species differences in biovolume were driven by functional redundancy at the clump level. In addition, the relative biovolume of species was related to species functional distinctiveness at the clump level. The authors concluded from their study that species with similar traits are selected under given environmental conditions.*

*This study is topical – especially in the current context of intense forcing on biodiversity - and can be an interesting contribution to improve our knowledge on how species are organised/distributed in nature and whether it is possible to determine species distributions, in different kind of ecosystems, on the basis of a theoretical framework. While this paper has a potential to be accepted, some important*

*shortcomings have to be clarified or fixed to make it suitable for publication. I here summarise my main concerns:*

1. *Please revise the conclusion in the abstract as no relation with environmental conditions had been clearly investigated; or I have missed that information.*

**Answer:** We now included in the manuscript explicit information about the relationship between phytoplankton and environmental conditions. We also modified the abstract accordingly.

2. *It is quite unclear to me what is behind the term “clump”. As defined by the authors, a clump is an aggregation of species. However, it could be useful to provide a more detailed and precise definition of clumpy patterns, to mention whether they are stable or episodic patterns, to discuss the relation between species’ niche and clumps (e.g. Segura et al. 2011), etc...*

**Answer:** We now added more information on the definition of the clump, in the abstract, introduction and also an operative definition in the M&M section. For each community, the niche axis was divided into equally spaced segments of one unit of  $\log_2$  volume, because they exhibit more clearly the structure of clumps and gaps. This led to 14 size-segments that were tested separately for larger entropy values with respect to a uniform distribution (now explained in the methods). Hence, following (Fort et al. 2010), we considered the significant segments as the clumps, and in the case of two or more consecutive segments being significant they were considered as a single clump. This is in agreement with the theory which suggest that the niche axis should comprise clumps followed by gaps. Please, find the information about the clump definition in the methods section lines 207– 224.

3. *While the Introduction section provide useful information for the readers, more attention should be devoted to the different theory (for example, by defining the main differences between the ecological sensu Hutchinson and the niche following the neutral theory); I sincerely trust that it could help to clarify the main message of the paper.*

**Answer:** We have included information about the niche-based and neutral-based view of community assembly now in the lines 52-60.

*My feeling when reading this section was that more precise information would help to strengthen the paper; when previous works are cited, it could be helpful to be more precise; e.g. “from a wide range*

*of ecosystems”: which kind of ecosystems? Has this theory been tested on different kind of ecosystems (e.g. polar, temperate, etc...) or the same kind of ecosystems in different areas?*

**Answer:** We thank you for the comment. We have modified the text to improve clarity, including specific examples where emergent neutrality has been detected. Lines 65-68.

In the Introduction section, some sentences need to be redrafted for clarity; e.g. *“These predictions were tested theoretically at both steady (Fort et al. 2010) and fluctuating resource conditions (Sakavara et al. 2018), but beyond robust theoretical results, empirical evidence of EN is still scarce.”* I do not understand the dichotomy made by the authors. The same applies for: *“However, even small differences in organismal traits are linked to species performance in the niche space (McGill et al. 2006; Violle et al. 2007).”*

*“[...] disclose the trigger of stochasticity driving the species coexistence”*: Please, clarify this sentence.

Please clarify the following sentence (e.g. by providing an example from the literature): *“However, although two functionally redundant species most likely show high similarity in trait combination, functional distinctiveness is not directly linked to redundancy”*.

**Answer:** We rephrased the sentences. Now in lines 71-74, 75-80, and 86-91.

*“Despite the importance of body size, considering only a single trait as a proxy for the niche might obscure clumpy patterns (D’Andrea et al. 2018)”*: I disagree with the word ‘obscure’. To me, using only a single trait, depending on which trait, only allow to discern a small deterministic part of how species are arranged in an ecosystem, but using too many dimensions may obscure the patterns because of numerical bias such as the influence of multicollinearity or the use of non-discriminant dimensions. Note that *“D’Andrea et al. 2018”* has been cited two times in the reference list at the end of the paper.

**Answer:** Thank you for your suggestion. There is always a tradeoff between synthesis and realism, and this has been a cornerstone discussion in ecology. Clumps were generally defined based on a single “master” trait (i.e., Body size). It has been proposed that the use of a single trait may impair (“obscure”) the discover of hidden niches (Barabás et al. 2013). Here we use a multidimensional classification of phytoplankton (MBFG) and explicit measure of multiple traits to study the differences among species within and among clumps. We believe that only this way we can bring new empirical evidence on the quasi-neutrality state of species within clumps. The references were corrected, thank you for pointing to this.

*“[...] by comparing species’ functional differences in high resolution [...]”*: Please, clarify ‘in high resolution’

“[...] because competition should be stronger within clumps rather than among clumps [...]”: Please, add a reference to that statement.

**Answer:** As the result of the modification of the introduction, these sentences were removed.

4. My major comment (the most critical point to me) is about the clarity/level of detail of the analyses and on how the results are presented in their current form. In essence, the ‘Results’ section is too short without a required level of explanations.

*Please consider providing more information on (1) the dataset used in this study (Are environmental parameters - and which - considered in the study?) and (2) how analyses have been performed to help readers interpret the results. A clearer guidance from the authors on how analyses were performed would help readers to better understand the paper and would allow researchers to reproduce analyses. I had some difficulties to understand the different steps of the methodology. This highly influence the understanding of the results and therefore of the ‘Discussion’ section. A map to locate the different sampling sites, and brief explanations on the biological data (only phytoplankton samples) would be useful.*

**Answer:** The methods section is now entirely re-written to provide the links between the analysis and the hypotheses. Furthermore, we have included one new analysis in the actual version to test the relationship between species traits and the environment. A map of the sampling stations is now included, as well as the explanation of the environmental variables measured, and phytoplankton samples taken and their analysis.

“Here, we measured species individual [...]”: It is not clear to me if the measure were performed on all individuals or on only a fraction of the samples.

**Answer:** We measured at least 20 individuals of each species whenever possible to estimate the body size (i.e., mean individual volume), surface, surface volume ratio, and maximum linear dimension. Other traits were observed during the abundance estimation, precisely, the presence of flagella, mucilage, siliceous exoskeleton, and aerotopes. We have now included these details in the text lines 177-185.

“Functional traits were used to classify the species into MBFG”: a table, in the main text, to resume the number of species per MBFG, site and season would be useful. This applies for the result: the authors did not mention enough the main/typical species that were affected into each (the three main) MBFG in order to make a better link between their results and the functioning of the study area.

**Answer:** We added this information to the results section. Lines 261-301 and Table 1.

*“The niche axis was divided into equally spaced fractions”*: what could be the consequence of this fractioning on the results?

**Answer:** The referee points to a relevant issue. However, the results are robust to moderate changes in the number of classes. We partitioned the niche axis to identify the clump sections. This method was tested theoretically (Fort et al. 2010) and empirically (Segura et al. 2011, 2013). In this paper, we considered each unit of  $\log_2$  volume as a segment of the niche axis because it revealed best the distribution of the clumps, and also agrees with previous observations (Segura et al. 2011).

*“[...] we applied linear regressions between species functional distinctiveness (FDist) and relative biovolume (RBiov) standardized to range between zero and one at clump-level”*. My apologies if I am wrong, but I cannot see the results from linear regressions. In the same way, I do not understand how was evaluated *“the relationship between RBiov and FDist at clump-level across river stretches and seasons”*.

**Answer:** We have now included a table with the results of the linear models in the main text. Differently from the previous version, the  $F_{Dist}$  and biovolume values are not standardized. Functional distinctiveness was calculated using the global species pool (whole species data), so the position of the species in the multidimensional space does not change across time and space. Therefore, the values of distinctiveness are comparable between clumps and it is possible to include both clumps into a single analysis, increasing the number of points per model. Biovolumes were log-transformed ( $\log_{10}$ ) to approximate normality but are not standardized in this version. Spatial and temporal clustering of the data were analyzed separately to avoid nested design on the data. Please find this information now in the lines 355-363 and the Table 2.

*The authors should explain better Figures 1 and 2. Figure 1 could be clarified and improved, for example by splitting the spatial representation from the temporal representation by means of two Figures instead of 5 panels; it could help to clarify the reading.*

**Answer:** We have now divided the figure 2 into two figures (Figure 4 and 5 in the new text).

The captions of the figures were re-written.

*“There were two significant (randomization test,  $p < 0.05$ , Figure 2) clumps [...]”*: how was determined the significance?

**Answer:** We now explain the randomization test. Please find this information in the lines 207-224.

*“The first aggregation (clump I, 24 species) comprises species ranging from 9 to 10 in log2 volume [...]”*. Please, provide 2-3 species in order to establish a better link with ecology”.

**Answer:** We included the name of the species for each clump.

*“The remaining results [...]”*. Please clarify what are the ‘remaining’ results.

*“The mean biovolume of species within clumps differed [...]”*. How was quantified/determined this difference?

**Answer:** The results section was re-written, and these sentences were removed.

*“According to the Mantel tests [...]”*. Results from the Mantel tests were not shown in the text.

**Answer:** We included a table with the results from the Mantel tests in the main text (Table 3).

*Figure 2 is not exploited at all. No explanation was provided for the different components (panels). For example, how were obtained the factorial plans? What is the meaning of the size of the circles on the right panels and how were they calculated? How was calculated the significance of the clumps (results of the tests are not shown in the text)? It is not clear how were obtained the PCoAxes. What is the meaning of looking at “the relative contribution of the six species with the highest biovolume” while overlooking the species that contributed most to the multivariate analyses?*

**Answer:** We have modified the figure 2, now it is divided into figure 4 and 5 (Lines 318 and 340). The captions have a brief explanation of each panel, and we added a reference for the size of the circle in the middle panels. For a matter of spacing we only show the six species with the highest biovolume within the segments, for each of the significant segments.

Furthermore, in the text, we focused on the two most abundant species as both of them hold the highest abundance at clump-level in different seasons and river stretches. We did not discuss the biovolume of the species that were not inside the significant clumps, as the main aim of this paper is to discuss the emergent neutrality through different angles. Moreover, the

most abundant species were mostly found within significant clumps. The significance of the clumps was tested following the steps described in the methods section (Lines 207-224). The axis of the PCoA were obtained following the steps described in the methods section (Lines 245-247).

*To resume, because results are not detailed enough and only partially presented, it is difficult to really understand how the conclusion is supported by the numerical analyses;*

**Answer:** We have re-written the results section as a whole. Please see in the lines 261-367.

*5. While interesting points are developed in the 'Discussion' section, I cannot see the link between the results and the conclusion, and why, for example, "results does (do) not support alternative explanations as pure Neutrality"; the same applies for several points highlighted in the Discussion, and the one I mentioned for the Abstract section (i.e. the relation with environmental conditions). For example:*

*- "Present results showed that the clumps in body size are a conspicuous feature of phytoplankton community structure in riverine systems across seasons and river stretches"*

**Answer:** Our results show a multimodal distribution of species biovolume along the body size axis irrespectively to the season (Figure 4) or river stretches (Figure 5) that we studied. This agrees with previous observations (Segura et al. 2011, 2013) and suggests that clumps are evident features of the phytoplankton community also in riverine ecosystems.

*- "[...] reflecting adaptations to survive in turbulent and mixed conditions". A parallel with species ecological traits must be done to better understand that statement.*

**Answer:** We have now extended our results to include trait x environment relationships.

Please find new text in the lines 261-302.

*- "[...] Our results suggest a strong effect of stabilizing mechanisms at the overall community level as size differences increase". Could the authors explain better which analyses/results allowed us to conclude, this is not evident as results were not detailed enough.*

**Answer:** We have rephrased the sentence. See the new paragraph Lines 404-428

*- "Studies evidencing EN using phytoplankton as the model community in coastal, estuarine, and here, riverine ecosystems presented fluctuations in environmental conditions (Segura et al. 2011, 2013), which is in agreement with recent modeling results (Sakavara et al. 2018)". Please clarify and document such as statement.*

**Answer:** We have rephrased the sentence, now in the lines 394-403.

*“Despite larger species showing lower growth rates (Kruk et al. 2010), the elongated shape provides advantage under turbulent conditions (Reynolds et al. 1994) favoring these species under high flow conditions on nutrient rich environments (Irwin et al. 2006)”*. It could be very helpful, for a better understanding of the paper, to establish clearer link among phytoplankton species encountered in the study area, species traits, functional groups and environmental conditions. Note that parameters representing environmental conditions are available (*“A detailed description of the biological and physico-chemical sampling methodology [...]”*), but not clearly described nor related to phytoplankton communities.

**Answer:** We have now included a new analysis to provide the link between species, functional traits and the environment (Lines 186-205). The results are detailed in the lines 261-302.

Note that the following sentence is incomplete: *“[...] the HDH does not predicts any particular trait distribution and (Vergnon et al. 2009; Ingram et al. 2018).”*

**Answer:** We fixed the sentence. Line 380. Thank you.

## **Reviewer #2**

*note to the editor: I am the so-called “naïve” reviewer on this topic.*

*It was difficult to review it because of the lack of lines and page numbers.*

*The authors test empirically the emergent neutrality hypothesis for the phytoplankton community, which has been mainly tested based on the theoretical level but never empirically. The study used data from a tropical river.*

*The topic is original and interesting, but major revisions are needed. My main concern would be that a lot of results were discussed in the discussion section without being clearly exposed in the results section. This section would clearly deserve more attention and details. The statistical methods used to test the neutrality hypothesis are original but a lot of them are not explained or described.*

**Answer:** Thank you very much. The methods, results and discussion are now mostly re-written in order to increase clarity and facilitate main conclusions drawn from the results.

## **Abstract**

*It would have been interested to give more details on the data in the abstract (e.g. multidimensional trait similarity is too vague)*

**R:** We rephrased the abstract with specific information.

*“two stable main clumps”*: what does “stable” mean here?



**Answer:** We meant that they were observed along the whole period and in the different regions of the river.

## **Introduction**

*“limited observational studies from a wide range of ecosystems”: a bit paradoxical, please explain; precise which ecosystems? what are the main results of the theoretical studies?*

**Answer:** We have modified the text to improve clarity including typical examples where evidence for EN was found.

*“due to their highly **speciose** communities”: typo to correct*

**Answer:** Thank you. Fixed.

*Functional difference and distinctiveness: better explain here the differences between the two indices*

**Answer:** We have improved this section to explain the differences. The functional redundancy of a species is a pairwise metric that measures the proximity of species within the functional space, while the functional distinctiveness measures the degree that each species differs from the mean trait value.

*River Continuum and Flood Pulse: better explain the link between these two concepts and your study*

**Answer:** The river continuum concept and flood pulse theories explain the distribution of multiple taxa along a river channel. Such theories highlight the role of physical and chemical variables driving the distribution of phytoplankton abundance, and the biotic interactions are often discarded. However, the distribution of phytoplankton based on species size, and the balance between neutrality and niche-based processes is not deeply investigated. We, therefore, believe that investigating the riverine phytoplankton communities under the lens of emergent neutrality predictions would help to unveil the role of habitat filtering and stochasticity on phytoplankton communities. We modified the text to include this concept (line ~117).

*I am not sure to fully understand H2; what is the reproductive success for phytoplankton?*

**Answer:** We have modified this sentence to make it clearer.

## **Methods**

*A map of the sampling stations would be necessary*

*“We set three river stretches”: Are the stretches the same along time? How the nine stations are related to the stretches? Maybe add the profiles in suppl. information file?*

**Answer:** The river stretches were defined based on the elevational profile of the river and, therefore, remains constant through time. We included a map that shows how the stations are distributed across the river stretches (Figure 1).

*“A detailed description of the biological and physico-**chemical** sampling methodology....”: typo to correct*

*Sampling is not clear: which device? How were the traits measured? How many individuals per sample were measured?*

*Please better show the number of samples and replicates used in this analysis*

### Statistical methods:

*This section is globally clear, but a lot of details are needed to fully understand what was used (e.g. how the functional space was obtained).*

*We better understand the difference between functional difference and distinctiveness in the methods section (but better explain it in the introduction section).*

### Clumpy pattern

*I did not fully understand how and why the niche axis was divided.*

### Functional distinctiveness

*Please better explain how the functional space was obtained.*

We thank for the comments, we have improved the M&M section, adding relevant details to fully reproduce the work and calculations.

**Answer:**

- (i) we added new analysis (RLQ and fourth-corner methods) to study the relationship between species traits and environmental variables. We believe that this could help readers to better understand the discussion about the observed patterns.
- (ii) All the analyses and statistical methods applied have their equations and reasoning extensively explained in the methods section.
- (iii) The methodology to identify clumps and divide species volume into equally spaced segments is now described in the Methods section

## Results

*MBFG groups: large size overlapping; a maybe naïve question arises: is it thus relevant to use biovolume as the main niche axis while it is so widely distributed among MBFG groups?*

*The three groups comprised 87% of a total of 150 recorded species; and in terms of abundance?*

*Why are there different species in the clumps by stretches and seasons? Why showing two clumps while one or three clumps seemed to appear sometimes?*

*Figure S1: please consider some points which seem to be an outlier?*

*The lack of significance in clump II comes from a smaller replicates number*

*The results section would deserve more details. A lot of results were discussed in the discussion section without being clearly shown or explained. This led to a global feeling of too speculative results.*

**Answer:** Following the reviewer's suggestions, we have modified the whole results section to be clearer. We believe the results are now informative enough to sustain the discussion section.

- (i) Regarding body size, this is exactly what we are testing in our study. There is overlap in body size among groups, but nevertheless, it seems to be a good proxy for phytoplankton niche preferences. It is worth noting that there is partial overlap among groups, and it is a relevant characteristic, but the use of other dimensions allows for a

better understanding of the ecological processes shaping phytoplankton community structure. This finding was also outlined by Kruk et al 2010, Segura et al. (2011); etc.

- (ii) The clumps are now explained in detail. We divided the  $\log_2$ -volume axis into equally spaced octave bins (we first  $\log_2$  transformed species volume) and tested whether each segment was significantly higher than random reshuffling of species biovolume. A segment significantly higher in diversity than chance is considered a clump (more details in the methods section);
- (iii) We have now included a table with the results of the linear model in the main text (in substitution to the Figure S1). Differently from the previous version, the  $F_{\text{Dist}}$  and biovolume values were not standardized. Please see response to R1. Spatial and temporal clustering of the data were analyzed separately to avoid nesting design on the data. Please find this information now in the lines 355 – 363 and the Table 2.

## **Discussion**

*“The HDH does not predict any particular trait distribution and”; please finish the sentence*

**Answer:** This sentence was corrected in the text. Thank you.

*In the second paragraph, this is not clear why the TDH could not be one other explanatory theory of the clumpy aggregations found in the study.*

**Answer:** The TDH suggests that the environment filters organisms towards self-organization into modes that vary across scales, leading to discontinuities in their body size distributions. Thus, the distribution of the modes should be landscape-dependent and not a feature from the organisms themselves. This means that species from different modes are not expected to differ in their physical and behavioral characteristics (here functional traits), and modes are not expected to be similar in landscapes that differ environmentally (e.g. rivers vs. lakes). However,

our results show that species from different modes (i.e., clumps) differ in their trait composition, which is observed by the MBFG composition of the modes. Furthermore, the body size structure of our data from a riverine ecosystem is similar to the body size structure observed in data from brackish waters. This suggests that the body size structure is more a feature from the organism than the landscape, and points to the emergent neutrality as a plausible mechanism to explain our findings. We have included this explanation in the discussion section.

*“not surprisingly.....”: an example of results that is not clearly shown in the Results section*

*“.... we observed a niche-based driven assembly process at clump I”: please better explain the results leading to this conclusion.*

**Answer:** We have re-written the discussion section. This sentence was removed and explained in a better way. Now in the lines 404-428

*The predation as a limiting factor controlling body size is not discussed.*

**Answer:** We did not discuss the predation as a controlling factor because top-down effects are not usually described as important to phytoplankton in riverine ecosystems (Reynolds et al. 1994; Reynolds 2000; Everbecq et al. 2001).

## **References**

Barabás G, D’Andrea R, Rael R, et al (2013) Emergent neutrality or hidden niches? *Oikos*.

<https://doi.org/10.1111/j.1600-0706.2013.00298.x>

Everbecq E, Gosselain V, Viroux L, Descy J-P (2001) Potamon: a dynamic model for predicting phytoplankton composition and biomass in lowland rivers. *Water Research* 35:901–912. [https://doi.org/10.1016/S0043-1354\(00\)00360-2](https://doi.org/10.1016/S0043-1354(00)00360-2)

Fort H, Scheffer M, van Nes E (2010) The clumping transition in niche competition: A robust critical phenomenon. *Journal of Statistical Mechanics: Theory and Experiment* 2010:.  
<https://doi.org/10.1088/1742-5468/2010/05/P05005>

Reynolds CS (2000) Hydroecology of river plankton: the role of variability in channel flow. *Hydrological Processes* 14:3119–3132. [https://doi.org/10.1002/1099-1085\(200011/12\)14:16/17<3119::AID-HYP137>3.0.CO;2-6](https://doi.org/10.1002/1099-1085(200011/12)14:16/17<3119::AID-HYP137>3.0.CO;2-6)

Reynolds CS, Descy JP, Padisák J (1994) Are phytoplankton dynamics in rivers so different from those in shallow lakes? *Hydrobiologia* 289:1–7.  
<https://doi.org/10.1007/BF00007404>

Segura AM, Calliari D, Kruk C, et al (2011) Emergent neutrality drives phytoplankton species coexistence. *Proceedings of the Royal Society B: Biological Sciences* 278:2355–2361. <https://doi.org/10.1098/rspb.2010.2464>

Segura AM, Kruk C, Calliari D, et al (2013) Competition drives clumpy species coexistence in estuarine phytoplankton. *Scientific Reports* 3:1–6. <https://doi.org/10.1038/srep01037>