# Clumpy coexistence in phytoplankton: The role of functional similarity in community assembly

Caio Graco-Roza [1, 4]*, Angel M Segura [2], Carla Kruk [3], Patrícia Domingos [4], Janne Soininen [1], Marcelo Manzi Marinho[4]

[1] University of Helsinki, Department of Geosciences and Geography, PO Box 64, FI00014, Helsinki, Finland.

[2] Modelización Estadística de Datos e Inteligencia Artificial (MEDIA) CURE-Rocha, Universidad de la República, Uruguay.

[3] Sección Limnología, IECA, Facultad de Ciencias, Universidad de la República, Uruguay.

[4] Laboratory of Phytoplankton of Ecology and Physiology, Department of Plant Biology, University of Rio de Janeiro State, Rua São Francisco Xavier 524—PHLC Sala 511a, 20550-900, Rio de Janeiro, Brazil. Telephone: +552123340822

*correspondence author: caio.roza@helsinki.fi

**Abstract**

Emergent neutrality (EN) suggests that species must be sufficiently similar or sufficiently different in their niches to avoid interspecific competition. Such a scenario results in a transient pattern with clumps and gaps of species abundance along the niche axis (*e.g.*, represented by body size). From this perspective, clumps are clusters of species with negligible fitness differences, and therefore, are subjected to stochastic abundance fluctuations. Plankton is an excellent model system for developing and testing ecological theories, especially those related to size structure and species coexistence. We tested EN predictions using the phytoplankton community along the course of a tropical river considering (i) body size structure, (ii) functional clustering of species in terms of morphology-based functional groups (MBFG), and (iii) the functional similarity among species concerning their functional traits. Two main clumps in the body size axis (clump I and II) were conspicuous through time and were detected in different stretches of the river. Clump I comprised medium-sized species from the MBFGs IV, V, and VI while clump II included large-bodied species from the MBFGs V and VI. Pairwise differences in species biovolume correlated with species functional similarity when the whole species pool was considered, but not among species within the same clump. Although clumps comprised multiple MBFGs, the dominant species within the clump belonged always to the same MBFG. Also, within-clump species biovolume increased with functional distinctiveness considering both seasons and stretches, except the lower course. These results suggest that species within clumps behave in a quasi-neutral state, but even minor shifts in trait composition may affect species biovolume. In sum, our findings point that EN belongs to the plausible mechanisms explaining community assembly in river ecosystems.

*Keywords*: emergent neutrality; functional distinctiveness; functional similarity; species coexistence

## Introduction

Understanding the mechanisms promoting species coexistence and shaping community structure has been a long-standing goal in community ecology. The former idea that the number of coexisting species is limited by the number of growth-limiting resources or niche dimensions (Gause 1936, Hardin 1960) and its derivate idea, "*the paradox of the plankton"* (Hutchinson 1957), have been widely explained in terms of endogenous and exogenous Spatio-temporal mechanisms (Roy and Chattopadhyay 2007). Trait-based approaches are useful to test this matter due to their potential to generalize patterns beyond species' identity, especially because traits influence the species' ability to acquire resources and persist through environmental changes (McGill et al. 2006, Díaz et al. 2013, 2016). Nonetheless, the niche-based theory proposes that the environment filters community composition through species' ecological requirements, which can be perceived through species' traits. Also, intra- and inter-specific interactions potentially drive community assembly, in local communities (Götzenberger et al. 2012). In contrast, the more recent neutral theory suggests that diversity results from random dispersal, speciation, and extinction rates with no role of niche differences in species coexistence (Hubbell 2001). This type of dynamics should then result in a random distribution of functional traits along environmental gradients (Kraft et al. 2008, Cornwell and Ackerly 2009).

More recently, it was shown that community organization is driven by eco-evolutionary processes such as speciation and nutrient uptake kinetics resulting in groups comprising different species with similar ecological requirements (Gravel et al. 2006, Scheffer and van Nes 2006, Hubbell 2006). This finding led to the 'emergent neutrality hypothesis' (EN; Holt 2006) that has been supported by observational studies, e.g., for phytoplankton from brackish waters (Segura et al. 2011), birds from the North of Mexico (Thibault et al. 2011) and beetles

69 at the global scale (Scheffer et al. 2015). EN suggests that species must be sufficiently
70 similar, and thus, behave neutrally, or different enough in their niches to avoid competition.
71 Such a scenario would result in species-rich aggregations or clumps along the niche axis
72 (Scheffer and van Nes 2006, Vergnon et al. 2009, Fort et al. 2010). Modelling studies have
73 shown that such predictions apply for both steady environmental conditions (Fort et al. 2010),
74 and also fluctuating resource conditions (Sakavara et al. 2018). Empirical evidence about EN
75 is still scarce, however (Scheffer et al. 2018).

76 The clumpy pattern arises from the exceedingly slow displacement rate of species under
77 intense competition, that is, species within the same clump overlap in their niche such that the
78 displacement rate of competing species is similar to the competition at the intraspecific level,
79 leading to stochastic fluctuations in species abundances through time (Scheffer et al. 2018).
80 Thus, the number of clumps corresponds to the number of species to be expected to stably
81 coexist at equilibrium, but the identity of the dominant species is expected to be random
82 among the clump residents. However, the assignment of species to clumps is challenged by
83 the fact that trait differences among species are continuous (Villéger et al. 2008) and the
84 threshold to include a species within a clump varies with the statistical approach that is
85 applied (Segura et al. 2011, D'Andrea et al. 2019). Therefore, it is difficult to state
86 empirically whether species behave neutrally within clumps (i.e., when the strength of
87 interspecific interactions equals the intraspecific interactions) or if results are an artefact of
88 clump construction.

89 Zooming in on the uniqueness of trait combinations of species, i.e. functional distinctiveness,
90 within clumps may advance our comprehension of biotic interactions and move towards a
91 measurable value of similarity at which species coexistence is driven stochastically.
92 Functional distinctiveness reflects the non-shared functions among species within a given

93 species pool (Violle et al. 2017), mirroring the concept of functional similarity (Pavoine et al.
94 2017). However, functional distinctiveness is not directly linked to functional similarity at the
95 pairwise level (Coux et al. 2016, Ricotta et al. 2016, Violle et al. 2017). For example, two
96 species may be equally distinct, i.e. the degree to which a species differs from all the others
97 within the species pool concerning their functional traits, and still not be similar in their trait
98 composition at a pairwise level (Coux et al. 2016). This suggests that both pairwise functional
99 similarity and group-based functional distinctiveness are complementary metrics to assess the
100 role of trait combination in community assembly. To this end, phytoplankton communities
101 are useful for biodiversity theory testing due to their species-rich communities, rapid
102 responses (in human time-scales) and well-characterized relationships between morphology
103 and physiological and ecological responses (Litchman and Klausmeier 2008, Kruk and
104 Segura 2012, Litchman et al. 2012).

105 Body size is considered a master ecological trait and it is often used to characterize species
106 niche differences (Downing et al. 2014). In phytoplankton, the body size is related to
107 physiology and life-history (Litchman and Klausmeier 2008), photosynthetic processes
108 (Marañón 2008), nutrient uptake kinetics (Litchman et al. 2010) and other eco-evolutionary
109 processes, e.g. the relationship among predation rates, nutrient uptake and organisms body
110 size (Sauterey et al. 2017). Although body size may relate to different processes, using a
111 single trait as a proxy for niche differences may not evidence species differences generated
112 by hidden/unknown niche axes (i.e. ecological dimensions of the niche) and impair the
113 understanding of clumpy patterns (Barabás et al. 2013, D'Andrea et al. 2018). The use of
114 multiple traits emerges as a powerful tool to disentangle phytoplankton functional structure
115 and evaluate competing hypotheses (Reynolds et al. 2014, Chen et al. 2015, Bortolini and
116 Bueno 2017, Aquino et al. 2018). Morphology-based functional groups (MBFG)
117 classification of phytoplankton species (Kruk et al. 2010) is a multidimensional combination

of morphological traits that cluster organisms into seven groups with similar physiology and ecological responses, potentially overcoming the limitations of using a single trait dimension only. Assessing the functional distinctiveness of species within the same functional cluster (e.g., clumps, MBFGs) could help to study the existence of functional equivalence (i.e., neutrality) among species. Overall, the functional similarity among species is a useful tool to compare species in a multidimensional space, particularly because the environment may filter different functional traits across space and time (D'Andrea et al. 2020).

Rivers are highly heterogeneous systems characterized by a continuous water flow that affects the ecosystem's morphology (e.g., meandering), sedimentation patterns, organisms' dispersal, and more specifically the phytoplankton abundance and distribution (Reynolds and Descy 1996, Wetzel 2001). Several theories, e.g., the River Continuum (Vannote et al. 1980) and Flood Pulse (Junk et al. 1989) concepts explain the longitudinal distribution and abundance of riverine phytoplankton communities. However, an explicit study of communities' body size structure and species coexistence under EN in riverine ecosystems is lacking. For example, phytoplankton species should attain higher biomass at the middle reaches or in the upper reaches of low-gradient stretches (Descy et al. 2017). Also, competition rates vary along the river course because water turbulence reduces the likelihood of biotic interactions (Reynolds et al. 1994), meaning that clumpy coexistence may not be observed in riverine phytoplankton. Alternatively, if functional trait combinations of species within the local species pool result from eco-evolutionary processes (Scheffer et al. 2015), the clumpy pattern should also be apparent in riverine phytoplankton communities. Here, we push forward three hypotheses to be tested in a tropical river by investigating phytoplankton community size structure both seasonally and spatially. We expect that:

**$H_1$** – There are peak aggregations of species abundance (i.e., clumps) along the body size axis of phytoplankton in the river that remain constant across space and time as a result of eco-evolutionary processes.

**$H_2$** – Pairwise-differences in species abundances increase with functional dissimilarity at the community-level but not at the clump level because species within the same clump behave in a quasi-neutral state. Thus, the dominance within clump varies stochastically between species as fitness differences are negligible.

**$H_3$** – Species abundance increases with functional distinctiveness with respect to other species within the clumps. Although abundance fluctuates stochastically at the pairwise level, the number of species bearing similar trait combinations may affect the likelihood of the interactions within clumps. Therefore, species with the most distinct trait combinations concerning their clump peers are less likely to share the same ecological requirements, and by consequence, attain higher abundance.

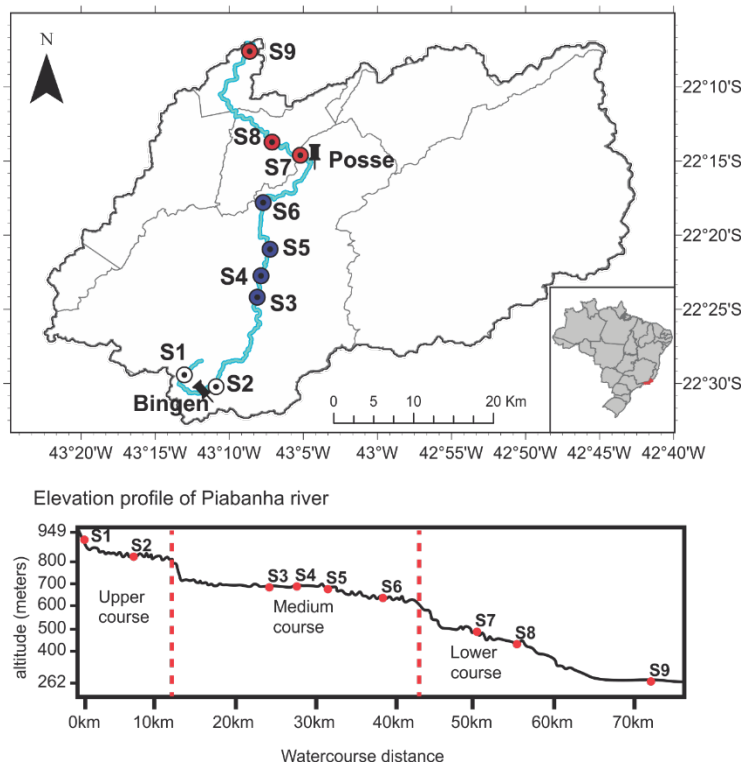
## Methods

### *Study area*

Samples were taken monthly at nine stations along the Piabanha river between May 2012 and April 2013. Piabanha river is in the tropical region of Brazil and has a drainage basin of approximately 4500 km² (Figure 1). The headwater is on Petrópolis at 1546m altitude and drains to the medium valley of Paraíba do Sul river crossing three cities and with agricultural activities in their watershed. We set three river stretches (lower, medium, and upper courses) based on the location of steep slopes the river elevation profile (Figure 1). Data from two meteorological stations (Bingen and Posse; Figure 1), located close to the sampling sites, were used to measure rainfall. We analyzed meteorological data up to three days before each

164 sampling campaign. We classified seasons as a dry season (May - October) and a wet season
165 (November – April) based on the rainfall data.

166




167

168 **Figure 1. Map of the study area.** The watershed area of the Piabanha river showing the river
169 course (blue line), the meteorological stations Bingen and Posse, and the sampling sites are
170 coloured according to river stretches (white circles = upper course, blue circles = medium
171 course, red circles = lower course). The vertical dotted red line in the elevational profile
172 figure indicates the locations of steep slopes used to define the boundaries of the river
173 stretches.

174 *Sampling and sample analysis*

175 In the field, we measured temperature (°C), dissolved oxygen (DO, mg $L^{-1}$), and turbidity by
176 a multiparameter probe sonde (YSI model 600 QS). Water discharge (WD, $m^3$ $s^{-1}$) was

177 measured with the SonTek RiverSurveyor – M9. Furthermore, water samples were taken and
178 kept frozen (one or 2 weeks) until the laboratory analysis for ammonium ($N{\cdot}NH_4^+$, mg $L^{-1}$),
179 nitrate ($N{\cdot}NO_3^-$,mg $L^{-1}$), nitrite ($N{\cdot}NO_2^-$,mg $L^{-1}$), total phosphorus (TP, mg $L^{-1}$) and
180 soluble reactive phosphorus (SRP, mg $L^{-1}$). Ammonium, nitrite, and nitrate were summed up
181 and are expressed as dissolved inorganic nitrogen (DIN, mg $L^{-1}$). The water samples were
182 filtered (except for total phosphorus analysis) using borosilicate filters (Whatman GF/C), and
183 nutrient concentrations were measured following APHA (2005). A complete description of
184 the spatial and temporal patterns of the environmental variables measured in the Piabanha
185 river can be found in Graco-Roza et al. (2020).

186 *Phytoplankton samples*

187 Subsurface samples of phytoplankton were collected with a bottle of 200 mL and fixed with
188 Lugol. In the laboratory, phytoplankton species were identified, and population densities
189 were estimated under an inverted microscope (Olympus CKX41) (Utermöhl 1958). At least
190 100 individuals of the dominant species were counted in each sample (Lund et al. 1958,
191 Uhelingher 1964). Biovolume ($mm^3$ $L^{-1}$) of phytoplankton species was estimated by
192 multiplying the density of each population (ind. $L^{-1}$) by the average individual volume of the
193 species (V, $\mu m^3 org^{-1}$). The volume of each species was estimated by measuring geometrical
194 dimensions and approximating to defined geometrical forms following Hillebrand et al.
195 (1999). Geometrical dimensions were measured in 20 organisms from each species (when
196 possible) and the average was used to characterize individual body size (volume). We recall
197 that biovolume represents the biomass density and volume is an organism's trait. Species'
198 surface area (S, $\mu m^2$) was estimated, the maximum linear dimension (MLD, μm) was
199 measured, and the presence of aerotopes, mucilage, flagella, and siliceous exoskeletal
200 structures. We then used the volume and surface area of the species to estimate the individual

surface volume ratio (SV). Species were then classified into MBFG according to Kruk et al. (2010), based on the above mentioned morphological traits. This classification included the following seven groups: (I) small organisms with high SV, (II) small, flagellated organisms with siliceous exoskeletal structures, (III) large filaments with aerotopes, (IV) organisms of medium size lacking specialised traits, (V) flagellates unicells with medium to large size, (VI) non-flagellated organisms with siliceous exoskeletons and (VII) large mucilaginous colonies. For further details on MBFG classification, we refer to Kruk et al. (2010) or Segura et al. (2013a).

*Statistical analyses*

Statistical analyses were performed on R v.4.0.4 (R Core Team 2020) using the packages 'ade4' v.1.7.16 (Chessel et al. 2004, Dray and Dufour 2007, Dray et al. 2007, Bougeard and Dray 2018, Thioulouse et al. 2018), 'FD' v.1.0.12 (Laliberte et al. 2010, Laliberté et al. 2014), the suite of packages 'tidyverse' v.1.3.0 (Wickham et al. 2019), and the package 'vegan' v.2.5.7 (Oksanen et al. 2020).

*Traits-environment relationship*

We tested the relationship between morphological traits and the environmental variables using a three-table ordination (RLQ) combined with a fourth-corner analysis (Dray et al. 2014) (Figure S1). Both RLQ and fourth-corner methods require the information from three tables; a data frame including the measurements of environmental variables across the sampling sites (R table), a matrix containing species abundances or occurrences across the sampling sites (L table), and a data frame comprising the trait values for each species (Q table). Also, both methods rely on the analysis of the fourth-corner matrix, crossing the information between tables R and Q, weighted by table L. The RLQ analysis (Legendre et al.

1997) provides ordination scores to summarize the joint structure among the three tables, but it does not allow the identification of traits or environmental variables contributing significantly to the structure. The fourth corner method (Dolédec et al. 1996) tests the significance of bivariate associations between each trait and environmental variables but disregards the covariance among traits or environmental variables. Here, we combined the RLQ analysis with the fourth corner method by applying the fourth corner method to the output of the RLQ analysis instead of the original raw values (Dray et al. 2014). By doing this, we summarized the main patterns in the multivariate space and tested the global significance of the trait–environment relationships using the $S_{RLQ}$ multivariate statistic and the fourth corner sequential testing procedure (Dray and Legendre 2008). Applying the fourth corner method in the output of the RLQ determines (a) the relationship between individual traits and RLQ environmental scores (known as environmental gradients), and (b) the relationship between environmental variables and RLQ traits scores (known as trait syndromes) (Dray et al. 2014).

Before applying the RLQ method, we first log-transformed ($\log_{10}$ x+1) species biovolume, species traits (SV, MLD, and V), and environmental variables (except pH and temperature). A correspondence analysis (Benzécri 1973) was performed on the L table using the function dudi.ca from 'ade4', and a Hill-Smith analysis (Hill and Smith 1976) on the R and Q tables separately using the function dudi.hill from 'ade4'. We used Hill-smith analysis because both R and Q table included categorical or binary variables. The RLQ analysis was conducted in the output of the ordinations using the function rlq from 'ade4'. We tested the significance of the joint structure among the RLQ tables using the fourth-corner method with a stepwise permutation procedure of 999 permutations using the function rlq.fourthcorner from 'ade4'. The null hypothesis that given fixed traits, species abundances are independent of environmental conditions was evaluated by permuting sites (rows of tables L or R) while

keeping the species traits (table Q) fixed. The null hypothesis that given fixed environmental conditions, species abundances are independent of functional traits was evaluated by permuting species (columns of table L or rows of table Q) while the environmental conditions (R table) were kept fixed (Dray et al. 2014). Rejecting both null hypotheses imply that tables R, L, and Q are significantly linked. Because the fourth-corner analysis explores one trait and one environmental variable at a time, multiple statistical tests are performed simultaneously increasing the probability of type I error (i.e. false significant associations), thus we adjusted p-values for multiple testing using the false discovery rate method (Benjamini and Hochberg 1995). We divided the value of the fourth-corner correlation by the square-root of the first eigenvalue of the correspondence analysis of the L matrix, which is the maximum possible value (Peres-Neto et al. 2017).

*Clumpy patterns*

To test for the existence of peak aggregations of species biovolume along the body size axis of phytoplankton - $\mathbf{H_1}$, we analyzed the community structure in each season (dry and wet) and river stretches (upper, medium, and lower course) (Figure S1). First, the individual volume of species was log-transformed ($log_2$) and used as the main niche axis (X= $log_2$ volume). Hence, we divided the niche axis into equally spaced segments (one segment per unit $log_2$ volume) and for each segment (j), we estimated the Shannon entropy (H) using the biovolume of the observed species (Fort et al. 2010, Segura et al. 2011). The entropy index was defined as:

$$H_j = \sum_{i=1}^{n} p_i \, log_2(p_i) \quad (1)$$

where $p_i$ is the fraction of biovolume of species $i$ in the community of $n$ species. Finally, we tested the significance of the entropy ($H$) by comparing the observed $H$ against an expected uniform distribution under the null hypothesis of homogeneous $H$. For this, we created 1000

272 communities by sampling the volume of species from a random uniform distribution bounded
273 by observed individual volumes. Then, each species had a biovolume assigned to it, which
274 was taken from randomization of the observed biovolume matrix, keeping both the empirical
275 species rank-biovolume pattern and total biovolume in the sample. For each segment, the
276 observed *H* was compared with the distributions of *H* generated under the null hypothesis,
277 with significance defined according to standard 5% criterion (Fort et al. 2010, Segura et al.
278 2011). Finally, we considered a significant segment or two consecutive significant segments
279 as a clump.

280 *Functional dissimilarity*

281 To test whether differences in species biovolume increases with functional dissimilarity – **$H_2$**,
282 we first calculated the functional dissimilarity and the differences in biovolume among pairs
283 of species using the whole community and using only the species from the significant clumps
284 separately (Figure S1). The functional dissimilarity was obtained by calculating Gower's
285 general dissimilarity coefficient on all the species functional traits, that is, all the traits that
286 showed a significant ($p < 0.05$) relationship with the environmental gradients in the fourth
287 corner method. The dissimilarity coefficient was estimated using the function gowdis
288 from'FD'. We used Gower's dissimilarity (*Gd*) because it can handle mixed variable types
289 (continuous, binary, ordinal, and categorical traits). *Gd* defines a distance value $d_{jk}$ between
290 two species as

291 $$d_{jk} = \frac{1}{N}\sum_{i=1}^{n}\left|\frac{(x_{ij} - x_{ik})}{\max(x_i) - \min(x_i)}\right| \quad (2)$$

292 where, *N* is the number of functional traits considered, $x_{ij}$ the value of trait *i* for species *j*, and
293 $x_{ik}$ the value of the trait *i* for species *k*. Therefore, *Gd* = d*jk* = functional dissimilarity. We thus

tested **$H_2$** by conducting Mantel tests with 999 randomizations on the matrices of functional dissimilarity and differences in biovolume using the function mantel from 'vegan'. We performed the Mantel test considering: a) all species present in a given season or river stretch, and b) separately for the species of each significant clump that were present in a given season or river stretch.

*Functional distinctiveness ($F_{Dist}$)*

To test whether species biovolume increases with functional distinctiveness at the clump-level – **$H_3$**, we estimated the functional distinctiveness ($F_{Dist}$) as the Euclidean distance of a species to the average trait position (centroid) in the multidimensional functional space for the set of species of each of the significant clumps using the equations proposed by Anderson (2006). First, we applied a Principal Coordinates Analysis (PCoA) in the species-by-traits data table using Gower's dissimilarity (*Gd*) and obtained species coordinates in the functional space using all the axes from the PCoA. Hence, $F_{Dist}$ was calculated as

$$F_{Dist} = \sqrt{\Delta^2\left(u_{ij}^{+}, c_{i'j'}^{+}\right) - \Delta^2\left(u_{ij}^{-}, c_{i'j'}^{-}\right)} \quad (3)$$

where $\Delta^2$ is the squared Euclidean distance between $u_{ij}$, the principal coordinate for the *j*th species in the *i*th clump, and $c_i$, the coordinate of the centroid for the *i*th clump. The super-scripted '+' and '-' indicate the real and imaginary parts respectively (see Anderson 2006, for details). We did not weight the clump-centroid by species biovolume because it would artificially give higher distinctiveness for less abundant species and bias our analysis. Besides, we calculated $F_{Dist}$ using only species from the significant clumps and normalized the $F_{Dist}$ to range between zero and one by diving the actual $F_{Dist}$ values by the $F_{Dist}$ of the most distinct species of the clump. We tested the **$H_3$**, by modelling the relationship between species biovolume and $F_{Dist}$ using linear models. We used $\log_{10}$ biovolume as the dependent

variable, with $F_{Dist}$ and Clump (i.e. the clump to which a species belong) as the independent variables for each season and river stretch separately.

## Results

Our samples included 150 species that were classified in six (MBFG I, III, IV, V, VI, and VII) from the seven MBFGs based on their functional traits (Table 1). MBFGs IV, V, and VI included 87% of the total number of species. Species from MBFG IV included filamentous, colonial, and unicellular species ranging from 21 µm³ to 8181 µm³ lacking specialized morphological traits (e.g. flagella, siliceous exoskeletal structures). MBFG V comprised unicellular flagellated species ranging in volume from 31µm³ to 31864 µm³, and MBFG VI included unicellular and chain-forming species with a siliceous exoskeletal body that ranged in volume from 48µm³ to 19045µm³.

**Table 1. Distribution of species among the morphological-based functional groups.**

| MBFG | Number of species | Representative taxa |
|---|---|---|
| I | 9 | *Chroococcales* sp., *Chroococcus* sp. |
| III | 3 | *Limnothrix* sp. |
| IV | 60 | *Pseudanabaena limnetica*, *Pseudanabaena catenata* |
| V | 13 | *Euglena* sp., *Cryptomonas* sp. |
| VI | 57 | *Cymbella* sp., *Synedra* sp. |
| VII | 8 | *Dictyosphaerium* sp. |
| Total | 150 | |

Regarding the trait-environment relationship, the first two RLQ axes preserved well the variance of the ordinations and explained altogether 84.37 % of the variation, with 66.71% corresponding to the first axis alone. The $S_{RLQ}$ statistic indicated a significant global relationship between trait syndromes and environmental gradients ($r = 0.18$, p-value < 0.01). Mainly, the first trait syndrome (RLQ axis $1_{Trait}$) correlated significantly with the first environmental gradient (RLQ axis $1_{Environment}$; $r = 0.17$, $p < 0.01$) while the second trait

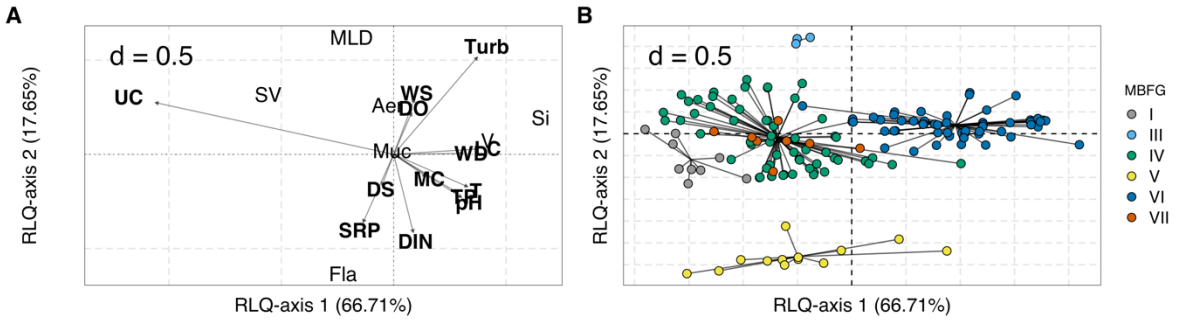
syndrome (RLQ axis $2_{Trait}$) correlated significantly with the second environmental gradient (RLQ axis $2_{Environment}$; $r = 0.12$, $p < 0.01$ ). Yet, there was no significant relationship between the first trait syndrome and the second environmental gradient, nor between the second trait syndrome and the first environmental gradient (Table 2).

**Table 2. Combined fourth-corner-RLQ analysis to test the relationship between functional syndromes (RLQ axis$_{Trait}$) and environmental gradients (RLQ axis$_{Environment}$).**

| | RLQ axis $1_{Trait}$ | RLQ axis $2_{Trait}$ |
|---|---|---|
| | Pearson's *r* (adjusted p-value) | |
| RLQ axis $1_{Environment}$ | **0.17 (< 0.01)** | 0.00 (1.00) |
| RLQ axis $2_{Environment}$ | 0.00 (1.00) | **0.12 (< 0.01)** |

Results of the RLQ showed that most of the biovolume-based variation in trait syndromes were related to the flow regime of the river Piabanha. The first RLQ axis summarised the spatial gradient in the environmental conditions, specifically the increase in turbidity, temperature, pH, and water discharge from the upper to the lower courses while the second RLQ axis summarised the temporal gradient with the smaller nutrient concentrations and higher water turbidity in the wet season (Figure 2A). Noteworthy, the spatial and temporal gradients in abiotic conditions coupled with the distribution of species from different MBFGs. The MBFGs I, III, and IV attained higher abundances in the upper course, contrasting with MBFG VI that showed the highest abundances at the lower course (Figure 2B). Regarding the temporal gradient, MBFG V had higher abundances during the dry season (Figure 2B). Indeed, the fourth-corner method showed that the first trait syndrome correlated positively with turbidity ($r = 0.16$, $p = 0.01$), temperature ($r = 0.14$, $p = 0.02$), pH ($r = 0.13$, $p = 0.03$), water discharge ($r = 0.13$, $p = 0.03$) and total phosphorus ($r = 0.12$, $p = 0.04$), and negatively with the upper course ($r = -0.21$, $p < 0.01$; Figure 2C). Besides, the second trait

syndrome correlated negatively with dissolved inorganic nitrogen ($r$ = -0.11 , p = 0.04 ) and positively with turbidity ($r$ = 0.13, p = 0.01; Figure 2C). For the spatial environmental gradient, there was a positive correlation with the species volume ($r$ = 0.15 , p = 0.01) and the presence of siliceous exoskeletal structures ($r$ = 0.22, p < 0.01), and a negative correlation with surface volume ratio ($r$ = -0.19, p < 0.01; Figure 2D). For the temporal environmental gradient, there was a positive significant correlation with species maximum linear dimension ($r$ = 0.12, p = 0.01), and a negative significant correlation with the presence of flagella ($r$ = -0.12, p = 0.01; Figure 2D).




C

Trait syndromes

| Environmental variables | RLQ axis $1_{Trait}$ | RLQ axis $2_{Trait}$ |
|---|---|---|
| WD | **0.13** | 0.001 |
| T | **0.138** | -0.045 |
| pH | **0.127** | -0.059 |
| Turb | **0.156** | **0.134** |
| SRP | -0.058 | -0.094 |
| DIN | 0.036 | **-0.108** |
| DO | 0.033 | 0.058 |
| TP | **0.118** | -0.052 |
| UC | **-0.209** | 0.033 |
| MC | 0.06 | -0.03 |
| LC | 0.108 | 0.004 |
| WS | 0.029 | 0.058 |
| DS | -0.029 | -0.058 |

D

Environmental gradient

| Traits | RLQ axis $1_{Environment}$ | RLQ axis $2_{Environment}$ |
|---|---|---|
| Aer | -0.009 | 0.05 |
| Muc | -0.003 | 0.004 |
| Si | **0.22** | 0.037 |
| Fla | -0.075 | **-0.121** |
| MLD | -0.063 | **0.12** |
| SV | **-0.187** | 0.06 |
| V | **0.141** | 0.014 |

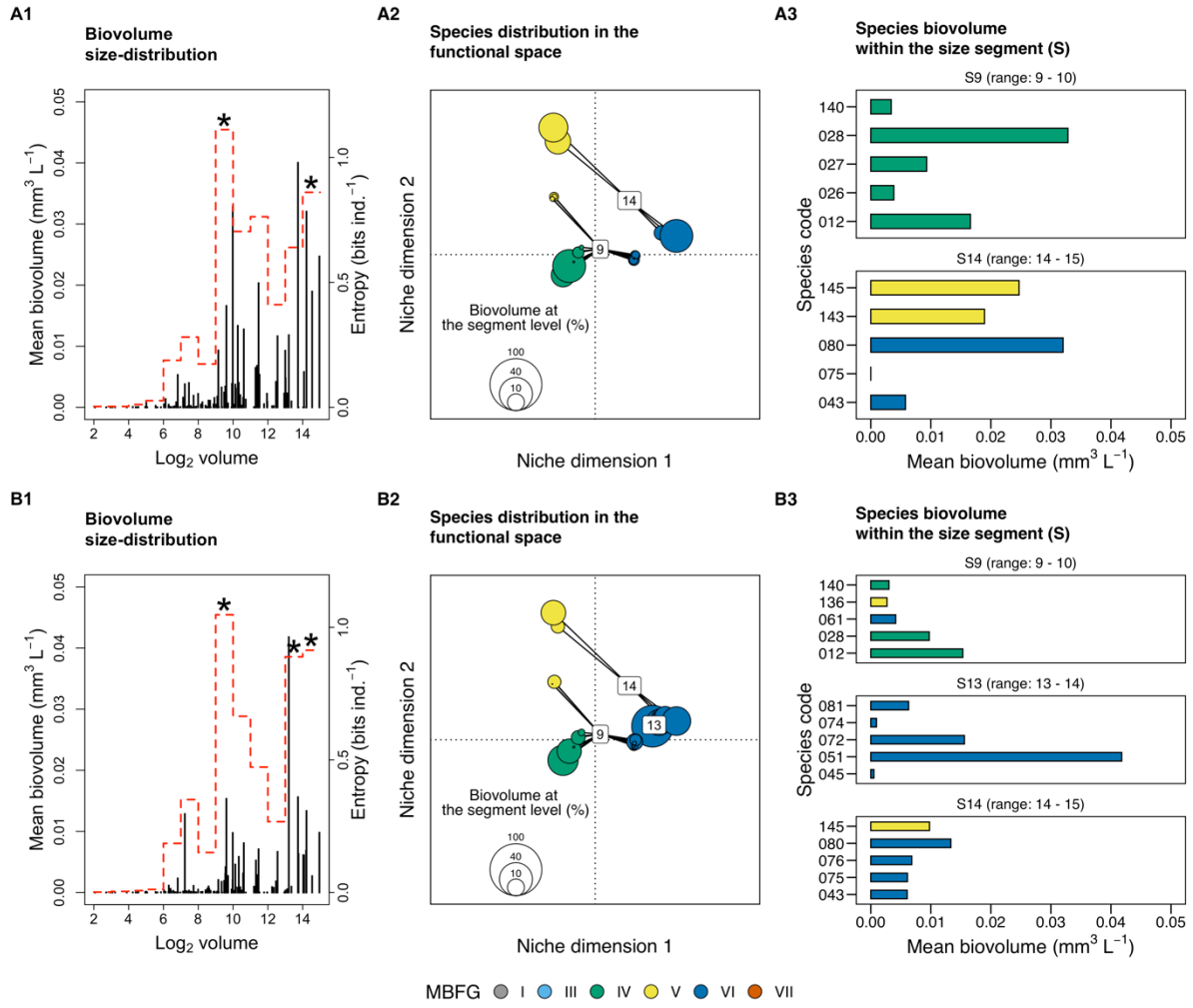
**Figure 2. Results of the (A, B) RLQ ordination and (C, D) hypothesis testing through fourth-corner analysis.** A) The relationships between species traits and environmental variables. B) The distribution of species in the functional space. Each point in the ordination plot represents the position of a species modelled according to its traits on RLQ axes 1 and 2. The black lines connect the species to the centroid of its morphology-based functional groups - MBFG. Colours represent MBFGs, C) The correlation between species traits and the environmental gradients (RLQ axis $_{environment}$), and D) the relationship between environmental

variables and the trait syndromes (RLQ axis $_{trait}$). The grey boxes in C and D indicate significant relationships, and the values within the boxes indicate Pearson's *r*. Aer, aerotopes; Muc, mucilage; Si, siliceous exoskeletal structures; Fla, flagella; MLD, maximum linear dimension; SV, surface volume ratio; V, volume; WD, water discharge; T, temperature; Turb, turbidity; SRP, soluble reactive phosphorus; DIN, dissolved inorganic nitrogen; DO, dissolved oxygen; TP, total phosphorus; WS, wet season; DS, dry season; UC, upper course; MC, medium course; LC, lower course.

Overall, species volume ranged from 4.19 µm³ to 31864 µm³ totalling 14 equally spaced segments (S) of volume along the niche axis. From the 14 segments, three of them showed significant ($p < 0.05$) entropy values, specifically S9, S13, and S14. This resulted in a biovolume aggregation (i.e. clumps) in two regions of the niche axis considering both temporal (Figure 3; Column 1) and spatial categories (Figure 4; Column 1). The first clump included 24 species from the MBFGs IV, V, and VI at the range of S9 (512µm³ - 1024µm³), particularly eight species from MBFG IV (e.g. *Pseudanabaena catenata* and *P. limnetica*), four species from the MBFG V (e.g. *Strombomonas* sp., *cf Cryptomonas* sp.), and 12 species from MBFG VI (e.g*., Fragillaria capuccina* var. gracilis, *Achnantes cf. rupestoides*). The second clump (hereafter Clump II) included six species at the range of S14, being two species from MBFG V (i.e. *Euglena* sp.) and four species from MBFG VI (e.g. *Pinnularia* sp., *Synedra* sp.). Moreover, during the wet season, the S13 also had significant entropy values and six more species from MBFG VI (e.g., *Achnantes inflata*, *Cymbella* sp.) were included in the clump II.

Species from the same MBFG tended to cluster in the functional space even if they belonged to different clumps (Figure 3 – 4; Column 2). Yet, species within the same MBFG did not attain the highest abundance in more than one clump (Figure 3 – 4; Column 3). The mean biovolume

394 of species within clumps differed between seasons, but the identity of the most abundant
395 species did not vary (Figure 3 – Column 3). *Pseudanabaena* sp. (spp. 028) and *P. catenata*
396 (spp. 012) had the highest biovolumes of clump I at both dry (Figure 3 – A2) and wet (Figure
397 3 – B3) seasons. Within clump II, *Synedra* sp. (spp. 080) attained the highest biovolume during
398 the dry season (Figure 3 – A3) while *Cymbella* sp. (spp. 051) had the highest biovolume in the
399 wet season (Figure 3 – B3).




400

401 **Figure 3. Distribution of phytoplankton biovolume along the body size axis, the**
402 **ordination of species from the significant size segments (S) in the functional space, and**
403 **the mean biovolume of the five most abundant species of each significant size segment**
404 **during the (A) dry and (B) wet seasons of the Piabanha river, RJ.** (1) Stem plots show size

distribution in the sampling sites of the river Piabanha. Each stem represents a species with its body size (in $\log_2$) plotted on the abscissa and the mean biovolume plotted on the ordinate. The red dotted line indicates the entropy value of each size segment (i.e., unit of $\log_2$ volume) and the asterisk highlights the significant entropy values tested through 1000 randomizations. (2) The species of the corresponding significant size segment are ordinated in the functional space. The size of the circles represents the species contribution to the total biovolume of the size segment, the black line connects species to the centroid (see equation 3), and the number in the centroid of the clump indicates the size segment. (3) Bar plots show the biovolume of the five most abundant species from each significant size segment. Species are coloured according to their morphology-based functional groups (MBFG). The code for species can be found in the supplementary material, Table S1.

Regarding the river stretches, only the clump I had significant entropy values for species from the S9 (Figure 4 – A1), with *Pseudanabaena* sp. (spp. 028) and *P. catenata* (spp. 012) contributing most of the biovolume (Figure 4 – A3). At the medium course, both clumps I and II had significant entropy values with *Pseudanabaena* sp. (spp. 028) attaining the highest biovolume within clump I, and *Synedra* sp. (spp. 080) attaining the highest biovolume within clump II (Figure 4 – C3). At the lower course, only clump II had significant entropy values at the S14 with *Synedra* sp. (spp. 080) as the most representative species (Figure 4 – C3).

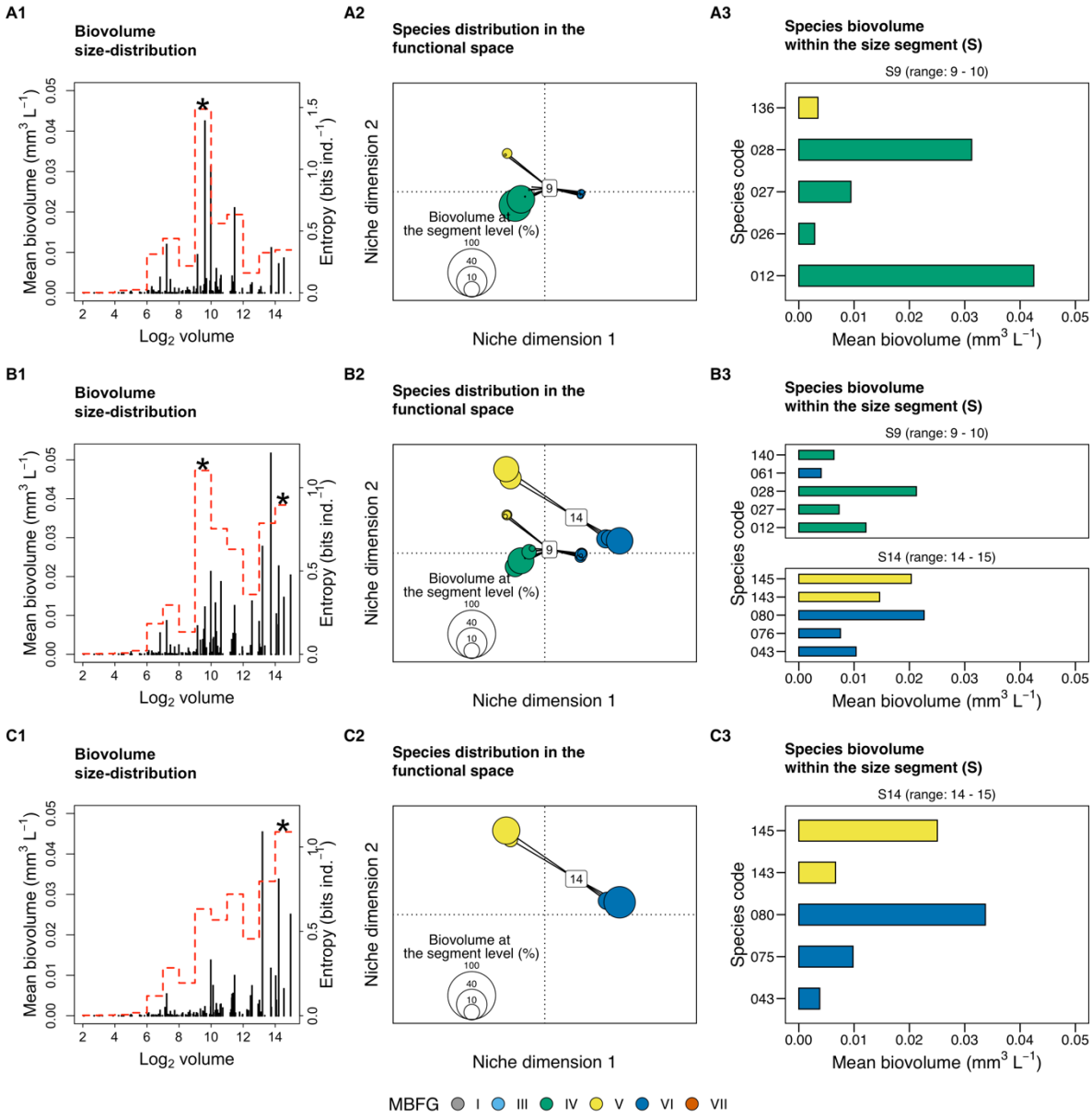





**Figure 4. Distribution of phytoplankton biovolume along the body size axis, the ordination of species from the significant size segments (S) in the functional space, and the mean biovolume of the five most abundant species of each significant size segment at the (A) upper, (B) medium, and (C) lower courses of the Piabanha river, RJ.** (1) Stem plots show size distribution in the sampling sites of the river Piabanha. Each stem represents a species with its body size (in $\log_2$) plotted on the abscissa and the mean biovolume plotted on

the ordinate. The red dotted line indicates the entropy value of each size segment (i.e., unit of $log_2$ volume) and the asterisk highlights the significant entropy values tested through 1000 randomizations. (2) The species of the corresponding significant size segment are ordinated in the functional space. The size of the circles represents the species contribution to the total biovolume of the size segment, the black line connects species to the centroid (see equation 3), and the number in the centroid of the clump indicates the size segment. (3) Bar plots show the biovolume of the five most abundant species from each significant size segment. Species are coloured according to their morphology-based functional groups (MBFG). The code for species can be found in the supplementary material, Table S1.

Mantel tests showed that species' pairwise differences in biovolume correlated with functional dissimilarity irrespectively of the season or river stretch when the whole community was analyzed (Table 3). The correlation was highest during the wet season (Mantel $r = 0.23$, $p < 0.01$) and at the lower course (Mantel $r = 0.26$, $p < 0.01$). For the clump-level pairwise differences, we only found a significant correlation at the upper course (Mantel $r = 0.23$, $p < 0.02$; Table 3). In contrast, functional distinctiveness at clump level presented a significant positive relationship for both the dry season ($\beta = 16.44$, $R^2 = 0.40$, $p < 0.01$) and the wet season ($\beta = 18.32$, $R^2 = 0.40$, $p < 0.01$), and also for the upper ($\beta = 5.68$, $R^2 = 0.40$, $p < 0.01$) and medium ($\beta = 14.46$, $R^2 = 0.34$, $p = 0.02$) courses (Table 4), indicating that species with the most distinct trait combinations within the clumps also attain the highest biovolume. Essentially, such pattern was observed only for the species within clump I, except during the wet season where species from clump II also showed a significant positive relationship ($\beta = -17.83$, $p = 0.02$; Table 4)

**Table 3. Mantel correlation results.** Mantel correlation between the differences in species biovolume and functional dissimilarity for the whole community, and separately for the species

within significant clumps (see Figure 3 – 4) along the seasons (dry and wet) and river stretches (upper, medium, and lower courses). The species number of each stratum (whole community or clumps) are given. The relationships were tested for significance using 999 permutations, whenever possible. Bold values indicate significant correlations ($p < 0.05$).

| *Seasons and river stretches* | *Stratum* | *Number of species* | *Mantel r* | *p-value (permutations)* |
|---|---|---|---|---|
| *Dry season* | | | | |
| | Whole community | 135 | **0.254** | **< 0.01 (999)** |
| | Clump I | 22 | 0.13 | 0.09 (999) |
| | Clump II | 4 | -0.14 | 0.71 (23) |
| *Wet Season* | | | | |
| | Whole community | 123 | **0.29** | **< 0.01 (999)** |
| | Clump I | 20 | 0.06 | 0.24 (999) |
| | Clump II | 11 | -0.16 | 0.69 (999) |
| *Upper course* | | | | |
| | Whole community | 100 | **0.17** | **< 0.01 (999)** |
| | Clump I | 17 | **0.23** | **0.02 (999)** |
| *Medium course* | | | | |
| | Whole community | 135 | **0.21** | **< 0.01 (999)** |
| | Clump I | 23 | 0.06 | 0.22 (999) |
| | Clump II | 5 | -0.16 | 0.65 (119) |

| | | | | |
|---|---|---|---|---|
| *Lower course* | | | | |
| | Whole community | 122 | **0.33** | **<0.01 (999)** |
| | Clump II | 5 | -0.23 | 0.75 (119) |

**Table 4. Linear model results.** Regression parameters of the relationship between species biovolume and functional distinctiveness at the clump level. The coefficients are shown along with the p-values of each independent variable.

| | Dependent variable: l5og$_{10}$ Biovolume | | | | |
|---|---|---|---|---|---|
| Independent variables | Dry season | Wet season | Upper course | Medium course | Lower course |
| $F_{Dist}$ | 16.443<br>p = 0.007 | 18.328<br>p = 0.009 | 5.682<br>p = 0.004 | 14.461<br>p = 0.020 | 9.649<br>p = 0.059 |
| Clump II | 12.286<br>p = 0.305 | 17.734<br>p = 0.017 | | 11.613<br>p = 0.298 | 5.774<br>p = 0.476 |
| $F_{Dist}$ × Clump II | -11.905<br>p = 0.343 | -17.835<br>p = 0.025 | | -11.12<br>p = 0.345 | -4.567<br>p = 0.591 |
| Intercept | -18.390<br>p = 0.002 | -20.420<br>p = 0.003 | -7.259<br>p < 0.001 | -16.636<br>p = 0.006 | -12.522<br>p = 0.013 |
| Observations | 26 | 31 | 17 | 28 | 26 |
| Adjusted $R^2$ | 0.408 | 0.397 | 0.405 | 0.343 | 0.546 |
| F Statistic | 6.752<br>(df = 3; 22) | 7.580<br>(df = 3; 27) | 11.874<br>(df = 1; 15) | 5.703<br>(df = 3; 24) | 11.006<br>(df = 3; 22) |

## Discussion

Present results showed that (i) the clumps in body size are a conspicuous feature of phytoplankton community structure in the Piabanha river across seasons and river stretches;

462 (ii) species within clumps showed a random distribution of biomass concerning their pairwise
463 functional dissimilarity, but not at the whole-community level; and (iii) species biovolume
464 generally increases for species far apart from the centroid of multivariate trait space (i.e.,
465 functional distinctiveness) within clumps. Altogether these results support the Emergent
466 neutrality hypothesis and show that studying species beyond pairwise interactions helps to
467 explain the biomass distribution of functionally similar species, paving the way to analyze
468 intra-clumps trait distributions.

469 Multimodal aggregation of species biovolume along body size axis only points to the
470 integration of niche-based processes and neutrality driving community assembly (Vergnon et
471 al. 2009), supporting $\mathbf{H_1}$. Alternative hypothesis such as pure neutrality (Hubbell 2001) or
472 high dimensional hypothesis (HDH, Clark et al. 2007) are not supported by present results
473 because pure neutrality predicts a uniform distribution of species biovolume and traits along
474 the niche axis (Hubbell 2001), and the HDH does not predict any particular trait distribution
475 (Vergnon et al. 2009, Ingram et al. 2018).

476 One alternative theory that is likely to explain clumpy aggregations is Holling's textural
477 hypothesis (Holling 1992), which suggests that multimodal species size distribution is the
478 result of environmental constraints. Our results do not support the textural hypothesis, as
479 river stretches and seasons were markedly different in hydrology, nutrient concentrations, and
480 other relevant descriptors of riverine landscapes fluxes, but that was not reflected in the stable
481 clumpy size structure of the phytoplankton registered in the present study (Figure 3). The
482 stability found in the clumps agrees with empirical results registered in Segura et al. (2011,
483 2013b) and theoretical findings on the location of clumps (Fort et al. 2009). However, the
484 morphological trait composition of species in the different clumps reflected different
485 environmental templates. The dominant species from the clump I belonged to MBFG IV

(Figure 3 – 4) and presented the highest biovolume under low-flow and high nutrient conditions (Figure 2), which is in line with previous findings for this MBFG (Chen et al. 2015). Within the second clump (II), the most abundant species belonged to MBFG VI (Figure 3 – 4) and had their highest biovolume under high-flow and turbid conditions (Figure 2), in line with the ecology of siliceous organisms (diatoms) able to cope with turbulent environments (Bortolini and Bueno 2017). The empirical evidence is consistent with studies from coastal and estuarine environments (Segura et al. 2011, 2013b) and is in line with recent modelling results suggesting that clumpy patterns arise in environments subjected to resources fluctuation (Sakavara et al. 2018), such as rivers. The trade-off between resources among competing species (Tilman 1982), which is a required ingredient for the emergence of clumps should be further explored.

The analysis of multiple trait dimensions, combining morphology-based functional groups (MBFG) and quantitative distance metrics helped to describe the changes in species traits within clumps. We found that the species from the same clump are distributed across multiple MBFGs but only species from the same MBFG attain the highest biovolume within a clump, reinforcing that body size is a good proxy for niche differences of the species (Blanckenhorn 2000). MBFGs helped to detect differences at a finer degree because they synthesize multiple trait dimensions as suggested previously to understand community organization (D'Andrea et al. 2018). Given that species within the same MBFG share similar ecological strategies (Kruk et al. 2010, Kruk and Segura 2012), under EN premises they should also perform similarly (Scheffer et al. 2018). This is in line with the significant functional dissimilarity observed at the whole community level but not for each clump separately, agreeing with $\mathbf{H_2}$. However, the effects of traits in species' fitness are context-dependent, whereas the use of traits for assigning MBFGs is static, therefore, testing the significance of traits given the observed

environmental conditions might help to unveil community assembling processes at an even finer degree (Kremer et al. 2017).

We also outlined the role of functional similarity in community assembly by studying the effects of functional distinctiveness on species biovolume at the clump level. Within clump I, species biovolume increased with functional distinctiveness, but this pattern was weaker within clump II (Table 4). It may be that such patterns stem from the fact that phytoplankton growth-rate decreases with body size while increases with surface volume ratio, providing that populations of smaller species (such as in clump I) are less sensitive to losses by flushing rates (Kruk et al. 2010). On the other hand, large-sized species have often an elongated shape that provides an advantage under turbulent conditions with low light availability (Reynolds et al. 1994). Differently, aerotopes and mucilage are useful to reach the surface in deep stratified lakes but are not key in small rivers or streams were turbulent fluxes dominate and these traits are not useful to recuperate the position in the water column. Furthermore, in rivers, large-bodied phytoplanktonic species are often randomly introduced from different habitats (e.g. periphyton or epiphyton) (Wang et al. 2014, Descy et al. 2017), which has also been found true for the Piabanha river especially under high flow conditions (Graco-Roza et al. 2020). This mechanism can help to explain the weak relationship between functional distinctiveness and biovolume within clump II, which explain the niche overlap found in large-sized species as the result of immigration and emigration out of the pelagic zone.

Emergent neutrality results from eco-evolutionary processes that lead species selection towards a limited number of functional groups (Scheffer and van Nes 2006). This implies that the clumps observed here are not likely a result of competitive exclusion at the Piabanha river, but a convergent evolution of competing species over time (Macarthur and Levins 1967). Therefore, even when the competition rates are relaxed due to sufficient nutrient

534 supply, some other limiting factors that are not consumed by biotic organisms such as heat
535 energy or turbulence determine species biovolume. Looking at species differences at a high-
536 order level (instead of pairwise differences) helped to detect the effects of trait composition
537 on the biovolume distribution in quasi-neutral clumps. Our results showed that it is possible
538 to predict the biovolume of species within clumps, but only when immigration from different
539 habitats are relaxed and biotic interactions are more likely to occur. Therefore, our findings
540 partially agree with **$H_3$** - there is a positive relationship between species abundance and
541 species functional distinctiveness within clumps, but the environmental conditions seem to
542 play a key role in the outcome.

543 In summary, we provided evidence of both neutral and niche mechanisms driving planktonic
544 community assembly and support the view that emergent neutrality is a likely mechanism to
545 explain species coexistence in an open and environmentally heterogeneous ecosystem. The
546 use of MBFG classification and functional space to describe species within clumps revealed
547 that under the same size range, species with a greater degree of functional similarity
548 unpredictably alternate their dominance. The position and dominance of the clumps were
549 related to the environmental conditions, but the biovolume of species within the clumps was
550 better predicted by functional distinctiveness than by pairwise functional similarity. This
551 addresses the difficulty to avoid the ghost of hidden niches (Barabás et al. 2013) and also
552 provides evidence from multiple angles that point to EN as a plausible mechanism in shaping
553 species coexistence in riverine landscapes.

554 **Acknowledgements**

555 CGR PhD scholarship was funded by Fudação de Apoio a Pesquisa do Estado do Rio de
556 Janeiro (FAPERJ), by Coordenação de Aperfeiçoamento de Pessoal de Nível

Superior (CAPES), and by Ella and Georg Ehrnrooth foundation. MMM was partially supported by CNPq (303572/2017-5).

## References

Anderson, M. J. 2006. Distance-based tests for homogeneity of multivariate dispersions. Biometrics 62:245–253.

APHA, A. 2005. WEF: Standard Methods for the Examination of Water and Wastewater: Centennial Edition. Washington, DC.

Aquino, C. A. N., J. C. Bortolini, C. C. R. Favaretto, N. Y. Sebastien, and N. C. Bueno. 2018. Functional phytoplankton distribution predicts the environmental variability between two subtropical rivers. Revista Brasileira de Botanica 41:835–847.

Barabás, G., R. D'Andrea, R. Rael, G. Meszéna, and A. Ostling. 2013. Emergent neutrality or hidden niches? Oikos.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological) 57:289–300.

Benzécri, J.-P. 1973. L'analyse des données. Dunod Paris.

Blanckenhorn, W. U. 2000. The Evolution of Body Size: What Keeps Organisms Small? The Quarterly Review of Biology 75:385–407.

Bortolini, J. C., and N. C. Bueno. 2017. Temporal dynamics of phytoplankton using the morphology-based functional approach in a subtropical river. Revista Brasileira de Botanica 40:741–748.

Bougeard, S., and S. Dray. 2018. Supervised Multiblock Analysis in {R} with the {ade4} Package. Journal of Statistical Software 86:1–17.

Chen, N., L. Liu, Y. Li, D. Qiao, Y. Li, Y. Zhang, and Y. Lv. 2015. Morphology-based

582 classification of functional groups for potamoplankton. Journal of Limnology 74:559–
583 571.
584 Chessel, D., A.-B. Dufour, and J. Thioulouse. 2004. The {ade4} Package -- {I}: One-Table
585 Methods. R News 4:5–10.
586 Clark, J. S., M. Dietze, S. Chakraborty, P. K. Agarwal, I. Ibanez, S. LaDeau, and M.
587 Wolosin. 2007. Resolving the biodiversity paradox. Ecology Letters 10:647–659.
588 Cornwell, W. K., and D. D. Ackerly. 2009. Community assembly and shifts in plant trait
589 distributions across an environmental gradient in coastal California. Ecological
590 Monographs 79:109–126.
591 Coux, C., R. Rader, I. Bartomeus, and J. M. Tylianakis. 2016. Linking species functional
592 roles to their network roles. Ecology letters 19:762–770.
593 D'Andrea, R., J. Guittar, J. P. O'Dwyer, H. Figueroa, S. J. Wright, R. Condit, and A. Ostling.
594 2020. Counting niches: Abundance-by-trait patterns reveal niche partitioning in a
595 Neotropical forest. Ecology 101.
596 D'Andrea, R., A. Ostling, and J. P. O'Dwyer. 2018. Translucent windows: how uncertainty in
597 competitive interactions impacts detection of community pattern. Ecology Letters
598 21:826–835.
599 D'Andrea, R., M. Riolo, and A. M. Ostling. 2019. Generalizing clusters of similar species as
600 a signature of coexistence under competition. PLOS Computational Biology
601 15:e1006688.
602 Descy, J. P., F. Darchambeau, T. Lambert, M. P. Stoyneva-Gaertner, S. Bouillon, and A. V.
603 Borges. 2017. Phytoplankton dynamics in the Congo River. Freshwater Biology 62:87–
604 101.
605 Díaz, S., J. Kattge, J. H. C. Cornelissen, I. J. Wright, S. Lavorel, S. Dray, B. Reu, M. Kleyer,
606 C. Wirth, I. Colin Prentice, E. Garnier, G. Bönisch, M. Westoby, H. Poorter, P. B.

607 Reich, A. T. Moles, J. Dickie, A. N. Gillison, A. E. Zanne, J. Chave, S. Joseph Wright,
608 S. N. Sheremet Ev, H. Jactel, C. Baraloto, B. Cerabolini, S. Pierce, B. Shipley, D.
609 Kirkup, F. Casanoves, J. S. Joswig, A. Günther, V. Falczuk, N. Rüger, M. D. Mahecha,
610 and L. D. Gorné. 2016. The global spectrum of plant form and function. Nature
611 529:167–171.
612 Díaz, S., A. Purvis, J. H. C. Cornelissen, G. M. Mace, M. J. Donoghue, R. M. Ewers, P.
613 Jordano, and W. D. Pearse. 2013. Functional traits, the phylogeny of function, and
614 ecosystem service vulnerability. Ecology and Evolution 3:2958–2975.
615 Dolédec, S., D. Chessel, C. J. F. ter Braak, and S. Champely. 1996. Matching species traits to
616 environmental variables: a new three-table ordination method. Environmental and
617 Ecological Statistics 3:143–166.
618 Downing, A. S., S. Hajdu, O. Hjerne, S. A. Otto, T. Blenckner, U. Larsson, and M. Winder.
619 2014. Zooming in on size distribution patterns underlying species coexistence in Baltic
620 Sea phytoplankton. Ecology Letters 17:1219–1227.
621 Dray, S., P. Choler, S. Dolédec, P. R. Peres-Neto, W. Thuiller, S. Pavoine, and C. J. F. Ter
622 Braak. 2014. Combining the fourth-corner and the RLQ methods for assessing trait
623 responses to environmental variation. Ecology 95:14–21.
624 Dray, S., and A.--B. Dufour. 2007. The {ade4} Package: Implementing the Duality Diagram
625 for Ecologists. Journal of Statistical Software 22:1–20.
626 Dray, S., A.-B. Dufour, and D. Chessel. 2007. The {ade4} Package -- {II}: Two-Table and
627 {K}-Table Methods. R News 7:47–52.
628 Dray, S., and P. Legendre. 2008. Testing the species traits environment relationships: The
629 fourth-corner problem revisited. Ecology 89:3400–3412.
630 Fort, H., M. Scheffer, and E. H. van Nes. 2009. The paradox of the clumps mathematically
631 explained. Theoretical Ecology 2:171–176.

632 Fort, H., M. Scheffer, and E. Van Nes. 2010. The clumping transition in niche competition: A
633 robust critical phenomenon. Journal of Statistical Mechanics: Theory and Experiment
634 2010.
635 Gause, G. F. 1936. The Struggle for Existence. Soil Science 41:159.
636 Götzenberger, L., F. de Bello, K. A. Bråthen, J. Davison, A. Dubuis, A. Guisan, J. Lepš, R.
637 Lindborg, M. Moora, M. Pärtel, L. Pellissier, J. Pottier, P. Vittoz, K. Zobel, and M.
638 Zobel. 2012. Ecological assembly rules in plant communities-approaches, patterns and
639 prospects. Biological Reviews 87:111–127.
640 Graco-Roza, C., J. B. O. J. B. O. Santos, V. L. M. V. L. M. Huszar, P. Domingos, J.
641 Soininen, and M. M. M. M. Marinho. 2020. Downstream transport processes modulate
642 the effects of environmental heterogeneity on riverine phytoplankton. Science of The
643 Total Environment 703:135519.
644 Gravel, D., C. D. Canham, M. Beaudet, and C. Messier. 2006. Reconciling niche and
645 neutrality: The continuum hypothesis. Ecology Letters 9:399–409.
646 Hardin, G. 1960. The competitive exclusion principle. Science 131:1292–1297.
647 Hill, M. O., and A. J. E. Smith. 1976. PRINCIPAL COMPONENT ANALYSIS OF
648 TAXONOMIC DATA WITH MULTI-STATE DISCRETE CHARACTERS. TAXON
649 25:249–255.
650 Hillebrand, H., C. D. Dürselen, D. Kirschtel, U. Pollingher, and T. Zohary. 1999. Biovolume
651 calculation for pelagic and benthic microalgae. Journal of Phycology 35.
652 Holling, C. S. 1992. Cross-Scale Morphology, Geometry, and Dynamics of Ecosystems.
653 Ecological Monographs 62:447–502.
654 Holt, R. D. 2006. Emergent neutrality. Trends in Ecology and Evolution 21:531–533.
655 Hubbell, S. P. 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Page
656 Monographs in Population Biology. Princeton University Press, Princeton, Oxford.

657 Hubbell, S. P. 2006. Neutral theory and the evolution of ecological equivalence. Ecology
658 87:1387–1398.
659 Hutchinson, G. E. 1957. Concluding Remarks. Cold Spring Harbor Symposia on Quantitative
660 Biology 22:415–427.
661 Ingram, T., R. Costa-Pereira, and M. S. Araújo. 2018. The dimensionality of individual niche
662 variation. Ecology 99:536–549.
663 Junk, W. J., P. B. Bayley, and R. E. Sparks. 1989. The flood pulse concept in river-floodplain
664 systems. Canadian special publication of fisheries and aquatic sciences 106:110–127.
665 Kraft, N. J. B., R. Valencia, and D. D. Ackerly. 2008. Functional traits and niche-based tree
666 community assembly in an Amazonian forest. Science 322:580–582.
667 Kremer, C. T., A. K. Williams, M. Finiguerra, A. A. Fong, A. Kellerman, S. F. Paver, B. B.
668 Tolar, and B. J. Toscano. 2017. Realizing the potential of trait-based aquatic ecology:
669 New tools and collaborative approaches. Limnology and Oceanography 62:253–271.
670 Kruk, C., V. L. M. Huszar, E. T. H. M. Peeters, S. Bonilla, L. Costa, M. LüRling, C. S.
671 Reynolds, and M. Scheffer. 2010. A morphological classification capturing functional
672 variation in phytoplankton. Freshwater Biology 55:614–627.
673 Kruk, C., and A. M. Segura. 2012. The habitat template of phytoplankton morphology-based
674 functional groups. Hydrobiologia 698:191–202.
675 Laliberte, E., P. Legendre, E. Laliberté, and P. Legendre. 2010. A distance-based framework
676 for measuring functional diversity from multiple traits. Ecology 91:299–305.
677 Laliberté, E., P. Legendre, B. Shipley, and M. E. Laliberté. 2014. Package 'FD.' Measuring
678 functional diversity from multiple traits, and other tools for functional ecology.
679 Legendre, P., R. G. Galzin, and M. L. Harmelin-Vivien. 1997. Relating behavior to habitat:
680 Solutions to the fourth-corner problem. Ecology 78:547–562.
681 Litchman, E., K. F. Edwards, C. A. Klausmeier, and M. K. Thomas. 2012. Phytoplankton

682 niches, traits and eco-evolutionary responses to global environmental change. Marine
683 Ecology Progress Series 470:235–248.
684 Litchman, E., and C. A. Klausmeier. 2008. Trait-Based Community Ecology of
685 Phytoplankton. Annual Review of Ecology, Evolution, and Systematics 39:615–639.
686 Litchman, E., P. de Tezanos Pinto, C. A. Klausmeier, M. K. Thomas, and K. Yoshiyama.
687 2010. Linking traits to species diversity and community structure in phytoplankton.
688 Hydrobiologia 653:15–28.
689 Lund, J. W. G., C. Kipling, and E. D. Le Cren. 1958. The inverted microscope method of
690 estimating algal numbers and the statistical basis of estimations by counting.
691 Hydrobiologia 11:143–170.
692 Macarthur, R., and R. Levins. 1967. The Limiting Similarity, Convergence, and Divergence
693 of Coexisting Species. The American Naturalist 101:377–385.
694 Marañón, E. 2008. Inter-specific scaling of phytoplankton production and cell size in the
695 field. Journal of Plankton Research.
696 McGill, B. J., B. J. Enquist, E. Weiher, and M. Westoby. 2006. Rebuilding community
697 ecology from functional traits. Trends in Ecology and Evolution 21:178–185.
698 Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin,
699 R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs, and H. Wagner.
700 2020. vegan: Community Ecology Package.
701 Pavoine, S., M. B. Bonsall, A. Dupaix, U. Jacob, and C. Ricotta. 2017. From phylogenetic to
702 functional originality: Guide through indices and new developments. Ecological
703 Indicators 82:196–205.
704 Peres-Neto, P. R., S. Dray, and C. J. F. ter Braak. 2017. Linking trait variation to the
705 environment: critical issues with community-weighted mean correlation resolved by the
706 fourth-corner approach. Ecography 40:806–816.

R Core Team. 2020. R: A Language and Environment for Statistical Computing. Vienna, Austria.

Reynolds, C., and J. Descy. 1996. The production biomass and structure of phytoplankton in large rivers. Arch. Hydrobiol. 113:161–187.

Reynolds, C. S., J. Alex Elliott, and M. A. Frassl. 2014. Predictive utility of trait-separated phytoplankton groups: A robust approach to modeling population dynamics. Journal of Great Lakes Research 40:143–150.

Reynolds, C. S., J. P. Descy, and J. Padisák. 1994. Are phytoplankton dynamics in rivers so different from those in shallow lakes? Hydrobiologia 289:1–7.

Ricotta, C., F. de Bello, M. Moretti, M. Caccianiga, B. E. L. Cerabolini, and S. Pavoine. 2016. Measuring the functional redundancy of biological communities: a quantitative guide. Methods in Ecology and Evolution 7:1386–1395.

Roy, S., and J. Chattopadhyay. 2007. Towards a resolution of 'the paradox of the plankton': A brief overview of the proposed mechanisms. Ecological Complexity 4:26–33.

Sakavara, A., G. Tsirtsis, D. L. Roelke, R. Mancy, and S. Spatharis. 2018. Lumpy species coexistence arises robustly in fluctuating resource environments. Proceedings of the National Academy of Sciences of the United States of America 115:738–743.

Sauterey, B., B. Ward, J. Rault, C. Bowler, and D. Claessen. 2017. The implications of eco-evolutionary processes for the emergence of marine plankton community biogeography. American Naturalist 190:116–130.

Scheffer, M., and E. H. van Nes. 2006. Self-organized similarity, the evolutionary emergence of groups of similar species. Proceedings of the National Academy of Sciences of the United States of America 103:6230–6235.

Scheffer, M., E. H. Van Nes, and R. Vergnon. 2018. Toward a unifying theory of biodiversity. Proceedings of the National Academy of Sciences of the United States of

America 115:639–641.

Scheffer, M., R. Vergnon, E. H. van Nes, J. G. M. Cuppen, E. T. H. M. Peeters, R. Leijs, and A. N. Nilsson. 2015. The Evolution of Functionally Redundant Species; Evidence from Beetles. PLOS ONE 10:e0137974.

Segura, A. M., D. Calliari, C. Kruk, D. Conde, S. Bonilla, and H. Fort. 2011. Emergent neutrality drives phytoplankton species coexistence. Proceedings of the Royal Society B: Biological Sciences 278:2355–2361.

Segura, A. M., C. Kruk, D. Calliari, and H. Fort. 2013a. Use of a morphology-based functional approach to model phytoplankton community succession in a shallow subtropical lake. Freshwater Biology 58:504–512.

Segura, A. M., C. Kruk, D. Calliari, F. García-Rodriguez, D. Conde, C. E. Widdicombe, and H. Fort. 2013b. Competition drives clumpy species coexistence in estuarine phytoplankton. Scientific Reports 3:1–6.

Thibault, K. M., E. P. White, A. H. Hurlbert, and S. K. M. Ernest. 2011. Multimodality in the individual size distributions of bird communities. Global Ecology and Biogeography.

Thioulouse, J., S. Dray, A.--B. Dufour, A. Siberchicot, T. Jombart, and S. Pavoine. 2018. Multivariate Analysis of Ecological Data with {ade4}. Springer.

Tilman, D. 1982. Resource Competition and Community Structure. Page Monographs in population biology. Princeton University Press.

Uhelingher, V. 1964. Étude statistique dês méthodes de dénobrement planctonique. Archives des Sciences 77:121–123.

Utermöhl, H. 1958. Zur Vervollkommnung der quantitativen Phytoplankton-Methodik. SIL Communications, 1953-1996 9:1–38.

Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell, and C. E. Cushing. 1980. The River Continuum Concept. Canadian Journal of Fisheries and Aquatic Sciences 37:130–

137.

Vergnon, R., N. K. Dulvy, and R. P. Freckleton. 2009. Niches versus neutrality: Uncovering the drivers of diversity in a species-rich community. Ecology Letters 12:1079–1090.

Villéger, S., N. W. H. H. Mason, and D. Mouillot. 2008. New multidimensional functional diversity indices for a multifaceted framework in functional ecology. Ecology 89:2290–2301.

Violle, C., W. Thuiller, N. Mouquet, F. Munoz, N. J. B. Kraft, M. W. Cadotte, S. W. Livingstone, and D. Mouillot. 2017. Functional Rarity: The Ecology of Outliers. Trends in Ecology and Evolution 32:356–367.

Wang, C., X. Li, Z. Lai, Y. Li, A. Dauta, and S. Lek. 2014. Patterning and predicting phytoplankton assemblages in a large subtropical river. Fundamental and Applied Limnology / Archiv für Hydrobiologie 185:263–279.

Wetzel, R. G. 2001. Limnology: Lake and River Ecosystems. Third edition. Academic Press.

Wickham, H., M. Averick, J. Bryan, W. Chang, L. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. Pedersen, E. Miller, S. Bache, K. Müller, J. Ooms, D. Robinson, D. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, and H. Yutani. 2019. Welcome to the Tidyverse. Journal of Open Source Software 4:1686.