

---

## Drivers and ecological consequences of dominance in periurban phytoplankton communities using networks approaches

Escalas Arthur <sup>1,4</sup>, Catherine Arnaud <sup>1</sup>, Maloufi Selma <sup>1</sup>, Cellamre Maria <sup>1,2</sup>, Hamlaoui Sahima <sup>1</sup>, Yepremian Claude <sup>1</sup>, Louvard Clarisse <sup>1</sup>, Trousselier Marc <sup>3</sup>, Bernard Cécile <sup>1,\*</sup>

<sup>1</sup> UMR 7245 MCAM, CNRS-MNHN, Muséum National D'Histoire Naturelle, 12 Rue Buffon, CP 39, 75231, Paris Cedex 05, France

<sup>2</sup> Phyto-Quality, 15 Rue Pétrarque, 75116, Paris, France

<sup>3</sup> UMR 9190 MARBEC, CNRS-Université de Montpellier-IRD-IFREMER, Place Eugène Bataillon, 34095, Montpellier Cedex 5, France

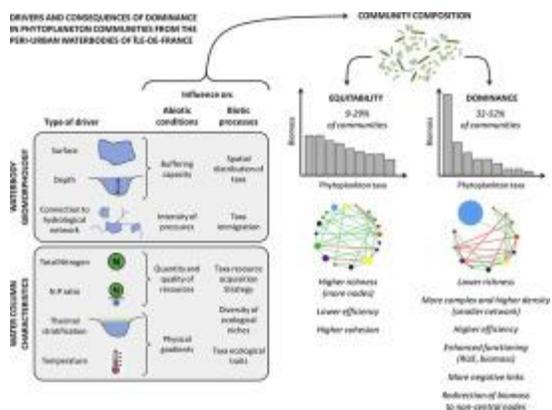
\* Corresponding author : Cécile Bernard, email address : [cecile.bernard@mnhn.fr](mailto:cecile.bernard@mnhn.fr)

---

### Abstract :

Evaluating the causes and consequences of dominance by a limited number of taxa in phytoplankton communities is of huge importance in the current context of increasing anthropogenic pressures on natural ecosystems. This is of particular concern in densely populated urban areas where usages and impacts of human populations on water ecosystems are strongly interconnected. Microbial biodiversity is commonly used as a bioindicator of environmental quality and ecosystem functioning, but there are few studies at the regional scale that integrate the drivers of dominance in phytoplankton communities and their consequences on the structure and functioning of these communities. Here, we studied the causes and consequences of phytoplankton dominance in 50 environmentally contrasted waterbodies, sampled over four summer campaigns in the highly-populated Île-de-France region (IDF). Phytoplankton dominance was observed in 32–52% of the communities and most cases were attributed to Chlorophyta (35.5–40.6% of cases) and Cyanobacteria (30.3–36.5%). The best predictors of dominance were identified using multinomial logistic regression and included waterbody features (surface, depth and connection to the hydrological network) and water column characteristics (total N, TN:TP ratio, water temperature and stratification). The consequences of dominance were dependent on the identity of the dominant organisms and included modifications of biological attributes (richness, cohesion) and functioning (biomass, RUE) of phytoplankton communities. We constructed co-occurrence networks using high resolution phytoplankton biomass and demonstrated that networks under dominance by Chlorophyta and Cyanobacteria exhibited significantly different structure compared with networks without dominance. Furthermore, dominance by Cyanobacteria was associated with more profound network modifications (e.g. cohesion, size, density, efficiency and proportion of negative links), suggesting a stronger disruption of the structure and functioning of phytoplankton communities in the conditions in which this group dominates. Finally, we provide a synthesis on the relationships between environmental drivers, dominance status, community attributes and network structure.

## Graphical abstract



## Highlights

► Periurban lakes were used to study causes and consequences of phytoplankton dominance. ► The main drivers of dominance in phytoplankton communities were identified using MLR. ► Chlorophyta and Cyanobacteria were responsible for most cases of dominance. ► Dominance impacts on community structure and functioning are taxa-specific. ► Cyanobacteria disrupts more phytoplankton co-occurrence networks than other taxa.

**Keywords :** Dominance, Phytoplankton, Co-occurrence network, Community cohesion, Community functioning, Periurban waterbodies

## **KEYWORDS**

Dominance, Phytoplankton, Co-occurrence network, Community cohesion, Community functioning, Periurban waterbodies

## **1. INTRODUCTION**

Continental aquatic ecosystems are considered among the most vulnerable to the combined pressure of anthropogenic activities and climate change (Adrian et al., 2009). In the recent decades, these pressures have led to an increase in the frequency and intensity of phytoplankton blooms in lakes all around the world (Aguilera et al., 2017b; Almanza et al., 2018; Beaver et al., 2018; Dokulil and Teubner, 2000; Moura et al., 2018; Ndlela et al., 2016; Paerl and Huisman, 2008; Paerl and Otten, 2016). Phytoplankton blooms correspond to rapid increases of primary producer's biomass, often associated with the dominance of only a handful of taxa. Even if blooms are naturally occurring phenomena, repeated dominance events have been associated with reduced number of species in eukaryotic and prokaryotic communities in lake ecosystems (Bagatini et al., 2014; Louati et al., 2015; Toporowska and Pawlik-Skowrońska, 2014; J. R. Yang et al., 2017). In addition, the decomposition of phytoplankton blooms has been shown to generate greenhouse gases (CO<sub>2</sub>, CH<sub>4</sub>, N<sub>2</sub>O) and release dissolved nutrients that can retroactively favor climate change and eutrophication (Li et al., 2017). Further, biomass accumulation into a single group of primary producers can modify the resource use efficiency of phytoplankton and the transfer of biomass to zooplankton communities (Filstrup et al., 2014; Tian et al., 2017). Ultimately this may generate ecosystem-wide modifications of food web dynamics leading to trophic collapse (Filstrup et al., 2014; Ullah et al., 2018), impacting on the long term the numerous services that lakes provide to human populations (Vaughn, 2010). This is of particular concern in densely populated urban areas where usages and impacts of human populations on lakes ecosystems are strongly interconnected. Surprisingly, studies at the regional scale on the drivers of dominance in phytoplankton communities are still relatively scarce (Almanza et al., 2018) or were mostly focused on Cyanobacteria (Beaver et al., 2018; Marion et al., 2017; O'Farrell et al., 2019) and sometimes on only one or two species (Bonilla et al., 2012; Marmen et al., 2016). Moreover, there is

no comparison of the consequence of dominance by various organisms on the structure and functioning of phytoplankton communities.

There is increasing evidences that the capacity of microbial systems to support ecological functions and resist environmental forcing are emerging properties arising from the interactions between many taxa (Dai et al., 2018; Faust and Raes, 2012; Goldford et al., 2018; Peura et al., 2015; Shi et al., 2016). Network-based approaches provide an integrated representation of such inter-taxa associations in microbial communities (Barberán et al., 2012; Deng et al., 2012) and were thus recently proposed as potential indicators of environmental quality (Karimi et al., 2017). Similarly, other approaches have been developed to estimate the amount of cohesion experienced by microbial communities, with the underlying idea that this constitutes a fundamental yet neglected facet of the biodiversity of microbial communities (Dai et al., 2018; Danczak et al., 2018; Herren and McMahon, 2018, 2017). Most studies on planktonic networks in lakes focused on bacterial communities or were based on sequencing data (Van Goethem et al., 2017; Woodhouse et al., 2016; C. Yang et al., 2017; Zhao et al., 2016). Additionally, there are few studies related to the impact of dominance on the structure of phytoplankton networks (Carey et al., 2017; Moe et al., 2016), despite the fact that recent studies have shown the under-appreciated role of evenness (the inverse of dominance) in shaping microbial co-occurrence networks (Faust et al., 2015; L. Liu et al., 2019; M. Liu et al., 2019; Xue et al., 2018). In addition, the ability of network approaches to provide meaningful ecological information and to reflect the impact of dominance on communities is still a matter of debate (Röttjers and Faust, 2018).

In this study, we wanted to address two main objectives. In a first time, to identify the factors driving the dominance by a limited number of taxa in phytoplankton communities from periurban waterbodies located in a highly-populated region and representing contrasted environmental conditions (*e.g.* size, depth, watershed characteristics, anthropogenic pressure). In a second time, to determine the consequences of dominance on the structure and functioning of phytoplankton communities, and more particularly on the characteristics of co-occurrence networks. To do so, we analyzed phytoplankton communities across four summer campaigns in 50 waterbodies located in the IDF region and estimated the absolute biomass of taxa composing them. Then, we

combined variables at various spatial scales (water column, waterbody, catchment) to identify the drivers of dominance and analyzed its consequences on the biological and ecological characteristics of these communities. Our approach is based on the hypothesis that dominance should be triggered by a combination of drivers directly related to the intensity of human pressures on these ecosystems with consequences on the structure and functioning of phytoplankton communities depending on the dominant phytoplankton group.

## **2. MATERIAL AND METHODS**

### **2.1. Study area, sampling and *in situ* data acquisition**

A stratified sampling strategy was used to select 50 waterbodies (Figure A1) representative of the contrasted environmental conditions observed in the 248 waterbodies of IDF with a surface area >5 ha (Catherine et al., 2008). According to the chlorophyll a-based OECD definition (OCED, 1982), 6% of the selected waterbodies are oligotrophic, 24% mesotrophic, 26% eutrophic and 44% hypereutrophic (Catherine et al., 2010). Sampling was conducted over two weeks in summers 2006, 2011, 2012 and 2013, where we sampled 50, 48, 49 and 49 waterbodies, respectively, providing a total of 196 phytoplankton samples. To integrate spatial heterogeneity, each waterbody was sampled in three stations and each station was sampled at three depths using a 5L Niskin water sampler. The nine samples per waterbody were pooled for microscopy analyses. Ammonium ( $\text{NH}_4^+$ ), orthophosphate ( $\text{PO}_4^{3-}$ ), total nitrogen (TN) and total phosphorus (TP) analyses were carried out using colorimetric methods previously described (Beck et al., 1992). Nitrate ( $\text{NO}_3^-$ ) was measured using a DX600 ion chromatograph equipped with an AS14 Ion Pack analytical column (Dionex Corp., Westmont, IL). Dissolved oxygen concentration, water temperature, depth and pH were measured using a multiparameter Sea-Bird SBE 19 Seacat Profiler (Sea-Bird Electronics Inc., WA). The values for each of the three sampling stations were averaged to obtain a single value per waterbody.

## 2.2. Phytoplankton data

Phytoplankton characterization was done in triplicate from the pooled nine samples from each waterbody using an inverted microscope (Nikon Eclipse TS100) following the Utermöhl method (CEN (European Committee for Standardization), 2006). A minimum of 500 units (single cells, colonies and filaments) were counted in each sample. To estimate the biovolume, each identified taxon was associated to a geometric shape (Hillebrand, 1999; Sun and Liu, 2003). For each waterbody-campaign combination, we measured the dimensions on 30 individuals from each taxon representing more than 5% of the total abundance, and reported it in a database that was used to infer the biovolume of those taxa when they represented less than 5% of the total abundance. The biovolume of rare taxa was taken from the HELCOM phytoplankton check list (Olenina et al., 2006). Phytoplankton carbon biomass was estimated from biovolumes assuming a density of  $1 \text{ g.cm}^{-3}$  as suggested elsewhere (Filstrup et al., 2014; Holmes et al., 1969).

We estimated three proxies of phytoplankton communities functioning. Total community biomass was estimated as the summed phytoplankton carbon biomass in the community (Filstrup et al., 2014). Resource use efficiency (RUE) was estimated as community biomass divided by the amount of available resources, both for TP and TN (Ptacnik et al., 2008). Although the sampling strategy and the counting procedure were shown to provide accurate estimates of local taxa richness (Maloufi et al., 2016), direct estimation of taxa richness might be biased by low detection limit of rare taxa in communities under a strong gradient of dominance. To safeguard against such a bias, community taxa richness was estimated from individuals counts using abundance-based Chao extrapolation method (function *estimateR* from the R package *vegan v 3.4.4*; (Chiu et al., 2014; Oksanen et al., 2016).

## 2.3. Determination of taxa connectedness and community cohesion

We used the recently developed *cohesion* framework (Herren and McMahon, 2018, 2017) to describe the potential of individual taxon for inter-taxa associations (*i.e.* connectedness) but also the overall degree of connectivity within communities (*i.e.* cohesion). First, we estimated the pairwise correlation matrices between

taxa across all communities, separately for positive and negative correlations. Then, we used a null model to account for bias in these correlations due to the skewed distribution of taxa biomass (Herren and McMahon, 2017). We followed the authors recommendations to select appropriated type of null model and after considering taxa distribution across and within communities we choose to maintain taxa average biomass while randomizing the data. Then, “expected” correlations generated by the null model were subtracted from the original correlation matrices to obtain corrected correlations matrices. For each taxon, the average positive and negative corrected correlation corresponded to their positive and negative connectedness values, respectively. Connectedness was estimated for the 227 taxa (out of the 506 detected) that appeared in at least four out of the 196 sampled communities. For each community, cohesion was calculated as the summed connectedness of taxa composing the community, weighted by their respective biomass. Because positive and negative connectedness are calculated separately there are two metrics of cohesion, a positive and a negative one.

Differences in connectedness between eight phytoplankton phyla (*i.e.* Bacillariophyta, Chlorophyta, Cryptophyta, Cyanobacteria, Dinophyta, Euglenophyta, Chrysophyta and Xantophyta) were tested using Kruskal-Wallis rank sum test and Dunn test (R package *dunn.test*; Dinno, 2017).

#### **2.4. Definition and estimation of dominance**

The first step to identify the drivers and consequences of dominance in phytoplankton communities was to identify groups of communities under various dominance scenarii. In other terms, we needed to identify (i) whether communities were under dominance by a reduced set of taxa or not, and (ii) what was the identity of the dominant taxon. In a first time, for each community we calculated a dominance index that corresponded to 1 minus the Pielou evenness index estimated using biomass matrices ( $\text{evenness} = H' / \log(S)$ ; with  $H'$  = Shannon-Weiner diversity and  $S$  = taxa richness). This dominance index ranged from 0 when the biomass distribution across taxa was perfectly even to 1 when community biomass corresponded to a single taxon. Then, this index was used to separate communities not under dominance (*i.e.* with index < threshold) from those under dominance (index > threshold). In a second time, communities considered under dominance in step 1 were

grouped according to the identity of the dominant phyla, that is the phyla whose relative biomass was above a selected threshold. In order to avoid biases related to the selection of a unique threshold for each of these steps, we used a range of values for the dominance (0.45, 0.5, 0.55) and relative biomass (0.45, 0.5, 0.55) thresholds, which generated nine thresholds combinations that were used to separate communities into various dominance groups (*i.e.* not-dominated and dominated by different phyla, Table A1). Data were analyzed in a similar way for each of the nine thresholds combinations and the results were summarized across all combinations to identify the global trends in our data.

## **2.5. Identification of the drivers of dominance in phytoplankton communities**

To identify the environmental drivers of dominance in the IDF region, we used classification models and variables at the water column, waterbody and catchment scales to predict the dominance group in which communities belong to. The set of predictor variables was defined on the basis of their assumed contribution to the environmental conditions experienced within the waterbodies and modified from (Table A2, Catherine et al., 2010). At the water column scale, nutrient concentration (total N and P) were used to reflect the quantity of available resources, while water temperature and thermal stratification were used to reflect physical conditions of the water column. At the waterbody scale, several predictors related to their characteristics were included. Mean depth and waterbody surface were used as variables reflecting the capacity of the lake to dilute nutrient loadings. Waterbodies position within the landscape was reflected by their altitude, their connection to the hydrological network (*i.e.* the network of channels connecting small rivers and waterbodies throughout the IDF region) or their location in a regularly flooded area. At the catchment scale, the ratio between catchment and waterbody size was used to reflect the loading ‘potential’ of a catchment system relative to the waterbody buffering capacity (Almanza et al., 2018). The density of drainage connections within catchments was considered as it affects the catchment’s ability to transport nutrients. Land use variables such as the proportion of catchment surface classified as forest, agricultural and urban, suburban or industrial referred as impervious cover were included as they constitute variables known to affect both the quantity and nature of loading. Variables at

the catchment scale were estimated using the Carthage 3.0 hydrological database (IGN-MATE, 2005) and the MOS databases (“www.iau-idf.fr”).

Two classification approaches were compared, random forest (RF using the *randomForest* function from the R package *randomForest* v 3.4.4, with n = 1000 trees, Liaw and Wiener, 2002), which has been shown to accurately predict the eutrophication level of waterbodies in the region (Catherine et al., 2010) and multinomial logistic regression (MLR) using the *multinom* function from the R package *nnet* v 3.4.0 (Venables and Ripley, 2002). Three models were tested (Table A3). Model 1 was the more complex and aimed at predicting the classification of communities into six groups that corresponded to the absence of dominance or the dominance by Bacillariophyta, Chlorophyta, Cyanobacteria, Dinophyta or other phyla. Model 2 focused on predicting the dichotomy between dominated versus not-dominated communities. Model 3, focused on communities under dominance and aimed at predicting the identity of the dominating taxon. For each model the RF and MLR methods were compared based on the proportion of communities accurately classified in their respective group (*i.e.* confusion matrices). The predictor variables that contributed to the classification success were subsequently identified using likelihood ratio Chi-square test (O’Farrell et al., 2019; Venables and Ripley, 2002).

## **2.6. Determination of the consequences of dominance on community-level properties**

To determine the consequence of dominance at the community level we compared the communities under various dominance scenarii in terms of community biomass, estimated taxa richness, RUE (separately for N- and P-based RUE) and cohesion (positive and negative). As widely different numbers of communities composed the different dominance groups, global and pairwise differences between groups were tested using non-parametric Kruskal-Wallis rank sum test and Dunn test, respectively (R package *dunn.test*; Dinno, 2017). Community biomass and RUE were log transformed before analyses.

## 2.7. Construction and characterization of co-occurrence networks

Co-occurrence networks were used to summarize the impact of dominance by a limited number of species on the structure of phytoplankton communities (Deng et al., 2012). Albeit promising, network approaches are not devoid of limitations (Röttjers and Faust, 2018). First, a large number of replicates are required to create a single network as its construction relies on co-variations in the biomass of taxa across communities. Hence, only the three groups with the largest number of communities were analyzed using this approach: no-dominance ( $17 < n < 56$  depending on the selected thresholds), dominance by Chlorophyta ( $22 < n < 39$ ) and dominance by Cyanobacteria ( $22 < n < 31$ ). Second, microbial data are compositional by nature (Faust and Raes, 2012) and are thus prone to spurious correlations (Jackson, 1997; Lovell et al., 2011). To limit these compositional biases, we used absolute taxa biomass to estimate taxa association. Third, microbial matrices contain a large proportion of zeros, a phenomenon referred as the data sparsity problem (Paulson et al., 2013). Hence, organisms absent from too many samples are often excluded from the analysis in search for a trade-off between the amount of available data and their reliability. Here, networks from each group were constructed using taxa detected in more than 12.5% of communities, which corresponded to a compromise between the need to keep taxa with a maximum number of observations to accurately estimate their co-occurrence and the need to keep enough taxa in the analysis to construct networks that are representative of the communities observed in the field (we also tried with 7.5, 10, 12.5, 15, 20 and 25%, Röttjers and Faust, 2018). Another issue lies in the choice of the metric used to estimate taxa association, which should be made to reduce as much as possible the number of false-positives (Karimi et al., 2017). Here we used the Pearson correlation, which is the standard in microbial networks studies and has been successfully used in association with Random Matrix Theory in soils (Deng et al., 2012; Wang et al., 2015), rhizosphere (Shi et al., 2016) and lakes (J. R. Yang et al., 2017; Zhao et al., 2016). This index assumes linear relationships between taxa biomass and is sensitive to data sparsity and compositional issues (Kurtz et al., 2015). However, we are confident that we considered all the ways to reduce the impact of methodological biases on our correlation-based analyses. Then, Random Matrix Theory (RMT) was used to objectively identify a cutoff determining which associations were kept in the final

network (Luo et al., 2006). Networks were generated using the Molecular Ecological Network Analyses (MENA) pipeline (Deng et al., 2012) and represented using Cytoscape 3.6.0 (Shannon et al., 2003).

Networks are composed of nodes, which correspond to individual taxon, connected by links (or edges), that represent significant associations between nodes. Networks structure was characterized using indexes derived from the graph theory (Pavlopoulos et al., 2011) and recently suggested as potential bioindicators of the state of a system (Karimi et al., 2017). At the node level, we estimated two centrality indexes describing the importance of nodes in the network. Node *degree* was defined as the number of links to this node while node *betweenness* was defined as the number of geodesics (*i.e.* shortest path between two nodes) passing through the node. Betweenness reflects how central and influent a node may be in the network by being on the paths relating other nodes in the network. We also tested whether the biomass of a taxon was related to its centrality in the network using correlation tests (*cor.test* function in R). Nodes *clustering coefficient* was used to describe how well a node was connected with its neighbors. At the network level, we estimated the proportion of positive and negative links, along with global network properties (Deng et al., 2012). *Average geodesic distance* corresponds to the average length of the shortest path between every pairs of nodes in the network. *Geodesic efficiency* reflects the size of the network while network *density* reflects its complexity and corresponded to the ratio between realized and potential links.

### 3. RESULTS

#### 3.1. Characteristics of phytoplankton communities in the waterbodies of the IDF region

A total of 506 phytoplankton taxa corresponding to 181 genera were identified across the 196 samples (four campaigns with 48 to 50 waterbodies), 72.5 % were classified at the species level and the remaining 27.5 % at the genus level. The generic term taxa will be used in the text to refer to the diversity unit used in this study. Community richness ranged from four to 213 taxa per community, with an average of  $42.6 \pm 27.8$ . The proportion of communities identified as dominated by a reduced number of taxa during summer in the IDF region ranged from 31.6 to 51.5%, depending on thresholds combinations used to define groups (Table A1). The

proportion of communities not under dominance was more variable and ranged from nine to 29%. Two phylum, the Chlorophyta and the Cyanobacteria, represented most of the cases of dominance, with 35.5-40.6% and 30.3-36.5% of dominance cases, respectively. The most dominant Chlorophytes taxa included *Coelastrum polychordum*, *Botryococcus* sp., *Pediastrum boryanum*, *Pediastrum simplex*, *Pediastrum duplex* and *Pandorina* sp.. The most dominant Cyanobacteria taxa included *Aphanizomenon flos-aquae*, *Dolichospermum* sp., *Aphanizomenon klebahnii*, *Planktothrix agardhii*, *Dolichospermum flos-aquae* and *Microcystis aeruginosa*. The other phyla dominated in a smaller number of cases: Dinophyta (9.4-11.0%), Bacillariophyta (7.3–9.0%) and the remaining other groups (8.8-12.5%).

### **3.2. Association potential of phytoplankton organisms**

The inter-taxa association potential (*i.e.* connectedness) was estimated for the 227 most occurring taxa and significant differences were observed across phyla. The effect of phylum was stronger for positive than for negative connectedness (Kruskal-Wallis test, p.value < 0.001 and 0.039, respectively, supplementary Figure A2 and Table A4). Pairwise tests revealed that Cryptophyta and Dinophyta exhibited significantly stronger negative connectedness than other phyla (Dunn test, p.value < 0.05, Table A3). However, these two groups exhibited highly variable negative connectedness and were represented by a limited number of taxa (n = 4 and 5, respectively). In terms of positive connectedness, Cyanobacteria differed significantly from five out of the seven other phyla (Dunn test, p.value < 0.05, Table A5), with the exception of Xantophyta and Bacillariophyta. This latter phyla was significantly different than Chlorophyta. These results position Cyanobacteria apart from the other phytoplankton phyla, with lower positive inter-taxa association.

### **3.3. Environmental drivers of dominance in phytoplankton communities**

Classification approaches were used to identify the environmental factors that determined the dominance status in phytoplankton communities (cf. Table A3 for a description of data). For each of the three tested models, multinomial logistic regression (MLR) was more accurate than random forest (RF) to determine the dominance

status of phytoplankton communities using the selected predictor variables (Table 1). Indeed, MLR accurately classified communities in 63, 69 and 80% of the cases while RF accurately classified communities in 43, 67 and 57% of the cases, for models 1, 2 and 3, respectively. This was mostly due to the low accuracy of RF in categories with low sample sizes (*i.e.* Bacillariophyta, Dinophyta and Others categories). Interestingly, both models that included a no-dominance category (1 and 2) exhibited a low classification accuracy for this category.

The predictors that contributed in the accuracy of MLR models included variables at various scales, from the water column physical and chemical features to the catchment characteristics (Table 2). Overall, the variables with the highest number of significant contributions were the total N concentration followed by the water temperature, the TN:TP ratio and the connection of the waterbody to the hydrological network, while water thermal stratification, depth and waterbody surface contributed less frequently. Regarding the quantitative contribution to classification accuracy (*i.e.*  $\chi^2$  values), the connection of the waterbody to the hydrological network, the water temperature and TN:TP ratio were the prominent variables. Variables representing the state of the water column and the characteristics of the waterbody were mostly influencing classification accuracy in models that included the identity of the dominant taxon (models 1 and 3). In model 2, only TN:TP ratio was significant across the nine tested dominance thresholds.

### **3.4. Structure and functioning of phytoplankton communities under various dominance scenarii**

We compared several community-level properties across dominance scenarii to determine whether dominance and the identity of the dominant group were associated with differences in the structure and functioning of phytoplankton communities (Figure 1). All the tested variables significantly differed across groups (Kruskal-Wallis test,  $p$ .value < 0.05, Table 3), but a more detailed picture was provided by pairwise comparison (Table A6). In terms of total community biomass, we observed a clear dichotomy, with the communities dominated by Bacillariophyta, Chlorophyta or Cyanobacteria reaching a significantly higher biomass than communities from the others groups. Regarding community richness, the effect of dominance appeared more contrasted and dependent on the identity of the dominant phylum (Figure 1). On one hand, Chlorophyta-dominated communities exhibited a higher richness compared to all the groups, with the exception

of Bacillariophyta. On the other hand, communities dominated by Cyanobacteria and Others organisms exhibited lower richness than the no-dominance and the Bacillariophyta groups. In terms of RUE, two groups clearly stood out, Chlorophyta and Cyanobacteria, which were the only ones to differ significantly from the Others and no-dominance groups. This was particularly striking for the use of nitrogen resources (*i.e.* RUE<sub>N</sub>). In terms of community cohesion, the Cyanobacteria-dominated communities were apart from the others, exhibiting significantly lower negative and positive cohesion (Figure 1). To conclude, the consequence of dominance on community structure and functioning appeared taxa-specific and dependent on the considered community-level property. Overall, dominance by Cyanobacteria showed the strongest effect on all the studied variables and tend to separate these communities from others, notably in terms of cohesion. Dominance by Chlorophyta also has strong effects on community functioning while Bacillariophyta dominance seemed to have a lower, albeit significant effect.

### **3.5. Characteristics of phytoplankton co-occurrence networks in various dominance scenarii**

We constructed co-occurrence networks to summarize the structure of phytoplankton communities in various scenarii of dominance by a limited number of taxa (Figure 2 and Figure A3). This was done for the nine tested combinations of dominance index and relative biomass thresholds and for the three groups for which we had the most replicates: dominance by Chlorophyta or Cyanobacteria and absence of dominance. The pairwise similarity cutoffs estimated using RMT were similar among the three groups and across the nine thresholds combinations, with average values of  $0.33\pm 0.02$ ,  $0.32\pm 0.01$  and  $0.32\pm 0.03$  for Chlorophyta, Cyanobacteria and no-dominance, respectively. Comparison of observed networks topological properties with those of randomized networks indicated that networks structures were non-random and unlikely due to chance. The observed networks exhibited a ratio of clustering coefficient to geodesic distance higher than 1, which is thought to be characteristic of small-world networks (Humphries and Gurney, 2008). This ratio was the highest in the no-dominance networks ( $2.1\pm 0.9$ ), then decreased under dominance with intermediate and low values for the Chlorophyta- ( $1.5\pm 0.3$ ) and Cyanobacteria-dominated ( $1.2\pm 0.1$ ) networks.

The observed numbers of nodes and links decreased from the Chlorophyta-dominated networks, the no-dominance and the Cyanobacteria-dominated networks (Table 4). Network density was the highest in the Cyanobacteria-dominated networks while the Chlorophyta and no-dominance networks exhibited lower and more similar density. Two indexes reflecting the size of the networks, geodesic efficiency and average geodesic distance, respectively increased and decreased from no-dominance, Chlorophyta- and Cyanobacteria-dominated networks. The highest proportion of negative links was observed in the Cyanobacteria-dominated networks with, on average,  $48.0 \pm 7.5\%$  of negative links. No-dominance ( $34.7 \pm 24.8\%$ ) and Chlorophyta-dominated ( $34.5 \pm 14.7\%$ ) networks exhibited lower proportions of negative links. The wide range of values observed in the no-dominance networks (5.3 to 71.4% of negative links) was associated with a more variable number of communities used for constructing networks in this group. It is worth noting that the most extreme values (*i.e.* 63.3 and 71.4% of negative links) were observed in networks constructed with the smallest number of communities, that is with  $n = 17, 20$  and  $20$  communities.

In terms of average node degree, the three networks differed significantly (in eight out of nine thresholds combinations, Kruskal,  $p.value < 0.05$ , Table S8), with the networks under dominance exhibiting higher average node degree than the no-dominance network (Dunn,  $p.value < 0.05$ ). The three networks differed significantly in terms of node betweenness (for every threshold combination, Kruskal,  $p.value < 0.001$ ) and Cyanobacteria-dominated networks exhibited systematically lower values compared with other groups. In terms of clustering coefficients, the three networks differed significantly (in eight out of nine thresholds combinations, Kruskal,  $p.value < 0.05$ ), with Cyanobacteria-dominated networks exhibiting higher values than other groups.

Then, we tested whether the biomass of a taxon was related to its centrality in the network using Pearson correlation (*i.e.* degree and betweenness, Table 5 and Figure A4) and all the estimated correlation coefficients were negative. In the absence of dominance centrality and biomass appeared unrelated. Betweenness was never significantly related to biomass across the tested threshold combinations while degree and biomass were significantly related three times. As observed for the proportion of negative links, these significant relationships were only observed in networks constructed with the lowest number of observations. In Chlorophyta-dominated

networks, betweenness was never significantly related to biomass, while node degree was related to biomass three times and only with the most stringent dominance threshold, *i.e.* 0.55. In Cyanobacteria-dominated networks, betweenness was significantly related to biomass four times, while node degree was significantly related to biomass eight out of nine times.

## **4. DISCUSSION**

### **4.1. Drivers of dominance in phytoplankton communities from the periurban waterbodies of the IDF region**

The best predictors of the dominance status in phytoplankton communities of the periurban waterbodies of the IDF region during summer were related to human pressures on these ecosystems. Dominance by a limited number of taxa was frequently observed (31.6-51.5%, depending on the thresholds used) and was mostly attributed to two phyla, Chlorophyta (35.5-40.6% of the cases) and Cyanobacteria (30.3-36.5%). Dominance was determined by a combination of (i) water column characteristics, that are related to the intensity of local anthropogenic pressures (TN and TN:TP) and meteorological conditions (water temperature and stratification), and (ii) the capacity of waterbodies to buffer these pressures (surface, depth) or the potential for immigration of taxa from other lakes including harmful algae (connection to the hydrological network). Our results support previous investigations showing that high N concentrations (Almanza et al., 2018; Beaver et al., 2018; Bonilla et al., 2012; Marion et al., 2017; Paerl and Otten, 2016; Persaud et al., 2015) and warmer waters (Dokulil and Teubner, 2000; Paerl and Huisman, 2008; Wagner and Adrian, 2009) favor the dominance of Cyanobacteria in phytoplankton communities at large scales. This is thought to be related to the particular ecological traits of these organisms (Mantzouki et al., 2016), such as higher increase in growth rates per unit of temperature compared with other groups (Carey et al., 2012), their ability to fix atmospheric N<sub>2</sub>, or to regulate their buoyancy in stratified water columns associated with warmer waters (Rinke et al., 2010). Regarding dominant Chlorophyta taxa, the most prominent genera (Coelastrum, Pediastrum and Pandorina) were known for their wide spatial distribution and their high abundance in meso to eutrophic freshwater ecosystems. Additionally, these genera

exhibited similar features. They are known for their colonial lifestyle and for being hard to graze by zooplankton due to a combination of thick cellulose cell walls and the production of a gelatinous envelope (Porter, 1977). Further, these genera are known for self regulating their vertical position and are typically blooming in early summer, often preceding blooms of Cyanobacteria (Salmaso, 2000; Salmaso et al., 2015).

The classification accuracy of MLR models was remarkable (63 to 80% depending on the models). This approach has been compared with RF for classifying soil types and exhibited either higher (Bernhardt-Barry et al., 2018) or lower performances (Camera et al., 2017). In a more similar context, (O'Farrell et al., 2019) obtained accuracy values comparable to ours while trying to classify Cyanobacteria into different ecological strategies. Additionally, MLR models were more accurate in predicting the presence than the absence of dominance, which contradicts results from the above-mentioned study (O'Farrell et al., 2019). This might be due to the selected variables that, in our case, were chosen for their expected influence on environmental characteristics within the waterbodies. Another explanation could be that a non-dominated phytoplankton community can be observed in a wide array of environmental conditions and be composed of many combinations of organisms. On the contrary, dominance implies that particular organisms have been selected by a narrower set of environmental conditions. In that sense, it appears less complex to identify favorable than unfavorable conditions to dominance.

#### **4.2. Considerations regarding the use of networks approaches**

Networks-based approaches provide an integrated and explicit representation of microbial communities, opening the way for a better understanding of their structure and functioning (Faust and Raes, 2012; Karimi et al., 2017). Here we assessed and compared the co-occurrence patterns of phytoplankton organisms under the dominance by two different phyla, Chlorophyta and Cyanobacteria, and in the absence of dominance. The constructed networks represent co-variations in the biomass of phytoplankton organisms across space and time, which could result from similar response to environmental drivers or from inter-dependencies taking the form of biological interactions (Röttjers and Faust, 2018). While positive associations might result from similar niche or

mutualism, negative associations might reflect non-overlapping niche or antagonism. Analyses of macro-ecological networks have shown that co-occurrences may better reflect niche preferences than biotic interactions (Freilich et al., 2018). This is an important point to keep in mind, as trying to infer associations using network analysis in highly heterogeneous environments might produce erroneous results (Röttgers and Faust, 2018). Here, we tried to avoid such bias by not considering networks as representations of biological interactions but rather as a synthetic representation of the structure of phytoplankton communities in various ecological contexts (in our case phytoplankton dominance scenarios).

#### **4.3. Relationships between phytoplankton communities structure, functioning and co-occurrence networks**

There is still no consensus about the ability of network-based analyses to reflect biologically and/or ecologically meaningful properties of environmental communities (Röttgers and Faust, 2018). In the following sections, we discuss whether the observed modifications in the structure and functioning of phytoplankton communities were reflected in co-occurrence networks.

The co-occurrence networks observed in this study exhibited “small-world” characteristics (low geodesic, high average degree and high efficiency) and this type of network tends to be considered as robust to perturbations and the random loss of nodes (Peura et al., 2015; Watts and Strogatz, 1998). Interestingly, the test proposed by Humphries and Gurney (2008) suggested that under dominance networks departed from the typical “small-world” network.

Community richness was highly influenced by the level of dominance and the identity of dominant taxa. Dominance by Bacillariophyta and Chlorophyta was associated with a higher richness, while dominance by Cyanobacteria and other groups was associated with lower richness compared with the no-dominance communities. This was well reflected in the number of nodes composing the networks. (Aguilera et al., 2017a; Dokulil and Teubner, 2000; Litchman, 2003). Similarly, observations of reduced number of taxa in microbial communities dominated by Cyanobacteria are legions (Bagatini et al., 2014; Filstrup et al., 2014; Holland and

Kinnear, 2013; Louati et al., 2015; Toporowska and Pawlik-Skowrońska, 2014; C. Yang et al., 2017; J. R. Yang et al., 2017; Zhao et al., 2016).

Although composed of fewer nodes and links, the Cyanobacteria-dominated networks can be considered as more complex than the others as exhibiting a higher link density and higher average number of links (degree) per node. Similar reductions in size and increased complexity have been reported in networks associated with *Microcystis aeruginosa* colonies (C. Yang et al., 2017) and after inoculation of freshwater mesocosms with the large colonial Cyanobacteria *Gloetrichia echinulata* (Carey et al., 2017). This later study reported that increasing densities *G. echinulata* increased the richness, biomass, evenness and composition stability of phytoplankton communities. However, this study was done in oligotrophic waters and the effects of *G. echinulata* were mostly attributed to its ability to release N and P in the environment (Carey et al., 2014). Consequently, these results might be specific to the conditions tested in this study and not generalizable to a wider context. On the contrary, by studying 50 waterbodies that represent heterogeneous and mostly non-oligotrophic environmental conditions, we observed that the increasing presence of Cyanobacteria was associated with higher biomass along with lower richness and evenness. In addition, the lowest evenness, richness and number of nodes in the networks were associated with higher proportions of negative links and corresponded to conditions of dominance by Cyanobacteria. Similar results were reported in a large-scale analysis on the role of dominance in shaping microbial co-occurrence networks (Faust et al., 2015). The proportions of negative links (~ 50%) observed under cyanobacterial dominance appear relatively high regarding the values reported for other networks in the literature: 12% (C. Yang et al., 2017), < 20% (Shi et al., 2016), < 5% (Van Goethem et al., 2017), 23% (Hu et al., 2017) and 37 to 44% (Jones et al., 2018). Similarly, many negative associations have been reported in bacterial co-occurrence networks during a *Dolichospermum-Microcystis* bloom (Guedes et al., 2018; Woodhouse et al., 2016).

The modifications of structure observed in conditions of dominance were translated in modifications of community functioning, with higher biomass and efficiency of resource use (N and P) under dominance, as observed in other studies (Filstrup et al., 2014; Gamfeldt and Hillebrand, 2011; Hodapp et al., 2015; Lehtinen et

al., 2017). Similarly, we observed higher efficiency in networks under dominance, as suggested by shorter distance between nodes and higher geodesic efficiency. In addition, community functioning depended on the identity of the dominant taxa, with Bacillariophyta, Chlorophyta and Cyanobacteria reaching the highest biomass when dominant. Interestingly, Bacillariophyta reached high biomass without exhibiting high efficiency but rather when total N concentrations were very high. On the contrary, Chlorophyta and Cyanobacteria dominance were associated with elevated efficiency ( $RUE_N$  and  $RUE_P$ ) and were observed in low TN:TP ratio conditions but also across a wide range of TN and TP concentrations. This suggests that these two groups are able to bloom in a wider array of environmental conditions which can explain why they were responsible for most of the cases of dominance.

Cohesion metrics were developed to reflect the degree of connectivity in a microbial community (Dai et al., 2018; Herren and McMahon, 2018) or its complexity/interconnectedness (Danczak et al., 2018). In addition, stronger cohesion (negative or positive) is thought to reflect the stabilizing effect of biotic interactions on community composition (Dai et al., 2018; Herren and McMahon, 2018, 2017). We observed that community cohesion was the lowest and differed the most from the no-dominance communities under cyanobacterial dominance. This agree with studies reporting few and weak positive correlations of cyanobacterial nodes with other phyla (C. Yang et al., 2017; Zhao et al., 2016) and suggests that these communities should be less stable. However, other studies at the regional scale reported higher temporal and spatial stability in community composition under the combined pressure of climate change, eutrophication and cyanobacterial dominance (Filstrup et al., 2014; Monchamp et al., 2018). Taken together, these results show that phytoplankton communities under Cyanobacteria dominance exhibit a lower cohesion and less interactions between taxa but at nonetheless compositionally stable.

A significant negative relationship was observed between the biomass of a taxon and its centrality in the networks (degree or betweenness) in conditions of Cyanobacteria dominance, suggesting that most of the community biomass was redirected toward the outer part of the network. These relationships were driven by cyanobacterial nodes, which exhibited a low number of connections (low degree) or were located on the edges of

the networks (low betweenness) while representing medium to high biomass. Cyanobacteria could redirect community biomass without participating in inter-taxa interactions. Experimental studies testing the relationship between biodiversity and functioning in phytoplankton communities have shown that Cyanobacteria exhibited a “high selection - low complementarity” effect on community biomass (Behl et al., 2011; Gamfeldt and Hillebrand, 2011). In other terms, they influence biomass through their high productivity and competitiveness rather than their ability to positively interact with other species.

To summarize, the characteristics of co-occurrence networks seemed to reflect the modifications of phytoplankton communities structure (reduced richness, lower cohesion) and functioning (higher biomass and RUE) observed in conditions of dominance. Reduction in community richness was reflected in the number of nodes composing the network and might be explained by the fact that environmental conditions allowing cyanobacterial dominance are suitable to fewer taxa than conditions allowing the dominance by other groups (*i.e.* environment filtering). Dominance was associated with more complex but also smaller and more efficient networks, which exhibited enhanced community functioning. In addition, more efficient community functioning seemed to arise through selection effect (Gamfeldt and Hillebrand, 2011), as suggested by the negative relationship between centrality and biomass in conditions of Cyanobacteria dominance. Further, the potential of organisms for inter-taxa association (*i.e.* connectedness) was reflected in their degree in the network and dominance by organisms with low potential corresponded to lower community cohesion. The isolation of a taxon, represented by its low centrality and connectedness, is not necessarily deleterious and may even be beneficial if it reflects its ability to fill a specific niche space for which no direct competitor exists (Peura et al., 2015). Several studies have shown that cyanobacterial dominance generates environmental conditions that are favorable to its maintenance and can consequently thus be observed for prolonged periods (Corcoran and Boeing, 2012; Filstrup et al., 2014; Scheffer et al., 1997). This might be related to specific cyanobacterial traits involved in their high biomass development (*e.g.* nitrogen fixation, specific CO<sub>2</sub> fixation, buoyancy). Another hypothesis might be related to the effect of competition/exclusion mechanisms on community composition (Aguilera et al., 2017a; Dokulil and Teubner, 2000; Litchman, 2003). For instance, Cyanobacteria are more efficient than other groups for light-harvesting in low-light conditions generated by high biomass (Litchman,

2003; Schwaderer et al., 2011) and are able to release a wide variety of allelochemical compounds (Dias et al., 2017; Holland and Kinnear, 2013; Leflaive and Ten-Hage, 2007). To conclude, it appears that the ecological particularities of this group allow their dominance in waterbodies under anthropogenic pressure, and this resulted in enhanced community functioning through selection effect. Once established, dominance creates environmental conditions favorable to its maintenance and we can hypothesize that the resulting network (more complex, higher density and proportion of negative links) reflect the fact the remaining taxa are involved in more interactions (*e.g.* commensalism, mutualism, competition) in order to occupy the few available niches and scavenge the remaining resources.

## **5. CONCLUSIONS**

We used a multidisciplinary approach that combined high resolution biomass reconstruction, dominance modeling using predictor variables at various scales (catchment, waterbody and water column), biodiversity and functioning estimation at the community level and co-occurrence networks. This approach was successful as it provides a synthetic perspective on the causes and consequences of dominance in phytoplankton communities from 50 waterbodies in the highly populated IDF region (Figure 3). Our results support the idea that the phytoplankton composition and dominance status are directly related to (i) the intensity of human pressures on these ecosystems, and their impact on water column characteristics (total N, TN:TP ratio, water temperature and stratification), but also to (ii) their capacity to buffer such pressures, which depends on their surface, depth and connection to the hydrological network). The consequences of dominance were dependent on the identity of the dominant taxa, and conditions leading to Cyanobacteria dominance challenge more strongly the structure and functioning of phytoplankton communities than conditions favoring the dominance by other groups. Finally, once established this dominance appear to create condition favorables to its self-maintenance as suggested by worldwide increases in the frequency and intensity of cyanobacterial blooms.

## **AKNOWLEDGEMENTS**

The authors would like to thank people that help during field campaigns. The authors acknowledge funding from the French National Research Agency (ANR, [www.agence-nationale-recherche.fr](http://www.agence-nationale-recherche.fr)) through the CYANOTOX (ANR-007-SEST-05) and PULSE (ANR-10-CEPL-0010) projects. This work was also supported by the CNRS through a PhD grant awarded to S. Maloufi (PED grant for students originating from developing countries). We thank the anonymous reviewers for their useful comments on the earlier draft of the manuscript.

## REFERENCES

- Adrian, R., O'Reilly, C.M., Zagarese, H., Baines, S.B., Hessen, D.O., Keller, W., Livingstone, D.M., Sommaruga, R., Straile, D., Van Donk, E., Weyhenmeyer, G.A., Winder, M., 2009. Lakes as sentinels of climate change. *Limnol. Oceanogr.* 54, 2283–2297. [https://doi.org/10.4319/lo.2009.54.6\\_part\\_2.2283](https://doi.org/10.4319/lo.2009.54.6_part_2.2283)
- Aguilera, A., Aubriot, L., Echenique, R.O., Salerno, G.L., Brena, B.M., Pérez, M., Bonilla, S., 2017a. Synergistic effects of nutrients and light favor Nostocales over non-heterocystous cyanobacteria. *Hydrobiologia* 794, 241–255. <https://doi.org/10.1007/s10750-017-3099-1>
- Aguilera, A., Haakonsson, S., Martin, M.V., Salerno, G.L., Echenique, R.O., 2017b. Bloom-forming cyanobacteria and cyanotoxins in Argentina: A growing health and environmental concern. *Limnologia* 69, 103–114. <https://doi.org/10.1016/j.limno.2017.10.006>
- Almanza, V., Pedreros, P., Dail Laughinghouse, H., Féllez, J., Parra, O., Azócar, M., Urrutia, R., 2018. Association between trophic state, watershed use and blooms of cyanobacteria in south-central Chile. *Limnologia* 1–12. <https://doi.org/10.1016/j.limno.2018.11.004>
- Bagatini, I.L., Eiler, A., Bertilsson, S., Klaveness, D., Tessarolli, L.P., Vieira, A.A.H., 2014. Host-specificity and dynamics in bacterial communities associated with bloom-forming freshwater phytoplankton. *PLoS One* 9. <https://doi.org/10.1371/journal.pone.0085950>
- Barberán, A., Bates, S.T., Casamayor, E.O., Fierer, N., 2012. Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J.* 6, 343–351. <https://doi.org/10.1038/ismej.2011.119>
- Beaver, J.R., Tausz, C.E., Scotese, K.C., Pollard, A.I., Mitchell, R.M., 2018. Environmental factors influencing the quantitative distribution of microcystin and common potentially toxigenic cyanobacteria in U.S. lakes and reservoirs. *Harmful Algae* 78, 118–128. <https://doi.org/10.1016/j.hal.2018.08.004>
- Beck, A.E., Shen, P.Y., Beltrami, H., Mareschal, J.C., Šafanda, J., Sebagenzi, M.N., Vasseur, G., Wang, K., 1992. A comparison of five different analyses in the interpretation of five borehole temperature data sets. *Glob. Planet. Change* 6, 101–112. [https://doi.org/10.1016/0921-8181\(92\)90029-A](https://doi.org/10.1016/0921-8181(92)90029-A)
- Behl, S., Donval, A., Stiborb, H., 2011. The relative importance of species diversity and functional group diversity on carbon uptake in phytoplankton communities. *Limnol. Oceanogr.* 56, 683–694. <https://doi.org/10.4319/lo.2011.56.2.0683>

- Bernhardt-Barry, M., Jyotishka, D., Wood, C., Price, J., 2018. Predicting Soil Type from Non-destructive Geophysical Data using Bayesian Statistical Methods. Fayetteville.
- Bonilla, S., Aubriot, L., Soares, M.C.S., González-Piana, M., Fabre, A., Huszar, V.L.M., Lüring, M., Antoniadou, D., Padisák, J., Kruk, C., 2012. What drives the distribution of the bloom-forming cyanobacteria *Planktothrix agardhii* and *Cylindrospermopsis raciborskii*? FEMS Microbiol. Ecol. 79, 594–607. <https://doi.org/10.1111/j.1574-6941.2011.01242.x>
- Camera, C., Zomeni, Z., Noller, J.S., Zissimos, A.M., Christoforou, I.C., Bruggeman, A., 2017. A high resolution map of soil types and physical properties for Cyprus: A digital soil mapping optimization. Geoderma 285, 35–49. <https://doi.org/10.1016/j.geoderma.2016.09.019>
- Carey, C.C., Brown, B.L., Cottingham, K.L., 2017. The cyanobacterium *Gloeotrichia echinulata* increases the stability and network complexity of phytoplankton communities. Ecosphere 8. <https://doi.org/10.1002/ecs2.1830>
- Carey, C.C., Cottingham, K.L., Weathers, K.C., Brentrup, J.A., 2014. Experimental blooms of the cyanobacterium *Gloeotrichia echinulata* increase phytoplankton biomass, richness and diversity in an oligotrophic lake 36, 364–377. <https://doi.org/10.1093/plankt/fbt105>
- Carey, C.C., Ibelings, B.W., Hoffmann, E.P., Hamilton, D.P., Brookes, J.D., 2012. Eco-physiological adaptations that favour freshwater cyanobacteria in a changing climate. Water Res. 46, 1394–1407. <https://doi.org/10.1016/j.watres.2011.12.016>
- Catherine, A., Mouillot, D., Escoffier, N., Bernard, C., Troussellier, M., 2010. Cost effective prediction of the eutrophication status of lakes and reservoirs. Freshw. Biol. 55, 2425–2435. <https://doi.org/10.1111/j.1365-2427.2010.02452.x>
- Catherine, A., Troussellier, M., Bernard, C., 2008. Design and application of a stratified sampling strategy to study the regional distribution of cyanobacteria (Ile-de-France, France). Water Res. 42, 4989–5001. <https://doi.org/10.1016/j.watres.2008.09.028>
- CEN (European Committee for Standardization), 2006. Water quality - Guidance standard on the enumeration of phytoplankton using inverted microscopy (Utermöhl technique), Statute. Brussels.
- Chiu, C., Wang, Y., Walther, B.A., Chao, A., 2014. An improved nonparametric lower bound of species richness via a modified good–turing frequency formula. BIOMETRIC Methodol. <https://doi.org/10.1111/biom.12200>
- Corcoran, A.A., Boeing, W.J., 2012. Biodiversity Increases the Productivity and Stability of Phytoplankton Communities. PLoS One 7, 1–9. <https://doi.org/10.1371/journal.pone.0049397>
- Dai, W., Chen, J., Xiong, J., 2018. Concept of microbial gatekeepers: Positive guys? Appl. Microbiol. Biotechnol. <https://doi.org/10.1007/s00253-018-9522-3>
- Danczak, R.E., Johnston, M.D., Kenah, C., Slattery, M., Wilkins, M.J., 2018. Microbial Community Cohesion Mediates Community Turnover in Unperturbed Aquifers. mSystems 3, 1–15. <https://doi.org/10.1128/mSystems.00066-18>

- Deng, Y., Jiang, Y.H., Yang, Y., He, Z., Luo, F., Zhou, J., 2012. Molecular ecological network analyses. *BMC Bioinformatics* 13, 113. <https://doi.org/10.1186/1471-2105-13-113>
- Dias, F., Antunes, J.T., Ribeiro, T., Azevedo, J., Vasconcelos, V., Leão, P.N., 2017. Cyanobacterial allelochemicals but not cyanobacterial cells markedly reduce microbial community diversity. *Front. Microbiol.* 8, 1495. <https://doi.org/10.3389/fmicb.2017.01495>
- Dinno, A., 2017. *dunn.test*: Dunn's Test of Multiple Comparisons Using Rank Sums.
- Dokulil, M.T., Teubner, K., 2000. Cyanobacterial dominance in lakes. *Hydrobiologia* 438, 1–12. <https://doi.org/10.1023/A:1004155810302>
- Faust, K., Lima-Mendez, G., Lerat, J.S., Sathirapongsasuti, J.F., Knight, R., Huttenhower, C., Lenaerts, T., Raes, J., 2015. Cross-biome comparison of microbial association networks. *Front. Microbiol.* 6, 1–13. <https://doi.org/10.3389/fmicb.2015.01200>
- Faust, K., Raes, J., 2012. Microbial interactions: From networks to models. *Nat. Rev. Microbiol.* 10, 538–550. <https://doi.org/10.1038/nrmicro2832>
- Filstrup, C.T., Hillebrand, H., Heathcote, A.J., Harpole, W.S., Downing, J.A., 2014. Cyanobacteria dominance influences resource use efficiency and community turnover in phytoplankton and zooplankton communities. *Ecol. Lett.* 17, 464–474. <https://doi.org/10.1111/ele.12246>
- Freilich, M.A., Wieters, E., Broitman, B.R., Marquet, P.A., Navarrete, S.A., 2018. Species co-occurrence networks: Can they reveal trophic and non-trophic interactions in ecological communities? *Ecology* 99, 690–699. <https://doi.org/10.1002/ecy.2142>
- Gamfeldt, L., Hillebrand, H., 2011. Effects of total resources, resource ratios, and species richness on algal productivity and evenness at both metacommunity and local scales. *PLoS One* 6, e21972. <https://doi.org/10.1371/journal.pone.0021972>
- Goldford, J.E., Lu, N., Bajić, D., Estrela, S., Tikhonov, M., Sanchez-Gorostiaga, A., Segrè, D., Mehta, P., Sanchez, A., 2018. Emergent simplicity in microbial community assembly. *Science* (80-. ). 361, 469–474. <https://doi.org/10.1126/science.aat1168>
- Guedes, I.A., Rachid, C.T.C.C., Rangel, L.M., Silva, L.H.S., Bisch, P.M., Azevedo, S.M.F.O., Pacheco, A.B.F., 2018. Close link between harmful cyanobacterial dominance and associated bacterioplankton in a tropical eutrophic reservoir. *Front. Microbiol.* 9. <https://doi.org/10.3389/fmicb.2018.00424>
- Herren, C.M., McMahon, K.D., 2018. Keystone taxa predict compositional change in microbial communities. *Environ. Microbiol.* 20, 2207–2217. <https://doi.org/10.1111/1462-2920.14257>
- Herren, C.M., McMahon, K.D., 2017. Cohesion: A method for quantifying the connectivity of microbial communities. *ISME J.* 11, 2426–2438. <https://doi.org/10.1038/ismej.2017.91>
- Hillebrand, H., 1999. Biovolume Calculation for Palagic and Benthic Microalgae. *J. Phycol.* 424, 403–424.
- Hodapp, D., Meier, S., Muijsers, F., Badewien, T.H., Hillebrand, H., 2015. Structural equation modeling approach to the diversity-productivity relationship of Wadden Sea phytoplankton. *Mar. Ecol. Prog. Ser.* 523, 31–40. <https://doi.org/10.3354/meps11153>

- Holland, A., Kinnear, S., 2013. Interpreting the Possible Ecological Role(s) of Cyanotoxins: Compounds for Competitive Advantage and/or Physiological Aide? *Mar. Drugs* 11, 2239–2258. <https://doi.org/10.3390/md11072239>
- Holmes, R., Norris, R., Smayda, T., Wood, E., 1969. Collection, fixation, identification, and enumeration of phytoplankton standing stock., in: Anonymous (Ed.), *Recommended Procedures for Measuring the Productivity of Plankton Standing Stock and Related Oceanic Properties*. Washington DC, pp. 17–46. <https://doi.org/10.1128/AEM.02566-14>
- Hu, A., Ju, F., Hou, L., Li, J., Yang, X., Wang, H., Mulla, S.I., Sun, Q., Bürgmann, H., Yu, C.P., 2017. Strong impact of anthropogenic contamination on the co-occurrence patterns of a riverine microbial community. *Environ. Microbiol.* 19, 4993–5009. <https://doi.org/10.1111/1462-2920.13942>
- Humphries, M.D., Gurney, K., 2008. Network “small-world-ness”: A quantitative method for determining canonical network equivalence. *PLoS One* 3. <https://doi.org/10.1371/journal.pone.0002051>
- Jackson, D.A., 1997. Compositional Data in Community Ecology : The Paradigm or Peril of Proportions ? *Ecology* 78, 929–940.
- Jones, A.C., Hambright, K.D., Caron, D.A., 2018. Ecological Patterns Among Bacteria and Microbial Eukaryotes Derived from Network Analyses in a Low-Salinity Lake. *Microb. Ecol.* 75, 917–929. <https://doi.org/10.1007/s00248-017-1087-7>
- Karimi, B., Alain, P., Nicolas, M., Bernard, N., Gilbert, D., Ranjard, L., Maron, P.A., Chemidlin-Prevost Boure, N., Bernard, N., Gilbert, D., Ranjard, L., 2017. Microbial diversity and ecological networks as indicators of environmental quality. *Environ. Chem. Lett.* 15, 265–281. <https://doi.org/10.1007/s10311-017-0614-6>
- Kurtz, Z.D., Müller, C.L., Miraldi, E.R., Littman, D.R., Blaser, M.J., Bonneau, R. a., 2015. Sparse and Compositionally Robust Inference of Microbial Ecological Networks. *PLoS Comput. Biol.* 11, 1–25. <https://doi.org/10.1371/journal.pcbi.1004226>
- Leflaive, J., Ten-Hage, L., 2007. Algal and cyanobacterial secondary metabolites in freshwaters: A comparison of allelopathic compounds and toxins. *Freshw. Biol.* 52, 199–214. <https://doi.org/10.1111/j.1365-2427.2006.01689.x>
- Lehtinen, S., Tamminen, T., Ptacnik, R., Andersen, T., 2017. Phytoplankton species richness, evenness, and production in relation to nutrient availability and imbalance. *Limnol. Oceanogr.* 62, 1393–1408. <https://doi.org/10.1002/lno.10506>
- Li, Z., Wang, G., Yang, Y., Yan, X., Sun, H., Shi, A., Wang, M., Xu, X., Wu, S., 2017. Climate warming and cyanobacteria blooms: Looks at their relationships from a new perspective. *Water Res.* 125, 449–457. <https://doi.org/10.1016/j.watres.2017.09.008>
- Liaw, A., Wiener, M., 2002. Classification and Regression by Random Forest. *R news* 2/3, 18–22. <https://doi.org/10.1177/154405910408300516>
- Litchman, E., 2003. Competition and coexistence of phytoplankton under fluctuating light: Experiments with two cyanobacteria. *Aquat. Microb. Ecol.* 31, 241–248. <https://doi.org/10.3354/ame031241>

- Liu, L., Chen, H., Liu, M., Yang, J.R., Xiao, P., Wilkinson, D.M., Yang, J., 2019. Response of the eukaryotic plankton community to the cyanobacterial biomass cycle over 6 years in two subtropical reservoirs. *ISME J.* <https://doi.org/10.1038/s41396-019-0417-9>
- Liu, M., Liu, L., Chen, H., Yu, Z., Yang, J.R., Xue, Y., Huang, B., Yang, J., 2019. Community dynamics of free-living and particle-attached bacteria following a reservoir *Microcystis* bloom. *Sci. Total Environ.* 660, 501–511. <https://doi.org/10.1016/j.scitotenv.2018.12.414>
- Louati, I., Pascault, N., Debroas, D., Bernard, C., Humbert, J.F., Leloup, J., 2015. Structural diversity of bacterial communities associated with bloom-forming freshwater cyanobacteria differs according to the cyanobacterial genus. *PLoS One* 10. <https://doi.org/10.1371/journal.pone.0140614>
- Lovell, D., Muller, W., Taylor, J., Zwart, A., Helliwell, C., Pawlowsky-Glahn, V., Buccianti, A., 2011. Proportions, percentages, ppm: do the molecular biosciences treat compositional data right?, in: *Compositional Data Analysis: Theory and Applications*. pp. 1–13.
- Luo, F., Zhong, J., Yang, Y., Scheuermann, R.H., Zhou, J., 2006. Application of random matrix theory to complex networks. *Phys. Lett. A* 357, 420–423. [https://doi.org/10.1007/978-3-319-17037-4\\_6](https://doi.org/10.1007/978-3-319-17037-4_6)
- Maloufi, S., Catherine, A., Mouillot, D., Louvard, C., Couté, A., Bernard, C., Troussellier, M., 2016. Environmental heterogeneity among lakes promotes hyper  $\beta$ -diversity across phytoplankton communities. *Freshw. Biol.* 61, 633–645. <https://doi.org/10.1111/fwb.12731>
- Mantzouki, E., Visser, P.M., Bormans, M., Ibelings, B.W., 2016. Understanding the key ecological traits of cyanobacteria as a basis for their management and control in changing lakes. *Aquat. Ecol.* 50, 333–350. <https://doi.org/10.1007/s10452-015-9526-3>
- Marion, J.W., Zhang, F., Cutting, D., Lee, J., 2017. Associations between county-level land cover classes and cyanobacteria blooms in the United States. *Ecol. Eng.* 108, 556–563. <https://doi.org/10.1016/j.ecoleng.2017.07.032>
- Marmen, S., Aharonovich, D., Grossowicz, M., Blank, L., Yacobi, Y.Z., Sher, D.J., 2016. Distribution and habitat specificity of potentially-toxic *Microcystis* across climate, land, and water use gradients. *Front. Microbiol.* 7. <https://doi.org/10.3389/fmicb.2016.00271>
- Moe, S.J., Haande, S., Couture, R.M., 2016. Climate change, cyanobacteria blooms and ecological status of lakes: A Bayesian network approach. *Ecol. Modell.* 337, 330–347. <https://doi.org/10.1016/j.ecolmodel.2016.07.004>
- Monchamp, M.E., Spaak, P., Domaizon, I., Dubois, N., Bouffard, D., Pomati, F., 2018. Homogenization of lake cyanobacterial communities over a century of climate change and eutrophication. *Nat. Ecol. Evol.* 2, 317–324. <https://doi.org/10.1038/s41559-017-0407-0>
- Moura, A.D.N., Aragão-Tavares, N.K.C., Amorim, C.A., 2018. Cyanobacterial blooms in freshwater bodies from a semiarid region, northeast Brazil: A review. *J. Limnol.* 77, 179–188. <https://doi.org/10.4081/jlimnol.2018.1646>

- Ndlela, L.L., Oberholster, P.J., Van Wyk, J.H., Cheng, P.H., 2016. An overview of cyanobacterial bloom occurrences and research in Africa over the last decade. *Harmful Algae* 60, 11–26. <https://doi.org/10.1016/j.hal.2016.10.001>
- O’Farrell, I., Motta, C., Forastier, M., Polla, W., Otaño, S., Meichtry, N., Devercelli, M., Lombardo, R., 2019. Ecological meta-analysis of bloom-forming planktonic Cyanobacteria in Argentina. *Harmful Algae* 83, 1–13. <https://doi.org/10.1016/j.hal.2019.01.004>
- OCED, 1982. Eutrophication of waters. Monitoring, assesement and control. Paris.
- Oksanen, J., Blanchet, F.G, Kindt, R., Legendre, P., Minchin, P.R., O’Hara, R., 2016. Vegan: community ecology package. R Packag. version.
- Olenina, I., Hajdu, S., Edler, L., Wasmund, N., Busch, S., Göbel, J., Gromisz, S., Huseby, S., Huttunen, M., Jaanus, A., Kokkonen, P., Ledaine, I., Niemkiewicz, E., 2006. Biovolumes and size-classes of phytoplankton in the Baltic Sea, *Baltic Sea Environment Proceedings*. [https://doi.org/HELCOM Balt.Sea Environ. Proc. No. 106, 144pp](https://doi.org/HELCOM%20Balt.Sea%20Environ.%20Proc.%20No.%20106,%20144pp).
- Paerl, H.W., Huisman, J., 2008. Climate: Blooms like it hot. *Science* (80-. ). 320, 57–58. <https://doi.org/10.1126/science.1155398>
- Paerl, H.W., Otten, T.G., 2016. Duelling “CyanoHABs”: Unravelling the environmental drivers controlling dominance and succession among diazotrophic and non-N<sub>2</sub>-fixing harmful cyanobacteria. *Environ. Microbiol.* 18, 316–324. <https://doi.org/10.1111/1462-2920.13035>
- Paulson, J.N., Colin Stine, O., Bravo, H.C., Pop, M., 2013. Differential abundance analysis for microbial marker-gene surveys. *Nat. Methods* 10, 1200–1202. <https://doi.org/10.1038/nmeth.2658>
- Pavlopoulos, G.A., Secrier, M., Moschopoulos, C.N., Soldatos, T.G., Kossida, S., Aerts, J., Schneider, R., Bagos, P.G., 2011. Using graph theory to analyze biological networks. *BioData Min.* 4, 10. <https://doi.org/10.1186/1756-0381-4-10>
- Persaud, A.D., Paterson, A.M., Dillon, P.J., Winter, J.G., Palmer, M., Somers, K.M., 2015. Forecasting cyanobacteria dominance in Canadian temperate lakes. *J. Environ. Manage.* 151, 343–352. <https://doi.org/10.1016/j.jenvman.2015.01.009>
- Peura, S., Bertilsson, S., Jones, R.I., Eiler, A., 2015. Resistant Microbial Cooccurrence Patterns Inferred by Network Topology. *Appl. Environ. Microbiol.* 81, 2090–2097. <https://doi.org/10.1128/AEM.03660-14>
- Ptacnik, R., Solimini, A.G., Andersen, T., Tamminen, T., Brettum, P., Lepisto, L., Willen, E., Rekolainen, S., 2008. Diversity predicts stability and resource use efficiency in natural phytoplankton communities. *Proc. Natl. Acad. Sci.* 105, 5134–5138. <https://doi.org/10.1073/pnas.0708328105>
- Rinke, K., Yeates, P., Rothhaupt, K.O., 2010. A simulation study of the feedback of phytoplankton on thermal structure via light extinction. *Freshw. Biol.* 55, 1674–1693. <https://doi.org/10.1111/j.1365-2427.2010.02401.x>
- Röttgers, L., Faust, K., 2018. From hairballs to hypotheses-biological insights from microbial networks. *FEMS Microbiol. Rev.* 42, 761–780. <https://doi.org/10.1093/femsre/fuy030>

- Scheffer, M., Rinaldi, S., Gragnani, A., Mur, L.R., van Nes, E., 1997. On the dominance of filamentous blue-green algae in shallow, turbid lakes. *Ecology* 78, 272–282. <https://doi.org/10.1371/journal.ppat.1003731>
- Schwaderer, A.S., Yoshiyama, K., De Tezanos Pinto, P., Swenson, N.G., Klausmeier, C.A., Litchman, E., 2011. Eco-evolutionary differences in light utilization traits and distributions of freshwater phytoplankton. *Limnol. Oceanogr.* 56, 589–598. <https://doi.org/10.4319/lo.2011.56.2.0589>
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. <https://doi.org/10.1101/gr.1239303>
- Shi, S., Nuccio, E.E., Shi, Z.J., He, Z., Zhou, J., Firestone, M.K., Johnson, N., 2016. The interconnected rhizosphere: High network complexity dominates rhizosphere assemblages. *Ecol. Lett.* 19, 926–936. <https://doi.org/10.1111/ele.12630>
- Sun, J., Liu, D., 2003. Geometric models for calculating cell biovolume and surface area for phytoplankton. *J. Plankton Res.* 25, 1331–1346. <https://doi.org/10.1093/plankt/fbg096>
- Tian, W., Zhang, H., Zhang, J., Zhao, L., Miao, M., Huang, H., 2017. Biodiversity effects on resource use efficiency and community turnover of plankton in Lake Nansihu, China. *Environ. Sci. Pollut. Res.* 24, 11279–11288. <https://doi.org/10.1007/s11356-017-8758-2>
- Toporowska, M., Pawlik-Skowrońska, B., 2014. Four-year study on phytoplankton biodiversity in a small hypertrophic lake affected by water blooms of toxigenic cyanobacteria. *Polish J. Environ. Stud.* 23, 491–499.
- Ullah, H., Nagelkerken, I., Goldenberg, S.U., Fordham, D.A., 2018. Climate change could drive marine food web collapse through altered trophic flows and cyanobacterial proliferation. *PLoS Biol.* 16, e2003446. <https://doi.org/10.1371/journal.pbio.2003446>
- Van Goethem, M.W., Makhalanyane, T.P., Cowan, D.A., Valverde, A., 2017. Cyanobacteria and Alphaproteobacteria may facilitate cooperative interactions in niche communities. *Front. Microbiol.* 8, 2099. <https://doi.org/10.3389/fmicb.2017.02099>
- Vaughn, C.C., 2010. Biodiversity Losses and Ecosystem Function in Freshwaters: Emerging Conclusions and Research Directions. *Bioscience* 60, 25–35. <https://doi.org/10.1525/bio.2010.60.1.7>
- Venables, W.N., Ripley, B.D., 2002. *Modern Applied Statistics with S*, Springer. ed. New York. <https://doi.org/10.1007/978-0-387-21706-2>
- Wagner, C., Adrian, R., 2009. Cyanobacteria dominance: Quantifying the effects of climate change. *Limnol. Oceanogr.* 54, 2460–2468. [https://doi.org/10.4319/lo.2009.54.6\\_part\\_2.2460](https://doi.org/10.4319/lo.2009.54.6_part_2.2460)
- Wang, F., Liang, Y., Jiang, Y., Yang, Y., Xue, K., Xiong, J., Zhou, J., Sun, B., 2015. Planting increases the abundance and structure complexity of soil core functional genes relevant to carbon and nitrogen cycling. *Sci. Rep.* 5, 14345. <https://doi.org/10.1038/srep14345>
- Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of ‘small-world’ networks. *Nature* 393, 440–442. <https://doi.org/10.1038/30918>

- Woodhouse, J.N., Kinsela, A.S., Collins, R.N., Bowling, L.C., Honeyman, G.L., Holliday, J.K., Neilan, B.A., 2016. Microbial communities reflect temporal changes in cyanobacterial composition in a shallow ephemeral freshwater lake. *ISME J.* 10, 1337–1351. <https://doi.org/10.1038/ismej.2015.218>
- [www.iau-idf.fr](http://www.iau-idf.fr) [WWW Document], n.d. URL <http://www.iau-idf.fr/en/paris-region/portraying-paris-region/land-use-map>
- Xue, Y., Chen, H., Yang, J.R., Liu, M., Huang, B., Yang, J., 2018. Distinct patterns and processes of abundant and rare eukaryotic plankton communities following a reservoir cyanobacterial bloom. *ISME J.* 12, 2263–2277. <https://doi.org/10.1038/s41396-018-0159-0>
- Yang, C., Wang, Q., Simon, P.N., Liu, J., Liu, L., Dai, X., Zhang, X., Kuang, J., Igarashi, Y., Pan, X., Luo, F., 2017. Distinct network interactions in particle-associated and free-living bacterial communities during a *Microcystis aeruginosa* bloom in a plateau lake. *Front. Microbiol.* 8, 1–15. <https://doi.org/10.3389/fmicb.2017.01202>
- Yang, J.R., Lv, H., Isabwe, A., Liu, L., Yu, X., Chen, H., Yang, J., 2017. Disturbance-induced phytoplankton regime shifts and recovery of cyanobacteria dominance in two subtropical reservoirs. *Water Res.* 120, 52–63. <https://doi.org/10.1016/j.watres.2017.04.062>
- Zhao, D., Shen, F., Zeng, J., Huang, R., Yu, Z., Wu, Q.L., 2016. Network analysis reveals seasonal variation of co-occurrence correlations between Cyanobacteria and other bacterioplankton. *Sci. Total Environ.* 573, 817–825. <https://doi.org/10.1016/j.scitotenv.2016.08.150>

## TABLES CAPTION

### **Table 1: Accuracy of classification approaches for the identification of dominance status in phytoplankton communities**

This table represents the proportion of communities under dominance by a given group that were correctly classified as dominated by this group. For each model and each group, accuracy values were averaged (mean±sd) across the nine combinations of dominance thresholds. MLR: multinomial logistic regression; RF: random forest.

### **Table 2: Environmental drivers of dominance in phytoplankton communities**

The three models are multinomial logistic regressions (MLR). Average  $\chi^2$  values and number of times the variable participated significantly in the classification were estimated across the nine threshold combinations.

### **Table 3: Global test for differences across dominance groups for various community-level properties**

The  $\chi^2$  and p.values presented are averages estimated across the nine thresholds combinations.

### **Table 4: Characteristics of co-occurrence networks**

The presented index values are mean±sd estimated across the nine thresholds combinations.

### **Table 5: Relationship between biomass and centrality in phytoplankton co-occurrence networks**

	<b>MLR</b>	<b>RF</b>
<b>Model 1</b>	<b>0,63 ± 0,07</b>	<b>0,43 ± 0,03</b>
Bacillariophyta	0,84 ± 0,16	0,17 ± 0,16
Chlorophyta	0,68 ± 0,07	0,50 ± 0,16
Cyanobacteria	0,61 ± 0,06	0,55 ± 0,12
Dinophyta	0,61 ± 0,23	0,00 ± 0,00
Others	0,73 ± 0,19	0,33 ± 0,07
No dominance	0,59 ± 0,06	0,39 ± 0,20
<b>Model 2</b>	<b>0,69 ± 0,08</b>	<b>0,67 ± 0,12</b>
Dominance	0,73 ± 0,09	0,84 ± 0,13
No dominance	0,58 ± 0,17	0,19 ± 0,16
<b>Model 3</b>	<b>0,80 ± 0,14</b>	<b>0,57 ± 0,02</b>
Bacillariophyta	0,98 ± 0,06	0,31 ± 0,26
Chlorophyta	0,79 ± 0,14	0,74 ± 0,03
Cyanobacteria	0,81 ± 0,14	0,72 ± 0,06
Dinophyta	0,72 ± 0,22	0,00 ± 0,00
Others	0,83 ± 0,16	0,37 ± 0,06

Sheet1

Predictors	Model 1		Model 2		Model 3		
	Average $\chi^2$	number of p.value < 0.05	Average $\chi^2$	number of p.value < 0.05	Average $\chi^2$	number of p.value < 0.05	
<b>Water column</b>	Total Nitrogen	9,4	4	0,2	0	6,4	4
	Total Phosphorus	4,2	0	0,7	0	4,5	1
	N:P ratio	9,3	2	1,5	1	10,0	3
	Temperature	9,6	3	1,1	0	11,0	4
	Thermal stratification of lake water	7,0	1	1,3	0	9,6	3
<b>Waterbody</b>	Surface of waterbodies	7,0	1	1,0	0	7,9	2
	Depth of waterbodies	5,8	1	1,5	0	10,0	3
	Altitude of waterbodies	7,3	1	0,6	0	2,8	0
	Connection with hydrological network	12,9	3	0,7	0	11,0	2
	Waterbody is on a flooded area	2,8	0	0,1	0	1,6	0
<b>Catchment</b>	Surface of catchment	4,4	0	0,3	0	4,1	1
	Catchment to waterbody surface ratio	5,3	0	0,4	0	5,4	0
	Drainage intensity	6,7	0	1,1	0	3,7	0
	Percentage of impervious surface	0,6	0	0,0	0	2,2	1
	Percentage of agricultural surface	1,8	0	0,0	0	1,7	1
	Percentage of forested surface	0,4	0	0,0	0	0,0	0

Sheet1

<i>Kruskal–Wallis test</i>	df	$\chi^2$	p.value	number of p.value <0.05
log(Biomass)	5	23,4	0,000	9
Richness	5	29,2	0,000	9
RUE <sub>P</sub>	5	13,8	0,023	8
RUE <sub>N</sub>	5	24,3	0,000	9
Negative cohesion	5	43,8	0,000	9
Positive cohesion	5	44,8	0,000	9
Evenness	5	72,1	0,000	9

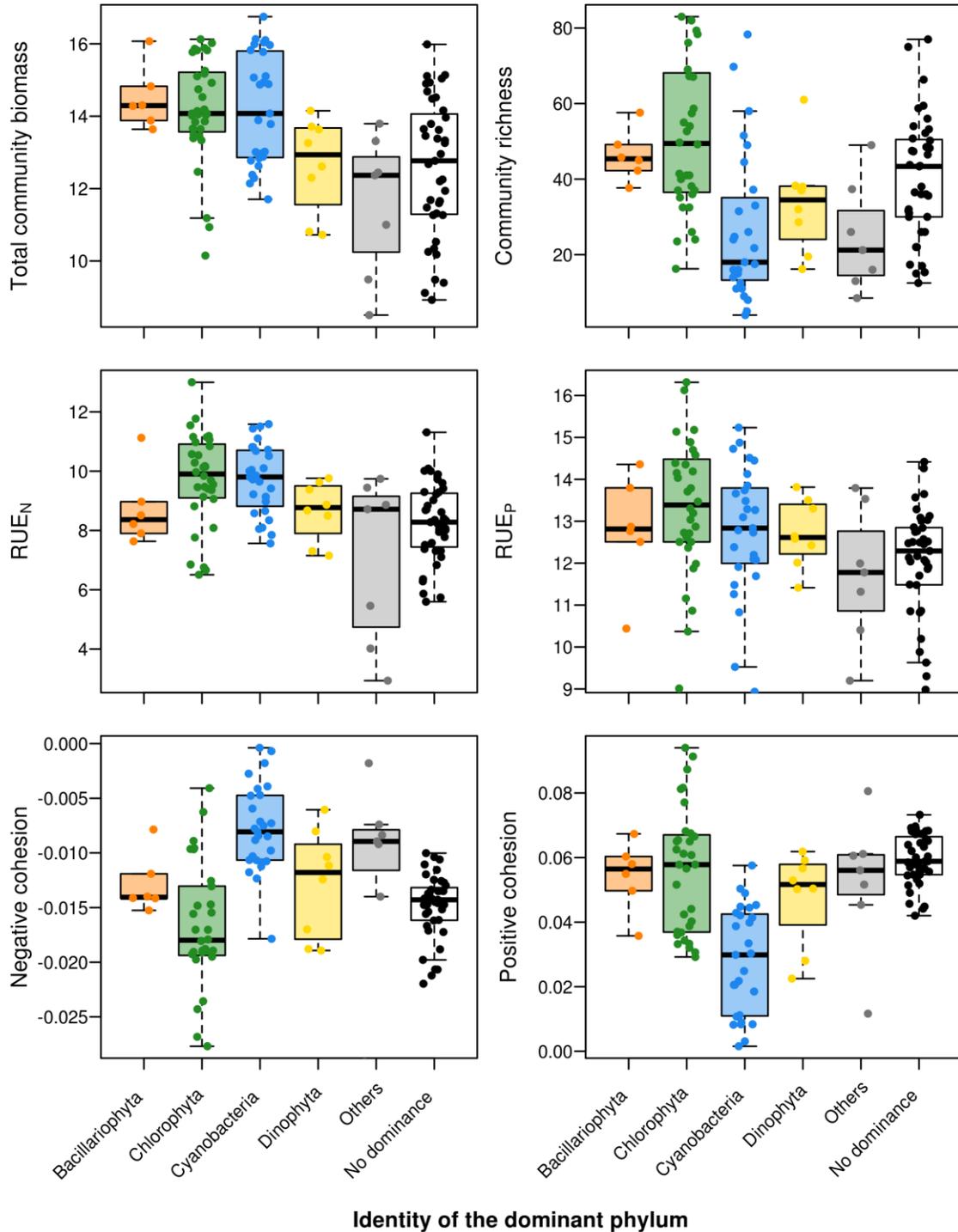
	<b>No dominance</b>	<b>Chlorophyta dominated</b>	<b>Cyanobacteria dominated</b>
Number of communities used to construct networks	35.7 ± 14.3	30.4 ± 7.0	26.1 ± 3.4
<b>Global network characteristics</b>			
Number of nodes	89 ± 10	106 ± 4	64 ± 7
Number of links	838 ± 611	1312 ± 595	728 ± 232
% of negative links	34.7 ± 24.8	34.5 ± 14.7	48.0 ± 7.5
Density	0.20 ± 0.11	0.23 ± 0.1	0.35 ± 0.08
Geodesic efficiency	0.54 ± 0.1	0.58 ± 0.07	0.66 ± 0.05
Average geodesic distance	2.20 ± 0.44	1.98 ± 0.24	1.76 ± 0.15
<b>Nodes characteristics</b>			
Average degree	18.1 ± 10.8	24.7 ± 10.7	22.3 ± 5.8
Average betweenness	52.2 ± 18.2	50.9 ± 12.4	24.1 ± 5.0
Average clustering coefficient	0.58 ± 0.08	0.55 ± 0.04	0.65 ± 0.05

Sheet1

	<b>df</b>	<b>stat</b>	<b>r</b>	<b>p.value</b>	<i>Number of significant cases</i>
<b><i>Node degree</i></b>					
Not dominated	87	-1,87	-0,18	0,375	3
Chlorophyta	104	-1,75	-0,17	0,232	3
Cyanobacteria	62	-2,69	-0,32	0,023	8
<b><i>Node betweenness</i></b>					
Not dominated	87	-1,08	-0,12	0,305	0
Chlorophyta	104	-1,07	-0,10	0,364	0
Cyanobacteria	62	-1,53	-0,19	0,194	4

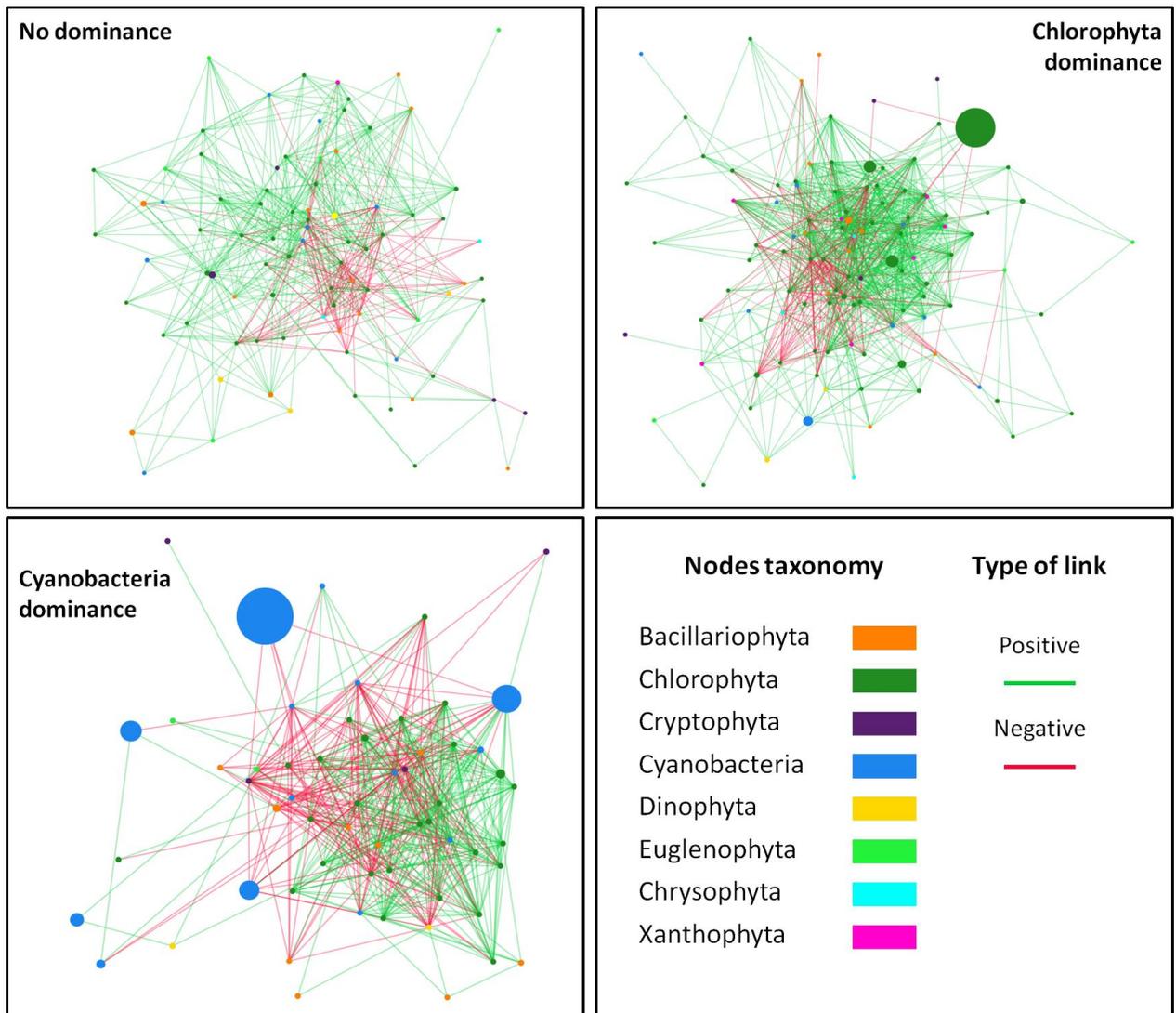
### Figure 1: Biological characteristics of phytoplankton communities in various dominance contexts

Here, and for illustrative purpose, we present the case where both the dominance and relative biomass thresholds were set to 0.5. Communities were first separated according to the estimated value of the dominance index, *i.e.* no dominance (index < 0.5) or under dominance (index > 0.5). Then communities under dominance were grouped according to which phyla was dominating the community (*i.e.* with a relative biomass > 0.5). The box and whiskers plots represent the median (black line) the first and fourth quartiles (colored boxes) and the 95% confidence interval (whiskers). The dots represent actual data points.



**Figure 2: Structures of phytoplankton co-occurrence networks under various dominance contexts**

Nodes correspond to taxa and links correspond to significantly positive (green) and negative (red) associations. Nodes are colored according to their taxonomy. Nodes size represents their average biomass in the communities composing the network. These networks were constructed based on communities grouping determined using both dominance and relative biomass thresholds equal to 0.5.



**Figure 3: Drivers and consequences of dominance in phytoplankton communities**

This figure synthesizes the results from the paper (cf. section 4.3 of the main text and conclusion).

