The habitat template of phytoplankton morphology-based functional groups

2 AUTHORS:

Carla Kruk
University of the Republic, Uruguay
32 PUBLICATIONS   **1,118** CITATIONS

SEE PROFILE

Angel Marley
University of the Republic, Uruguay
20 PUBLICATIONS   **68** CITATIONS

SEE PROFILE

PHYTOPLANKTON

The habitat template of phytoplankton morphology-based functional groups

Carla Kruk · Angel M. Segura

Abstract The identification of the main factors driving phytoplankton community structure is essential to understand and adequately manage freshwater ecosystems. We hypothesize that differences in morphological traits reflect phytoplankton functional properties that will be selected under particular environmental conditions, namely their *habitat template*. We apply a morphology-based functional groups (MBFG) approach to classify phytoplankton organisms and define each group template. We use machine learning techniques to classify a large number of phytoplankton communities and environmental variables from different climate zones and continents. Random forest analysis explained well the distribution of most groups' biovolume and the selected variables reflected ecological preferences according to morphology. By means of a classification tree it was also possible to identify thresholds of the environmental variables promoting groups dominance in different lakes. For example group III (filaments with aerotopes and high surface/volume including potentially toxic species) was dominant when light attenuation coefficient was >3.9 m^{-1} and total nitrogen was $>2,800$ μg l^{-1}. We demonstrate that morphology captures ecological preferences of phytoplankton groups and provides empirical values to describe their habitat template.

Keywords Morphological traits · Functional groups · Random forest · CART · Environmental change

Carla Kruk and Angel M. Segura have contributed equally to this article.

Guest editors: N. Salmaso, L. Naselli-Flores, L. Cerasino, G. Flaim, M. Tolotti & J. Padisák / Phytoplankton responses to human impacts at different scales: 16th workshop of the International Association of Phytoplankton Taxonomy and Ecology (IAP)

C. Kruk (✉)
Laboratory of Ethology, Ecology and Evolution, Instituto de Investigaciones Biológicas Clemente Estable, Avenida Italia 3318, 11600 Montevideo, Uruguay
e-mail: ckruk@yahoo.com

C. Kruk
Functional Ecology of Aquatic Systems, Limnology, IECA, Facultad de Ciencias, Universidad de la República, Iguá 4225, 11400 Montevideo, Uruguay

A. M. Segura
Functional Ecology of Aquatic Systems, Oceanography and Marine Ecology, IECA, Facultad de Ciencias, Universidad de la República, Iguá 4225, 11400 Montevideo, Uruguay

Introduction

Phytoplankton is essential for the functioning of our planet as it accounts for half of earth's primary production (Falkowski et al., 2003; Arrigo, 2005).

Also many problems of water quality are caused by phytoplankton with potentially serious implications for human and ecosystems health (Huisman et al., 2005). The identification of the main biotic and abiotic factors controlling phytoplankton in lakes is thus essential for the adequate management of freshwater ecosystems (Peretyatko et al., 2007).

Aggregated estimators of phytoplankton communities (e.g. total biomass) may work to describe overall community responses to varying environmental conditions (Vollenweider, 1976; Scheffer et al., 2003). However, phytoplankton species differ widely in their responses to environmental change, including their way of resources acquisition (light and nutrients) to grow, and the way of avoidance of mortality (washout, sedimentation and grazing) (Margalef, 1978; Reynolds, 1984a; Naselli-Flores et al., 2007). These features can be combined to describe the species habitat template (sensu Southwood, 1977). This concept views the habitat as a template on which evolution forges characteristic species traits (Southwood, 1988), and can be used to predict community organization (Keddy, 1992). Habitat templates have been built up for phytoplankton for different species, combining traits and environmental gradients (Margalef, 1978; Reynolds, 1988; Reynolds et al., 2002; Salmaso & Padisák, 2007).

In this vein, trait-based approaches have been increasingly applied to explain and predict the response of phytoplankton species to environmental conditions both in marine and continental aquatic systems. Well-known examples are the Plankton Ecology Group (PEG) model (Sommer, 1989) that predicts the seasonal succession in temperate lakes and the Margalef mandala (1978) that explains the main strategies and mechanisms for marine plankton in terms of a trade-off between r and K-selected traits. More recently, models based on functional traits have been shown to capture phytoplankton distribution in the world's oceans quite well (Le Quéré et al., 2005; Follows et al., 2007). These and other examples illustrate that clustering species based on their functional traits makes sense to summarise their response to environmental change.

Morphology-based functional groups

Morphological traits are relatively easy to measure and have clear relationships with the functional properties of phytoplankton (Lewis, 1976; Reynolds,

1984b; Naselli-Flores et al., 2007; Kruk et al., 2010). The morphology-based functional groups (MBFGs) approach clusters organisms in seven groups in terms of morphological traits (e.g. volume and the presence of flagella) independently from the organism's taxonomic affiliation (Kruk et al., 2010) (Fig. 1). In turn, significant differences in growth rate, sinking rates, demographic properties and competitive ability have been shown for these groups (Kruk et al., 2010; Segura et al., 2011, 2010). For example, group I represents small, high surface to volume ratio (S/V) organisms, with high growth rate and low sinking, and with better competitive ability at the beginning of temporal succession. The groups are also well predicted by environmental variables independently from geographical location (Kruk et al., 2011). Based on the morphological traits of each MBFG, potential ecological performance in terms of resources acquisition and avoidance of loss processes (consumption and sinking) have been derived (Table 4 in Kruk et al., 2010). However, the analysis of each MBFG environmental preferences has not been yet established. We hypothesize that differences in morphological traits among MBFG reflect phytoplankton functional properties which will be selected under particular environments. Therefore, MBFG anticipate phytoplankton habitat template.

In this article, we aim to link each MBFG to its habitat template using information from a very large number of phytoplankton communities and environmental variables from different climate zones and continents. We used random forest (RF) regression to evaluate which environmental variables explained best each MBFG biovolume distribution among lakes, and classification trees to detect the particular environmental threshold favouring the dominance of each MBFG. We also re-validate the classification in MBFG evaluating its power using classification trees.

Materials and methods

Fundamentals of the MBFG classification

Phytoplankton organisms are distinguished in seven MBFG groups based on eight morphological traits identified for each organism at the light microscopy (Fig. 1) (Kruk et al., 2010). Group I includes small

Table 1 Average and range for the analysed environmental variables and the seven MBFGs biovolume (I–VII)

Variables	Mean (range)
Temp (°C)	17.2 (0.4–33.0)
Z_{mix} (m)	2.2 (0.1–17.0)
K_D (m^{-1})	3.9 (0.4–43.6)
TN (µg l^{-1})	2424 (35–37,928)
TP (µg l^{-1})	191 (0.0–10,086)
RSi (µg l^{-1})	3492 (0.0–23,533)
TZ (org l^{-1})	1644 (0.5–26,319)
I (mm^3 l^{-1}), $N = 211$	4.8 (0.0–2453)
II (mm^3 l^{-1}), $N = 472$	0.6 (0.0–18)
III (mm^3 l^{-1}), $N = 446$	7.0 (0.0–1,798)
IV (mm^3 l^{-1}), $N = 675$	8.9 (0.0–3,173)
V (mm^3 l^{-1}), $N = 856$	3.3 (0.0–152)
VI (mm^3 l^{-1}), $N = 819$	10.4 (0.0–3,367)
VII (mm^3 l^{-1}), $N = 583$	6.3 (0.0–987)

N number of non zero values

organisms with high S/V. Group II clusters small flagellated organisms with siliceous exoskeletal structures. Group III represents large filaments with aerotopes. Organisms of medium size lacking specialized traits are included in group IV. Group V gathers unicellular or colonial flagellates of medium to large size. Non-flagellated organisms with siliceous exoskeletons are in group VI, and group VII includes large mucilaginous colonies.

The continuous traits included in the classification are: volume (V, µm^3), surface/volume (S/V, µm^{-1}) and maximum linear dimension (MLD, µm), and are calculated based on organisms geometrical approximations following Hillebrand et al. (1999). For all lakes, the organisms are considered as the unit (unicell, colony or filament). For colonial organisms with mucilage, V and S calculations are made for whole colonies including mucilage. The categorical traits incorporated are the presence or otherwise of flagella, mucilage, siliceous exoskeletal structures and aerotopes. To classify the organisms into the seven MBFG the estimation of continuous traits and the presence of the categorical traits has to be noted for each relevant organism for the original sample and not based on the species names. The absence of any of the traits after inspections at the larger magnifications should not be included even if expected based on the taxonomic classification of the species.

A code in the R software (R, 2011) to classify phytoplankton organisms into MBFG according to individual morphological traits is provided in the Supplementary material online. Using the provided code and R software, which can be free downloaded (http://www.r-project.org/) the user can upload a matrix with information about the organisms
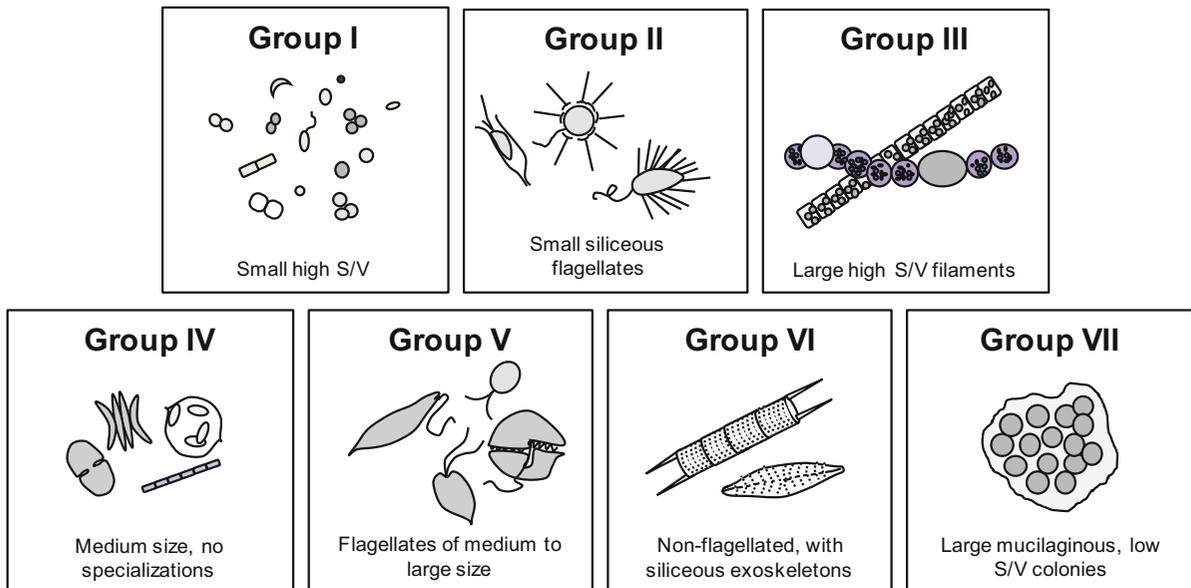

Fig. 1 Schematic representation of the seven MBFGs (Kruk et al., 2010) including a brief description of their morphology. *S* surface, *V* volume

morphological traits and will obtain a new matrix with the original organisms classified into MBFGs (Kruk et al., 2010).

Database

In order to evaluate the MBFG environmental preferences, we compiled a database of 711 species and lake environmental variables from 211 lakes with a total number of 925 samples located within four climate zones in South America, Europe and North America, and covering a wide range of environmental characteristics (Kosten et al., 2009; Kruk et al., 2009, 2011) (Table 1). For 107 of the lakes, information was obtained from published (De León, 2000; Mazzeo et al., 2003) and unpublished sources (V.L.M. Huszar, personal communication; 1999 Dutch multi-lake survey Gerben van Geest & Frank Roozen, personal communication). The remaining 104 lakes were sampled during 2005–2006 by standard procedures, described by Kosten et al. (2009) and Kruk et al. (2009). Of the total, 150 lakes were sampled only once, while 61 were sampled at least once every season. Both seasonal and snapshot-sampling strategies were conducted in all the climatic regions and across the whole trophic spectrum. The sampling and sample-analysis protocols were comparable among the sampled lakes and in the published and unpublished sources from where we extracted the information. Most lakes were sampled at random points integrating the water column and covering the whole lake area. Light attenuation in the water column, temperature (Temp, °C) and oxygen profiles were measured in situ at noon. Light attenuation coefficient (K_D, m^{-1}) and water column mixing depth (Z_{mix}, m) were calculated from in situ measurements. Total nitrogen (TN, μg l^{-1}), total phosphorus (TP, μg l^{-1}) and soluble reactive silicate (RSi, μg l^{-1}) were estimated using standard procedures. For zooplankton abundance determination (TZ, org ml^{-1}), 2 l of lake water were filtered through a 50-μm sieve and preserved in a 4% formaldehyde solution (details on sample analysis in Kosten et al., 2009). Phytoplankton samples were fixed in Lugol's solution.

Phytoplankton traits and biovolume

Phytoplankton populations (individuals ml^{-1}) were counted in random fields using the settling technique (Utermöhl, 1958). We examined the samples at multiple magnifications and counted until we reached at least 100 individuals of the most frequent species (Lund et al., 1958). Organisms between 5 and 100 μm were counted at 400×, larger organisms were counted at 200×, and organisms between 5 and 2 μm were counted at 1,000×. We did not include species strongly associated with periphytic communities. Organism dimensions, including MLD were estimated for V and S calculations. The presence of aerotopes, flagella, mucilage and siliceous exoskeletal structures were noted for each relevant organism. Population biovolume (mm^3 l^{-1}) was calculated as the individual volume of the species multiplied by the abundance of individuals. More than 80% of the samples were analyzed by the same group of scientists, using the same identification keys, a common protocol and fluid communication. The species were classified into the seven MBFG and their biovolumes were summed per sample (Table 1).

Statistical analyses

Machine learning

Natural systems generally do not meet statistical assumptions (e.g. normality, homoscedasticity). They are often of high order, non-linear and sometimes show abrupt shifts (Levin, 1992; McGill et al., 2006) which challenges the interpretation of classical statistical techniques (e.g. general lineal models). A number of highly computational statistical methods have recently emerged from the machine learning literature including classification trees and RF (De'ath & Fabricius, 2000; Cutler et al., 2007). These methods can cope with small sample size as compared to the number of variables (small n large p problems), complex interactions, and even with highly correlated predictor variables. See De'ath & Fabricius (2000) and Cutler et al. (2007) for a discussion of these statistical methods in an ecological framework. Despite, these problems are common in phytoplankton ecological studies, the application of the mentioned statistical methods is scarce (Zhao et al., 2008).

Classification and regression trees (CARTs)

In a standard regression situation, we aim to model the response variable based on one or several predictor variables. For example, classical multiple regressions

defines the linear combination of predictors that best explain the response variable in terms of explanatory power. In a different way, a tree is constructed by recursive binary partitioning of the response variable into regions that are increasingly homogeneous (i.e. nodes) until no improvement is possible. This final nodes are called leafs. In regression trees, at each node, the predictor variable that results in the most homogeneous partition of the response variable (measured by the sum of squared errors, SSE) is selected based on an optimization process (Breiman, 2001). This keep on going until no longer reduction of SSE is achieved. Similarly, the process can be performed for classification trees, with the aim of developing rules for assigning current and new observations into the classes using numerical and/or categorical predictors. These methods are easily interpretable and provide simple yes ($>$) or no ($<$) decision trees (De'ath & Fabricius, 2000).

Random forest

RFs are based on the combination of predictions made by many regression or classification trees to a specific data-set. The method selects many (e.g. 1,000) samples with replacement of the data (i.e. bootstrap samples) and fit a tree for a portion of the re-sampled data-sets. In each of the re-sampled data-sets, a small random number of predictor variables becomes available for the binary partitioning at each node until a full grown tree is constructed. For each bootstrap sample the best tree, as defined previously, is used to predict the data not used for the tree construction (i.e. out-of-bag data, OOB). Accuracy and error rates are computed for each observation using the OOB predictions and then averaged over all the observations. The importance of the predictor variables is assessed by randomly permuting the OOB observations, and then the modified OOB data is passed down the tree to obtain new predictions. The difference between the mean squared error (MSE) of the original and permuted OOB data, divided by its standard error, is a measure of the importance of the predictor variable (Cutler et al., 2007).

Validation of the MBFG classification using CART

We evaluated the validity and accuracy of Kruk et al. (2010) classification rules using an unconstrained classification tree. The nine organism traits defined by Kruk et al. (2010) were used as explanatory variables to classify the original 711 species into the seven MBFG. We then evaluated the accuracy of the classification based on the number of organisms well classified as compared to total number of organisms classified.

Habitat template of the MBFG

We used RF in the randomForest package (R, 2011) to evaluate the explained variance and the importance of environmental variables (K_D, RSi, Temp, TN, TP, TZ and Z_{mix}) in explaining each MBFG biovolume. The inorganic nutrient concentrations (soluble reactive phosphorus and dissolved nitrogen forms) were not considered in the analysis because they are highly variable, dependent on phytoplankton consumption and cause–effect relation are hard to disentangle. Therefore, using dissolved nutrients can lead to unclear relationships. The positive values of group's biovolume were $\log_{10}$ transformed. For each MBFG we constructed 1,000 regression random trees to compose the forest. Three environmental variables were randomly selected for each node of the 1,000 constructed trees. For each tree, the MSE on the OOB portion of the data was recorded. Then the same was done after permuting each predictor variable. Importance of a predictor variable was defined as the average over all trees of the difference between the two MSE normalized by the standard deviation of the differences (Cutler et al., 2007).

Environmental thresholds for MBFG dominance

We evaluated the environmental thresholds for dominance of the MBFG by means of a classification tree (CART). The lakes were classified as dominated by a particular group, when the groups accounted for at least 80% of the total phytoplankton biovolume in that lake. Then we used CART to classify lakes with a particular group dominating (I–VII; categorical variable) according to the environmental variables defined previously. After the full tree was constructed, we pruned it back to avoid overfitting. We did so by minimizing the cross-validated error (De'ath & Fabricius, 2000).

Results

Validation of the MBFG classification

CART classification of organisms based on the morphological traits was able to correctly separate ca. 97% of the organisms into the original MBFG. The number of miss-classified organisms was <3% (20 cases over the total 711 cases). The misclassified cases corresponded originally to groups I, III, IV and VII. Most of the misclassified species were related with group I, either classified as members of other groups or erroneously classified as group I.

Environmental template of the MBFG

Average and range for the environmental variables and the seven MBFGs biovolume (I–VII) are shown in Table 1. The average explained variance for the seven groups biovolume based on K_D, RSi, Temp, TN, TP, TZ and Z_{mix} was 42% (Fig. 2). Six of the seven groups showed an important explained variance reaching at least 34.3%. Group VI had the highest explained variance (57.7%) and group II the lowest (10.3%). All variables were selected as important explaining at least one group. Group I was mostly related to TN, followed by TP and K_D. Group II was not well explained by any of the variables, still RSi and Temp had the larger importance. The most important variables explaining group III were TP and K_D. TN and TZ explained group IV. Group V had a high explained variance by Temp, TZ and TN. Group VI had the highest explained variance including Temp, TN and TZ. Finally, the most important variables explaining group VII were RSi and Z_{mix}.

Environmental thresholds for MBFG dominance

From 925 lakes-cases, 147 presented one MBFG reaching >80% of total biovolume. Groups I and II were dominant only in four and three lake-samples, respectively. Group III dominated in 15 lake-samples, group IV in 11, group V in 37, VI in 58 and in 22 lake-samples group VII dominated. We then labelled each lake according to the dominant MBFG and classified them according to the environmental variables. The pruned tree had a complexity parameter of 0.01, a relative error of 0.18 and a cross-validated error of 0.85 (Fig. 3).
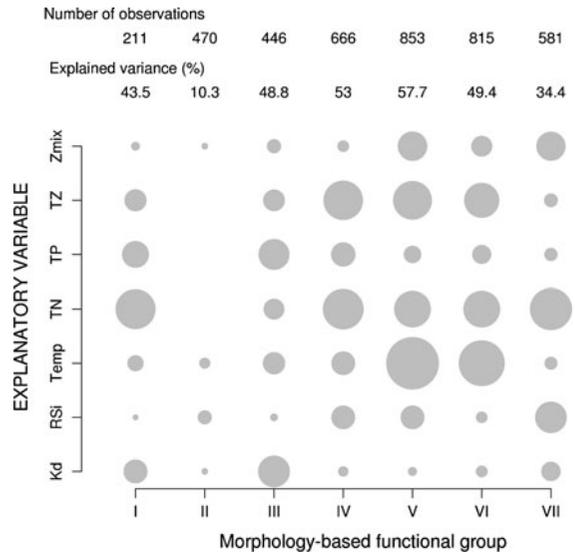

Fig. 2 Results of the RF analysis for the seven MBFGs of phytoplankton (MBFG, I–VII) based on seven environmental explanatory variables. Number of observations and average explained variance (%) for each MBFG are shown. The importance of each environmental variable in explaining each MBFG is proportional to circle diameter. K_D light attenuation coefficient, *RSi* reactive silicate, *Temp* temperature, *TN* total nitrogen, *TP* total phosphorus, *TZ* total zooplankton and Z_{mix} depth of the mixing zone

Lakes dominated by groups III–VII were adequately discriminated according to the environmental variables by the tree (Fig. 3). Light attenuation coefficient (K_D) was the first selected variable with a threshold value of 3.9 m^{-1} in the first node, the root node. The next two selected variables were TN to the left and Temp to the right, with threshold values of 2,800 μg l^{-1} and 24°C, respectively. The further selected variables included Z_{mix} (1.5 m), and reactive silicate (RSi, 1730 μg l^{-1}), and again K_D, Temp and RSi. TP and TZ were not important according to the analysis.

We now describe the environmental conditions driving groups dominance from the left to the right of the constructed tree (Fig. 3). Looking at the root node, to the left, lakes dominated by group III presented elevated K_D (>3.9 m^{-1}) and high TN (>2,800 μg l^{-1}). Lakes dominated by group V showed elevated K_D (>3.9 m^{-1}), TN lower than 2,800 μg l^{-1} and Temps <20°C. From the root node to the right, group VII clearly dominated at K_D <3.9 and Temp between 24 and 29°C. At K_D <3.9 m^{-1} and Temp <24°C, when RSi was below 1,730 μg l^{-1}, lakes were dominated by

Fig. 3 Classification tree showing the environmental variables explaining the dominance (80% over total biovolume in mm^3 l^{-1}) of MBFGs (MBFG: I–VII). In each node, the environmental variable and its threshold value are shown. K_D light attenuation coefficient, *RSi* reactive silicate, *Temp* temperature, *TN* total nitrogen and Z_{mix} depth of the mixing zone. The height of the branch is proportional to the variance explained by that split. At the end of each branch a histogram with bars representing the number of cases where a specific MBFG (I–VII: *left to right*) was found dominant is included. The MBFG with more cases as dominant is shown below each histogram
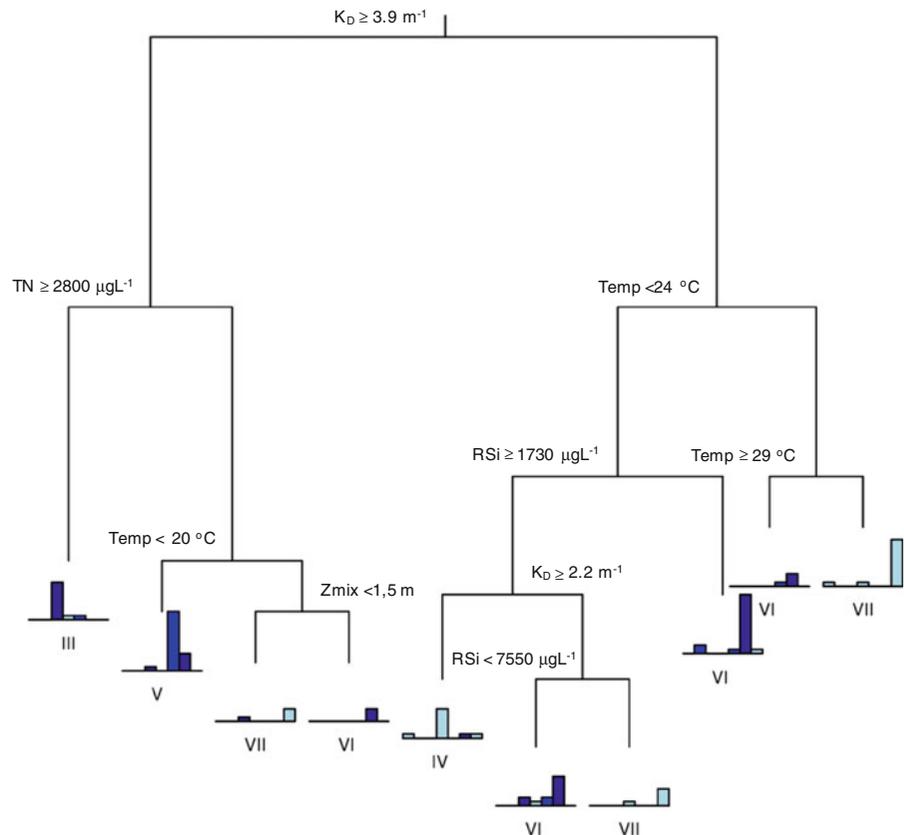
group VI. Similarly, lakes with dominance of group IV presented K_D between 2.2 and 3.9 m^{-1} and Temp <24 with RSi above 1,730 μg l^{-1}. In the rest of the leafs, there was a mix of lakes dominated by different groups that could not be separated according to the included environmental variables, which is the case for some subset of lakes dominated by group VI. On the contrary, lakes dominated by groups III–V where rather homogenous in their environmental settings.

Discussion

Differences in morphological traits among phytoplankton MBFG reflected well their different habitat templates. Main ecological processes including resources acquisition (light and nutrients), evasion of loss processes (mixing and zooplankton), as well as Temp, were represented in the results. The application of the MBFG approach was useful in reducing the diversity of species to a diversity of functions (sensu Dray & Legendre, 2008). This reduction of complexity provides an efficient tool to explore the effects of

environmental changes on phytoplankton independently from geographical location and specific composition.

Though we used a huge database covering more than 900 lake-cases we obtained a relative high mean explained variance of MBFG biovolume (34%) in comparison with other ecological studies. Møller & Jennions (2002) found that biological studies, even experimental ones, often only explain a very small amount of variance (R^2: 0.3–29%). Therefore, ecological models with explained variances >30% might be considered good predictive tools.

RFs do not assume linear relationships among variables, thus allowing to correctly addressing the often nonlinear interactions occurring in phytoplankton communities (Zhao et al., 2008). However, RFs are black-box models, in which the fitted relationships cannot be written in the form of an equation. RF are data dependent and do not inform if the effect is positive or negative. This precludes an easy transference to managers or scientists (Cutler et al., 2007). To increase the applicability of the model, we constructed a classification tree, which allowed us to know the

specific environmental thresholds determining the dominance of a group in a lake. We successfully separated most of the group-dominated lakes according to the environmental variables. The identification of environmental thresholds is needed to understand ecosystems responses to environmental changes (Scheffer et al., 2001a; Bayley et al., 2007), including climate change and trophic interactions (Scheffer et al., 2001b) and forecast phytoplankton community changes (Roelke, 2000; Carpenter et al., 2009). Also the identification of thresholds is of paramount importance for the implementation of adequate management programs and water quality guidelines (i.e. Chorus & Bartram, 1999). TP and TZ were not selected in the CART to explain MBFG dominance which seems to be counterintuitive. This might have been the result of either a redundancy of the variables (i.e. TP and TN) or that the variable is particularly not important as a threshold for the 80% dominance (i.e. TZ). Below we describe in detail each group's habitat template and its relation with organism's morphological traits.

Group I: small organisms with high S/V

TP and TN were the most important variables explaining the distribution of group I biovolume. According to their morphology these organisms are r-selected (Pianka, 1970), have effective resources acquisition and high specific growth rate in resource-saturated and limited environments (Raven, 1998; Callieri & Stockner, 2002; Kruk et al., 2010). Group I biovolume might increase with total nutrients, while it's relative importance over total phytoplankton declines with increasing total nutrients (Raven, 1998; Bell & Kalff, 2001). Still, these organisms can be dominant in a wide variety of trophic conditions (Callieri & Stockner, 2002; Izaguirre et al., 2003) for example in flushed ecosystems or during transitional stages (Carrick et al., 1993; Reynolds, 2006; Kruk et al., 2010). Our data-set consisted on a low number of lakes dominated by group I which precludes further analysis of the lakes allocating their dominance.

Group II: small flagellated organisms with siliceous exoskeletal structures

Based on their morphology, we expected moderate resources gathering ability, moderate vulnerability to

consumption and low to moderate sinking losses (Kruk et al., 2010). In general this group representatives have low optimum Temps (Kim et al., 2009; Jansson et al., 2010) and are more important in cold oligotrophic conditions (Kristlansen & Takahashi, 1982; Izaguirre et al., 2003), as well as in mesotrophic clear-water plant dominated lakes (Reynolds et al., 2002). Low explained variance can be a consequence of poor representation of lakes with this group. Also other environmental variables as is the case of pH and conductivity might have improved the explained variance (Siver & Hamer, 1989). In addition to the morphological traits originally considered, the production of resistant propagules and the facultative mixotrophy might improve their description (Sandgren, 1988).

Group III: large filaments with aerotopes

Light attenuation coefficient and TP were the main drivers of group III biovolume. Given their morphology, the species in this group may be mostly characterized as K-selected (Pianka, 1970) with high to moderate saturating nutrient concentration, and low losses due to consumption and sinking (Kruk et al., 2010). These features along with high S/V confer a greater tolerance to limiting light conditions (Naselli-Flores et al., 2007) and result in the success of these organisms in low-light high-trophic status environments (Padisák & Reynolds, 1998; Kruk et al., 2002; Reynolds et al., 2002; Bonilla et al., 2011). Furthermore, dominance of this group occurred mostly at very high K_D (>3.9 m^{-1}) and TN (>2,800 µg l^{-1}) values. This is in accordance with other studies implying that members of this group can succeed in turbid, eutrophic lakes as originally proposed for Oscillatoriales (Scheffer et al., 1997) and discussed also for Nostocales (Bonilla et al., 2011).

Group IV: organisms of medium size lacking specialized traits

The most important variables explaining group IV were TN and TZ. The small size and high quality as food (Sterner & Elser, 2002) of many of the species in this group (e.g. *Chlorella* sp.) make them liable to high grazing losses. The expected moderate tolerances to limiting resources (Kruk et al., 2010), including nutrients and light, might result in a positive relation

of biovolume with lower TN concentrations, and higher light in the water column (K_D in Fig. 3). Group IV dominated at the lowest values of light attenuation coefficient in the water column ($K_D < 2.2$ m^{-1}) and Temps <24°C. This combination of variables: low nutrient, high zooplankton abundance, low light attenuation might indicate the success of this group under good water quality conditions or during transitional ecosystem stages (Reynolds et al., 2002).

Group V: unicellular flagellates of medium to large size

Temp, TZ and TN explained the distribution of group V biovolume among lakes. According to their morphology, these organisms would have moderate aptitude to resources gathering and high to moderate vulnerability to consumption (Kruk et al., 2010). Motility that permits effective nutrient foraging in conjunction with the production of cysts might increase tolerance of lower nutrient conditions (Reynolds et al., 2002). In addition, the capacity of some species to benefit from mixotrophy and phagotrophy implies a means of tolerating conditions of reduced availability of dissolved nutrients and limiting light conditions (Graham & Wilcox, 2000). Their relatively high maximum linear dimension and the presence of flagella may give substantial tolerance to grazing by all but the specialised zooplankton (Reynolds, 1997). These organisms achieved dominance at low light ($K_D > 3.9$) and TN (<2,800 µg l^{-1}). Their moderate size and surface to volume ratio and the possession of flagella reduces high sinking losses therefore tolerates large mixings zones with lower light ($Z_{mix} > 1.5$ m) that can be characteristic of deeper meso to eutrophic lakes. These conditions are probably achieved not in the warmer seasons (Temp < 20°C) (Reynolds et al., 2002).

Group VI: non-flagellated organisms with siliceous exoskeletons

Temp, TZ and TN were the variables better explaining group VI biovolume. Their morphology indicates moderate resources gathering properties, with silicate requirements and moderate vulnerability to consumption (Kruk et al., 2010). Siliceous walls increase sinking but have advantages against certain types of grazers and viral infections (Smetacek, 2001).

However, these organisms can suffer more than other groups from fungal infections, especially chytrids (Ibelings et al., 2004) and a wide variety of protists to crustacean zooplankton can successfully feed on diatoms (Hamm et al., 2003). Despite this group shows a wide range of responses to trophic status (Reynolds et al., 2002), its members are better competitors at lower temperatures (Tilman, 1982; Tilman et al., 1986). Owing to their high cell density and lack of motility, these organisms are rapidly excluded from illuminated waters, explaining their dominance in lakes with lower light attenuation than other groups ($K_D < 3.9$ m^{-1}). The obligate presence of a siliceous wall is probably the main constraining trait of these species (Kruk et al., 2010). Coincidently this group dominated in a close relationship with RSi values. High biovolume of group VI species might diminish RSi concentration in lakes. However, cause and effect cannot be disentangled from studies based on correlations.

Group VII: large mucilaginous colonies

Total RSi and mixing zone depth were the main environmental variables explaining the biovolume of group VII. RSi is highly correlated to the effect of the catchment area and nutrient access from it (Conley, 2002; Kruk et al., 2009). Large size and volume, and low surface to volume ratio, should tend to make species sensitive to low resource supply (Kruk et al., 2010). Therefore, we would expect an increase of this group in high-trophic status lakes, with larger catchment areas and probably deeper than shallow lakes. The presence of mucilage, along with lipids and aerotopes in the larger colonies, gives controllable buoyant properties (Walsby & Reynolds, 1980). However, water mixing can affect them negatively disrupting scums in the water surface (Chorus & Bartram, 1999). Dominance was attained with $K_D < 4.0$ m^{-1} and Temp between 24 and 28°C. The preference for high temperatures might indicate stratified and stable water column and is consistent with dominance this group in several tropical shallow lakes (Ganf & Viner, 1973; Huisman et al., 2005).

Here, we were able to describe the habitat template of most MBFG in terms of biovolume distribution among lakes. The conditions for 80% dominance of groups III, IV and V were also well discriminated by the environmental variables and constituted

homogeneous groups. However, the discriminatory power of the conditions for dominance of groups I and II was low, and groups VI and VII dominated in various environmental conditions. Deep, oligotrophic, high altitude lakes and lakes from extreme environments (i.e. polar and salty lakes) should be included to improve the results presented here. Also the validity of this approach to explain the temporal replacement of MBFG needs to be further analysed (but see Pacheco et al., 2010; Segura et al., 2010). Assuming community structure is shaped by similar processes, it would be a challenge to apply the MBFG classification to phytoplankton communities from coastal open waters and the ocean.

Probably, if we aim at predicting specific nuisance species blooms, we will have to probe deeper into the mechanisms ruling the dynamics of algal communities. Indeed, while it is encouraging that MBFGs appear to be relatively predictable (Kruk et al., 2011), it still makes a difference which species will dominate within the groups, as features such as toxicity and edibility may still differ quite a bit within groups (Chorus & Bartram, 1999; Sterner & Elser, 2002; Huisman et al., 2005). In this vein, the MBFG approach as we have presented it has some weaknesses and does not invalidate other approaches (Carpenter et al., 1993; Reynolds et al., 2002; Le Quéré et al., 2005; Salmaso & Padisák, 2007; Padisák et al., 2009), as does not rule out the classical taxonomical approach. However, morphological traits as a surrogate for species taxonomy allow researchers to produce more general models and generalize to ecosystems with very different taxonomic composition (Keddy, 1992; McGill et al., 2006; Dray & Legendre, 2008). This will likely bring us closer to the goal of predicting and managing nuisance algal blooms and to arrive to a verifiable quantitative method of describing community structure and change with other more "ecological" purposes.

References

Arrigo, K. R., 2005. Marine microorganisms and global nutrient cycles. Nature 437: 349–355.

Bayley, S. E., I. F. Creed, G. Z. Sass & A. S. Wong, 2007. Frequent regime shifts in trophic states in shallow lakes on the Boreal Plain: alternative "unstable" states? Limnology and Oceanography 52: 2002–2012.

Bell, T. & J. Kalff, 2001. The contribution of picophytoplankton in marine and freshwater systems of different trophic status and depth. Limnology and Oceanography 46: 1243–1248.

Bonilla, S., L. Aubriot, M. C. S. Soares, M. González-Piana, A. Fabre, V. L. M. Huszar, M. Lürling, D. Antoniades, J. Padisák & C. Kruk, 2011. What drives the distribution of the bloom forming cyanobacteria *Planktothrix agardhii* and *Cylindrospermopsis raciborskii*? FEMS Microbiology Ecology 79(3): 594–607.

Breiman, L., 2001. Random forests. Machine Learning 45: 15–32.

Callieri, C. & J. Stockner, 2002. Freshwater autotrophic picoplankton: a review. Journal of Limnology 61: 1–14.

Carpenter, S. R., J. A. Morrice, J. J. Elser, A. S. Amand & N. A. Mackay, 1993. Phytoplankton community dynamics. In Carpenter, S. R. & J. F. Kitchell (eds), The Trophic Cascade in Lakes. Cambridge University Press, Cambridge: 189–209.

Carpenter, S. R., W. A. Brock, J. J. Cole & M. L. Pace, 2009. Leading indicators of phytoplankton transitions caused by resource competition. Theoretical Ecology 2: 139–148.

Carrick, H. J., F. J. Aldridge & C. L. Schelske, 1993. Wind induces phytoplankton biomass and composition in a shallow, productive lake. Limnology and Oceanography 38: 1179–1192.

Chorus, I. & J. Bartram, 1999. Toxic cyanobacteria in water. A guide to their public health consequences, monitoring and management. Chapman & Hall, London.

Conley, D. J., 2002. Terrestrial ecosystems and the global biogeochemical silica cycle. Global Biogeochemical Cycles 16: 1121.

Cutler, D. R., T. C. Edwards Jr, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson & J. J. Lawler, 2007. Random forests for classification in ecology. Ecology 88: 2783–2792.

De León, L., 2000. Composición y dinámica de la comunidad fitoplanctónica de un embalse subtropical (Salto Grande, Uruguay-Argentina). Tesis de Maestría, Universidad de Concepción, Concepción.

De'ath, G. & K. E. Fabricius, 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81: 3178–3192.

Dray, S. & P. Legendre, 2008. Testing the species traits–environment relationships: the fourth-corner problem revisited. Ecology 89: 3400–3412.

Falkowski, P. G., E. A. Laws, R. T. Barber & J. W. Murray, 2003. Phytoplankton and their role in primary, new, and export production. In Fasham, M. J. R. (ed.), Ocean Biogeochemistry: The Role of the Ocean Carbon Cycle in Global Change. Springer, Berlin: 99–121.

Follows, M. J., S. Dutkiewicz, S. Grant & S. W. Chisholm, 2007. Emergent biogeography of microbial communities in a model ocean. Science 315: 1843–1846.

Ganf, G. & A. Viner, 1973. Ecological stability in a shallow equatorial lake (L. George, Uganda). Proceedings of the Royal Society B 184: 321–346.

Graham, L. E. & L. W. Wilcox, 2000. Algae. Prentice-Hall, Upper Saddle River.

Hamm, C. E., R. Merkel, O. Springer, P. Jurkojc, C. Maler, K. Prechtel & V. Smetacek, 2003. Architecture and material properties of diatom shells provide effective mechanical protection. Nature 421: 841–843.

Hillebrand, H., C. Dürselen, D. Kirschtel, T. Zohary & U. Pollingher, 1999. Biovolume calculation for pelagic and benthic microalgae. Journal of Phycology 35: 403–424.

Huisman, J., H. C. P. Matthijs & P. M. Visser, 2005. Harmful Cyanobacteria. Springer, Dordrecht.

Ibelings, B. W., A. De Bruin, M. Kagami, M. Rijkeboer, M. Brehm & E. Van Donk, 2004. Review of host parasite interactions between freshwater phytoplankton and chytrid fungi (chytridiomycota). Journal of Phycology 40: 437–453.

Izaguirre, I., L. Allende & M. C. Marinone, 2003. Comparative study of the planktonic communities of three lakes of contrasting trophic status at Hope Bay (Antarctic Peninsula). Journal of Plankton Research 25: 1079–1097.

Jansson, M., A. Jonsson, A. Andersson & J. N. Karlsson, 2010. Biomass and structure of planktonic communities along an air temperature gradient in subarctic Sweden. Freshwater Biology 55: 691–700.

Keddy, P. A., 1992. A pragmatic approach to functional ecology. Functional Ecology 6: 621–626.

Kim, H. J., L. K. Lee & S. H. Kim, 2009. Effect of nutrients and light intensity on growth of Mallomonas caudata (Synurophyceae). Nordic Journal of Botany 27: 516–522.

Kosten, S., G. Lacerot, E. Jeppesen, D. Da Motta Marques, E. H. Van Nes, N. Mazzeo & M. Scheffer, 2009. Effects of submerged vegetation on water clarity across climates. Ecosystems 12: 1117–1129.

Kristlansen, J. & E. Takahashi, 1982. Chrysophyceae: introduction and bibliography. In Rosowski, J. & B. Parker (eds), Selected Papers in Phycology 2. Phycological Society of America, Brandenburg: 698–704.

Kruk, C., N. Mazzeo, G. Lacerot & C. S. Reynolds, 2002. Classification schemes for phytoplankton: a local validation of a functional approach to the analysis of species temporal replacement. Journal of Plankton Research 24: 901–912.

Kruk, C., L. Rodríguez-Gallego, M. Meerhoff, F. Quintans, G. Lacerot, N. Mazzeo, F. Scasso, J. Paggi, E. T. H. M. Peeters & M. Scheffer, 2009. Determinants of biodiversity in subtropical shallow lakes (Atlantic coast, Uruguay). Freshwater Biology 54: 2628–2641.

Kruk, C., V. L. M. Huszar, E. T. H. M. Peeters, S. Bonilla, L. Costa, M. Lürling, C. S. Reynolds & M. Scheffer, 2010. A morphological classification capturing functional variation in phytoplankton. Freshwater Biology 55: 614–627.

Kruk, C., E. T. H. M. Peeters, E. H. Van Nes, V. L. M. Huszar, L. S. Costa & M. Scheffer, 2011. Phytoplankton community composition can be predicted best in terms of morphological groups. Limnology and Oceanography 56: 110–118.

Le Quéré, C., S. P. Harrison, I. C. Prentice, E. T. Buitenhuis, O. Aumont, L. Bopp, H. Claustre, L. C. Da Cunha, R. Geider, X. Giraud, C. Klaas, K. E. Kohfeld, L. Legendre, M. Manizza, T. Platt, R. B. Rivkin, S. Sathyendranath, J. Uitz, A. J. Watson & D. Wolf-Gladrow, 2005. Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. Global Change Biology 11: 2016–2040.

Levin, S. A., 1992. The problem of pattern and scale in ecology. Ecology 73(1943): 1967.

Lewis, W. M. J., 1976. Surface/volume ratio: implications for phytoplankton morphology. Science 192: 885–887.

Lund, J. W. G., C. Kipling & E. D. Le Cren, 1958. The inverted microscope method of estimating algal numbers and the statistical basis of estimations by counting. Hydrobiologia 11: 143–170.

Margalef, R., 1978. Life-forms of phytoplankton as survival alternatives in an unstable environment. Oceanologica Acta 1: 493–509.

Mazzeo, N., L. Rodríguez-Gallego, C. Kruk, M. Meerhoff, J. Gorga, G. Lacerot, F. Quintans, M. Loureiro, D. Larrea & F. García-Rodríguez, 2003. Effects of Egeria densa Planch. beds on a shallow lake without piscivorous fish. Hydrobiologia 506–509: 591–602.

McGill, B., B. J. Enquist, E. Weiher & M. Westoby, 2006. Rebuilding community ecology from functional traits. Trends in Ecology & Evolution 21: 178–185.

Møller, A. P. & M. D. Jennions, 2002. How much variance can be explained by ecologists and evolutionary biologists. Oecologia 132: 492–500.

Naselli-Flores, L., J. Padisák & M. Albay, 2007. Shape and size in phytoplankton ecology: do they matter? Hydrobiologia 578: 157–161.

Pacheco, J. P., C. Iglesias, M. Meerhoff, C. Fosalba, G. Goyenola, F. Teixeira-De Mello, S. García, M. Gelós & F. García-Rodríguez, 2010. Phytoplankton community structure in five subtropical shallow lakes with different trophic status (Uruguay): a morphology based approach. Hydrobiologia 646: 187–197.

Padisák, J. & C. S. Reynolds, 1998. Selection of phytoplankton associations in Lake Balaton, Hungary, in response to eutrophication and restoration measures, with special reference to the cyanoprokaryotes. Hydrobiologia 384: 41–53.

Padisák, J., L. O. Crossetti & L. Naselli-Flores, 2009. Use and misuse in the application of the phytoplankton functional classification: a critical review with updates. Hydrobiologia 621: 1–19.

Peretyatko, A., S. Teissier, J. J. Symoens & L. Triest, 2007. Phytoplankton biomass and environmental factors over a gradient of clear to turbid peri-urban ponds. Aquatic Conservation: Marine and Freshwater Ecosystems 17: 584–601.

Pianka, E. R., 1970. On r and K selection. American Naturalist 104: 592–597.

R, 2011. Development Core Team R: a language and environment for statistical computing, R Foundation for Statistical Computing, Vienna.

Raven, J. A., 1998. Small is beautiful: the picophytoplankton. Functional Ecology 12: 503–513.

Reynolds, C. S., 1984a. Phytoplankton periodicity: the interaction of form, function and environmental variability. Freshwater Biology 14: 111–142.

Reynolds, C. S., 1984b. The Ecology of Freshwater Phytoplankton. Cambridge University Press, Cambridge.

Reynolds, C. S., 1988. Functional morphology and the adaptive strategies of freshwater phytoplankton. In Sandgren, C. D. (ed.), Growth and Reproductive Strategies of Freshwater Phytoplankton. Cambridge University Press, New York: 388–433.

Reynolds, C. S., 1997. Vegetation Process in the pelagic: a model for ecosystem theory. Excellence in Ecology. Ecology Institute, Luhe.

Reynolds, C. S., 2006. Ecology of Phytoplankton. Cambridge University Press, Cambridge.

Reynolds, C., V. Huszar, C. Kruk, L. Naselli-Flores & S. Melo, 2002. Towards a functional classification of the freshwater phytoplankton. Journal of Plankton Research 24: 417–428.

Roelke, D. L., 2000. Copepod food-quality threshold as a mechanism influencing phytoplankton succession and accumulation of biomass, and secondary productivity: a modeling study with management implications. Ecological Modelling 134: 245–274.

Salmaso, N. & J. Padisák, 2007. Morpho-functional groups and phytoplankton development in two deep lakes (Lake Garda, Italy and Lake Stechlin, Germany). Hydrobiologia 578: 97–112.

Sandgren, C. D. (ed.), 1988. Growth and Reproductive Strategies of Freshwater Phytoplankton. Cambridge University Press, Cambridge: 442.

Scheffer, M., S. Rinaldi, A. Gragnani, L. Mur & E. H. Van Nes, 1997. On the dominance of filamentous Cyanobacteria in shallow, turbid lakes. Ecology 78: 272–282.

Scheffer, M., S. Carpenter, J. A. Frey, C. Folke & B. Walker, 2001a. Catastrophic shifts in ecosystems. Nature 413: 591–596.

Scheffer, M., D. Straile, E. H. Van Nes & H. Hosper, 2001b. Climatic warming causes regime shifts in lake food webs. Limnology and Oceanography 46: 1780–1783.

Scheffer, M., S. Rinaldi, J. Huisman & F. J. Weissing, 2003. Why plankton communities have no equilibrium: solutions to the paradox. Hydrobiologia 491: 9–18.

Segura, A., C. Kruk, D. Calliari & H. Fort, 2010. Trait-based approach disentangles core features of phytoplankton succession. Morphology captures function in phytoplankton. A large-scale analysis of phytoplankton communities in relation to their environment. Ph.D. Thesis, pp. 75–90. University of Wageningen, Wageningen, The Netherlands.

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

Siver, P. A. & J. S. Hamer, 1989. Multivariate statistical analysis of the factors controlling the distribution of scaled chrysophytes. Limnology and Oceanography 34: 368–381.

Smetacek, V., 2001. A watery arms race. Nature 411: 745.

Sommer, U., 1989. Plankton Ecology: Succession in Plankton Communities. Springer, Berlin.

Southwood, T. R. E., 1977. Habitat, the templet for ecological strategies? Journal of Animal Ecology 46: 337–365.

Southwood, T. R. E., 1988. Tactics, strategies and templets. Oikos 52: 3–18.

Sterner, R. W. & J. J. Elser, 2002. Ecological Stoichiometry: The Biology of Elements from Molecules to the Biosphere. Princeton University Press, Princeton, NJ.

Tilman, D., 1982. Resource Competition and Community Structure. Princeton University Press, Princeton, NJ.

Tilman, D., R. Kiesling, R. Sterner & S. Tilman, 1986. Green, blue-green and diatom algae: taxonomic differences in competitive ability for phosphorus, silicon, and nitrogen. Archiv fur Hydrobiologie 106: 473–485.

Utermöhl, H., 1958. Zur Vervollkommnung der quantitativen phytoplankton-methodik. Mitteilungen Internationale Vereinigung für Theoretische und Angewandte Limnologie 9: 1–38.

Vollenweider, R. A., 1976. Advances in defining critical loading levels for phosphorus in lake eutrophication. Memorie dell' Istituto Italiano di Idrobiologia 33: 53–83.

Walsby, A. E. & C. S. Reynolds, 1980. Sinking and floating. In Morris, I. G. (ed.), The Physiological Ecology of Phytoplankton. Blackwell Scientific Publications, Oxford: 371–412.

Zhao, J., M. Ramin, V. Cheng & G. B. Arhonditsis, 2008. Competition patterns among phytoplankton functional groups: how useful are the complex mathematical models? Acta Oecologica 33: 324–344.