Factors controlling the community structure of picoplankton in contrasting marine environments

Jose Luis Otero-Ferrer¹, Pedro Cermeño², Antonio Bode⁶, Bieito Fernández-Castro¹,³, Josep M. Gasol²,⁵, Xosé Anxelu G. Morán⁴, Emilio Marañon¹, Víctor Moreira-Coello¹, Marta M. Varela⁶, Marina Villamaña¹, and Beatriz Mouriño-Carballido¹

¹Departamento de Ecoloxía e Bioloxía Animal, Universidade de Vigo, Vigo, Spain
²Institut de Ciències del Mar, Consejo Superior de Investigaciones Científicas, Barcelona, Spain
³Departamento de Oceanografía, Instituto de investigacións Mariñas (IIM-CSIC), Vigo, Spain
⁴King Abdullah University of Science and Technology (KAUST), Read Sea Research Center, Biological and Environmental Sciences and Engineering Division, Thuwal, Saudi Arabia
⁵Centre for Marine Ecosystem Research, School of Sciences, Edith Cowan University, WA, Perth, Australia
⁶Centro Oceanográfico de A Coruña, Instituto Español de Oceanografía (IEO), A Coruña, Spain

Correspondence: Jose Luis Otero-Ferrer (jootero@uvigo.es)

Received: 27 April 2018 – Discussion started: 4 June 2018
Revised: 4 October 2018 – Accepted: 10 October 2018 – Published: 26 October 2018

Abstract. The effect of inorganic nutrients on planktonic assemblages has traditionally relied on concentrations rather than estimates of nutrient supply. We combined a novel dataset of hydrographic properties, turbulent mixing, nutrient concentration, and picoplankton community composition with the aims of (i) quantifying the role of temperature, light, and nitrate fluxes as factors controlling the distribution of autotrophic and heterotrophic picoplankton subgroups, as determined by flow cytometry, and (ii) describing the ecological niches of the various components of the picoplankton community. Data were collected at 97 stations in the Atlantic Ocean, including tropical and subtropical open-ocean waters, the northwestern Mediterranean Sea, and the Galician coastal upwelling system of the northwest Iberian Peninsula. A generalized additive model (GAM) approach was used to predict depth-integrated biomass of each picoplankton subgroup based on three niche predictors: sea surface temperature, averaged daily surface irradiance, and the transport of nitrate into the euphotic zone, through both diffusion and advection. In addition, niche overlap among different picoplankton subgroups was computed using nonparametric kernel density functions. Temperature and nitrate supply were more relevant than light in predicting the biomass of most picoplankton subgroups, except for Prochlorococcus and low-nucleic-acid (LNA) prokaryotes, for which irradiance also played a significant role. Nitrate supply was the only factor that allowed the distinction among the ecological niches of all autotrophic and heterotrophic picoplankton subgroups. Prochlorococcus and LNA prokaryotes were more abundant in warmer waters (> 20°C) where the nitrate fluxes were low, whereas Synechococcus and high-nucleic-acid (HNA) prokaryotes prevailed mainly in cooler environments characterized by intermediate or high levels of nitrate supply. Finally, the niche of picoeukaryotes was defined by low temperatures and high nitrate supply. These results support the key role of nitrate supply, as it not only promotes the growth of large phytoplankton, but it also controls the structure of marine picoplankton communities.

1 Introduction

Picoplankton, including archaea, bacteria, and picoeukaryotes, are the smallest (cell diameter < 2 µm) and most abundant organisms in marine ecosystems. Photosynthetic picoplankton often dominate marine phytoplankton biomass and primary production in oligotrophic tropical and subtropical regions (Chisholm, 1992), whereas they are typically a minor component in nutrient-replete coastal environments,
usually dominated by large-sized plankton species (Finkel et al., 2010; Marañón, 2015). However, due to the large temporal and spatial variability in the structure and composition of the microbial community in shelf seas (Sherr et al., 2005), picoplankton, together with nanoplankton, can dominate the microbial community under certain conditions (Morán, 2007; Espinoza-González et al., 2012). In addition, picoplankton contributes overwhelmingly to the recycling of organic matter (Azam et al., 1983; Fenchel, 2008), and previous studies suggest that photosynthetic picoplankton could also play a role in the export of carbon to the deep ocean (Richardson and Jackson, 2007; Lomas and Moran, 2011; Guidi et al., 2015). As a result, picoplankton is considered a key component of the current carbon cycle and likely more important in future climate warming scenarios (Laufkötter et al., 2016). When analyzed using flow cytometric techniques, two genera of picocyanobacteria (Prochlorococcus and Synechococcus), one or two subgroups of autotrophic picoeukaryotes of different sizes (small and large), and two subgroups of heterotrophic prokaryotes, based on their high (HNA) or low nucleic acid (LNA) content, can be distinguished (Gasol and del Giorgio, 2000; Marie and Partensky, 2006). Although closely related phylogenetically, Synechococcus and Prochlorococcus exhibit distinct physiological traits (Moore et al., 1995), divergent evolutionary strategies (Scanlan and West, 2002), and disparate geographic distributions (Zubkov et al., 2000). Prochlorococcus tend to be restricted to relatively warm (above 15 °C) and nutrient-poor waters, extending from the surface down to 150 m, along the 40° N–40° S latitudinal band (Partensky et al., 1999b; Johnson et al., 2006). Synechococcus exhibit a wider geographic and thermal distribution, including high-nutrient waters and occasionally reaching polar latitudes (Paulsen et al., 2016); their vertical distribution is shallower than that of Prochlorococcus (Partensky et al., 1999a; Li, 2002). The contribution of picoeukaryotes to picoplankton biomass is generally smaller than the contribution of picocyanobacteria (Zubkov et al., 2000; Buitenhuis et al., 2012), except in coastal regions where their contribution usually increases (Grob et al., 2007). In general, LNA prokaryotes dominate heterotrophic prokaryotic biomass in the oligotrophic open ocean, whereas HNA cells dominate in coastal regions (Li et al., 1995; Bouvier et al., 2007). These contrasting spatial distributions suggest that the picoplankton subgroups occupy differential ecological niches or, according to the classical definition proposed by Hutchinson (1957), distinct multidimensional hyper-volumes of environmental factors in which viable populations develop. By describing the overlaps of environmental factors, realized niche partitioning can be defined, and the factors controlling the distribution of picoplankton subgroups can be identified. However, despite decades of experimental and field observations, the relative importance of the factors driving the variability in the growth and the spatial distribution of picoplankton subgroups remains largely unknown.

Aside from the effect of trophic controls, the distribution of microbial plankton is primarily determined by seawater temperature, light, and nutrients (Li, 2009, 2007; Barton et al., 2015). Quantifying their relative influence on the spatial and temporal distribution of the different picoplankton subgroups is complicated by the fact that the abovementioned factors are often correlated in the ocean (Finkel et al., 2010). This shortcoming can be circumvented by using experimental approaches in the laboratory, where the influence of each independent factor is isolated. Alternatively, it can be approached by combining large datasets of hydrographic and biological observations collected from contrasting marine environments, which allow us to characterize the suite of variables that best define the organism’s ecological niches.

In order to study the significance of temperature and nutrient concentrations in determining the contribution of pico-phytoplankton to total phytoplankton biomass and production, Agawin et al. (2000) reviewed the available literature from oceanic and coastal estuarine areas. Although the number of observations for which both temperature and nutrient concentration were available was too small to statistically separate their effects, these authors hypothesized that the dominance of picoplankton in warm, oligotrophic waters was due to differences between picophytoplankton and larger cells in their capacity to use nutrients and in their intrinsic growth rate. Bouman et al. (2011) investigated how vertical stratification controls the community structure of picophytoplankton in subtropical regions. According to their results, photosynthetic picoeukaryotes dominate in weakly stratified waters, whereas in strongly stratified waters, Prochlorococcus cyanobacteria are prevalent. More recently, Flombaum et al. (2013), using a compilation of flow cytometry data from all major ocean regions, concluded that Prochlorococcus and Synechococcus abundance distributions were controlled by temperature and photosynthetically active radiation (PAR, 400–700 nm), discarding the role of nitrate concentration. However, in tropical and subtropical domains, the most well-represented regions in the study of Flombaum et al. (2013), surface nitrate is almost depleted and the variability in its concentration can be widely disconnected from changes in its availability for phytoplankton, which depends more on the supply from deeper waters by turbulent diffusion (Mouriño-Carballido et al., 2016). It is also believed that fine-scale turbulence can enhance the nutrient uptake and subsequent growth of larger phytoplankton (Lazier and Mann, 1989; Karp-Boss et al., 1996; Guasto et al., 2012), especially in regions with low nutrient levels and strong grazing pressure (Barton et al., 2014).

As far as we know, only one study has previously used estimates of nitrate availability, derived from observations of microstructure turbulence, to investigate the role of nutrient availability in controlling the composition of picoplankton communities (Mouriño-Carballido et al., 2016). These authors, using local data from the northwestern Mediterranean Sea, found that different autotrophic picophytoplank-
ton subgroups exhibit contrasting responses to nitrate supply and that as a result the ratio of prokaryotic to picoeukaryotic photoautotrophic biomass decreased with increasing nitrate supply. However, whether these patterns are general and widespread in the ocean remains largely uncertain, given that no concomitant datasets including measurements of turbulent diffusion nutrient flux and picoplankton subgroup structure have been available to date.

Here we extend the analysis described in Mouriño-Carballido et al. (2016) by combining a dataset of picoplankton community composition, hydrographic properties, turbulent mixing, and inorganic nutrient concentrations collected at a total of 97 stations. Observations were made in contrasting environments of the Atlantic Ocean in order to quantify the role of temperature, light, and nitrate availability in the composition of the picoplankton community and to describe the ecological niches of each picoplankton subgroup.

2 Materials and methods

This study includes data collected at 97 stations from three contrasting environments covering the tropical and subtropical Atlantic Ocean (T), the northwestern Mediterranean Sea (M), and the Galician coastal upwelling ecosystem (G), between October 2006 and December 2015 (Table 1 and Fig. 1). Two cruises (CARPOS October–November 2006 and TRYNITROP April–May 2008) sampled 26 stations located in the tropical and subtropical Atlantic Ocean. Three other cruises carried out in the Mediterranean Sea (FAMOSO1 March 2009, FAMOSO2 April–May 2009, and FAMOSO3 September 2009) sampled 19 stations during three contrasting hydrographic conditions, covering from winter mixing to summer stratification. Finally, 52 stations were sampled in the Galician coastal upwelling ecosystem during the HERCULES1 July 2010, HERCULES2 September 2011, HERCULES3 July 2012, DISTRAL February–November 2012, ASIMUTH June 2013, CHAOS August 2013, and NICANOR February 2014–December 2015 cruises. Additional information about the sampling design of these cruises is presented in Aranguren-Gassis et al. (2011, CARPOS), Mouriño-Carballido et al. (2011, TRYNITROP), Mouriño-Carballido et al. (2016, FAMOSO), Cermeño et al. (2016, DISTRAL), Villamaña et al. (2017, CHAOS), and Moreira-Coello et al. (2017, NICANOR).

At each station, information about hydrographic properties, turbulent mixing, nitrate concentration, and picoplankton community composition was collected. Light conditions for each sampling station were considered to be the 5-day averaged daily surface PAR obtained from satellite data (http://globcolour.info; last access: 18 October 2018). Light attenuation coefficients were obtained from vertical profiles of PAR estimated with LI-COR sensors using the Beer–Lambert law equation (Kirk, 1994). Depth of the euphotic layer was calculated as the depth at which PAR was 1% of its surface value. For those cruises for which PAR profiles were not available (ASIMUTH, CHAOS, and NICANOR), which sampled stations in the outer part of the Galician rias, the depth of the photic layer was calculated by con-
considering light attenuation coefficients derived from surface chlorophyll \( a \) data, following the algorithms proposed by Morel et al. (2007) for Case-1 waters \( \log_{10} Z_{\text{eu}} = 1.524 - 0.460 [\text{Chl}]_{\text{surf}} - 0.00051 [\text{Chl}]_{\text{surf}}^2 + 0.0282 [\text{Chl}]_{\text{surf}}^{3} \). A comparison of the estimation of the base of the euphotic zone by using the Morel et al. (2007) equation and the data collected by a radiometer during the HERCULES cruise is shown in Fig. A1a.

2.1 Hydrography and turbulent mixing

Hydrographic properties and turbulent mixing were derived from a microstructure turbulent profiler (Prandke and Stips, 1998, MSS) equipped with a high-precision conductivity–temperature–depth (CTD) probe, two microstructure shear sensors (type PNS06), and also a sensor to measure the horizontal acceleration of the profiler. Measurements of dissipation rates of turbulent kinetic energy \( (\varepsilon) \) were conducted to the bottom, or to 137–323 m over deep waters (Table 1). The number of microstructure turbulence profiles used for computing nitrate fluxes at each station were always deployed successively. Sets include 2–11 in the tropical and subtropical Atlantic (37 ± 18 min), 6–7 in the Mediterranean (76 ± 22 min), and 3–402 in the Galician coastal upwelling (65 ± 246 min) (Fig. A1b in Appendix). Episodic bursts of turbulence can induce episodic inputs of nutrient supply, which can be easily missed in sets of a low number of profiles. In coastal regions where short-term variability in mixing processes is expected to be higher, our dataset includes two high-frequency samplings carried out in the outer part of Ría de Vigo (Galician upwelling ecosystem) in August 2013 (CHAOS cruises). During these cruises two 25 h series of turbulent microstructure and current observations were carried out during spring and neap tides. Turbulent kinetic energy dissipation at the interface between upwelled and surface waters was enhanced by 2 orders of magnitude during the ebbs, as the result of the interplay of the bidirectional upwelling circulation and the tidal current shear (Fernández-Castro et al., 2018). Diffusive nitrate fluxes due to the enhanced dissipation observed during spring tide could be responsible for about half of the phytoplankton primary production estimated in this system during periods of upwelling relaxation–stratification (Villamaña et al., 2017).

The profiler was balanced to have negative buoyancy and a sinking velocity of 0.4 to 0.7 m s\(^{-1}\). The frequency of data sampling was 1024 Hz. The sensitivity of the shear sensors was checked after each use. Due to significant turbulent generation close to the ship, only the data below 5 m (HERCULES1, HERCULES2, HERCULES3, DISTRAL, ASIMUTH, CHAOS, and NICANOR) and 10 m (CARPOS, TRYNTROP, FAMOSO1, FAMOSO2, FAMOSO3) were considered reliable. Data processing and calculation of dissipation rates of \( (\varepsilon) \) were carried out with the commercial software MSSpro. The squared Brunt–Väisälä frequency \( (N^2) \) was computed from the CTD profiles according to the equation

\[
N^2 = - \left( \frac{g}{\rho_w} \right) \left( \frac{\partial \rho}{\partial z} \right) (s^{-2}),
\]

(1)

where \( g \) is the acceleration due to gravity (9.8 m s\(^{-2}\)), \( \rho_w \) is seawater density (1025 kg m\(^{-3}\)), and \( \partial \rho/\partial z \) is the vertical potential density gradient. Vertical diffusivity \( (K_z) \) was estimated as

\[
K_z = \Gamma \frac{\varepsilon}{N^2} (m^2 s^{-1}),
\]

(2)

where \( \Gamma \) is the mixing efficiency, here considered to be 0.2 (Osborn, 1980).

2.2 Nutrient supply

Samples for the determination of nitrate (\( \text{NO}_3 \)) + nitrite (\( \text{NO}_2 \)) were collected from 5 ± 2 (Galician coastal upw-
welling), 7 ± 1 (Mediterranean), and 11 ± 2 (tropical and subtropical Atlantic Ocean) different depths in rinsed polyethylene tubes and stored frozen at −20 °C until analysis on land, according to standard methods using the automated colorimetric technique (Hansen and Koroleff., 1999). The frequency histogram of sampling depths collected for nitrate concentration in each region is indicated in Fig. A1c, whereas the maximum sampling depth at which the microstructure turbulence profiler was deployed is indicated in Table 1. At one station carried out during the NICANOR cruises, where nitrate concentrations were not available, concentration values were obtained by using a nitrate-density relationship built by using all samples (n = 52) collected during the NICANOR sampling period. The relationship showed a linear behavior (NO\textsubscript{3} = 9.7788 × σ − 256.38; Adj-\(r^2 = 0.87; p < 0.001\)) for density ranging between 26.1 and 27.1 kg m\textsuperscript{-3} (Fig. A1d).

Vertical diffusive fluxes of nitrate into the euphotic zone were calculated following Fick’s law as

\[
\text{Flux}_{\text{NO}_3} = \bar{K}_z \Delta \text{NO}_3, \tag{3}
\]

where \(\Delta \text{NO}_3\) is the nitrate vertical gradient obtained by linear fitting of nitrate concentrations in the nitracline, determined as a region of approximately maximum and constant gradient, and \(\bar{K}_z\) is the averaged turbulent mixing over the same depth interval. In the Galician coastal upwelling, nitrate diffusive fluxes were estimated over a fixed depth interval using the same procedure (10–40 m) except at the shallowest stations at which we compute the surface–bottom flux.

Most stations carried out in the Galician coastal upwelling were conducted inside three different rias (Ría de Vigo, Ría de Pontevedra, and Ría de A Coruña). The rias are coastal embayments affected by seasonal wind-driven coastal upwelling of cold, nutrient-rich North Atlantic Central Water (Wooster et al., 1976; Fraga, 1981; Alvarez-Salgado et al., 1993). The Galician rias, despite being, in general, longer and narrower than many open bays in upwelling areas, they resemble them in that its primary hydrographic and circulation features are determined by the extension of wind-driven flow on the external continental shelf throughout the bay (Alvarez-Salgado et al., 2010). Fertilization in the rias occurs essentially by coastal upwelling, with fresh and rain water inputs being residual (2 %) (Fernández et al., 2016). The total nitrate supply in the Galician rias was computed as the sum of nitrate vertical diffusion plus nitrate vertical advection due to coastal upwelling. A simplified estimate of nitrate supply through vertical advection due to upwelling was computed considering the Galician rias as single boxes divided into two layers (Álvarez-Salgado et al., 1993), the deeper one influenced by upwelled inflowing waters and the surface layer dominated by the outgoing flow. Assuming that the bottom layer volume is conservative and stationary, the vertical advective flux \(Q_Z, \text{m}^3\text{s}^{-1}\) would be equivalent to the incoming bottom flux \(Q_B, \text{m}^3\text{s}^{-1}\), computed as the product of the upwelling index \(I_W, \text{m}^3\text{s}^{-1}\text{km}^{-1}\) and the lengths of the mouth of the rias (ca. 10–11.5 km). \(I_W\) was averaged over the 3-day period before each cruise from wind data recorded by meteorological buoys located in Cabo Vilán (HERCULES, NICANOR) and Cabo Silleiro (DISTRAL, ASIMUTH, CHAOS, ASIMUTH), or modeled by the Fleet Numerical Meteorology and Oceanography Center (FNMOC) model when buoy data were not available (http://www.indicedeafioramento.ieo.es; last access: 18 October 2018). Finally, the transport of nitrate into the euphotic zone through vertical advection was computed as

\[
\text{NO}_3 \text{ Advection flux} = \frac{Q_Z}{A_{\text{basin}}}[\text{NO}_3]_D, \tag{4}
\]

where \(A_{\text{basin}}\) is the surface area of the Galician rias, \(Q_Z\) is the vertical advective flux, and \([\text{NO}_3]_D\) is the averaged nitrate concentration at the base of the euphotic layer. \(A_{\text{basin}}\) is 141 km\textsuperscript{2} for Ría de Pontevedra (ASIMUTH), 174 km\textsuperscript{2} for Ría de Vigo (CHAOS, ASIMUTH, DISTRAL), and 145 km\textsuperscript{2} for Ría de A Coruña (HERCULES, NICANOR) (see Villamaña et al., 2017; Villamaña et al., 2017; Moreira-Coello et al., 2017; Moreira-Coelho et al., 2017, for details).

### 2.3 Flow cytometry

Picoplankton samples (1.8 mL) for the determination of picoplankton abundance and cell properties were taken from three to nine depths and measured immediately onboard (TRYNITROP), or preserved with paraformaldehyde plus glutaraldehyde (\(P + G\)) and frozen at −80 °C until analysis in the laboratory (the other cruises). Unfortunately, due to problems with sample preservation, only heterotrophic or autotrophic picoplankton subgroup data were available for the DISTRAL and ASIMUTH cruises, respectively. Two aliquots from the same sample were used for the study of picophytoplankton (0.6 mL) and heterotrophic prokaryotes (0.4 mL), analyzed at high (mean 60 µL min\textsuperscript{-1}) and low (mean 18 µL min\textsuperscript{-1}) flow rates for 4 and 2 min, respectively. Before the analysis, the DNA of heterotrophic prokaryotes was stained with nucleic-acid-specific fluorescent dye (SYTO-13 or SYBR1). A FACSCalibur flow cytometer (Becton, Dickinson and Company) equipped with a laser emitting at 488 nm was used to measure and count picoplankton. Autotrophic cells were separated into two groups of cyanobacteria (\textit{Synechococcus} and \textit{Prochlorococcus}) and one group of small picoeukaryotes, based on their fluorescence and light scatter signals (size scatter, SSC), as explained in Calvo-Díaz and Morán (2006). Two groups of heterotrophic prokaryotes (LNA and HNA) were distinguished based on their relative green fluorescence, which was used as a proxy for nucleic acid content (Gasol and del Giorgio, 2000; Bouvier et al., 2007).

In order to estimate biovolume (BV), we used an empirical calibration between SSC and cell diameter (Calvo-Díaz and Morán, 2006), assuming spherical shape for all groups. The following volume-to-carbon conversion fac-
tors were used for picoautotrophic groups: 230 fg C BV for *Synechococcus*, 240 fg C BV for *Prochlorococcus*, and 237 fg C BV for picoeukaryotes (Worden et al., 2004). For bacteria, BV was converted into carbon biomass by using the allometric relationship: 108.8 fg C BV\(^{0.898}\) (Gundersen et al., 2002). More details about the processing and analysis of flow cytometry samples are provided in Calvo-Díaz and Morán (2006) (Calvo-Díaz and Morán, 2006, TRYNITROP), Gomes et al. (2015) (Gomes et al., 2015, FAMOSO), Villamaña et al. (2017) (Villamaña et al., 2017, CHAOS), and Moreira-Coello et al. (2017) (Moreira-Coello et al., 2017, NICANOR). Abundance data obtained at different depths for each station were combined to compute depth-integrated biomass for the photic layer.

2.4 Generalized additive models

A generalized additive model (GAM) approach was used to predict depth-integrated biomass of each picoplankton subgroup, the contribution of LNA prokaryotes to heterotrophic picoplankton, the cyanobacteria-to-picoeukaryote ratio, and the autotrophic-to-heterotrophic ratio based on observations and estimates of three environmental factors: sea surface temperature (SST), daily surface PAR, and the transport of nitrate into the euphotic zone (NO\(_3\) flux), including both diffusive and advective processes. GAMs assume that the effect of each predictor on the response variable can be described by smooth functions whose effects are additive. Due to the large number of zero observations, data overdispersion, and the need for a single parsimonious model to make predictions for a large number of groups, we assumed that the depth-integrated biomass of each picoplankton subgroup, relative contribution values, and biomass ratios all followed negative binomial distributions. Those niche descriptors that did not follow normal distributions were log transformed. The complete model structure for the biomass of each picoplankton subgroup was

\[ y_j = 1 + s(\text{SST}) + s(\text{PAR}) + s(\log(\text{NO}_3\text{Flux})) + \text{Error}, \]  

where \( y_j \) represents the depth-integrated biomass for each picoplankton subgroup \( j \), and \( s \) a cubic regression spline used for fitting the observations to the model (Wood, 2006).

Generalized models include a function linking the mean value of \( y_j \) and the predictors. For those response variables that followed a negative binomial distribution, the used link function was the natural logarithm. The LNA contribution to total heterotrophic prokaryotes was adjusted using a Gaussian distribution and an identify link (Wood et al., 2016). The inclusion of the different predictors to explain the response variable (the biomass of each picoplankton subgroup, its relative contribution, and biomass ratios) was assessed via stepwise model selection using the minimum Akaike information criterion (Hastie and Tibshirani, 1993).

Smooth terms were tested using a Bayesian test (Marra and Wood, 2012) to prevent overfitting. GAMs were evaluated based on explanatory power (explained variance) and goodness of fit (GOF). GOF was assessed via quantile–quantile (QQ) plots of Pearson residuals (provided in SM Fig A2). All calculations were carried out using the mgcv package (Wood, 2011) in R (R Core Team, 2015).

2.5 Niche overlap analysis

The estimation of niche overlap among different picoplankton subgroups based on nonparametric kernel density functions (NO\(_K\)) was calculated following Mouillot et al. (2005):

\[ \text{NO}_{K_{i,j,t}} = 1 - \frac{1}{2} \int |f_{it}(x) - f_{jt}(x)| \, dx, \]  

where \( \text{NO}_{K_{i,j,t}} \) is the niche overlap between picoplankton subgroups \( i \) and \( j \) for the environmental factor \( t \), and \( f_{it} \) and \( f_{jt} \) are the kernel population density functions of factor \( t \) for species \( i \) and \( j \), respectively. In order to correct the correlation among niche predictors, we used the estimator in a dependent sample (EDS) proposed by Kark et al. (2002).

To assess the statistical niche differences between subgroups, null model permutation tests were performed to verify whether the niche overlaps were significantly lower than 100% (Geange et al., 2011). When the contribution of depth-integrated biomass for each picoplankton subgroup exceeded that expected by chance (one-third for autotrophic and one-half for heterotrophic picoplankton), niche predictors for each station were selected. Statistical null distributions (the distribution of the statistic test under the null hypothesis of no niche differentiation) were generated by calculating pseudo-values through randomly permuting group labels in the corresponding dataset over 10,000 runs. The distributions of the average niche overlaps for the null model were then computed. Niche overlap calculations and associated null model tests were performed using the density function and the source code provided as supporting information in Geange et al. (2011). All calculations were performed using R (R Core Team, 2015).

3 Results

3.1 Environmental variables and picoplankton biomass

Our database covered a wide environmental gradient from oligotrophic to eutrophic conditions. Stations sampled in the tropical and subtropical Atlantic Ocean (T) were, on average, characterized by warm surface waters (26 ± 2 °C, mean ± SD) where the supply of nitrate through vertical diffusion from deeper waters (0.7 ± 1.6 mmol N m\(^{-2}\) d\(^{-1}\)) and surface chlorophyll \( a \) were low (0.1 ± 0.1 mg m\(^{-3}\)) (Table 2 and Fig. 2). The Mediterranean Sea, sampled from March to September, was characterized by cooler surface waters (16 ± 4 °C) and intermediate values of both nitrate vertical diffusive supply (41 ± 113 mmol N m\(^{-2}\) d\(^{-1}\)) and also surface chlorophyll \( a \) (0.9 ± 0.9 mg m\(^{-3}\)). Finally, the stations...
sampled in the Galician coastal upwelling system, which included year-round samples, were characterized by relatively cold surface waters (16 ± 2 °C), enhanced rates of nitrate supply (30 ± 46 mmol N m⁻² d⁻¹), and high values of surface chlorophyll a (2.2 ± 2.5 mg m⁻³). No statistically significant differences were observed in averaged surface PAR among the three regions.

Differences in picoplankton biomass and composition were also observed among the three domains. Averaged photic layer depth-integrated total picoplankton biomass (including both LNA and HNA prokaryotes, Prochlorococcus, Synechococcus, and picoeukaryotes) was higher in the tropical and subtropical Atlantic (1052 ± 215 mg C m⁻²) and the Mediterranean (1038 ± 485 mg C m⁻²), compared to the Galician coastal upwelling (216 ± 36 mg C m⁻²) (Table 2). In the tropical and subtropical Atlantic, the contribution of Prochlorococcus to total picoplankton biomass was 41 %, followed by LNA (27 %) and HNA (22 %) prokaryotes, with smaller contributions of Synechococcus and picoeukaryotes (< 5 % each). In the Mediterranean, picoplankton biomass was on average dominated by Synechococcus (50 %), followed by LNA and HNA prokaryotes (~ 17 % for each group), picoeukaryotes (10 %), and Prochlorococcus (5 %). Finally, HNA (55 %) and LNA (21 %) prokaryotes dominated in the Galician coastal upwelling (G).

Table 2. Mean ± standard deviation of sea surface temperature (SST), surface photosynthetic active radiation (PAR), mixed-layer depth (MLD), photic layer depth (1 % PAR), surface nitrate concentration (sNO₃), nitrate gradient, vertical diffusivity (K), nitrate supply (NO₃ flux), surface chlorophyll (sChl a), photic layer depth-integrated chlorophyll a (Chl a); biomass (B), abundance (A), and contribution (C) to total picoplankton biomass (Total Pico B); and surface abundance (s) of LNA and HNA bacteria, Prochlorococcus, Synechococcus and picoeukaryotes computed for the tropical and subtropical Atlantic Ocean (T), the Mediterranean (M), and the Galician coastal upwelling (G). MLD was estimated from an increase in water column density of 0.125 Kg m⁻³ relative to surface values. A nonparametric one-way ANOVA (Kruskal–Wallis) was performed to test the null hypothesis that independent groups come from the same distribution. The Bonferroni multiple comparison test was applied posteriori to analyze the differences between every pair of groups (* p < 0.05; ** p < 0.01; *** p < 0.001).

<table>
<thead>
<tr>
<th>Variables (units)</th>
<th>T</th>
<th>M</th>
<th>G</th>
<th>KW p value</th>
<th>Post hoc Bonferroni</th>
</tr>
</thead>
<tbody>
<tr>
<td>SST (°C)</td>
<td>26 ± 2</td>
<td>16 ± 4</td>
<td>16 ± 2</td>
<td>&lt;0.001***</td>
<td>T &gt; G &gt; M</td>
</tr>
<tr>
<td>PAR (E m⁻² d⁻¹)</td>
<td>43 ± 23</td>
<td>42 ± 13</td>
<td>49 ± 17</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>MLD (m)</td>
<td>61 ± 30</td>
<td>61 ± 71</td>
<td>57 ± 10</td>
<td>&lt;0.001**</td>
<td>T, M &gt; G</td>
</tr>
<tr>
<td>sNO₃ (µmol m⁻³)</td>
<td>200 ± 7</td>
<td>241 ± 1036</td>
<td>161 ± 1604</td>
<td>&lt;0.001***</td>
<td>T &lt; M, G</td>
</tr>
<tr>
<td>NO₃ gradient (µmol m⁻³)</td>
<td>146 ± 18</td>
<td>90 ± 40</td>
<td>105 ± 100</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>K (m² s⁻¹) × 10⁻³</td>
<td>0.2 ± 0.3</td>
<td>5.3 ± 13.9</td>
<td>0.5 ± 0.8</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>NO₃ flux (µmol m⁻² d⁻¹)</td>
<td>0.7 ± 1.6</td>
<td>41 ± 113</td>
<td>30 ± 46</td>
<td>&lt;0.001***</td>
<td>T &lt; M &lt; G</td>
</tr>
<tr>
<td>sChl a (µg ml⁻¹)</td>
<td>0.1 ± 0.1</td>
<td>0.9 ± 0.9</td>
<td>2.2 ± 2.5</td>
<td>&lt;0.01**</td>
<td>T &lt; M &lt; G</td>
</tr>
<tr>
<td>Chl a (µg ml⁻²)</td>
<td>31 ± 6</td>
<td>70 ± 99</td>
<td>81 ± 66</td>
<td>&lt;0.001***</td>
<td>T &lt; G</td>
</tr>
<tr>
<td>sLNA A (cell ml⁻¹)</td>
<td>4.4 ± 2.4</td>
<td>3.7 ± 2.5</td>
<td>2.1 ± 1.1</td>
<td>&lt;0.001***</td>
<td>T &gt; M &gt; G</td>
</tr>
<tr>
<td>sHNA A (cell ml⁻¹)</td>
<td>3.0 ± 1.8</td>
<td>4.0 ± 4.5</td>
<td>3.6 ± 2.3</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>sProch A (cell ml⁻¹)</td>
<td>1.14 ± 0.13</td>
<td>2.2 ± 4.4</td>
<td>1.0 ± 2.8</td>
<td>&lt;0.001***</td>
<td>T &gt; M, G</td>
</tr>
<tr>
<td>sSynech A (cell ml⁻¹)</td>
<td>18 ± 66</td>
<td>75 ± 81</td>
<td>5.7 ± 6.9</td>
<td>&lt;0.001***</td>
<td>T &lt; M &lt; G</td>
</tr>
<tr>
<td>sPicoEuk A (cell ml⁻¹)</td>
<td>2.5 ± 9.4</td>
<td>6.8 ± 8.4</td>
<td>5.7 ± 6.9</td>
<td>&lt;0.001***</td>
<td>T &lt; M, G</td>
</tr>
<tr>
<td>LNA A (cell ml⁻²)</td>
<td>40 ± 20</td>
<td>22 ± 8</td>
<td>6.4 ± 4</td>
<td>&lt;0.001***</td>
<td>T, M &gt; G</td>
</tr>
<tr>
<td>HNA A (cell ml⁻²)</td>
<td>27 ± 1</td>
<td>22 ± 1</td>
<td>9.4 ± 0.8</td>
<td>&lt;0.001***</td>
<td>T &gt; M &gt; G</td>
</tr>
<tr>
<td>Prochla A (cell ml⁻²)</td>
<td>156 ± 121</td>
<td>10 ± 23</td>
<td>0.5 ± 1</td>
<td>&lt;0.001***</td>
<td>T &gt; M, G</td>
</tr>
<tr>
<td>Synech A (cell ml⁻²)</td>
<td>7 ± 15</td>
<td>50 ± 49</td>
<td>2 ± 2</td>
<td>&lt;0.001***</td>
<td>T &gt; M &gt; G</td>
</tr>
<tr>
<td>PicoEuk A (cell ml⁻²)</td>
<td>1.7 ± 3</td>
<td>2.8 ± 2</td>
<td>1 ± 2</td>
<td>&lt;0.001***</td>
<td>T &lt; M &gt; G</td>
</tr>
<tr>
<td>LNA B (mg C ml⁻²)</td>
<td>253 ± 105</td>
<td>170 ± 97</td>
<td>43 ± 34</td>
<td>&lt;0.001***</td>
<td>T &gt; M &gt; G</td>
</tr>
<tr>
<td>HNA B (mg C ml⁻²)</td>
<td>216 ± 127</td>
<td>168 ± 105</td>
<td>108 ± 73</td>
<td>0.02*</td>
<td>T &gt; M &gt; G</td>
</tr>
<tr>
<td>Prochla B (mg C ml⁻²)</td>
<td>482 ± 516</td>
<td>36 ± 84</td>
<td>1.3 ± 4</td>
<td>&lt;0.001***</td>
<td>T &gt; M, G</td>
</tr>
<tr>
<td>Synech B (mg C ml⁻²)</td>
<td>43 ± 83</td>
<td>576 ± 530</td>
<td>19 ± 26</td>
<td>&lt;0.001***</td>
<td>T, M &gt; G</td>
</tr>
<tr>
<td>PicoEuk B (mg C ml⁻²)</td>
<td>59 ± 102</td>
<td>86 ± 59</td>
<td>43 ± 59</td>
<td>&lt;0.001***</td>
<td>T &lt; M &gt; G</td>
</tr>
<tr>
<td>Total Pico B (mg C ml⁻²)</td>
<td>1052 ± 215</td>
<td>1038 ± 485</td>
<td>216 ± 36</td>
<td>&lt;0.001***</td>
<td>T, M &gt; G</td>
</tr>
<tr>
<td>LNA C (%)</td>
<td>27 ± 10</td>
<td>18 ± 8</td>
<td>21 ± 9</td>
<td>&lt;0.001***</td>
<td>T &gt; M, G</td>
</tr>
<tr>
<td>HNA C (%)</td>
<td>22 ± 12</td>
<td>17 ± 10</td>
<td>55 ± 15</td>
<td>&lt;0.001***</td>
<td>T, M &lt; G</td>
</tr>
<tr>
<td>Prochla C (%)</td>
<td>41 ± 16</td>
<td>5 ± 12</td>
<td>1 ± 2</td>
<td>&lt;0.001***</td>
<td>T, M &lt; G</td>
</tr>
<tr>
<td>Synech C (%)</td>
<td>4 ± 5</td>
<td>50 ± 24</td>
<td>6 ± 7</td>
<td>&lt;0.001***</td>
<td>T, G &gt; M</td>
</tr>
<tr>
<td>PicoEuk C (%)</td>
<td>5 ± 5</td>
<td>10 ± 9</td>
<td>11 ± 9</td>
<td>&lt;0.001***</td>
<td>T &lt; M, G</td>
</tr>
</tbody>
</table>
6206 J. L. Otero-Ferrer et al.: Factors controlling picoplankton community structure

Figure 2. Box-and-whisker plots of sea surface temperature (SST), surface photosynthetic active radiation (PAR), nitrate supply (NO₃ Flux), surface chlorophyll a concentration (sChl a), and contribution to total picoplankton biomass of low (LNA) and high (HNA) nucleic acid content bacteria, Prochlorococcus (Proch), Synechococcus (Synech), and small picoeukaryotes (PicoEuk) computed for the tropical and subtropical Atlantic Ocean (T), the Mediterranean (M), and the Galician coastal upwelling (G). In each box, the central mark indicates the median, the notches the 95% confidence interval for the median, and the bottom and top edges of the box the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using white circles.

followed by picoeukaryotes (11%), Synechococcus (6%), and Prochlorococcus (1%).

Vertical distributions of temperature, nitrate concentration, and the biomass of autotrophic and heterotrophic picoplankton groups for each domain are shown in Fig. A3.

3.2 The role of environmental factors in picoplankton composition

In order to explore the role of temperature, light, and nitrate supply in the composition of the picoplankton community, we first used generalized linear models to investigate simple linear relationships among each of these factors and the depth-integrated biomass of each picoplankton subgroup, the contribution of LNA prokaryotes to heterotrophic picoplankton biomass, the ratio of cyanobacteria (Prochlorococcus and Synechococcus) to picoeukaryote, and the autotrophic-to-heterotrophic picoplankton ratio (Fig. 3 and Table 3). All picoplankton groups except picoeukaryotes exhibited statistically significant relationships with SST. This relationship was negative for Synechococcus and positive for all the other subgroups (Fig. 3). Only LNA and HNA prokaryotes and Synechococcus exhibited significant, positive relationships with surface radiation. All groups except the picoeukaryotes were negatively correlated with nitrate fluxes. The contribution of LNA prokaryotes to heterotrophic picoplankton biomass only exhibited a significant negative relationship with nitrate fluxes, whereas the cyanobacteria-to-picoeukaryote ratio was positively correlated with surface temperature and negatively correlated with nitrate fluxes. Finally, the ratio of autotrophic to heterotrophic biomass was not linearly correlated with any of the studied environmental factors. In order to exclude cross correlation among the three environmental factors and consider the possibility of nonlinear relationships, we subsequently fitted the data to GAMs (Fig. 4 and Table 3). Temperature was the only factor included in the models built for predicting the depth-integrated biomass of all picoplankton subgroups. HNA prokaryotes exhibited a positive relationship with temperature above 19°C, whereas Prochlorococcus and LNA prokaryotes showed a nearly sigmoid curve re-
Figure 3. Pair scatter plots representing the relationship among log-transformed depth-integrated biomass for each picoplankton subgroup, the contribution of bacteria with a low nucleic acid content to heterotrophic picoplankton biomass (%LNA), the ratio of cyanobacteria \((Prochlorococcus + Synechococcus)\) to picoeukaryote depth-integrated biomass \((CB / PicoEuK)\), and the ratio of autotrophic to heterotrophic picoplankton biomass \((Auto / hetero)\) versus sea surface temperature (SST), surface photosynthetically active radiation (PAR), and nitrate flux \((NO_3\) flux). Significant linear relationships are indicated as solid \((p \text{ value} < 0.01)\) black lines. Samples collected at different regions are indicated as red dots (tropical and subtropical Atlantic Ocean), green squares (Mediterranean), and blue diamonds (Galician coastal upwelling).

The relationship with a transition between ca. 16 and 25°C (Fig. 4). The relationship between the biomass of both \(Synechococcus\) and picoeukaryotes and temperature showed a negative trend until \(\sim 20°C\) and remained relatively constant above this temperature. PAR was included in the models of all picoplankton groups except picoeukaryotes. Whereas the biomass of \(Prochlorococcus\) exhibited a saturation-type relationship with PAR, heterotrophic prokaryotes and \(Synechococcus\) showed a linear positive relationship. Finally, only LNA prokaryotes, \(Synechococcus\), and picoeukaryotes exhibited statistically significant relationships with nitrate supply. This relationship was negative for LNA prokaryotes and \(Synechococcus\), whereas picoeukaryotes showed a unimodal function, peaking at \(\sim 1\) mmol \(NO_3\) m\(^{-2}\) d\(^{-1}\). Nitrate flux was the only factor selected in the models to predict the contribution of LNA prokaryotes to heterotrophic picoplankton biomass and both the cyanobacteria-to-picoeukaryote biomass ratio and the autotrophic-to-heterotrophic biomass ratio. This relationship was negative in the three models. Temperature was also negatively correlated with the contribution of LNA prokaryotes to heterotrophic biomass and the ratio of autotrophic to heterotrophic biomass. Finally, PAR showed a positive correlation with the contribution of LNA prokaryotes to heterotrophic biomass and the cyanobacteria-to-picoeukaryote biomass ratio.

3.3 Ecological niches for picoplankton groups

By using nonparametric kernel density functions, we investigated the overlapping in the ecological niches of the autotrophic and heterotrophic picoplankton subgroups de-
Figure 4. GAM-predicted effects of the response variables (biomass and contribution of picoplankton subgroups) as a smooth function of sea surface temperature (SST), photosynthetically active radiation (PAR), and nitrate flux (NO$_3$ flux). All terms were centered at zero. Significant linear relationships are indicated as solid ($p$ value < 0.01) black lines. Shaded regions represent the 95% confidence intervals of the smooth spline functions. Intercept values were 4.6 (LNA), 5.1 (HNA), 2.1 (Prochlorococcus), 3.6 (Synechococcus), 3.7 (picoeukaryotes), 36.4 (contribution of LNA to heterotrophic picoplankton, %LNA), 1.4 (cyanobacteria-to-picoeukaryote ratio, CB/PicoEuk), and −0.1 (autotrophic-to-heterotrophic biomass ratio, Auto/hetero).

Defined by using the three variables previously considered together with surface nitrate concentration (Fig. 5). Photic layer depth-integrated biomass of each picoplankton group was used for this analysis. These results revealed three ecological niches in the distribution of picoplankton subgroups. Prochlorococcus and LNA prokaryotes were more abundant in warm waters, where nitrate supply was low. HNA prokaryotes and Synechococcus dominated in cooler regions with medium to high nitrate supply, and picoeukaryotes were more abundant in cold waters with high nitrate supply. A large degree of overlapping of the ecological niches for all picoplankton subgroups was observed when only surface light was considered. For each picoplankton subgroup Table 4 shows the partial weighted overlap of the ecological niches defined by the four factors: SST, PAR, nitrate flux, and surface nitrate concentration. According to these data only nitrate supply enabled a statistically significant separation of the niches of both heterotrophic (HNA and LNA prokaryotes) and autotrophic (Prochlorococcus, Synechococcus, and picoeukaryotes) picoplankton subgroups. Although the minimum overlap between Prochlorococcus and the other autotrophic picoplankton subgroups was also well defined by temperature, only nitrate supply could statistically distinguish the niche partitioning between the two groups of heterotrophic prokaryotes ($p < 0.05$) and between Synechococcus and picoeukaryotes ($p < 0.1$).
Table 3. Simple ($R^2$) and adjusted squared correlation coefficients (Adj-$R^2$) for simple linear regression and multiple generalized additive models (GAMs) built to predict depth-integrated biomass for each picoplankton subgroup, the contribution of LNA bacteria to total heterotrophic picoplankton biomass (% LNA), the ratio of cyanobacteria ($Prochlorococcus + Synechococcus$) to picoeukaryote depth-integrated biomass (CB / PicoEuK), and the ratio of autotrophic (CB + PicoEuK) to heterotrophic bacteria (LNA + HNA) biomass based on sea surface temperature (SST), surface photosynthetically active radiation (PAR), and nitrate supply (NO$_3$ flux). Negative binomial distribution was assumed. Multiple model selection was based on stepwise regression and the Akaike information criterion (see Sect. 2). Only significant ($p$ value < 0.05) results are shown. Percentage of total effects represents the contribution of each environmental factor to the variability explained by each GAM model (see Sect. 2).

<table>
<thead>
<tr>
<th>Group</th>
<th>$R^2$ simple linear</th>
<th>Adj-$R^2$ multiple regression</th>
<th>Percentage of total effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SST</td>
<td>PAR</td>
<td>NO$_3$ flux</td>
</tr>
<tr>
<td>LNA</td>
<td>0.39</td>
<td>&lt; 0.01</td>
<td>0.34</td>
</tr>
<tr>
<td>HNA</td>
<td>0.47</td>
<td>0.05</td>
<td>0.18</td>
</tr>
<tr>
<td>Proch</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td></td>
</tr>
<tr>
<td>Synech</td>
<td>0.11</td>
<td>&lt; 0.01</td>
<td></td>
</tr>
<tr>
<td>PicoEuK</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% LNA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CB / PicoEuK</td>
<td>0.25</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>Auto / hetero</td>
<td>0.29</td>
<td>0.39</td>
<td></td>
</tr>
</tbody>
</table>

Figure 5. Kernel density estimates of LNA and HNA bacteria, $Prochlorococcus$, $Synechococcus$, and picoeukaryotes based on the considered niche descriptors: sea surface temperature (SST), surface photosynthetically active radiation (PAR), nitrate flux (NO$_3$ flux), and surface concentration ($s$NO$_3$).

4 Discussion

4.1 Environmental factors and ecological niches

Picoplankton community composition and concurrent estimates of nitrate supply into the euphotic zone from highly contrasting marine environments allowed us to conclude that SST and nitrate supply are the main factors controlling the variability in the biomass of different subgroups, whereas surface light generally played a minor role. As far as we know, only one study had previously investigated the role of these environmental factors in the distribution of, in this case, the two major groups of cyanobacteria. By using a large flow cytometry dataset from all major ocean regions, Flombaum et al. (2013) concluded that temperature and light were the most important predictors of the abundances of $Prochlorococcus$ and $Synechococcus$, with nitrate availability exerting a negligible effect. Although this conclusion seems to be contradictory with the results presented here, some important differences between these studies should be noted. Firstly, Flombaum et al. (2013) used bulk estimates of sea-
Table 4. Partial weighted niche overlap (%) for each environmental factor and picoplankton subgroup. sNO$_3$ represents surface nitrate concentration. Asterisks denote the existence of significant differences among niches (*p < 0.1, **p < 0.05, ***p < 0.01, ****p < 0.001).

<table>
<thead>
<tr>
<th></th>
<th>LNA</th>
<th>HNA</th>
<th>Proch</th>
<th>Synech</th>
<th>PicoEuk</th>
</tr>
</thead>
<tbody>
<tr>
<td>SST</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LNA</td>
<td>100</td>
<td></td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HNA</td>
<td>85</td>
<td>100</td>
<td>85</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Proch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synech</td>
<td>9****</td>
<td>100</td>
<td>9****</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>PicoEuk</td>
<td>5****</td>
<td>84</td>
<td>84</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>PAR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LNA</td>
<td>100</td>
<td></td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HNA</td>
<td>80</td>
<td>100</td>
<td>80</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Proch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synech</td>
<td>80</td>
<td>100</td>
<td>80</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>PicoEuk</td>
<td>74**</td>
<td>94</td>
<td>74**</td>
<td>94</td>
<td>100</td>
</tr>
<tr>
<td>NO$_3$ Flux</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LNA</td>
<td>100</td>
<td></td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HNA</td>
<td>69**</td>
<td>100</td>
<td>69**</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Proch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synech</td>
<td>100</td>
<td></td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PicoEuk</td>
<td>14****</td>
<td>77*</td>
<td>14****</td>
<td>77*</td>
<td>100</td>
</tr>
<tr>
<td>sNO$_3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LNA</td>
<td>100</td>
<td></td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HNA</td>
<td>73*</td>
<td>100</td>
<td>73*</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Proch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synech</td>
<td>100</td>
<td></td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PicoEuk</td>
<td>29****</td>
<td>89</td>
<td>29****</td>
<td>89</td>
<td>100</td>
</tr>
</tbody>
</table>

Water nitrate concentration as a proxy for nitrate availability in the euphotic zone. However, in near-steady-state systems such as the subtropical gyres, where diffusive nutrient supply into the euphotic zone is slow, nitrate concentrations are kept close to the detection limit due to phytoplankton uptake. For this reason, nitrate concentrations and actual nitrate supply into the euphotic zone in the vast oligotrophic regions are often largely disconnected (Mouriño-Carballido et al., 2011, 2016). Moreover, whereas Flombaum et al. (2013) used Prochlorococcus and Synechococcus abundances determined at several depths in the upper 200 m, we used depth-integrated biomass of both autotrophic and heterotrophic picoplankton subgroups in the photic layer.

Although our results point to both temperature and nitrate supply as important factors controlling the distribution of the picoplankton subgroups (Fig. 4 and Table 3), nitrate supply was the only factor that allowed the distinction among the ecological niches of autotrophic and heterotrophic picoplankton subgroups in the photic layer.

Although previous studies have revealed that Prochlorococcus may have acquired the ability to use nitrate by horizon-
tal gene transfer, their photosynthetic activity primarily relies on regenerated forms of nitrogen (Moore et al., 2002; Malmstrom et al., 2013). Our results support this view and substantiate that, after controlling for the concurrent effects of light and seawater temperature, Prochlorococcus biomass was uncorrelated with nitrate fluxes. However, it is important to note that we could not discriminate between high-light (HL) and low-light (LL) ecotypes and that the presence of nitrate reductase seems to be more relevant in LL (Martin et al., 2009; Berube et al., 2014). Evolutionary adaptation to light-limiting conditions makes Prochlorococcus the most efficient light harvesters among Earth’s photosynthetic organisms (Morel et al., 1993). Their competitive ability under light-limiting conditions could explain the negative effect of light as a predictor for Prochlorococcus biomass. Ultimately, the photo-physiological strategy of Prochlorococcus leads to (i) thermal sensitivity of photosystem II (Mackey et al., 2013), which limits its fundamental niche to temperatures greater than 15 °C (Moore et al., 1995), and (ii) high sensitivity to ultraviolet (UV) radiation in surface waters (Llabrés et al., 2010; Mackey et al., 2013; Sommaruga et al., 2005). This could explain that, after removing the effect of light, our data analysis revealed that the effect of temperature on Prochlorococcus biomass showed a sigmoid relationship as temperature increased.

Synechococcus is able to use both new and regenerated forms of nitrogen (Moore et al., 2002; Mulholland and Lomas, 2008), which largely explains its wider geographical distribution range (Flölbbaum et al., 2013). The fact that it is more abundant at intermediate levels of nitrate supply is consistent with the lower intracellular nitrogen quota of Synechococcus relative to Prochlorococcus and hence their higher growth rate under saturating nutrient conditions (Marañón et al., 2013). Conversely, the large affinity of Prochlorococcus to acquire nutrients (Partensky and Garcarz, 2010) and absorb light under severe nutrient- and light-limiting conditions (Mella-Flores et al., 2012) precludes the supremacy of Synechococcus in warm and stratified oligotrophic systems (Moore et al., 2007). Although Prochlorococcus and Synechococcus are not very different in cell size and they usually coexist in oligotrophic regions, differences in adaptation to light conditions and UV stress lead to segregate their maximal distributions across space (vertical segregation) and through time (Chisholm, 1992; Mella-Flores et al., 2012).

Picoeukaryotes, like Synechococcus, also exhibited a negative relationship with seawater temperature, again reflecting the superior competitive ability of Prochlorococcus under severe nutrient-limiting conditions (Moore et al., 2007). The relative dominance of cyanobacteria in oligotrophic systems results from the fact that cyanobacteria are less negatively affected by nutrient diffusion limitation than picoeukaryotes (Chisholm, 1992). It is widely accepted that small cells are at an advantage over large cells in stratified open-ocean environments, where nutrient recycling dominates biogeochemical fluxes (Raven, 1998). First, the surface-to-volume ratio increases with decreasing cell size, which narrows the nutrient diffusion boundary layer around the cell and facilitates the acquisition of nutrients in nutrient impoverished environments. Second, small-sized cells have lower sinking rates than their larger counterparts, which allow them to extend their chances of survival in the euphotic layer (Smayda, 1980; Chisholm, 1992; Kiørboe, 1993).

Our analysis indicates that among the picophytoplankton, picoeukaryotes were the most responsive to nutrient fluxes. This is consistent with experimental observations under laboratory-controlled conditions revealing that, within the picoplankton size range, the maximum attainable growth rate increases with increasing cell size (Raven, 1994; Marañón et al., 2013). This positive relationship between maximum growth rate and cell size in the picophytoplankton to small nanophytoplankton size range has been explained as a trade-off between intracellular nitrogen quotas (N requirements) and mass-specific nitrate uptake rates (N uptake) (Marañón et al., 2013). Whereas nitrogen uptake rate exhibits an isometric relationship with cell size, smaller picoplankton cells have substantially larger intracellular nitrogen quotas, which reduce their capability to maximize carbon-specific growth rates. Conversely, high maximum growth rates represent an advantage for picoeukaryotes, compared to any other organism, as they attenuate the effect of loss processes such as predation or the washout of plankton communities in highly dynamic, turbulent systems (Sherr et al., 2005; Echevarría et al., 2009). For instance, microzooplankton is thought to maintain the biomass of their prey under tight control, and thus slight variations in picophytoplankton growth rate may substantially alter the resulting biomass of the different picophytoplankton subgroups (Chen et al., 2009).

The unimodal relationship observed between the biomass of picoeukaryotes and nitrate supply (Fig. 4) could seem at first contradictory with the rising hypothesis proposed by Barber and Hiscock (2006), which describes that improved growth conditions benefit all phytoplankton size classes, including picoplankton. In this regard, Brewin et al. (2014) by using data collected along the Atlantic Meridional Transect cruises showed that < 2 µm size-fractionated chlorophyll was positively correlated with total chlorophyll only until a value of 1 mg m⁻³, and then it did not show any positive relationship with total chlorophyll. It is also important to note that surface abundance of picoplankton subgroups reported in our study, which are consistent with previous observations (Zubkov et al., 2000; Frojan et al., 2014; Teira et al., 2015), did show higher surface abundance of picoeukaryotes in the Galician coastal upwelling and the Mediterranean compared to the tropical and subtropical Atlantic (Table 2 and Fig. A3). However, this pattern was diluted when depth-integrated biomasses were computed since the lower limit for the integration (the base of the photic zone) was much shallower in the coastal upwelling domain (ca. 37 m) compared to the Mediterranean (ca. 62 m) and the tropical and subtropical regions (ca. 109 m).
Heterotrophic prokaryotes also use dissolved inorganic nutrients, including nitrate, for growth (Kirchman, 2000). Consistent with this, Gasol et al. (2009) showed a positive relationship between prokaryotic abundance and a proxy for nutrient supply in a latitudinal gradient across the Atlantic. They did not partition this effect on the two subgroups that can universally be differentiated among bacteria and archaea. Our results suggest that LNA prokaryotes respond less markedly to nutrient fluxes than HNA prokaryotes. The effect of nitrate supply on the biomass of LNA prokaryotes showed a linear negative relationship as nitrate supply increases, perhaps associated with their ability to survive under nutrient starving conditions (Mary et al., 2008). Under such conditions, proteorhodopsin-containing LNA prokaryotes (e.g., example SAR11) can use energy from light (Mary et al., 2008; Pinhassi et al., 2016), improving their competitiveness against non-proteorhodopsin-containing prokaryotes. Consistent with this idea, our results showed a positive relationship between the biomass of LNA prokaryotes and PAR. Li et al. (2004) already proposed the ubiquity of this bottom-up control of prokaryotic abundance in oligotrophic environments (< 1 mg Chl m⁻³). Therefore, we believe that the underlying cause for the clear niche difference between LNA and HNA prokaryotes is more the difference in the suite of genes (Schattenhofer et al., 2011) than in cell size (Morán et al., 2015).

5 Outlook

Picoeukaryotes are, to a large extent, analogous to cyanobacteria. Both are microorganisms that can perform autotrophic photosynthesis, and both have a significant impact on marine primary production. However, cyanobacteria are typically larger than picoeukaryotes, with the exception of the most abundant picoeukaryote, the coccolithophore *Emiliania huxleyi*. The role of picoeukaryotes in marine primary production is less well understood than that of cyanobacteria. However, recent studies have suggested that picoeukaryotes may play an important role in the transfer of carbon to the deep ocean and in the recycling of organic matter (Azam et al., 1983; Fenchel, 1992; Agawin et al., 2000), contributing overwhelmingly to the recycling of organic matter.
Figure A1. (a) Pair scatter plot representing the relationship between the euphotic zone depth ($Z_{eu}$) computed using the Morel et al. (2007) equation and the data collected by a radiometer during the HERCULES cruises; the solid line represents 1:1 relationship. (b) Frequency histogram of the number of turbulence profiles deployed at each station and domain. (c) Frequency histograms of the number of samples collected for the determination of nitrate concentration at each station and domain: tropical and subtropical Atlantic Ocean (red), the Mediterranean (green), and the Galician coastal upwelling (blue). (d) Pair scatter plot representing the relationship between nitrate concentration and density built by using all samples collected during the NICANOR sampling period.
Figure A2. Quantile–quantile (QQ) plots between the observations and the selected GAM models for each picoplankton subgroup, the contribution of LNA to heterotrophic picoplankton (%LNA), the cyanobacteria-to-picocoeukaryote ratio (CB / PicoEuk), and the autotrophic-to-heterotrophic biomass ratio (Auto / hetero). The y axes represent the Pearson residuals and the x axes the negative binomial theoretical quantiles. Solid red lines indicate the theoretical quantile of the models and grey shadows the 95% confidence intervals.
Figure A3. Vertical distribution of temperature (Temp), nitrate (NO$_3$), and picoplankton biomass of autotrophic (Phyto) and heterotrophic (Bacteria) groups for each domain: tropical and subtropical Atlantic Ocean (T), the Mediterranean (M), and the Galician coastal upwelling (G). Points represent raw data and the solid line the locally weighted scatter plot smoothing (LOESS). Dashed lines indicate 95% confidence intervals. Dot and line color intensity indicates the number of overlapping observations.
Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/bg-15-6199-2018-supplement.

Author contributions. JO, PC, and BM designed the research. JO analyzed the data, and JO and BM prepared the paper with contributions from all the co-authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank the officers and crew of the research vessels Hespérides, Sarmiento de Gamboa, Mytilus, Ramon Margalef and Lara for their help during the cruises. We are also very grateful to Fátima Eiroa for the flow cytometry analysis in the NICANOR, HERCULES1, HERCULES2, HERCULES3, and ASIMUTH cruises. Finally, we would like to thank Julia Uitz and the two anonymous reviewers for their valuable comments on the paper.

This research was supported by the Spanish Ministry of Economy and Competitiveness (MINECO) through projects CSTM2012-30680 to Beatriz Mouriño, CSTM2008-06261-C03-01 to Mikel Latasa, REN2003-09532-C03-01 to Ramiro Varela Benvenuto, CSTM2004-05174-C02 to Emilio Marañón, and CSTM2011-25035 to Pedro Cermeño; by the Galician government through grants 09MMA027604PR to Manuel Ruiz Villareal and 25035 to Pedro Cermeño; by the Galician government through the time series project RADIALES coordinated by Antonio Bode and by the 7th Framework Programme Oceanografia (IEO) through the time series project RADIALES coordinated by Antonio Bode and by the 7th Framework Programme of the European Commission through grant FP7 SPACE.2010.1.1-01 261860 to Manuel Ruiz Villareal. Jose Luis Otero Ferrer acknowledges the receipt of a FPI contract from MINECO (CTM2012-30680) and Bieito Fernández Castro a Juan de La Cierva Research Fellowship (FJCI-6412015-25712, Ministerio de Economía y Competitividad, Spanish government).

Edited by: Julia Uitz
Reviewed by: two anonymous referees

References


Cermeño, P., Chouciño, P., Fernández-Castro, B., Figueiras, F. G., Marañón, E., Marrasé, C., Mouriño-Carballido, B., Pérez-Lorenzo, M., Rodríguez-Ramos, T., Teixeira,


R Core Team: R: A Language and Environment for Statistical Computing, available at: http://www.r-project.org/ (last access: 23 October 2018), 2015.


