LIMNOLOGY AND OCEANOGRAPHY
Letters

ASLO
Open Access

Limnology and Oceanography Letters 2021

doi: 10.1002/lol2.10188

DATA ARTICLE

Prochlorococcus, *Synechococcus*, and picoeukaryotic phytoplankton abundances in the global ocean

Natalia Visintini[1,2,3]*, ***Adam C. Martiny***[4,5], ***Pedro Flombaum***[1,2,3,6]

[1]Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Buenos Aires, Argentina; [2]CONICET – Universidad de Buenos Aires, Centro de Investigaciones del Mar y la Atmósfera (CIMA), Buenos Aires, Argentina; [3]CNRS–IRD–CONICET–UBA, Instituto Franco-Argentino para el Estudio del Clima y sus Impactos (UMI 3351 IFAECI), Buenos Aires, Argentina; [4]Department of Earth System Science, University of California, Irvine, California; [5]Department of Ecology and Evolutionary Biology, University of California, Irvine, California; [6]Facultad de Ciencias Exactas y Naturales, Departamento de Ecología, Genética y Evolución, Universidad de Buenos Aires, Buenos Aires, Argentina

Scientific Significance Statement

Prochlorococcus, *Synechococcus*, and picoeukaryotic phytoplankton dominate oligotrophic oceans, contributing at least 10% of global primary productivity, are central to global biogeochemical cycles and, thus, there is growing interest to include them in Earth System Models. However, global datasets of these groups are sparse in time and space representing a strong limitation to validate biogeochemical models. Here, we provide a global compilation of in situ observations along with ancillary environmental variables, a global climatology, and a projected global abundance in future climate scenarios aimed to overcome these limitations.

Abstract

Marine picophytoplankton is the most abundant photosynthetic group on Earth; however, it is still underrepresented in dynamic ecosystem models. Major constraints for understanding its role in the ecosystem at a global scale are sparse data and lack of a baseline description of its distribution. Here, we present three datasets to

*Correspondence: natalia.visintini@cima.fcen.uba.ar

Associate editor: Kendra Spence Cheruvelil

Author Contribution Statement: PF and ACM compiled the in situ dataset and conducted the analysis for the future climate data record. NV conducted the data analysis for the monthly climatology and wrote the draft of the typescript with inputs from all authors.

Data Availability Statement: Data record 1: https://hdl.handle.net/1912/25514 DOI: 10.1575/1912/bco-dmo.793451.1. Data record 2: https://hdl.handle.net/1912/25779 DOI: 10.26008/1912/bco-dmo.811147.1. Data record 3: https://hdl.handle.net/1912/25782 DOI: 10.26008/1912/bco-dmo.793690.1, https://hdl.handle.net/1912/25780 DOI: 10.26008/1912/bco-dmo.793776.1, https://hdl.handle.net/1912/25781 DOI: 10.26008/1912/bco-dmo.793847.1. *Measurements*: Collection of flow cytometry counts for *Prochlorococcus, Synechococcus*, and picoeukaryotic phytoplankton and also microscope counts for *Synechococcus*, with ancillary environmental information from in situ observations. *Technology Types*: Data record 1: Compilation of observations; Data record 2: Quantitative niche modelling approach; Data record 3: Quantitative niche modelling approach under four climate scenarios. *Temporal range*: Data record 1: 1987 to 2008; Data record 2: January to December (averaged for years 1900 to 2017); Data record 3: 1901 to 2100. *Frequency or sampling interval*: Data record 1: Snapshot in time and time series; Data record 2: Monthly; Data record 3: Annual. *Spatial scale*: Data record 1: Site-based; Data record 2: Global; Data record 3: Global. *Grant sponsor information*: Financial support for this work was provided by NSF Division of Ocean Sciences awards OCE-1559002 and OCE-1848576, Agencia Nacional de Promoción Científica y Tecnológica (PICT-2017-3020), and Universidad de Buenos Aires (UBACyT 20020170100620BA).

Additional Supporting Information may be found in the online version of this article.

assess the global abundance of the principal groups of picophytoplankton, *Prochlorococcus*, *Synechococcus*, and picoeukaryotic phytoplankton: (1) a compilation of 109,045 field observations with ancillary environmental data, (2) a global monthly climatology of 1° grids from 0 to 200 m, and (3) four climate scenarios projections, from the Coupled Model Intercomparison Project 5, spanning years 1901 to 2100. Together this set of observational and modeled data can improve our understanding of the role of picophytoplankton in the global ecosystem.

Background and motivation

Picophytoplankton, comprising three major groups, *Prochlorococcus*, *Synechococcus*, and picoeukaryotic phytoplankton, is the smallest sized and most abundant phytoplankton component on Earth (Fuhrman and Campbell 1998; Partensky et al. 1999). These groups are dominant in the oligotrophic ocean (Flombaum et al. 2013), but cover all ocean environments (Raven 1998; Partensky et al. 1999; Scanlan et al. 2009) and are found from surface waters down to 150–200 m depth (Flombaum et al. 2013). Picophytoplankton is thought to contribute at least 10% of global net primary production that largely stays in the photic layer because of small cell size and high buoyancy (Raven 1998; Le Quéré et al. 2005), yet, new evidence suggested that picophytoplankton contribution to carbon export may be larger than expected (Stukel et al. 2013; Guidi et al. 2016). Thus, considering picophytoplankton in Earth System Models may improve our understanding of global biogeochemical cycles (Hood et al. 2006).

Biogeochemical models represent the complex structure of ecosystems grouping organisms into functional types and connecting these biological entities with chemical, biological, and physical processes. Phytoplankton is usually represented by few cell size classes (two or three) with different effects on nutrient concentrations, which may result in inaccurate biogeochemical processes (Gruber and Doney 2010). This reduced complexity allows high control and understanding of marine biogeochemical processes; however, highly simplistic representations may provide limited insights (Gruber and Doney 2010; Kwiatkowski et al. 2014) that should be considered during the analysis of results (Emerson and Hedges 2008).

Phytoplankton diversity affects the marine cycles of carbon, nitrogen, and phosphorus through distinct metabolic needs and amount of carbon exported to the deep ocean (Weber and Deutsch 2010). For example, the variability of C/N/P ratios showed a strong latitudinal pattern driven in part by the biological diversity of plankton assemblages, and in the nutrient-depleted subtropical North Atlantic Ocean, the high elemental ratios were partly explained by the higher ratios in marine *Prochlorococcus* and *Synechococccus* that dominated this region (Martiny et al. 2013*a*). Furthermore, in the future, ocean warming and reduced nutrients are expected to benefit *Prochlorococcus* and *Synechococccus*, at the expense of bigger cell size groups, increasing their domains and altering, in turn, carbon export (Flombaum et al. 2013; Martiny et al. 2013*a*). Yet, *Prochlorococcus* and *Synechococccus* high C/P ratios could in part smooth the effects of cell size decrease in carbon export (Martiny et al. 2013*a*). Therefore, biogeochemical models may be improved by more explicitly considering the biological uniqueness of phytoplankton groups (Martiny et al. 2013*b*; Flombaum et al. 2020).

To incorporate *Prochlorococcus*, *Synechococcus*, and picoeukaryotic phytoplankton in ocean models, it is key to understand which environmental variables shape their biogeography. Patterns for picophytoplankton can be obtained from in situ observations, remote sensing, and quantitative niche models. Regional in situ observations provided detailed information on abundance and associated environmental variables as nutrients (Guo et al. 2014), temperature (Morán et al. 2010), and light (Malmstrom et al. 2010), and described seasonal and spatial distribution patterns (Blanchot et al. 2001; Guo et al. 2014; Amorim et al. 2016). Previous global in situ data compilation for picophytoplankton specified information on flow cytometry cell counts, but lacked data on environmental variables (Buitenhuis et al. 2012). Remote sensing methods, based on chlorophyll *a* and other pigments, were used to distinguish *Prochlorococcus* and *Synechococcus* from other dominant groups (Alvain et al. 2005), providing real-time measurements at large scales for the ocean surface, and were also used to estimate the size distribution of phytoplankton, including picophytoplankton in the water column (Uitz et al. 2006; Lange et al. 2018). Still, since some accessory pigments are shared by different taxonomic groups, and organisms may cover a wide range of sizes, the pigment-based approach may have been inaccurate to reflect the actual phytoplankton community structure (Bracher et al. 2017; Mouw et al. 2017). Instead, quantitative niche models, based on the relationship between abundance and environmental variables, identified photosynthetic active radiation (PAR), temperature, and nutrients as predictive variables for picophytoplankton cell abundance (Morán et al. 2010; Flombaum et al. 2013, 2020). Thus, quantitative niche models for *Prochlorococcus*, *Synechococcus*, and picoeukaryotic phytoplankton, plus inputs of PAR, temperature and nutrients, can be used as an alternative to estimate their mean abundances, global scale patterns, and to project changes in future climates (Flombaum et al. 2013, 2020).

To understand the role of picophytoplankton in the Earth system, it is key to address existing data gaps in observations and modeling. Here, we contribute with an in situ observations dataset that includes environmental information that was used to derive quantitative niche models (Flombaum et al. 2013, 2020), a novel modeled monthly climatology dataset, and a novel set of modelled projections in future climates for *Prochlorococcus*, *Synechococcus*, and picoeukaryotic phytoplankton.

Data description

Our dataset is integrated by three independent global datasets that comprise past observations, mean present estimates, and future projections of picophytoplankton principal groups: *Prochlorococcus*, *Synechococcus*, and picoeukaryotic phytoplankton. The first dataset comprises 41,912 in situ observations of *Prochlorococcus*, 44,949 of *Synechococcus*, and 22,184 of picoeukaryotic phytoplankton along with in situ measurements of environmental variables, distributed in major ocean basins (Fig. 1). Data cover the latitudinal range from 81°N to 69°S, down to a depth of 400 m, presenting 80% of the observations in the northern hemisphere and regions with intensive research efforts (North Atlantic Ocean, North Pacific Ocean). Although most of the observations of *Prochlorococcus*, *Synechococcus*, and picoeukaryotic phytoplankton were from tropical regions (53%, 51%, and 69%, respectively, located between 35°N and S) temperature range was well represented with 42%, 45%, and 30% of samples below 15°C (Fig. 1b).

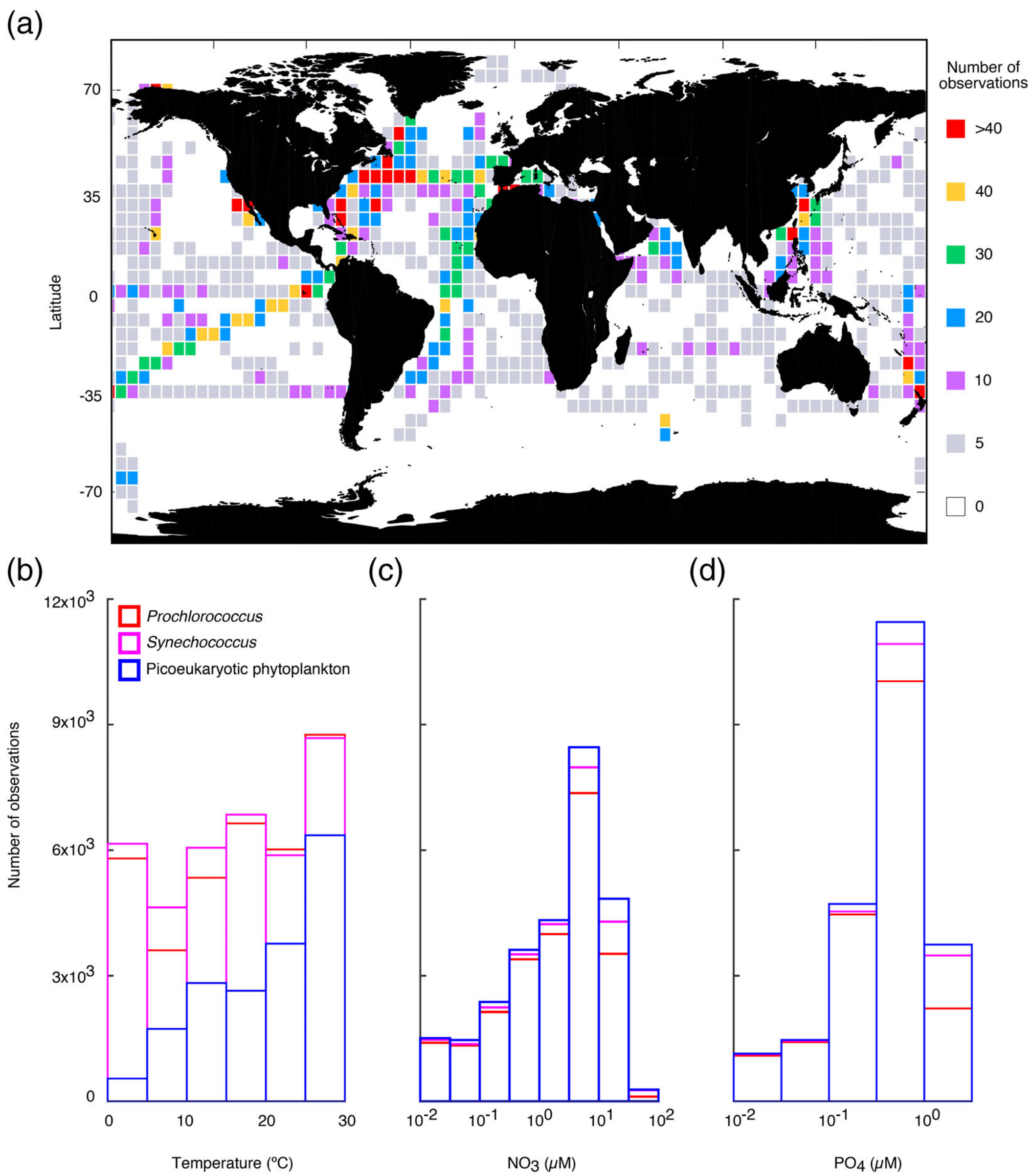

Fig. 1. Geographic distribution and associated ancillary data for in situ observations. Sampling sites and number of observations in 5° grids (a). Number of observations for temperature (b), nitrate (c), and phosphate (d) associated with each group.

The second dataset is a monthly global climatology of *Prochlorococcus*, *Synechococcus*, and picoeukaryotic phytoplankton abundances predicted in cell counts for a 1° grid from the surface down to 200 m depth. Using quantitative niche models and the monthly mean of temperature, nitrate and PAR as inputs, we generated the climatology dataset for each group (Fig. 2). The mean annual concentration showed three areas of elevated concentration (highest 10%): a band between 20°N and S where *Prochlorococcus* and *Synechococcus* dominated, two bands around 40°N and S where *Synechococcus* and picoeukaryotic phytoplankton dominated, and an area above 50°N and S where picoeukaryotic phytoplankton dominated (Fig. 3d–f). Global abundance seasonal variation ranked for *Prochlorococcus*, *Synechococcus*, and picoeukaryotic phytoplankton in increasing amplitude, reaching 6.5%, 15.6%, and 28.0% difference between highest and lowest abundance, respectively (Fig. 4a).

The third dataset synthesized distributions and global abundances for the three groups from 1901 to 2100 (Figs. 3d–f and 4b–d). Similarly, projected cell abundance was the combination of quantitative niche models and inputs from the climate scenarios defined in the Coupled Model Intercomparison Project (CMIP5) (Fig. 2). Major projected changes included increase concentration in tropical areas and expansion towards higher latitudes (Fig. 4d–f). As a whole, the three datasets represented observations, baseline conditions, and modeled projections for the three main groups constituting picophytoplankton.

Data record 1

The variables and units for the in situ observations dataset are listed in Table 1.

The dataset file "global-Prochlorococcus-Synechococcus-and-picoeukaryotic-phytoplankton" (February, 2020 version) was uploaded to Biological and Chemical Oceanography Data Management Office (BCO-DMO, https://www.bco-dmo.org/) in csv format. A header containing all the fields listed in table 1 is available in the data file.

Data record 2

The variables and units for the monthly climatology dataset are listed in Table 2.

Each file is named after a group; every file follows the same name pattern. The name indicates the group as PRO for *Prochlorococcus*, as SYN for *Synechococcus*, and as PEUK for picoeukaryotic phytoplankton, followed by the data record name and month number (i.e., "PRO_climatology_01" indicates the file contains the global abundance climatology of *Prochlorococcus* for January). Annual mean is also identified with the group's prefix.

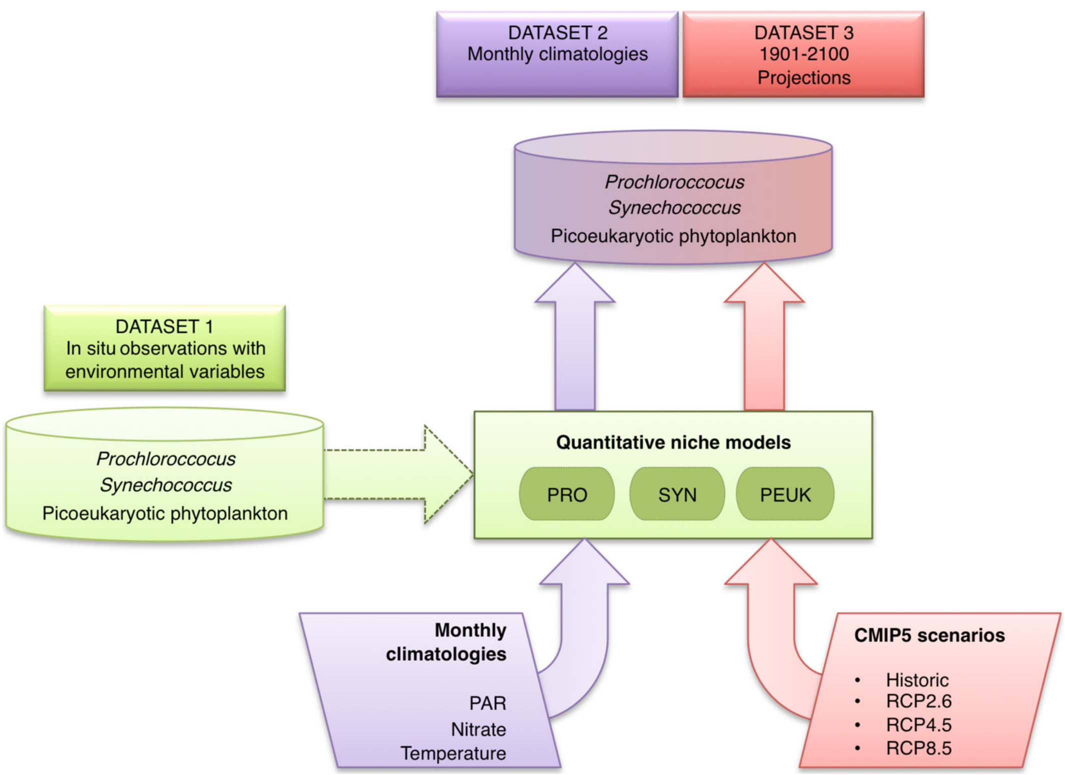

Fig. 2. Flow diagram showing the three datasets: in situ (dataset 1), climatology (dataset 2) and projections for climate change (dataset 3). The in situ dataset was used to establish niche models for each group (Flombaum et al. 2013, 2020). We generated a monthly climatology dataset using PAR, temperature, and nitrate climatologies as inputs for the niche models. Similarly, we obtained the projections for four climates scenarios using nitrate and temperature from an ensemble of five global circulation models.

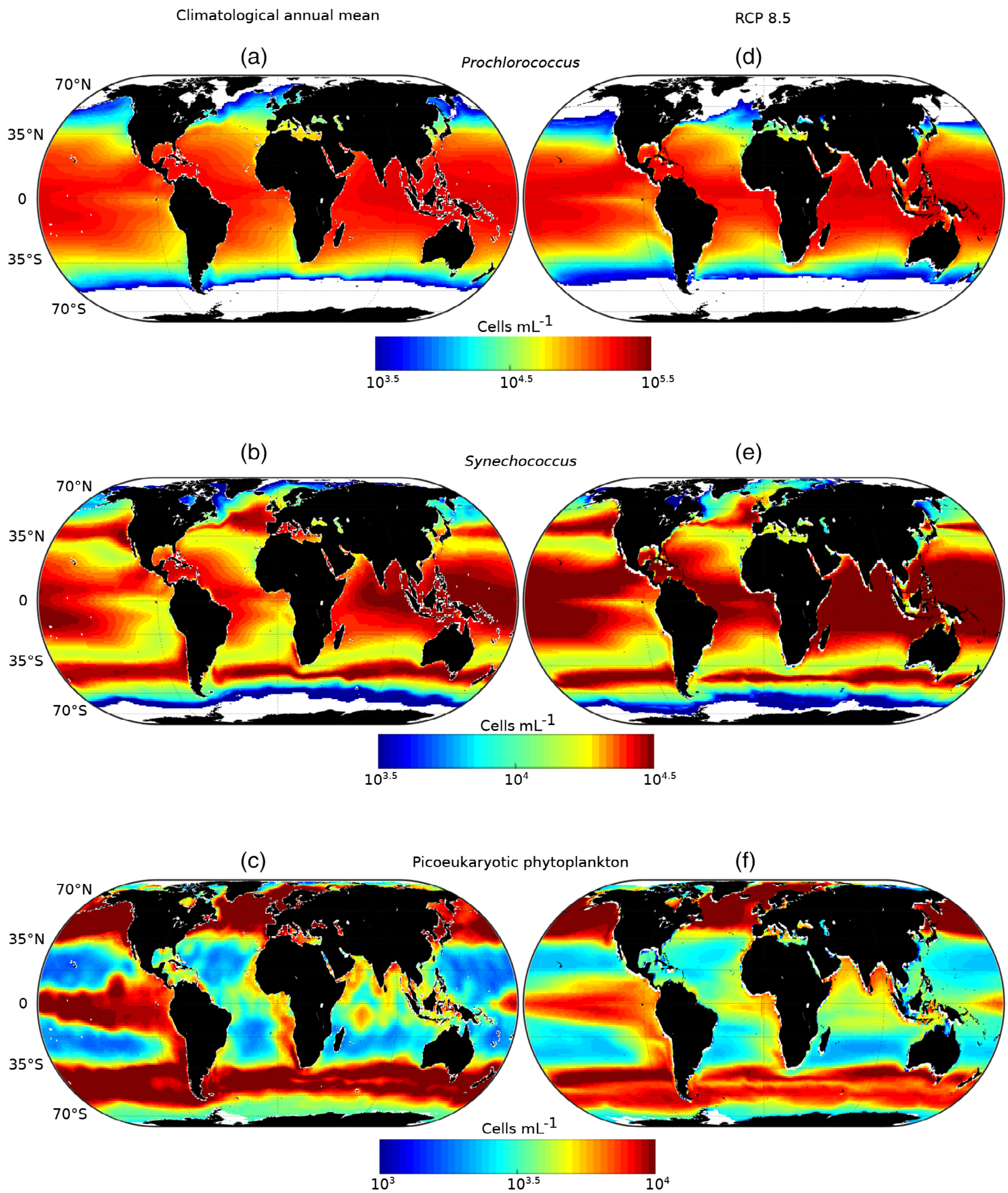

Fig. 3. Estimated abundance for the shallower 50 m of the water column for *Prochlorococcus* (a, d), *Synechococcus* (b, e), and picoeukaryotic phytoplankton (c, f). Maps on the right are the mean of the five models for the CMIP5 scenario RCP 8.5 for the year 2100.

Each file is a matrix of 41,088 rows and 27 columns, which holds the global ocean abundance in units of cells mL^{-1} for each grid location (1° grid) and depth. The first column is latitude, the second is longitude and then each column corresponds to the abundance concentration from 0 to 200 m, following the same depth bins as the World Ocean Atlas (WOA, https://www.nodc.noaa.gov/OC5/woa13/), with 5 m intervals for the first hundred meters and 25 m intervals for the next hundred.

The dataset files (May, 2020 version) in csv format were uploaded to the Biological and Chemical Oceanography Data

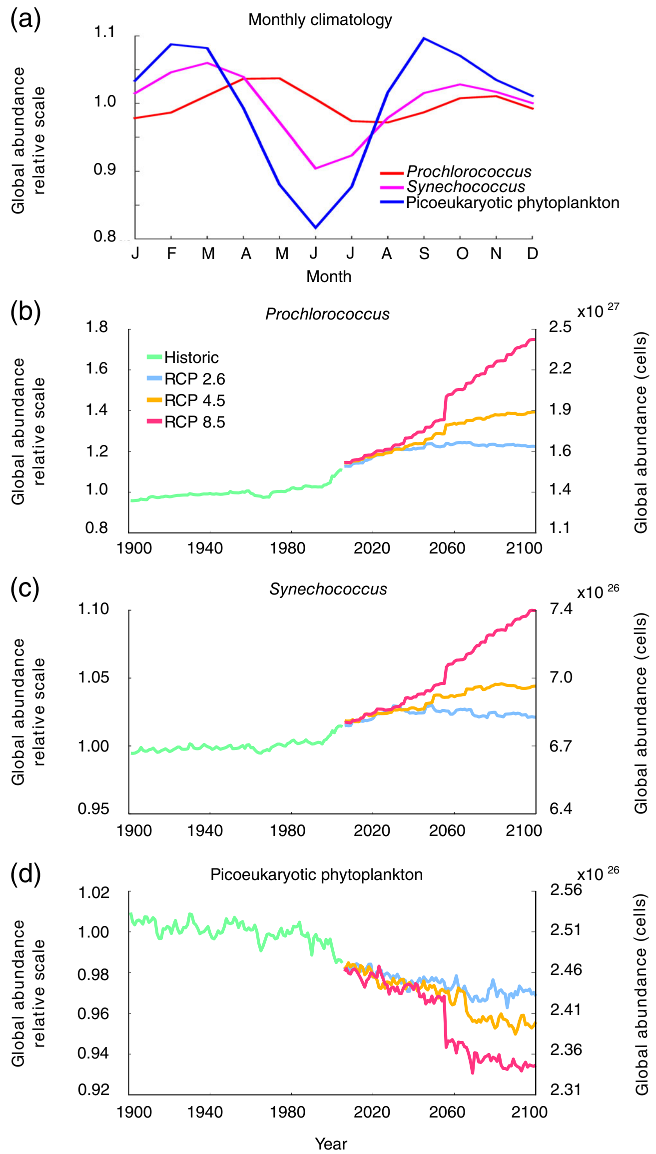

Fig. 4. Global abundance temporal variation for major groups of picophytoplankton. Global monthly abundance (a), and the historic and three future climate scenarios projections for *Prochlorococcus* (b), *Synechococcus* (c), and picoeukaryotic phytoplankton (d).

Management Office (BCO-DMO). A header containing all the fields listed in table 2 is available in each data file. The zip file named "picophytoplankton_climatology" contains all the files of this data record.

Data record 3

The climate change dataset variables and units are listed in Table 3.

This data record is composed of three files: "pro-syn-peuk-cc-global-abundance-mean," "pro-syn-peuk-cc-global-abundance-std," and "pro-syn-peuk-cc-surface." The first file holds the abundance yearly mean in cells units, for each group and climate scenario tested with five global circulation models. The matrix is of two dimensions arranged in 105 rows and 62 columns. Each row corresponds to a different year, and columns are related to group, scenario, and tested model. The first column shows the years (1901–2005) for the historic scenario and the second column indicates the years (2006–2100)

Table 1. Variables and units in the in situ observations dataset.

Variable	Unit
Latitude	−90 to 90
Longitude	0–360
Year	YYYY
Day	Julian
Temperature	°C
Nitrate and nitrite	μM
Phosphate	μM
Depth	m
Prochlorococcus	cells mL^{-1}
Synechococcus	cells mL^{-1}
Picoeukaryotic phytoplankton	cells mL^{-1}

Table 2. Variables and units in the climatology dataset.

Variable	Unit
Latitude	−90 to 90
Longitude	0–360
Depth	0–200 m
Prochlorococcus	cells mL^{-1}
Synechococcus	cells mL^{-1}
Picoeukaryotic phytoplankton	cells mL^{-1}

Table 3. Variables and units in climate change dataset.

Variable	Unit
Latitude (for surface file)	−90 to 90
Longitude (for surface file)	0–360
Year (abundance mean and SD files)	From 1901 to 2005 and from 2006 to 2100
Prochlorococcus	cells (cells mL^{-1} for surface file)
Synechococcus	cells (cells mL^{-1} for surface file)
Picoeukaryotic phytoplankton	cells (cells mL^{-1} for surface file)
Scenarios	Historic, RCP2.6, RCP4.5, RCP8.5
Models	GFDL-ESM2G, HadGEM2-ES, IPSLCM5A-MR, MPI-ESM-LR, NorESM1-ME

for scenarios RCP 2.6, RCP 4.5, and RCP 8.5. The next 20 columns correspond to *Prochlorococcus* abundance for each scenario tested with every model, and it then continues in the same order for *Synechococcus* and picoeukaryotic phytoplankton. The second file holds the abundance standard deviation in cell units arranged in an identical way.

The file "pro-syn-peuk-cc-surface" contains a 2D-matrix of global surface abundance (50 m) and distribution for each group, scenario and circulation model for the last 30 years of the 21st century. The first column is latitude and the second is longitude. Subsequent columns are arranged in an identical way to previous files in this record.

The dataset files (February, 2020 version) in csv format were uploaded to the Biological and Chemical Oceanography Data Management Office (BCO-DMO). Headers containing all the fields listed in table 3 are available in each file.

Methods

Observations data record (data citation 1)

We compiled a database for *Prochlorococcus*, *Synechococcus*, and picoeukaryotic phytoplankton cell counts along with environmental data (Martiny and Flombaum 2020*a*). In situ observations from public repositories and previous studies are listed in Supplementary Table 1. Environmental data included temperature, nitrate plus nitrite and phosphate, and ancillary information on depth, year, and Julian day.

We considered flow cytometry counts for the three groups, and included microscope counts for *Synechococcus*. No standardization was attempted. We did not consider *Prochlorococcus* microscopy cell counts because of their weak autofluorescence. We imposed a nitrate minimum of 10^{-2} μM to avoid issues with detection limits.

Climatology data record (data citation 2)

We generated a monthly climatology for *Prochlorococcus*, *Synechococcus*, and picoeukaryotic phytoplankton cell abundance for the global ocean in a 1° grid from 0 to 200 m depth (Visintini et al. 2020). We used niche models established for each group (Flombaum et al. 2013, 2020) together with PAR and temperature, for *Prochlorococcus* and *Synechococcus*, plus nitrate, for picoeukaryotic phytoplankton. The explained variances (R^2) of the quantitative niche models of *Prochlorococcus*, *Synechococcus*, and picoeukaryotic phytoplankton were 0.66, 0.35, and 0.46, respectively (Flombaum et al. 2013, 2020).

Downwelling irradiance was based on the climatologies of PAR (NASA Goddard Space Flight Center Ocean Ecology Laboratory Ocean Biology Processing Group 2018*a*) and K_{d490} diffuse attenuation coefficient (Modis-Aqua 0.083° grid) (NASA Goddard Space Flight Center Ocean Ecology Laboratory Ocean Biology Processing Group 2018*b*), and then averaged to fit a 1° grid. Monthly global temperature and nitrate statistical means were obtained from WOA for 1° grids and 200 m of water column depth (Garcia et al. 2013; Levitus et al. 2015; Locarnini et al. 2013).

We did not generate abundance data for ranges of PAR, temperature, and nutrient concentration that fell outside niche model boundaries. Boundaries for PAR, temperature and nitrate were 10^{-4} and $10^{1.8}$ E m^{-2} d^{-1}, 0 and 30°C, and 10^{-2} μM with no upper threshold value, respectively.

Climate change data record (data citation 3)

We projected *Prochlorococcus*, *Synechococcus*, and picoeukaryotic phytoplankton abundance (Martiny and Flombaum 2020*b*,*c*,*d*) in climate change scenarios (Historic, RCP 2.6, RCP 4.5, and RCP 8.5) using five global circulation models (GFDL-ESM2G, HadGEM2-ES, IPSL-CM5A-MR, MPI-ESM-LR, and NorESM1-ME) included in the Coupled Model Intercomparison Project (Taylor et al. 2012).

The global circulation models provided yearly temperature and nitrate concentration, while we considered PAR constant along time. PAR boundaries for *Prochlorococcus* and *Synechococcus* niche models were 10^{-4} and $10^{1.8}$ E m^{-2} d^{-1}, and 10^{-3} and $10^{1.8}$E m^{-2} d^{-1} for the picoeukaryotic phytoplankton model. Global abundance represented the sum of cells number in the entire ocean, accounting for differences in grid size from low to high latitudes. Sea surface cell concentration represented the annual average of the first 50 m per grid.

Technical validation

For the in situ observations dataset, all available values were considered. For the climatology dataset we do not present abundance data below 1000 cells mL^{-1} for *Prochlorococcus* and *Synechococcus* and below 500 cells mL^{-1} for picoeukaryotic phytoplankton as we rounded abundances to those significant figures. In the climate change datasets, note that from years 2056 to 2065 the scenario RCP 4.5 does not contain data. For all datasets NaN and nd describe no data.

Data use and recommendations for reuse

Datasets can be used to parameterize and validate biogeochemical models that consider this group of phytoplankton, and to calibrate or support remote sensing approaches meant to detect major groups of picophytoplankton. Within the in situ observations dataset, there are time-series observations that can be used to analyze abundance temporal variations for some coordinates, while the climatology dataset enables the evaluation of the global abundance variability of *Prochlorococcus*, *Synechococcus*, and picoeukaryotic phytoplankton. The climatology dataset can be used as a reference for in situ observations at a regional scale as well. The dataset for future climates can be used to summarize projected changes and to contrast with other methods of projections, such as dynamic ecosystem models.

Data and metadata are available at https://doi.org/10.1575/1912/bco-dmo.793451.1, https://doi.org/10.26008/1912/bco-dmo.811147.1, https://doi.org/10.26008/1912/bco-dmo.793690.1, https://doi.org/10.26008/1912/bco-dmo.793776.1 and https://doi.org/10.26008/1912/bco-dmo.793847.1

References

Alvain, S., C. Moulin, Y. Dandonneau, and F.-M. Bréon. 2005. Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. Deep Sea Res. Part I Oceanogr. Res. Pap. **52**: 1989–2004. doi:10.1016/j.dsr.2005.06.015.

Amorim, A. L., P. León, J. M. Mercado, D. Cortés, F. Gómez, S. Putzeys, S. Salles, and L. Yebra. 2016. Controls of picophytoplankton abundance and composition in a highly dynamic marine system, the Northern Alboran Sea (Western Mediterranean). J. Sea Res. **112**: 13–22. doi:10.1016/j.seares.2016.02.005.

Blanchot, J., J. M. André, C. Navarette, J. Neveux, and M. H. Radenac. 2001. Picophytoplankton in the equatorial Pacific: Vertical distributions in the warm pool and in the high nutrient low chlorophyll conditions. Deep. Res. Part I Oceanogr. Res. Pap. **48**: 297–314. doi:10.1016/S0967-0637(00)00063-7.

Bracher, A., and others. 2017. Obtaining phytoplankton diversity from ocean color: A scientific roadmap for future development. Front. Mar. Sci. **4**: 1–15. doi:10.3389/fmars.2017.00055.

Buitenhuis, E. T., and others. 2012. Picophytoplankton biomass distribution in the global ocean. Earth Syst. Sci. Data **4**: 37–46. doi:10.5194/essd-4-37-2012.

Emerson, S. R., and J. I. Hedges. 2008, *Chemical oceanography and the marine carbon cycle*. Cambridge University Press. doi:10.1017/CBO9780511793202.

Flombaum, P., and others. 2013. Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. Proc. Natl. Acad. Sci. USA **110**: 9824–9829. doi:10.1073/pnas.1307701110.

Flombaum, P., W.-L. Wang, F. W. Primeau, and A. C. Martiny. 2020. Global picophytoplankton niche partitioning predicts overall positive response to ocean warming. Nat. Geosci. **13**: 116–120. doi:10.1038/s41561-019-0524-2.

Fuhrman, J. A., and U. Campbell. 1998. Microbial microdiversity. Nature **393**: 410–411. doi:10.1038/30839.

Garcia, H. E., R. A. Locarnini, T. P. Boyer, J. I. Antonov, O. K. Baranova, M. M. Zweng, J. R. Reagan, and D. R. Johnson. 2013. World ocean atlas 2013. Volume 4: Dissolved inorganics nutrients (phosphate, nitrate, silicate) S. Levitus and A. Mishonov [eds.]. NOAA Atlas NESDIS 76: 25.

Gruber, N., and S. C. Doney. 2010. Ocean biogeochemistry and ecology, modeling of. Encycl. Ocean Sci. **2**: 89–104. doi:10.1016/B978-012374473-9.00741-4.

Guidi, L., and others. 2016. Plankton networks driving carbon export in the oligotrophic ocean. Nature **532**: 465–470. doi:10.1038/nature16942.

Guo, C., H. Liu, L. Zheng, S. Song, B. Chen, and B. Huang. 2014. Seasonal and spatial patterns of picophytoplankton growth, grazing and distribution in the East China Sea. Biogeosciences **11**: 1847–1862. doi:10.5194/bg-11-1847-2014.

Hood, R. R., and others. 2006. Pelagic functional group modeling: Progress, challenges and prospects. Deep. Res. Part II Top. Stud. Oceanogr. **53**: 459–512. doi:10.1016/j.dsr2.2006.01.025.

Kwiatkowski, L., and others. 2014. iMarNet: An ocean biogeochemistry model intercomparison project within a common physical ocean modelling framework. Biogeosciences **11**: 7291–7304. doi:10.5194/bg-11-7291-2014.

Lange, P., R. Brewin, G. Dall'Olmo, G. A. Tarran, S. Sathyendranath, M. V. Zubkov, and H. Bouman. 2018. Scratching beneath the surface: A model to predict the vertical distribution of *Prochlorococcus* using remote sensing. Remote Sens. (Basel) **10**: 847. doi:10.3390/rs10060847.

Levitus Sydney, Boyer Tim P., Garcia Hernan E., Locarnini Ricardo A., Zweng Melissa M., Mishonov Alexey V., Reagan James R., Antonov John I., Baranova Olga K., Biddle Mathew, Hamilton Melanie, Johnson Daphne R., Paver Christopher R., Seidov Dan. 2015. World Ocean Atlas 2013 (NCEI Accession 0114815). Temperature and nutrients datasets. NOAA National Centers for Environmental Information. [accessed on 2017 April 24]. doi:10.7289/v5f769gt

Locarnini, R. A., and others. 2013. World ocean atlas 2013. Vol. 1: Temperature. S. Levitus and A. Mishonov [eds.]. NOAA Atlas NESDIS 73: 40.

Malmstrom, R. R., A. Coe, G. C. Kettler, A. C. Martiny, J. Frias-Lopez, E. R. Zinser, and S. W. Chisholm. 2010. Temporal dynamics of *Prochlorococcus* ecotypes in the Atlantic and Pacific oceans. ISME J. **4**: 1252–1264. doi:10.1038/ismej.2010.60.

Martiny, A. C., and P. Flombaum. 2020*a*. Global observations *Prochlorococcus*, *Synechococcus*, and picoeukaryotic phytoplankton with ancillary environmental data from 1987 to 2008. Biological and Chemical Oceanography Data Management Office. doi:10.1575/1912/bco-dmo.793451.1

Martiny, A. C., and P. Flombaum. 2020*b*. *Prochlorococcus*, *Synechococcus*, and picoeukaryotic phytoplankton yearly mean global abundance for four CMIP5 climate scenarios using an ensemble of five circulation models. Biological and Chemical Oceanography Data Management Office. doi:10.26008/1912/bco-dmo.793690.1

Martiny, A. C., and P. Flombaum. 2020*c*. *Prochlorococcus*, *Synechococcus*, and picoeukaryotic phytoplankton yearly standard deviation from mean global abundance estimated for four CMIP5 climate scenarios using an ensemble of five circulation models. Biological and Chemical Oceanography Data Management Office. doi:10.26008/1912/bco-dmo.793776.1

Martiny, A. C., and P. Flombaum. 2020*d*. *Prochlorococcus*, *Synechococcus*, and picoeukaryotic phytoplankton for four future climatic scenarios and five Earth System Models (cell/ml) in a global 1x1 grid for the ocean surface (50m). Biological and Chemical Oceanography Data Management Office. doi:10.26008/1912/bco-dmo.793847.1

Martiny, A. C., C. T. A. A. Pham, F. W. Primeau, J. A. Vrugt, J. K. Moore, S. A. Levin, and M. W. Lomas. 2013*a*. Strong latitudinal patterns in the elemental ratios of marine plankton and organic matter. Nature **6**: 279–283. doi:10.1038/ngeo1757.

Martiny, A. C., J. A. Vrugt, F. W. Primeau, and M. W. Lomas. 2013*b*. Regional variation in the particulate organic carbon to nitrogen ratio in the surface ocean. Global Biogeochem. Cycles **27**: 723–731. doi:10.1002/gbc.20061.

Morán, X. A. G., Á. López-Urrutia, A. Calvo-Díaz, and W. K. W. LI. 2010. Increasing importance of small phytoplankton in a warmer ocean. Glob. Chang. Biol. **16**: 1137–1144. doi:10.1111/j.1365-2486.2009.01960.x.

Mouw, C. B., and others. 2017. A consumer's guide to satellite remote sensing of multiple phytoplankton groups in the global ocean. Front. Mar. Sci. **4**: 41. doi:10.3389/fmars.2017.00041.

NASA Goddard Space Flight Center Ocean Ecology Laboratory Ocean Biology Processing Group. 2018*a*. Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua. Level3 Mapped Photosynthetically Available Radiation Data. NASA OB.DAAC, Greenbelt, MD, USA. [accessed 2017 September 18]. doi:10.5067/aqua/modis/l3m/par/2018.

NASA Goddard Space Flight Center Ocean Ecology Laboratory Ocean Biology Processing Group. 2018*b*. Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua. Level 3 Mapped Downwelling Diffuse Attenuation Coefficient Data. NASA OB.DAAC, Greenbelt, MD, USA. [accessed on 2017 September 18]. doi:10.5067/aqua/modis/l3m/kd/2018

Partensky, F., W. R. Hess, and D. Vaulot. 1999. Prochlorococcus, a marine photosynthetic prokaryote of global significance. Microbiol. Mol. Biol. Rev. **63**: 106–127.

Le Quéré, C., and others. 2005. Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. Glob. Chang. Biol. **11**: 2016–2040. doi:10.1111/j.1365-2486.2005.1004.x.

Raven, J. A. 1998. The twelfth Tansley lecture. Small is beautiful: The picophytoplankton. Funct. Ecol. **12**: 503–513. doi:10.1046/j.1365-2435.1998.00233.x.

Scanlan, D. J., and others. 2009. Ecological genomics of marine Picocyanobacteria. Microbiol. Mol. Biol. Rev. **73**: 249–299. doi:10.1128/MMBR.00035-08.

Stukel, M. R., M. Décima, K. E. Selph, D. A. A. Taniguchi, and M. R. Landry. 2013. The role of Synechococcus in vertical flux in the Costa Rica upwelling dome. Prog. Oceanogr. **112–113**: 49–59. doi:10.1016/j.pocean.2013.04.003.

Taylor, K. E., R. J. Stouffer, G. A. Meehl, K. E. Taylor, R. J. Stouffer, and G. A. Meehl. 2012. An overview of CMIP5 and the experiment design. Bull. Am. Meteorol. Soc. **93**: 485–498. doi:10.1175/BAMS-D-11-00094.1.

Uitz, J., H. Claustre, A. Morel, and S. B. Hooker. 2006. Vertical distribution of phytoplankton communities in open ocean: An assessment based on surface chlorophyll. J. Geophys. Res. **111**: C08005. doi:10.1029/2005JC003207.

Visintini, N., A. C. Martiny, and P. Flombaum. 2020. Prochlorococcus, Synechococcus and picoeukaryotic phytoplankton abundance climatology in the global ocean from quantitative niche models. Biological and Chemical Oceanography Data Management Office. [accessed 2020 May 19]. doi:10.26008/1912/bco-dmo.811147.1

Weber, T., and C. Deutsch. 2010. Ocean nutrient ratios governed by plankton biogeography. Nature **467**: 550–554. doi:10.1038/nature09403.

Acknowledgments

We thank the OBPG at NASA and the many contributing researchers for the oceanographic data and Coupled Model Intercomparison Project (CMIP5) model outputs. We would also like to thank M. Florio and the anonymous reviewers for comments that greatly improved this typescript. Financial support for this work was provided by Universidad de Buenos Aires UBACyT 20020170100620BA and Agencia Nacional de Promoción Científica y Tecnológica PICT-2017-3020 (PF), and US National Science Foundation OCE-1559002 and OCE-1848576 (ACM). NV was supported by the doctoral Scholarship programs of Universidad de Buenos Aires and Doctoral Scholarship Program of Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

Conflict of interest

The authors declared no potential conflicts of interest.

Submitted 25 June 2020
Revised 28 January 2021
Accepted 12 February 2021