Present and future global distributions of the marine Cyanobacteria Prochlorococcus and Synechococcus

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1073/pnas.1307701110">http://dx.doi.org/10.1073/pnas.1307701110</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>National Academy of Sciences (U.S.)</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Tue Nov 13 14:58:08 EST 2018</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/88271">http://hdl.handle.net/1721.1/88271</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher’s policy and may be subject to US copyright law. Please refer to the publisher’s site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
Present and future global distributions of the marine Cyanobacteria Prochlorococcus and Synechococcus

Pedro Flombaum\textsuperscript{a,b}, José L. Gallegos\textsuperscript{a}, Rodolfo A. Gordillo\textsuperscript{a}, José Rincón\textsuperscript{a}, Lina L. Zabala\textsuperscript{b}, Nianzhi Jiao\textsuperscript{c}, David M. Karl\textsuperscript{d,e}, William K. W. Li\textsuperscript{a}, Michael W. Lomas\textsuperscript{a}, Daniele Veneziano\textsuperscript{a}, Carolina S. Vera\textsuperscript{b}, Jasper A. Vrugt\textsuperscript{a,h}, and Adam C. Martiny\textsuperscript{a,i,1}

Departments of \textsuperscript{a}Earth System Science, \textsuperscript{b}Civil Engineering, and \textsuperscript{c}Ecology and Evolutionary Biology, University of California, Irvine, CA 92697; \textsuperscript{d}Centro de Investigaciones del Mar y la Atmósfera, Departamento de Ciencias de la Atmósfera y los Océanos, and Instituto Franco Argentina sobre Estudios del Clima y sus Impactos, Consejo Nacional de Investigaciones Científicas y Tecnológicas y Universidad de Buenos Aires, 1428 Buenos Aires, Argentina; \textsuperscript{e}Institute of Microbes and Ecosphere, State Key Lab for Marine Environmental Sciences, Xiamen University, Xiamen 361005, People’s Republic of China; \textsuperscript{f}Center for Microbial Oceanography: Research and Education (C-MORE), University of Hawaii, Honolulu, HI 96822; \textsuperscript{g}Fisheries and Oceans Canada, Bedford Institute of Oceanography, Dartmouth, NS, Canada B2Y 4A2, \textsuperscript{h}Bigelow Laboratory for Ocean Sciences, East Boothbay, ME 04544; and \textsuperscript{i}Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139

Contributed by David M. Karl, April 25, 2013 (sent for review January 22, 2013)

The Cyanobacteria Prochlorococcus and Synechococcus account for a substantial fraction of marine primary production. Here, we present quantitative niche models for these lineages that assess present and future global abundances and distributions. These niche models are the result of neural network, nonparametric, and parametric analyses, and they rely on >35,000 discrete observations from all major ocean regions. The models assess cell abundance based on temperature and photosynthetically active radiation, but the individual responses to these environmental variables differ for each lineage. The models estimate global biogeographic patterns and seasonal variability of cell abundance, with maxima in the warm oligotrophic gyres of the Indian and the western Pacific Oceans and minima at higher latitudes. The annual mean global abundances of Prochlorococcus and Synechococcus are $2.9 \pm 0.1 \times 10^{27}$ and $7.0 \pm 0.3 \times 10^{25}$ cells, respectively. Using projections of sea surface temperature as a result of increased concentration of greenhouse gases at the end of the 21st century, our niche models projected increases in cell numbers of 29% and 14% for Prochlorococcus and Synechococcus, respectively. The changes are geographically uneven but include an increase in area. Thus, our global niche models suggest that oceanic microbial communities will experience complex changes as a result of projected future climate conditions. Because of the high abundances and contributions to primary production of Prochlorococcus and Synechococcus, these changes may have large impacts on ocean ecosystems and biogeochemical cycles.

Results

Global Distributions of Cyanobacteria. To predict the global distributions of Prochlorococcus and Synechococcus, we first compiled a global dataset. The dataset consisted of 35,129 and 57,699 observations of abundance, respectively, and contained data from a total of 103 cruises and time series with samples from all major ocean regions (Fig. L4 and Table S1). We also compiled additional data on habitat temperature, light [described as photosynthetically active radiation (PAR)], and nutrients (i.e., nitrate and phosphate concentration). This extensive dataset covered nearly the full environmental range of these variables (Fig. 1B) as well as their combinations (Fig. S1). We then applied a three-step approach to formulate quantitative niche models (Fig. S2). First, we identified the environmental factors that explained most of the variance using artificial neural network (ANN) models. Temperature and PAR were found to be important for predicting both Prochlorococcus and Synechococcus (ANN, $R^2_{\text{PAR}} = 0.66$ and $R^2_{\text{syn}} = 0.35$, whereas nutrients explained little additional

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

\textsuperscript{1}To whom correspondence may be addressed. E-mail: dkarl@hawaii.edu or amartiny@uci.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1307701110/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1307701110

PNAS Early Edition | 1 of 6
variance (Table S2). Second, we predicted cell abundance for each combination of temperature and PAR using nonparametric local regression models and found that the performances were similar to the ANN models ($R^2_{Pro} = 0.68$ and $R^2_{Syn} = 0.42$). Third, we closely reproduced the nonparametric predictions using more robust parametric regression models (Fig. S3A). The parametric models for Prochlorococcus and Synechococcus explained a high fraction of the total variances based on temperature and PAR ($R^2_{Pro} = 0.67$ and $R^2_{Syn} = 0.34$) (Table S3). Thus, the variances of the parametric models were comparable with both the ANN and nonparametric models. A subsequent sensitivity analysis using split sampling showed that the model parameters, $R^2$, and predicted abundance were robust relative to the selection of input data and cruises (Fig. S4A–C). No regional bias was detected for either lineage (Fig. S4D).

Using the parametric models, we then estimated the global quantitative distributions of Prochlorococcus and Synechococcus. On average, Prochlorococcus was most abundant in the warm oligotrophic waters, especially the Indian and western Pacific Ocean subtropical gyres (maximum = $2.5 \times 10^5$ and $2.1 \times 10^5$ cells mL$^{-1}$, respectively) (Fig. 2A), and the highest monthly average value estimated was $2.8 \times 10^5$ cells mL$^{-1}$. The mean abundance declined from $4.4 \times 10^4$ to $2.8 \times 10^3$ cells mL$^{-1}$ from lower latitude waters (30°N to 30°S) to the latitudinal band 30° to 40° (either N or S) and was low at latitudes above 40°. The analysis also suggested lower abundances in areas dominated by cold surface currents and strong upwelling. The median abundance of Prochlorococcus in these areas was $1.0 \times 10^5$ cells mL$^{-1}$, about one order of magnitude lower than in the Indian and western Pacific oligotrophic waters. The niche model predicted that Synechococcus was also most abundant in the Indian and western Pacific Oceans (maximum = $3.4 \times 10^4$ and $4.0 \times 10^4$ cells mL$^{-1}$, respectively), had a peak at mid-latitudes, and had low abundance in the Arctic and Southern Oceans (Fig. 2B). For this lineage, we observed a much smaller

![Fig. 1. Observations of Prochlorococcus and Synechococcus cell abundance. (A) Geographic distribution of samples used in this study. (B) Prochlorococcus and Synechococcus as a function of temperature, PAR, nitrate, and phosphate concentration. Symbol color represents the number of overlapping observations.](image)

![Fig. 2. Present global distribution of Prochlorococcus and Synechococcus abundance. (A) Prochlorococcus and (B) Synechococcus mean annual abundances at the sea surface.](image)
decline in abundance in cold current or upwelling regions. We also found strong seasonal variations in both the distributions and abundances of *Prochlorococcus* and *Synechococcus*. The latitudinal range of cell abundance above 10^7 cells mL^-1 for *Prochlorococcus* shifted 18.0°N and 8.1°S for *Synechococcus* shifted 23.3°N and 6.0°S (Fig. S5). The global mean annual abundance for *Prochlorococcus* and *Synechococcus* was 2.9 ± 0.1 \times 10^7 and 7.0 ± 0.3 \times 10^6 cells, respectively. The highest/lowest global mean monthly abundance was April/August (3.0/2.7 \times 10^7 cells) for *Prochlorococcus* and March/July (7.5/6.4 \times 10^6 cells) for *Synechococcus*.

Next, we estimated sea surface abundances and distributions for the end of the 20th and 21st centuries using climate model projections of sea surface temperatures (18). The projection for the 21st century used the Representative Concentration Pathway (RCP) 4.5 scenario, which was based on a stabilization of the radiative forcing of greenhouse gases at what is equivalent to 650 ppm CO₂. Because of the rise in temperature, the multimodel ensemble mean projected a 28.7% and 13.9% increase in *Prochlorococcus* and *Synechococcus* abundance, respectively. However, projections slightly differed in the total increase with a range of 18.4–35.4% and 9.8–16.4% for *Prochlorococcus* and *Synechococcus*, respectively. We also found that these projected changes were geographically uneven (Fig. 3 A and B). The multimodel ensemble mean for *Prochlorococcus* abundance increased 25.8 ± 7.0% in a band between 12°N and 26°S and 37.9 ± 10.4% up to 40° (Fig. 3A). For *Synechococcus*, the increase was 20.0 ± 4.5% in the tropical band (20°N to 20°S) as well as a slight expansion at high latitudes (Fig. 3B). However, for *Synechococcus*, the biggest changes were observed from ~30°N to 60°N or ~30°S to 60°S. In a zone centered on 35°N or 35°S, we found a decline in abundance. Although most significant in the Pacific Ocean, we also detected a zone centered on 45°N with a large increase in abundance. Projections of the four climate models largely covaried in the increase in cell number within the tropics, and in the poleward shifts of the abundance peak at mid-latitudes (Fig. 3 C and D). In a north–south transect along the Atlantic Ocean, the models depicted two consistent crests around the equator for both lineages (Fig. 3 C and D). For *Synechococcus*, the models also showed a peak around 45° of latitude (Fig. 3D). In this transect, the uncertainty caused by climate model differences for *Prochlorococcus* abundance was larger in the Northern Hemisphere and the Equator, and it decreased toward the southern distribution limit (Fig. 3C). For *Synechococcus*, the uncertainty was associated with differences in the projected abundance and the specific latitude for the secondary peak (Fig. 3D). All four models projected a slight expansion of distribution areas with high cell abundance (here defined as 10^7 cells mL^-1) of 3.3 ± 0.4°N and 2.5 ± 0.2°S for *Prochlorococcus* and 5.6 ± 2.1°N and 2.5 ± 0.3°S for *Synechococcus* (Fig. 3 C and D). Thus, we project an overall increase of both *Prochlorococcus* and *Synechococcus* but with substantial regional variations in cell abundances.

**Prochlorococcus** and **Synechococcus** Niches. We next identified the niches for *Prochlorococcus* and *Synechococcus* along temperature, PAR, and nutrient gradients (Figs. 1B and 4). The response to temperature for *Prochlorococcus* could be divided into three zones. Below 8 °C, *Prochlorococcus* was present in 230 of 7,213 (3.2%) field samples (Fig. S3B), and this low likelihood of observing *Prochlorococcus* was reflected in the niche model (Fig. S6A). Between 8 °C and 13 °C, the probability of finding *Prochlorococcus* increased from 20% to 80%, and the estimated relationship between cell abundance and temperature was the strongest irrespective of PAR (Fig. S6A). Above 13 °C, the probability of finding *Prochlorococcus* increased from 80% to ~100%, and both the observed and estimated abundances changed from 10^7 to 10^10 cells mL^-1 (Fig. 4A and S6A). As expected, the abundance was also strongly related to PAR intensity (Fig. 4B and D).
Synechococcus abundance was also strongly linked to temperature, but the shape of the relationship was different compared with Prochlorococcus (Fig. 4A and C). Synechococcus was largely absent in subzero waters and found to have a maximum in abundance at 10 °C. Above 10 °C, the abundance declined and displayed a local minimum at ~20 °C. Above this temperature, there was a small increase in cell concentration. We observed a nonlinear positive relationship of Synechococcus abundance with PAR (i.e., flat/positive/flat). This relationship was dependent on temperature, because the parametric regression model failed to identify a link between abundance and PAR at low temperatures (Fig. S6B). We also observed detectable cell abundance in deep, low PAR samples (depth > 100 m, mean PAR = 0.15 E m⁻² d⁻¹, mean abundance = 172 cells mL⁻¹) (Fig. S3C). Similar to Prochlorococcus, we did not find a clear relationship between nutrient concentration and Synechococcus abundance after controlling for temperature and PAR (Fig. S6C).

Discussion

Our analyses build on the summed effort of 35 y of research cruises and time series and provide a basis for quantifying the global abundance of Prochlorococcus and Synechococcus within their realized niche. Our dataset differs from previous efforts (21) by having a wider geographic and environmental representation, and it includes multiple sample sets from the eastern and western Pacific, southern Atlantic, and southern Indian Oceans. We also recognize that slightly different techniques have been used to count cells, but the large size of our dataset and our sensitivity analyses suggest that the estimates are robust against such possible biases.

We find mean annual averages of 2.9 × 10²⁷ and 7.0 × 10²⁰ cells for Prochlorococcus and Synechococcus, respectively, which are in good agreement with the near surface estimate of 2.9 × 10²⁷ cells of total prokaryotic autotrophs (22). Our estimates suggest that marine Cyanobacteria constitute about 10% of the total ocean marine picoplankton in the upper 200 m (assuming 3.6 × 10²⁸ cells) (22). We predict the global net primary production of 4 and 8 Gr C y⁻¹ [8.5% and 16.7% of ocean net primary production (1)] for Prochlorococcus and Synechococcus respectively, using estimates of hourly carbon fixation rates (3). We acknowledge a high uncertainty of the global variation in fixation rates, but our results suggest that marine Cyanobacteria could be responsible for ~25% of ocean net primary productivity. This value is within the range of local observations of ocean productivity for the two lineages (23).

We observe complex changes in the abundances and distributions of Prochlorococcus and Synechococcus as a result of the climate change projected by the end of 21st century using the RCP4.5 scenario. As proposed previously (24), we expect an overall increase in cell abundances of Prochlorococcus and Synechococcus. However, the changes in abundance are geographically differentiated. Mid-latitude regions can experience large shifts in community structure, including regions with a large increase in both lineages but also places with a decline in Synechococcus abundance (around 35°N or 35°S). At low latitudes, we expect that both lineages increase their abundance. Our analysis supports an overall increase in Cyanobacteria abundance explained primarily by the increase in areas where Prochlorococcus and Synechococcus are currently abundant and secondarily by the geographical expansion of these lineages. Large model uncertainties in the distribution and abundance of cells for future projections are mostly at high latitudes. However, the sign of change is consistent among the four climate models. It is also worth noting that other scenarios predict a much larger increase in the atmospheric CO₂ concentration and resulting ocean temperature. Thus, it is quite likely that marine ecosystems may experience even larger changes in the phytoplankton community structure than suggested here.

It can be challenging to differentiate between the individual effects of environmental factors because of a high degree of covariance in the ocean (25). Yet, the diversity and range of our large dataset provides an unprecedented opportunity to determine the importance of the different environmental variables. We find that temperature is the main control of the regional distributions of both Prochlorococcus and Synechococcus, but there are important quantitative differences. For Prochlorococcus, the inclusion of observations with zero cell abundance is key to delineate the sharp geographic limits and identify a steeper than previously thought response to temperature between 10 °C and 15 °C (26). Synechococcus displays a similar positive step response to temperature but with a threshold around 0 °C instead of 10 °C. This lower temperature boundary results in predicting significant Synechococcus populations in arctic waters but an absence in subzero waters around Antarctica—consistent with observations (27, 28). In addition, we observe a local minimum around 20 °C, which is

![Fig. 4. Prochlorococcus and Synechococcus observations and quantitative niche model for temperature and PAR. Cell abundance as a function of (A and C) temperature at constant PAR (10⁻¹ ± 0.005 E m⁻² d⁻¹) and (B and D) PAR at constant temperature (20 ± 0.05 °C). Symbol color represents the number of overlapping observations, and the lines show the model output.](Image)
Supported by specific latitudinal transects (11, 29). This intermediate temperature minimum could be the result of ecotype replacements with different temperature optima or a competitive interaction with Prochlorococcus, which displays a large increase in abundance at this temperature (Fig. 4 A and C and Fig. S6 A). PAR displays a complex control on Prochlorococcus and Synechococcus. At high PAR intensities (>10 E m⁻² d⁻¹), such as in tropical surface waters, we detect a 31% decline in cell abundance for Prochlorococcus. This observation is corroborated by quantitative PCR (26) and thus, likely not an artifact of counting populations with low chlorophyll content with flow cytometry. Photoinhibition or UV radiation damage can result in lower growth rates and overall lower abundances at these light levels (30). Alternatively, these cells receive a lower flux of nutrients compared with lower depths and PAR. At low PAR levels, we found a substantial number of Prochlorococcus and Synechococcus cells. Culture studies predict that both lineages have a compensation light level at about 0.06 E m⁻² d⁻¹ (19). The compensation light suggests that cells are unable to support autotrophic growth below this PAR level. Perhaps cells have been vertically advected out of the photic zone rather than actively growing below the compensation depth. Alternatively, some uncultured ecotypes may be capable of heterotrophy or growing at extremely low light levels. We posit vertical advection, because it is consistent with a suggested downward flux of pico-phytoplankton (5, 31).

Supported by the ANN and regression analyses, nutrient concentrations seem to have a limited influence on the biogeography of both Prochlorococcus and Synechococcus (Fig. S6 C and Tables S2 and S3). The limited role of nutrients is in agreement with the presence of Prochlorococcus and Synechococcus ecotypes that are adapted to different nutrient levels and thus, enable each lineage as a whole to proliferate over a wide range of nutrient concentrations (20, 32, 33). However, it is in contrast to the commonly held view that Prochlorococcus is at a competitive disadvantage in elevated nutrient environments (13). It is important to recognize that we may underestimate the bioavailable nutrients, because most of the nitrogen and phosphate in oligotrophic waters are linked to dissolved organic matter. In addition, we have analyzed the relationship between abundance and nutrient stocks rather than fluxes. These two aspects could influence the observed relationship between nutrient availability and abundance, and therefore, it is possible that nutrient limitation is not the primary control on Prochlorococcus and Synechococcus abundance in other situations. Nevertheless, the widespread presence of Prochlorococcus in elevated nutrient environments suggests that the linkage between abundance and nutrient availability is more complex for this lineage than previously thought.

Niche-based models have been widely used to predict the response of plant species to future climate change (17). Niche models express species abundance (or just presence) as a function of physical variables of the environment. They are based on collections of observations and their associated metadata describing the environment. These models represent the realized niche of species, because they reflect the use of resources as the net outcome of bottom-up and top-down processes. In this way, niche models do not address mechanisms such as predation or biotic interactions, which also can influence distribution and abundance (17). Projections of species distribution in future climate scenarios assume that species presence and abundance will be linked to similar factors as under current climate condition. We then use these principles to predict the global abundances and distributions of Prochlorococcus and Synechococcus. Our models provide a quantitative basis for predicting and understanding shifts in Prochlorococcus and Synechococcus distributions and suggest that future changes in ocean temperature can cause large-scale changes in phytoplankton community structures. Because of the high abundance of Prochlorococcus and Synechococcus, these changes can have a large impact on ocean ecosystems and global biogeochemical cycles.

Materials and Methods

Dataset. All analyses were done using Matlab (Mathworks). We obtained Prochlorococcus and Synechococcus observations from available public repositories and primary sources (Fig. 1 and Table S1). We only considered cell counts by flow cytometry for Prochlorococcus because of their weak autofluorescence, but we also included microscope counts for Synechococcus. Samples covered a latitudinal range from 81°N to 69°S up to 400 m depth; 79% of the samples were collected in the Northern Hemisphere. Ancillary temperature/nitrate/phosphate records were available for all but 3,600/8,000/14,000 observations, respectively, which we complemented with 1° monthly depth-dependent averages from the World Ocean Atlas (www.nodc.noaa.gov). To avoid issues with detection limits, we imposed a minimum nutrient concentration of 0.01 μM. We calculated surface PAR (0.8 d averaged, 0.047° grid cell) using SeaWiFS and MODIS observations. Downward PAR was calculated using the attenuation coefficient kₐₙₜ from MODIS and SeaWiFS (http://oceancolor.gsfc.nasa.gov) and corrected for chlorophyll a (34). Numerical results illustrated that the use of absolute PAR was better suited for our models than the average light intensity through the mixed layer (SI Materials and Methods, section 1.1). Because the averaging in the mixed layer involved a number of parameters to estimate the mixed layer depth that did not result in a noticeable improvement in the predictive power of PAR. We discarded infrequently high and low values for all variables (Fig. S1). The dataset is available at http://gdps.cima.fcen.uba.ar/.

ANN and Local Regression Models. We used back-propagation feed-forward ANN analyses to identify which input variables explained the observed Prochlorococcus and Synechococcus abundances. Temperature, PAR, nitrate, and phosphate concentrations and their possible combinations were tested as predictors (Table S2). For both lineages, temperature and PAR exhibited the highest predictive powers, and they were subsequently included in the regression models. We next applied local regression models with temperature and PAR as independent variables. Local abundances were estimated by fitting a linear regression to 200 observations in the neighborhood of each temperature and PAR combination. This local analysis provided the shape of the response curve of the cell abundance to environmental variables and highlighted thresholds on both axes. Based on the shapes and thresholds, we formulated parametric regression models (SI Materials and Methods, section 1.2).

Parametric Regression Model. The parametric regression model for Prochlorococcus cell abundance (Cₚₚ) consists of two components: an indicator variable I for the probability of presence (P) and cell abundance C:

$$ Cₚₚ = P × C + $$  

[1]

We assumed that Cₚₚ and C are on log₁₀ scale. To obtain P = P(foot T = 1) and Cₚₚ, we used log₁₀-transformed PAR and temperature (T) as input variables. For P, we used a logistic fit with two different temperature thresholds (Eq. 2):

$$ P = a + b × Pₚₚ × c × T + c₁ × T + i₁ + c₂ × T + i₂. $$  

[2]

where I₁(T) = 0 if T ≤ 8 (13) and I₂(T) = 1 otherwise. For Cₚₚ, we used nonzero observations and predicted the cell abundance of Prochlorococcus as follows:

$$ Cₚₚ = d + e × Pₚₚ + e₁ × Pₚₚ × d × T + f × T + f₁ × T^2 + i₃. $$  

[3]

where I₃(T) = 0 if PAR ≤ 10 (20) and I₃(T) = 1 otherwise. The regression model proposed for Synechococcus cell abundance (Cₚₚ) included temperature and PAR as explanatory variables (Eq. 4). We assumed Cₚₚ on a log₁₀ scale, unless stated differently. We defined a region of linear dependence (−2 < PAR < 0.1 log₁₀ E d⁻¹ m⁻²) surrounded by two regions without influence of PAR. The model assumed a nonlinear interaction between temperature and PAR:

$$ Cₚₚ = m + n × T + n₁ × T^2 + b₁ × T + b₂ × T^2 + b₃ + (b₄ × (1 - PAR × b₅)) × I₅, $$  

[4]

where I₅(T) = 0 if T ≤ 20 (5) and I₅(T) = 1 otherwise; I₅(T) = 0 if −2 ≤ PAR ≤ 0 and I₅(T) = 1 otherwise; and I₅(T) = 0 if T ≤ 5×(1 + PAR × b₆) ≤ 0 and I₅(T) = 1 otherwise. Fitted parameter values are listed in Table S4, and additional details are in SI Materials and Methods, section 1.3.

Sensitivity Analysis. We evaluated the sensitivity of the parametric model to input data in three ways. First, a jackknife analysis showed that the variance in each parameter decreased as a function of observations, but the means were stable (Fig. S4 A). The reduced number of samples influenced the spread of the determination coefficient and the variance of the mean global abundance.
but not their mean values (Fig. 5A and C). Second, we performed a split sample test to evaluate the influence of individual cruises or time series on the parametric regression models. This analysis would reveal possible biases originating from different measurement methods. The split sampling was done by stepwise removing datasets. This split-sampling approach hardly affected the coefficients of our regression models (Fig. 5A and D) except for ID12 (35). ID12 included 11,056 samples with zero Prochlorococcus abundance (set as zero in the log10 scale) and thus, influenced the logistic model (Eq. 2 and Fig. S6A). Third, we tested for any possible geographical bias contrasting observed and expected cell abundances (Fig. S4D).

**Present and Future Global Abundances and Distributions.** To estimate the mean global cell abundances of *Prochlorococcus* and *Synechococcus*, we used as input to our parametric regression models monthly average temperatures from the World Ocean Atlas 2005 (1° × 1° resolution) (36) and B- and D-average PAR and KD490 values derived from satellite data (SeaWiFS 0.083° × 0.083°) (37) (SI Materials and Methods, section 1.4). To estimate the sea surface distributions (Fig. 3 A and B), we used sea surface temperatures and PAR satellite data (SeaWiFS 0.083° × 0.083°) (37) as input variables in our regression models (SI Materials and Methods, section 1.4). To estimate abundance changes in the context of climate change, we used as inputs to our monthly PAR from MODIS (averaged 2002–2011) and sea surface temperature simulations from an ensemble of four global circulation models (CSIRO Mk3 6 0, CanESM2, HadGEM2 ES, and GISS E2 H) included in the Coupled Model Intercomparison Project (18). From the RCP4.5 scenario (38), we obtained the monthly sea surface temperature means of the last 30 y of the 20th and 21st century to represent present and future climate conditions, respectively. The RCP4.5 scenario predicts long-term global emissions of greenhouse gases, short-lived species, and land use changes, which stabilized radiative forcing at 4.5 Wm−2 or the equivalent of 650 ppm CO2. The percentage change in cell abundance was calculated based on the difference between the annual averages of future and present estimates. Uncertainty associated with each value represented the difference in the projection of each model (multimodel ensemble mean ± 1 SD; n = 4 models).

**ACKNOWLEDGMENTS.** We thank the many contributing researchers for the oceanographic data and Coupled Model Intercomparison Project (CMIP5) model outputs. Financial support for this work was provided by the Agencia Nacional de Promoción Científica y Tecnológica and Consejo Nacional de Investigaciones Científicas y Técnicas of Argentina (P.F.), University of California at Irvine Undergraduate Research Opportunity Program awards (to J.L.G. and R.A.G.), and a National Aeronautics and Space Administration Inspiring the Next Generation of Earth Explorers grant to the Earth System Science as well as California Alliance for Minority Participation at University of California at Irvine (to J.L.G., R.A.G., and J.R.), National Major Scientific Research Program and Natural Science Foundation of China (N.J.), National Science Foundation Emerging Frontiers program (D.M.K.), the Gordon and Betty Moore Foundation (D.M.K.), the National Science Foundation Dimensions of Biodiversity and Biological Oceanography programs (M.W.L. and A.C.M.), and the University of California at Irvine Environment Institute (J.A.V. and A.C.M.).