Acta Oceanol. Sin., 2022, Vol. 41, No. 8, P. 111–119
https://doi.org/10.1007/s13131-021-1956-2
http://www.aosocean.com
E-mail: ocean2@hyxb.org.cn

Top-down control on major groups of global marine diazotrophs

Hua Wang[1, 2], Ya-Wei Luo[1, 2*]

[1] State Key Laboratory of Marine Environmental Science, Xiamen University, Xiamen 361102, China

[2] College of Ocean and Earth Sciences, Xiamen University, Xiamen 361102, China

Received 16 July 2021; accepted 20 October 2021

Abstract

Dinitrogen (N_2) fixed by a group of prokaryotes (diazotrophs) is the dominant process adding bioavailable nitrogen into the ocean. Although it has been intensively studied how N_2 fixation is controlled by resources (bottom-up factors), it is unclear whether the grazing (top-down control) effectively impacts growth and distribution of different diazotroph groups. In this study, we evaluate this question by conducting log-log regression of diazotroph biomass onto corresponding N_2 fixation rates in the global ocean. The slope of the regression for *Trichodesmium* is ~0.8, indicating that a small portion of the increase in N_2 fixation does not accumulate as its biomass. This leads to a conclusion that *Trichodesmium* is under a substantial top-down control, although bottom-up control still dominates. We also analyze the residuals of the regression in the North Atlantic, concluding that free trichomes of *Trichodesmium* are subject to stronger top-down control than its colonies. The weak correlation between the biomass and N_2 fixation of unicellular cyanobacterial diazotrophs indicates that the degree of top-down control on this type of diazotrophs varies greatly. The analyses obtain unrealistic results for diatom-diazotroph assemblages due to complicated nitrogen sources of these symbioses. Our study reveals the variability of top-down control among different diazotroph groups across time and space, suggesting its importance in improving our understandings of ecology of diazotrophs and predictions of N_2 fixation in biogeochemical models. Measurements of size-specific N_2 fixation rates and growth rates of different diazotroph groups can be useful to more reliably analyze the top-down control on these key organisms in the global ocean.

Key words: marine diazotrophs, nitrogen fixation, top-down control, bottom-up control, *Trichodesmium*, diatom-diazotroph assemblages

Citation: Wang Hua, Luo Ya-Wei. 2022. Top-down control on major groups of global marine diazotrophs. Acta Oceanologica Sinica, 41(8): 111–119, doi: 10.1007/s13131-021-1956-2

1 Introduction

Marine biological dinitrogen (N_2) fixation is an important process in which special groups of microorganisms, termed diazotrophs, reduce N_2 gas to bioavailable nitrogen (N). That newly fixed N contributes up to a half of total external bioavailable N input to the euphotic zone, supporting a substantial portion of marine new production and leading to carbon export (Tyrrell, 1999; Zehr and Kudela, 2010).

Autotrophic cyanobacteria are the major diazotrophs in the ocean (Karl et al., 2002; Zehr, 2011), although diverse photoheterotrophic or heterotrophic diazotrophs have also been detected while their contribution to global N_2 fixation is still controversial (Bombar et al., 2016; Zehr and Capone, 2020). The autotrophic diazotrophic cyanobacteria include three major groups. *Trichodesmium*, the dominant diazotrophic group contributing a half of total N_2 fixation of the global ocean (Gruber, 2008), can exist as free filament trichomes consisting of approximately a hundred of cells (Capone et al., 1997; LaRoche and Breitbarth, 2005). Hundreds of *Trichodesmium* trichomes can further form large colonies in shapes of puffs or tufts. Most *Trichodesmium* trichomes are less than 10 μm wide and can be as long as several hundred micrometers, while the diameter of *Trichodesmium* colonies can be as large as several millimeters (Luo et al., 2012). The second group is diatom-diazotroph assemblages (DDAs), in which N_2-fixing heterocyst consisting of several heterocystous cyanobacteria (mostly of the genera *Richelia* and *Calothrix*) form symbioses with photosynthetic diatoms *Rhizosolenia*, *Hemiaulus* and *Chaetoceros* (Villareal, 1992; Carpenter et al., 1992). The width of these DDAs ranges ~10–100 μm (Foster et al., 2011). The most lately found diazotrophic bacterial group is the unicellular N_2-fixing cyanobacteria (UCYN), including three major subgroups UCYN-A, UCYN-B (*Crocosphaera)* and UCYN-C (Zehr, 2011). UCYN-A was also found to form symbiosis with a prymnesiophyte or coccolithophore species (*Braarudosphaera bigelowii*) (Thompson et al., 2012; García-Gómez et al., 2016). The diameter of UCYN ranges in 1–8 μm (Webb et al., 2009; Moisander et al., 2010; García-Gómez et al., 2016).

To study what controls marine N_2 fixation, much more attentions have focused on bottom-up factors, such as phosphorus, iron, light, temperature and physical environment, than on top-down controls from zooplankton grazing (e.g., Stukel et al., 2014; Hunt et al., 2016). However, statistical analyses show that bottom-up controls can only explain ~60% or less the spatial distribution of marine N_2 fixation rates and diazotrophic abundance in the global ocean (Luo et al., 2014; Tang and Cassar, 2019; Tang et al., 2019). Although early study showed some *Trichodesmium* species can be toxic to zooplankton (Hawser et al., 1992), analyses using quantitative polymerase chain reaction (qPCR) of gut

Foundation item: The National Natural Science Foundation of China under contract Nos 41890802 and 42076153.
*Corresponding author, E-mail: ywluo@xmu.edu.cn

contents in copepods and imaging flow cytometry of food vacuoles in dinoflagellates and ciliates reveal that *Trichodesmium*, DDA and UCYNs are all ingested by these zooplankton (Scavotto et al., 2015; Hunt et al., 2016; Conroy et al., 2017; Dugenne et al., 2020). Natural abundance $\delta^{15}N$ isotope data indicate *Trichodesmium* can support up to 60% of the food source of macrozooplankton in oligotrophic regions (Holl et al., 2007). Dilution grazing experiments also show that UCYN-B can be grazed at a rate up to 0.7 d^{-1} (Wilson et al., 2017). The above limited evidence cannot give clear conclusion if the top-down control on diazotrophs is substantial, particularly on large spatial scale.

The possible top-down controls on diazotrophs can potentially impact not only their spatial distribution and temporal variability, but also the fate of the newly fixed N in either being recycled to support primary production, or directly export. However, it is still unclear whether the top-down or the bottom-up control dominates marine diazotrophs on the global scale. It is important to answer that question not only in understanding the ecology of diazotroph but also in constraining marine ecosystem models. For example, some models assume the grazing rate on diazotrophs is lower than that on non-diazotrophic phytoplankton in order to compensate the lower growing rate set in the model for diazotrophs (Keller et al., 2012; Paulsen et al., 2017), while other ideal model experiments propose equal importance of top-down and bottom-up factors on diazotrophs (Stukel et al., 2014; Wang et al., 2019).

Furthermore, considering the wide range of size in different diazotrophs as mentioned above, they may be subject to different grazing pressure. For example, in a mesocosm experiment, zooplankton directly ingest UCYN-C but not *Trichodesmium* and DDAs (Hunt et al., 2016), while co-existence of a DDA bloom and the low $\delta^{15}N$ signatures of zooplankton in the oligotrophic North Atlantic Ocean indicates the DDA ingestion by zooplankton (Montoya et al., 2002). Therefore, top-down control can also impact the community composition of diazotrophs. If only considering their sizes, we speculate that the strongest grazing pressure is on UCYNs, the moderate on DDAs, and the weakest on *Trichodesmium*.

It is difficult to directly compare the top-down versus the bottom-up controls on microorganism on large scale. An indirect method has been applied to evaluate the relative importance of these two types of controls on marine bacteria by evaluating whether their biomass increases proportionally with their production (Billen et al., 1990; Pace and Cole, 1994; Dufour and Torréton, 1996; Raveh et al., 2015). In this study, we adopt that framework to analyze the strength of top-down control on major diazotrophic groups by conducting log-log regression of diazotroph biomass onto their N_2 fixation rates in the global ocean. Paired historical measurements of N_2 fixation rates and diazotrophic abundance from microscope counts and *nifH* copies were identified and diazotroph abundance is converted to biomass considering cell size of diazotrophs and their symbioses. The reliability of the results is also evaluated.

2 Methods

2.1 *Framework of comparing top-down and bottom-up controls*

The framework initially proposed by Billen et al. (1990) for aquatic bacteria linearly regresses logarithm of bacterial biomass onto logarithm of their production (Fig. 1). A slope of the regression closes to 1 indicates that the increase in the resource supply (reflecting as the increase in production) is proportionally transformed to biomass. However, if the slope is less than 1, it indicates that some of elevated biomass supported by increased resource is grazed. That is, the slope of 1 demonstrates a full bottom-up control on the evaluated organisms, while the top-down control gradually increases with the lowering slope.

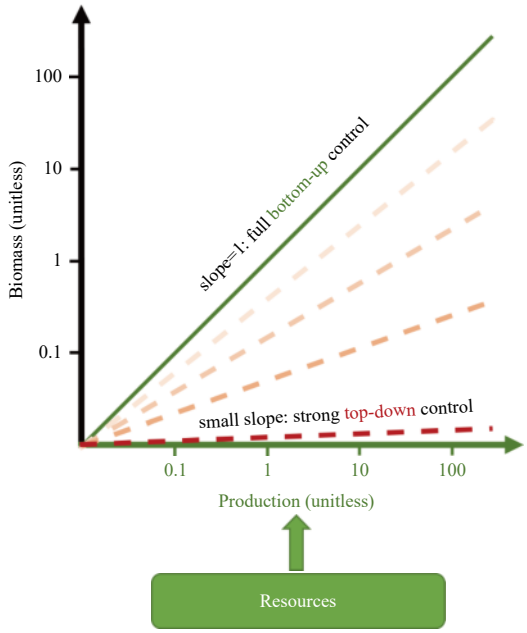

Fig. 1. Schematic diagram of determining whether organisms are under top-down or bottom-up control through regression analysis of logarithmic biomass and productivity.

In this study, we used this framework to study strength of top-down control on diazotrophs by using N_2 fixation rate as the variable representing production. That is, we assumed that diazotrophs fulfill their N requirement mostly from N_2 fixation.

2.2 *Diazotrophic abundance and N_2 fixation data*

Paired measurements of diazotrophic abundance and N_2 fixation rates were identified with existing datasets (Luo et al., 2012; Tang and Cassar, 2019) or collected from new literature (Table S1). We paired the abundance and N_2 fixation rates only when they were measured using the same samples. All the zero-value data were excluded because they needed to be log-transformed in the analyses (Fig. 1). The raw data are included in supplementary Excel datasheet of the paper.

One type of the abundance data is the directly-counted abundance of *Trichodesmium* trichomes or colonies using standard light or epifluorescence microscopy. Those data were reported in volumetric or depth-integrated units (Table S1). There are 268 and 222 pairs of *Trichodesmium* abundance and N_2 fixation rate data in volumetric and depth-integrated units, respectively (Table S1). We could not find sufficient paired N_2 fixation rates for directly-counted DDA abundance which therefore was not included in our analyses. Due to their small cell sizes and limitation of filtration, UCYNs were rarely directly counted and were not analyzed either.

Copies of *nifH* gene, the key gene encoding the iron protein component of the nitrogenase enzyme, were used to derive abundance of all diazotroph groups (Luo et al., 2012; Tang and Cassar, 2019), in which multiple *nifH* copies in diazotroph and DDA cells (White et al., 2018) were also considered (see below in biomass conversion section). All the *nifH* data were volumetric (Table S1).

Selected N_2 fixation rates paired for direct counting of *Trichodesmium* abundance were mostly reported for both 10-μm filtered and whole-water samples. The difference of the two rates was used as the N_2 fixation rate of *Trichodesmium*. Several studies only providing whole-water N_2 fixation rates were also included in our analyses (Table S1) because the studies reported *Trichodesmium* as the dominant diazotrophs (Text S1). There were 268 N_2 fixation rate data paired to the direct-count volumetric *Trichodesmium* abundance and 222 data to the depth-integrated abundance (Table S1). To pair *nifH*-based abundance data, only 157 whole-water N_2 fixation rates were available (Table S1) and there was no N_2 fixation rates for filtered samples.

2.3 *Biomass conversion factor*

Directly counted and *nifH*-based abundance was converted to carbon biomass using estimated factors (Table 1). The counted *Trichodesmium* trichomes and colonies were first converted to number of cells assuming 100 cells per trichome and 200 trichomes per colony (LaRoche and Breitbarth, 2005; Luo et al., 2012). The carbon biomass of 300 pg C per *Trichodesmium* cell was adopted from Luo et al. (2012), which is based on a cell volume-to-biomass relationship (Verity et al., 1992):

$$C = 0.433 \times V^{0.863}, \quad (1)$$

where C is cell carbon biomass (pg/cell, in terms of C); V is cell volume (μm^3).

UCYN abundance is derived from its *nifH* copies assuming 1 *nifH* per cell (Luo et al., 2012). The biomass conversion factors for UCYN-B and UCYN-C were adopted from Luo et al. (2012) (Table 1), while that for UCYN-A was estimated differently because they form symbioses. To evaluate top-down versus bottom-up controls, the symbioses need to be considered as a whole organism as they contribute to production and being grazed together. UCYN-A has two major clades including UCYN-A1 and UCYN-A2, which associate to hosts of prymnesiophyte and *B. bigelowii*, respectively (Thompson et al., 2014). Meanwhile, most *nifH* data of UCYN-A, we collected did not identify clade. Carbon biomasses of the diazotroph and the hosts were calculated from cell volume (assuming spheric shape) using Eq. (1). UCYN-A1 and its host are both smaller than UCYN-A2 and its host, respectively (Cabello et al., 2016; Cornejo-Castillo et al., 2016, 2019) (Table S2). However, each prymnesiophyte cell is often found to host only 1 UCYN-A1 cell, while a *B. bigelowii* cell can host 5–10 UCYN-A2 cells (Cornejo-Castillo et al., 2019). Therefore, if assuming each 8 UCYN-A2 cells per symbiosis, the factor converting UCYN-A2 abundance to symbiosis biomass reduces to 1.0–6.0 pg C per UCYN-A2 cell, closer to that for UCYN-A2 (1.7–2.4 pg C per UCYN-A1 cell) (Table S2). Furthermore, A1 is the dominant clade in UCYN-A (Thompson et al., 2014). We therefore used a uniform conversion factor of 2 pg C per UCYN-A cell (Table 1) based on estimates from UCYN-A1.

Table 1. Carbon biomass conversion factors from abundance

Type	Diazotrophic group	Conversion factor	Unit
Direct count	*Trichodesmium*	300	pg/cell, in terms of C
nifH	*Trichodesmium*	30	pg C per *nifH* gene copy
	UCYN-A	2	pg C per *nifH* gene copy
	UCYN-B	20	pg C per *nifH* gene copy
	UCYN-C	10	pg C per *nifH* gene copy
	Richelia-Hemiaulus	43.5	pg C per *nifH* gene copy
	Richelia-Rhizosolenia	50.1	pg C per *nifH* gene copy

The most widely distributed DDAs include the symbiosis formed by *Richelia intracellularis* and diatom *Hemiaulus* spp. or *Rhizosolenia hebetata*. Similarly, the biomass of whole DDA symbioses was used in our analyses. Average biomasses of these two host diatoms were reported at 733 pg/cell and 2 193 pg/cell (in terms of C), respectively (Barton et al., 2013). Each *Richelia* trichome has been estimated to consist of 1 heterocyst of 22 pg/cell (in terms of C) and 10 vegetative cells of 9 pg/cell (in terms of C) (Luo et al., 2012), i.e., 11 cells with 112 pg C in total. Based on previous observations (Villareal, 1992; Yeung et al., 2012), we further assumed 2 and 5 *Richelia* trichomes associated with each *Hemiaulus* and *Rhizosolenia* cell, respectively. Above estimates give normalized symbiosis biomass at 43.5 pg C per *Richelia* cell and 50.1 pg C per *Richelia* cell for *Richelia-Hemiaulus* and *Richelia-Rhizosolenia*, respectively (Table 1).

Although large number of *nifH* copies in each *Trichodesmium* and *Richelia* cells was suggested (White et al., 2018), it is still uncertain due to limitation of qPCR method and limited samples. Indeed, in our analyses, the number of *Trichodesmium nifH* copies had to be divided by 10 when estimating *Trichodesmium* abundance, otherwise the calculated cell specific N_2 fixation rates would be too low. Hence, we reduced the biomass conversion factor for *Trichodesmium* to 30 pg C per *nifH* gene copy, one order of magnitude less than that for direct counting data (Table 1). However, that is not the case for *Richelia-Hemiaulus* and *Richelia-Rhizosolenia*. Nevertheless, using different number of *nifH* gene copies per diazotrophic cell changes all the estimated biomass proportionally, which, however, does not change the slope of regression in our evaluation of top-down versus bottom-up controls because the biomass is log-transformed (Fig. 1).

2.4 *Analysis of intensity of top-down control on diazotrophs*

All the carbon biomass was converted to N biomass using Redfield ratio, so that specific growth rate can be easily calculated.

For the dataset using directly-counted *Trichodesmium* abundance, the *Trichodesmium* biomass and *Trichodesmium* N_2 fixation rates (some were whole-water rates as discussed above) were used in the analyses.

For the datasets using *nifH* gene copies, only the total UCYN biomass was used. That is, we analyzed top-down control for three groups, *Trichodesmium*, UYCN and DDA. As only whole-water N_2 fixation rates were available in these samples, we therefore paired each whole-water N_2 fixation rate to the dominant diazotroph group, defined here as one of the three groups that contributed more than 2/3 of the total diazotrophic biomass. If there was no dominant group, the data were discarded.

The slope of the fitted regression line represents average strength of top-down control on each diazotroph group. Meanwhile, a positive residual, i.e., higher biomass than the regression line, indicates that the diazotroph at that site is subject to a weaker top-down control than the average level, while a negative residual indicates a stronger top-down control. We then used the residuals to reveal a pattern of top-down control on *Trichodesmium* in the North Atlantic, where we had more data than other regions.

3 Results

3.1 *Distributions of diazotroph abundance, biomass and nitrogen fixation*

The directly-counted *Trichodesmium* data and paired N_2 fixa-

tion rates distribute mostly in the Northern Hemisphere, particularly in the North Atlantic Ocean (Figs 2a–d), while those *nifH*-based data distribute more widely in the global ocean. In terms of biomass, *Trichodesmium* dominates the diazotroph community mostly in western side of tropical and subtropical Atlantic, DDA dominates mostly in tropical Atlantic, while UCYN can dominate across the global ocean (Figs 2e, f).

3.2 *Analyses based on Trichodesmium direct-count data*

Using the directly-counted *Trichodesmium* data, the slope of the log-log regression of biomass to N_2 fixation rates is 0.82±0.05 and 0.74±0.05 for volumetric and depth-integrated data, respectively (Fig. 3). In the previous work for aquatic bacteria, it suggests that the strengths of top-down and bottom-up controls are comparable when the slope is ~0.5 (Dufour and Torréton, 1996). Hence, the slope values obtained here indicate *Trichodesmium* is mainly controlled by bottom-up factor while top-down control, although weaker than bottom-up control, also exists. The high explained variance with R^2=0.78 and R^2=0.76 for volumetric and depth-integrated data, respectively, suggests that the relative strength of the two types of control is relatively stable across the ocean. The wide range of N_2 fixation rate (6–7 orders of magnitude) and biomass (5–6 orders of magnitude) data (Fig. 3) partly support the robustness of our analyses.

Nevertheless, the residuals of the regression are up to ±2 orders of magnitude in *Trichodesmium* biomass (Fig. 3), indicating that the strength of top-down control can vary in different site or time. In the North Atlantic, the residuals tend to be negative in its eastern region and positive in the west (Fig. 4), which suggests that *Trichodesmium* is under stronger top-down control in the former regions than in the later.

In the eastern North Atlantic, the nutrients in the surface waters are generally replete due to strong upwelling (Neuer et al., 2007). However, the phytoplankton biomass rarely accumulates in this area because the high grazing rate (0.15–1.29 d^{-1}) is comparable to the phytoplankton growth rate (0.11–1.60 d^{-1}) (Cáceres et al., 2013). The abundance of *Trichodesmium* in this area is low, while UCYN-A is the dominant diazotroph and has much higher abundance than *Trichodesmium* (Agawin et al., 2014; Benavides et al., 2016; Fonseca-Batista et al., 2019). In other words, the environmental condition in this area is acceptable for diazotrophs in general but not for *Trichodesmium*. Interestingly, most *Trichodesmium* in the region exists as free trichomes (Benavides et al., 2011; Singh et al., 2019), which can be more easily grazed because of their smaller size than *Trichodesmium* colonies. Additionally, there is evidence that *Trichodesmium* tend to release toxin after forming colonies (Detoni et al., 2016). We thus speculate that the strong grazing on the free trichomes is one of the main reasons that *Trichodesmium* is under stronger top-down control in the eastern North Atlantic.

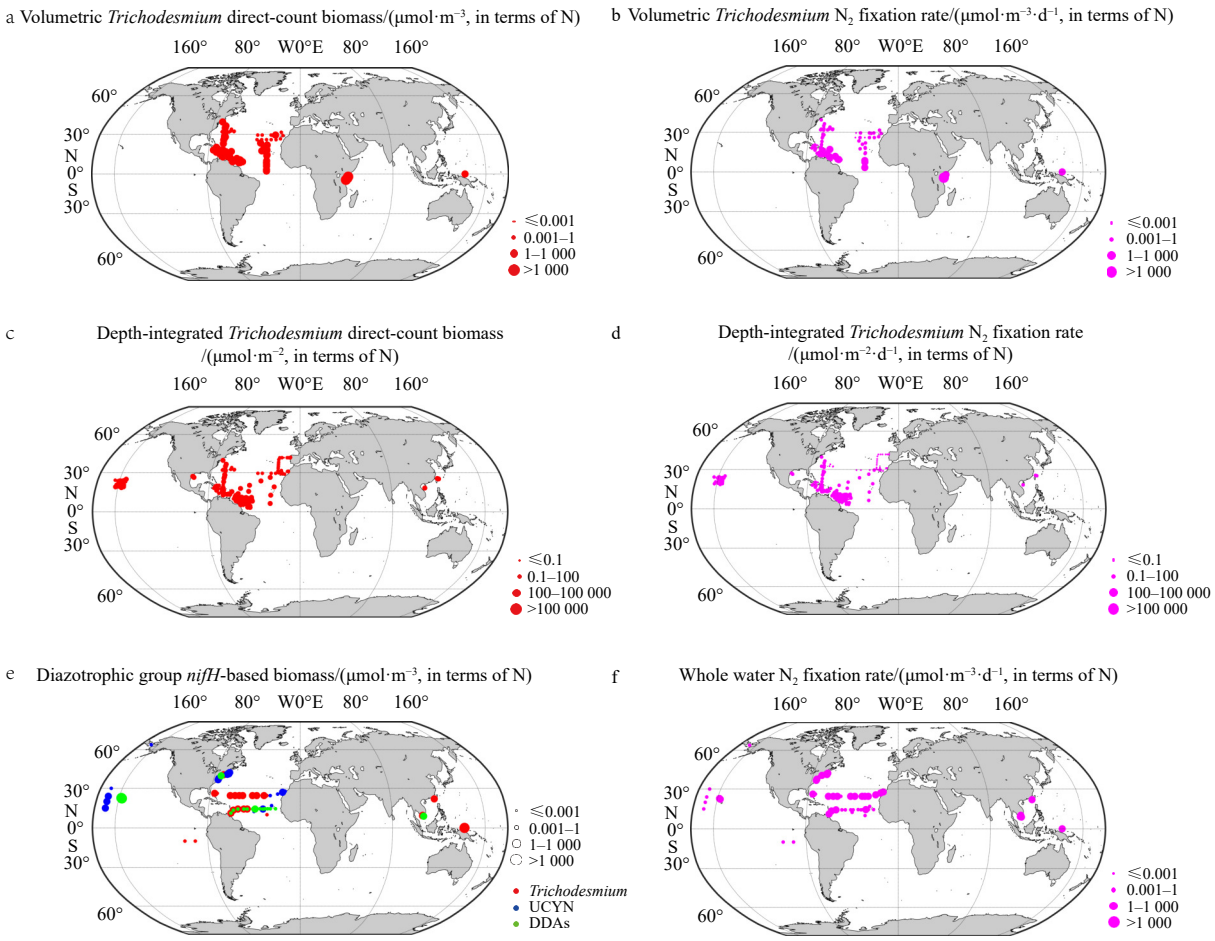

Fig. 2. Spatial distributions of collected diazotrophic data. Volumetric (a) and depth-integrated (c) *Trichodesmium* biomass using directly counted abundance data. Diazotroph biomass derived from *nifH*-based abundance, with the color representing the dominant group (note that some data points overlap spatially) (e). In b, d and f, N_2 fixation rates paired to diazotrophic biomass in a, c and e, respectively.

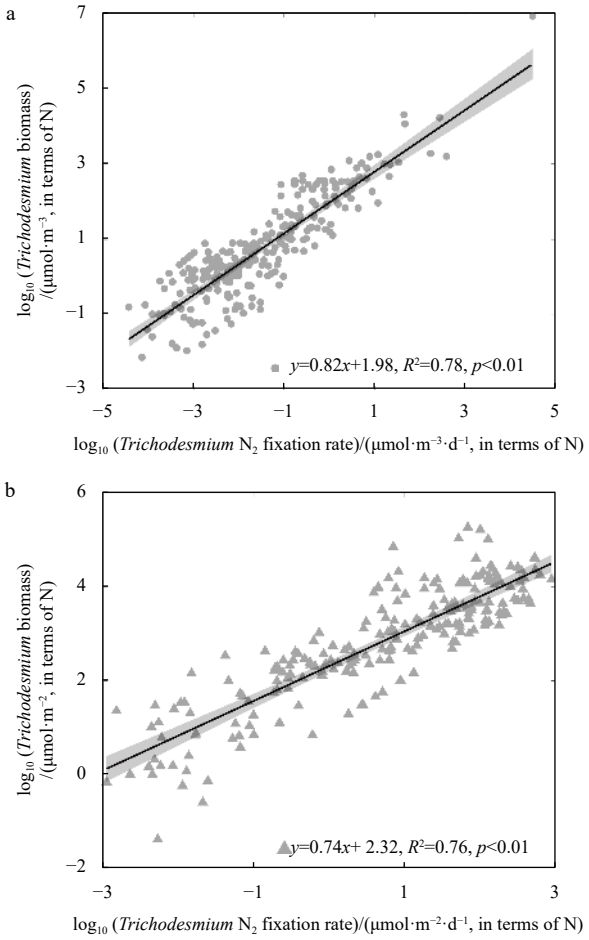

Fig. 3. The log-log regression of *Trichodesmium* biomass on its N_2 fixation rates. The *Trichodesmium* biomass data are based on directly counted abundance. a. volumetric; b. depth-integrated data. The grey area near the regression line represents 68% ($\pm 1\sigma$) confidence interval of the regression.

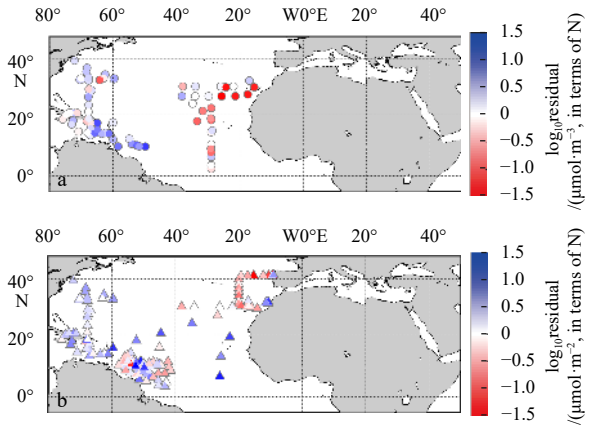

Fig. 4. The residuals of the log-log regression of Fig. 3 in North Atlantic. Positive residuals indicate weaker top-down control than average and negative residuals indicate stronger top-down control. Both the analyses using volumetric (a) and depth-integrated (b) data are shown.

3.3 *Analyses based on nifH data*

The analyses involving *nifH* data use total diazotroph biomass and whole-water N_2 fixation rates, while the data are separated into three categories according to the dominant diazotroph group (see Section 2). The log-log regression was then conducted to each category. By doing these dominant group-based analyses, we expect at least that the comparison of the slopes can improve our understandings of relative degree of top-down control on different diazotroph groups. The values of the slopes can be comparable only when the correlation between the biomass and production is substantial (i.e., R^2 is not low) and the slope is in a reasonable range (between 0–1).

The slope of the regression using *nifH*-based *Trichodesmium*-dominant data (n=54) is 0.84±0.24 (Fig. 5), comparable to those using the directly-counted *Trichodesmium* data (Fig. 3), confirming the existence of top-down control on *Trichodesmium*. The lower R^2 (0.48) in the *nifH*-based regression than those (0.78 and 0.76) using the direct-count data is expected, considering that the fraction of *Trichodesmium* in the total diazotroph group varies, although always dominant (>2/3), in the former regression across different sites.

The regression of the UCYN-dominant data (n=30) gives a slope of 0.92±0.82 (Fig. 5). However, the large uncertainty of the slope, the low R^2 of 0.14, and the p value of regression closer to 0.05 (Fig. 5) suggest that the estimated slope of 0.92 should not be used to identify the degree of top-down control on UCYN. Instead, the results indicate that the top-down and bottom-up controls on UCYN varies greatly in space and time.

The regression of DDA-dominant data (n=32) has the highest slope of 1.74±0.57 among the three groups (Fig. 5). R^2 value of 0.50 and p<0.01 suggest the regression is acceptable to analyze the top-down control versus bottom-up control on DDAs. The high slope suggests that DDAs are not under substantial top-down controls. However, the slope is much higher than its maximal theoretical value of 1 (Fig. 1), which will be discussed later.

4 Discussion

4.1 *Comparison of top-down control on different groups*

Our study using both directly counted abundance and *nifH*-based data indicates that *Trichodesmium* in the global ocean is overall under both top-down and bottom-up controls, although the former is weaker than the latter (Figs 3 and 5). That can result from both facts that *Trichodesmium* can be the food source and habitats for copepods, dinoflagellates and their larvae (e.g., Guo and Tester, 1994; O'Neil et al., 1996; Lugomela et al., 2002; Sheridan et al., 2002; Conroy et al., 2017), while can also be toxic particularly when it forms colonies and blooms (Hawser et al., 1992; Detoni et al., 2016). It is thus possible that the degree of top-down control on *Trichodesmium* depends on its form: Top-down control is stronger on *Trichodesmium* free trichomes than on its colonies, which is also partly supported by our analysis in the North Atlantic Ocean (Fig. 4). The *Trichodesmium* blooms that frequently occur in the tropical and subtropical oceans thus can contribute substantially to export (Capone et al., 1998), although the newly fixed N can still directly release from *Trichodesmium* cells to support other phytoplankton (Mulholland, 2007).

Our analyses suggest that DDAs are under very weak top-down control (Fig. 5), which is consistent to the findings in the North Pacific Subtropical Gyre where summertime increases of DDA contribute a strong pulse of export to as deep as 4 000 m (Karl et al., 2012). It is indeed not surprising as in the ocean, particularly in tropical and subtropical regions, diatoms are one of the major contributors to the carbon export to the deep ocean (Falkowski et al., 2004; Siegel et al., 2014, 2016), which has no reason to be different when they are associated to diazotrophs.

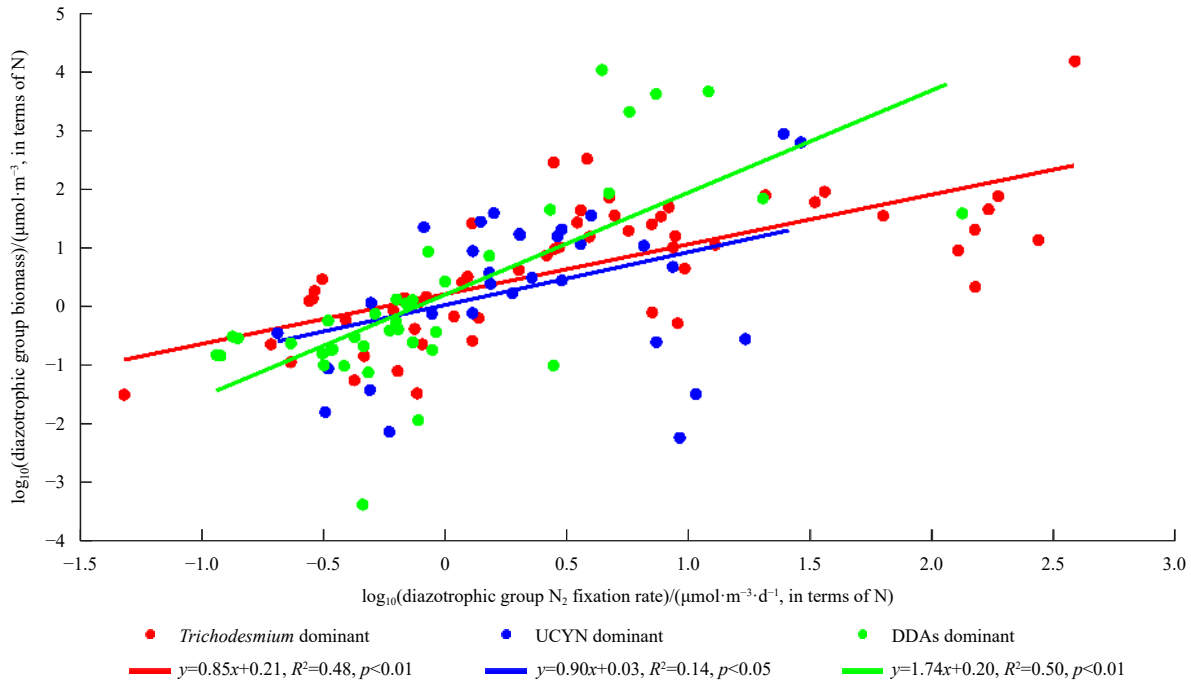

Fig. 5. The log-log regression of *nifH*-based diazotroph biomass to N_2 fixation rate. The data are separated into three categories according to dominant diazotroph group including *Trichodesmium* (red) (n=54), unicellular N_2-fixing cyanobacteria (UCYN, blue) (n=30) and diatom-diazotroph assemblages (DDA, green) (n=32).

However, the regression slope (1.7) is much higher than its theoretical upper bound (1.0), raising the question whether the top-down/bottom-up analysis framework used in this study is applicable to test DDAs. Although the high slope can attribute to the insufficient number of data (Fig. 5), it is also possible that the increased supply of newly fixed N from diazotroph to the host diatom can enhance the competitive advantage of the host over other phytoplankton. The host diatom can then take up more other dissolved inorganic N (DIN) directly from environment. If the mechanism proposed above is true, the value of the regression slope then becomes useless for DDAs, unless the relationship between the rate of N_2 fixation by the diazotroph and the rate of DIN uptake by the host diatom can be resolved. For instance, although our analyses suggest DDAs should not be under top-down control, high detection of *Richelia-Hemiaulus* symbiosis was found in gut contents of calanoid copepods in the Amazon River plume (Conroy et al., 2017).

Our analyses also tentatively suggest that the degree of top-down control is highly variable on UCYN because of relatively weak correlation between biomass and N_2 fixation (Fig. 5), although our number of data points is limited (n=30). This speculation is consistent to a finding that the grazing rate on UCYN-B varies greatly between 0 to 0.7 d^{-1} (Wilson et al., 2017). Some large size class of UCYN-B releases large amount of exopolymer secretions (Sohm et al., 2011), which can act as a grazing deterrent to zooplankton (Liu and Buskey, 2000) and reduce top-down control. Lastly, the large variation in top-down may attribute to the different composition of subgroups of UCYN, which certainly needs more data to explore.

4.2 *Caveats*

There are many uncertainties when applying the top-down/bottom-up framework to diazotrophs. The framework assumes that the samples represent systems in steady state or at least quasi-steady state. However, it is inevitable that some samples were collected when the system is far away from steady state. For example, even if grazing is negligible, low biomass and high N_2 fixation can be sampled at the onset of a diazotroph bloom, leading to a false judgement that the diazotroph is under strong top-down control. An opposite situation can occur during the declining period after a bloom when the samples tend to show as being bottom-up controlled. Nevertheless, the temporal variations of N_2 fixation may be much smaller than its spatial variation of up to 6 orders of magnitude in the global ocean (Figs 3 and 5; Luo et al., 2012). Our analyses on the global ocean scale then can be expected to give first-order results comparing bottom-up and top-down controls. In fact, in most previous studies in evaluating global marine N_2 fixation rate and their controlling mechanisms, almost all samples have been included in analyses, thus implicitly assuming the widely collected data were mostly sampled from steady-state system.

Although some interesting patterns emerge from our analyses, the available paired abundance and N_2 fixation data are still limited with large regions undersampled (Fig. 2 and Table S1). When using *nifH* data, because only whole-water N_2 fixation rates are available, uncertainties are introduced when separating data into three categories according to the dominant diazotroph group, because the relative contribution of the dominant group to total biomass and whole-water N_2 fixation varies across different sites. That also limits us to conduct analyses on individual subgroups of UCYN, although their cell sizes are different (Table S2; Luo et al., 2012) and may be grazed by different predators.

Because the cell-specific growth rates of the diazotrophs were mostly not measured in our dataset, N_2 fixation rate was used to represent the variation in production of diazotrophs. Although it is known that diazotrophs can also take up other forms of DIN (nitrate and ammonium) (Holl and Montoya, 2005; Dekaezemacker and Bonnet, 2011), we assume that the niche for

marine diazotrophs is the environment with low level of DIN, where diazotrophs can compete over other phytoplankton. Therefore, we consider that the sampled diazotrophs depend on N_2 fixation as their main N source. Nevertheless, if the relative contribution of DIN to diazotroph N uptake is relatively stable or even stochastic, the slope of regression is unlikely biased. Additionally, if DIN would be a substantial N source of *Trichodesmium*, then our analysis would underestimate its production and would wrongly move the data points leftward in the log-log regression of biomass and production (Fig. 3), leading to overestimated residuals. If it would be true, a positive relationship between the residuals and environmental nitrate concentration should emerge assuming *Trichodesmium* could use more nitrate under eutrophic condition. However, the nitrate concentration (World Ocean Atlas 2018; Boyer et al., 2018) is weakly correlated to the residuals from analysis using volumetric *Trichodesmium* data (R^2=0.06) (Fig. S1a) and is uncorrelated to the residuals from that using the depth-integrate data (p>0.3) (Fig. S1b), indicating that overall DIN is not a substantial source for *Trichodesmium* at least in our data samples. Lastly, the unrealistically high slope found in the regression analyses for DDAs can attribute to the increased DIN uptake of host diatom stimulated by stronger N_2 fixation. That is, if there is a systematic relationship between N_2 fixation and DIN uptake, the framework is not applicable.

5 Conclusion

In this study, we use log-log regression between diazotroph biomass and N_2 fixation rates to identify the top-down control on different diazotroph groups in the global ocean. Our results confirm that *Trichodesmium* is subject to certain degree of top-down control, particularly when in free trichomes, although bottom-up control is stronger. The analyses also tentatively reveal the highly varying top-down control on unicellular diazotrophs. The indication of the results for diazotroph-diatom symbiosis, however, is unclear considering the intervention of DIN uptake by host diatoms. To improve the reliability of the conclusions, we suggest that future field studies should measure not only size-specific N_2 fixation rate but also the growth rate of diazotrophs. The variability of top-down control in different diazotroph groups and forms revealed in this study urges studies on this question, which can greatly contribute to improving the simulation and prediction of N_2 fixation in biogeochemical models.

References

Agawin N S R, Benavides M, Busquets A, et al. 2014. Dominance of unicellular cyanobacteria in the diazotrophic community in the Atlantic Ocean. Limnology and Oceanography, 59(2): 623–637, doi: 10.4319/lo.2014.59.2.0623

Barton A D, Finkel Z V, Ward B A, et al. 2013. On the roles of cell size and trophic strategy in North Atlantic diatom and dinoflagellate communities. Limnology and Oceanography, 58(1): 254–266, doi: 10.4319/lo.2013.58.1.0254

Benavides M, Agawin N S R, Arístegui J, et al. 2011. Nitrogen fixation by *Trichodesmium* and small diazotrophs in the subtropical northeast Atlantic. Aquatic Microbial Ecology, 65(1): 43–53, doi: 10.3354/ame01534

Benavides M, Moisander P H, Daley M C, et al. 2016. Longitudinal variability of diazotroph abundances in the subtropical North Atlantic Ocean. Journal of Plankton Research, 38(3): 662–672, doi: 10.1093/plankt/fbv121

Billen G, Servais P, Becquevort S. 1990. Dynamics of bacterioplankton in oligotrophic and eutrophic aquatic environments: bottom-up or top-down control? Hydrobiologia, 207(1): 37–42,doi: 10.1007/BF00041438

Bombar D, Paerl R W, Riemann L. 2016. Marine non-cyanobacterial diazotrophs: moving beyond molecular detection. Trends in Microbiology, 24(11): 916–927, doi: 10.1016/j.tim.2016.07.002

Boyer T P, Garcia H E, Locarnini R A, et al. 2018. World Ocean Atlas 2018. NOAA National Centers for Environmental Information. https://accession.nodc.noaa.gov/NCEI-WOA18[2021-08-30]

Cabello A M, Cornejo-Castillo F M, Raho N, et al. 2016. Global distribution and vertical patterns of a prymnesiophyte-cyanobacteria obligate symbiosis. The ISME Journal, 10(3): 693–706, doi: 10.1038/ismej.2015.147

Cáceres C, Taboada F G, Höfer J, et al. 2013. Phytoplankton growth and microzooplankton grazing in the subtropical northeast Atlantic. PLoS One, 8(7): e69159, doi: 10.1371/journal.pone.0069159

Capone D G, Subramaniam A, Montoya J P, et al. 1998. An extensive bloom of the N_2-fixing cyanobacterium *Trichodesmium erythraeum* in the central Arabian Sea. Marine Ecology Progress Series, 172: 281–292, doi: 10.3354/meps172281

Capone D G, Zehr J P, Paerl H W, et al. 1997. *Trichodesmium*, a globally significant marine cyanobacterium. Science, 276(5316): 1221–1229, doi: 10.1126/science.276.5316.1221

Carpenter E J, Capone D G, Rueter J R. 1992. Marine Pelagic Cyanobacteria: *Trichodesmium* and Other Diazotrophs. Dordrecht: Springer

Conroy B J, Steinberg D K, Song B, et al. 2017. Mesozooplankton graze on cyanobacteria in the amazon river plume and western tropical North Atlantic. Frontiers in Microbiology, 8: 1436, doi: 10.3389/fmicb.2017.01436

Cornejo-Castillo F M, Cabello A M, Salazar G, et al. 2016. Cyanobacterial symbionts diverged in the late Cretaceous towards lineage-specific nitrogen fixation factories in single-celled phytoplankton. Nature Communications, 7: 11071, doi: 10.1038/ncomms11071

Cornejo-Castillo F M, del Carmen Muñoz-Marín M, Turk-Kubo K A, et al. 2019. UCYN-A3, a newly characterized open ocean sublineage of the symbiotic N_2-fixing cyanobacterium *Candidatus* Atelocyanobacterium thalassa. Environmental Microbiology, 21(1): 111–124, doi: 10.1111/1462-2920.14429

Dekaezemacker J, Bonnet S. 2011. Sensitivity of N_2 fixation to combined nitrogen forms (NO_3^- and NH_4^+) in two strains of the marine diazotroph *Crocosphaera watsonii* (Cyanobacteria). Marine Ecology Progress Series, 438: 33–46, doi: 10.3354/meps09297

Detoni A M S, Costa L D F, Pacheco L A, et al. 2016. Toxic *Trichodesmium* bloom occurrence in the southwestern South Atlantic Ocean. Toxicon, 110: 51–55, doi: 10.1016/j.toxicon.2015.12.003

Dufour P H, Torréton J P. 1996. Bottom-up and top-down control of bacterioplankton from eutrophic to oligotrophic sites in the tropical northeastern Atlantic Ocean. Deep-Sea Research Part I: Oceanographic Research Papers, 43(8): 1305–1320, doi: 10.1016/0967-0637(96)00060-X

Dugenne M, Henderikx Freitas F, Wilson S T, et al. 2020. Life and death of *Crocosphaera* sp. in the Pacific Ocean: Fine scale predator–prey dynamics. Limnology and Oceanography, 65(11): 2603–2617, doi: 10.1002/lno.11473

Falkowski P G, Koblfzek M, Gorbunov M, et al. 2004. Development and application of variable chlorophyll fluorescence techniques in marine ecosystems. In: Papageorgiou G C, Govindjee, eds. Chlorophyll *a* Fluorescence: A Signature of Photosynthesis. Dordrecht: Springer, 757–778

Fonseca-Batista D, Li X F, Riou V, et al. 2019. Evidence of high N_2 fixation rates in the temperate northeast Atlantic. Biogeosciences, 16(5): 999–1017, doi: 10.5194/bg-16-999-2019

Foster R A, Kuypers M M M, Vagner T, et al. 2011. Nitrogen fixation and transfer in open ocean diatom-cyanobacterial symbioses. The ISME Journal, 5(9): 1484–1493, doi: 10.1038/ismej.2011.26

García-Gómez C, Mata M T, Van Breusegem F, et al. 2016. Low-steady-state metabolism induced by elevated CO_2 increases resilience to UV radiation in the unicellular green-algae *Dunaliella tertiolecta*. Environmental and Experimental Botany, 132: 163–174, doi: 10.1016/j.envexpbot.2016.09.001

Gruber N. 2008. The marine nitrogen cycle: overview and challenges.

In: Capone D G, Bronk D A, Mulholland M R, et al., eds. Nitrogen in the Marine Environment. 2nd ed. San Diego: Academic Press, 1–50

Guo C Z, Tester P A. 1994. Toxic effect of the bloom-forming *Trichodesmium* sp. (cyanophyta) to the copepod Acartia tonsa. Natural Toxins, 2(4): 222–227, doi: 10.1002/nt.2620020411

Hawser S P, O'Neil J M, Roman M R, et al. 1992. Toxicity of blooms of the cyanobacterium *Trichodesmium* to zooplankton. Journal of Applied Phycology, 4(1): 79–86, doi: 10.1007/BF00003963

Holl C M, Montoya J P. 2005. Interactions between nitrate uptake and nitrogen fixation in continuous cultures of the marine diazotroph *Trichodesmium* (cyanobacteria). Journal of Phycology, 41(6): 1178–1183, doi: 10.1111/j.1529-8817.2005.00146.x

Holl C M, Villareal T A, Payne C D, et al. 2007. *Trichodesmium* in the western Gulf of Mexico: $^{15}N_2$-fixation and natural abundance stable isotopic evidence. Limnology and Oceanography, 52(5): 2249–2259, doi: 10.4319/lo.2007.52.5.2249

Hunt B P V, Bonnet S, Berthelot H, et al. 2016. Contribution and pathways of diazotroph-derived nitrogen to zooplankton during the VAHINE mesocosm experiment in the oligotrophic New Caledonia Lagoon. Biogeosciences, 13(10): 3131–3145, doi: 10.5194/bg-13-3131-2016

Karl D M, Church M J, Dore J E, et al. 2012. Predictable and efficient carbon sequestration in the North Pacific Ocean supported by symbiotic nitrogen fixation. Proceedings of the National Academy of Sciences of the United States of America, 109(6): 1842–1849, doi: 10.1073/pnas.1120312109

Karl D, Michaels A, Bergman B, et al. 2002. Dinitrogen fixation in the world's oceans. Biogeochemistry, 57–58: 47–98, doi: 10.1023/A:1015798105851

Keller D P, Oschlies A, Eby M. 2012. A new marine ecosystem model for the University of Victoria Earth System Climate Model. Geoscientific Model Development, 5(5): 1195–1220, doi: 10.5194/gmd-5-1195-2012

LaRoche J, Breitbarth E. 2005. Importance of the diazotrophs as a source of new nitrogen in the ocean. Journal of Sea Research, 53(1–2): 67–91, doi: 10.1016/j.seares.2004.05.005

Liu Hongbin, Buskey E J. 2000. The exopolymer secretions (EPS) layer surrounding *Aureoumbra lagunensis* cells affects growth, grazing, and behavior of protozoa. Limnology and Oceanography, 45(5): 1187–1191, doi: 10.4319/lo.2000.45.5.1187

Lugomela C, Lyimo T J, Bryceson I, et al. 2002. *Trichodesmium* in coastal waters of Tanzania: diversity, seasonality, nitrogen and carbon fixation. Hydrobiologia, 477(1–3): 1–13, doi: 10.1023/A:1021017125376

Luo Ya-Wei, Doney S C, Anderson L A, et al. 2012. Database of diazotrophs in global ocean: abundance, biomass and nitrogen fixation rates. Earth System Science Data, 4(1): 47–73, doi: 10.5194/essd-4-47-2012

Luo Ya-Wei, Lima I D, Karl D M, et al. 2014. Data-based assessment of environmental controls on global marine nitrogen fixation. Biogeosciences, 11(3): 691–708, doi: 10.5194/bg-11-691-2014

Moisander P H, Beinart R A, Hewson I, et al. 2010. Unicellular cyanobacterial distributions broaden the oceanic N_2 fixation domain. Science, 327(5972): 1512–1514, doi: 10.1126/science.1185468

Montoya J P, Carpenter E J, Capone D G. 2002. Nitrogen fixation and nitrogen isotope abundances in zooplankton of the oligotrophic North Atlantic. Limnology and Oceanography, 47(6): 1617–1628, doi: 10.4319/lo.2002.47.6.1617

Mulholland M R. 2007. The fate of nitrogen fixed by diazotrophs in the ocean. Biogeosciences, 4(1): 37–51, doi: 10.5194/bg-4-37-2007

Neuer S, Cianca A, Helmke P, et al. 2007. Biogeochemistry and hydrography in the eastern subtropical North Atlantic gyre. Results from the European time-series station ESTOC. Progress in Oceanography, 72(1): 1–29, doi: 10.1016/j.pocean.2006.08.001

O'Neil J M, Metzler P M, Glibert P M. 1996. Ingestion of $^{15}N_2$-labelled *Trichodesmium* spp. and ammonium regeneration by the harpacticoid copepod Macrosetella gracilis. Marine Biology, 125(1): 89–96, doi: 10.1007/BF00350763

Pace M L, Cole J J. 1994. Comparative and experimental approaches to top-down and bottom-up regulation of bacteria. Microbial Ecology, 28(2): 181–193, doi: 10.1007/BF00166807

Paulsen H, Ilyina T, Six K D, et al. 2017. Incorporating a prognostic representation of marine nitrogen fixers into the global ocean biogeochemical model HAMOCC. Journal of Advances in Modeling Earth Systems, 9(1): 438–464, doi: 10.1002/2016ms000737

Raveh O, David N, Rilov G, et al. 2015. The temporal dynamics of coastal phytoplankton and bacterioplankton in the eastern Mediterranean Sea. PLoS ONE, 10(10): e0140690, doi: 10.1371/journal.pone.0140690

Scavotto R E, Dziallas C, Bentzon-Tilia M, et al. 2015. Nitrogen-fixing bacteria associated with copepods in coastal waters of the North Atlantic Ocean. Environmental Microbiology, 17(10): 3754–3765, doi: 10.1111/1462-2920.12777

Sheridan C C, Steinberg D K, Kling G W. 2002. The microbial and metazoan community associated with colonies of *Trichodesmium* spp. : a quantitative survey. Journal of Plankton Research, 24(9): 913–922, doi: 10.1093/plankt/24.9.913

Siegel D A, Buesseler K O, Behrenfeld M J, et al. 2016. Prediction of the export and fate of global ocean net primary production: The EXPORTS science plan. Frontiers in Marine Science, 3: 22, doi: 10.3389/fmars.2016.00022

Siegel D A, Buesseler K O, Doney S C, et al. 2014. Global assessment of ocean carbon export by combining satellite observations and food-web models. Global Biogeochemical Cycles, 28(3): 181–196, doi: 10.1002/2013gb004743

Singh A, Gandhi N, Ramesh R. 2019. Surplus supply of bioavailable nitrogen through N_2 fixation to primary producers in the eastern Arabian Sea during autumn. Continental Shelf Research, 181: 103–110, doi: 10.1016/j.csr.2019.05.012

Sohm J A, Edwards B R, Wilson B G, et al. 2011. Constitutive extracellular polysaccharide (EPS) production by specific isolates of *Crocosphaera watsonii*. Frontiers in Microbiology, 2: 229, doi: 10.3389/fmicb.2011.00229

Stukel M R, Coles V J, Brooks M T, et al. 2014. Top-down, bottom-up and physical controls on diatom-diazotroph assemblage growth in the Amazon River plume. Biogeosciences, 11(12): 3259–3278, doi: 10.5194/bg-11-3259-2014

Tang Weiyi, Cassar N. 2019. Data-driven modeling of the distribution of diazotrophs in the global ocean. Geophysical Research Letters, 46(21): 12258–12269, doi: 10.1029/2019gl084376

Tang Weiyi, Li Zuchuan, Cassar N. 2019. Machine learning estimates of global marine nitrogen fixation. Journal of Geophysical Research, 124(3): 717–730, doi: 10.1029/2018JG004828

Thompson A, Carter B J, Turk-Kubo K, et al. 2014. Genetic diversity of the unicellular nitrogen-fixing cyanobacteria UCYN-A and its prymnesiophyte host. Environmental Microbiology, 16(10): 3238–3249, doi: 10.1111/1462-2920.12490

Thompson A W, Foster R A, Krupke A, et al. 2012. Unicellular cyanobacterium symbiotic with a single-celled eukaryotic alga. Science, 337(6101): 1546–1550, doi: 10.1126/science.1222700

Tyrrell T. 1999. The relative influences of nitrogen and phosphorus on oceanic primary production. Nature, 400(6744): 525–531, doi: 10.1038/22941

Verity P G, Robertson C Y, Tronzo C R, et al. 1992. Relationships between cell volume and the carbon and nitrogen content of marine photosynthetic nanoplankton. Limnology and Oceanography, 37(7): 1434–1446, doi: 10.4319/lo.1992.37.7.1434

Villareal T A. 1992. Marine nitrogen-fixing diatom-cyanobacteria symbioses. In: Carpenter E J, Capone D G, Rueter J G, eds. Marine Pelagic Cyanobacteria: *Trichodesmium* and Other Diazotrophs. Dordrecht: Springer, 163–175

Wang Weilei, Moore J K, Martiny A C, et al. 2019. Convergent estimates of marine nitrogen fixation. Nature, 566(7743): 205–211, doi: 10.1038/s41586-019-0911-2

Webb E A, Ehrenreich I M, Brown S L, et al. 2009. Phenotypic and genotypic characterization of multiple strains of the diazotrophic cyanobacterium, *Crocosphaera watsonii*, isolated from the open ocean. Environmental Microbiology, 11(2): 338–348, doi: 10.1111/j.1462-2920.2008.01771.x

White A E, Watkins-Brandt K S, Church M J. 2018. Temporal variability of *Trichodesmium* spp. and diatom-diazotroph assemblages in the North Pacific subtropical gyre. Frontiers in Marine Science, 5: 27, doi: 10.3389/fmars.2018.00027

Wilson S T, Aylward F O, Ribalet F, et al. 2017. Coordinated regulation of growth, activity and transcription in natural populations of the unicellular nitrogen-fixing cyanobacterium *Crocosphaera*. Nature Microbiology, 2(9): 17118, doi: 10.1038/nmicrobiol.2017.118

Yeung L Y, Berelson W M, Young E D, et al. 2012. Impact of diatom-diazotroph associations on carbon export in the Amazon River plume. Geophysical Research Letters, 39(18): L18609, doi: 10.1029/2012GL053356

Zehr J P. 2011. Nitrogen fixation by marine cyanobacteria. Trends in Microbiology, 19(4): 162–173, doi: 10.1016/j.tim.2010.12.004

Zehr J P, Capone D G. 2020. Changing perspectives in marine nitrogen fixation. Science, 368(6492): eaay9514, doi: 10.1126/science.aay9514

Zehr J P, Kudela R M. 2010. Nitrogen cycle of the open ocean: from genes to ecosystems. Annual Review of Marine Science, 3: 197–225, doi: 10.1146/annurev-marine-120709-142819

Supplementary information:

Fig. S1. The relationship between nitrate concentrations and the residuals of the log-log regression of Fig. 3.

Table S1. Data source of paired diazotrophic abundance and N_2 fixation rates.

Table S2. The cell size of UCYN-A1, UCYN-A2 and their hosts (prymnesiophyte and *Braarudosphaera bigelowii*, respectively) and estimated carbon biomass of symbiosis normalized to number of UCYN cells.

Text S1. N_2 fixation rates of whole-water samples were treated as the rates of *Trichodesmium* (Table S1) in following data sources.

The supplementary information is available online at https://doi.org/10.1007/s13131-021-1956-2 and www.aosocean.com. The supplementary information is published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.