~~Version~~Global oceanic diazotroph database version 2 and elevated estimate of ~~the~~ global oceanic ~~diazotroph database~~N_2 fixation

~~Zhibo Shao[1,#], Yangchun Xu[1,#], Hua Wang[1], Weicheng Luo[1], Lice Wang[1], Yuhong Huang[1], Ya-Wei Luo[1]~~
[1]Zhibo Shao[1,#], Yangchun Xu[1,#], Hua Wang[1], Weicheng Luo[1], Lice Wang[1], Yuhong Huang[1], Nona Sheila R. Agawin[2], Ayaz Ahmed[3], Mar Benavides[4,5], Mikkel Bentzon-Tilia[6], Ilana Berman-Frank[7], Hugo Berthelot[8], Isabelle C. Biegala[4], Mariana B. Bif[9], Antonio Bode[10], Sophie Bonnet[4], Deborah A. Bronk[11], Mark V. Brown[12], Lisa Campbell[13], Douglas G. Capone[14], Edward J. Carpenter[15], Nicolas Cassar[16,17], Bonnie X. Chang[18], Dreux Chappell[19], Yuh-ling Lee Chen[20], Matthew J. Church[21], Francisco M. Cornejo-Castillo[22], Amália Maria Sacilotto Detoni[23], Scott C. Doney[24], Cecile Dupouy[4], Marta Estrada[22], Camila Fernandez[25,26], B. Fernández-Castro[27], Debany Fonseca-Batista[28], Rachel A. Foster[29], Ken Furuya[30], Nicole Garcia[4], Kanji Goto[31], Jesús Gago[32], Mary Rose Gradoville[33], M. Robert Hamersley[34], Britt A. Henke[35], Cora Hörstmann[4], Amal Jayakumar[36], Zhibing Jiang[37], Shuh-Ji Kao[1], Dave Karl[38], Leila R. Kittu[39], Angela N. Knapp[40], Sanjeev Kumar[41], Julie LaRoche[42], Hongbin Liu[43], Jiaxing Liu[44], Caroline Lory[45], Carolin R. Löscher[46], Emilio Marañón[47], Lauren F Messer[48], Matthew M. Mills[49], Wiebke Mohr[50], Pia H. Moisander[51], Claire Mahaffey[52], Robert Moore[53], Beatriz Mouriño-Carballido[47], Margaret R. Mulholland[54], Shin-ichiro Nakaoka[55], Joseph A. Needoba[56], Eric J. Raes[57], Eyal Rahav[58], Teodoro Ramírez-Cárdenas[59], Christian Furbo Reeder[4], Lasse Riemann[60], Virginie Riou[61], Julie C. Robidart[62], V.V.S.S. Sarma[63], Takuya Sato[64], Himanshu Saxena[41], Corday Selden[65], Justin R. Seymour[66], Dalin Shi[1], Takuhei Shiozaki[67], Arvind Singh[41], Rachel E. Sipler[12], Jun Sun[68,69], Koji Suzuki[70], Kazutaka Takahashi[71], Yehui Tan[44], Weiyi Tang[36], Jean-Éric Tremblay[72], Kendra Turk-Kubo[35], Zuozhu Wen[1], Angelicque E. White[38], Samuel T. Wilson[73], Takashi Yoshida[65], Jonathan P. Zehr[35], Run Zhang[1], Yao Zhang[1], Ya-Wei Luo[1]

[1] State Key Laboratory of Marine Environmental Science and College of Ocean and Earth Sciences, Xiamen University, Xiamen, Fujian ~~361102~~, China

[2] Marine Ecology and Systematics (MarES) research group, University of the Balearic Islands, Palma de Mallorca, Spain

[3] Environment and life Science Research Centre, Kuwait Institute for Scientific Research, Salmiya, Kuwait

[4] Aix Marseille Univ, Université de Toulon, CNRS, IRD, MIO, UM 110, 13288, Marseille, France

[5] Turing Center for Living Systems, Aix-Marseille University, 13009 Marseille, France

[6] Department for Biotechnology and Biomedicine, Technical University of Denmark, Lyngby, Denmark

[7] Department of Marine Biology, Leon H. Charney School of Marine Sciences, University of Haifa, Haifa, Israel

[8] Ifremer, DYNECO, Plouzané, France

[9] Monterey Bay Aquarium Research Institute, Moss Landing, California, USA

[10] Oceanographic Center of A Coruña, Spanish Institute of Oceanography (IEO-CSIC), A Coruña, Spain

[11] Bigelow Laboratory for Ocean Sciences, East Boothbay, Maine, USA

[12] Climate Change Cluster, University of Technology Sydney, Sydney NSW, Australia

[13] Department of Oceanography, Texas A&M University, College Station, Texas, USA

[14] Department of Biological Sciences, Marine and Environmental Biology Section, University of Southern California, Los Angeles, California, USA

[15] College of Science and Engineering, San Francisco State University, San Francisco, California, USA
[16] Division of Earth and Ocean Sciences, Nicholas School of the Environment, Duke University, Durham, North Carolina, USA
[17] CNRS, Université de Brest, IRD, Ifremer, LEMAR, Plouzané, France
[18] Vesta, PBC, Southampton, New York, USA
[19] College of Marine Science, University of South Florida, Florida, USA
[20] Department of Oceanography, National Sun Yat-sen University, Kaohsiung, Taiwan
[21] Flathead Lake Biological Station, University of Montana, Polson, Montana, USA
[22] Institute of Marine Sciences (ICM-CSIC), Barcelona, Spain
[23] Institute of Marine Sciences of Andalucía (ICMAN), Consejo Superior de Investigaciones Científicas (CSIC), Campus Río San Pedro, Puerto Real, Spain
[24] Department of Environmental Sciences, University of Virginia, Charlottesville, Virginia, USA
[25] CNRS Observatoire océanologique, Banyuls sur mer, France
[26] Center for Oceanographic Research COPAS Coastal, Universidad de Concepción, Chile
[27] Ocean and Earth Science, National Oceanography Centre, University of Southampton, Southampton, UK
[28] Department of Oceanography, Dalhousie University, Halifax, Nova Scotia, Canada
[29] Department of Ecology, Environment, and Plant Sciences, Stockholm University, Stockholm, Sweden
[30] Institute of Plankton Eco-engineering, Soka University, Hachioji, Tokyo, Japan
[31] Graduate School of Environmental Science, Hokkaido University, Kita-Ku, Sapporo, Japan
[32] Spanish Institute of Oceanography (IEO-CSIC), Centro Oceanografico de Vigo, Spain
[33] Columbia River Inter-Tribal Fish Commission, Portland, Oregon, USA
[34] Environmental Studies, Soka University of America, Aliso Viejo, California, USA
[35] Ocean Sciences Department, University of California at Santa Cruz, Santa Cruz, California, USA
[36] Department of Geosciences, Princeton University, Princeton, New Jersey, USA
[37] Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou, Zhejiang, China
[38] Department of Oceanography, University of Hawai'i at Mānoa, Honolulu, Hawaii, USA
[39] Marine Biogeochemistry, GEOMAR Helmholtz Centre for Ocean Research Kiel, Düstern, Kiel, Germany
[40] Department of Earth, Ocean, & Atmospheric Science, Florida State University, Tallahassee, Florida, USA
[41] Geosciences Division, Physical Research Laboratory, Ahmedabad, India
[42] Department of Biology, Dalhousie University, Halifax, Nova Scotia, Canada
[43] Department of Ocean Science, The Hong Kong University of Science and Technology, Hong Kong, China
[44] Key Laboratory of Tropical Marine Bio-resources and Ecology, Guangdong Provincial Key Laboratory of Applied Marine Biology, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou, Guangdong, China
[45] French National Research Institute for Sustainable Development, IRD, Marseille, France

[46] Department of Biology, DIAS, University of Southern Denmark, Odense, Denmark
[47] Centro de Investigación Mariña da Universidade de Vigo (CIM-UVigo), Departamento de Ecoloxía e Bioloxía Animal, Universidade de Vigo, Campus Lagoas-Marcosende, Vigo, Spain
[48] Division of Biological and Environmental Sciences, Faculty of Natural Sciences, University of Stirling, Stirling, Scotland, UK
[49] Earth System Science, Stanford University, Stanford, California, USA
[50] Max Planck Institute for Marine Microbiology, Bremen, Germany
[51] Department of Biology, University of Massachusetts Dartmouth, Dartmouth, Massachusetts, USA
[52] Department of Earth, Ocean and Ecological Sciences, University of Liverpool, Liverpool, UK
[53] Department of Oceanography, Dalhousie University, Halifax, Nova Scotia, Canada
[54]Department of Ocean and Atmospheric Sciences, Old Dominion University, Norfolk, Virginia, USA
[55] Center for Global Environmental Research, National Institute for Environmental Studies, Tsukuba, Japan
[56] OHSU-PSU School of Public Health, Oregon Health and Science University Portland, Portland, Oregon, USA
[57] Flourishing Oceans, Minderoo Foundation, Broadway, Nedlands, WA, Australia
[58] Israel Oceanographic and Limnological Research, National Institute of Oceanography, Haifa, Israel
[59] Centro Oceanográfico de Málaga, Instituto Español de Oceanografía (IEO, CSIC), Fuengirola, Spain
[60] Department of Biology, University of Copenhagen, Helsingør, Denmark
[61] Analytical, Environmental and Geo-Chemistry & Earth System Sciences, Vrije Universiteit Brussel, Brussels, Belgium
[62] National Oceanography Centre, Southampton, UK
[63] CSIR-National Institute of Oceanography, Regional Cente Waltair, Visakhapatnam, India
[64] Institute for Chemical Research, Kyoto University, Kyoto, Japan
[65] Department of Marine and Coastal Sciences, Rutgers University, New Brunswick New Jersey, USA
[66] Climate Change Cluster, University of Technology Sydney, Sydney, New South Wales, Australia
[67] Atmosphere and Ocean Research Institute, The University of Tokyo, Chiba, Japan
[68] Research Centre for Indian Ocean Ecosystem, Tianjin University of Science and Technology, Tianjin, China
[69] College of Marine Science and Technology, China University of Geosciences (Wuhan), Wuhan, Hubei, China
[70] Faculty of Environmental Earth Science, Hokkaido University, Sapporo, Japan
[71] Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan
[72] Québec-Océan and Takuvik, Department of Biology, Laval University, Québec, Canada
[73] School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, UK

[#] These authors contributed equally.

Correspondence to: Ya-Wei Luo (ywluo@xmu.edu.cn)

Abstract. Marine diazotrophs convert dinitrogen (N_2) ~~in seawater~~gas into bioavailable nitrogen (N), ~~contributing approximately half of the external input of bioavailable N to~~ supporting life in the global ocean. ~~A~~ In 2012, the first version of the global oceanic diazotroph database (version 1) was ~~previously~~ published ~~in 2012 (Luo et al., 2012).~~ Here, we ~~compiled~~present an updated version ~~2~~ of the database ~~by adding 23,095 *in situ* measurements of marine diazotrophic abundance and N_2 fixation rates published in the past decade,~~(version 2), significantly increasing the number of *in situ* diazotrophic measurements from 13,565 to 55, 286. Data points for N_2 fixation rates ~~and microscopic and qPCR based~~, diazotrophic cell abundance, and *nifH* gene copy abundance ~~data~~have increased by ~~140%, 26%~~184%, 86%, and ~~443~~809%, respectively. ~~Although~~Version 2 includes two new datasheets for the ~~updated database expanded spatial coverage considerably, particularly in the Indian Ocean, the data~~*nifH* gene copy abundance of non-cyanobacterial diazotrophs and cell-specific N_2 fixation rates. The measurements of N_2 fixation rates approximately follow a log-normal distribution ~~was still not uniform and most data were sampled in the surface Pacific and Atlantic Oceans. By summing the arithmetic means of the N_2 fixation rates in each ocean basin, the updated database substantially increased the estimate of~~ in both version 1 and version 2. However, version 2 considerably extends both the left and right tails of the distribution. Consequently, when estimating global oceanic N_2 fixation ~~from 137 ± 9 Tg N yr^{-1}~~ rates using the ~~old database to 260 ± 20 Tg N yr^{-1} (mean ± standard error). However, using~~ geometric means ~~instead, the updated database gave an estimate of global oceanic N_2 fixation (60 Tg N yr^{-1}~~ of different ocean basins, version 1~~)~~ and version 2 yield similar ~~to that estimated from the old database (62 Tg N yr^{-1}), while the new estimate had a larger uncertainty (confidence intervals~~rates (43–57 versus 45–63 Tg N yr^{-1}; ranges based on one geometric standard error). In contrast, when using arithmetic means, version 2 suggests a significantly higher rate of 223±30 Tg N yr^{-1} (mean±standard error~~: 47—107~~; same hereafter) compared to version 1 (74±7 Tg N yr^{-1}). Specifically, substantial rate increases are estimated for the South Pacific Ocean (88±23 versus 20±2 Tg N yr^{-1}), primarily driven by measurements in the southwestern subtropic, and the North Atlantic Ocean (40±9 versus 10±2 Tg N yr^{-1}). Moreover, version 2 estimates the N_2 fixation rate in the Indian Ocean to be 35±14 Tg N yr^{-1} ~~versus 52—73 Tg N yr^{-1}), mostly attributable to elevated uncertainties in the Pacific Ocean. An analysis comparing~~ , which could not be estimated using version 1 due to limited data availability. Furthermore, a comparison of N_2 fixation rates ~~measured~~obtained through different measurement methods at the same months ~~and location (1° × 1° grids) showed that the new $^{15}N_2$ dissolution method obtained N_2 fixation rates higher than~~ , locations, and depths reveals that the conventional $^{15}N_2$ bubble method yields lower rates in ~~65% of~~69% cases~~, with this percentage increasing when~~ compared to the ~~N_2 fixation rates were high (> approximately 3 μmol N m^{-3} d^{-1} using the~~new $^{15}N_2$ dissolution method~~). With greatly increased data points, this~~. This updated version ~~2~~ of the ~~global oceanic diazotrophic~~ database can ~~support~~facilitate future studies in marine ecology and biogeochemistry. The database is stored at the Figshare repository ~~(https://doi.org/10.6084/m9.figshare.21677687).~~(https://doi.org/10.6084/m9.figshare.21677687) (Shao et al., 2022).

1 Introduction

Dinitrogen (N_2) fixation is ~~conducted~~a process carried out by ~~a group of microbes, termed~~ select prokaryotes (diazotrophs~~, to convert inert~~) capable of converting N_2 ~~gases~~gas, which is not usable by most organisms, into bioavailable nitrogen (N). In the sunlit surface ocean, where dissolved inorganic forms of N ~~nutrients are largely~~ such as nitrate (NO_3^-) and ammonium (NH_4^+) are scarce, N_2 fixation plays an important role in ~~fertilizing~~providing N that can contribute to primary production, particularly in oligotrophic regions ~~(Gruber et al., 2008; Wang et al., 2019), and balances N loss processes such as~~ (Wang et al., 2019; Gruber, 2008). Globally, N_2 fixation serves to compensate, at least partially, for fixed N removed via denitrification and anammox ~~(Deutsch et al., 2007; Gruber, 2019)~~(Deutsch et al., 2007; Gruber, 2019).

~~Marine diazotrophs include three main types of autotrophic cyanobacteria (Zehr, 2011): (1) nonheterocystous filamentous cyanobacteria, *Trichodesmium*; (2) heterocystous cyanobacteria, *Richelia* or *Calothrix*, forming diazotroph diatom associations (DDAs); and (3) unicellular cyanobacteria (UCYNs).~~

~~Diazotrophic abundance can be directly obtained by counting their cells using microscopes or be estimated from their copies of *nifH* using qPCR assays (Church et al., 2005). *NifH* copies have been more frequently measured than microscopic cell counting, particularly in the past decade, and can be more useful when evaluating the abundance of different diazotrophic groups. However, caution must be taken because there are discrepancies between cell count based and *nifH* based diazotrophic abundance (Luo et al., 2012), largely attributed to large variations in the number of *nifH* copies in the genomes of marine diazotrophs, particularly *Trichodesmium* and heterocystous cyanobacteria (Sargent et al., 2016; White et al., 2018).~~

~~Considering the key role of N_2 fixation in oceans, a database compiling up-to-date measurements of N_2 fixation and diazotrophic abundances is essential in studying marine ecology and biogeochemistry. For example, the estimated global marine N_2 fixation rate ranges from 15 to 238 Tg N yr^{-1} using different methods (Zehr and Capone, 2021), which is commonly thought to be much lower than the estimated nitrogen loss (126 – 481 Tg N yr^{-1}) from denitrification and anammox~~ Marine diazotrophs include three main types of cyanobacteria (Zehr, 2011): (1) nonheterocystous filamentous cyanobacteria (e.g., *Trichodesmium*); (2) heterocystous cyanobacteria like *Richelia*, which may form diatom-diazotroph associations (DDAs); and (3) unicellular cyanobacteria (UCYNs). Non-cyanobacterial diazotrophs (NCDs) have also been widely detected in the ocean (Bombar et al., 2016; Delmont et al., 2021; Moisander et al., 2017). However, the contribution of NCDs to marine N_2 fixation has not been directly quantified, despite a few studies that have reported N_2 fixation by putative NCDs at the cellular level (Harding et al., 2022; Bentzon-Tilia et al., 2015a).

Diazotroph abundance has been estimated from *nifH* gene copies using qPCR assays (Church et al., 2005b) or droplet digital PCR (ddPCR) (Gradoville et al., 2017). The abundance of some cyanobacterial diazotrophs can also be obtained by counting them directly using microscopy-based techniques and in some cases flow cytometry. A recent work combined an image recognition pipeline with molecular mapping of the *nifH* gene to quantify diazotrophs in the Tara Oceans dataset (Karlusich et al. 2021). *NifH* gene copies have been more frequently measured than microscopy-based cell counts and can be more useful when evaluating the abundance of different diazotrophic groups. Caution must be taken because there can be

discrepancies between cell-count-based and *nifH*-based diazotrophic abundances (~~Zehr and Capone, 2020; Gruber~~Luo et al., ~~2008; Zhang et al., 2020~~2012)~~. One possible reason for this imbalance is inaccurate estimation of global marine N_2 fixation due to limited spatio-temporal coverage of measurements and questionable N_2 fixation assays (White et al., 2020).~~

, a finding largely attributed to large variations in the number of *nifH* copies per diazotroph cell, thus far observed particularly in *Trichodesmium* and heterocystous cyanobacteria (Sargent et al., 2016; White et al., 2018; Karlusich et al., 2021). However, a recent regional study spanning over 200 km in the North Pacific Subtropical Gyre has found a statistically significant linear correlation between the abundances of the *nifH* gene and cell counts in the UCYN-B (i.e., *Crocosphaera*) (linear slope = 1.82) and heterocystous cyanobacteria (*Richelia* and *Calothrix*; linear slope from 1.51-2.58) but not in *Trichodesmium* (Gradoville et al., 2022). A recent discussion highlighted the influence of the uncertainty in gene copy conversion to biomass and the need for further investigations on how to best take advantage of gene copy data for global diazotroph biogeography modelling purposes (Meiler et al., 2022; Zehr and Riemann, 2023); however, there is an agreement that quantifying gene counts is a powerful tool for studying marine diazotroph distributions (Meiler et al., 2023; Zehr and Riemann, 2023). Meiler et al., (2023) proposed a number of topics of study for this field moving forward; Gradoville et al. (2022) concluded that "we hope that future studies report *nifH*:cell and explore the mechanisms controlling this ratio." Both gene based and microscopy cell counts have innate biases, which should be elucidated in future studies.

Given the importance of N_2 fixation to ocean ecology and biogeochemistry, it is imperative that a database of up-to-date N_2 fixation and diazotrophic abundance measurements be maintained. Currently, global estimates of marine fixed N inputs via N_2 fixation rate mostly ranges from 100 to 170 Tg N yr^{-1} (see summary in Zhang et al., 2020). This value, together with other bioavailable N sources to the ocean including riverine input and atmospheric deposition, is considerably lower than estimates of N losses from the ocean such as denitrification, anammox and sediment burial (Zhang et al., 2020; Gruber, 2008; Zehr and Capone, 2021). While the overestimation of the N losses cannot be ruled out, one of possible reasons for this imbalance is the inaccurate estimation of global marine N_2 fixation due to limited spatio-temporal coverage of rate measurements and the different methods employed in N_2 fixation assays (White et al., 2020). Another possible reason is the limited knowledge of ecological niches of N_2 fixing organisms. Over the last decade, the realm of marine N_2 fixation has been expanded to include numerous non-paradigmatic habitats. Coastal (Mulholland et al., 2012; Bentzon-Tilia et al., 2015b; Mulholland et al., 2019; Tang et al., 2020; Turk-Kubo et al., 2021), subpolar (Sato et al., 2021; Shiozaki et al., 2018c), and even polar ocean regions (Blais et al., 2012; Sipler et al., 2017; Harding et al., 2018; Shiozaki et al., 2020) have demonstrated N_2 fixation. Notably, N_2 fixation in aphotic waters remains debated (Bonnet et al., 2013; Farnelid et al., 2013; Selden et al., 2021a; Rahav et al., 2013b; Hamersley et al., 2011; Benavides et al., 2018a; Moisander et al., 2017). Other studies have also suggested that NCDs may be significant contributors to marine N_2 fixation (Shiozaki et al., 2014b; Turk-Kubo et al., 2022; Geisler et al., 2020; Delmont et al., 2021; Karlusich et al., 2021; Bombar et al., 2016; Moisander et al., 2017) and may occupy different niches from cyanobacterial diazotrophs (Shao and Luo, 2022).

Luo et al. (2012) compiled the first global oceanic diazotrophic database including *in situ* measurements of N_2 fixation rates and cell-count-based and *nifH*-based diazotrophic abundance. ~~In the past decade, many more measurements have been~~

reported, extending diazotrophic activities to new habitats such as subpolar (Sato et al., 2021; Shiozaki et al., 2018b) and even polar regions (Shiozaki et al., 2020; Harding et al., 2018). Several years agolater, two studies supplemented the database with a collection of some newly measuredreported diazotrophic data (Tang and Cassar, 2019; Tang et al., 2019)(Tang and Cassar, 2019; Tang et al., 2019), although a substantial amount of additional data still neededremained to be included.

Here, we present an updated version of the global oceanic diazotrophic database with data not yet compiled. We describe the database information, a summary of the data updates, measurement methods and data distribution. Furthermore, we conduct a first-order estimation of the global oceanic N_2 fixation rate using the updated version of the database. We also analyzed the discrepancy in In light of the aforementioned concerns of *nifH*:cell and various N_2 fixation assays and the relationshipmethods (*see* **Section 2.3**), we also discuss the significance of employing different methodological approaches to estimate N_2 fixation rates and abundance metrics. We use the data available in the database to analyze the discrepancies between N_2 fixation rates using $^{15}N_2$ bubble and dissolution methods, and compare the observed ranges of *nifH* gene copies and diazotrophic cell abundance based on the existing data.

2 Data and methods

2.1 Database summary

This study updated the original global oceanic diazotrophic database of Luo et al. (2012) (version 1 hereafter) with new *in situ* measurements of N_2 fixation rates and abundanceabundances of diazotrophic cells and *nifH* gene copies. The Together there were 55, 286 diazotrophic data points in the updated database (version 2) included 23,095 new data points, including 6,902 hereafter) (**Tables 1–3**), including 13,565 data points from version 1 (Luo et al., 2012), 6,736 measured in 2012 – 2018 and collectedcompiled by two previous studies (Tang et al., 2019; Tang and Cassar, 2019) and 16,193(Tang et al., 2019; Tang and Cassar, 2019), 26,597 data points measured in 20121979 – 2023 and collectedcompiled by this study (**Table 1**), and 8,388 NCD data mostly from Turk-Kubo et al. (2022) (see below). In version 2 of the database, some errors in the datasets of Tang et al. (2019)Tang et al. (2019) (mostly caused inby unit conversions) were also corrected.

Version 2 of the database is was composed of six main sub-databases: (1) 7,5159,231 volumetric N_2 fixation rates (4,2005,853 new data points) (**Tables 1 & 4**); (2); () 2) 2,248,590 depth-integrated N_2 fixation rates (1,497 new data points) (**Tables 1 & 2**); (3) 6,016 volumetric cell counts (1,130805 new data points) (**Tables 1 & 4**); (3); (4) 1,291 depth-integrated) 9,040 volumetric cell counts (360abundances (4,154 new data points) (**Tables 1 & 3**); (5) 17,1432 & 5); (4) 1,784 depth-integrated cell abundances (859 new data points) (**Tables 2 & 5**); (5) 29,655 volumetric *nifH* copies (13,994 new data points) (**Tables 1 & 4**); and (6) 2,356 depth integrated *nifH* copies (1,914gene copy abundances (26,506 new data points) (**Tables 1 & 43 & 6**); and (6) 2,986 depth-integrated *nifH* gene copy abundances (2,544 new data points) (**Tables 3 & 6**). Please be aware that 2,416 N_2 fixation rates were measured with incubation periods less than 24 hours; they were listed in separate spreadsheets in the database for reasons discussed in **Section 2.3**. Additionally we included a compiled NCD dataset (Turk-Kubo et al., 2022) in the database, which contained 7,919 *nifH* gene copy abundances of primarily the most studied phylotype

NCD Gamma A (Shao and Luo, 2022; Langlois et al., 2015), also referred to as 24774A11 (Moisander et al., 2012) and UMB (Bird et al., 2005), as well as other phylotypes, and updated that compilation with 469 additional *nifH* gene copy abundances of NCDs published more recently (Turk-Kubo et al., 2021; Sato et al., 2022; Moore et al., 2018; Reeder et al., 2022; Wen et al., 2022; Bonnet et al., 2023). We also collected 468 cell-specific *in situ* N_2 fixation rates and added them to version 2 (**Table 7**).

~~More recently, a unique group of non-cyanobacterial diazotrophs (NCDs) carrying *nifH* (gene encoding N_2-fixing enzyme, nitrogenase) has been widely found (Moisander et al., 2017; Zehr et al., 1995; Zehr, 1998), although direct evidence of N_2 fixation by NCDs is still limited (Harding et al., 2022). We therefore did not include NCD data in our database, while those who would be interested can use two recently published NCD datasets (Shao and Luo, 2022; Turk-Kubo et al., 2022).~~

~~Depth-integrated data were either provided directly in published papers or calculated for those vertical profiles with at least 3 volumetric data points in each profile. A profile was integrated from the sea surface to the depth of the deepest datum, while using the value of the shallowest datum to represent the level in the upper layer. We ignored those rates of N_2 fixation deeper than 200 m when calculating depth integrals, because they often had low vertical resolutions.~~

~~As in the original database, the data in version 2 were grouped into three taxonomic categories: *Trichodesmium*, UCYN and heterocystous cyanobacteria. The UCYN abundance data were further grouped into UCYN-A, UCYN-B, and UCYN-C, and the heterocystous cyanobacterial abundance was grouped into *Richelia* and *Calothrix*. N_2 fixation rates were measured for size groups or in whole seawater samples.~~ Depth-integrated data were either provided directly in published papers or calculated as part of this study for those vertical profiles with at least 3 volumetric data points in each profile. The measurements within a profile were first interpolated linearly with depth, with the shallowest datum representing the level between the sea surface and the depth of that datum. The profile was then integrated from the sea surface to the deepest recorded measurement. Most vertical profiles of N_2 fixation rates were measured within the euphotic zone, with a few studies extending measurements to several hundred meters or deeper. In these cases, we only integrated to the deepest data point above 200 m, taking into account the scarcity of aphotic N_2 fixation measurements in the global ocean and their controversial contribution to the global budget (Benavides et al., 2018a). As a result, it was possible that certain measurements below the euphotic zone but above 200 m were included in the integration. However, these measurements would typically have minimal impact on the depth-integrated N_2 fixation rates due to their low rates and limited vertical extent in this range.

N_2 fixation rates were measured for whole seawater samples, for different size fractions (> 10 μm and < 10 μm), or specifically for *Trichodesmium* and heterocystous cyanobacteria. When whole-water N_2 fixation rates were not reported, total N_2 fixation rates were calculated as the sum of the N_2 fixation rates of ~~different~~available groups.

The cyanobacterial diazotrophic abundance data in version 2 were grouped into three taxonomic categories: *Trichodesmium*, UCYN, and heterocystous cyanobacteria. The UCYN abundance data were further grouped into UCYN-A, UCYN-B, and UCYN-C. Four sublineages of UCYN-A, including UCYN-A1, UCYN-A2, UCYN-A3, and UCYN-A4, have been identified (Thompson et al., 2014; Farnelid et al., 2016). UCYN-A1 and UCYN-A2 have significant distinctions in the sizes and species of their symbiotic hosts, with the former living in relatively smaller hosts (Thompson et al., 2014; Martínez-

Pérez et al., 2016; Cornejo-Castillo et al., 2016). Hence, in addition to recording the total *nifH* gene copy abundance of UCYN-A in our database, the *nifH* gene copy abundances of its sublineages were also included if reported. Heterocystous cyanobacterial abundance was grouped into *Richelia intracelluaris* (het-1 and het-2, associated with *Hemiaulus* and *Rhizosolenia*, respectively) and *Richelia rhizosoleniae* (het-3, named *Calothrix* sp. before, associated with *Chaetoceros*) (Foster et al., 2022b).

Sampling information (latitude, longitude, depth and time) was provided for each data point. Physical, chemical and biological parameters, including temperature, salinity, and concentrations of nitrate, phosphate, iron and chlorophyll *a*, were also included when ~~they were~~ available.

Table 1. Summary of number of data points for N_2 fixation rates~~, cell-count-based abundance and *nifH*-based abundance.~~ by category. Measurements with incubation periods of 24 hours or of shorter than 24 hours are summarized separately.

-	Original database		New data added in ~~Version~~version 2				~~Sum~~sum	
			Tang et al., 2019	~~Tang & Cassar, 2019~~	This study			
Volumetric N_2 fixation rate								
	24 h	< 24 h	24 h	< 24 h	24 h	< 24 h	24 h	< 24 h
Trichodesmium	~~689~~	~~145~~677		~~83~~	~~9176~~		6	677
~~UCYN~~	~~275~~		~~124~~					~~399~~
Heterocystous	~~205~~	~~30~~185		~~83~~	~~318~~			185
< 10 µm	228	28	75		265	6	568	34
> 10 µm	54	36	9	21	51	6	114	63
Whole seawater	~~2,146~~ 1,743	427	1,~~32~~ ~~2~~169	171	~~2,413~~ 3,782	~~5,881~~2 92	6,694	890
Total	~~3,315~~ 2,025	1,~~621~~353	1,25 3	~~2,579~~192	4,104	304	7,~~515~~ 382	1,849
Proportion in version 2	~~48.1~~2 1.9%	~~21~~14.6%	13.6 %	2.1%	44.5 %	~~34~~3.3 %		
Depth-integrated N_2 fixation rate								
	24 h	< 24 h	24 h	< 24 h	24 h	< 24 h	24 h	< 24 h
Trichodesmium	~~2804~~0	~~89~~206	81	~~45~~8	~~414~~	9	121	223
~~UCYN~~	~~46~~		~~18~~				~~1~~	~~65~~
Heterocystous	1	65	~~92~~80	12	~~20~~	~~177~~	81	77
< 10 µm	28	18	7	12	21	2	56	32
> 10 µm	3	32			21	2	24	34
Whole seawater	~~360~~28 5	~~544~~107	500	~~702~~53	~~16069~~ 56	41	1,741	201
~~Total~~	~~751~~		~~743~~				~~768~~	~~2262~~
~~Proportion~~	~~33.2 %~~		~~32.8%~~				~~34.0%~~	
~~**Volumetric cell-count-based data**~~								
~~*Trichodesmium*~~	~~3,274~~						~~645~~	~~3,919~~
~~UCYN~~							~~85~~	~~85~~
~~Heterocystous~~	~~1,612~~						~~400~~	~~2,012~~
~~Total~~	~~4,886~~						~~1,130~~	~~6,016~~
~~Proportion~~	~~81.2 %~~						~~18.8%~~	

Depth-integrated cell-count-based data									
Trichodesmium	626							241	867
UCYN								19	19
Heterocystous	305							100	405
Total	931							360	1,291
Proportion	72.1%							27.9%	
Volumetric *nifH*-based data									
Trichodesmium	758					770		2,382	3,910
UCYN	1,792					2,640		4,822	9,254
Heterocystous	599					505		2,875	3,979
Total	3,149					3915		10,079	17,143
Proportion	18.4%					22.8%		58.7%	
Depth-integrated *nifH*-based data									
Trichodesmium	105					123		297	525
UCYN	263					418		609	1,290
Heterocystous	74					82		385	541
Total	44235 7		428		6236 68	1,29185	998	54	2,356 023	567
Proportion in version 2	1813. 8%	16.5%		26.4%	5425. 8%	3.3%	38.5 %	2.1%		

280 **Table 2.** Summary of number of data points of N_2 fixation rates in version 2 of the database, including whole seawaterfor diazotrophic cell abundances. UCYNs include UCYN-A, UCYN-B and volumetric measurements for *Trichodesmium*, unicellular unclassified UCYNs. Heterocystous cyanobacteria include Het-1, Het-2 and heterocystous diazotrophsHet-3.

	Original database	New data added to version 2	Sum
Volumetric cell abundances			
Trichodesmium	3,274	2,812	6,086
UCYN		139	139
Heterocystous cyanobacteria	1,612	1203	2,815
Total	4,886	4,154	9,040
Proportion in version 2	54.1%	45.9%	

Reference	Region	*Tricho*	UC	Heter	Whole
Depth-integrated cell abundances					

- *desmium*	Y N	o- eystous	Sea water			
Agawin et al. (2013)*Trichodesmium*			Subtropical Atlantic620	21692	1,312	17
Ahmed et al. (2017)UCYN			E Arabian Sea	19	19	5[a]
Benavides et al. (2013)Heterocystous			Subtropical N Atlantic305	15148	15453	15
Benavides et al. (2016a)			Mediterranean Sea		10	
Benavides et al. (2018a)			Tropical SW Pacific		59	
Benavides et al. (2022)			S Pacific		38	
Bentzon-Tilia et al. (2015)			Baltic Sea		23	23[a]
Berthelot et al. (2017)			Tropical W Pacific		48	12[a]
Bhavya et al. (2016)			Arabian Sea		4	
Biegala and Raimbault (2008)			SW Pacific		9	2
Blais et al. (2012)			Arctic Ocean		18	12
Bombar et al. (2011)			South China Sea		15	
Bombar et al. (2015)			Subtropical N Pacific		20	2
Bonnet et al. (2013)			Tropical SW Pacific			8[a]
Bonnet et al. (2015)			SW Pacific			30[a]

Bonnet et al. (2018)	Tropical SW Pacific	102	14
Böttjer et al. (2017)	Subtropical N Pacific	243	108[a]
Chang et al. (2000)	S East China Sea		7[a]
Chang et al. (2019)	Tropical SE Pacific	37	
Dekaezemacker et al. (2013)	Tropical SE Pacific	43	10
Fernandez et al. (2015)	Central Chile Upwelling System	84	14[a]
Fernández-Castro et al. (2015)	Atlantic, Pacific and Indian Oceans		43[a]
Fonseca-Batista et al. (2017)	E Atlantic	56	14
Fonseca-Batista et al. (2019)	Temperate NE Atlantic	46	10[a]
Foster et al. (2009)	Red Sea	26	
Gandhi et al. (2011)	E Arabian Sea	28	7[a]
Garcia et al. (2007)	SW Pacific		1[a]
González et al. (2014)	Southern Ocean	8	
Gradoville et al. (2020)	N Pacific	20	
Großkopf et al. (2012)	Atlantic Ocean	39	17
Hallstrøm et al. (2022)	NE Atlantic	59	11[a]

Halm et al. (2012)	S Pacific	43	43		43	10[a]
Harding et al. (2018)	Arctic Ocean				38	
Hörstmann et al. (2021)	S Indian Ocean; Southern Ocean				13	
Ibello et al. (2010)	Mediterranean Sea				21	7[a]
Jayakumar et al. (2017)	Tropical NE Pacific				32	7[a]
Knapp et al. (2016)	Tropical SE Pacific					6[a]
Konno et al. (2010)	NW Pacific					16[a]
Kromkamp et al. (1997)	Indian Ocean					4[a]
Krupke et al. (2013)	Subtropical N Atlantic				3	
Krupke et al. (2014)	N Atlantic	42	42	30	42	44[a]
Krupke et al. (2015)	Subtropical NE Atlantic				1	
Kumar et al. (2017)	E Arabian Sea				12	3
Kumari et al. (2022)	Bay of Bengal				97	18[a]
Lee Chen et al. (2014)	South China Sea					24[a]
Li et al. (2020)	N South China Sea; East China Sea				68	15[a]

Liu et al. (2020)	South China Sea	25	5[a]
Loescher et al. (2014)	Pacific Ocean	30	5[a]
Loick-Wilde et al. (2015)	Amazon River		36[a]
Loick-Wilde et al. (2019)	Tropical W Pacific	8	
Lory et al. (2022)	Tropical SW Pacific	5	
Löscher et al. (2016)	Tropical SW Pacific	225	31[a]
Löscher et al. (2020)	Bay of Bengal	18	
Lu et al. (2018)	Equatorial W Pacific	3	3[a]
Martínez-Pérez et al. (2016)	Tropical N Atlantic	84	14
Messer et al. (2016)	S Pacific	27	
Mouriño-Carballido et al. (2011)	Atlantic Ocean		20[a]
Mulholland et al. (2019)	NW Atlantic	402	242[a]
Needoba et al. (2007)	Temperate N Pacific	2	1[a]
Raes et al. (2020)	S Pacific	55	
Rahav et al. (2013a); Rahav et al. (2015)	Red Sea and E Mediterranean Sea	62	10
Rahav et al. (2013b); Rahav et al. (2013c)	Mediterranean Sea	8	

Rahav et al. (2016)	Mediterranean Sea				3[a]
Reeder et al. (2022)	S Baltic Sea			15	5
Riou et al. (2016)	N Atlantic	24	24		6
Sahoo et al. (2021)	Bay of Bengal				6[a]
Sarma et al. (2020)	Bay of Bengal			2	
Sato et al. (2021)	Subarctic Sea of Japan; Sea of Okhotsk			31	3
Saxena et al. (2020)	Bay of Bengal			32	8
Selden et al. (2019)	Tropical NE Pacific			8	16[a]
Shiozaki et al. (2013)	W Pacific			50	10
Shiozaki et al. (2014a)	Indian Ocean			42	
Shiozaki et al. (2014b)	SW Pacific			26	6[a]
Shiozaki et al. (2015a)	NW Pacific			73	11[a]
Shiozaki et al. (2015b)	N Pacific			112	22[a]
Shiozaki et al. (2017)	N Pacific			74	15
Shiozaki et al. (2018b)	W Arctic Ocean			84	21[a]
Shiozaki et al. (2020)	Antarctic Coast			53	15[a]
Sipler et al. (2017)	Arctic Ocean			8	
Sohm et al. (2011)	S Atlantic			12	3[a]

~~Subramaniam et al. (2008)~~	~~Tropical N Atlantic~~						~~242~~[a]
~~Subramaniam et al. (2013)~~	~~Atlantic Ocean~~			~~96~~			~~24~~[a]
~~Tang et al. (2020)~~	~~N Atlantic~~			~~15~~			
~~Turk-Kubo et al. (2012)~~	~~Tropical N Atlantic~~			~~27~~			~~7~~
~~Wang et al. (2021)~~	~~NW Atlantic~~			~~85~~			
~~Wasmund et al. (2015)~~	~~S Atlantic~~						~~66~~[a]
~~Watkins-Brandt et al. (2011)~~	~~N Pacific~~						~~1~~[a]
~~Wen et al. (2022)~~	~~Tropical NW Pacific~~			~~143~~			~~22~~[a]
~~White et al. (2018)~~	~~Subtropical N Pacific~~	~~83~~	~~83~~	~~62~~			~~51~~[a]
~~Wilson et al. (2012)~~	~~N Pacific~~			~~9~~			4[a]
~~Wilson et al. (2017)~~	~~Subtropical N Pacific~~			~~33~~			
~~Wu et al. (2021)~~	~~Eastern Indian Ocean~~			~~48~~			~~7~~
~~Yogev et al. (2011)~~	~~E Mediterranean Sea~~			~~16~~			~~32~~[a]
~~Zhang et al. (2015)~~	~~South China Sea~~			~~82~~			~~11~~
~~Zhang et al. (2019)~~	~~Tropical NW Pacific~~			~~87~~			~~9~~[a]
Total	925		~~228~~859	~~124~~1,784	~~11 3~~	~~373 5~~	~~149 7~~
Proportion in version 2	51.9%		48.1%				

Table 3. Summary of number of data points for *nifH* gene copy abundances. UCYNs include UCYN-A1, UCYN-A2, UCYN-B and UCYN-C. Heterocystous cyanobacteria include Het-1, Het-2 and Het-3.

	Original database	New data added to version 2		Sum
		Tang & Cassar, 2019	This study	
Volumetric *nifH* gene copy abundances				
Trichodesmium	758	770	3,165	4,693
UCYN	1,792	2,640	6,903	11,309
Heterocystous cyanobacteria	599	505	4,135	5,239
NCDs			8,388	8,388
Total	3,149	3915	22,591	29,655
Proportion in version 2	10.6%	13.2%	76.2%	
Depth-integrated *nifH* gene copy abundances				
Trichodesmium	105	123	408	636
UCYN	263	418	871	1552
Heterocystous	74	82	642	798
Total	442	623	1,921	2,986
Proportion in version 2	14.8%	20.9%	64.3%	

285

Table 4. Summary of new data points of N_2 fixation rates added to the version 2 of the database.

Reference	Region	*Tricho-desmium*	Hetero-cystous	< 10 µm Diazotrophs	> 10 µm Diazotrophs	Whole Seawater	Depth-integrated data
Part 1. Incubation periods of 24 hours							
Ahmed et al. (2017)	E Arabian Sea					19	5[a]
Benavides et al. (2016a)	Mediterranean Sea					10	
Benavides et al. (2018a)	Tropical SW Pacific					59	
Benavides et al. (2022a)	Tropical SW Pacific					38	
Benavides et al. (2017)	SW Pacific					2	
Benavides et al. (2021)	S Pacific					41	
Benavides et al. (2022b)	S Pacific	6				6	2

Reference	Region	*Tricho-desmium*	Hetero-cystous	< 10 µm Diazotrophs	> 10 µm Diazotrophs	Whole Seawater	Depth-integrated data
Bentzon-Tilia et al. (2015b)	Baltic Sea					23	23[a]
Berthelot et al. (2017)	Tropical W Pacific					48	12[a]
Biegala and Raimbault (2008)	SW Pacific			12	12	12	9
Blais et al. (2012)	Arctic Ocean					18	12
Bombar et al. (2015)	Subtropical N Pacific					20	2
Bonnet et al. (2013)	Tropical SE Pacific						8[a]
Bonnet et al. (2018)	Tropical SW Pacific					102	14
Bonnet et al. (2015)	SW Pacific			126		128	30[a]
Bonnet et al. (2023)	Subtropical S Pacific					84	14
Böttjer et al. (2017)	Subtropical N Pacific					243	108[a]
Cerdan-Garcia et al. (2021)	subtropical N Atlantic					15	
Chang et al. (2019)	Tropical SE Pacific					37	
Chen et al. (2019)	W Pacific Ocean					95	16
Dekaezemacker et al. (2013)	Tropical SE Pacific					43	10
Dugenne et al. (2023)	Subtropical N Pacific					30	5
Fernandez et al. (2015)	Central Chile Upwelling System					55	14[a]
Fernández-Castro et al. (2015)	Atlantic, Pacific and Indian Oceans					177	43[a]
Fonseca-Batista et al. (2017)	E Atlantic					56	14
Fonseca-Batista et al. (2019)	Temperate NE Atlantic					46	10[a]
Foster et al. (2009)	Red Sea					26	
Foster et al. (unpublished data)	E tropical S Pacific					23	5
Garcia et al. (2007)	SW Pacific						1[a]
Gradoville et al. (2020)	N Pacific					20	

Reference	Region	*Tricho-desmium*	Hetero-cystous	< 10 µm Diazotrophs	> 10 µm Diazotrophs	Whole Seawater	Depth-integrated data
Gradoville et al. (2017)	S Pacific; N Pacific					30	5
Großkopf et al. (2012)	Atlantic Ocean					39	17
Hallstrøm et al. (2022)	NE Atlantic					59	11[a]
Harding et al. (2018)	Arctic Ocean					38	
Harding et al. (2022)	Subtropical N Pacific					7	
Hörstmann et al. (2021)	S Indian Ocean; Southern Ocean					13	
Ibello et al. (2010)	Mediterranean Sea					21	14[a]
Jayakumar et al. (2017)	Tropical NE Pacific					32	7
Jiang et al. (2023)	East China Sea and Southern Yellow Sea					97	29[a]
Kittu et al. (2023)	Tropical SE Pacific					103	21
Knapp et al. (2016)	Tropical SE Pacific						6[a]
Konno et al. (2010)	NW Pacific						16[a]
Krupke et al. (2015)	Subtropical NE Atlantic					1	
Kumari et al. (2022)	Bay of Bengal					97	18[a]
Landou et al. (2023)	Red Sea					72	22[a]
Li et al. (2020)	N South China Sea; East China Sea					68	15[a]
Liu et al. (2020)	South China Sea					25	5[a]
Loescher et al. (2014)	Tropical SE Pacific					30	5[a]
Loick-Wilde et al. (2015)	Amazon River						54[a]
Loick-Wilde et al. (2019)	Tropical W Pacific					8	
Lory et al. (2022)	Tropical SE Pacific					5	
Löscher et al. (2016)	Tropical SW Pacific					225	31+4[a]
Löscher et al. (2020)	Bay of Bengal					18	
Lu et al. (2018)	Equatorial W Pacific					3	3[a]
Martínez-Pérez et al. (2016)	Tropical N Atlantic					84	14
Messer et al. (2016)	S Pacific					27	
Messer et al. (2021)	S Australian Gulf System			10		10	

Reference	Region	*Tricho-desmium*	Hetero-cystous	< 10 µm Diazotrophs	> 10 µm Diazotrophs	Whole Seawater	Depth-integrated data
Mills et al. (2020)	California Current System					4	
Moreira-Coello et al. (2017)	the coastal NW Iberian upwelling			30			10[a]
Mulholland et al. (2019)	NW Atlantic					402	242[a]
Needoba et al. (2007)	Temperate N Pacific					2	
Palter et al. (2020)	Gulf stream					7	
Raes et al. (2014)	E Indian					31	
Raes et al. (2020)	S Pacific					118	
Rahav et al. (2013b); Rahav et al. (2015)	Red Sea and E Mediterranean Sea					62	10
Rahav et al. (2013a)	Mediterranean Sea					8	
Rahav et al. (2016)	Mediterranean Sea						3[a]
Reeder et al. (2022)	S Baltic Sea					15	5
Riou et al. (2016)	N Atlantic					24	6
Sarma et al. (2020)	Bay of Bengal					2	
Sato et al. (2021)	Subarctic Sea of Japan; Sea of Okhotsk					31	3
Sato et al. (2022)	E Indian					73	18[a]
Saulia et al. (2020)	Tropical SW Pacific					71	71[a]
Selden et al. (2019)	Tropical NE Pacific					8	16[a]
Selden et al. (2021b)	NW Atlantic					93	26[a]
Selden et al. (2021a)	Tropical SE Pacific					125	19
Shiozaki et al. (2013)	W Pacific					50	10
Shiozaki et al. (2014a)	SW Pacific			40		42	
Shiozaki et al. (2014b)	Indian Ocean			26		26	6[a]
Shiozaki et al. (2015a)	NW Pacific					73	11
Shiozaki et al. (2015b)	N Pacific					112	22[a]
Shiozaki et al. (2017)	N Pacific					74	15[a]
Shiozaki et al. (2018b)	W Arctic					84	21[a]
Shiozaki et al. (2018c)	S Pacific					65	15[a]
Shiozaki et al. (2020)	Antarctic Coast					53	15
Singh et al. (2017)	Tropical NE Atlantic					52	13

Reference	Region	*Tricho-desmium*	Hetero-cystous	< 10 µm Diazotrophs	> 10 µm Diazotrophs	Whole Seawater	Depth-integrated data
Sipler et al. (2017)	Arctic Ocean					8	
Sohm et al. (2011)	S Atlantic					12	3[a]
Subramaniam et al. (2008)	Tropical N Atlantic						242[a]
Subramaniam et al. (2013)	Atlantic Ocean					96	24[a]
Tang et al. (2020)	N Atlantic					15	
Turk-Kubo et al. (2012)	Tropical N Atlantic			27			7
Turk-Kubo et al. (2021)	Southern California Current System			21		64	14
Wasmund et al. (2015)	S Atlantic						66[a]
Watkins-Brandt et al. (2011)	N Pacific						1[a]
Wen et al. (2022)	Tropical NW Pacific					143	22[a]
White et al. (2018)	Subtropical N Pacific					43	13[a]
Wilson et al. (2012)	N Pacific					9	4[a]
Wilson et al. (2017)	Subtropical N Pacific					33	
Wu et al. (2021)	Eastern Indian Ocean			48	48	48	7
Yogev et al. (2011)[b]	E Mediterranean Sea					16	32[a]
Zhang et al. (2015)	South China Sea					82	11
Zhang et al. (2019)	Tropical NW Pacific					87	9[a]
Part 2. Incubation period less than 24 hours							
Agawin et al. (2013)	Subtropical Atlantic				21	17	
Benavides et al. (2013b)	subtropical N Atlantic					38	
Benavides et al. (2014)	the coastal Namibian upwelling system					14	3
Bhavya et al. (2016)	Arabian Sea					4	
Biegala and Raimbault (2008)	SW Pacific			6	6	6	6
Bombar et al. (2011)	South China Sea					15	
Fernandez et al. (2015)	Central Chile Upwelling System					29	

Reference	Region	*Tricho-desmium*	Hetero-cystous	< 10 µm Diazotrophs	> 10 µm Diazotrophs	Whole Seawater	Depth-integrated data
Foster et al. (2013)	Subtropical N Pacific					3	
Foster et al. (2022a)	Tropical NW Atlantic					45	9
Foster et al. (unpublished data)	N Atlantic					24	5
Gandhi et al. (2011)	E Arabian Sea					28	7[a]
Halm et al. (2012)	S Pacific					43	10[a]
Kromkamp et al. (1997)	Indian Ocean						9[a]
Krupke et al. (2013)	Subtropical N Atlantic					6	
Krupke et al. (2014)	N Atlantic					42	44[a]
Kumar et al. (2017)	E Arabian Sea					12	3
Chen et al. (2014)	South China Sea						24[a]
Sahoo et al. (2021)	Bay of Bengal						6[a]
Saxena et al. (2020)	Bay of Bengal					32	8[a]
Singh et al. (2019)	E Arabian Sea					20	5[a]
Wang et al. (2021)	NW Atlantic					85	
Total		6	0	346	87	5414	1805

[a] Data are reported by data providers as depth-integrated N_2 fixation rates (unlabelled data computed ~~from~~ by integrating profiles of volumetric N_2 fixation rate data).

[b] N_2 fixation rate incubation time during 24-30 hrs.

Table ~~3~~5. Summary of new data points of cell-count-based ~~abundance in~~ abundances added to the version 2 of the database~~, including volumetric measurements for *Trichodesmium*, unicellular~~. The data were measured using microscopy-based method (method A), TSA/CARD-FISH (method B), flow cytometer (method C) or image recognition (method D). UCYNs include UCYN-A, UCYN-B and ~~heterocystous diazotrophs.~~unclassified UCYNs. Heterocystous cyanobacteria include Het-1, Het-2 and Het-3.

~~Reference~~	~~Region~~	~~Cell-count-based abundance~~				
Reference	Region	Method~~*Tricho-desmium*~~	*Tricho-desmium*~~Unicellular~~	~~*Richelia*~~UCYN	~~*Calothrix*~~Heterocystous cyanobacteria	Depth-integrated data
Biegala and Raimbault (2008)	SW Pacific	B		15		

~~Reference~~	~~Region~~	~~Cell-count-based abundance~~				
Reference	Region	Method~~*Tricho-desmium*~~	*Tricho-desmium*~~Unicellular~~	*Richelia*UCYN	~~*Calothrix*~~Heterocystous cyanobacteria	Depth-integrated data
~~Bif and Yunes (2017)~~Bif and Yunes (2017)	S Atlantic	~~16~~A	16			
~~Campbell et al. (2005)~~Campbell et al. (2005)	SW Pacific	A	462		259	33[a]
~~Detoni et al. (2016)~~Detoni et al. (2016)	S Atlantic	~~14~~A	14			
Dugenne et al. (2023); Gradoville et al. (2022)	N Pacific Subtropic Gyre	C	4	4	7	
~~Dupouy et al. (2011)~~Dupouy et al. (2011)	SW Pacific	~~18~~A	18			
Estrada et al. (2016)	Global	A	407		407	
~~Fernández et al. (2010)~~Fernández et al. (2010)	Global	A				40[a]
Foster et al. (2022a)	W tropical N Atlantic	A			37	9
Foster et al. (unpublished data)	N Atlantic	A			54	
~~Hegde et al. (2008)~~Hegde et al. (2008)	Bay of Bengal	~~135~~A	135			
~~Holl et al. (2007)~~Holl et al. (2007)	N Atlantic	A				10[a]
Jiang et al. (2017)	E China Sea	A	1174			252[a]
Jiang et al. (2023)	E China Sea	A	39		39	78[a]
~~Krupke et al. (2013)~~Krupke et al. (2013)	N Atlantic	B	~~9~~	9		
Le Moal and Biegala (2009)	Mediterranean Sea	B		17		
Le Moal et al. (2011)	Mediterranean Sea	B		18		
Lory et al. (2022)	S Pacific	~~3~~A	3			
~~Lu et al. (2018)~~Lu et al. (2018)	W ~~Equatorial~~Eq. Pacific	~~2~~A	2			
Martínez-Pérez et al. (2016)	Tropical N Atlantic	A	~~56~~	56		14

~~Reference~~	~~Region~~	~~Cell-count-based abundance~~				
Reference	Region	Method~~*Tricho-desmium*~~	*Tricho-desmium*~~Unicellular~~	*~~Richelia~~*UCYN	~~*Calothrix*~~Heterocystous cyanobacteria	Depth-integrated data
~~Masotti et al. (2007)~~Masotti et al. (2007)	~~Southwestern~~SW Pacific	~~20~~A	20			5
~~Mompeán et al. (2013)~~Mompeán et al. (2013)	N Atlantic	A				43[a]
Mompeán et al. (2016)	Global	A				141[a]
~~Pierella Karlusich et al. (2021)~~Karlusich et al. (2021)	Global	~~46~~D	46	~~46~~	~~35~~81	
~~Riou et al. (2016)~~Riou et al. (2016)	N Atlantic	B	~~20~~	20		5
~~Sahu et al. (2017)~~Sahu et al. (2017)	Bay of Bengal	~~14~~A	14			
~~Shiozaki et al. (2013)~~Shiozaki et al. (2013)	W Pacific	~~10~~A	10	~~12~~	12	
Shiozaki et al. (2015a)	NW Pacific	~~60~~A	60			10
~~Subramaniam et al. (2008)~~Subramaniam et al. (2008)	N Atlantic	A				162[a]
Tenório et al. (2018)	SW Pacific	A	81			19[a]
White et al. (2018)	N Pacific	~~83~~A	83	~~83~~	83	38
~~Wu et al. (2021)~~Wu et al. (2021)	Bay of Bengal	~~224~~A	224	~~224~~	224	
Total		~~645~~	~~85~~2812	~~365~~139	~~35~~1203	~~360~~859

[a] Data are reported by data providers as depth-integrated cell~~-count-based~~ abundance (unlabelled ~~data~~depth-integrated abundances computed from volumetric ~~cell-count-based abundance~~data).

Table 46. Summary of new data points of *nifH* based abundance in gene copy abundances added to the version 2 of the database, including volumetric measurements for *Trichodesmium*, unicellular. UCYNs include UCYN-A1, UCYN-A2, UCYN-B and heterocystous diazotrophsUCYN-C. Heterocystous cyanobacteria include Het-1, Het-2 and Het-3.

References	Region	*nifH* based abundances			
Reference	Region	*Trichodesmium*Trichodesmium	UnicellularUCYN	Heterocystous cyanobacteria	Depth-integrated data
Benavides et al. (2016a)	N Atlantic	13	30	15	
Bentzon-Tilia et al. (2015)Bentzon-Tilia et al. (2015b)	Baltic Sea		20		
Berthelot et al. (2017)	W Tropical Pacific	64	256	64	96
Bombar et al. (2011)	SSouth China Sea	18	36	18	
Bombar et al. (2015)Bombar et al. (2015)	N Pacific				32
Bonnet et al. (2015)	SW Pacific	87	261	87	84
Bonnet et al. (2023)	SW Pacific	66	132		44
Cabello et al. (2020)	Monterey Bay		200		
Confesor et al. (2022)[b]	W Florida Shelf	67			
Cerdan-Garcia et al. (2021)	N Atlantic	7	7		
Chen et al. (2019)	W Pacific	103	381	177	123
Cheung et al. (2020)	N Pacific	519	519		
Cheung et al. (2022)Cheung et al. (2022)	WesternW Bering Sea		58	29	
Church and Zehr (2020)Church and Zehr (2020)	N Pacific	968	1936	1936	605
Church et al. (2008)Church et al. (2008)	N Pacific				60
Detoni et al. (2022)Detoni et al. (2022)	WSSW Atlantic	70	140	70	72
Dugenne et al. (2023); Gradoville et al. (2022)	N Pacific Subtropic Gyre	72	216	216	112
Foster et al. (unpublished data)	South China Sea	99	224	350	158
Gradoville et al. (2020)Gradoville et al. (2020)	N Pacific	43	85	28	
Hallstrøm et al. (2022)Hallstrøm et al. (2022)	NE Atlantic				42[a]

References	Region	*nifH* based abundances			
Reference	Region	*Trichodesmium*	UCYN	Heterocystous cyanobacteria	Depth-integrated data
Halm et al. (2012)	S Pacific Gyre	8	16		
Hamersley et al. (2011)	S California Bight	6	12	6	
Harding et al. (2018)	Arctic Ocean		39		
Hashimoto et al. (2016)	Seto Inland Sea		176		
Henke et al. (2018)	W Tropical S Pacific		142		
Krupke et al. (2013)	N Atlantic		24		3
Liu et al. (2020)	SSouth China Sea	49	98		33
Lory et al. (2022)	W Tropical S Pacific	3	3		
Lu et al. (2018)	W Tropical Pacific	3	6	3	
Martínez-Pérez et al. (2016)	N Tropical Atlantic	84	252	84	70
Messer et al. (2021)	S Australian Gulf		20		
Mills et al. (2020)	Coast of S California	4	12	4	
Moisander et al. (2014)	S Pacific	174	348	174	92
Moore et al. (2018)	Tropical Atlantic	104	312	208	
Moreira-Coello et al. (2017)	N Pacificthe coastal NW Iberian upwelling		20		20[a]
Palter et al. (2020)	Gulf stream	24	24		
Ratten et al. (2015)	N Atlantic	9	27	9	10
Reeder et al. (2022)	Baltic Sea		15	15	
Sato et al. (2021)	Subarctic Sea		31		3
Sato et al. (2022)	Eastern Indian Ocean	73	73		36
Saulia et al. (2020)	SW Pacific	71	213	143	
Scavotto et al. (2015)			2		
Selden et al. (2021b)	Atlantic Bight	23	69	23	

References	Region	*nifH* based abundances			
Reference	Region	*Tricho desmium**Trichodesmium*	UnicellularUCYN	Heterocystous cyanobacteria	Depth-integrated data
Selden et al. (2022)	Arctic Ocean		40		
Shiozaki et al. (2014a)Shiozaki et al. (2014b)	Arabian Sea	26	52		18
Shiozaki et al. (2014b)Shiozaki et al. (2014c)	S China Sea	171	342		72[a]
Shiozaki et al. (2015a)	Temperate N Pacific	73	146		33
Shiozaki et al. (2017)	N Pacific	74	222	74	90
Shiozaki et al. (2018a)	Kuroshio	46	138	46	
Shiozaki et al. (2018b)Shiozaki et al. (2018b)	W Arctic		84		21
Shiozaki et al. (2018c)	S Pacific	94	285	95	95
Shiozaki et al. (2020)Shiozaki et al. (2020)	Antarctic sea ice		53		
Sohm et al. (2011)Sohm et al. (2011)	S Atlantic Gyre		58		
Stenegren et al. (2017)	W Tropical N Atlantic			235	61
Stenegren et al. (2018)Stenegren et al. (2018)	W Tropical S Pacific	108	402	120	108
Tang et al. (2020)Tang et al. (2020)	N Atlantic	42	42		
Turk-Kubo et al. (2014)Turk-Kubo et al. (2014)	E Tropical S Pacific	60	159	57	53
Turk-Kubo et al. (2021)	Coast of S California	190	588	202	135
Wen et al. (2017)Wen et al. (2017)	W Pacific	22	44	22	
Wen et al. (2022)Wen et al. (2022)	W Pacific	130	390	130	110[a]
White et al. (2018)	N Pacific				34
Wu et al. (2019)Wu et al. (2019)	Bay of Bengal	68	63		19
Total		30823935	74629543	32814640	19142544

[a] Data are reported by data providers as depth-integrated *nifH* based abundance gene copy abundances (unlabelled datadepth-integrated abundances computed from volumetric *nifH* based abundancedata).

~~**2.2 Nitrogen fixation rates**~~

~~Marine N_2 fixation rates are commonly measured using the acetylene reduction assay or $^{15}N_2$ assimilation method (Mohr et al., 2010; Montoya et al., 1996; Capone, 1993). The acetylene reduction assay estimates gross N_2 fixation rates indirectly by using a factor of 3:1 (Stewart et al., 1967; Bhavya et al., 2019) or 4:1 (Zehr and Capone, 2021) to convert acetylene reduction rates to N_2 fixation rates. When using the $^{15}N_2$ assimilation method, samples are incubated in seawater with $^{15}N_2$ labelled gas; the $^{15}N/^{14}N$ ratio of particulate nitrogen is measured at the beginning and at end of incubation to calculate the N_2 fixation rate (Capone and Montoya, 2001). Compared to the $^{15}N_2$ assimilation method, the acetylene reduction assay is easier to conduct, but acetylene used in the assay can potentially impact the metabolic activities of diazotrophs (Bhavya et al., 2019). The conversion factor between acetylene reduction and N_2 fixation is also controversial (Flett et al., 1976; Giller, 1987; Hardy et al., 1973). The $^{15}N_2$ assimilation method has higher sensitivity and requires a shorter incubation time than the acetylene reduction assay, particularly when N_2 fixation is low (Montoya et al., 1996). The $^{15}N_2$ assimilation method, however, needs to concentrate cells for signal detection, which can potentially damage cells and underestimate N_2 fixation rates (Bhavya et al., 2019). Hence, the $^{15}N_2$ assimilation method only measures the fixed N in particulate forms and ignores the N that is fixed but then excreted by diazotrophs during incubation, which, however, can theoretically be counted by the acetylene reduction assays (Mulholland, 2007).~~

~~The conventional $^{15}N_2$ assimilation method was conducted by bubbling $^{15}N_2$ labelled gas. It was later found to be difficult to reach complete solubility equilibrium over a short incubation time, leading to serious underestimations of N_2 fixation rates (Mohr et al., 2010; Großkopf et al., 2012; Wannicke et al., 2018). This "bubble" method was then modified by dissolving $^{15}N_2$-labelled gas in seawater for an adequate period to ensure that it reached solubility equilibrium (Mohr et al., 2010). Recently, this dissolution method was reported to have risks of introducing heavy metal pollution and affecting the growth of diazotrophs (White et al., 2020).~~

[b] *rnpB* gene copies were determined. **Table 7.** Summary of data points of cell-specific N_2 fixation rates added to the version 2 of the database. The rates were measured either by using the combination of CARD-FISH and nanoSIMS (method A), or via the measurements of bulk N_2 fixation rates incubated with known number of diazotrophic cells (method B) (*see* **Section 2.3**). Note that all the data were reported as N_2 fixation rates per cell, except for Filella et al. (2022) in which biomass-normalized rates in unit of d^{-1} were reported.

Reference	Region	Method	*Trichodesmium*	UCYN-A	UCYN-A1	UCYN-A2	UCYN-B	*Richelia*	*Calothrix*	Unclassified Cyanobacteria	NCDs
Benavides et al. (2022a)	Tropical SW Pacific	A	6								
Benavides et al. (2017)	SW Pacific	A	2								
Bonnet et al. (2018)	Tropical SW Pacific	A	3				2				

Reference	Region	Method	*Trichodesmium*	UCYN-A	UCYN-A1	UCYN-A2	UCYN-B	*Richelia*	*Calothrix*	Unclassified Cyanobacteria	NCDs
Filella et al. (2022)	S Pacific Gyre	A	12				12				
Foster et al. (2011)	N Pacific	A					2	18	2		
Foster et al. (2013)	N Pacific	A					6				
Foster et al. (2022a)	Tropical NW Atlantic	A						39			
Gradoville et al. (2020)	N Pacific	A			5						
Gradoville et al. (2021)	N Pacific	A			17						
Harding et al. (2018)	Arctic Ocean	A				2					
Harding et al. (2022)	Subtropical N Pacific	A								40	34
Krupke et al. (2013)	Subtropical N Atlantic	A		4			2				
Krupke et al. (2015)	Subtropical NE Atlantic	A		1							
Martínez-Pérez et al. (2016)	Tropical N Atlantic	A	101		57	10					
Mccarthy and Carpenter (1979)	N Atlantic	B	24								
Mills et al. (2020)	California Current System	A			15	9					
(Turk-Kubo et al., 2021)	Southern California Current System	A			26	17					
Total			148	10	115	38	24	57	2	40	34

2.2 Quality control

The data of N_2 fixation rates and diazotrophic abundance in the database spanned over several orders of magnitude. Extremely high rate and abundance values of both usually occurred during algal blooms, and zero values indicated that diazotrophic activity was below detection or truly absent at the sampling time and stations. The positive-value data were first logarithmically transformed and then analyzed for outliers, considering that they were approximately log-normally distributed (**Fig. S1-S5**). For each parameter, we used Chauvenet's criterion to identify suspicious outliers whose probability of deviation from the means is lower than $1/2n$, where n is the number of data points (Glover et al., 2011). Because N_2 fixation rates and diazotroph abundances in the ocean can be extremely low, this filtering only applied to data on the high side. Although these outliers (labelled in the database) could be true values, we flagged them to remind users for caution.

2.3 Nitrogen fixation rate data

The commonly used methods for marine N_2 fixation rates include $^{15}N_2$ tracer methods and acetylene reduction assay (Mohr et al., 2010; Montoya et al., 1996; Capone, 1993). However, in the last decade, the community has turned largely to the use of $^{15}N_2$ tracer methods. The acetylene reduction assay estimates gross N_2 fixation rates indirectly from the reduction of acetylene to ethylene. Theoretical conversion factors of 3:1 or 4:1 have been used to convert acetylene reduction rates to N_2 fixation rates (Postgate, 1998; Capone, 1993; Wilson et al., 2012), although a wide range of conversion factors from 0.93 to 56 have been reported (e.g., Mague et al., 1974; Graham et al., 1980; Montoya et al., 1996; Capone et al., 2005; Mulholland et al., 2006; Wilson et al., 2012). When using the $^{15}N_2$ tracer method, samples are incubated in seawater with $^{15}N_2$ gas; the $^{15}N/^{14}N$ ratio of particulate nitrogen is measured at the beginning and at end of the incubation to calculate the N_2 fixation rate (Capone and Montoya, 2001). Most measurements using the $^{15}N_2$ tracer method only counted the fixed N in particulate forms and ignored the N that was fixed but then excreted by diazotrophs in form of dissolved organic N (DON) during incubation, which could theoretically be counted by the acetylene reduction assays (Mulholland, 2007). In some studies using the $^{15}N_2$ tracer method, this missing N was counted by also measuring the ^{15}N enrichment in DON (Berthelot et al., 2017; Benavides et al., 2013a; Berthelot et al., 2015; Benavides et al., 2013b).

Compared to the $^{15}N_2$ tracer method, the acetylene reduction assay needs a shorter incubation time. However, in addition to the uncertainty in converting ethylene production to N_2 fixation, the purity of acetylene gas, trace ethylene contamination, and the Bunsen gas solubility coefficient of produced ethylene can also affect the accuracy of estimated N_2 fixation rates (Hyman and Arp, 1987; Breitbarth et al., 2004; Kitajima et al., 2009). Acetylene used in the assay can even impact the metabolic activities of diazotrophs (Giller, 1987; Hardy et al., 1973; Flett et al., 1976; Staal et al., 2001). Moreover, the acetylene reduction assay needs to pre-concentrate cells for signal detection when diazotrophic biomass is low, which may lead to underestimated N_2 fixation rates by perturbing cells during concentration and filtration (e.g., Capone et al., 2005; Barthel et al., 1989; Staal et al., 2007). In recent years, the acetylene reduction assay has undergone significant advancements.

The sensitivity of ethylene detection has been improved by utilizing a reduced gas analyzer (Wilson et al., 2012) and by using highly purified acetylene gas to minimize the ethylene background (Kitajima et al., 2009). However, the preparation of high-purity acetylene with low level of ethylene contamination remains a challenge. More recently, a new method named Flow-through incubation Acetylene Reduction Assays by Cavity ring-down laser Absorption Spectroscopy (FARACAS) has been introduced for high-frequency measurements of aquatic N_2 fixation (Cassar et al., 2018). This method involves continuous flow-through incubations and spectral monitoring of the acetylene reduction to ethylene. By employing short-duration flow-through incubations without cell preconcentration, potential artifacts are minimized. This approach also allows for near real-time estimates, enabling adaptive sampling strategies.

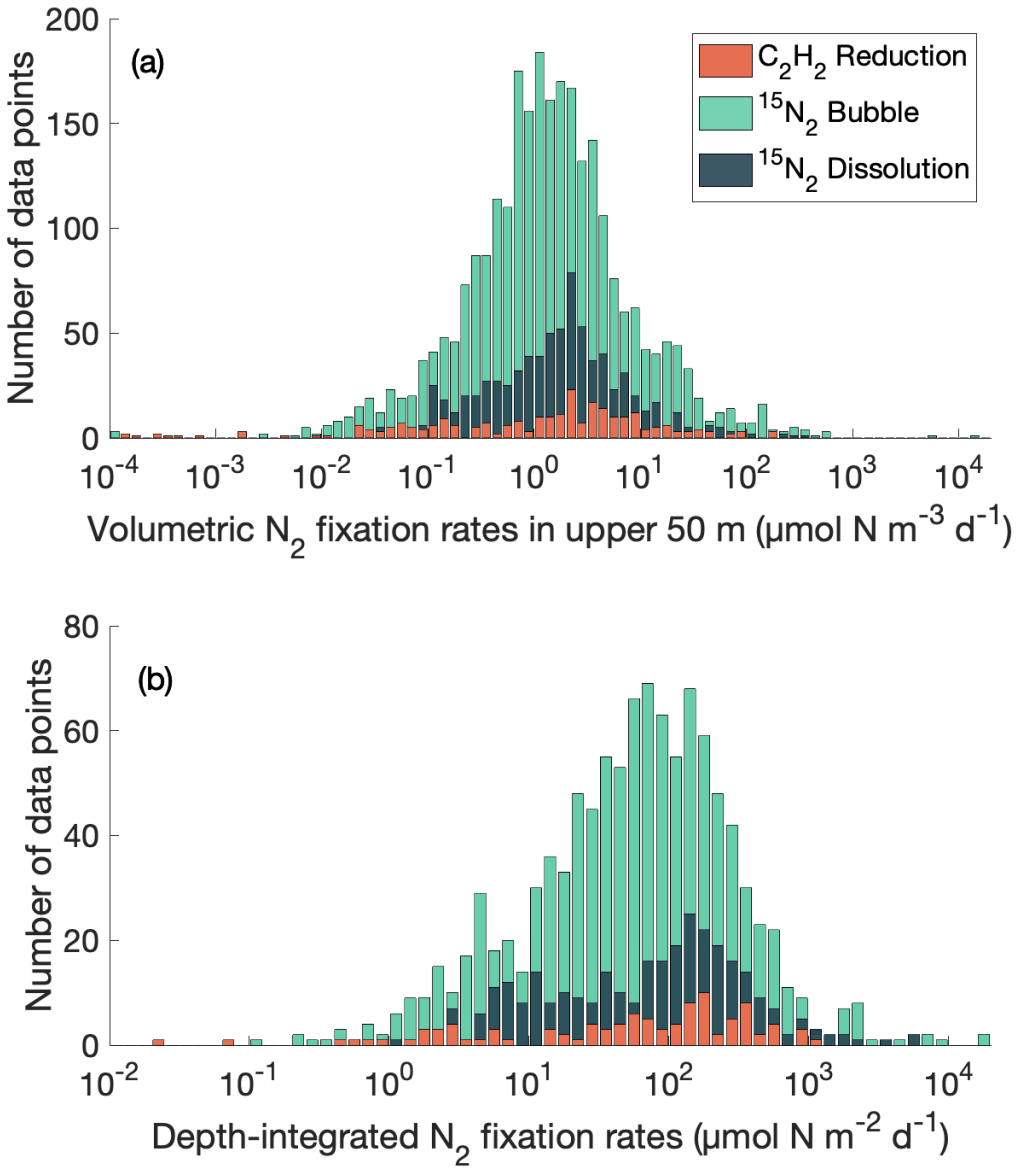
The original $^{15}N_2$ tracer method involved addition of a known volume of $^{15}N_2$-labelled bubbles to the incubation bottle (named *original $^{15}N_2$ bubble method* hereafter). However, this method was later found to underestimate rates because N_2 gas solubility is low and tracer additions take a long time to equilibrate (Mohr et al., 2010; Großkopf et al., 2012; Jayakumar et al., 2017). To address this issue, the *$^{15}N_2$ dissolution method* has been employed, which involves pre-preparing $^{15}N_2$-enriched seawater to maintain a constant $^{15}N_2$ atom% enrichment throughout the incubation (Mohr et al., 2010), similar to the method described in Glibert and Bronk (1994). However, the $^{15}N_2$ dissolution method does not always yield higher N_2 fixation rates than the original $^{15}N_2$ bubble method (Table S4 in Großkopf et al., 2012; Saulia et al., 2020); it is still not conclusive what control the magnitude of the underestimation (if it exists) by the original $^{15}N_2$ bubble method. Compared to the original $^{15}N_2$ bubble method, the $^{15}N_2$ dissolution method is more susceptible to the introduction of contaminants (e.g., metals) during the preparation of the $^{15}N_2$ inoculum due to its more complex process, which can alter the diazotrophic activities and abundance, thereby impacting the accuracy of N_2 fixation measurements (Dabundo et al., 2014; Klawonn et al., 2015). For example, Needoba et al. (2007) reported that a low but detectable amount of Fe^{3+} contamination can be measured when protecting the needle of the gas-tight syringe with a commercially available tubing. Additionally, pH and other chemical properties of the inoculum may be altered during its preparation, further affecting the measurements of N_2 fixation. Despite these limitations, the $^{15}N_2$ dissolution method remains the predominant assay for measuring N_2 fixation rate due to its ability to satisfy the fundamental assumption of constant $^{15}N_2$ atom% enrichment over the incubation period.

More recently, a modified $^{15}N_2$ bubble method, known as the *$^{15}N_2$ bubble release method*, has been proposed as an alternative to the $^{15}N_2$ dissolution method (Klawonn et al., 2015; Chang et al., 2019; Selden et al., 2019). This method involves adding $^{15}N_2$ gas to the incubation bottles and mixing for a brief period (~15 min) to facilitate $^{15}N_2$ equilibration, then removing the gas bubble. Compared to the original $^{15}N_2$ bubble method, the $^{15}N_2$ bubble release method ensures a uniform $^{15}N_2$ atom% enrichment throughout the incubation. Moreover, it causes less interference with the incubation matrix than the $^{15}N_2$ dissolution method. However, the mixing of incubation bottles required to stimulate gas dissolution has been suggested to negatively affect diazotrophs, although no robust studies have yet been performed to assess this criticism (Wannicke et al., 2018; White et al., 2020). Moreover, the $^{15}N_2$ bubble release method requires a handling step and additional costs for preparing tracers may be another challenge for researchers (White et al., 2020). Ultimately White et al. (2020) “advise employing either the

dissolution or bubble release method, whichever is best suited to the specific research objectives and logistical constraints" with additional recommendations on the need for determination of detection limits for all rate measurements.

We compared volumetric N_2 fixation rates in the upper 50 m and depth-integrated N_2 fixation rates in the database measured using the ~~above three methods~~acetylene reduction assays, the original $^{15}N_2$ bubble method and the $^{15}N_2$ dissolution method, and found that they ~~basically~~ span a similar range ~~of magnitude~~ (**Fig. 1**). ~~The results~~Meanwhile, in the analysis for volumetric N_2 fixation rates in upper 50 m, the peak of the log-normal distributions of ~~a~~the measurements using the $^{15}N_2$ dissolution method was approximately double that of the original $^{15}N_2$ bubble method (**Fig. 1a**). The measurements using the $^{15}N_2$ bubble release method were limited to several study sites and their distribution was thus not presented in this study. A further analysis comparing the original $^{15}N_2$ bubble method and the $^{15}N_2$ dissolution ~~methods~~method will be ~~discussed~~presented later.

~~Most N_2 fixation rates were originally reported as daily rates. For those reported hourly N_2 fixation rates in the daytime, we converted them to daily rates by multiplying by 12 hours. Some studies also reported hourly rates at night. We multiplied these night rates by 12 hours and added them to the daytime rates.~~


(a)
C_2H_2 Reduction
$^{15}N_2$ Bubble
$^{15}N_2$ Dissolution
Number of data points
200
150
100
50
0
10^{-4} 10^{-3} 10^{-2} 10^{-1} 10^{0} 10^{1} 10^{2} 10^{3} 10^{4}
Volumetric N_2 fixation rates in upper 50 m (μmol N m^{-3} d^{-1})
(b)
Number of data points
80
60
40
20
0
10^{-2} 10^{-1} 10^{0} 10^{1} 10^{2} 10^{3} 10^{4}
Depth-integrated N_2 fixation rates (μmol N m^{-2} d^{-1})

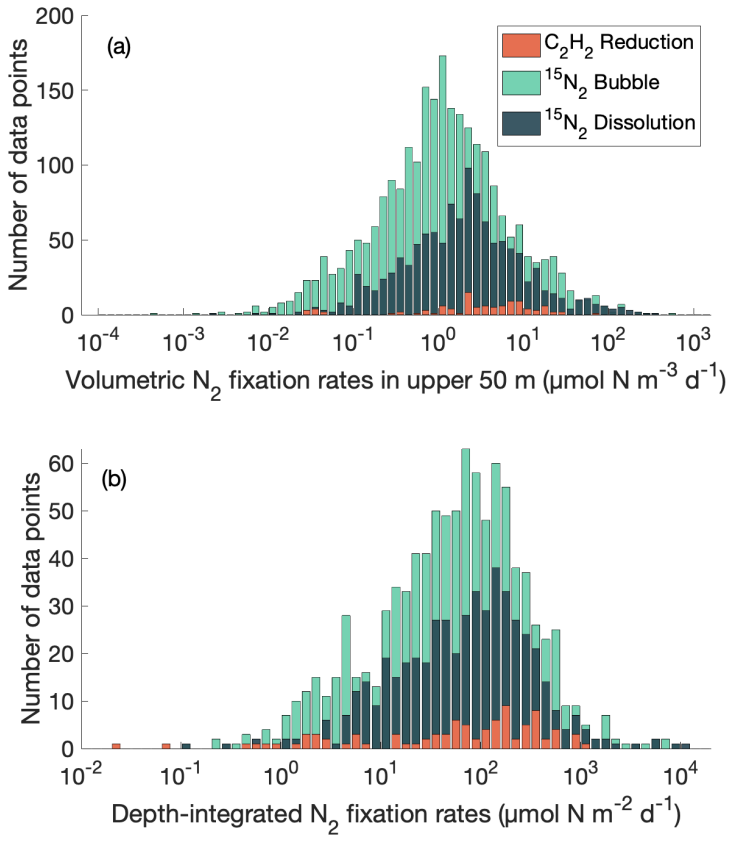


Figure 1. Distribution of N_2 fixation rates measured using ~~different methods.~~the acetylene (C_2H_2) reduction assays, the original $^{15}N_2$ bubble method and the $^{15}N_2$ dissolution method. (a) Volumetric data in upper 50 m; (b) depth-integrated data. Only rates measured with incubation periods of 24 hours are shown. Please note that the bars in plot do not represent cumulative data.

~~2.3~~

The majority of N_2 fixation rates (9,405) were measured with incubation periods of 24 hours and were reported as daily rates. In contrast, 2,416 samples were incubated for less than 24 hours and hourly N_2 fixation rates were reported. Diel cycles of N_2 fixation vary among samples and/or diazotrophic groups, and substantial errors may be introduced when extrapolating N_2 fixation rates incubated for less than 24 hours to daily rates (White et al., 2020). Therefore, the N_2 fixation rates measured with incubation periods of less than 24 hours were collected into separated datasheets in our database and were not used in further analyses within this study. Please note that the incubation periods of whole diurnal cycles (e.g., 24, 48, or 72 hours) were used in Konno et al. (2010). The samples in Yogev et al. (2011) were incubated between 24 to 30 hours. The reported daily N_2 fixation rates by these two studies were also included in the 24-hour datasheets and were used in our estimation of the global marine N_2 fixation rate (see below).

Cell-specific N_2 fixation rates of diazotrophs (or symbioses) were mostly measured using catalyzed reporter deposition fluorescence in-situ hybridization (CARD-FISH) and nanoscale secondary ion mass spectrometry (nanoSIMS), in combination with $^{15}N_2$ addition experiments (Mills et al., 2020; Berthelot et al., 2019). Using specific oligonucleotide probes, CARD-FISH enables the visualization and location of the regions of interest in diazotrophs at a single-cell level under epifluorescence microscope. This is subsequently prepared for the secondary electron image in nanoSIMS analysis. Importantly, the handling, fixation and processing of the samples with CARD-FISH, has been demonstrated to significantly impact the enrichment measured by nanoSIMS (see Musat et al. 2014; Woebken et al. 2015; Meyer et al. 2021). The nanoSIMS technique detects the enrichment of ^{15}N atoms in the targeted regions, allowing for the calculation of the cell-specific rate. Additionally, in one study, handpicked *Trichodesmium* colonies or trichomes were incubated, and the measured total N_2 fixation rates were normalized to number of cells (Mccarthy and Carpenter, 1979).

2.4 Estimation of the global marine N_2 fixation rate

Using these data, we performed a first-order estimation of the global marine N_2 fixation rate. In a previous study (Luo et al., 2012), version 1 was utilized to estimate the global marine N_2 fixation rate, which included all the depth-integrated N_2 fixation rates. However, in this study, we employed more rigorous criteria to estimate the global rate using both version 1 and version 2, taking into account the reliability of different N_2 fixation rate data discussed in the preceding section. Specifically, we exclusively used depth-integrated N_2 fixation rates that met the following criteria: (1) measurements were taken from whole seawater samples, (2) incubation periods of 24 hours were used, and (3) the three $^{15}N_2$-based methods were employed, although we acknowledged that the rates obtained using the original $^{15}N_2$ bubble method might be underestimated. N_2 fixation rates obtained through the acetylene reduction method were excluded from this estimate due to the significant uncertainties described above.

Applying these criteria, we selected 309 and 1,642 depth-integrated N_2 fixation rates from version 1 and version 2, respectively. The much more data in version 2 potentially provided more constraints on estimating global marine N_2 fixation. We applied Chauvenet's criterion to identify outliers, using the log-transformed values of the selected data (*see* **Section 2.2**). As a result, two high-value outliers were removed in version 1 (one in North Pacific and one in South Pacific) while no outliers were detected in version 2. This difference can be attributed to the larger number of data samples in version 2, which allowed for a more relaxed threshold in identifying outliers.

The estimation of the global marine N_2 fixation rate involved four steps. First, we calculated the arithmetic or geometric means of depth-integrated N_2 fixation rates within each 3° latitude × 3° longitude bin. Second, these mean values were further averaged using either arithmetic or geometric methods to determine the mean N_2 fixation rates for different ocean basins, which included the North Atlantic, South Atlantic, North Pacific, South Pacific, Indian, Arctic, Southern Oceans, and the Mediterranean Sea. Third, we multiplied the arithmetic or geometric mean of each basin by its respective area to estimate the total N_2 fixation rate for that specific basin, except when there was insufficient spatial coverage available. Finally, we obtained

the global marine N_2 fixation rate by summing up the individual rates calculated for each basin, with the errors associated with basin rates propagated properly (Glover et al., 2011).

In the first two steps, the geometric means were derived from positive N_2 fixation rates (NF_+): if μ and SE represented the mean and standard error of $\ln(NF_+)$, respectively, the geometric mean was e^{μ}. The confidence interval for the geometric mean, based on the standard error, ranged between e^{μ}/e^{SE} and $e^{\mu} \cdot e^{SE}$ (Thomas, 1979). To address the issue of not including zero-value N_2 fixation rates, we adjusted the geometric means by multiplying them with the percentage of zero-value data within each 3° × 3° bin (in the first step) or within each basin (in the second step).

2.5 Diazotrophic abundance data

Diazotroph cell ~~abundance was counted mainly~~abundances were determined by using standard light microscopy, and in some cases by using epifluorescence microscopy ~~with the aid of color excitation.~~. A recent study used machine learning techniques to detect and enumerate diazotrophs in a large dataset of microscopic images ~~(Pierella Karlusich et al., 2021).~~(Karlusich et al., 2021). In the original database, only the cell ~~counted abundance~~ abundances of *Trichodesmium* and heterocystous cyanobacteria ~~was~~were recorded. ~~The updated database~~Version 2 also included ~~a dataset~~datasets of enumerated abundance of all UCYN ~~A after staining~~ groups detecting them by TSA (Tyramid Signal Amplification)-FISH using the specific DNA probe UCYN-238 (Biegala and Raimbault, 2008; Le Moal and Biegala 2009; Le Moal et al., 2011; Riou et al. 2016). This method is also called CARD-FISH ~~assays~~and was used to specifically enumerated UCYN-A (Martínez-Pérez et al., 2016; Biegala and Raimbault, 2008; Le Moal et al., 2011). (**Table ~~3~~) ~~(Martínez-Pérez et al., 2016), although this unicellular diazotroph remained uncultivated.~~5**).

Cell abundance of *Trichodesmium* ~~cell count abundance in our database~~ was recorded as the number of trichomes per volume of water in our database, although it was also reported in ~~the literature~~some studies as the number of cells or colonies per volume of water. In the latter cases, the data were converted to trichomes per volume of water by using a commonly used factor of 200 (132–241) trichomes colony^{-1} ~~or 100 cells trichome~~$^{-1}$(Letelier and Karl, 1996), similar to the conversion used in the original database (Luo et al., 2012).

~~The abundance of heterocystous cyanobacterial cells was recorded in this database. When the number of symbioses was reported in several studies, we assumed that 2 and 5 *Richelia* heterocystous cells were associated with each *Hemiaulus* and *Rhizosolenia* cell, respectively (Villareal et al., 2011), and 5 *Calothrix* heterocystous cells were associated with each *Chaetoceros* cell (Tuo et al., 2021).~~

~~In measurements of *nifH* copies, different qPCR assays were designed to target specific diazotrophic groups (Church et al., 2005; Foster et al., 2007), mainly including *Trichodesmium*, UCYN subgroups (A, B and C) and heterocystous groups (**Table 4**). More recently, UCYN-A was found to have three sublineages, UCYN-A1, UCYN-A2, and UCYN-A3, with clade UCYN-A1 sharing the same genome as previously targeted UCYN-A (Thompson et al., 2014). UCYN-A1 and UCYN-A2 have significant distinctions in the sizes and species of their symbiotic hosts, with the former living in relatively smaller hosts~~

~~(Thompson et al., 2014; Martínez-Pérez et al., 2016; Cornejo-Castillo et al., 2016). The *nifH* copies of these two sublineages were included in our database. The *nifH* copies of UCYN-A3 were not listed separately in our database because there were only very few measurements reported. Three major associations of heterocystous groups were also marked for the *nifH* data, including *Richelia-Hemiaulus* (het-1), *Richelia-Rhizosolenia* (het-2) and *Calothrix-Chaetoceros* (het-3).~~

~~2.4 Quality control~~

~~The data of N_2 fixation rates and diazotrophic abundance in the database ranged by several orders of magnitude. Extremely high values usually occurred during algal blooms, and zero values indicated that diazotrophic activity was below detection or true absence at the sampling time and stations. The data (nonzero value) were first logarithmically transformed and then analyzed for outliers, considering that they were likely log-normally distributed (Fig. S1-S6). For each parameter, we used Chauvenet's criterion to identify suspicious outliers whose probability of deviation from the means is lower than $1/2n$, where n is the number of data points (Glover et al., 2011). Because N_2 fixation rates and diazotroph abundances in the ocean can be extremely low, this filtering only applies to data on the high side. Although these outliers (labelled in database) may be true values, we flagged them to remind users for caution. These outliers were also not used in our estimation of global ocean N_2 fixation rates.~~

~~3. Result~~

The abundance of heterocystous cyanobacterial cells was also recorded in this database. When the number of DDAs was reported in several studies, we assumed that 2 (reported range: 1–2) and 5 (reported range: 1–5) *Richelia* spp. filaments associated with each *Hemiaulus* and *Rhizosolenia* cell, respectively (Villareal et al., 2011; Caputo et al., 2019), and 5 (reported range: 3–10) *Richelia rhizosoleniae* filaments were associated with each *Chaetoceros* cell (Tuo et al., 2021; Caputo et al., 2019). *Richelia* have terminal heterocysts, and the number of vegetative cells varies depending on the host diatom. In *Hemiaulus* and *Chaetoceros* spp. diatoms, *Richelia* filaments are shorter (e.g., 3-4 vegetative cells), compared to in *Rhizosolenia*, the *Richelia* filaments are longer (e.g., 5-6 vegetative cells) (Foster et al., 2022b).

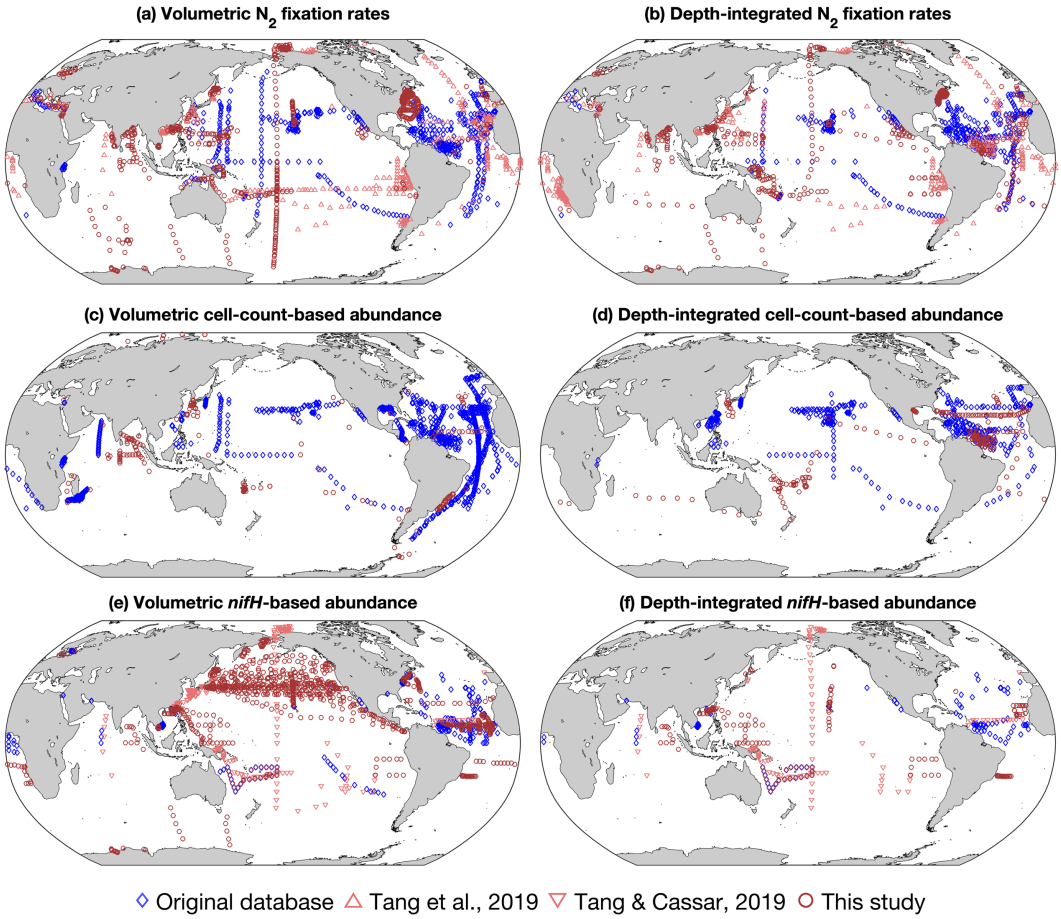
In measurements of *nifH* gene copy abundances, different qPCR or ddPCR assays were designed to target specific diazotrophic groups (Church et al., 2005a; Foster et al., 2007; Gradoville et al., 2017; Benavides et al., 2016a), mainly including *Trichodesmium*, UCYN subgroups (A1, A2, B, and C) and heterocystous groups (het-1, het-2, and het-3) (**Table 6**).

All the uncertainties reported in this paper reflect one standard error of the means unless specified otherwise.

3. Results

3.1 Data distribution

~~Version~~The version 2 of the database ~~approximately doubled the number of~~significantly expanded N_2 fixation rate ~~data in the original database~~measurements, filling spatial gaps particularly in the Indian Ocean and the Southern ~~Oceans (~~Hemisphere (**Table 1; Figs. 2a~~, ~~2b~~ b, 3a~~–b~~). The number of depth-integrated N_2 fixation rate measurements was tripled (**Table 1; Figs. 2b & 3b**). ~~The number of depth-integrated N_2 fixation rate data tripled, potentially providing more constraints on estimating global marine N_2 fixation (**Figs. 2b & 3b**). *NifH* based abundance consisted of the~~The largest fraction of new data~~, mostly in~~ derived from inclusion of *nifH* gene abundances, in particular data contributions from the Pacific and Atlantic Oceans (**Table 3; Figs. 2e~~, ~~f~~ ~~&~~, 3e~~, ~~f**). Compared to other parameters, ~~a relatively moderate amount of new cell count data was added and mainly distributed in the Indian and Atlantic Oceans (~~the new database contained only a modest increase in new cell abundances, mostly from the subtropical oceans (**Table 2; Figs. 2c~~, ~~d~~ ~~&~~, 3c~~, ~~d**). Overall, there ~~were~~remained more limited data on N_2 fixation and diazotrophic abundance in the ~~Southern Ocean~~Arctic and Southern Oceans, with a number of rate measurements reporting values below detection limits.


(a) Volumetric N_2 fixation rates
(b) Depth-integrated N_2 fixation rates
(c) Volumetric cell-count-based abundance
(d) Depth-integrated cell-count-based abundance
(e) Volumetric *nifH*-based abundance
(f) Depth-integrated *nifH*-based abundance
◇ Original database △ Tang et al., 2019 ▽ Tang & Cassar, 2019 ○ This study

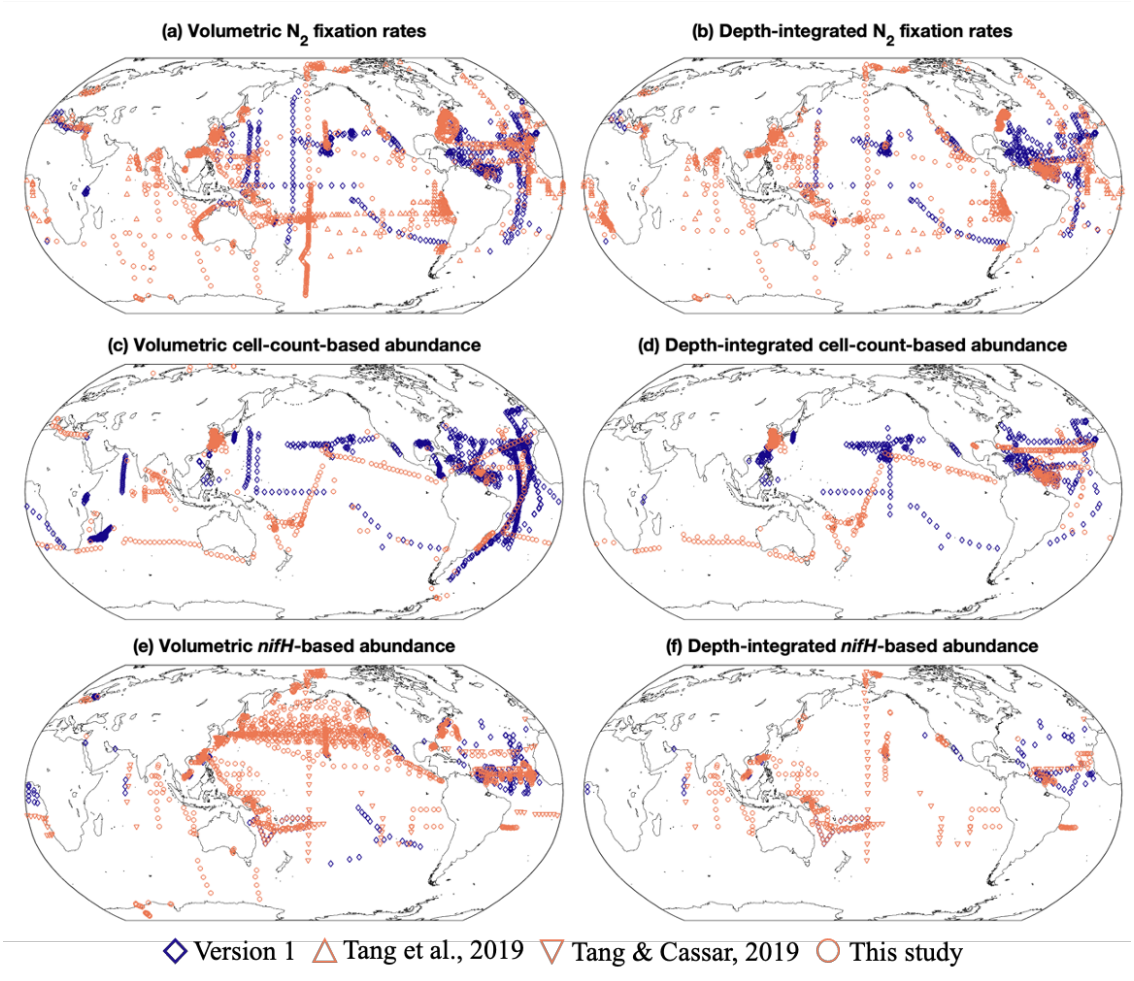
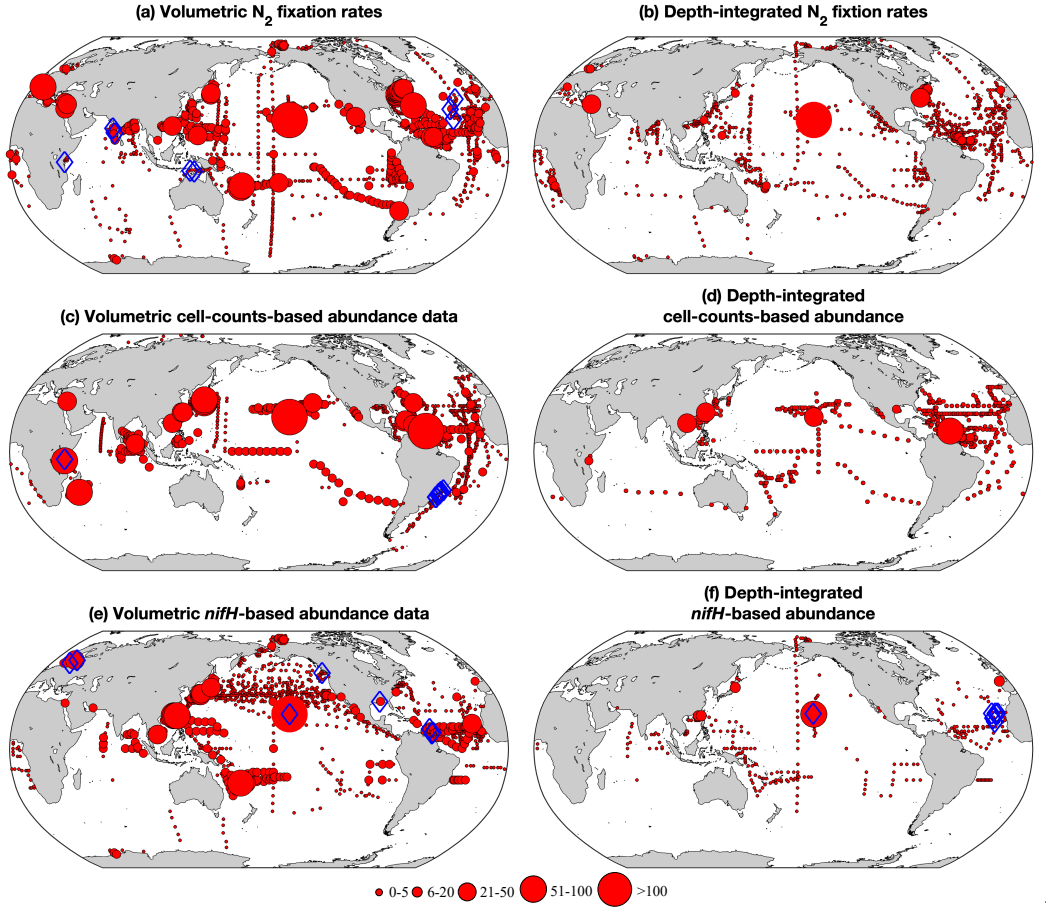


Figure 2. Spatial distribution of volumetric and depth-integrated number of data ~~points~~point in the version 2 of the diazotrophic database binned in 1~~°×1°°~~ latitude × 1° longitude grids. (**a-b**) N_2 fixation rates, (**c-d**) cell ~~counts~~abundance, and (**e-f**) *nifH* gene ~~copies~~copy abundance. The data sources include the original version of this database (Luo et al., 2012) (blue diamonds), two compiled datasets (Tang et al., 2019; Tang and Cassar, 2019) (~~red~~orange triangles) and this study (~~red~~orange circles).


(a) Volumetric N_2 fixation rates
(b) Depth-integrated N_2 fixtion rates
(c) Volumetric cell-counts-based abundance data
(d) Depth-integrated cell-counts-based abundance
(e) Volumetric *nifH*-based abundance data
(f) Depth-integrated *nifH*-based abundance
0-5
6-20
21-50
51-100
>100

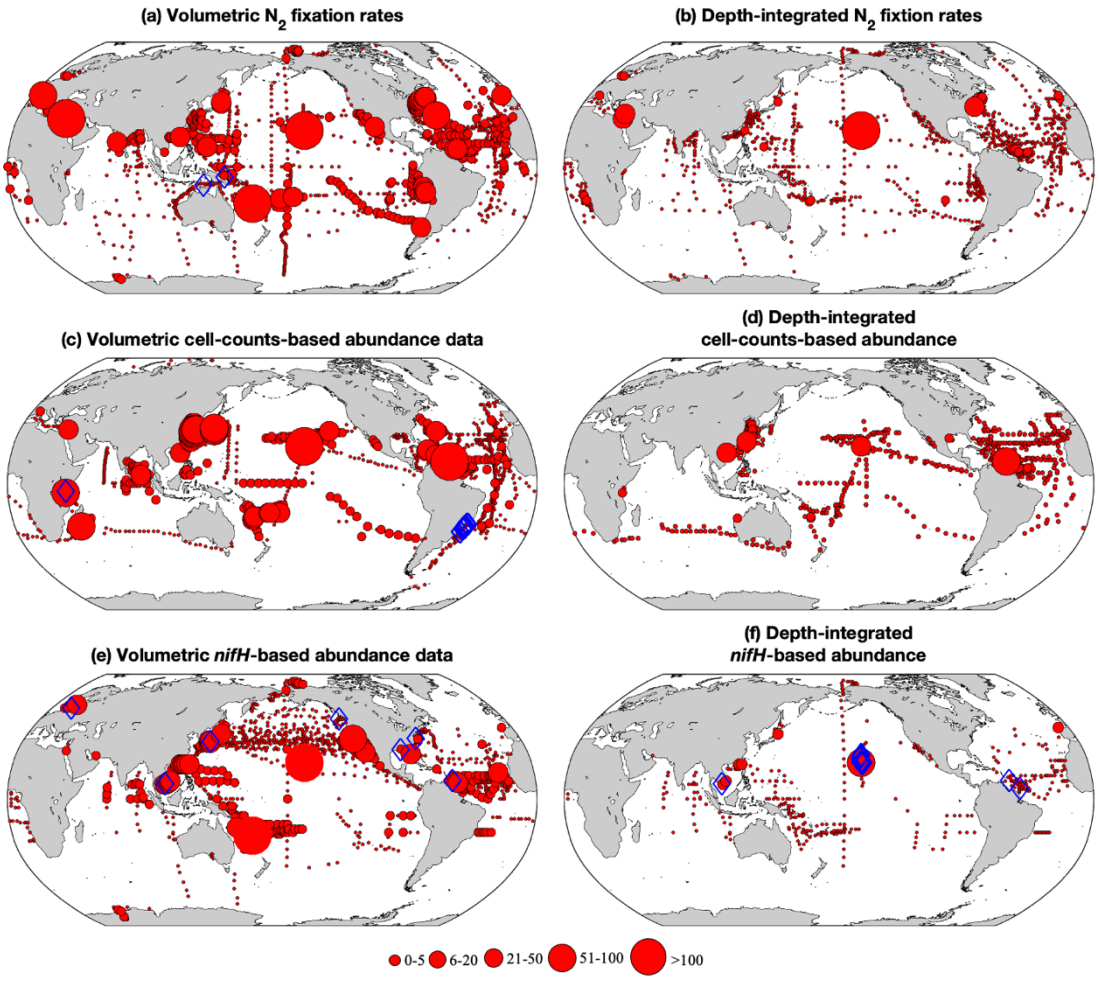
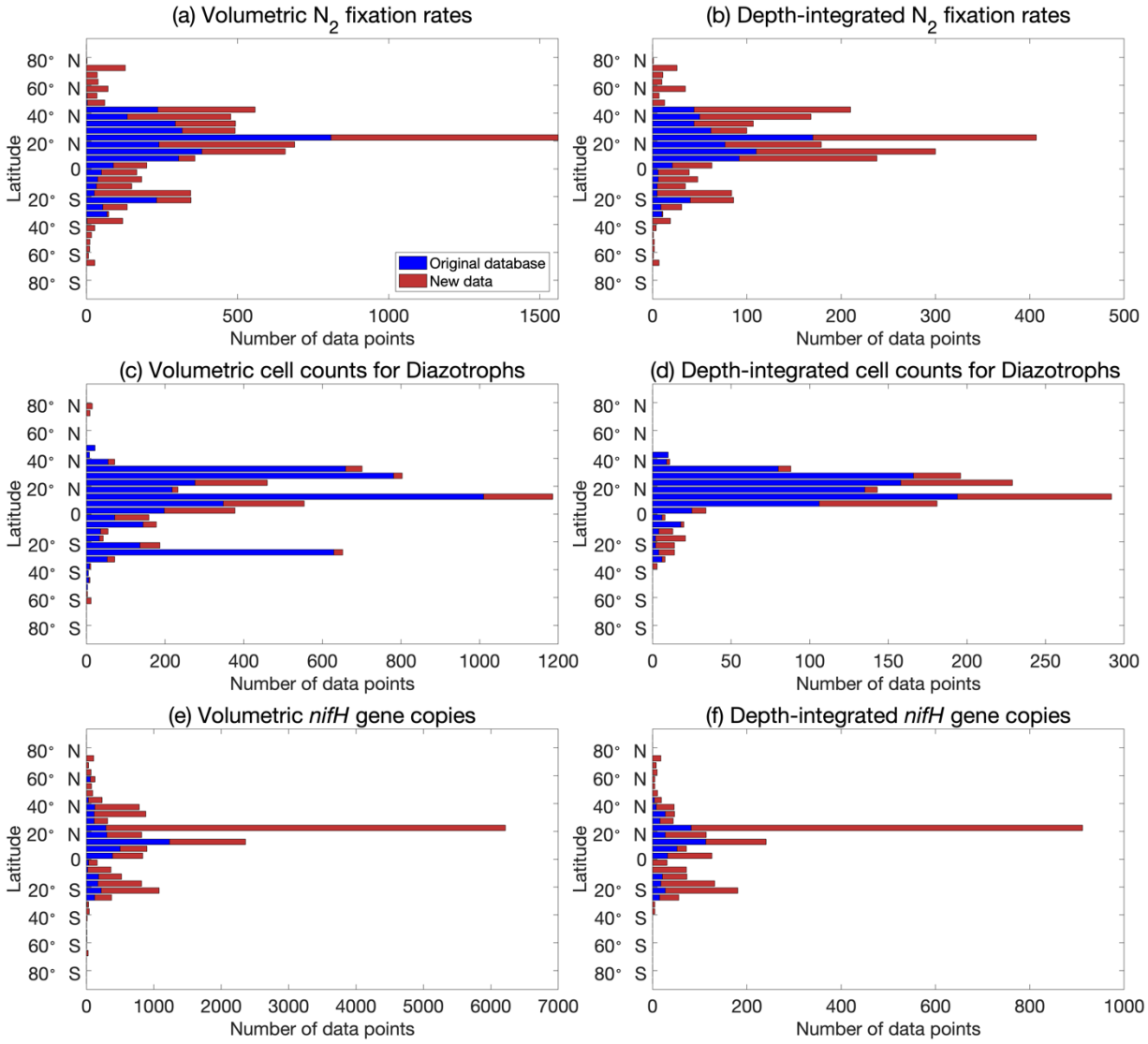


Figure 3. Spatial distribution of volumetric and depth-integrated number of data ~~points~~point binned in 1~~°×1°~~° latitude × 1° longitude grids for (**a-b**) N_2 fixation rates, (**c-d**) cell ~~counts~~abundance, and (**e-f**) *nifH* gene ~~copies~~copy abundance. The size of the circles represents the number of data points in each bin. The blue diamonds mark the location of outliers identified using Chauvenet's criterion.

Version 2 ~~of the database~~ added data at all latitudinal ranges (**Fig. 4**). In particular, version 2 extended the range of data from tropical and subtropical areas to include polar regions in the Arctic Ocean (Harding et al., 2018) ~~and Antarctic coast (Shiozaki et al., 2020).~~



and Antarctic coast (Shiozaki et al., 2020).

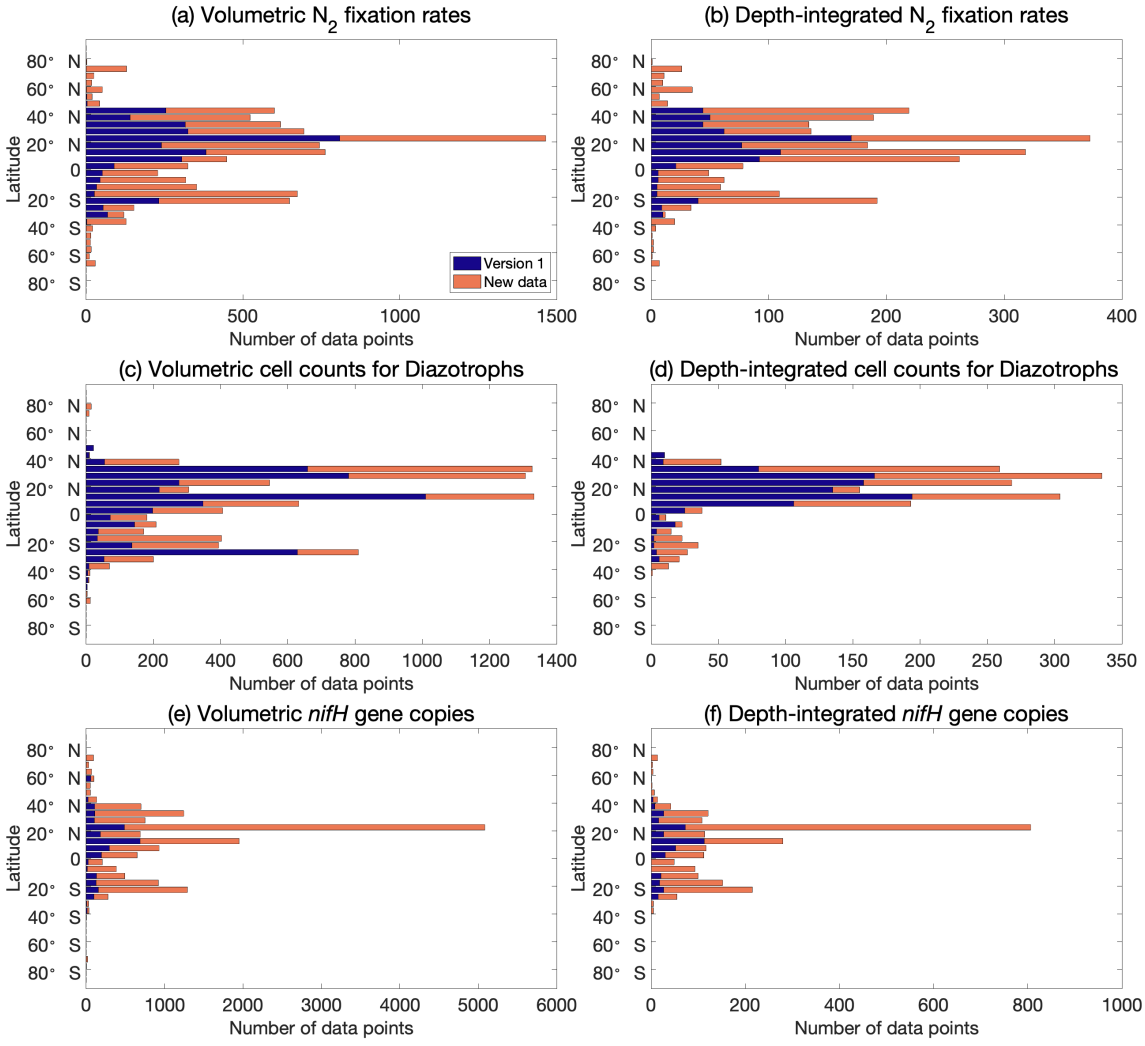
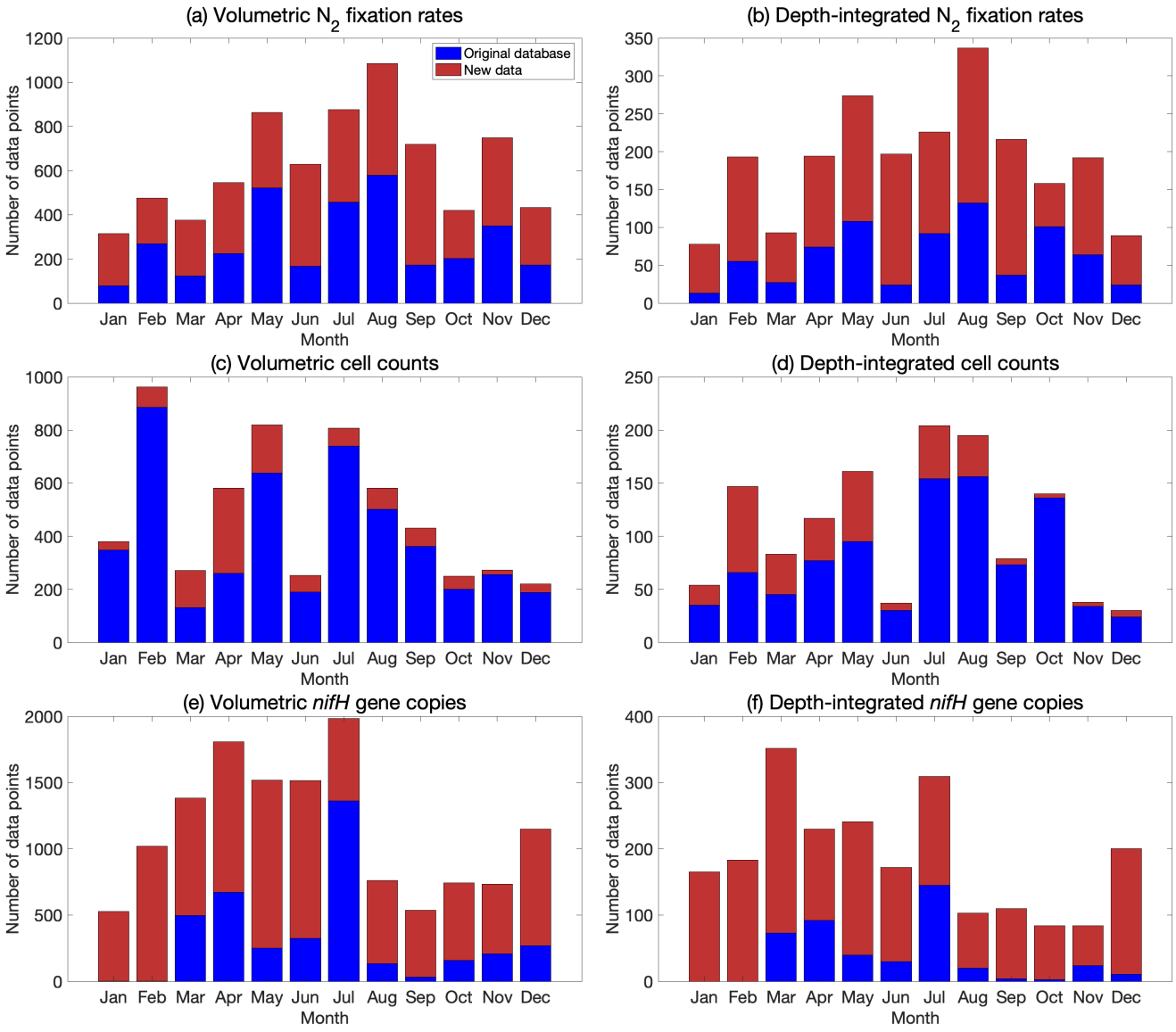


Figure 4. Latitudinal distribution of volumetric and depth-integrated ~~data including~~ (**a-b**) N_2 fixation rates, (**c-d**) cell ~~counts~~abundance, and (**e-f**) *nifH* gene ~~copies~~copy abundance, including the data in the ~~original~~version 1 of the database (blue) and the new data added ~~in~~to the version 2 of the database (~~red~~orange).

The data in version 2 ~~of~~reduce the ~~database were distributed more evenly~~difference in number of data points across months ~~than in the original database~~, especially for ~~the~~ *nifH* gene copies, in which substantially more ~~data~~samples were collected ~~from~~in January ~~to~~and February (**Fig. 5**). When considering seasons in both the Northern hemisphere and South Atlantic and Pacific the data were distributed more evenly (**Fig. 6**).


(a) Volumetric N_2 fixation rates
(b) Depth-integrated N_2 fixation rates
(c) Volumetric cell counts
(d) Depth-integrated cell counts
(e) Volumetric *nifH* gene copies
(f) Depth-integrated *nifH* gene copies
Original database
New data
Number of data points
Month
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec

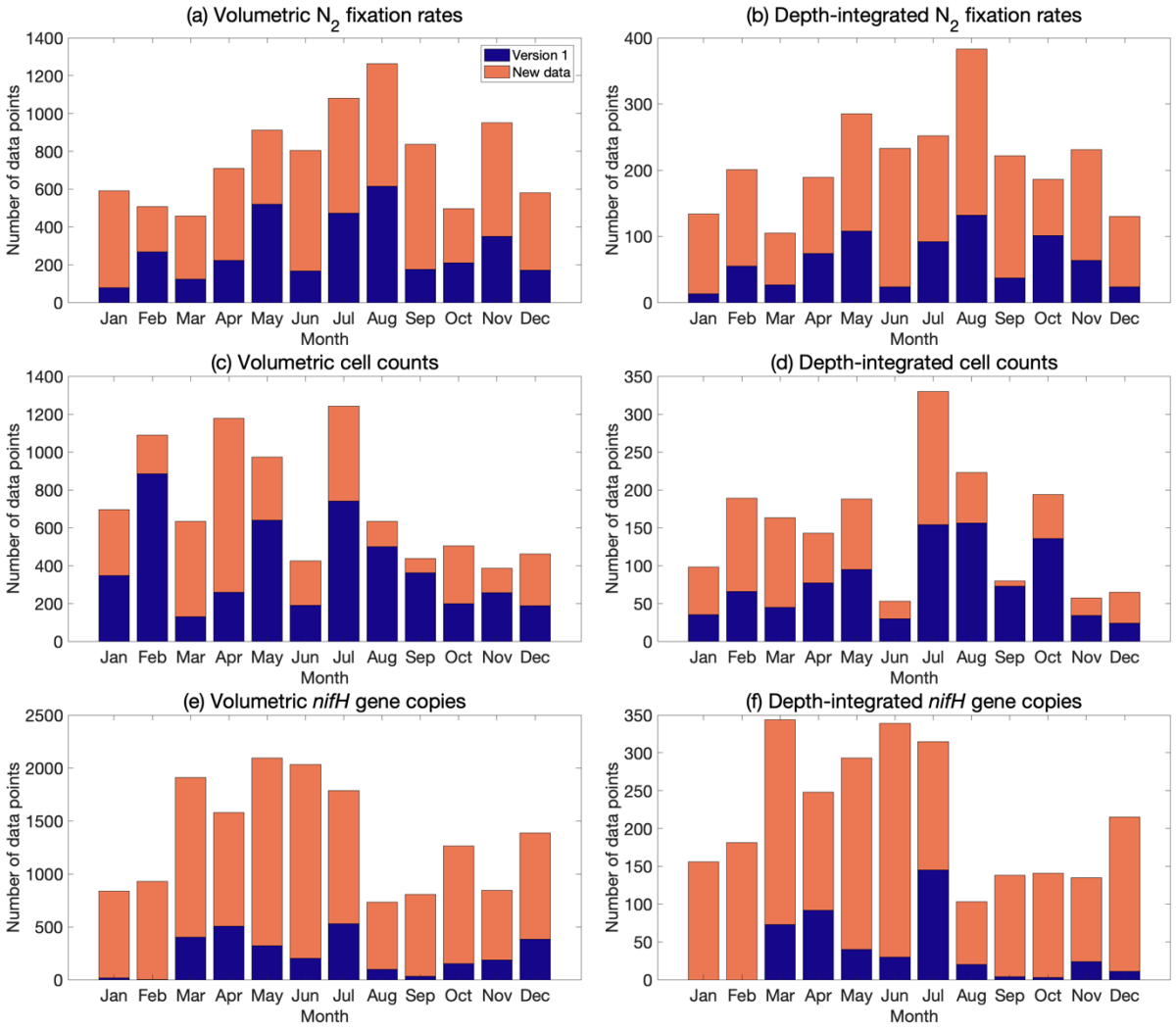


Figure 5. ~~Temporal~~Monthly distribution of volumetric and depth-integrated ~~data including~~ (**a-b**) N_2 fixation rates, (**c-d**) cell ~~counts~~abundance, and (**e-f**) *nifH* gene ~~copies~~copy abundance, including the data in the ~~original~~version 1 of the database (blue) and the new data added ~~in~~to the version 2 of the database (~~red~~orange).

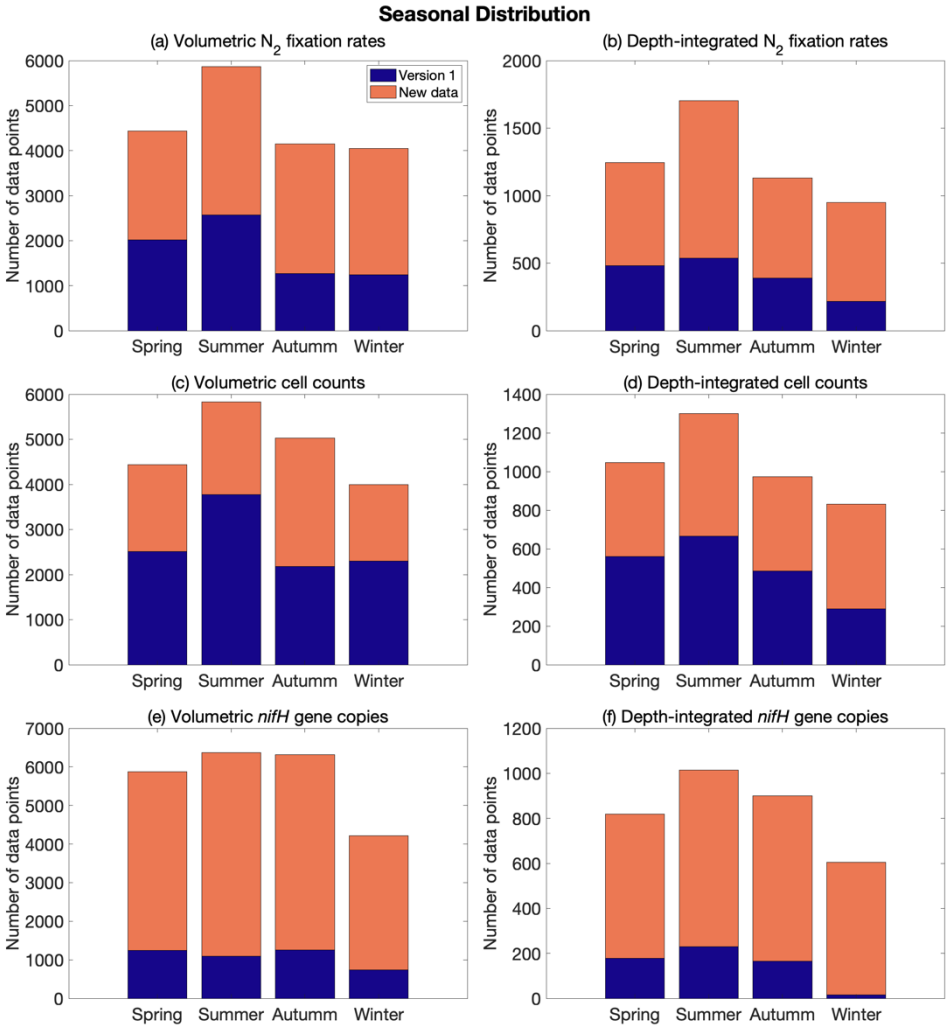
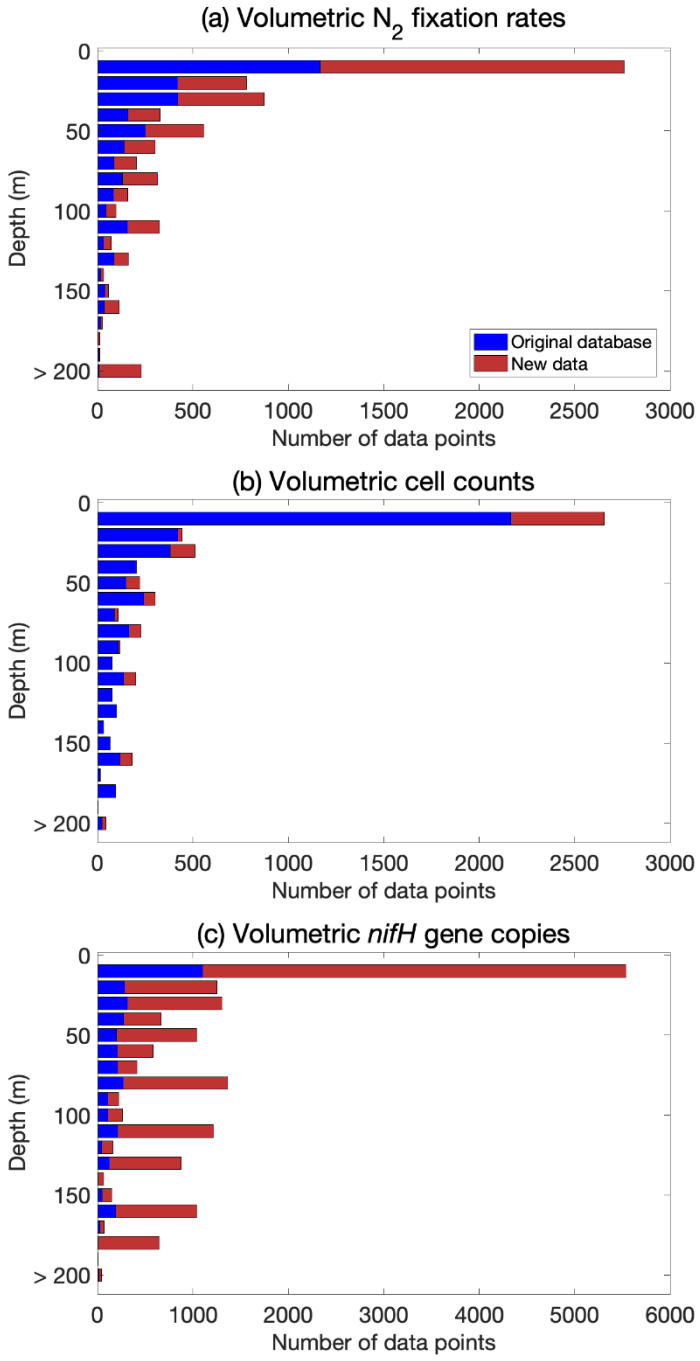


Figure 6. Seasonal distribution of volumetric and depth-integrated (**a-b**) N_2 fixation rates, (**c-d**) cell abundance, and (**e-f**) *nifH* gene copy abundance, including the data in the version 1 of the database (blue) and the new data added to the version 2 of the database (orange). Spring: March–May in the Northern Hemisphere and September–November in the Southern Hemisphere; summer: June–August in the Northern Hemisphere and December–February in the Southern Hemisphere; autumn: September–November in the Northern Hemisphere and March–May in the Southern Hemisphere; and winter: December–February in the Northern Hemisphere and June–August in the Southern Hemisphere.

Although most of the new data were measured in the near-surface waters, ~~there were~~ numerous *nifH* gene copy abundance data were also sampled in ~~other~~deeper layers in the euphotic zone (**Fig. ~~6~~7**). Additionally, active N_2 fixation and the existence of diazotrophs were found below the euphotic zone (e.g., depth > 200 m) ~~(Benavides et al., 2016b; Benavides et al., 2018b; Selden et al., 2019; Hamersley et al., 2011; Bonnet et al., 2013; Loescher et al., 2014) (**Fig. 6**~~(Benavides et al., 2016b;

Benavides et al., 2018b; Selden et al., 2019; Hamersley et al., 2011; Bonnet et al., 2013; Loescher et al., 2014; Benavides et al., 2015) (**Fig. 7**).



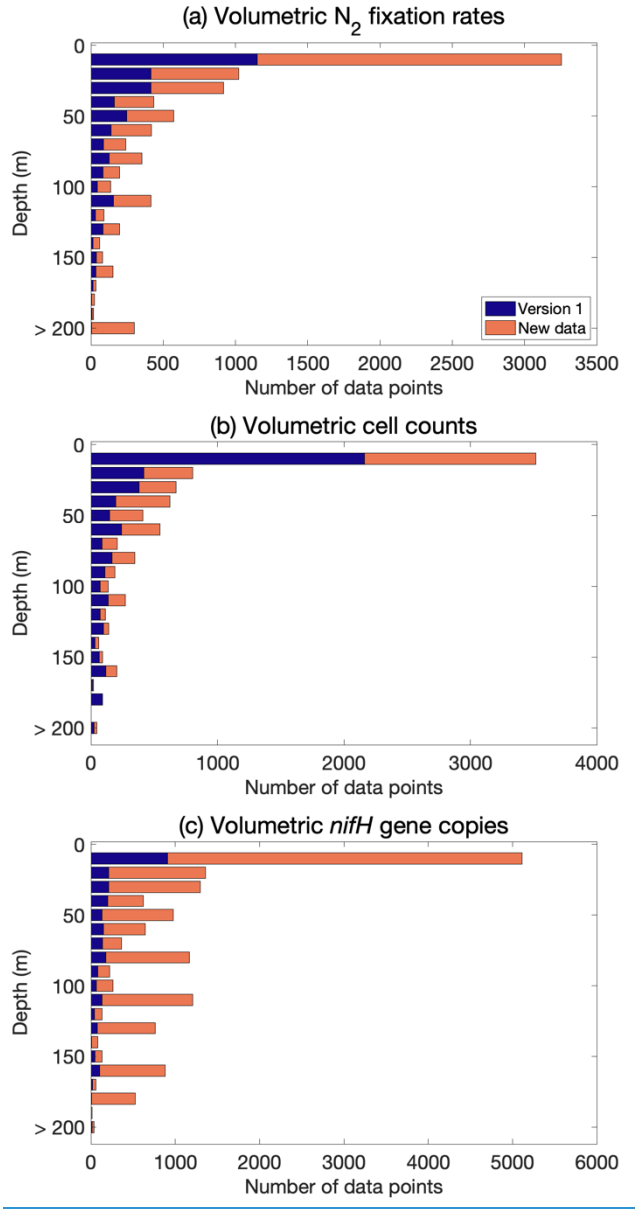
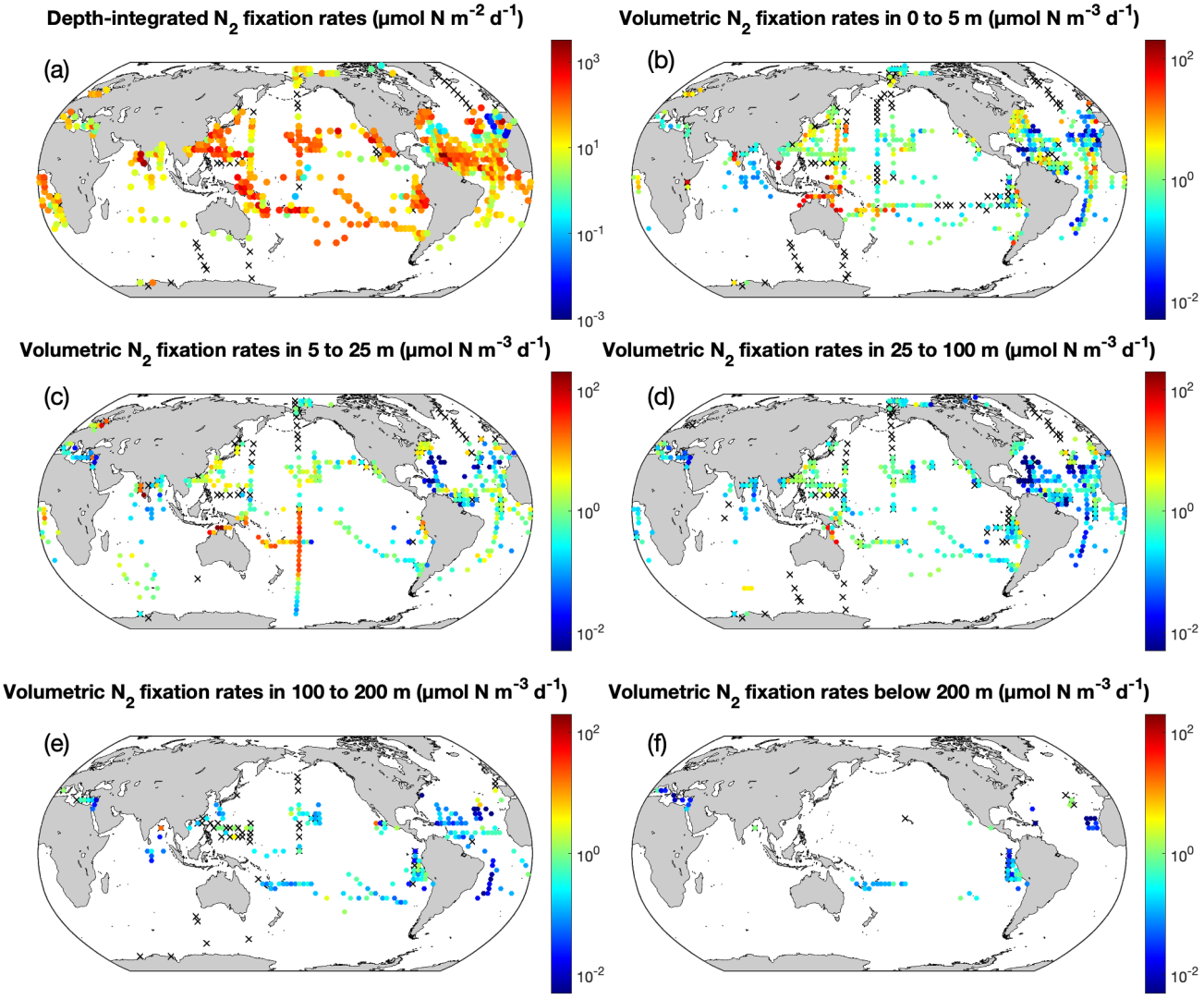


Figure ~~6~~7. Vertical distribution number of ~~data including~~ (**a**) N_2 fixation rates, (**b**) cell ~~counts~~abundance, and (**c**) *nifH* gene ~~copies~~copy abundance data, including the data in the ~~original~~version 1 of the database (blue) and the new data added ~~in~~to the version 2 of the database (~~red~~orange).

3.2 N_2 fixation rates

The volumetric N_2 fixation rates in 5 vertical layers and the depth-integrated N_2 fixation rates were binned in 3°~~×~~ latitude × 3° ~~grids~~longitude bins, and the ~~geometric~~arithmetic means in each bin are displayed (**Fig. ~~7~~8**). The depth-integrated N_2 fixation rates ranged ~~in order~~orders of magnitude, from 10^{-4} – 10^{3} µmol N m^{-2} d^{-1} (mostly ~~in order of~~from 1 ~~–100~~to 10^{2} µmol N m^{-2} d^{-}

^{1}) (**Fig. ~~7a~~8a**). Some high rates (i.e., $10^{2}-10^{3}$ μmol N m^{-2} d^{-1}) were found in the western Pacific Ocean, the regions near the ~~Hawaii~~Hawaiian Islands, and the western tropical Atlantic Ocean. Approximately ~~25~~10% of the depth-integrated N_2 fixation rates were ~~lower than~~< 1 μmol N m^{-2} d^{-1}, and were mainly ~~in~~from the North Atlantic ~~Ocean~~ and ~~in the~~ Indian ~~Ocean. Vertically~~Oceans. Within the water column, the N_2 fixation rates were ~~high~~highest in the upper 25 m (**Fig. ~~7b~~8b, c**), below which the rates rapidly decreased with depth (**Fig. ~~7d~~8d, e, f**). In the upper 25 m, ~~the~~volumetric N_2 fixation rates in the southwestern Pacific were higher than those in other areas, mostly ranging from 1 to 100 μmol N m^{-3} d^{-1}. ~~Note that some zero~~Undetectable N_2 fixation rates were reported mostly in subpolar regions, as well as in certain tropical and subtropical regions (**Fig. ~~7~~8**).


Depth-integrated N_2 fixation rates (μmol N m^{-2} d^{-1})
(a)
10^3
10^1
10^{-1}
10^{-3}
Volumetric N_2 fixation rates in 0 to 5 m (μmol N m^{-3} d^{-1})
(b)
10^2
10^0
10^{-2}
Volumetric N_2 fixation rates in 5 to 25 m (μmol N m^{-3} d^{-1})
(c)
10^2
10^0
10^{-2}
Volumetric N_2 fixation rates in 25 to 100 m (μmol N m^{-3} d^{-1})
(d)
10^2
10^0
10^{-2}
Volumetric N_2 fixation rates in 100 to 200 m (μmol N m^{-3} d^{-1})
(e)
10^2
10^0
10^{-2}
Volumetric N_2 fixation rates below 200 m (μmol N m^{-3} d^{-1})
(f)
10^2
10^0
10^{-2}

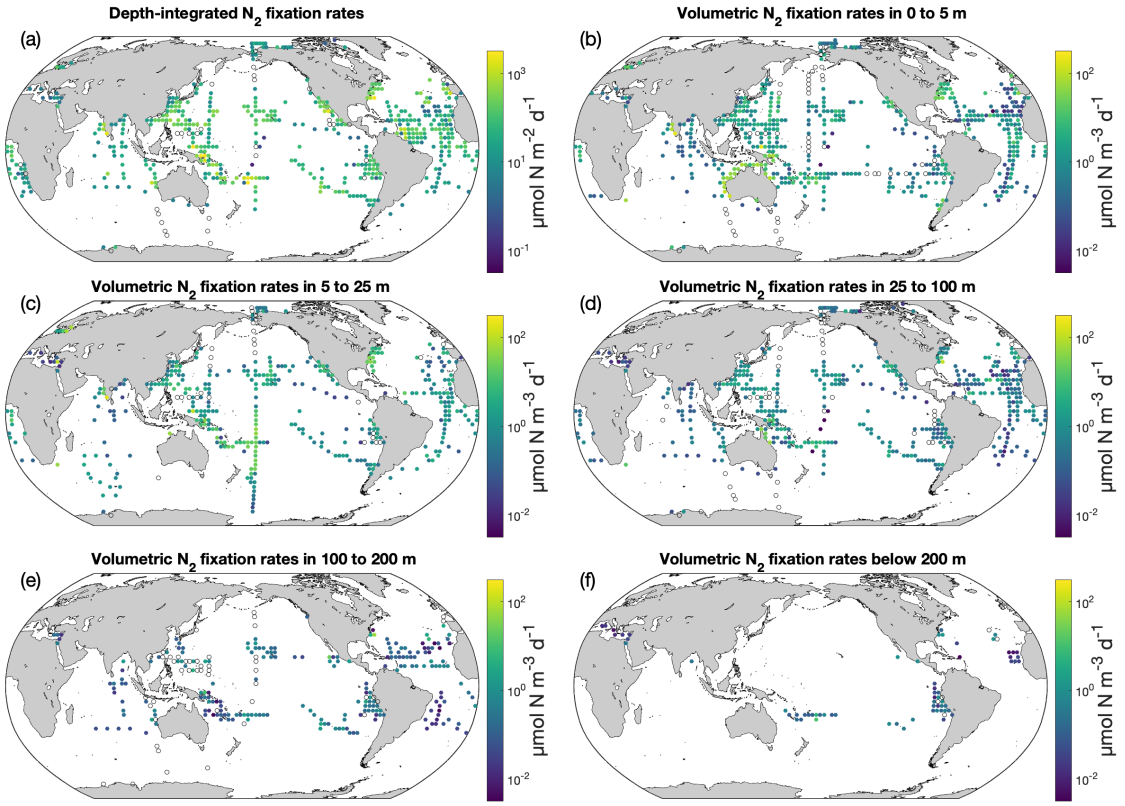


Figure ~~7~~8. N_2 fixation rates in the version 2 of the database. The panels show (**a**) depth-integrated data and volumetric data in (**b**) 0–5 m, (**c**) 5–25 m, (**d**) 25–100 m, (**e**) 100-200 m, and (**f**) below 200 m. For a clear demonstration, ~~data are binned to~~ arithmetic mean N_2 fixation rates in 3° latitude × 3° ~~grids and geometric means in each bin~~longitude bins are shown. Zero-value data are denoted as black ~~crosses.~~empty circles. Only rates measured with incubation periods of 24 hours are included.

Cell-specific N_2 fixation rates span a range from 10^{-4} to 10^{3} fmol N cell^{-1} d^{-1}, although mostly on the order of 10^{-2} to 10^{2} fmol N cell^{-1} d^{-1} (**Fig. 9**). The mean cell-specific N_2 fixation rates of *Trichodesmium*, UCYN-A2 and heterocystous cyanobacteria were one to two orders of magnitude higher than those of other diazotrophic groups (**Fig. 9 & Table S1**).

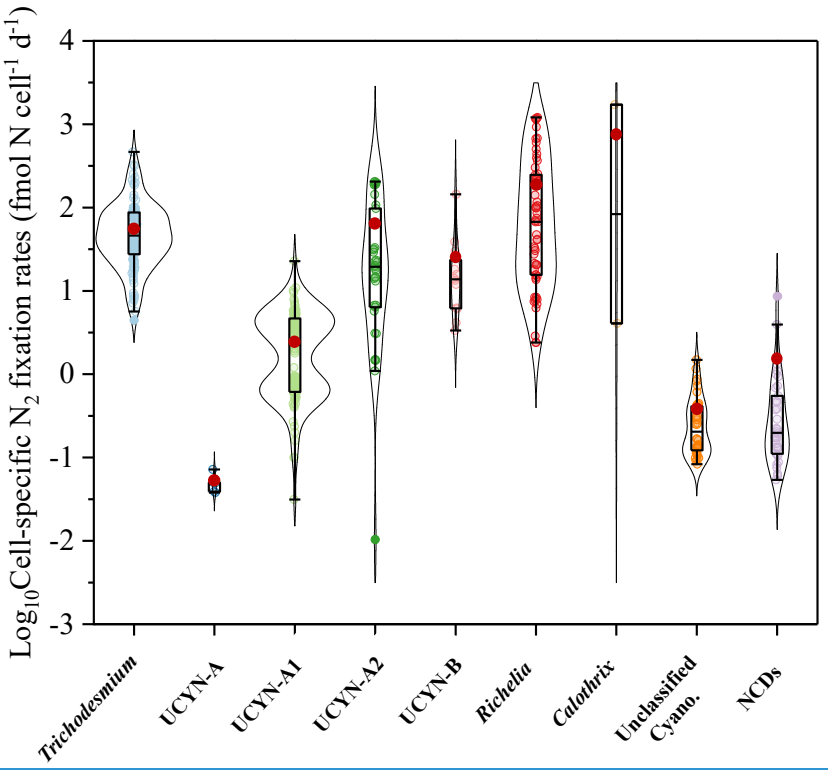


Figure 9. Violin plot of cell-specific N_2 fixation rates, including measurements for *Trichodesmium*, UCYN-A, UCYN-A1, UCYN-A2, UCYN-B, heterocystous cyanobacteria, unclassified cyanobacteria, and NCDs. The range of each box spans the 25th−75th percentile of data, the black line in each box is the median, and the red dot represents the arithmetic mean.

3.3 Diazotrophic abundance

The depth-integrated cell abundances and ~~upper 25 m~~ volumetric cell~~-count-based abundance was~~ abundances in upper 25 m are also shown ~~in geometric~~as the arithmetic means ~~of each~~in 3~~°×~~° latitude × 3~~° grid~~° longitude bins (**Fig. 8**~~**10**~~). *Trichodesmium* abundance generally decreased from the west to the east in the Atlantic Ocean (**Fig. 8**~~**a**~~**10a–b**). In the Pacific Ocean, *Trichodesmium* appeared more abundant in the west. The abundance data of ~~*Richelia* and *Calothrix*~~heterocystous diazotrophs were still scarce (**Fig. 8**~~**e**~~**10c, e**). The volumetric cell-count-based abundance data are also displayed in three additional depth intervals (**Fig. S6**).

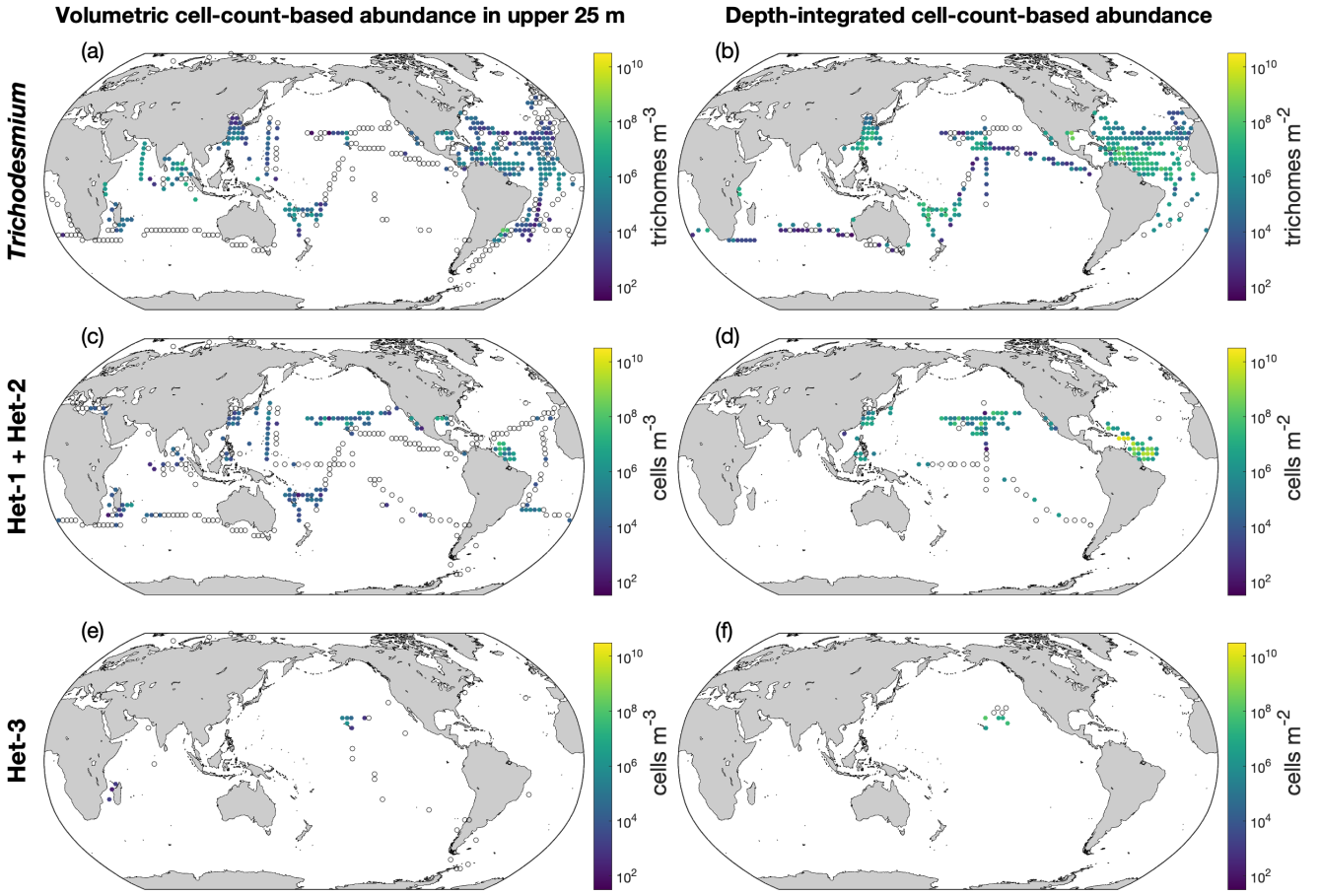
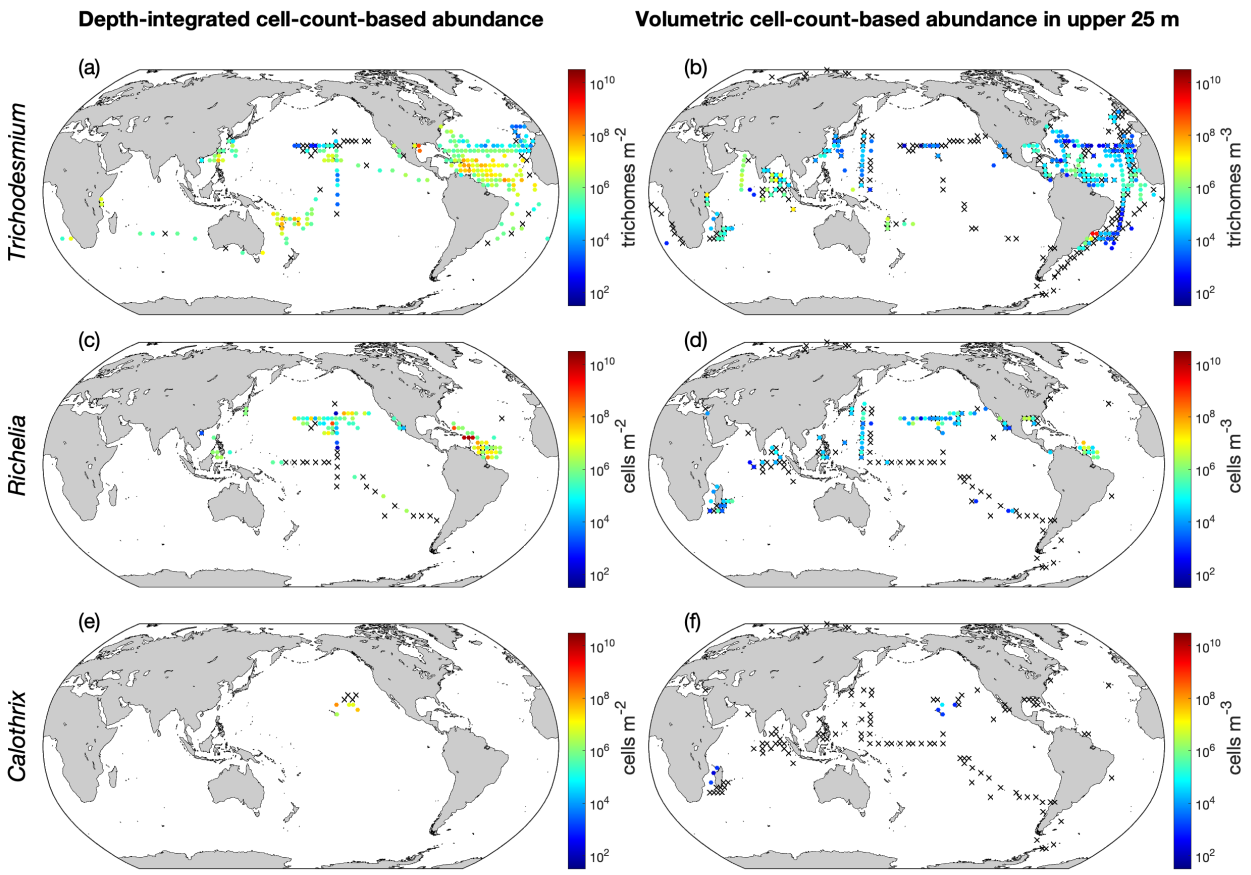


Figure 10. Depth-integrated cell abundances and volumetric cell abundances in upper 25 m in the version 2 of the database. The panels show (**a**–**b**) *Trichodesmium*, (**c**–**d**) het-1/2, and (**e**–**f**) het-3. For a clear demonstration, data are binned to 3° latitude × 3° longitude, and arithmetic means in each bin are shown. Zero-value data are denoted as open black circles.

Gene copies of *nifH* ~~*copies*~~ had better spatial coverage than the cell-count data (**Fig. 911**). Depth-integrated *Trichodesmium nifH* copies were also more abundant in the western Pacific and western Atlantic Oceans (**Fig. 9a11a**). Some high depth-integrated *nifH* abundance of UCYN-A and UCYN-B ~~was~~were also reported in the northwestern and southwestern Pacific Ocean (**Fig. 9c11c, e**). High *nifH* ~~abundance~~abundances of *Richelia* ~~was also~~were found in the ~~northwestern~~southwestern Pacific Ocean and western Atlantic Oceans (**Fig. 9i11i**). The *nifH* abundance data for UCYN-C and ~~*Calothrix*~~het-3 were sparse. The volumetric *nifH* abundance data are ~~also shown~~displayed in ~~four~~three depth intervals (**Fig. 911 & Fig. S8S7**). Almost all diazotrophs were more abundant in the upper 25 m than ~~below~~in deeper water.


Depth-integrated cell-count-based abundance
Volumetric cell-count-based abundance in upper 25 m
(a)
(b)
(c)
(d)
(e)
(f)
Trichodesmium
Richelia
Calothrix
trichomes m-2
trichomes m-3
cells m-2
cells m-3
10^10
10^8
10^6
10^4
10^2

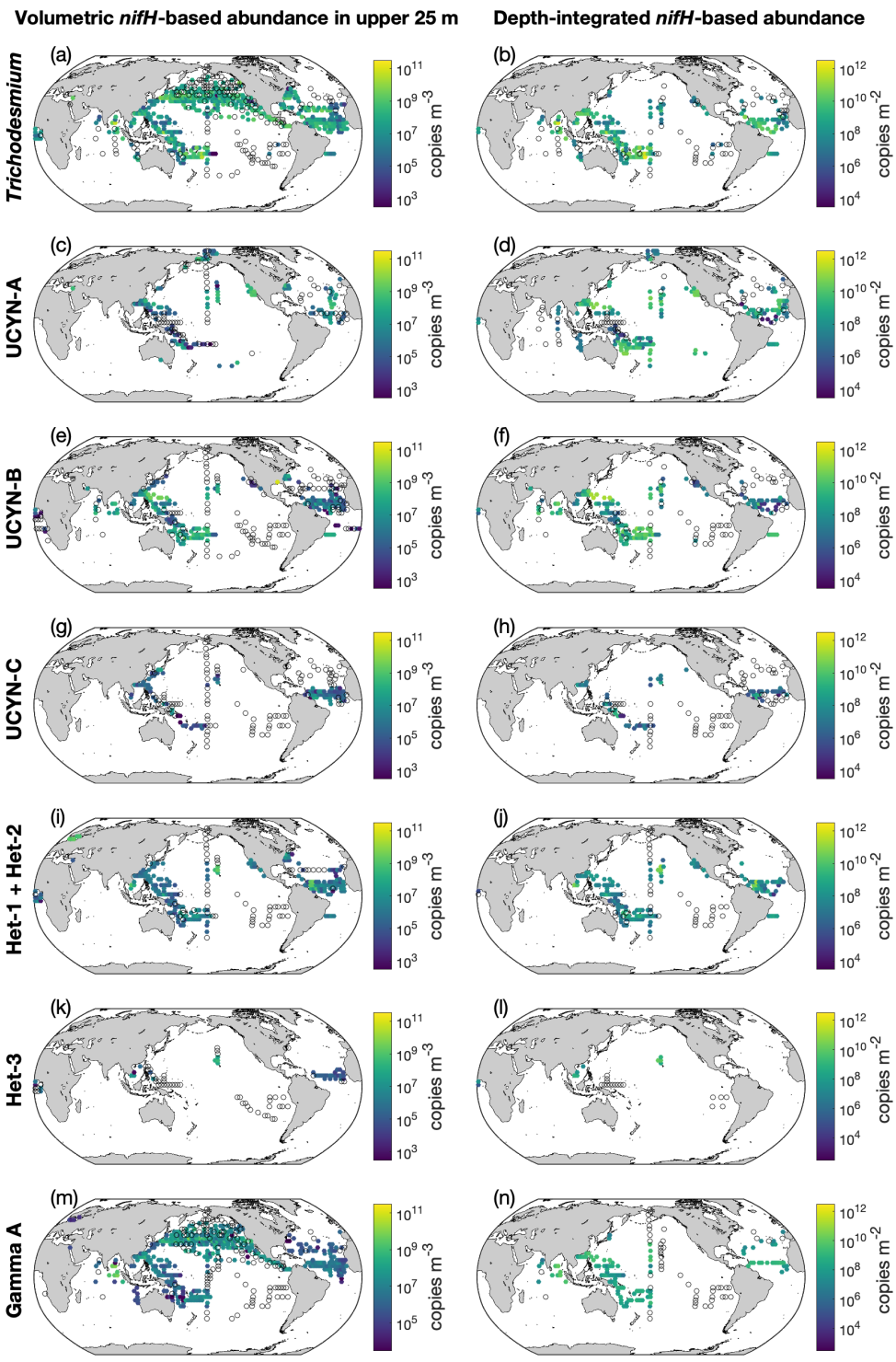
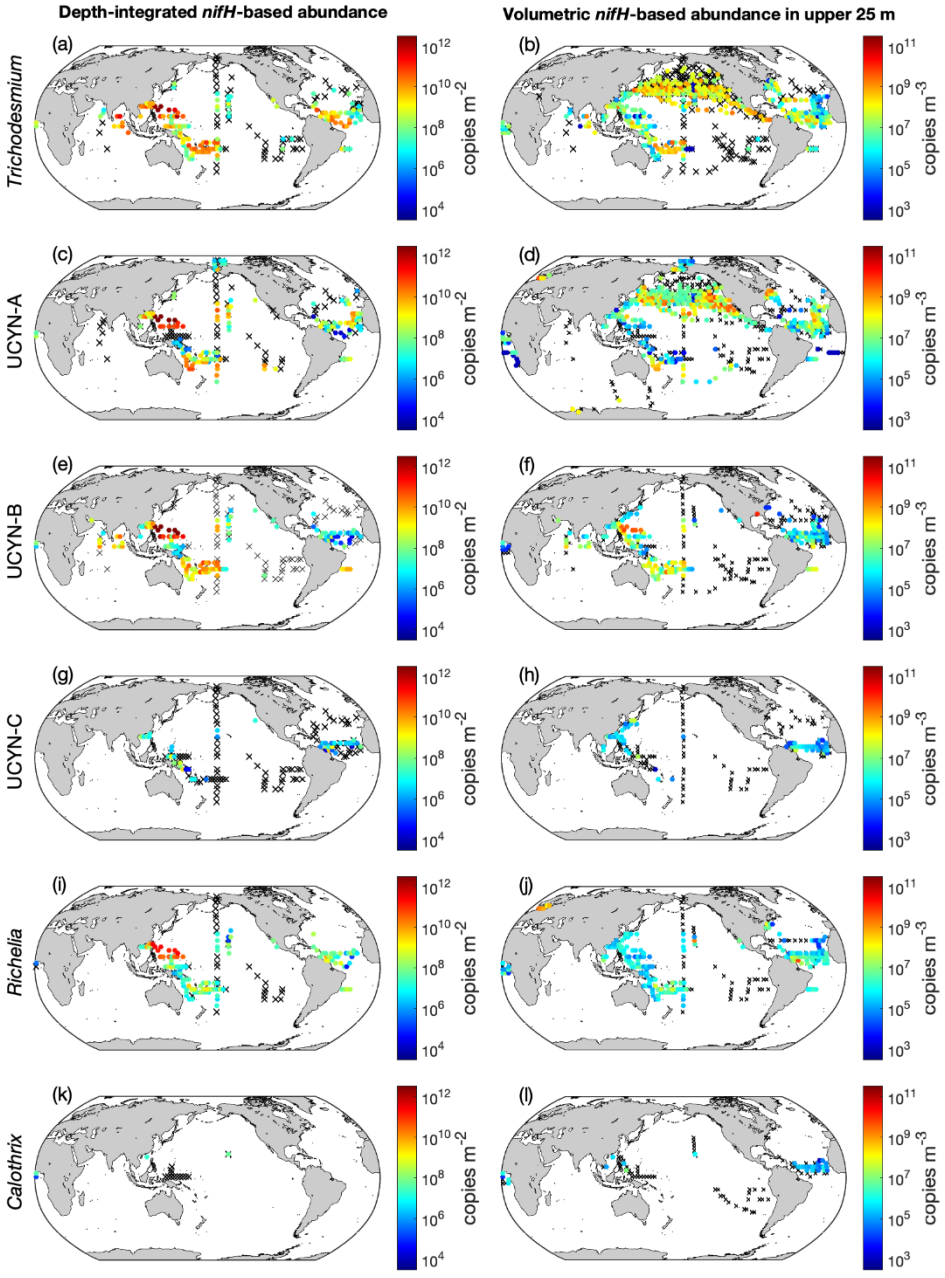

Volumetric nifH-based abundance in upper 25 m
Depth-integrated nifH-based abundance
Trichodesmium
UCYN-A
UCYN-B
UCYN-C
Het-1 + Het-2
Het-3
Gamma A
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)
(l)
(m)
(n)
copies m-3
copies m-2

Figure ~~**8. Depth-integrated cell-count-based abundance**~~**11.** Volumetric and ~~upper 25 m volumetric cell-count-based abundance~~depth-integrated *nifH* gene copy abundances in the version 2 of the database. For volumetric abundances, only data in the upper 25 m are shown. The panels show gene copy abundances of (**a–b**) *Trichodesmium*, (**c–d**) ~~*Richelia*~~UCYN-A, (**e–f**) UCYN-B, (**g–h**) UCYN-C, (**i–j**) het-1 + het-2, (**k–l**) het-3, and (~~**e–f**) *Calothrix*~~ **m–n**) Gamma A (an NCD phylotype). The depth-integrated data for Gamma A are not available. For a clear demonstration, data are binned to 3° latitude × 3°~~,~~ longitude and ~~geometric~~arithmetic means in each bin are shown. Zero-value data are denoted as open black ~~crosses. Note that the abundance of *Trichodesmium* is reported as the number of trichomes per square or cubic meter~~circles.



645

Figure 9. Depth-integrated *nifH* abundance and upper 25 m volumetric *nifH* abundance in version 2 of the database. The panels show (**a**–**b**) *Trichodesmium*, (**c**–**d**) UCYN-A, (**e**–**f**) UCYN-B, (**g**–**h**) UCYN-C, (**i**–**j**) *Richelia*, and (**k**–**l**) *Calothrix*. For a clear demonstration, data are binned to 3° × 3° and geometric means in each bin are shown. Zero-value data are denoted as black crosses.

3.4 First-order estimate of global ~~ocean~~oceanic N_2 fixation rate

~~Similar to that applied to the original database, we used version 2 of the database to conduct a first-order estimate of global ocean N_2 fixation rates (**Table 5**). The adequate data in version 2 supported the estimates of mean and total N_2 fixation rates in the Indian and Arctic Oceans, while the estimates for these two ocean basins were not available when using the original database~~ Compared to version 1, the spatial coverage of data in version 2, in terms of the fraction of 3°×3° bins, was greatly increased in all ocean basins (**Table 8**). The spatial data coverage was very low in the Southern and Arctic Oceans (1% and 2% of total bins, respectively) (**Table 8**) and we therefore did not estimate total N_2 fixation rates for these two basins. Please note that the inaccurate areas of the North and South Pacific Oceans used in estimating global oceanic N_2 fixation rate by ~~(~~Luo et al~~.~~, (2012)~~.~~ was corrected in this study (**Table 8**).

~~Version 2 estimated a much higher arithmetic mean of N_2 fixation rates in the South Pacific Ocean than that using the original database, while the difference did not exist in the geometric means in the same basin. Another large difference between the two versions of the database was the much lower geometric mean of N_2 fixation rates in the North Pacific. The global oceanic N_2 fixation rate was calculated by summing its geometric or arithmetic means in every ocean basin, with the associated errors being propagated (Glover et al., 2011; Luo et al., 2012). The estimates of global oceanic N_2 fixation based on geometric means were close using the original and version 2 databases (62 and 60 Tg N yr^{-1}, respectively), while those based on arithmetic means differed greatly (137 versus 260 Tg N yr^{-1}, respectively). This higher arithmetic mean-based estimate of global oceanic N_2 fixation was mainly due to the supplementation of the Indian Ocean, for which the estimate was unavailable when using the original database because of data limitations, and due to the higher rates in the South Pacific Ocean. It must be noted that high uncertainties were associated with the arithmetic means of N_2 fixation in the South Pacific and Indian Oceans when using version 2 (**Table 5**), indicating that more measurements were needed in these basins to better constrain the elevated estimates of global oceanic N_2 fixation.~~

Table ~~5~~8. First-order estimates of N_2 fixation rates ~~in~~ based on their arithmetic means in different ocean basins. Data are first binned to 3°~~×3°~~° latitude ×3° longitude grids before being used to calculate ~~geometric or~~ arithmetic means in each basin. The arithmetic means are multiplied by the basin areas to calculate the N_2 fixation rates of each basin. NQ: not quantified due to limited data points. ND: no data. The ~~numbers~~percentages in parentheses ~~after geometric means~~ are ~~confidence intervals estimated from~~fraction of the 3° ×3° bins in each basin that have measurements. The reported uncertainties are one standard error of ~~log-10 transformed data. Arithmetic means were reported with one standard error~~the mean.

Region	Number of binned data	Latitudinal range	Ocean area (× 10^{12} m^2)	~~Mean~~ Arithmetic mean N_2 fixation rate (μmol N m^{-2} d^{-1})	Areal sum of N_2 fixation rate (Tg N yr^{-1})

	Version 1	Version Original database 2	Version 2 1	Original database Version 2	Version 2 1	Version 2	Version 1	Version 2	Version 1	Version 2
	Original database	Version 2	Geometric	Arithmetic	Geometric		Arithmetic	Geometric	Arithmetic	Geometric	Arithmetic
North Atlantic	12547 (9%)	18011 6 (21%)	9.2 (7.0–12)° –55°N	170 ± 400°–55° N	18.9 (18.6 –19.1)37	180 ± 3337	1.7 (1.3–2. 2)55±9	32 ± 7.4213±4 6	3.98 (3.93– 4.03)1 0±2	38 ± 7.040±9
South Atlantic	1514 (4%)	5352 (15%)	7.9 (6.3–9.8) 40°S –0°	13 ± 4.445°S –0°	17 (6.9– 43)26	30 ± 5	1.1 (0.9–1. 13±4)	1.8 ± 0.630±5	2.7 (1.1–8 ±0.6.7)	4.6 ± 0.815±1
North Pacific	4534 (4%)	131143 (17%)	78 (67–90)0 °–55°N	120 ± 220°–55° N	50 (22–1 15)75	130 ± 2075	35 (30–41)111±17	56 ± 9.8144±2 8	20 (8.8–4 7)42±7	52 ± 8.255±11
South Pacific	2620 (2%)	95100 (12%)	64 (54–76)4 0°S–0°	130 ± 4645°S–0°	58 (30–1 15)63	240 ± 6669	24 (20–28)61±7	46 ± 17250±6 6	20 (10–4 0)±2	83 ± 88±23
IndiaIndia n Ocean	4ND	3447 (9%)	120 (29–490)	590 ± 32045°S– 25°N	17 (15–2 0)	270 ± 14056	NQND	NQ123± 50	4.9 (4.3–5 .6)ND	76 ± 3335±14
Mediterran ean Sea	101 (3%)	129 (23%)	18 (12–28)3 0°N–45° N	30°N–45 ± 21°N	9.8 (7.4– 13)2.5	120 ± 942.5	0.2 (0.1–0. 4)NQ	0.6 ± 0.35±1	0.2 (0.15– 0.26) NQ	2.4 ± 1.90.06± 0.02
Arctic Ocean	ND	17 (2%)	ND	ND	14 (8.5– 23)	19 ± 511	ND	ND23±5	0.8 (0.48– 1.3)N D	1.1 ± 0.28NQ
Southern Ocean	ND	10 (1%)	ND	ND	21 (7.7– 58)	8.6 ± 8.160	ND	ND9±8	6.5 (2.4–1 8)ND	2.7 ± 2.5NQ

Global Ocean	-	-	-	-	-	-	~~62 (52–73)~~	~~137 ± 9.2~~	~~60 (47–107)~~ 74 ±7	~~260 ± 20~~ 223±30

We first compared the N_2 fixation rates estimated based on arithmetic means between using version 1 and version 2 (**Table 8**). Using available data in version 2, the global N_2 fixation rate was determined to be 223±30 Tg N yr^{-1}, which was three times that obtained from version 1 (**Table 8**). The substantial increase was mostly driven by notable changes in the South Pacific, North Atlantic, and Indian Oceans. In the South Pacific Ocean, numerous high N_2 fixation rates were observed in the western subtropical region over the past decade (**Fig. 12**), resulting in a substantial increase of 68±23 Tg N yr^{-1} in the estimated N_2 fixation rate for this basin (**Table 8**). It is worth noting that these newly recorded measurements in the western subtropics of the South Pacific Ocean might even be underestimated since most of them were obtained using the original $^{15}N_2$ bubble method. In the North Atlantic Ocean, the estimated N_2 fixation rate also experienced an increase of 30±9 Tg N yr^{-1} for (**Table 8**), without any discernible pattern regarding the locations of the new high N_2 fixation measurements (**Fig. 13**). Furthermore, in the Indian Ocean, the improved data coverage in version 2 (**Fig. 8a**) supported the estimation of an N_2 fixation rate of 35±14 Tg N yr^{-1} for this basin (**Table 8**), which was not possible to calculate using version 1 due to insufficient data availability.

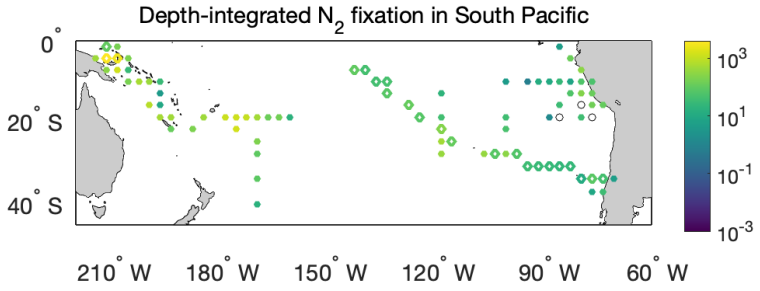


Figure 12. Depth-integrated N_2 fixation rates in the South Pacific Ocean (μmol N m^{-2} d^{-1}). The shown data are arithmetic mean rates in 3° latitude ×3° longitude bins. Empty diamonds and filled circles denote the existing data in the version 1 of the database and the new data added to version 2, respectively.

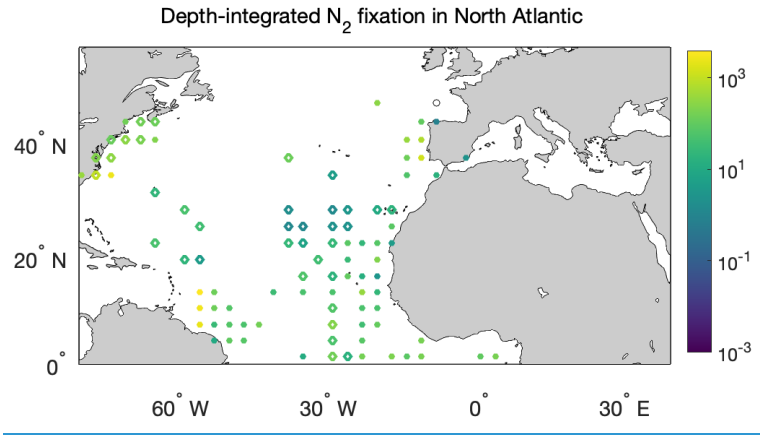


Figure 13. Depth-integrated N_2 fixation rates in the North Atlantic Ocean (μmol N m^{-2} d^{-1}). The shown data are arithmetic mean rates in 3° latitude ×3° longitude bins. Empty diamonds and filled circles denote the existing data in the version 1 of the database and the new data added to version 2, respectively.

However, when estimating the global marine N_2 fixation rate using geometric means, both version 1 and version 2 yielded similar rates of approximately 50 Tg N yr^{-1} (**Table 9**). The N_2 fixation rates in each basin tended to follow a log-normal distribution (**Fig. 14**), with the geometric mean aligning near the peak of the distribution. In the South Pacific Ocean, as discussed earlier, version 2 included a substantial number of newly observed high N_2 fixation rates, but it also incorporated a significant number of rates that were much lower than those in version 1 (**Fig. 14c**). This could be partially attributed to enhanced detection limits in measurements. Consequently, while version 2 yielded a much higher arithmetic mean N_2 fixation rate compared to version 1 for the South Pacific Ocean (**Table 8**), their geometric means remained quite similar (**Table 9**). In the North Pacific Ocean, for the same reasons, the arithmetic mean N_2 fixation rates obtained from both versions were very close, while the geometric mean from version 1 could be even higher than that from version 2 (**Tables 8 & 9; Fig. 14a**). These analyses reveal that, despite the similarity in geometric means of N_2 fixation rates obtained from both versions of the database, the higher arithmetic means in version 2 were not coincidental. Instead, they were a direct outcome of the improved measurement methods and the expanded spatial and temporal coverage of marine N_2 fixation over the past decade. Consequently, previous assessments of the global marine N_2 fixation rate were likely underestimated due to the absence of these new measurements.

Table 9. Same as Table 8 but based on the geometric means of N_2 fixation rates. The numbers in parentheses are estimated ranges based on one standard error of log-transformed N_2 fixation rates (*see* **Section 2.4**).

Region	Proportion of zero-value data		Geometric mean N_2 fixation rate (μmol N m^{-2} d^{-1})		Areal sum of N_2 fixation rate (Tg N yr^{-1})	
	Version 1	Version 2	Version 1	Version 2	Version 1	Version 2
North Atlantic	0%	5%	22 (18–26)	46 (39–54)	4.1 (3.3–5.0)	8.7 (7.4–10.1)

South Atlantic	0%	25%	8 (6–10)	15 (13–17)	1.1 (0.9–1.3)	2.3 (1.9–2.7)
North Pacific	3%	6%	73 (63–83)	45 (39–52)	27.8 (24.2–32.0)	17.3 (15.1–19.8)
South Pacific	0%	9%	52 (45–59)	51 (43–61)	16.6 (14.4–19.1)	18.0 (15.1–21.4)
Indian Ocean	ND	0%	ND	25 (20–31)	ND	7.1 (5.7–8.9)
Mediterranean Sea	0%	3%	NQ	3 (2–4)	NQ	0.04 (0.03–0.05)
Arctic Ocean	ND	2%	ND	14 (11–18)	ND	NQ
Southern Ocean	ND	70%	ND	4 (1–16)	ND	NQ
Global Ocean					50 (43–57)	53 (45–63)

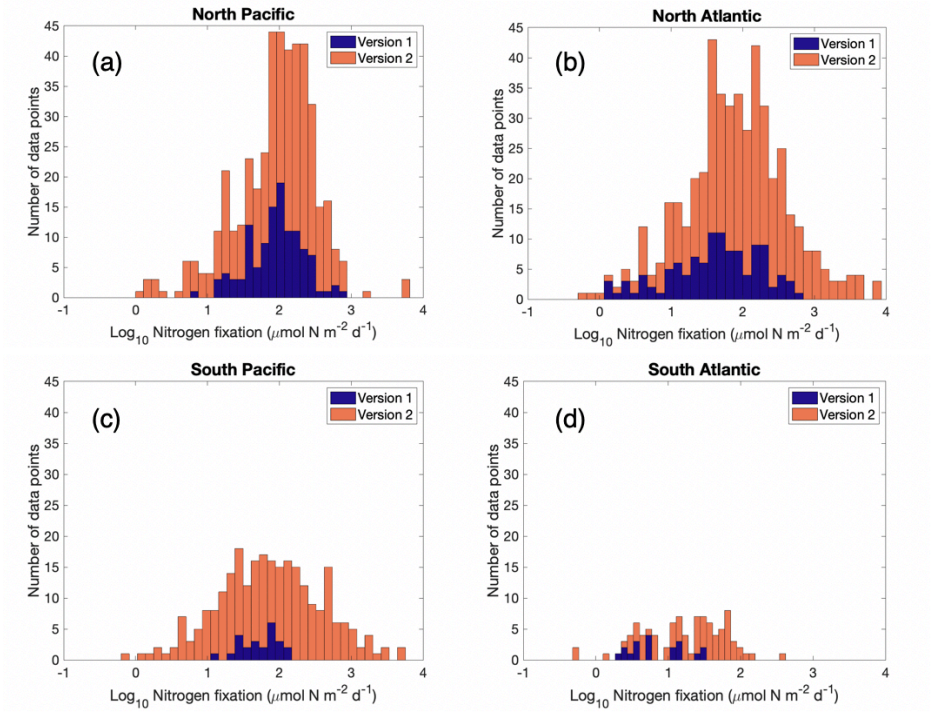


Figure 14. Comparison of the distribution of log-transformed N_2 fixation rates between the two versions of the database. Note that the zero-value data are not included because of the log-transformation. The comparison is performed for data in (**a**) North Pacific, (**b**) North Atlantic, (**c**) South Pacific, and (**d**) South Atlantic Oceans.

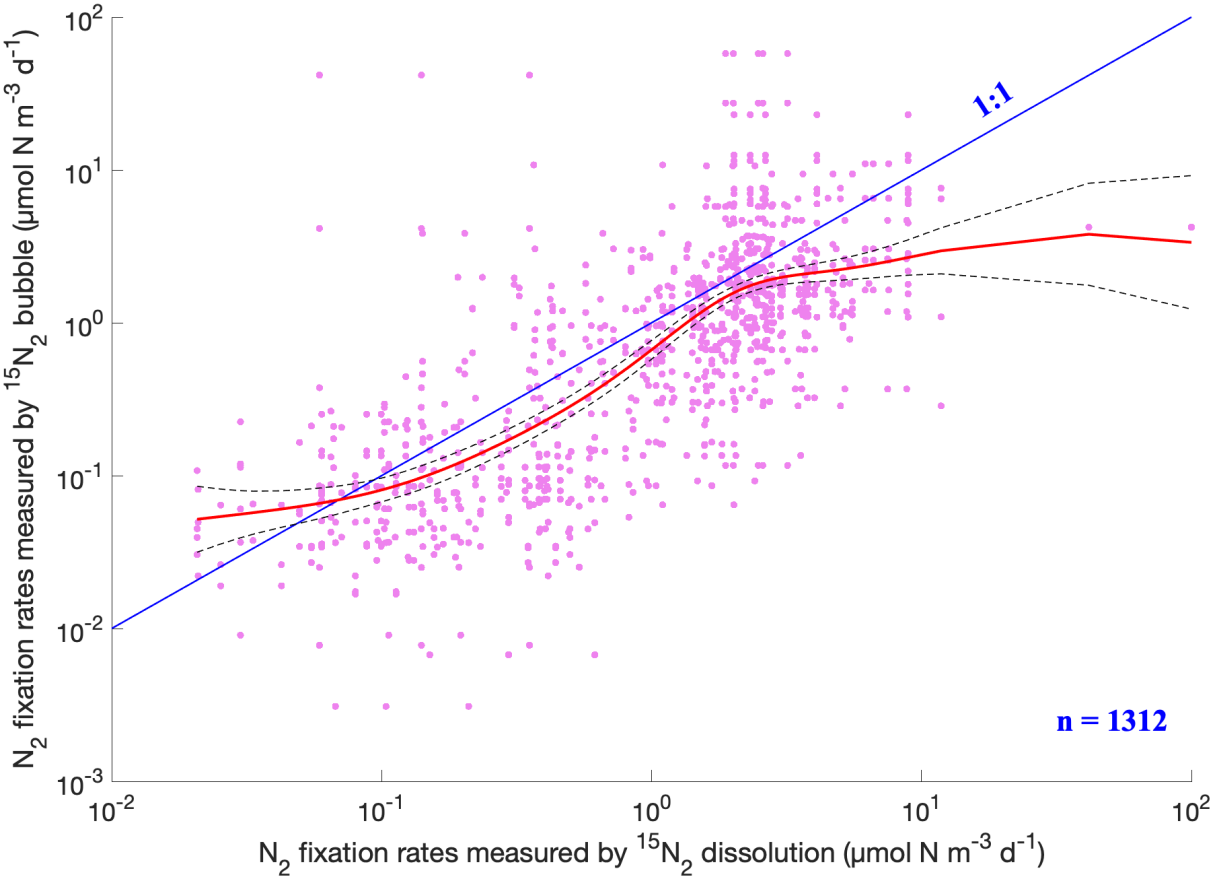
We must emphasize that this calculation simply used average N_2 fixation rates in different ocean basins, and therefore can only be considered as a first-order estimate. Furthermore, limited measurements have shown a large range of N_2 fixation rates in the Southern Ocean (**Fig. 8**). Considering its vast area, future measurements expanding coverage of N_2 fixation rates in the Southern Ocean (but see White et al. 2022) may help to better constrain the contribution of N_2 fixation to the N budget of the global ocean.

4. Discussion

4.1 Comparison of N_2 fixation measured using $^{15}N_2$ ~~bubbling~~bubble and dissolution methods

~~Compared to the $^{15}N_2$ dissolution method, the magnitude of underestimation in~~ To date, the discrepancy in N_2 fixation rates ~~acquired by the $^{15}N_2$ bubble method remains inconclusive. Here, we used data in our database to compare these two methods. First, as~~estimated using different $^{15}N_2$ tracer methods remains unclear. As shown above, the volumetric N_2 fixation rates obtained by ~~these two methods varied in a similar range of extent of magnitude (**Fig. 1**). The average N_2 fixation rate measured using~~ the original $^{15}N_2$ bubble method and the $^{15}N_2$ dissolution method ~~was~~ spanned a similar range (**Fig. 1**), while the average rates using the former method were significantly ~~higher (17%)~~lower than that measured using the $^{15}N_2$ ~~bubble~~latter method (one-tailed Wilcoxon test, p ~~<~~ <0.~~01). We also~~001, n = 2460 and 1128). With substantial data accumulated over the past decade, we further compared N_2 fixation rates ~~at the same location (1° × 1° grids) and months but measured by either the $^{15}N_2$ bubble or dissolution method~~measured using the two methods at close locations and sampling time, although the samples ~~measured by the two methods~~ were not identical. We first binned data collected from the same months, horizontal locations (3° latitude × 3° longitude) and depth intervals (0–5 m, 5–25 m, 25–100 m, and 100–200 m), and calculated the average rates for each method in each bin. The results showed that the original $^{15}N_2$ ~~dissolution~~bubble method produced ~~higher~~lower rates than the $^{15}N_2$ ~~bubble~~dissolution method in ~~65~~69% of the cases (**Fig. ~~10~~). The 13**). Furthermore, our analysis ~~using~~employing the generalized additive model (GAM) ~~further demonstrated~~revealed that the ~~underestimation by the~~ relationship between the rates measured using the original $^{15}N_2$ bubble method ~~tended to be exaggerated under high N_2 fixation (> 3 μmol N m^{-3} d^{-1})~~and those obtained through the $^{15}N_2$ dissolution method closely adhered to the 1:1 line, albeit with slightly lower values in the former (**Fig. ~~10~~), which15**). Please note that this slightly lower values can ~~be explained by the gas equilibrium time (Mohr et al., 2010; Wannicke et al., 2018; Jayakumar et al., 2017): Under low N_2 fixation, the~~still result in significant underestimation in measured N_2 fixation rates, because the GAM model was applied in a logarithmic space. It is crucial to reiterate that the rates being compared were derived from different samples, emphasizing the necessity for more future investigations that directly compare the two methods using the same samples with controlled parameters such as temperature, volume of injected $^{15}N_2$ and incubation volume. Despite this limitation, our analysis suggests that the extensive body of historical marine N_2 fixation rate data obtained through the original $^{15}N_2$ bubble method still holds a value, particularly in the examination of spatial and temporal variations in N_2 fixation.

We also used the same procedure to compare the N_2 fixation rates measured using the acetylene reduction assays and the $^{15}N_2$ tracer methods. However, there were insufficient pairs of data available for reliable comparisons (n = 16 for acetylene reduction versus the $^{15}N_2$ dissolution method; n = 6 for acetylene reduction versus original $^{15}N_2$ bubble method ~~can provide sufficient dissolved $^{15}N_2$ regardless of whether the gas reaches equilibrium; under high N_2 fixation, the $^{15}N_2$ bubble method cannot fulfil the requirement of dissolved $^{15}N_2$, resulting in relatively large underestimation.~~).



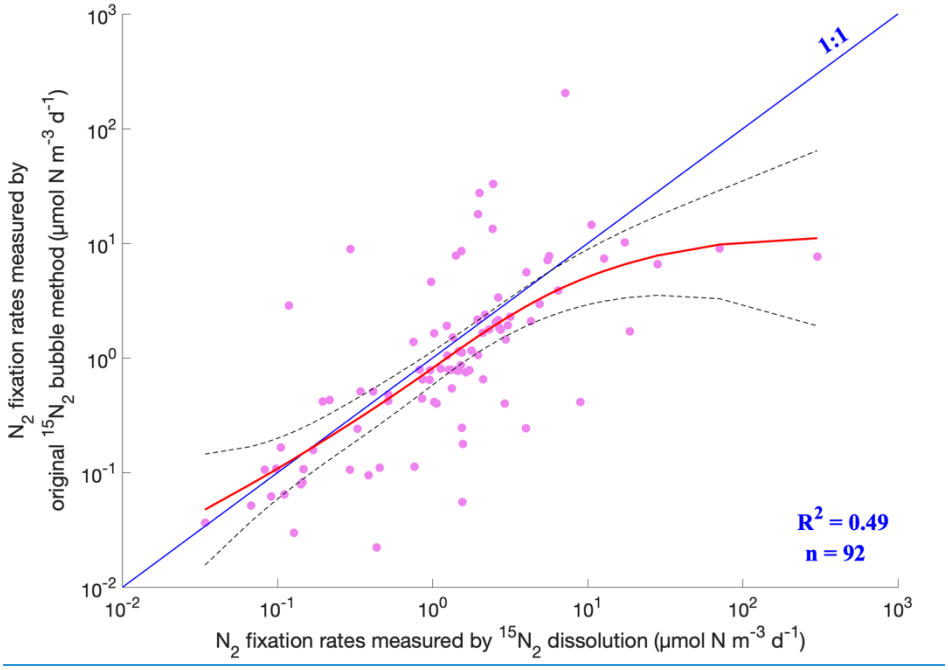
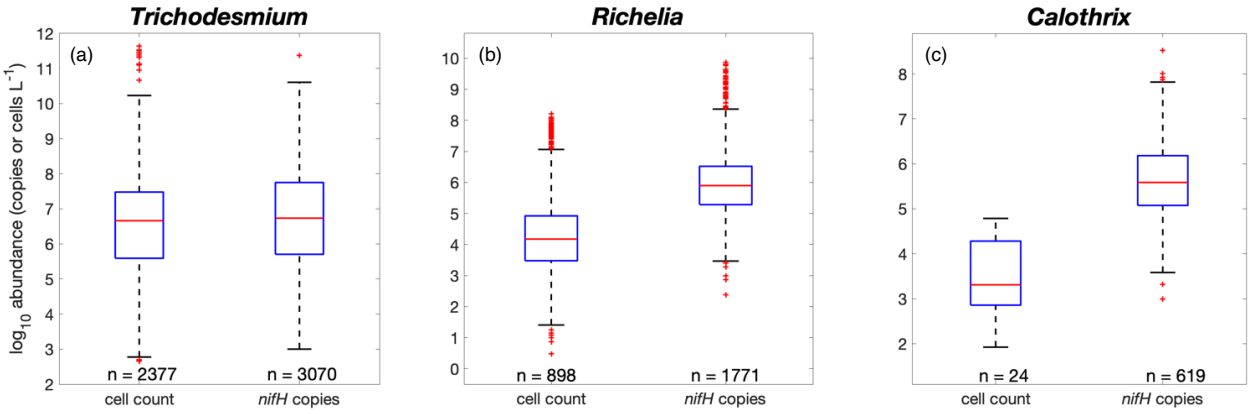


Figure ~~10~~15. Comparison of measured N_2 fixation rates using the original $^{15}N_2$ ~~dissolution~~bubble method and the $^{15}N_2$ ~~bubble assays. The blue line represents the 1:1 ratio of the two methods~~dissolution method. The pink dots are measurements. The fitted results of the two methods by the generalized additive model (GAM) and confidence intervals are represented by the red solid line and the dashed black lines, respectively. Only the N_2 fixation rates measured with incubation periods of 24 hours were included in this analysis.

4.2 Comparison between diazotrophic cell counts and *nifH* copies

Whether *nifH* copies can be used to infer diazotrophic abundance ~~remains debated~~and to study diazotrophic biogeography, while some challenges remain in conversion of gene counts to biomass, as a large range in the number of *nifH* copies per diazotrophic cell has been reported (**Table ~~S1~~**). ~~The reported ratios of *nifH* copies to cell numbers in~~**S2**). In version 2, we first converted *Trichodesmium* ~~and heterocystous cyanobacteria appeared larger than those~~trichome abundance to cell abundance using the same conversion factor of 100 cells trichome^{-1} as that used in ~~UCYNs (**Table S1**), possibly caused by large genome size in *Trichodesmium* (Sargent~~Luo et al~~., 2016~~. (2012) ~~and inclusion of the *nifH* gene in vegetative cells of heterocystous filaments~~. This conversion resulted in mean and variance of log-10 transformed *Trichodesmium* cell abundance ($10^{6.5\pm1.3}$ cells L^{-1}) very similar to that of *Trichodesmium nifH* gene copies ($10^{6.6\pm1.5}$ copies L^{-1}) (**Fig. 16a**). More recently, however, a much lower conversion factor of 13.2±2.3 cells trichome^{-1} was suggested for *Trichodesmium* based on larger sample sizes, although a very large range of 1.2–685 cells trichome^{-1} were reported (White et al., 2018)~~. In our database, the log-10 transformed abundance of *Trichodesmium* cell counts ($10^{6.5\pm1.2}$ cells L^{-1}) was only slightly lower than the log-10 transformed abundance of *Trichodesmium nifH* copies ($10^{6.6\pm1.6}$ copies L^{-1}) (**Fig. 11a**), while the difference was approximately two orders of magnitude~~

~~in heterocystous cyanobacteria *Richelia* and *Calothrix* (**Fig. 11b, c**). It must be noted that this simple analysis used all the data in our database, and the cell counts and *nifH* copies were measured for samples that were not in the same season and location. Much lower ratios of cell counts to *nifH* copies (1.51 – 2.58) were recently reported in heterocystous cyanobacteria and UCYN-B collected near the Hawaii Islands (Gradoville et al., 2022), demonstrating potentially large variance in these ratios.~~



. Hence, when a conversion factor of 10 cells trichome^{-1} was applied, the *Trichodesmium nifH* gene copy abundance was an order of magnitude higher than its cell abundance (**Fig. 16a**). This result was within the reported mean *nifH*:cell ratios for *Trichodesmium*, albeit based on sparse samples, on the order of 10 – 100 (**Table S2**). It is worth noting that there have been suggestions that the observed *nifH*:cell ratio for *Trichodesmium* may be overestimated due to methodological limitations (Gradoville et al., 2022). Our analyses underscore the importance of enumerating *Trichodesmium* cells, rather than solely focusing on trichomes, in future studies, as suggested by White et al. (2018). While counting all *Trichodesmium* cells may be impractical, it would be valuable to report the number of cells in random samples of *Trichodesmium* trichomes.

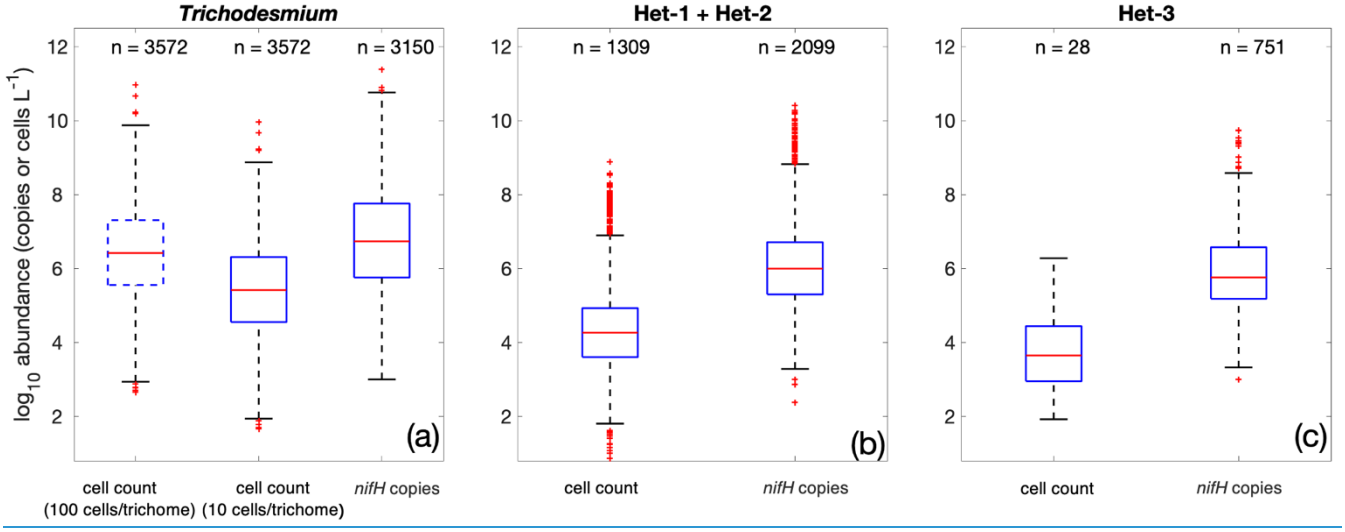


Figure ~~11~~16. Comparison of all cell-count and *nifH* gene copy abundance data in the database. The box plots show the median (central line), 25th and 75th percentiles ~~of log-10 transformed data~~ (upper and lower edges of the boxes), 5th and 95th percentiles (error lines) and outliers (red crosses~~).~~) of log-10 transformed data. The comparisons are conducted for (**a**) *Trichodesmium*, (**b**) het-1/2, (**c**) het-3. Note that two conversion factors of 10 and 100 cells trichome^{-1} are used for *Trichodesmium*.

The same analyses for heterocystous cyanobacteria showed that the *nifH* gene copy abundances were approximately two orders of magnitude greater than the cell abundances in terms of both mean and distribution (**Fig. 16b, c**). It must be noted that this simple analysis used all the data in our database. The limited *in situ* measurements for identical samples resulted in a mean *nifH*:cell ratio of 76 for heterocystous cyanobacteria (**Table S2**), consistent with our simple analysis.

In contrast, much lower *nifH*:cell ratios (1.51 – 2.58) were derived from regression analysis for heterocystous cyanobacteria and UCYN-B collected in the subtropical North Pacific (Gradoville et al., 2022). Considering these overall scarce measurements and the outcomes of our analysis, it is plausible that there is substantial variability in *nifH*:cell ratios. We expect that future studies, focusing on constraining these ratios and identifying mechanisms underlying variability in these ratios, will contribute to a more comprehensive understanding of the connection between *nifH* gene counts and diazotrophic cell abundance.

The application of qPCR assays for *nifH* based abundance (DNA) and expression (RNA) emerged as a critical step forward in our understanding of the distribution, abundance, and physiology (e.g., expression of *nifH*) of diazotrophs (Short and Zehr, 2005; Zehr and Riemann, 2023). Until then, estimating the abundances of diazotrophs were limited to those that could be identified by microscopy, e.g., *Trichodesmium*, heterocystous cyanobacteria (e.g., *Richelia*, *Calothrix, Anabaena, Nodularia, Aphanizomenon*), and some unicellulars (e.g., *Cyanothece*, later *Crocosphaera*). Thus, qPCR enabled the study of diazotrophic targets (and their activity) without the need to microscopy identify them, which came later as some diazotrophs would (and still) require application of FISH techniques for identification (Biegala and Raimbault, 2008). Additionally, qPCR allowed the study of *in situ* activity (gene expression) by diazotrophs without the need for cultivation. Although beyond the scope of the work presented here, important considerations should be taken into account when using microscopy and qPCR datasets (Table S3), for example, in application to biogeochemical models (Meiler et al., 2023).

4.3 Biomass conversion factor

For ~~the~~ possible further ~~use~~usage of cell-~~count or *nifH* based~~counted abundance data, here we suggest ~~factors to convert the abundance to~~ carbon biomass conversion factors for different diazotrophic groups (**Tables ~~6~~10 and ~~S2~~S4**). Most biomass conversion factors suggested here are the same as those used in Luo et al. (2012)~~, except for heterocystous cyanobacteria and~~, excluding UCYN-A and heterocystous cyanobacteria where new information has become available or additional consideration is necessary. A recent study has discovered a new symbiosis association between the unicellular diazotroph

(UCYN-C) and diatom *Epithemia* strains (Schvarcz et al., 2022). However, the conversion factor of UCYN-C could not be updated in this study due to insufficient information on the biovolumes of host cell.

The conversion factor for UCYN-A was updated because it has been found to live symbiotically with haptophyte *Braarudosphaera bigelowii* and relatives (Thompson et al., 2012; Hagino et al., 2013). Because the host and UCYN-A should function together, the host biomass is allocated to UCYN-A. It has been reported that each haptophyte cell hosts one UCYN-A1 cell (Cornejo-Castillo et al., 2019) or one UCYN-A2 cell (Suzuki et al., 2021). We used the empirically-derived equation (Verity et al., 1992):

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

to estimate biomass of UCYN-A and their hosts. The biomass of a UCYN-A1 cell with a diameter of 1 µm and a UCYN-A2 cell with a diameter of 1.6–3.3 µm (Cornejo-Castillo et al., 2019; Martínez-Pérez et al., 2016) equate to 0.2 pg C and 0.8–5.5 pg C, respectively. The biomasses of the host cell for UCYN-A1 or UCYN-A2 is 1.5–2.2 pg C or 6.8–43 pg C according to their reported cell diameters (2–2.3 µm or 3.6–7.3 µm), respectively (Martínez-Pérez et al., 2016; Cornejo-Castillo et al., 2019). Hence, the biomasses of the UCYN-A1 and the UCYN-A2 symbioses are 1.7–2.4 pg C and 7.6–48 pg C, respectively. After normalizing the symbiotic biomass to the number of UCYN cells in each symbiosis (1 for both UCYN-A1 and UCYN-A2), the biomass conversion factors are 1.7–2.4 pg C (UCYN-A1 cell)$^{-1}$ and 7.6–48 pg C (UCYN-A2 cell)$^{-1}$.

Because heterocystous cyanobacteria and their host diatoms form DDAs, similar to UCYN-A, we also suggest allocating the biomass of host diatoms to each associated diazotrophic cell (**Table S4**). The biomasses of heterocystous cells and vegetative cells in *Richelia* filaments were updated according to the cell dimension data reported in Caputo et al. (2019) using the same empirical equation above. The carbon biomass of host diatom cells was calculated using an empirical equation (Menden-Deuer and Lessard, 2000):

$$C = 0.117 \times V^{0.881}, \quad (2)$$

~~UCYN-A because new information has become available or additional consideration is necessary.~~

~~Because heterocystous cyanobacteria and their host diatoms form DDAs and need to function together, we suggest allocating the biomass of host diatoms to each associated diazotrophic cell (**Table S2**). The carbon biomass of host diatom cells was calculated using an empirical equation (Menden-Deuer and Lessard, 2000):~~

~~$C = 0.117 \times V^{0.881}$, (1)~~

where C is the diatom cell carbon biomass (pg C cell^{-1}), and V is the average cell biovolume (µm^3) of each diatom genus, for which values from a database ~~(Harrison et al., 2015)~~(Harrison et al., 2015) were used in this study (**Table ~~S2~~S4**). Each host diatom associates with multiple heterocysts. The numbers of *Richelia* heterocysts ~~in~~ associated with *Hemiaulus*, *Rhizosoleniae* and ~~*Rhizosoleni*~~*Chaetoceros* were observed to be ~~2 and 5 (Villareal et al., 2011; Yeung et al., 2012), and the~~within the range of 1–2, 1–5 and 3–10 respectively (Villareal et al., 2011; Yeung et al., 2012; Caputo et al., 2019), we selected both the maximum and minimum to do the estimation. The number of vegetative cells in each heterocyst ~~can range from 3 to 10 (Foster et al., 2011). We tried 2 and 5 *Calothrix* heterocysts in *Chaetocero*s when estimating the biomass conversion factor, although the values were unknown. The biomasses of heterocystous cells and vegetative cells (2–80 pg C cell^{-1} for *Richelia* and 5–20~~

~~pg C cell^{-1} for *Calothrix*)~~ were ~~adopted from Luo et al. (2012). Hence, the conversion~~ also updated according to Caputo et al. (2019). Conversion factors for DDAs ~~are~~were estimated by dividing the total biomass of each DDA by the number of associated heterocysts. ~~As a result, the~~Changes in the number of *Richelia* in *Rhizosoleniae* (1 or 5) would make a large variation in its conversion factor, possibly due to large host biomass, therefore we keep them both to let users take caution when using this conversion factor. The resulting biomass conversion factors of *Richelia-Hemiaulus* and *Richelia-~~Rhizosolenia~~Chaetoceros* associations were estimated to be 280 (range: 150–1250) and 430 (range: 10–1900) pg C heterocyst^{-1}, respectively (**Table ~~S2~~S4**), ~~As~~ as the number of filaments ~~(2 or 5)~~ did not have a large impact on the ~~factors of *Calothrix-Chaetoceros* associations, we recommend using 100 (range: 20–360) pg C heterocyst^{-1}~~ conversion factors.

~~The conversion factor for UCYN-A is also updated because it has been found to live symbiotically with prymnesiophyte or coccolitophore species (Thompson et al., 2012). Similar to DDAs, the host biomass is allocated to UCYN-A. It has been reported that each prymnesiophyte cell hosts one UCYN-A1 cell, and each coccolitophore hosts 5–10 UCYN-A2 cells (Cornejo-Castillo et al., 2019). The biomass of a UCYN-A1 cell with a diameter of 1 μm and a UCYN-A2 cell with a diameter of 1.6–3.3 μm (Cornejo-Castillo et al., 2019; Martínez-Pérez et al., 2016) is 0.2 and 0.8–5.5 pg C, respectively, by using an empirical equation (Verity et al., 1992):~~

~~$C = 0.433 \times V^{0.863}$, (2)~~

~~Using Equation (2), the biomass of a host prymnesiophyte or coccolitophore cell is 1.5–2.2 pg C or 6.8–43 pg C according to their reported cell diameters (2–2.3 μm and 3.6–7.3 μm, respectively) (Martínez-Pérez et al., 2016; Cornejo-Castillo et al., 2019). Hence, the biomass of the UCYN-A1 symbiosis and the UCYN-A2 symbiosis is 1.7–2.4 pg C and 7.6–48 pg C, respectively. After normalizing the symbiotic biomass to the number of UCYN cells in each symbiosis (1 for UCYN-A1 and 5–10 for UCYN-A2), the biomass conversion factors are 1.7–2.4 pg C (UCYN-A1 cell)$^{-1}$ and 0.8–9.6 pg C (UCYN-A2 cell)$^{-1}$. Thus, we recommend using a uniform conversion factor of 2 pg C cell^{-1} for these two clades of UCYN-A (**Table 7**) considering that UCYN-A1 is more frequently found and often in higher abundance than UCYN-A2 (Thompson et al., 2014).~~

~~**Table 6.** Biomass conversion factors for diazotrophs.~~

It is important to reiterate that these biomass conversion factors are only applicable to cell-count data. Attempting to convert *nifH* gene copies to biomass is not recommended due to significant uncertainties associated with *nifH*:cell, as previously discussed.

Table 10. Recommended carbon biomass conversion factors and their likely ranges for diazotrophic groups.

~~*Trichodesmium*~~	UCYN-~~A~~A1 (pg C cell^{-1})	UCYN-A2 (pg C cell^{-1})	UCYN-B (pg C cell^{-1})	UCYN-C (pg C cell^{-1})	**Het-1** *Richelia-Hemiaulus* (pg C heterocyst^{-1})	**Het-2** *Richelia-Rhizosolenia* (pg C heterocyst^{-1})	~~*Calothrix*~~**Het-3** *Richelia-Chaetoceros*

	Trichodesmium (pg C cell^{-1})							(pg C heterocyst^{-1})
Recommended	300	2	30	20	10	280350	450 (5 heterocyst DDA^{-1}) or 1900430 (1 heterocyst DDA^{-1})	10050
RangeLikely range	100–500	1–103	10-50	4–50	5–24	150–12501030	10–19–57001900	20–3609–300

5. Conclusions

This study, we updated the global oceanic diazotrophic database by Luo et al. (2012) by adding new measurements reported in the past decade. Although the spatial coverage of the data was greatly expanded by this effort, the data distribution is still uneven, with most measurements reported from the Pacific and Atlantic Oceans. Using the updated database, the estimation of global oceanic N_2 fixation based on rates in ocean basins was increased from 74±7 Tg N yr^{-1} to 217±29 Tg N yr^{-1}. This change is largely attributable to a new estimate for the Indian Ocean, and a much elevated estimate for the South Pacific Ocean that would account for ~40% of global N_2 fixation, this high estimation for the South Pacific Ocean is in line with its qualification as a 'hot spot' for diazotrophy (Messer et al., 2016; Bonnet et al., 2017), partly due to iron fertilization processes in this region (Bonnet et al., 2023). Due to data sparsity, our updated estimation did not include N_2 fixation in the Southern and Arctic Oceans. Furthermore, data were more concentrated in surface seawater, and a significant amount of data were measured with incubation periods shorter than a daily cycle (24 h), limiting reliable evaluations of depth-integrated N_2 fixation rates. Although this result suggests more balanced N inputs and losses in the global ocean than the previous estimate suggested, large uncertainties still exist. We also compared the N_2 fixation rates measured using addition of a bubble of labelled gas or addition of dissolving $^{15}N_2$ gases reported at the same location and month (not necessarily in identical samples). The results indicated that the original $^{15}N_2$ bubble method produces lower rates, on average, than the $^{15}N_2$ dissolution method, although in 69% of the former method even generated higher rates in approximately one third of cases. These results reveal that, despite the efforts in the past several decades of effort, the ocean is still undersampled in terms of the distribution of diazotrophs and level of N_2 fixation rate measurements. Our analyses suggest that prioritizing N_2 fixation measurements in the South Pacific Ocean and high northern latitudes can significantly

reduce the current uncertainty of N_2 fixation rates in the global ocean. Nevertheless, we believe that this updated diazotrophic database, ~~updated~~supplemented with ~~a large amount of~~enhanced data ~~accumulated in~~from the past decade, is timely and can ~~help~~be helpful to scientists ~~study~~studying the marine ~~ecology and biogeochemistry~~biogeochemical cycle of N.

Data availability.

~~The database is available in a data repository (https://doi.org/10.6084/m9.figshare.21677687) (Shao et al., 2022)~~

~~Author contributions.~~

~~YWL~~The database is available in a data repository (https://doi.org/10.6084/m9.figshare.21677687) (Shao et al., 2022)

Author contributions.

Y.-W. Luo conceived and designed the structure of the database ~~and supervised the study. All the authors~~. Z. Shao, Y. Xu, H. Wang, W. Luo, L. Wang, Y. Huang and Y.-W.Luo collected the data and updated the database. ~~ZS, YX, HW and YWL~~Z. Shao, Y. Xu, H. Wang, S. C. Doney and Y.-W. Luo analyzed data. Other authors contributed the data. ~~ZS, YX~~Z. Shao, Y. Xu and ~~YWL~~Y.-W. Luo wrote the first draft of the manuscript, and all authors revised the manuscript.

Competing interests.

The authors declare that they have no conflicts of interest.

Acknowledgments.

We would like to thank all the scientists ~~who collected and shared the sampling data.~~and crew who contributed to sample and measure these tremendous amounts of diazotrophic data in the past several decades. We also thank Christopher Somes and an anonymous reviewer for their constructive comments. This work was supported by the National Natural Science Foundation of China (grants 41890802 and 42076153).

References

Agawin, N. S. R., Tovar-Sánchez, A., De Zarruk, K. K., Duarte, C. M., and Agustí, S.: Variability in the abundance of *Trichodesmium* and nitrogen fixation activities in the subtropical NE Atlantic, Journal of Plankton Research, 35, 1126-1140, https://doi.org/10.1093/plankt/fbt059, 2013.

Ahmed, A., Gauns, M., Kurian, S., Bardhan, P., Pratihary, A., Naik, H., Shenoy, D. M., and Naqvi, S. W. A.: Nitrogen fixation rates in the eastern Arabian Sea, Estuarine, Coastal and Shelf Science, 191, 74-83, https://doi.org/10.1016/j.ecss.2017.04.005, 2017.

Barthel, K.-G., Schneider, G., Gradinger, R., and Lenz, J.: Concentration of live pico- and nanoplankton by means of tangential flow filtration, Journal of Plankton Research, 11, 1213-1221, https://doi.org/10.1093/plankt/11.6.1213, 1989.

Benavides, M., Bonnet, S., Berman-Frank, I., and Riemann, L.: Deep into oceanic N_2 fixation, ~~Front. Mar. Sci.,~~Frontiers in Marine Science, 5, 108, https://doi.org/10.3389/fmars.2018.00108, 2018a.

Benavides, M., Santana-Falcón, Y., Wasmund, N., and Aristegui, J.: Microbial uptake and regeneration of inorganic nitrogen off the coastal Namibian upwelling system, Journal of Marine Systems, https://doi.org/10.1016/j.jmarsys.2014.05.002, 2014.

Benavides, M., Agawin, N. S. R., Arístegui, J., Peene, J., and Stal, L. J.: Dissolved organic nitrogen and carbon release by a marine unicellular diazotrophic cyanobacterium, Aquat Microb Ecol, 69, 69-80, https://doi.org/10.3354/ame01621 2013a.

Benavides, M., Berthelot, H., Duhamel, S., Raimbault, P., and Bonnet, S.: Dissolved organic matter uptake by *Trichodesmium* in the Southwest Pacific, Scientific Reports, 7, 41315, https://doi.org/10.1038/srep41315, 2017.

Benavides, M., Moisander, P. H., Daley, M. C., Bode, A., and Aristegui, J.: Longitudinal variability of diazotroph abundances in the subtropical North Atlantic Ocean, J. Plankton Res., 38, 662-672, https://doi.org/10.1093/plankt/fbv121, 2016a.

Benavides, M., Bronk, D. A., Agawin, N. S. R., Pérez-Hernández, M. D., Hernández-Guerra, A., and Arístegui, J.: Longitudinal variability of size-fractionated N_2 fixation and DON release rates along 24.5°N in the subtropical North Atlantic, Journal of Geophysical Research: Oceans, 118, 3406-3415, https://doi.org/10.1002/jgrc.20253, ~~2013~~2013b.

Benavides, M., Moisander, P. H., Berthelot, H., Dittmar, T., Grosso, O., and Bonnet, S.: Mesopelagic N_2 fixation related to organic matter composition in the Solomon and Bismarck Seas (Southwest Pacific), Plos One, 10, 12, https://doi.org/10.1371/journal.pone.0143775, 2015.

Benavides, M., Conradt, L., Bonnet, S., Berman-Frank, I., Barrillon, S., Petrenko, A., and Doglioli, A.: Fine-scale sampling unveils diazotroph patchiness in the South Pacific Ocean, ISME Communications, 1, 3, https://doi.org/10.1038/s43705-021-00006-2, 2021.

Benavides, M., Caffin, M., Duhamel, S., Foster, R. A., Grosso, O., Guieu, C., Van Wambeke, F., and Bonnet, S.: Anomalously high abundance of *Crocosphaera* in the South Pacific Gyre, FEMS Microbiol Lett, 369, https://doi.org/10.1093/femsle/fnac039, ~~2022~~2022a.

Benavides, M., Shoemaker, K. M., Moisander, P. H., Niggemann, J., Dittmar, T., Duhamel, S., Grosso, O., Pujo-Pay, M., Helias-Nunige, S., Fumenia, A., and Bonnet, S.: Aphotic N_2 fixation along an oligotrophic to ultraoligotrophic transect in the western tropical South Pacific Ocean, Biogeosciences, 15, 3107-3119, https://doi.org/10.5194/bg-15-3107-2018, 2018b.

Benavides, M., Bonnet, S., Hernandez, N., Martinez-Perez, A. M., Nieto-Cid, M., Alvarez-Salgado, X. A., Banos, I., Montero, M. F., Mazuecos, I. P., Gasol, J. M., Osterholz, H., Dittmar, T., Berman-Frank, I., and Aristegui, J.: Basin-wide N_2 fixation in the deep waters of the Mediterranean Sea, Global Biogeochemical Cycles, 30, 952-961, https://doi.org/10.1002/2015gb005326, 2016b.

Benavides, M., Bonnet, S., Le Moigne, F. A. C., Armin, G., Inomura, K., Hallstrøm, S., Riemann, L., Berman-Frank, I., Poletti, E., Garel, M., Grosso, O., Leblanc, K., Guigue, C., Tedetti, M., and Dupouy, C.: Sinking *Trichodesmium* fixes nitrogen in the dark ocean, The ISME Journal, 16, 2398-2405, https://doi.org/10.1038/s41396-022-01289-6, 2022b.

Bentzon-Tilia, M., Severin, I., Hansen, L. H., and Riemann, L.: Genomics and Ecophysiology of Heterotrophic Nitrogen-Fixing Bacteria Isolated from Estuarine Surface Water, https://doi.org/10.1128/mbio.00929-15, 2015a.

Bentzon-Tilia, M., Traving, S. J., Mantikci, M., Knudsen-Leerbeck, H., Hansen, J. L., Markager, S., and Riemann, L.: Significant N_2 fixation by heterotrophs, photoheterotrophs and heterocystous cyanobacteria in two temperate estuaries, The ISME Journal, 9, 273-285, https://doi.org/10.1038/ismej.2014.119, ~~2015~~2015b.

Berthelot, H., Benavides, M., Moisander, P. H., Grosso, O., and Bonnet, S.: High-nitrogen fixation rates in the particulate and dissolved pools in the Western Tropical Pacific (Solomon and Bismarck Seas), Geophys. Res. Lett., 44, 8414-8423, https://doi.org/10.1002/2017gl073856, 2017.

Berthelot, H., Bonnet, S., Camps, M., Grosso, O., and Moutin, T.: Assessment of the dinitrogen released as ammonium and dissolved organic nitrogen by unicellular and filamentous marine diazotrophic cyanobacteria grown in culture, Front. Mar. Sci., 2, https://doi.org/10.3389/fmars.2015.00080, 2015.

Berthelot, H., Duhamel, S., L'Helguen, S., Maguer, J.-F., Wang, S., Cetinić, I., and Cassar, N.: NanoSIMS single cell analyses reveal the contrasting nitrogen sources for small phytoplankton, The ISME Journal, 13, 651-662, 10.1038/s41396-018-0285-8, 2019.

Bhavya, P. S., Kumar, S., Gupta, G. V. M., Sudheesh, V., Sudharma, K. V., Varrier, D. S., Dhanya, K. R., and Saravanane, N.: Nitrogen uptake dynamics in a tropical eutrophic estuary (Cochin, India) and adjacent coastal waters, Estuaries and Coasts, 39, 54-67, https://doi.org/10.1007/s12237-015-9982-y, 2016.

~~Bhavya, P. S., Min, J. O., Kim, M. S., Jang, H. K., Kim, K., Kang, J. J., Lee, J. H., Lee, D., Jo, N., Kim, M. J., Kim, Y., Lee, J., Lee, C. H., Bae, H., Yoo, H., Park, S., Yun, M. S., and Lee, S. H.: A review on marine N_2 fixation: Mechanism, evolution of methodologies, rates, and future concerns, Ocean Science Journal, 54, 515-528, https://doi.org/10.1007/s12601-019-0037-3, 2019.~~

Biegala, I. and Raimbault, P.: High abundance of diazotrophic picocyanobacteria (<3 µm) in a Southwest Pacific coral lagoon, ~~Aquat Microb Ecol~~Aquatic Microbial Ecology, 51, 45-53, https://doi.org/10.3354/ame01185, 2008.

Bif, M. and Yunes, J.: Distribution of the marine cyanobacteria *Trichodesmium* and their association with iron-rich particles in the South Atlantic Ocean, ~~Aquat Microb Ecol~~Aquatic Microbial Ecology, 78, 107-119, https://doi.org/10.3354/ame01810, 2017.

Bird, C., Martinez, M. J., O'Donnell, A. G., and Wyman, M.: Spatial distribution and transcriptional activity of an uncultured clade of planktonic diazotrophic γ-proteobacteria in the Arabian Sea, Applied and Environmental Microbiology, 71, 2079-2085, https://doi.org/10.1128/AEM.71.4.2079-2085.2005, 2005.

Blais, M., Tremblay, J. É., Jungblut, A. D., Gagnon, J., Martin, J., Thaler, M., and Lovejoy, C.: Nitrogen fixation and identification of potential diazotrophs in the Canadian Arctic, Global Biogeochemical Cycles, 26, 1-13, https://doi.org/10.1029/2011gb004096, 2012.

Bombar, D., Paerl, R. W., and Riemann, L.: Marine non-cyanobacterial diazotrophs: moving beyond molecular detection, Trends Microbiol., 24, 916-927, https://doi.org/10.1016/j.tim.2016.07.002, 2016.

Bombar, D., Moisander, P. H., Dippner, J. W., Foster, R. A., Voss, M., Karfeld, B., and Zehr, J. P.: Distribution of diazotrophic microorganisms and *nifH* gene expression in the Mekong River plume during intermonsoon, Mar. Ecol. Prog. Ser., 424, 39-55, https://doi.org/10.3354/meps08976, 2011.

Bombar, D., Taylor, C. D., Wilson, S. T., Robidart, J. C., Rabines, A., Turk-Kubo, K. A., Kemp, J. N., Karl, D. M., and Zehr, J. P.: Measurements of nitrogen fixation in the oligotrophic North Pacific Subtropical Gyre using a free-drifting submersible incubation device, ~~J.~~Journal of Plankton ~~Res.~~,Research, 37, 727-739, https://doi.org/10.1093/plankt/fbv049, 2015.

Bonnet, S., Caffin, M., Berthelot, H., and Moutin, T.: Hot spot of N_2 fixation in the western tropical South Pacific pleads for a spatial decoupling between N_2 fixation and denitrification, P Natl Acad Sci USA, 114, E2800 - E2801, https://doi.org/10.1073/pnas.1619514114, 2017.

Bonnet, S., Dekaezemacker, J., Turk-Kubo, K. A., Moutin, T., Hamersley, R. M., Grosso, O., Zehr, J. P., and Capone, D. G.: Aphotic N_2 ~~fixation~~Fixation in the Eastern Tropical South Pacific Ocean, PLoS ~~One~~ONE, 8, ~~14~~e81265, https://doi.org/10.1371/journal.pone.0081265, 2013.

Bonnet, S., Caffin, M., Berthelot, H., Grosso, O., Benavides, M., Helias-Nunige, S., Guieu, C., Stenegren, M., and Foster, R. A.: In-depth characterization of diazotroph activity across the western tropical South Pacific hotspot of N_2 fixation (OUTPACE cruise), Biogeosciences, 15, 4215-4232, https://doi.org/10.5194/bg-15-4215-2018, 2018.

Bonnet, S., Rodier, M., Turk-Kubo, K. A., Germineaud, C., Menkes, C., Ganachaud, A., Cravatte, S., Raimbault, P., Campbell, E., Quéroué, F., Sarthou, G., Desnues, A., Maes, C., and Eldin, G.: Contrasted geographical distribution of N_2 fixation rates and *nifH* phylotypes in the Coral and Solomon Seas (southwestern Pacific) during austral winter conditions, Global Biogeochemical Cycles, 29, 1874-1892, https://doi.org/10.1002/2015gb005117, 2015.

Bonnet, S., Guieu, C., Taillandier, V., Boulart, C., Bouruet-Aubertot, P., Gazeau, F., Scalabrin, C., Bressac, M., Knapp, A., Cuypers, Y., González-Santana, D., Forrer, H., Grisoni, J. M., Grosso, O., Habasque, J., Jardin-Camps, M., Leblond, N., Le Moigne, F., Lebourges-Dhaussy, A., and Tilliette, C.: Natural iron fertilization by shallow hydrothermal sources fuels diazotroph blooms in the ocean, Science (New York, N.Y.), 380, 812-817, http://dx.doi.org/10.1126/science.abq4654, 2023.

Böttjer, D., Dore, J. E., Karl, D. M., Letelier, R. M., Mahaffey, C., Wilson, S. T., Zehr, J., and Church, M. J.: Temporal variability of nitrogen fixation and particulate nitrogen export at Station ALOHA, Limnology and Oceanography, 62, 200-216, https://doi.org/10.1002/lno.10386, 2017.

Breitbarth, E., Mills, M. M., Friedrichs, G., and LaRoche, J.: The Bunsen gas solubility coefficient of ethylene as a function of temperature and salinity and its importance for nitrogen fixation assays, Limnology and Oceanography: Methods, 2, 282-288, https://doi.org/10.4319/lom.2004.2.282, 2004.

Cabello, A. M., Turk-Kubo, K. A., Hayashi, K., Jacobs, L., Kudela, R. M., and Zehr, J. P.: Unexpected presence of the nitrogen-fixing symbiotic cyanobacterium UCYN-A in Monterey Bay, California, Journal of Phycology, 56, 1521-1533, https://doi.org/10.1111/jpy.13045, 2020.

Campbell, L., Carpenter, E., Montoya, J., Kustka, A., and Capone, D.: Picoplankton community structure within and outside a *Trichodesmium* bloom in the southwestern Pacific Ocean, Vie et Milieu/Life & Environment, 185-195, https://hal.sorbonne-universite.fr/hal-03219052, 2005.

Capone, D. G.: Determination of nitrogenase activity in aquatic samples using the acetylene reduction procedure~~.~~, in: Handbook of Methods in Aquatic Microbial Ecology, ~~621~~ edited by: Kemp, P. F., Cole, J. J., Sherr, B. F., and Sherr, E. B., Lewis Publishers, Boca Raton, FL, 621–631, 1993.

Capone, D. G. and Montoya, J. P.: Nitrogen fixation and denitrification, ~~in:~~ Methods in Microbiology, ~~Academic Press~~30, 501-515, https://doi.org/10.1016/S0580-9517(01)30060-0, 2001.

Capone, D. G., Burns, J. A., Montoya, J. P., Subramaniam, A., Mahaffey, C., Gunderson, T., Michaels, A. F., and Carpenter, E. J.: Nitrogen fixation by *Trichodesmium* spp.: An important source of new nitrogen to the tropical and subtropical North Atlantic Ocean, Global Biogeochemical Cycles, 19, GB2024, https://doi.org/10.1029/2004GB002331, 2005.

Caputo, A., Nylander, J. A. A., and Foster, R. A.: The genetic diversity and evolution of diatom-diazotroph associations highlights traits favoring symbiont integration (vol 366, fny297, 2019), Fems Microbiology Letters, 366, 10.1093/femsle/fnz120, 2019.

Cassar, N., Tang, W., Gabathuler, H., and Huang, K.: Method for High Frequency Underway N(2) Fixation Measurements: Flow-Through Incubation Acetylene Reduction Assays by Cavity Ring Down Laser Absorption Spectroscopy (FARACAS), Anal Chem, 90, 2839-2851, https://doi.org/10.1021/acs.analchem.7b04977, 2018.

Cerdan-Garcia, E., Baylay, A., Polyviou, D., Woodward, E. M. S., Wrightson, L., Mahaffey, C., Lohan, M. C., Moore, C. M., Bibby, T. S., and Robidart, J. C.: Transcriptional responses of Trichodesmium to natural inverse gradients of Fe and P availability, The ISME Journal, 10.1038/s41396-021-01151-1, 2021.

Chang, B. X., Jayakumar, A., Widner, B., Bernhardt, P., Mordy, C. W., Mulholland, M. R., and Ward, B. B.: Low rates of dinitrogen fixation in the eastern tropical South Pacific, Limnology and Oceanography, 64, 1913-1923, https://doi.org/10.1002/lno.11159, 2019.

Chang, J., Chiang, K.-P., and Gong, G.-C.: Seasonal variation and cross-shelf distribution of the nitrogen-fixing cyanobacterium, *Trichodesmium*, in southern East China Sea, Continental Shelf Research, 20, 479-492, https://doi.org/10.1016/S0278-4343(99)00082-5, 2000.

Chen, L. Y.-L., Chen, H.-Y., Lin, Y.-H., Yong, T.-C., Taniuchi, Y., and Tuo, S.-H.: The relative contributions of unicellular and filamentous diazotrophs to N_2 fixation in the South China Sea and the upstream Kuroshio, Deep Sea Research Part I: Oceanographic Research Papers, 85, 56-71, https://doi.org/10.1016/j.dsr.2013.11.006, 2014.

Chen, M. M., Lu, Y. Y., Jiao, N. Z., Tian, J. W., Kao, S. J., and Zhang, Y.: Biogeographic drivers of diazotrophs in the western Pacific Ocean, Limnol Oceanogr, 64, 1403-1421, https://doi.org/10.1002/lno.11123, 2019.

Cheung, S., Liu, K., Turk-Kubo, K. A., Nishioka, J., Suzuki, K., Landry, M. R., Zehr, J. P., Leung, S., Deng, L., and Liu, H.: High biomass turnover rates of endosymbiotic nitrogen-fixing cyanobacteria in the western Bering Sea, Limnology and Oceanography Letters, 7, 501-509, https://doi.org/10.1002/lol2.10267, 2022.

Cheung, S. Y., Nitanai, R., Tsurumoto, C., Endo, H., Nakaoka, S., Cheah, W., Lorda, J. F., Xia, X. M., Liu, H. B., and Suzuki, K.: Physical forcing controls the basin-scale occurrence of nitrogen-fixing organisms in the North Pacific Ocean, Global Biogeochemical Cycles, 34, 9, https://doi.org/10.1029/2019GB006452, 2020.

Church, M. J. and Zehr, J.: Time series measurements of *nifH* gene abundances for several cyanobacteria in the subtropical North Pacific Ocean, Zenodo [dataset], https://doi.org/10.5281/zenodo.4728253, 2020.

Church, M. J., Jenkins, B. D., Karl, D. M., and Zehr, J. P.: Vertical distributions of nitrogen-fixing phylotypes at Stn ALOHA in the oligotrophic North Pacific Ocean, Aquat Microb EcolAquatic Microbial Ecology, 38, 3-14, https://doi.org/10.3354/ame038003, 20052005a.

Church, M. J., Björkman, K. M., Karl, D. M., Saito, M. A., and Zehr, J. P.: Regional distributions of nitrogen-fixing bacteria in the Pacific Ocean, Limnology and Oceanography, 53, 63-77, https://doi.org/10.4319/lo.2008.53.1.0063, 2008.

Church, M. J., Short, C. M., Jenkins, B. D., Karl, D. M., and Zehr, J. P.: Temporal Patterns of Nitrogenase Gene (*nifH*) Expression in the Oligotrophic North Pacific Ocean, Applied and Environmental Microbiology, 71, 5362-5370, https://doi.org/10.1128/aem.71.9.5362-5370.2005, 2005b.

Confesor, K. A., Selden, C. R., Powell, K. E., Donahue, L. A., Mellett, T., Caprara, S., Knapp, A. N., Buck, K. N., and Chappell, P. D.: Defining the Realized Niche of the Two Major Clades of Trichodesmium: A Study on the West Florida Shelf, Frontiers in Marine Science, 9, https://doi.org/10.3389/fmars.2022.821655, 2022.

Cornejo-Castillo, F. M., Muñoz-MarínMunoz-Marin, M. dD. C., Turk-Kubo, K. A., Royo-Llonch, M., Farnelid, H., Acinas, S. G., and Zehr, J. P.: UCYN-A3, a newly characterized open ocean sublineage of the symbiotic N_2N(2) -fixing cyanobacterium Candidatus Atelocyanobacterium thalassa, Environ- Microbiol-, 21, 111-124, https://doi.org/10.1111/1462-2920.14429, 2019.

Cornejo-Castillo, F. M., Cabello, A. M., Salazar, G., Sánchez-Baracaldo, P., Lima-Mendez, G., Hingamp, P., Alberti, A., Sunagawa, S., Bork, P., de Vargas, C., Raes, J., Bowler, C., Wincker, P., Zehr, J. P., Gasol, J. M., Massana, R., and Acinas, S. G.: Cyanobacterial symbionts diverged in the late Cretaceous towards lineage-specific nitrogen fixation factories in single-celled phytoplankton, Nat. Commun.,Nature Communications, 7, 11071, https://doi.org/10.1038/ncomms11071, 2016.

Dabundo, R., Lehmann, M. F., Treibergs, L., Tobias, C. R., Altabet, M. A., Moisander, P. H., and Granger, J.: The contamination of commercial $^{15}N_2$ gas stocks with ^{15}N-labeled nitrate and ammonium and consequences for nitrogen fixation measurements, PLoS One, 9, e110335, https://doi.org/10.1371/journal.pone.0110335, 2014.

Dekaezemacker, J., Bonnet, S., Grosso, O., Moutin, T., Bressac, M., and Capone, D. G.: Evidence of active dinitrogen fixation in surface waters of the eastern tropical South Pacific during El Nino and La Nina events and evaluation of its potential nutrient controls, Global Biogeochemical Cycles, 27, 768-779, https://doi.org/10.1002/gbc.20063, 2013.

Delmont, T. O., Pierella Karlusich, J. J., Veseli, I., Fuessel, J., Eren, A. M., Foster, R. A., Bowler, C., Wincker, P., and Pelletier, E.: Heterotrophic bacterial diazotrophs are more abundant than their cyanobacterial counterparts in metagenomes covering most of the sunlit ocean, The ISME Journal, https://doi.org/10.1038/s41396-021-01135-1, 2021.

Detoni, A. M. S., Ciotti, Á. M., Calil, P. H. R., Tavano, V. M., and Yunes, J. S.: *Trichodesmium* latitudinal distribution on the shelf break in the southwestern Atlantic Ocean during spring and autumn, Global Biogeochemical Cycles, 30, 1738-1753, https://doi.org/10.1002/2016gb005431, 2016.

Detoni, A. M. S., Subramaniam, A., Haley, S. T., Dyhrman, S. T., and Calil, P. H. R.: Cyanobacterial diazotroph distributions in the western South Atlantic, ~~Front. Mar. Sci.,~~Frontiers in Marine Science, 9, https://doi.org/10.3389/fmars.2022.856643, 2022.

Deutsch, C., Sarmiento, J. L., Sigman, D. M., Gruber, N., and Dunne, J. P.: Spatial coupling of nitrogen inputs and losses in the ocean, Nature, 445, 163-167, https://doi.org/10.1038/nature05392, 2007.

Dugenne, M., Gradoville, M., Church, M., Wilson, S., Sheyn, U., Harke, M., Björkman, K., Hawco, N., Hynes, A., Ribalet, F., Karl, D., DeLong, E., Dyhrman, S., Armbrust, E., John, S., Eppley, J., Harding, K., Stewart, B., Cabello, A., and Zehr, J.: Nitrogen Fixation in Mesoscale Eddies of the North Pacific Subtropical Gyre: Patterns and Mechanisms, Global Biogeochemical Cycles, 37, https://doi.org/10.1029/2022GB007386, 2023.

Dupouy, C., Benielli-Gary, D., Neveux, J., Dandonneau, Y., and Westberry, T. K.: An algorithm for detecting *Trichodesmium* surface blooms in the South Western Tropical Pacific, Biogeosciences, 8, 3631-3647, https://doi.org/10.5194/bg-8-3631-2011, 2011.

Estrada, M., Delgado, M., Blasco, D., Latasa, M., Cabello, A. M., Benítez-Barrios, V., Fraile-Nuez, E., Mozetič, P., and Vidal, M.: Phytoplankton across tropical and subtropical regions of the Atlantic, Indian and Pacific oceans, PLoS One, 11, e0151699, https://doi.org/10.1371/journal.pone.0151699, 2016.

Farnelid, H., Turk-Kubo, K., Muñoz-Marín, M. C., and Zehr, J. P.: New insights into the ecology of the globally significant uncultured nitrogen-fixing symbiont UCYN-A, Aquatic Microbial Ecology, 77, 125-138, https://doi.org/10.3354/ame01794, 2016.

Farnelid, H., Bentzon-Tilia, M., Andersson, A. F., Bertilsson, S., Jost, G., Labrenz, M., Jürgens, K., and Riemann, L.: Active nitrogen-fixing heterotrophic bacteria at and below the chemocline of the central Baltic Sea, The ISME Journal, 7, 1413-1423, https://doi.org/10.1038/ismej.2013.26, 2013.

~~Fernández-Castro, B., Mouriño-Carballido, B., Marañón, E., Chouciño, P., Gago, J., Ramírez, T., Vidal, M., Bode, A., Blasco, D., Royer, S.-J., Estrada, M., and Simó, R.: Importance of salt fingering for new nitrogen supply in the oligotrophic ocean, Nature Communications, 6, 8002, https://doi.org/10.1038/ncomms9002, 2015.~~

Fernández, A., Mouriño-Carballido, B., Bode, A., Varela, M., and Marañón, E.: Latitudinal distribution of Trichodesmium spp. and N_2 fixation in the Atlantic Ocean, Biogeosciences, 7, 3167-3176, https://doi.org/10.5194/bg-7-3167-2010, 2010.

Fernandez, C., González, M. L., Muñoz, C., Molina, V., and Farias, L.: Temporal and spatial variability of biological nitrogen fixation off the upwelling system of central Chile (35-38.5°S), J Geophys Res-Oceans, 120, 3330-3349, https://doi.org/10.1002/2014jc010410, 2015.

Fernández-Castro, B., Mouriño-Carballido, B., Marañón, E., Chouciño, P., Gago, J., Ramírez, T., Vidal, M., Bode, A., Blasco, D., Royer, S.-J., Estrada, M., and Simó, R.: Importance of salt fingering for new nitrogen supply in the oligotrophic ocean, Nature Communications, 6, 8002, https://doi.org/10.1038/ncomms9002, 2015.

Filella, A., Riemann, L., Van Wambeke, F., Pulido-Villena, E., Vogts, A., Bonnet, S., Grosso, O., Diaz, J. M., Duhamel, S., and Benavides, M.: Contrasting Roles of DOP as a Source of Phosphorus and Energy for Marine Diazotrophs, Frontiers in Marine Science, 9, https://doi.org/10.3389/fmars.2022.923765, 2022.

Flett, R. J., Hamilton, R. D., and Campbell, N. E. R.: Aquatic acetylene-reduction techniques: solutions to several problems, Canadian ~~Journal~~journal of ~~Microbiology~~microbiology, 22 1, 43-51, https://doi.org/10.1139/m76-006, 1976.

Fonseca-Batista, D., Dehairs, F., Riou, V., Fripiat, F., Elskens, M., Deman, F., Brion, N., Quéroué, F., Bode, M., and Auel, H.: Nitrogen fixation in the eastern Atlantic reaches similar levels in the ~~southern~~Southern and ~~northern hemisphere~~Northern Hemisphere, Journal of Geophysical Research: Oceans, 122, 587-601, https://doi.org/10.1002/2016jc012335, 2017.

Fonseca-Batista, D., Li, X., Riou, V., Michotey, V., Deman, F., Fripiat, F., Guasco, S., Brion, N., Lemaitre, N., Tonnard, M., Gallinari, M., Planquette, H., Planchon, F., Sarthou, G., Elskens, M., Laroche, J., Chou, L., and Dehairs, F.: Evidence of high N_2 fixation rates in the temperate northeast Atlantic, Biogeosciences, 16, 999-1017, https://doi.org/10.5194/bg-16-999-2019, 2019.

Foster, R. A., Paytan, A., and Zehr, J.: Seasonality of N_2 fixation and *nifH* gene diversity in the Gulf of Aqaba (Red Sea), Limnol Oceanogr, 54, 219-233, https://doi.org/10.4319/lo.2009.54.1.0219, 2009.

Foster, R. A., Sztejrenszus, S., and Kuypers, M. M. M.: Measuring carbon and N_2 fixation in field populations of colonial and free-living unicellular cyanobacteria using nanometer-scale secondary ion mass spectrometry, Journal of Phycology, 49, 502-516, https://doi.org/10.1111/jpy.12057, 2013.

Foster, R. A., Kuypers, M. M. M., Vagner, T., Paerl, R. W., Musat, N., and Zehr, J. P.: Nitrogen fixation and transfer in open ocean diatom-cyanobacterial symbioses, ~~Isme Journal~~ISME J, 5, 1484-1493, https://doi.org/10.1038/ismej.2011.26, 2011.

Foster, R. A., Subramaniam, A., Mahaffey, C., Carpenter, E. J., Capone, D. G., and Zehr, J. P.: Influence of the Amazon River plume on distributions of free-living and symbiotic cyanobacteria in the western tropical north Atlantic Ocean, Limnol Oceanogr, 52, 517-532, https://doi.org/10.4319/lo.2007.52.2.0517, 2007.

Foster, R. A., Tienken, D., Littmann, S., Whitehouse, M. J., Kuypers, M. M. M., and White, A. E.: The rate and fate of N_2 and C fixation by marine diatom-diazotroph symbioses, ISME J, 16, 477-487, https://doi.org/10.1038/s41396-021-01086-7, 2022a.

Foster, R. A., Villareal, T. A., Lundin, D., Waterbury, J. B., Webb, E. A., and Zehr, J. P.: *Richelia*, in: Bergey's Manual of Systematics of Archaea and Bacteria, John Wiley & Sons, Inc., in association with Bergey's Manual Trust, 1-17, https://doi.org/10.1002/9781118960608.gbm01520, 2022b.

Gandhi, N., Singh, A., Prakash, S., Ramesh, R., Raman, M., Sheshshayee, M. S., and Shetye, S.: First direct measurements of N_2 fixation during a *Trichodesmium* bloom in the eastern Arabian Sea, Global Biogeochemical Cycles, 25, 1-10, https://doi.org/10.1029/2010gb003970, 2011.

Garcia, N., Raimbault, P., and Sandroni, V.: Seasonal nitrogen fixation and primary production in the Southwest Pacific: nanoplankton diazotrophy and transfer of nitrogen to picoplankton organisms, Marine Ecology Progress Series, 343, 25-33, https://doi.org/10.3354/meps06882, 2007.

Geisler, E., Bogler, A., Bar-Zeev, E., and Rahav, E.: Heterotrophic nitrogen fixation at the hyper-eutrophic qshon river and estuary system, Front Microbiol, 11, 1370, https://doi.org/10.3389/fmicb.2020.01370, 2020.

Giller, K. E., Nambiar, P. T. C., Srinivasa Rao, B., Dart, P. J., and Day, J. M.: A comparison of nitrogen fixation in genotypes of 420 groundnut (Arachis hypogaea L.) using ^{15}N-isotope dilution, Biology and Fertility of Soils, https://doi.org/10.1007/BF00264341, 1987.

Glibert, P. M. and Bronk, D. A.: Release of Dissolved Organic Nitrogen by Marine Diazotrophic Cyanobacteria, *Trichodesmium* spp, Applied and environmental microbiology, 60, 3996-4000, https://doi.org/10.1128/aem.60.11.3996-4000.1994, 1994.

Glover, D. M., Jenkins, W. J., and Doney, S. C.: Modeling methods for marine science, ~~https://doi.org/10.1017/cbo9780511975721~~, Cambridge University Press, Cambridge, UK,2011.

~~González, M. L., Molina, V., Florez-Leiva, L., Oriol, L., Cavagna, A. J., Dehairs, F., Farias, L., and Fernandez, C.: Nitrogen fixation in the Southern Ocean: a case of study of the Fe-fertilized Kerguelen region (KEOPS II cruise), https://doi.org/10.5194/bgd-11-17151-2014, 2014.~~

Gradoville, M., Cabello, A., Wilson, S., Turk-Kubo, K., Karl, D., and Zehr, J.: Light and depth dependency of nitrogen fixation by the non-photosynthetic, symbiotic cyanobacterium UCYN-A, Environmental microbiology, 23, https://doi.org/10.1111/1462-2920.15645, 2021.

Gradoville, M. R., Dugenne, M., Hynes, A. M., Zehr, J. P., and White, A. E.: Empirical relationship between *nifH* gene abundance and diazotroph cell concentration in the North Pacific Subtropical Gyre, ~~J Phycol, 58, 829-833~~Journal of Phycology, https://doi.org/10.1111/jpy.13289, 2022.

Gradoville, M. R., Bombar, D., Crump, B. C., Letelier, R. M., Zehr, J. P., and White, A. E.: Diversity and activity of nitrogen-fixing communities across ocean basins, Limnology and Oceanography, 62, 1895-1909, https://doi.org/10.1002/lno.10542, 2017.

Gradoville, M. R., Farnelid, H., White, A. E., Turk‐Kubo, K. A., Stewart, B., Ribalet, F., Ferrón, S., Pinedo‐Gonzalez, P., Armbrust, E. V., Karl, D. M., John, S., and Zehr, J. P.: Latitudinal constraints on the abundance and activity of the cyanobacterium UCYN-A and other marine diazotrophs in the North Pacific, Limnology and Oceanography, 65, 1858-1875, https://doi.org/10.1002/lno.11423, 2020.

Graham, J. A., Argyle, M., and Furnham, A.: The goal structure of situations, European Journal of Social Psychology, 10, 345-366, https://doi.org/10.1002/ejsp.2420100403, 1980.

Großkopf, T., Mohr, W., Baustian, T., Schunck, H., Gill, D., Kuypers, M. M. M., Lavik, G., Schmitz, R. A., Wallace, D. W. R., and LaRoche, J.: Doubling of marine dinitrogen-fixation rates based on direct measurements, Nature, 488, 361-364, https://doi.org/10.1038/nature11338, 2012.

Gruber, N.: The marine nitrogen cycle: overview and challenges, in: Nitrogen in the marine environment, 2nd edition, edited by: Capone, D. G., Bronk, D. A., Mulholland, M. R., and Carpenter, E. J., Elsevier, Amsterdam, 1-50, https://doi.org/10.1016/B978-0-12-372522-6.00001-3, 2008.

Gruber, N.: A diagnosis for marine nitrogen fixation, Nature, 566, 191-193, https://doi.org/10.1038/d41586-019-00498-y, 2019.

~~Gruber, N., Capone, D. G., Bronk, D. A., Mulholland, M. R., and Carpenter, E. J.: The marine nitrogen cycle: overview and challenges, Nitrogen in the Marine Environment, 2nd Edition, Elsevier Academic Press Inc, 525 B Street, Suite 1900, San Diego, Ca 92101-4495 USA, 1-50 pp., https://doi.org/10.1016/b978-0-12-372522-6.00001-3, 2008.~~

Hagino, K., Onuma, R., Kawachi, M., and Horiguchi, T.: Discovery of an Endosymbiotic Nitrogen-Fixing Cyanobacterium UCYN-A in Braarudosphaera bigelowii (Prymnesiophyceae), PLOS ONE, 8, e81749, 10.1371/journal.pone.0081749, 2013.

Hallstrøm, S., Benavides, M., Salamon, E. R., Arístegui, J., and Riemann, L.: Activity and distribution of diazotrophic communities across the Cape Verde Frontal Zone in the Northeast Atlantic Ocean, Biogeochemistry, https://doi.org/10.1007/s10533-022-00940-w, 2022.

Halm, H., Lam, P., Ferdelman, T. G., Lavik, G., Dittmar, T., LaRoche, J., D'Hondt, S., and Kuypers, M. M. M.: Heterotrophic organisms dominate nitrogen fixation in the South Pacific Gyre, ISME J, 6, 1238-1249, https://doi.org/10.1038/ismej.2011.182, 2012.

Hamersley, M. R., Turk, K. A., Leinweber, A., Gruber, N., Zehr, J. P., Gunderson, T., and Capone, D. G.: Nitrogen fixation within the water column associated with two hypoxic basins in the Southern California Bight, Aquatic Microbial Ecology, 63, 193-205, https://doi.org/10.3354/ame01494, 2011.

Harding, K., Turk-Kubo, K. A., Sipler, R. E., Mills, M. M., Bronk, D. A., and Zehr, J. P.: Symbiotic unicellular cyanobacteria fix nitrogen in the Arctic Ocean, P Natl Acad Sci USA, 115, 13371-13375, https://doi.org/10.1073/pnas.1813658115, 2018.

Harding, K. J., Turk-Kubo, K. A., Mak, E. W. K., Weber, P. K., Mayali, X., and Zehr, J. P.: Cell-specific measurements show nitrogen fixation by particle-attached putative non-cyanobacterial diazotrophs in the North Pacific Subtropical Gyre, Nat Commun, 13, 6979, https://doi.org/10.1038/s41467-022-34585-y, 2022.

Hardy, R. W. F., Burns, R. C., and Holsten, R. D.: Applications of the acetylene-ethylene assay for measurement of nitrogen fixation, Soil Biology and Biochemistry, 5, 47-81, https://doi.org/10.1016/0038-0717(73)90093-X, 1973.

Harrison, P., Zingone, A., Mickelson, M., Lehtinen, S., Nagappa, R., Kraberg, A., Sun, J., McQuatters-Gollop, A., and Jakobsen, H.: Cell volumes of marine phytoplankton from globally distributed coastal data sets, ~~Estuar Coast Mar Sci~~Estuarine, Coastal and Shelf Science, 162, https://doi.org/10.1016/j.ecss.2015.05.026, 2015.

Hashimoto, R., Watai, H., Miyahara, K., Sako, Y., and Yoshida, T.: Spatial and temporal variability of unicellular diazotrophic cyanobacteria in the eastern Seto Inland Sea, Fisheries Science, 82, 459-471, https://doi.org/10.1007/s12562-016-0983-y, 2016.

Hegde, S., Anil, A., Patil, J., Mitbavkar, S., Krishnamurthy, V., and Gopalakrishna, V.: Influence of environmental settings on the prevalence of *Trichodesmium* spp. in the Bay of Bengal, Mar. Ecol. Prog. Ser., 356, 93-101, https://doi.org/10.3354/meps07259, 2008.

Henke, B. A., Turk-Kubo, K. A., Bonnet, S., and Zehr, J. P.: Distributions and abundances of sublineages of the N_2-Fixing Cyanobacterium *Candidatus* Atelocyanobacterium thalassa (UCYN-A) in the New Caledonian Coral Lagoon, Front Microbiol, 9, https://doi.org/10.3389/fmicb.2018.00554, 2018.

Holl, C. M., Villareal, T. A., Payne, C. D., Clayton, T. D., Hart, C., and Montoya, J. P.: *Trichodesmium* in the western Gulf of Mexico: $^{15}N_2$-fixation and natural abundance stable isotopic evidence, Limnology and Oceanography, 52, 2249-2259, https://doi.org/10.4319/lo.2007.52.5.2249, 2007.

Hörstmann, C., Raes, E. J., Buttigieg, P. L., Lo Monaco, C., John, U., and Waite, A. M.: Hydrographic fronts shape productivity, nitrogen fixation, and microbial community composition in the southern Indian Ocean and the Southern Ocean, Biogeosciences, 18, 3733-3749, https://doi.org/10.5194/bg-18-3733-2021, 2021.

Hyman, M. R. and Arp, D. J.: Quantification and removal of some contaminating gases from acetylene used to study gas-utilizing enzymes and microorganisms, Applied and environmental microbiology, 53, 298-303, https://doi.org/10.1128/aem.53.2.298-303.1987, 1987.

Ibello, V., Cantoni, C., Cozzi, S., and Civitarese, G.: First basin-wide experimental results on N_2 fixation in the open Mediterranean Sea, ~~Geophysical Research Letters, 37, n/a-n/a~~Geophys. Res. Lett., 37, https://doi.org/10.1029/2009gl041635, 2010.

Jayakumar, A., Chang, B. X., Widner, B., Bernhardt, P., Mulholland, M. R., and Ward, B. B.: Biological nitrogen fixation in the oxygen-minimum region of the eastern tropical North Pacific ocean, The ISME Journal, 11, 2356-2367, https://doi.org/10.1038/ismej.2017.97, 2017.

Jiang, Z., Chen, J., Zhou, F., Zhai, H., Zhang, D., and Yan, X.: Summer distribution patterns of Trichodesmium spp. in the Changjiang (Yangtze River) Estuary and adjacent East China Sea shelf, Oceanologia, 59, 248-261, https://doi.org/10.1016/j.oceano.2017.02.001, 2017.

Jiang, Z., Zhu, Y., Sun, Z., Zhai, H., Zhou, F., Yan, X., Zeng, J., Chen, J., and Chen, Q.: Enhancement of Summer Nitrogen Fixation by the Kuroshio Intrusion in the East China Sea and Southern Yellow Sea, Journal of Geophysical Research: Biogeosciences, 128, e2022JG007287, https://doi.org/10.1029/2022JG007287, 2023.

Karlusich, J. J. P., Pelletier, E., Lombard, F., Carsique, M., Dvorak, E., Colin, S., Picheral, M., Cornejo-Castillo, F. M., Acinas, S. G., Pepperkok, R., Karsenti, E., De Vargas, C., Wincker, P., Bowler, C., and Foster, R. A.: Global distribution patterns of marine nitrogen-fixers by imaging and molecular methods, Nat. Commun., 12, https://doi.org/10.1038/s41467-021-24299-y, 2021.

Kitajima, S., Furuya, K., Hashihama, F., Takeda, S., and Kanda, J.: Latitudinal distribution of diazotrophs and their nitrogen fixation in the tropical and subtropical western North Pacific, Limnology and Oceanography, 54, 537-547, https://doi.org/10.4319/lo.2009.54.2.0537, 2009.

Kittu, L. R., Paul, A. J., Fernández - Méndez, M., Hopwood, M. J., and Riebesell, U.: Coastal N_2 Fixation Rates Coincide Spatially With Nitrogen Loss in the Humboldt Upwelling System off Peru, Global Biogeochemical Cycles, 37, https://doi.org/10.1029/2022gb007578, 2023.

Klawonn, I., Lavik, G., Boning, P., Marchant, H. K., Dekaezemacker, J., Mohr, W., and Ploug, H.: Simple approach for the preparation of $^{15-15}N_2$-enriched water for nitrogen fixation assessments: evaluation, application and recommendations, Front Microbiol, 6, 769, https://doi.org/10.3389/fmicb.2015.00769, 2015.

Knapp, A. N., Casciotti, K. L., Berelson, W. M., Prokopenko, M. G., and Capone, D. G.: Low rates of nitrogen fixation in eastern tropical South Pacific surface waters, Proc Natl Acad Sci U S A, 113, 4398-4403, https://doi.org/10.1073/pnas.1515641113, 2016.

Konno, U., Tsunogai, U., Komatsu, D. D., Daita, S., Nakagawa, F., Tsuda, A., Matsui, T., Eum, Y.-J., and Suzuki, K.: Determination of total N_2 fixation rates in the ocean taking into account both the particulate and filtrate fractions, Biogeosciences, 7, 2369-2377, https://doi.org/10.5194/bg-7-2369-2010, 2010.

Kromkamp, J., De Bie, M., Goosen, N., Peene, J., Van Rijswijk, P., Sinke, J., and Duinevel, G. C. A.: Primary production by phytoplankton along the Kenyan coast during the SE monsoon and November intermonsoon 1992, and the occurrence of *Trichodesmium*, Deep Sea Research Part II: Topical Studies in Oceanography, 44, 1195-1212, https://doi.org/10.1016/s0967-0645(97)00015-5, 1997.

Krupke, A., Lavik, G., Halm, H., Fuchs, B. M., Amann, R. I., and Kuypers, M. M. M.: Distribution of a consortium between unicellular algae and the N_2 fixing cyanobacterium UCYN-A in the North Atlantic Ocean, Environmental Microbiology, 16, 3153-3167, https://doi.org/10.1111/1462-2920.12431, 2014.

Krupke, A., Mohr, W., Laroche, J., Fuchs, B. M., Amann, R. I., and Kuypers, M. M.: The effect of nutrients on carbon and nitrogen fixation by the UCYN-A–haptophyte symbiosis, The ISME Journal, 9, 1635-1647, https://doi.org/10.1038/ismej.2014.253, 2015.

Krupke, A., Musat, N., LaRoche, J., Mohr, W., Fuchs, B. M., Amann, R. I., Kuypers, M. M. M., and Foster, R. A.: In situ identification and N_2 and C fixation rates of uncultivated cyanobacteria populations, Systematic and Applied Microbiology, 36, 259-271, https://doi.org/10.1016/j.syapm.2013.02.002, 2013.

Kumar, P. K., Singh, A., Ramesh, R., and Nallathambi, T.: N_2 Fixation in the eastern Arabian Sea: probable role of heterotrophic diazotrophs, ~~Front. Mar. Sci.,~~Frontiers in Marine Science, 4, https://doi.org/10.3389/fmars.2017.00080, 2017.

Kumari, V. R., Ghosh, V. R. D., Rao, D. N., Krishna, M. S., and Sarma, V. V. S. S.: Nitrogen fixation in the western coastal Bay of Bengal: Controlling factors and contribution to primary production, Regional Studies in Marine Science, 53, 102410, https://doi.org/10.1016/j.rsma.2022.102410, 2022.

~~Lee Chen, Y.~~Landou, E., Lazar, B., LaRoche, J., Fennel, K., and Berman‐Frank, I.: Contribution of photic and aphotic N_2 fixation to production in an oligotrophic sea, Limnology and Oceanography, 68, 692-708, https://doi.org/10.1002/lno.12303, 2023.

Langlois, R., Grokopf, T., Mills, M., Takeda, S., and LaRoche, J.: Widespread distribution and expression of Gamma A (UMB), an uncultured, diazotrophic, gamma-proteobacterial *nifH* phylotype, Plos One, 10, 17, https://doi.org/10.1371/journal.pone.0128912, 2015.

Le Moal, M. and Biegala, I. C.: Diazotrophic unicellular cyanobacteria in the northwestern Mediterranean Sea: A seasonal cycle, Limnol Oceanogr, 54, 845-855, https://doi.org/10.4319/lo.2009.54.3.0845, 2009.

Le Moal, M., Collin, H., and Biegala, I. C.: Intriguing diversity among diazotrophic picoplankton along a Mediterranean transect: a dominance of rhizobia, Biogeosciences, 8, 827-840, https://doi.org/10.5194/bg-8-827-2011, 2011.

Letelier, R. and Karl, D.: Role of Trichodesmium spp. in the productivity of the subtropical North Pacific Ocean, Mar. Ecol. Prog. Ser., 133, 263-273, https://doi.org/10.3354/meps133263, 1996.

~~L., Chen, H.-Y., Lin, Y.-H., Yong, T.-C., Taniuchi, Y., and Tuo, S.-H.: The relative contributions of unicellular and filamentous diazotrophs to N_2 fixation in the South China Sea and the upstream Kuroshio, Deep Sea Research Part I: Oceanographic Research Papers, 85, 56-71, https://doi.org/10.1016/j.dsr.2013.11.006, 2014.~~

Li, L., Wu, C., Sun, J., Song, S., Ding, C., Huang, D., and Pujari, L.: Nitrogen fixation driven by mesoscale eddies and the Kuroshio Current in the northern South China Sea and the East China Sea, Acta Oceanol. Sin., 39, 30-41, https://doi.org/10.1007/s13131-020-1691-0, 2020.

Liu, J. X., Zhou, L. B., Li, J. J., Lin, Y. Y., Ke, Z. X., Zhao, C. Y., Liu, H. J., Jiang, X., He, Y. H., and Tan, Y. H.: Effect of mesoscale eddies on diazotroph community structure and nitrogen fixation rates in the South China Sea, Regional Studies in Marine Science, 35, 14, https://doi.org/10.1016/j.rsma.2020.101106, 2020.

Loescher, C. R., Großkopf, T., Desai, F. D., Gill, D., Schunck, H., Croot, P. L., Schlosser, C., Neulinger, S. C., Pinnow, N., Lavik, G., Kuypers, M. M. M., LaRoche, J., and Schmitz, R. A.: Facets of diazotrophy in the oxygen minimum zone waters off Peru, The Isme Journal, 8, 2180-2192, https://doi.org/10.1038/ismej.2014.71, 2014.

Loick-Wilde, N., Weber, S. C., Conroy, B. J., Capone, D. G., Coles, V. J., Medeiros, P. M., Steinberg, D. K., and Montoya, J. P.: Nitrogen sources and net growth efficiency of zooplankton in three Amazon River plume food webs, Limnology and Oceanography, 61, 460-481, https://doi.org/10.1002/lno.10227, 2015.

Loick-Wilde, N., Fernandez-Urruzola, I., Eglite, E., Liskow, I., Nausch, M., Schulz-Bull, D., Wodarg, D., Wasmund, N., and Mohrholz, V.: Stratification, nitrogen fixation, and cyanobacterial bloom stage regulate the planktonic food web structure, Glob Chang Biol, 25, 794-810, https://doi.org/10.1111/gcb.14546, 2019.

Lory, C., Van Wambeke, F., Fourquez, M., Barani, A., Guieu, C., Tilliette, C., Marie, D., Nunige, S., Berman-Frank, I., and Bonnet, S.: Assessing the contribution of diazotrophs to microbial Fe uptake using a group specific approach in the Western Tropical South Pacific Ocean, ISME Communications, 2, 41, https://doi.org/10.1038/s43705-022-00122-7, 2022.

Löscher, C. R., Mohr, W., Bange, H. W., and Canfield, D. E.: No nitrogen fixation in the Bay of Bengal?, Biogeosciences, 17, 851-864, https://doi.org/10.5194/bg-17-851-2020, 2020.

Löscher, C. R., Bourbonnais, A., Dekaezemacker, J., Charoenpong, C. N., Altabet, M. A., Bange, H. W., Czeschel, R., Hoffmann, C., and Schmitz, R.: N_2 fixation in eddies of the eastern tropical South Pacific Ocean, Biogeosciences, 13, 2889-2899, https://doi.org/10.5194/bg-13-2889-2016, 2016.

Lu, Y., Wen, Z., Shi, D., Chen, M., Zhang, Y., Bonnet, S., Li, Y., Tian, J., and Kao, S. J.: Effect of light on N_2 fixation and net nitrogen release of *Trichodesmium* in a field study, Biogeosciences, 15, 1-12, https://doi.org/10.5194/bg-15-1-2018, 2018.

Luo, Y. W., Doney, S. C., Anderson, L. A., Benavides, M., Berman-Frank, I., Bode, A., Bonnet, S., Bostrom, K. H., Bottjer, D., Capone, D. G., Carpenter, E. J., Chen, Y. L., Church, M. J., Dore, J. E., Falcon, L. I., Fernandez, A., Foster, R. A., Furuya, K., Gomez, F., Gundersen, K., Hynes, A. M., Karl, D. M., Kitajima, S., Langlois, R. J., LaRoche, J., Letelier, R. M., Maranon, E., McGillicuddy, D. J., Moisander, P. H., Moore, C. M., Mourino-Carballido, B., Mulholland, M. R., Needoba, J. A., Orcutt, K. M., Poulton, A. J., Rahav, E., Raimbault, P., Rees, A. P., Riemann, L., Shiozaki, T., Subramaniam, A., Tyrrell, T., Turk-Kubo, K. A., Varela, M., Villareal, T. A., Webb, E. A., White, A. E., Wu, J., and Zehr, J. P.: Database of diazotrophs in global ocean: abundance, biomass and nitrogen fixation rates, Earth Syst. Sci. Data, 4, 47-73, https://doi.org/10.5194/essd-4-47-2012, 2012.

Mague, T. H., Weare, N. M., and Holm-Hansen, O.: Nitrogen fixation in the North Pacific Ocean, Mar Biol, 24, 109-119, https://doi.org/10.1007/bf00389344, 1974.

Martínez‐Pérez, C., Mohr, W., Loscher, C. R., Dekaezemacker, J., Littmann, S., Yilmaz, P., Lehnen, N., Fuchs, B. M., Lavik, G., Schmitz, R. A., LaRoche, J., and Kuypers, M. M.: The small unicellular diazotrophic symbiont, UCYN-A, is a key player in the marine nitrogen cycle, Nat Microbiol, 1, 16163, https://doi.org/10.1038/nmicrobiol.2016.163, 2016.

Masotti, I., Ruiz-Pino, D., and Le Bouteiller, A.: Photosynthetic characteristics of *Trichodesmium* in the southwest Pacific Ocean: importance and significance, Mar. Ecol. Prog. Ser., 338, 47-59, https://doi.org/10.3354/meps338047, 2007.

McCarthy, J. J. and Carpenter, E. J.: Oscillatoria (*Trichodesmium*) Thiebautii (cyanophyta) the central North Atlantic Ocean, Journal of Phycology, 15, 75-82, https://doi.org/10.1111/j.1529-8817.1979.tb02965.x, 1979.

Meiler, S., Britten, G. L., Dutkiewicz, S., Moisander, P. H., and Follows, M. J.: Challenges and opportunities in connecting gene count observations with ocean biogeochemical models: Reply to Zehr and Riemann (2023), Limnol Oceanogr, 68, 1413-1416, https://doi.org/10.1002/lno.12363, 2023.

Meiler, S., Britten, G. L., Dutkiewicz, S., Gradoville, M. R., Moisander, P. H., Jahn, O., and Follows, M. J.: Constraining uncertainties of diazotroph biogeography from *nifH* gene abundance, Limnology and Oceanography, 67, 816-829, https://doi.org/10.1002/lno.12036, 2022.

Menden-Deuer, S. and Lessard, E. J.: Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton, ~~Limnol Oceanogr~~Limnology and Oceanography, 45, 569-579, https://doi.org/10.4319/lo.2000.45.3.0569, 2000.

Messer, L. F., Brown, M. V., Van Ruth, P. D., Doubell, M., and Seymour, J. R.: Temperate southern Australian coastal waters are characterised by surprisingly high rates of nitrogen fixation and diversity of diazotrophs, PeerJ, 9, e10809, https://doi.org/10.7717/peerj.10809, 2021.

Messer, L. F., Mahaffey, C., M Robinson, C., Jeffries, T. C., Baker, K. G., Bibiloni Isaksson, J., Ostrowski, M., Doblin, M. A., Brown, M. V., and Seymour, J. R.: High levels of heterogeneity in diazotroph diversity and activity within a putative hotspot for marine nitrogen fixation, The ISME Journal, 10, 1499-1513, https://doi.org/10.1038/ismej.2015.205, 2016.

Meyer, N. R., Fortney, J. L., and Dekas, A. E.: NanoSIMS sample preparation decreases isotope enrichment: magnitude, variability and implications for single‐cell rates of microbial activity, Environ. Microbiol., 23, 81-98, https://doi.org/10.1111/1462-2920.15264, 2021.

Mills, M. M., Turk-Kubo, K. A., van Dijken, G. L., Henke, B. A., Harding, K., Wilson, S. T., Arrigo, K. R., and Zehr, J. P.: Unusual marine cyanobacteria/haptophyte symbiosis relies on N_2 fixation even in N-rich environments, The ISME Journal, 14, 2395-2406, https://doi.org/10.1038/s41396-020-0691-6, 2020.

Mohr, W., Grosskopf, T., Wallace, D. W., and LaRoche, J.: Methodological underestimation of oceanic nitrogen fixation rates, PLoS One, 5, e12583, https://doi.org/10.1371/journal.pone.0012583, 2010.

Moisander, P. H., Serros, T., Paerl, R. W., Beinart, R. A., and Zehr, J. P.: Gammaproteobacterial diazotrophs and *nifH* gene expression in surface waters of the South Pacific Ocean, ISME J, 8, 1962-1973, https://doi.org/10.1038/ismej.2014.49, 2014.

Moisander, P. H., Benavides, M., Bonnet, S., Berman-Frank, I., White, A. E., and Riemann, L.: Chasing after non-cyanobacterial nitrogen fixation in marine pelagic environments, ~~Frontiers in Microbiology~~Front Microbiol, 8, 1736, https://doi.org/10.3389/fmicb.2017.01736, 2017.

Moisander, P. H., Zhang, R., Boyle, E. A., Hewson, I., Montoya, J. P., and Zehr, J. P.: Analogous nutrient limitations in unicellular diazotrophs and *Prochlorococcus* in the South Pacific Ocean, The Isme Journal, 6, 733-744, https://doi.org/10.1038/ismej.2011.152, 2012.

Mompeán, C., Bode, A., Benítez-Barrios, V. M., Domínguez-Yanes, J. F., Escánez, J., and Fraile-Nuez, E.: Spatial patterns of plankton biomass and stable isotopes reflect the influence of the nitrogen-fixer Trichodesmium along the subtropical North Atlantic, ~~J.~~Journal of Plankton ~~Res.~~,Research, 35, 513-525, ~~https://doi.org/10.1093/plankt/fbt011~~,10.1093/plankt/fbt011, 2013.

Mompeán, C., Bode, A., Latasa, M., Fernández-Castro, B., Mouriño-Carballido, B., and Irigoien, X.: The influence of nitrogen inputs on biomass and trophic structure of ocean plankton: a study using biomass and stable isotope size-spectra, Journal of Plankton Research, 38, 1163-1177, 10.1093/plankt/fbw052, 2016.

Montoya, J. P., Voss, M., Kahler, P., and Capone, D. G.~~: A simple, high-precision, high-sensitivity tracer assay~~; A Simple, High-Precision, High-Sensitivity Tracer Assay for N_2 ~~fixation, Applied and Environmental Microbiology~~Fixation, Appl Environ Microbiol, 62, 986-993, https://doi.org/10.1128/aem.62.3.986-993.1996, 1996.

Moore, R. M., Grefe, I., Zorz, J., Shan, S., Thompson, K., Ratten, J., and LaRoche, J.: On the relationship between hydrogen saturation in the tropical Atlantic Ocean and nitrogen fixation by the symbiotic diazotroph UCYN-A, J Geophys Res-Oceans, 123, 2353-2362, https://doi.org/10.1002/2017jc013047, 2018.

Moreira-Coello, V., Mourino-Carballido, B., Maranon, E., Fernandez-Carrera, A., Bode, A., and Varela, M. M.: Biological N_2 fixation in the upwelling region off NW Iberia: magnitude, relevance, and players, Front. Mar. Sci., 4, 303, https://doi.org/10.3389/fmars.2017.00303, 2017.

Mouriño-Carballido, B., Graña, R., Fernàndez, A., Bode, A., Varela, M., Domínguez, J. F., Escànez, J., De Armas, D., and Marañón, E.: Importance of N_2 fixation vs. nitrate eddy diffusion along a latitudinal transect in the Atlantic Ocean, Limnol Oceanogr, 56, 999-1007, https://doi.org/10.4319/lo.2011.56.3.0999, 2011.

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

Mulholland, M. R., Bernhardt, P. W., Heil, C. A., Bronk, D. A., and O'Neil, J. M.: Nitrogen fixation and release of fixed nitrogen by *Trichodesmium* spp. in the Gulf of Mexico, Limnology and Oceanography, 51, 1762-1776, https://doi.org/10.4319/lo.2006.51.4.1762, 2006.

Mulholland, M. R., Bernhardt, P. W., Widner, B. N., Selden, C. R., Chappell, P. D., Clayton, S., Mannino, A., and Hyde, K.: High rates of N_2 fixation in temperate, western North Atlantic coastal waters expand the realm of marine diazotrophy, Global Biogeochemical Cycles, 33, 826-840, https://doi.org/10.1029/2018gb006130, 2019.

Mulholland, M. R., Bernhardt, P. W., Blanco-Garcia, J. L., Mannino, A., Hyde, K., Mondragon, E., Turk, K., Moisander, P. H., and Zehr, J. P.: Rates of dinitrogen fixation and the abundance of diazotrophs in North American coastal waters between Cape Hatteras and Georges Bank, Limnology and Oceanography, 57, 1067-1083, https://doi.org/10.4319/lo.2012.57.4.1067, 2012.

Musat, N., Stryhanyuk, H., Bombach, P., Adrian, L., Audinot, J.-N., and Richnow, H. H.: The effect of FISH and CARD-FISH on the isotopic composition of 13C- and 15N-labeled Pseudomonas putida cells measured by nanoSIMS, Systematic and Applied Microbiology, 37, 267-276, https://doi.org/10.1016/j.syapm.2014.02.002, 2014.

Needoba, J. A., Foster, R. A., Sakamoto, C., Zehr, J. P., and Johnson, K. S.: Nitrogen fixation by unicellular diazotrophic cyanobacteria in the temperate oligotrophic North Pacific Ocean, Limnology and OceanographyLimnol Oceanogr, 52, 1317-1327, https://doi.org/10.4319/lo.2007.52.4.1317, 2007.

Pierella Karlusich, J. JPalter, J. B., Ames, E. J., Benavides, M., Goncalves Neto, A., Granger, J., Moisander, P. H., Watkins‐Brandt, K. S., and White, A. E.: High N_2 fixation in and near the Gulf Stream consistent with a circulation control on diazotrophy, Geophys. Res. Lett., 47, e2020GL089103, https://doi.org/10.1111/j.1365-2656.2010.01695.x, 2020.

Postgate, J. R.: The Fundamentals of Nitrogen Fixation, Cambridge University Press, New York1998.

, Pelletier, E., Lombard, F., Carsique, M., Dvorak, E., Colin, S., Picheral, M., Cornejo-Castillo, F. M., Acinas, S. G., Pepperkok, R., Karsenti, E., De Vargas, C., Wincker, P., Bowler, C., and Foster, R. A.: Global distribution patterns of marine nitrogen-fixers by imaging and molecular methods, Nature Communications, 12, https://doi.org/10.1038/s41467-021-24299-y, 2021.

Raes, E., van de Kamp, J., Bodrossy, L., Fong, A., Riekenberg, J., Holmes, B., Erler, D., Eyre, B., Weil, S.-S., and Waite, A.: N_2 fixation and new insights into nitrification from the ice-edge to the equator in the South Pacific Ocean, Front. Mar. Sci.,Frontiers in Marine Science, 7, https://doi.org/10.3389/fmars.2020.00389, 2020.

Raes, E. J., Waite, A. M., McInnes, A. S., Olsen, H., Nguyen, H. M., Hardman-Mountford, N., and Thompson, P. A.: Changes in latitude and dominant diazotrophic community alter N_2 fixation, Marine Ecology Progress Series, 516, 85-102, 10.3354/meps11009, 2014.

Rahav, E., Giannetto, M. J., and Bar-Zeev, E.: Contribution of mono and polysaccharides to heterotrophic N_2 fixation at the eastern Mediterranean coastline, Sci Rep-UkScientific Reports, 6, 27858, https://doi.org/10.1038/srep27858, 2016.

Rahav, E., Herut, B., Levi, A., Mulholland, M. R., and Berman-Frank, I.: Springtime contribution of dinitrogen fixation to primary production across the Mediterranean Sea, Ocean Science, 9, 489-498, https://doi.org/10.5194/os-9-489-2013, 2013a.

Rahav, E., Herut, B., Mulholland, M., Belkin, N., Elifantz, H., and Berman-Frank, I.: Heterotrophic and autotrophic contribution to dinitrogen fixation in the Gulf of Aqaba, Marine Ecology Progress Series, 522, 67-77, https://doi.org/10.3354/meps11143, 2015.

Rahav, E., Bar-Zeev, E., Ohayon, S., Elifantz, H., Belkin, N., Herut, B., Mulholland, M. R., and Berman-Frank, I.: Dinitrogen fixation in aphotic oxygenated marine environments, ~~Frontiers in Microbiology, 4, 11,~~ https://doi.org/10.3389/fmicb.2013.00227, 2013b.

~~Rahav, E., Herut, B., Mulholland, M. R., Voß, B., Stazic, D., Steglich, C., Hess, W. R., and Berman-Frank, I.: Contribution of dinitrogen fixation to bacterial and primary productivity in the Gulf of Aqaba (Red Sea), https://doi.org/10.5194/bgd-10-10327-2013, 2013c.~~

Ratten, J.-M., LaRoche, J., Desai, D. K., Shelley, R. U., Landing, W. M., Boyle, E., Cutter, G. A., and Langlois, R. J.: Sources of iron and phosphate affect the distribution of diazotrophs in the North Atlantic, Deep Sea Research Part II: Topical Studies in Oceanography, 116, 332-341, https://doi.org/10.1016/j.dsr2.2014.11.012, 2015.

Reeder, C. F., Stoltenberg, I., Javidpour, J., and Löscher, C. R.: Salinity as a key control on the diazotrophic community composition in the southern Baltic Sea, Ocean Science, 18, 401-417, https://doi.org/10.5194/os-18-401-2022, 2022.

Riou, V., Fonseca-Batista, D., Roukaerts, A., Biegala, I. C., Prakya, S. R., Magalhães Loureiro, C., Santos, M., Muniz-Piniella, A. E., Schmiing, M., Elskens, M., Brion, N., Martins, M. A., and Dehairs, F.: Importance of N_2-fixation on the productivity at the North-Western Azores Current/Front System, and the abundance of diazotrophic unicellular cyanobacteria, PLoS One, 11, e0150827, https://doi.org/10.1371/journal.pone.0150827, 2016.

Sahoo, D., Saxena, H., Nazirahmed, S., Kumar, S., Sudheer, A. K., Bhushan, R., Sahay, A., and Singh, A.: Role of eddies and N_2 fixation in regulating C:N:P proportions in the Bay of Bengal, Biogeochemistry, 155, 413-429, https://doi.org/10.1007/s10533-021-00833-4, 2021.

Sahu, B. K., Baliarsingh, S. K., Lotliker, A. A., Parida, C., Srichandan, S., and Sahu, K. C.: Winter thermal inversion and *Trichodesmium* dominance in north-western Bay of Bengal, Ocean Science Journal, 52, 301-306, https://doi.org/10.1007/s12601-017-0028-1, 2017.

Sargent, E. C., Hitchcock, A., Johansson, S. A., Langlois, R., Moore, C. M., LaRoche, J., Poulton, A. J., and Bibby, T. S.: Evidence for polyploidy in the globally important diazotroph *Trichodesmium*, FEMS microbiology letters, 363, fnw244, https://doi.org/10.1093/femsle/fnw244, 2016.

Sarma, V. V. S. S., Vivek, R., Rao, D. N., and Ghosh, V. R. D.: Severe phosphate limitation on nitrogen fixation in the Bay of Bengal, Continental Shelf Research, 205, 104199, https://doi.org/10.1016/j.csr.2020.104199, 2020.

Sato, T., Shiozaki, T., Taniuchi, Y., Kasai, H., and Takahashi, K.: Nitrogen fixation and diazotroph community in the subarctic Sea of Japan and Sea of Okhotsk, J Geophys Res-Oceans, 126, https://doi.org/10.1029/2020jc017071, 2021.

Sato, T., Shiozaki, T., Hashihama, F., Sato, M., Murata, A., Sasaoka, K., Umeda, S.-i., and Takahashi, K.: Low Nitrogen Fixation Related to Shallow Nitracline Across the Eastern Indian Ocean, Journal of Geophysical Research: Biogeosciences, 127, e2022JG007104, https://doi.org/10.1029/2022JG007104, 2022.

Saulia, E., Benavides, M., Henke, B., Turk-Kubo, K., Cooperguard, H., Grosso, O., Desnues, A., Rodier, M., Dupouy, C., Riemann, L., and Bonnet, S.: Seasonal Shifts in Diazotrophs Players: Patterns Observed Over a Two-Year Time Series in the New Caledonian Lagoon (Western Tropical South Pacific Ocean), Frontiers in Marine Science, 7, https://doi.org/10.3389/fmars.2020.581755, 2020.

Saxena, H., Sahoo, D., Khan, M. A., Kumar, S., Sudheer, A. K., and Singh, A.: Dinitrogen fixation rates in the Bay of Bengal during summer monsoon, Environmental Research Communications, 2, 051007, https://doi.org/10.1088/2515-7620/ab89fa, 2020.

Scavotto, R. E., Dziallas, C., Bentzon-Tilia, M., Riemann, L., and Moisander, P. H.: Nitrogen-fixing bacteria associated with copepods in coastal waters of the North Atlantic Ocean, Environmental Microbiology, 17, 3754-3765, https://doi.org/10.1111/1462-2920.12777, 2015.

Schvarcz, C. R., Wilson, S. T., Caffin, M., Stancheva, R., Li, Q., Turk-Kubo, K. A., White, A. E., Karl, D. M., Zehr, J. P., and Steward, G. F.: Overlooked and widespread pennate diatom-diazotroph symbioses in the sea, Nat. Commun., 13, 799, https://doi.org/10.1038/s41467-022-28065-6, 2022.

Selden, C. R., Mulholland, M. R., Widner, B., Bernhardt, P~~. W., Widner, B., Macías Tapia, A., Ji, Q., ,~~ and Jayakumar, A.~~: Dinitrogen~~ : Toward resolving disparate accounts of the extent and magnitude of nitrogen fixation ~~across physico-chemical gradients of the~~ in the Eastern Tropical ~~North~~South Pacific oxygen deficient zone, Limnology and Oceanography, 66, 1950-1960, https://doi.org/10.1002/lno.11735, 2021a.

Selden, C. R., Chappell, P. D., Clayton, S., Macías-Tapia, A., Bernhardt, P. W., and Mulholland, M. R.: A coastal N_2 fixation hotspot at the Cape Hatteras front: Elucidating spatial heterogeneity in diazotroph activity via supervised machine learning, Limnology and Oceanography, 66, 1832-1849, https://doi.org/10.1002/lno.11727, 2021b.

Selden, C. R., Mulholland, M. R., Bernhardt, P. W., Widner, B., Macías‐Tapia, A., Ji, Q., and Jayakumar, A.: Dinitrogen Fixation Across Physico‐Chemical Gradients of the Eastern Tropical North Pacific Oxygen Deficient Zone, Global Biogeochemical Cycles, 33, 1187-1202, ~~https://doi.org/10.1029/2019GB006242~~https://doi.org/10.1029/2019gb006242, 2019.

Selden, C. R., Einarsson, S. V., Lowry, K. E., Crider, K. E., Pickart, R. S., Lin, P., Ashjian, C. J., and Chappell, P. D.: Coastal upwelling enhances abundance of a symbiotic diazotroph (UCYN-A) and its haptophyte host in the Arctic Ocean, Front. Mar. Sci., 9, https://doi.org/10.3389/fmars.2022.877562, 2022.

Shao, Z. and Luo, Y.-W.: Controlling factors on the global distribution of a representative marine non-cyanobacterial diazotroph phylotype (Gamma A), Biogeosciences, 19, 2939-2952, https://doi.org/10.5194/bg-19-2939-2022, 2022.

Shao, Z., Xu, Y., Wang, H., Luo, W., Wang, L., Huang, Y., and Luo, Y.-W.: Version 2 of the global oceanic diazotroph database, Figshare [dataset], https://doi.org/10.6084/m9.figshare.21677687, 2022.

Shiozaki, T., Kodama, T., and Furuya, K.: Large-scale impact of the island mass effect through nitrogen fixation in the western South Pacific Ocean, Geophysical Research Letters, 41, 2907-2913, https://doi.org/10.4319/lo.2007.52.4.131710.1002/2014gl059835, 2014a.

Shiozaki, T., Kondo, Y., Yuasa, D., and Takeda, S.: Distribution of major diazotrophs in the surface water of the Kuroshio from northeastern Taiwan to south of mainland Japan, J. Plankton Res., 40, 407-419, https://doi.org/10.1093/plankt/fby027, 2018a.

Shiozaki, T., Nagata, T., Ijichi, M., and Furuya, K.: Nitrogen fixation and the diazotroph community in the temperate coastal region of the northwestern North Pacific, Biogeosciences, 12, 4751-4764, https://doi.org/10.5194/bg-12-4751-2015, 2015a.

Shiozaki, T., Ijichi, M., Kodama, T., Takeda, S., and Furuya, K.: Heterotrophic bacteria as major nitrogen fixers in the euphotic zone of the Indian Ocean, Global Biogeochemical Cycles, 28, 1096-1110, https://doi.org/10.1002/2014gb004886, ~~2014a~~2014b.

Shiozaki, T., Kodama, T., Kitajima, S., Sato, M., and Furuya, K.: Advective transport of diazotrophs and importance of their nitrogen fixation on new and primary production in the western Pacific warm pool, Limnol Oceanogr, 58, 49-60, https://doi.org/10.4319/lo.2013.58.1.0049, 2013.

Shiozaki, T., Fujiwara, A., Inomura, K., Hirose, Y., Hashihama, F., and Harada, N.: Biological nitrogen fixation detected under Antarctic sea ice, ~~Nat Geosci~~Nature Geoscience, 13, 729-+, https://doi.org/10.1038/s41561-020-00651-7, 2020.

Shiozaki, T., Chen, Y. L. L., Lin, Y. H., Taniuchi, Y., Sheu, D. S., Furuya, K., and Chen, H. Y.: Seasonal variations of unicellular diazotroph groups A and B, and T*richodesmium* in the northern South China Sea and neighboring upstream Kuroshio Current, Continental Shelf Research, 80, 20-31, https://doi.org/10.1016/j.csr.2014.02.015, ~~2014b~~2014c.

Shiozaki, T., Takeda, S., Itoh, S., Kodama, T., Liu, X., Hashihama, F., and Furuya, K.: Why is *Trichodesmium* abundant in the Kuroshio?, Biogeosciences, 12, 6931-6943, https://doi.org/10.4319/lo.2007.52.4.131710.5194/bg-12-6931-2015, 2015b.

Shiozaki, T., Fujiwara, A., Ijichi, M., Harada, N., Nishino, S., Nishi, S., Nagata, T., and Hamasaki, K.: Diazotroph community structure and the role of nitrogen fixation in the nitrogen cycle in the Chukchi Sea (western Arctic Ocean), Limnology and Oceanography, 63, 2191-2205, https://doi.org/10.1002/lno.10933, 2018b.

Shiozaki, T., Bombar, D., Riemann, L., Hashihama, F., Takeda, S., Yamaguchi, T., Ehama, M., Hamasaki, K., and Furuya, K.: Basin scale variability of active diazotrophs and nitrogen fixation in the North Pacific, from the tropics to the subarctic Bering Sea, Global Biogeochemical Cycles, 31, 996-1009, https://doi.org/10.1002/2017gb005681, 2017.

Shiozaki, T., Bombar, D., Riemann, L., Sato, M., Hashihama, F., Kodama, T., Tanita, I., Takeda, S., Saito, H., Hamasaki, K., and Furuya, K.: Linkage ~~between dinitrogen fixation~~Between Dinitrogen Fixation and ~~primary production~~Primary Production in the ~~oligotrophic~~Oligotrophic South Pacific Ocean, Global Biogeochemical Cycles, 32, 1028-1044, https://doi.org/10.1029/~~2017gb005869~~2017GB005869, 2018c.

Short, S. M. and Zehr, J. P.: Quantitative Analysis of *nifH* Genes and Transcripts from Aquatic Environments, in: Methods in Enzymology, Academic Press, 380-394, https://doi.org/10.1016/S0076-6879(05)97023-7, 2005.

Singh, A., Gandhi, N., and Ramesh, R.: 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, https://doi.org/10.1016/j.csr.2019.05.012, 2019.

Singh, A., Bach, L. T., Fischer, T., Hauss, H., Kiko, R., Paul, A. J., Stange, P., Vandromme, P., and Riebesell, U.: Niche construction by non-diazotrophs for N_2 fixers in the eastern tropical North Atlantic Ocean, Geophysical Research Letters, 44, 6904-6913, https://doi.org/10.1002/2017gl074218, 2017.

Sipler, R. E., Gong, D., Baer, S. E., Sanderson, M. P., Roberts, Q. N., Mulholland, M. R., and Bronk, D. A.: Preliminary estimates of the contribution of Arctic nitrogen fixation to the global nitrogen budget, Limnology and Oceanography Letters, 2, 159-166, https://doi.org/10.1002/lol2.10046, 2017.

Sohm, J. A., Hilton, J. A., Noble, A. E., Zehr, J. P., Saito, M. A., and Webb, E. A.: Nitrogen fixation in the South Atlantic Gyre and the Benguela ~~Upwelling System, Geophysical Research Letters,~~upwelling system, Geophys. Res. Lett., 38, ~~n/a-n/a~~L16608, https://doi.org/10.1029/~~2011gl048315~~2011GL048315, 2011.

Staal, M., Lintel-Hekkert, S. t., Harren, F., and Stal, L.: Nitrogenase activity in cyanobacteria measured by the acetylene reduction assay: a comparison between batch incubation and on-line monitoring, Environmental Microbiology, 3, 343-351, https://doi.org/10.1046/j.1462-2920.2001.00201.x, 2001.

Staal, M., te Lintel Hekkert, S., Jan Brummer, G., Veldhuis, M., Sikkens, C., Persijn, S., and Stal, L. J.: Nitrogen fixation along a north-south transect in the eastern Atlantic Ocean, Limnology and Oceanography, 52, 1305-1316, https://doi.org/10.4319/lo.2007.52.4.1305, 2007.

Stenegren, M., Caputo, A., Berg, C., Bonnet, S., and Foster, R. A.: Distribution and drivers of symbiotic and free-living diazotrophic cyanobacteria in the western tropical South Pacific, Biogeosciences, 15, 1559-1578, https://doi.org/10.5194/bg-15-1559-2018, 2018.

~~Stewart, W. D., Fitzgerald, G. P., and Burris, R. H.: In situ studies on N_2 fixation using the acetylene reduction technique, P Natl Acad Sci USA, 58, 2071-2078, https://doi.org/10.1073/pnas.58.5.2071, 1967.~~

Stenegren, M., Berg, C., Padilla, C., David, S.-S., Montoya, J., Yager, P., and Foster, R.: Piecewise Structural Equation Model (SEM) Disentangles the Environmental Conditions Favoring Diatom Diazotroph Associations (DDAs) in the Western Tropical North Atlantic (WTNA), Front Microbiol, 8, https://doi.org/10.3389/fmicb.2017.00810, 2017.

Subramaniam, A., Mahaffey, C., Johns, W., and Mahowald, N.: Equatorial upwelling enhances nitrogen fixation in the Atlantic Ocean, Geophysical Research Letters, 40, 1766-1771, https://doi.org/10.1002/grl.50250, 2013.

Subramaniam, A., Yager, P., Carpenter, E., Mahaffey, C., Björkman, K., Cooley, S., Kustka, A., Montoya, J., Sañudo-Wilhelmy, S., and Shipe, R.: Amazon River enhances diazotrophy and carbon sequestration in the tropical North Atlantic Ocean, ~~P Natl Acad Sci USA~~Proceedings of the National Academy of Sciences, 105, 10460-10465, https://doi.org/10.1073/pnas.0710279105, 2008.

Suzuki, S., Kawachi, M., Tsukakoshi, C., Nakamura, A., Hagino, K., Inouye, I., and Ishida, K.-i.: Unstable relationship between *Braarudosphaera bigelowii* (= *Chrysochromulina parkeae*) and its nitrogen-fixing endosymbiont, Frontiers in Plant Science, 12, https://doi.org/10.3389/fpls.2021.749895, 2021.

Tang, W., Cerdán-García, E., Berthelot, H., Polyviou, D., Wang, S., Baylay, A., Whitby, H., Planquette, H., Mowlem, M., Robidart, J., and Cassar, N.: New insights into the distributions of nitrogen fixation and diazotrophs revealed by high-resolution sensing and sampling methods, The ISME Journal, 14, 2514-2526, https://doi.org/10.1038/s41396-020-0703-6, 2020.

Tang, W. Y. and Cassar, N.: Data-driven modeling of the distribution of diazotrophs in the global ocean, Geophysical Research Letters, 46, 12258-12269, https://doi.org/10.1029/2019gl084376, 2019.

Tang, W. Y., Wang, S., Fonseca-Batista, D., Dehairs, F., Gifford, S., Gonzalez, A. G., Gallinari, M., Planquette, H., Sarthou, G., and Cassar, N.: Revisiting the distribution of oceanic N_2 fixation and estimating diazotrophic contribution to marine production, Nature Communications, 10, https://doi.org/10.1038/s41467-019-08640-0, 2019.

Tenório, M. M. B., Dupouy, C., Rodier, M., and Neveux, J.: *Trichodesmium* and other planktonic cyanobacteria in New Caledonian waters (SW tropical Pacific) during an El Niño episode, Aquatic Microbial Ecology, 81, 219-241, https://doi.org/10.3354/ame01873, 2018.

Thomas, B. L. K.: Geometric means and measures of dispersion, Biometrics, 35, 908-909, 1979.

Thompson, A., Carter, B. J., Turk-Kubo, K., Malfatti, F., Azam, F., and Zehr, J. P.: Genetic diversity of the unicellular nitrogen-fixing cyanobacteria UCYN-A and its prymnesiophyte host, Environ. Microbiol., 16, 3238-3249, https://doi.org/10.1111/1462-2920.12490, 2014.

Thompson, A. W., Foster, R. A., Krupke, A., Carter, B. J., Musat, N., Vaulot, D., Kuypers, M. M. M., and Zehr, J. P.: Unicellular cyanobacterium symbiotic with a single-celled eukaryotic alga, Science~~,~~ (New York, N.Y.), 337, 1546-1550, https://doi.org/10.1126/science.1222700, 2012.

Tuo, S.-h., Mulholland, M. R., Taniuchi, Y., Chen, H.-Y., Jane, W.-N., Lin, Y.-H., and Chen, Y.-l. L.: Trichome lengths of the heterocystous N_2-fixing cyanobacteria in the tropical marginal seas of the western north pacific, Frontiers in Marine Science, 8, https://doi.org/10.3389/fmars.2021.678607, 2021.

Turk-Kubo, K., Achilles, K., Serros, T., Ochiai, M., Montoya, J., and Zehr, J.: Nitrogenase (*nifH*) gene expression in diazotrophic cyanobacteria in the Tropical North Atlantic in response to nutrient amendments., Frontiers in Aquatic Microbiology, 3, 1-17, https://doi.org/10.3389/fmicb.2012.00386, 2012.

Turk-Kubo, K., Gradoville, M., Cheung, S., Cornejo Castillo, F. M., Harding, K., Morando, M., Mills, M., and Zehr, J.: Non-cyanobacterial diazotrophs: Global diversity, distribution, ecophysiology, and activity in marine waters, FEMS microbiology reviews, https://doi.org/10.1093/femsre/fuac046, 2022.

Turk-Kubo, K. A., Karamchandani, M., Capone, D. G., and Zehr, J. P.: The paradox of marine heterotrophic nitrogen fixation: abundances of heterotrophic diazotrophs do not account for nitrogen fixation rates in the Eastern Tropical South Pacific, Environ Microbiol, 16, 3095-3114, https://doi.org/10.1111/1462-2920.12346, 2014.

Turk-Kubo, K. A., Mills, M. M., Arrigo, K. R., van Dijken, G., Henke, B. A., Stewart, B., Wilson, S. T., and Zehr, J. P.: UCYN-A/haptophyte symbioses dominate N_2 fixation in the Southern California Current System, ISME Communications, 1, 42, https://doi.org/10.1038/s43705-021-00039-7, 2021.

Verity, P. G., Robertson, C. Y., Tronzo, C. R., Andrews, M. G., Nelson, J. R., and Sieracki, M. E.: Relationships between cell volume and the carbon and nitrogen content of marine photosynthetic nanoplankton, ~~Limnol Oceanogr~~Limnology and Oceanography, 37, 1434-1446, https://doi.org/10.4319/lo.1992.37.7.1434, 1992.

Villareal, T. A., Adornato, L., Wilson, C., and Schoenbaechler, C. A.: Summer blooms of diatom-diazotroph assemblages and surface chlorophyll in the North Pacific gyre: A disconnect, Journal of Geophysical Research, 116, https://doi.org/10.1029/2010jc006268, 2011.

Wang, S., Tang, W., Delage, E., Gifford, S., Whitby, H., González, A. G., Eveillard, D., Planquette, H., and Cassar, N.: Investigating the microbial ecology of coastal hotspots of marine nitrogen fixation in the western North Atlantic, ~~Sci Rep-Uk~~Scientific Reports, 11, https://doi.org/10.1038/s41598-021-84969-1, 2021.

Wang, W. L., Moore, J. K., Martiny, A. C., and Primeau, F. W.: Convergent estimates of marine nitrogen fixation, Nature, 566, 205-213, https://doi.org/10.1038/s41586-019-0911-2, 2019.

Wannicke, N., Benavides, M., Dalsgaard, T., Dippner, J. W., Montoya, J. P., and Voss, M.: New perspectives on nitrogen ~~fixation~~Fixation measurements using $^{15}N_2$ ~~gas, Front. Mar. Sci.,~~Gas, Frontiers in Marine Science, 5, https://doi.org/10.3389/fmars.2018.00120, 2018.

Wasmund, N., Struck, U., Hansen, A., Flohr, A., Nausch, G., Grüttmüller, A., and Voss, M.: Missing nitrogen fixation in the Benguela region, Deep Sea Research Part I: Oceanographic Research Papers, 106, 30-41, https://doi.org/10.1016/j.dsr.2015.10.007, 2015.

Watkins-Brandt, K., Letelier, R., Spitz, Y., Church, M., Böttjer, D., and White, A.: Addition of inorganic or organic phosphorus enhances nitrogen and carbon fixation in the oligotrophic North Pacific, Marine Ecology Progress Series, 432, 17-29, https://doi.org/10.3354/meps09147, 2011.

Wen, Z., Lin, W., Shen, R., Hong, H., Kao, S.-J., and Shi, D.: Nitrogen fixation in two coastal upwelling regions of the Taiwan Strait, ~~Sci Rep-Uk~~Scientific Reports, 7, 17601, https://doi.org/10.1038/s41598-017-18006-5, 2017.

Wen, Z., Browning, T. J., Cai, Y., Dai, R., Zhang, R., Du, C., Jiang, R., Lin, W., Liu, X., Cao, Z., Hong, H., Dai, M., and Shi, D.: Nutrient regulation of biological nitrogen fixation across the tropical western North Pacific, Science advances, 8, eabl7564-eabl7564, https://doi.org/10.1126/sciadv.abl7564, 2022.

White, A. E., Watkins-Brandt, K. S., and Church, M. J.: Temporal variability of *Trichodesmium* spp. and diatom-diazotroph assemblages in the North Pacific Subtropical Gyre, Front. Mar. Sci., 5, https://doi.org/10.3389/fmars.2018.00027, 2018.

White, A. E., Granger, J., Selden, C., Gradoville, M. R., Potts, L., Bourbonnais, A., Fulweiler, R. W., Knapp, A. N., Mohr, W., Moisander, P. H., Tobias, C. R., Caffin, M., Wilson, S. T., Benavides, M., Bonnet, S., Mulholland, M. R., and Chang, B. X.: A critical review of the $^{15}N_2$ tracer method to measure diazotrophic production in pelagic ecosystems, Limnology and Oceanography: Methods, 18, 129-147, https://doi.org/10.1002/lom3.10353, 2020.

Wilson, S. T., Böttjer, D., Church, M. J., and Karl, D. M.: Comparative assessment of nitrogen fixation methodologies, conducted in the oligotrophic North Pacific Ocean, Applied and Environmental Microbiology, 78, 6516-6523, https://doi.org/10.1128/aem.01146-12, 2012.

Wilson, S. T., Aylward, F. O., Ribalet, F., Barone, B., Casey, J. R., Connell, P. E., Eppley, J. M., Ferrón, S., Fitzsimmons, J. N., Hayes, C. T., Romano, A. E., Turk-Kubo, K. A., Vislova, A., Armbrust, E. V., Caron, D. A., Church, M. J., Zehr, J. P., Karl, D. M., and DeLong, E. F.: Coordinated regulation of growth, activity and transcription in natural populations of the unicellular nitrogen-fixing cyanobacterium *Crocosphaera*, ~~Nat. Microbiol~~Nature Microbiology, 2, 17118, https://doi.org/10.1038/nmicrobiol.2017.118, 2017.

Woebken, D., Burow, L. C., Behnam, F., Mayali, X., Schintlmeister, A., Fleming, E. D., Prufert-Bebout, L., Singer, S. W., Cortés, A. L., Hoehler, T. M., Pett-Ridge, J., Spormann, A. M., Wagner, M., Weber, P. K., and Bebout, B. M.: Revisiting N2 fixation in Guerrero Negro intertidal microbial mats with a functional single-cell approach, ISME J, 9, 485-496, https://doi.org/10.1038/ismej.2014.144, 2015.

Wu, C., Kan, J., Liu, H., Pujari, L., Guo, C., Wang, X., and Sun, J.: Heterotrophic bacteria dominate the diazotrophic community in the Eastern Indian Ocean (EIO) during pre-southwest monsoon, Microb Ecol, 78, 804-819, https://doi.org/10.1007/s00248-019-01355-1, 2019.

Wu, C., Sun, J., Liu, H., Xu, W., Zhang, G., Lu, H., and Guo, Y.: Evidence of the significant contribution of heterotrophic diazotrophs to nitrogen fixation in the Eastern Indian Ocean during pre-southwest monsoon period, Ecosystems, https://doi.org/10.1007/s10021-021-00702-z, 2021.

Yeung, L. Y., Berelson, W. M., Young, E. D., Prokopenko, M. G., Rollins, N., Coles, V. J., Montoya, J. P., Carpenter, E. J., Steinberg, D. K., Foster, R. A., Capone, D. G., and Yager, P. L.: Impact of diatom-diazotroph associations on carbon export in the Amazon River plume, ~~Geophys. Res. Lett.,~~Geophysical Research Letters, 39, https://doi.org/10.1029/2012GL053356, 2012.

Yogev, T., Rahav, E., Bar-Zeev, E., Man-Aharonovich, D., Stambler, N., Kress, N., Béjà, O., Mulholland, M. R., Herut, B., and Berman-Frank, I.: Is dinitrogen fixation significant in the Levantine Basin, East Mediterranean Sea?, Environmental Microbiology, 13, 854-871, https://doi.org/10.1111/j.1462-2920.2010.02402.x, 2011.

Zehr, J. P.: Nitrogen fixation by marine cyanobacteria, Trends Microbiol., 19, 162-173, https://doi.org/10.1016/j.tim.2010.12.004, 2011.

Zehr, J. P. and Capone, D. G.~~: Changing perspectives in marine nitrogen fixation, Science, 368, https://doi.org/10.1126/science.aay9514, 2020.~~

~~Zehr, J. P. and Capone, D. G.:~~ Marine nitrogen fixation~~,~~ Springer, https://doi.org/10.1007/978-3-030-67746-6, 2021.

~~Zehr, J. P., Mellon, M., Braun, S., Litaker, W., Steppe, T., and Paerl, H. W.: Diversity of heterotrophic nitrogen fixation genes in a marine cyanobacterial mat, Applied and Environmental Microbiology, 61, 2527-2532, https://doi.org/10.1111/j.1574-6968.1997.tb12589.x, 1995.~~

~~Zehr, J. P., Mellon, M. T., Zani, S: New nitrogen-fixing microorganisms detected in oligotrophic oceans by amplification of nitrogenase (*nifH*) genes, Appl Environ Microbiol, 64, https://doi.org/.10.1128/AEM.64.9.3444-3450.1998, 1998.~~

Zehr, J. P. and Riemann, L.: Quantification of gene copy numbers is valuable in marine microbial ecology: A comment to Meiler et al. (2022), Limnol Oceanogr, 68, 1406-1412, https://doi.org/10.1002/lno.12364, 2023.

Zhang, R., Zhang, D., Chen, M., Jiang, Z., Wang, C., Zheng, M., Qiu, Y., and Huang, J.: N_2 fixation rate and diazotroph community structure in the western tropical North Pacific Ocean, Acta Oceanol. Sin., 38, 26-34, https://doi.org/10.1007/s13131-019-1513-4, 2019.

Zhang, R., Chen, M., Yang, Q., Lin, Y., Mao, H., Qiu, Y., Tong, J., Lv, E., Yang, Z., Yang, W., and Cao, J.: Physical-biological coupling of N_2 fixation in the northwestern South China Sea coastal upwelling during summer, Limnology and Oceanography, 60, 1411-1425, https://doi.org/10.1002/lno.10111, 2015.

Zhang, X., Ward, B. B., and Sigman, D. M.: Global nitrogen cycle: critical enzymes, organisms, and processes for nitrogen budgets and dynamics, Chem Rev, 120, 5308-5351, https://doi.org/10.1021/acs.chemrev.9b00613, 2020.