Supplement of

Version 2 of the global oceanic diazotroph database

Zhibo Shao[1,#], Yangchun Xu[1,#], Hua Wang[1], Weicheng Luo[1], Lice Wang[1], Yuhong Huang[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

[#] These authors contributed equally.

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

Includes:

Fig. S1- Fig. S8

Table S1 & Table S2

References
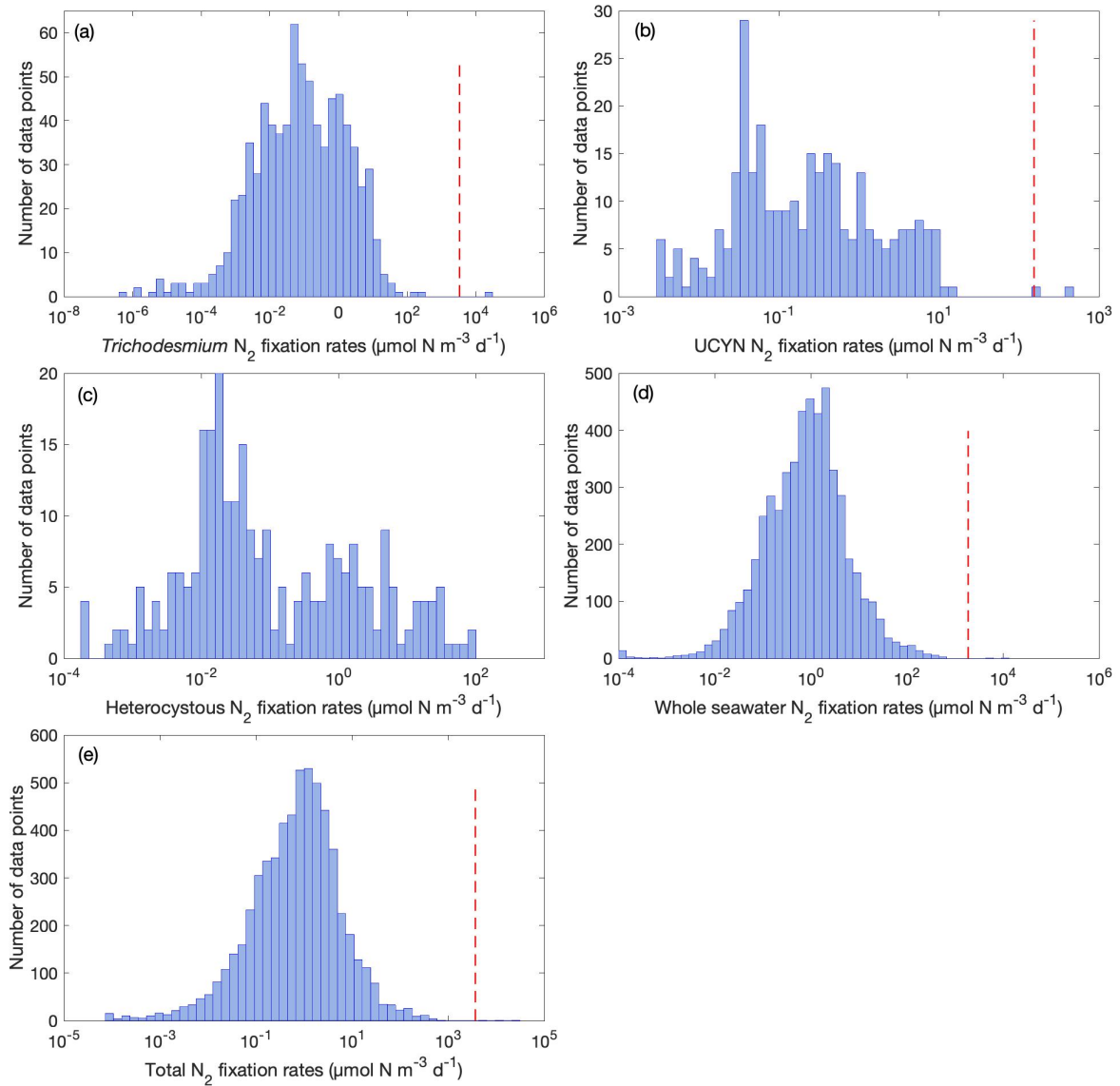
Figure S1. Histogram of volumetric N_2 fixation rates measurements for (a) *Trichodesmium* N_2 fixation, (b) UCYN N_2 fixation, (c) heterocystous N_2 fixation, (d) whole seawater N_2 fixation and (e) total N_2 fixation (non-zero values). Data values are on logarithmic scale. Red line denotes the results of quality control, indicating the critical values above which data are likely outliers.
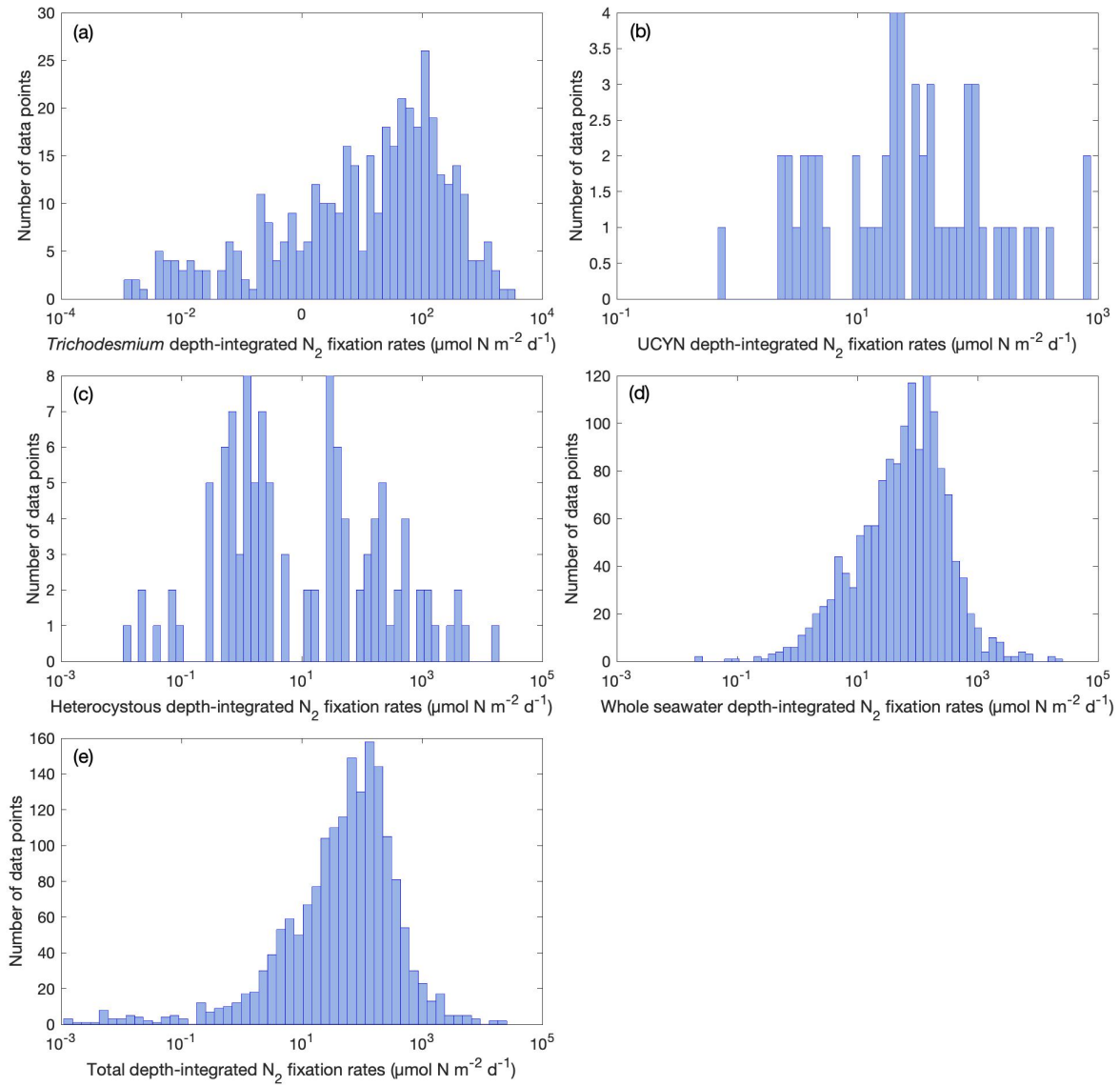
Figure S2. Histogram of depth-integrated N_2 fixation rates measurements for (a) *Trichodesmium* N_2 fixation, (b) UCYN N_2 fixation, (c) heterocystous N_2 fixation, (d) whole seawater N_2 fixation and (e) total N_2 fixation (non-zero values). Data values are on logarithmic scale.
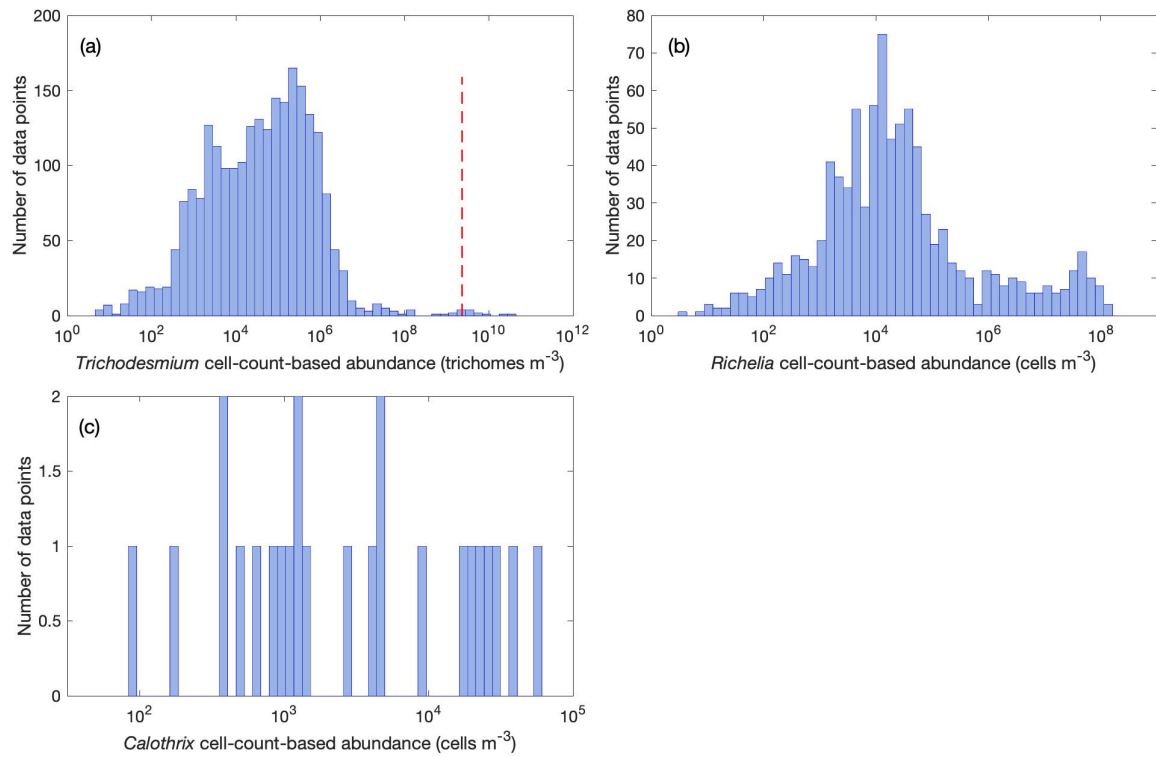
Figure S3. Histogram of volumetric cell counts measurements for (a) *Trichodesmium*, (b) UCYN-A, (c) *Richelia*, and (d) *Calothrix* (non-zero values). Data values are on logarithmic scale. Red line denotes the results of quality control, indicating the critical values above which data are likely outliers.
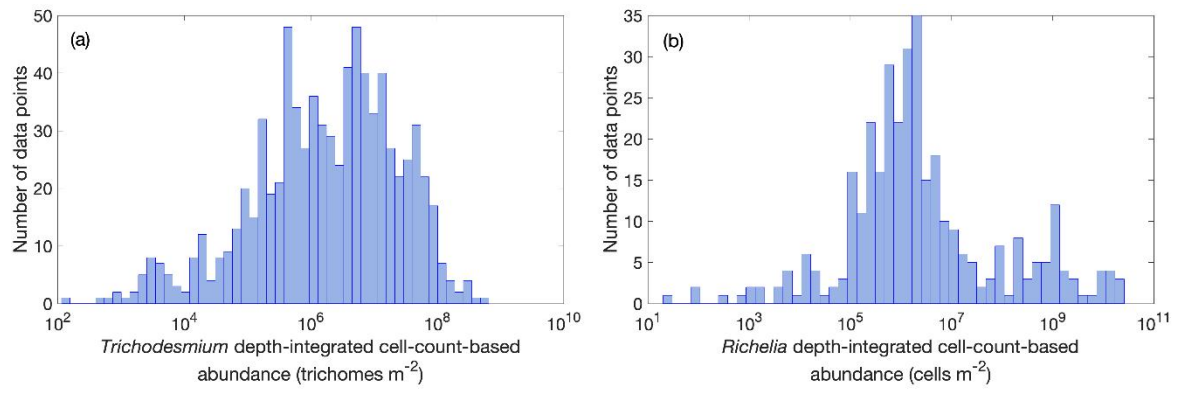
Figure S4. Histogram of depth-integrated cell counts measurements for (a) *Trichodesmium* and (b) *Richelia* (non-zero values). Data values are on logarithmic scale.
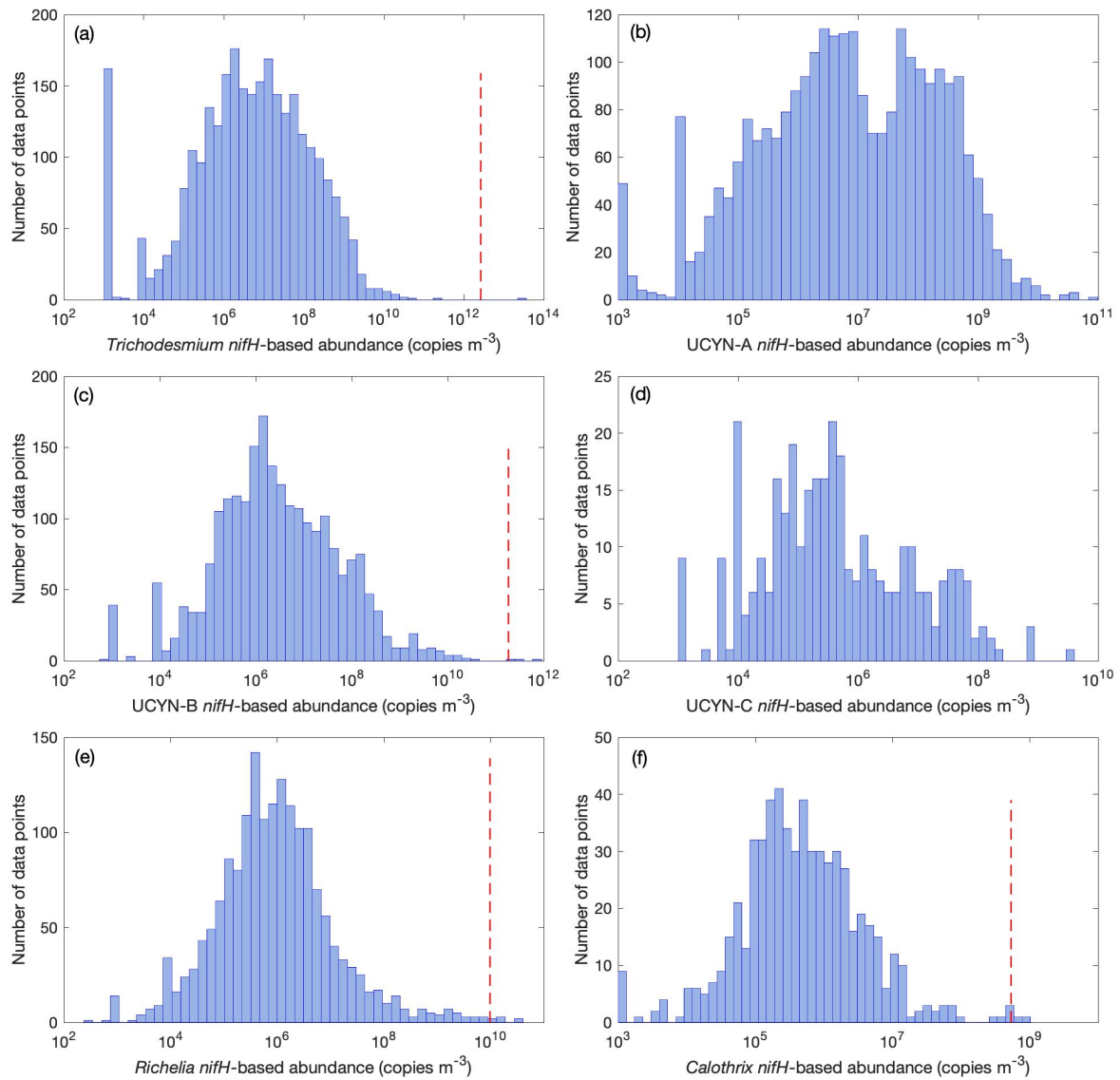
Figure S5. Histogram of volumetric *nifH* qPCR copies measurements for (a) *Trichodesmium*, (b) UCYN-A, (c) UCYN-B, (d) UCYN-C (e) *Richelia*, and (f) *Calothrix* (non-zero values). Data values are on logarithmic scale. Red line denotes the results of quality control, indicating the critical values above which data are likely outliers.
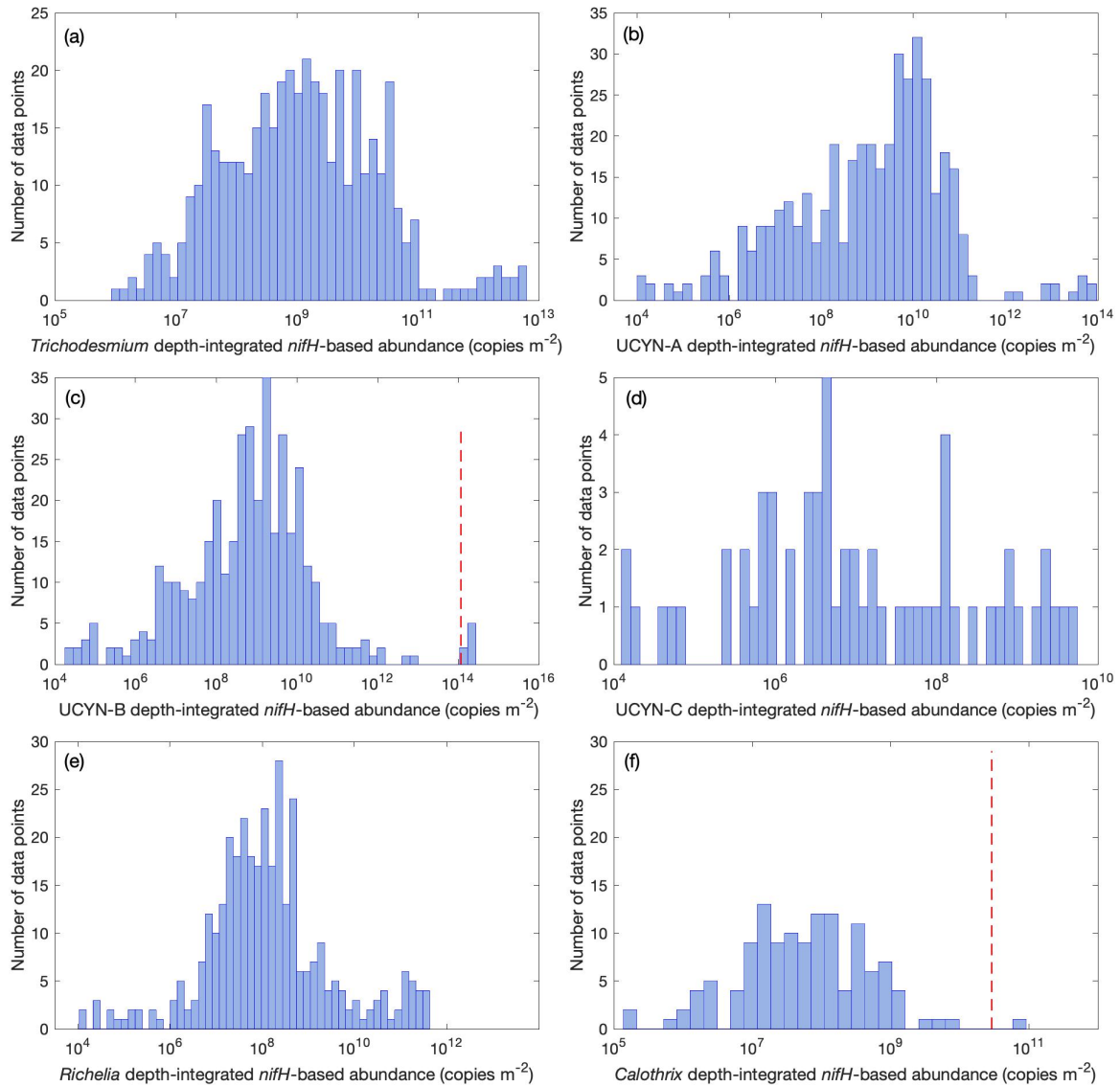
Figure S6. Histogram of depth-integrated *nifH* qPCR copies measurements for (a) *Trichodesmium*, (b) UCYN-A, (c) UCYN-B, (d) UCYN-C (e) *Richelia*, and (f) *Calothrix* (non-zero values). Data values are on logarithmic scale. Red lines denote the results of quality control, indicating the critical values above which data are likely outliers.
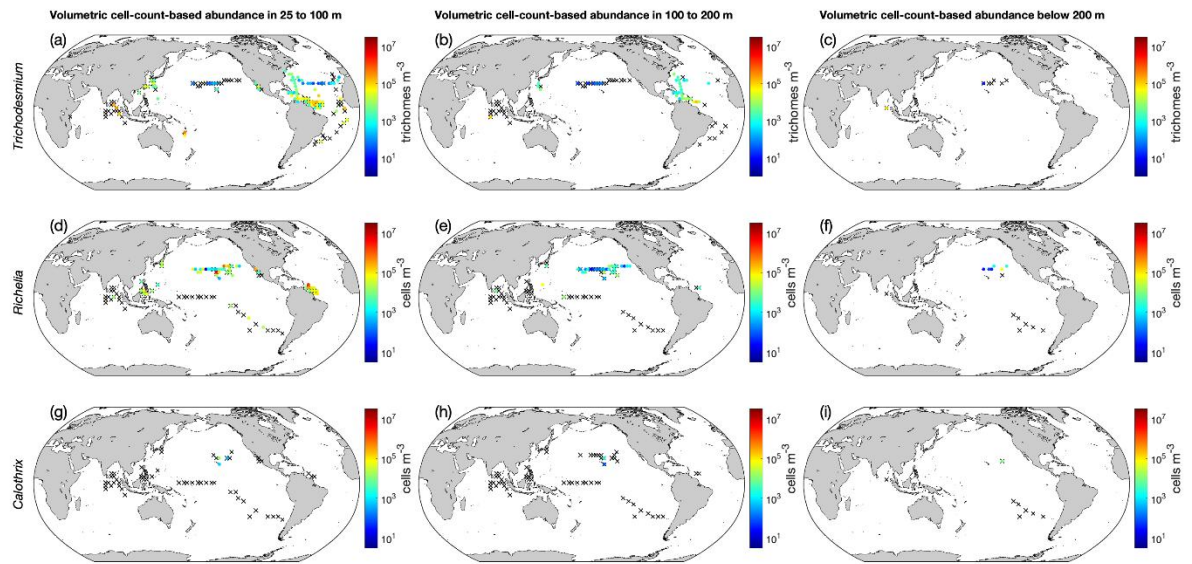
Figure S7. Volumetric cell-count-based abundance of diazotrophic groups. The panels show *Trichodemium* (a–c), *Richelia* (d–f) and *Calothrix* (g–i) volumetric data in depth ranges of 25–100 m (a, d, g), 100-200 m (b, e, h) and below 200 m (c, f, i). For clear demonstration, data are binned to 3° × 3° and geometric means in each bin are shown. Zero-value data are denoted as black crosses.
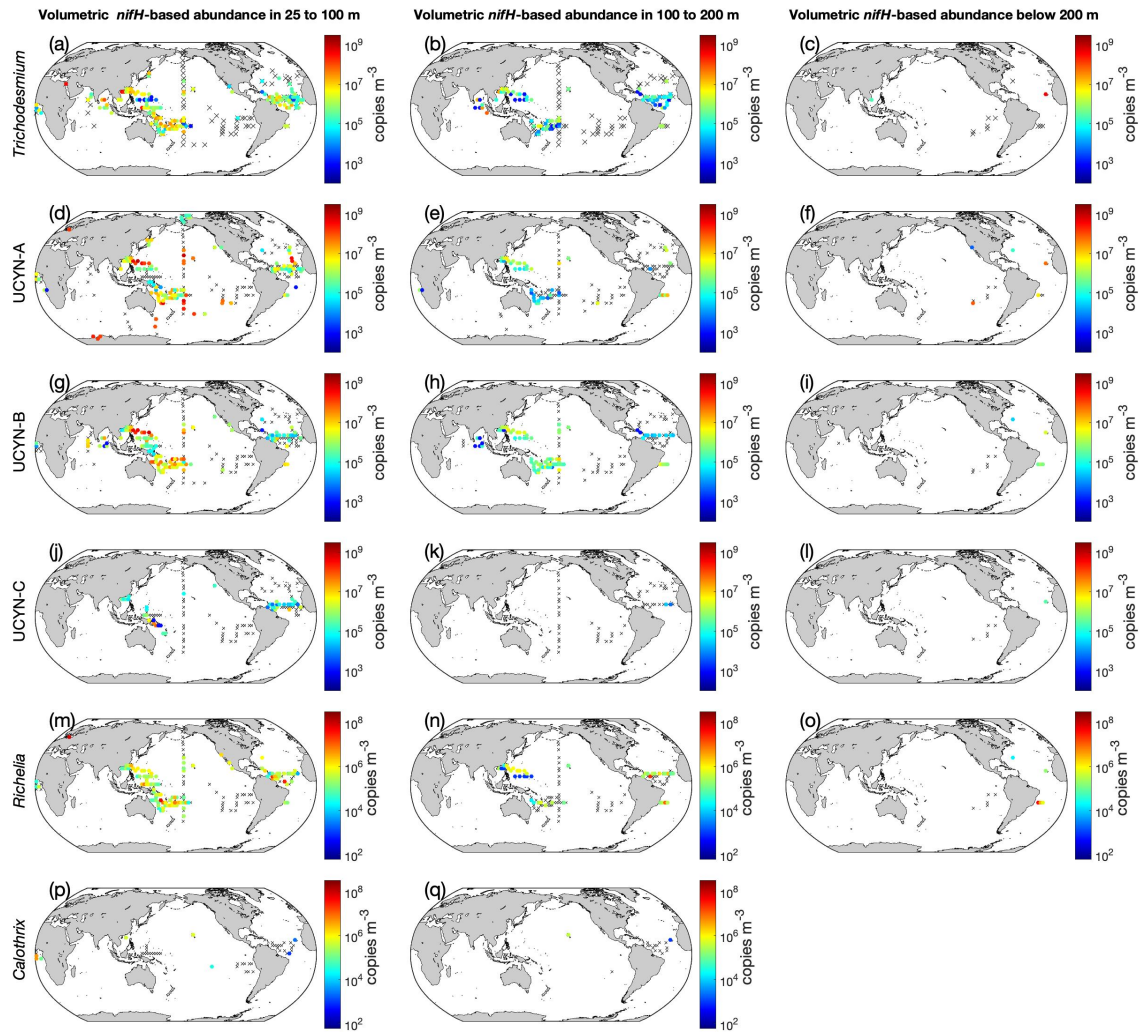
Figure S8. Volumetric *nifH* copies of different diazotrophic groups. The panels show *Trichodemium* (a–c), UCYN-A (d–f), UCYN-B (g–i), UCYN-C (j–l), *Richelia* (m–o) and *Calothrix* (p, q) volumetric data in depth ranges of 25–100 m (a, d, g, j, m, p), 100-200 m (b, e, h, k, n, q) and below 200 m (c, f, i, l, o). For clear demonstration, data are binned to 3° × 3° and geometric means in each bin are shown. Zero-value data are denoted as black crosses. *Calothrix* volumetric abundance has no data below 200 m.

Table S1. Summary of ratio of *nifH* gene copies to cell numbers in published paper. The numbers in parentheses are ranges of observations.

Reference	Species	Mean cells L^{-1}	Mean *nifH* gene copies L^{-1}	Mean *nifH* copies $cell^{-1}$
White et al. (2018)	*Trichodesmium*	1,130 (8 – 4,130)	207,000 (593 – 1,460,000)	148
	Het 1-3	190[a] (78 – 390)	21,200 (148 – 124,000)	76
Lu et al. (2018)	*Trichodesmium*	221,000 (19,000 – 423,000[b])	2,020,000 (42,300 – 40,400,000[b])	92
Sargent et al. (2016)[c]	*Trichodesmium*	(56 – 18,000)[d]	(0 – 270,000)[d]	12
		(0 – 19,000)[d]	(0 – 670,000)[d]	31
Krupke et al. (2013)	UCYN-A	134,000 (1,760 – 819,000)	746,000 (37,200 – 3,690,000)	14

[a] Heterocystous cell numbers were reported, here total cell numbers were used (heterocystous cell numbers multiply a factor of 6 assuming one filament contains one heterocystous cell and five vegetative cells).
[b] data collected during *Trichodesmium* bloom.
[c] two cruises in different seasons.
[d] no average data reported.

Table S2. Biomass conversion factors for heterocystous diazotroph-diatom associations. The numbers in parentheses are ranged used in the estimation.

Species	Diatom Biovolume (μm^3)	Biomass[a] (pg C cell⁻¹)	Vegetative cells per heterocyst	Heterocysts per DDA	Number of diazotrophic cells per DDA[b]	Diazotrophic cell biomass conversion (pg C cell⁻¹)	Diazotrophic biomass (pg C DDA⁻¹)	Total DDA biomass[c] (pg C DDA⁻¹)	DDA Biomass Conversion factor (pg C heterocyst⁻¹)
Richelia-Hemiaulus	10800 (7050−20900)	418 (287−748)		2	12 (8−22)	12 (2−80)	144 (16−1760)	562 (303−2508)	281 (152−1254)
Richelia-Rhizosolenia	57100 (439−188000)	1814 (25−5184)	5 (3−10)	5	30 (20−55)		360 (40−4400)	2174 (65−9584)	430 (13−1916)
Calothrix-Chaetoceros	1480 (300-4630)	73 (18−288)		2	12 (8−22)	12 (5−20)	144 (40−440)	217 (58−728)	110 (29−364)
				5	30 (20−55)		360 (100−1100)	433 (118−1388)	90 (24−278)

[a] Based on an empirical relationship (Menden-Deuer and Lessard, 2000) : biomass = $0.117 \times \text{biovolume}^{0.881}$.

[b] (1 + number of vegetative cells) × number of heterocyst.

[c] Diatom biomass + diazotrophic biomass.

Reference

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.

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.

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

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.

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, 27, https://doi.org/10.3389/fmars.2018.00027, 2018.