Supplementary Information of

**New insights into the distributions of nitrogen fixation and diazotrophs revealed by high-resolution sensing and sampling methods**

Weiyi Tang1,2, Elena Cerdán García3, Hugo Berthelot4,Despo Polyviou3, Seaver Wang1, Alison Baylay3, Hannah Whitby4, Hélène Planquette4, Matthew Mowlem5, Julie Robidart5,\*, Nicolas Cassar1,4,\*

1Division of Earth and Ocean Sciences, Nicholas School of the Environment, Duke University, Durham, NC 27708, USA

2Now at Department of Geosciences, Princeton University, NJ 08544, USA

3Department of Ocean and Earth Sciences, National Oceanography Centre, University of Southampton, European Way, SO14 3ZH, Southampton, United Kingdom

4Laboratoire des Sciences de l'Environnement Marin (LEMAR), UMR 6539 UBO/CNRS/IRD/IFREMER, Institut Universitaire Européen de la Mer (IUEM), Brest, France

5Ocean Technology and Engineering Group, National Oceanography Centre, European Way, SO14 3ZH, Southampton, United Kingdom

\*Corresponding to [j.robidart@noc.ac.uk](mailto:j.robidart@noc.ac.uk) and [Nicolas.Cassar@duke.edu](mailto:Nicolas.Cassar@duke.edu)

Supplementary Table 1. Primer and probe sequences used in *nifH* sequencing, qPCR and RT-qPCR analyses

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Target | Forward primer (5’-3’) | Probe (5’-3’) | Reverse primer (5’-3’) | Sources |
| *nifH1-*illumina | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG TGY GAY CCN AAR GCN GA |  |  | (1) and this study |
| *nifH2-*illumina |  |  | GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GAD NGC CAT CAT YTC NCC | (1) and this study |
| *nifH* *Trichodesmium* | GACGAAGTATTGAAGCCAGGTTTC | CATTAAGTGTGTTGAATCTGGTGGTCCTGAGC | CGGCCAGCGCAACCTA | (2) |
| *nifH*  UCYN-A | AGCTATAACAACGTTTTATGCGTTGA | TCTGGTGGTCCTGAGCCTGGA | ACCACGACCAGCACATCCA | (2) |
| *nifH*  *Crocosphaera* | CGTAATGCTCGAAGGGTTTGA | CAAGTGTGTAGAATCTGGTGGTCCTGAGCC | CACGACCAGCACAACCAACT | (2) |
| *pstS* *Trichodesmium* | ATGCACAAAACCCTGCCGAC | YAACCCYCTATACCCATATTCTCA | GAAATGCTAGAACTTCTSGAGAGG | This study |
| *pstS*  UCYN-A | ACGTTTACTTTGATTCGTCCC | AGCGTACTCGATCATGGTCTCCA | ATGAMCCTAAAAAAGCAATTG | This study |

Supplementary Table 2. Mean (standard deviations) of *nifH* gene copies, *nifH* transcripts and *pstS* transcripts of *Trichodesmium* and UCYN-A from 9 discrete noon samples. Standard deviations in the brackets. BD: below detection limits.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Stations | Latitude (°N) | Longitude (°W) | *Trichodesmium* (copies L-1) | | | UCYN-A (copies L-1) | | |
| *nifH* gene | *nifH* transcript | *pstS* transcript | *nifH* gene | *nifH* transcript | *pstS* transcript |
| 1 | 35.28 | -64.26 | 68798 (4118) | 34598 (10277) | 377 (216) | 617 (500) | BD (/) | BD (/) |
| 2 | 38.28 | -63.74 | 366702 (2684) | 4423684 (812297) | 17227 (3181) | 660 (268) | 7775 (758) | 1320 (/) |
| 3 | 41.88 | -64.94 | 184833 (18072) | 44156 (15748) | BD (/) | 734417 (46193) | 1018981 (398145) | 2453 (932) |
| 4 | 42.26 | -67.83 | 7503 (2105) | BD (/) | BD (/) | 1919667 (135665) | 1288149 (26483) | 2544 (/) |
| 5 | 40.17 | -71.32 | 988 (867) | BD (/) | BD (/) | 137658 (11592) | 3225045 (364139) | 5485 (3561) |
| 6 | 38.73 | -74.49 | 1522 (341) | BD (/) | BD (/) | 798778 (41485) | 1716301 (419210) | BD (/) |
| 7 | 36.21 | -71.7 | 397 (56) | 46029 (1570) | BD (/) | 750 (806) | 1161 (915) | 337 (/) |
| 8 | 34.84 | -68.02 | 1354 (/) | 1535 (369) | BD (/) | 336 (236) | BD (/) | BD (/) |
| 9 | 32.81 | -65.01 | 392 (112) | 8005 (3985) | BD (/) | 432 (110) | BD (/) | BD (/) |

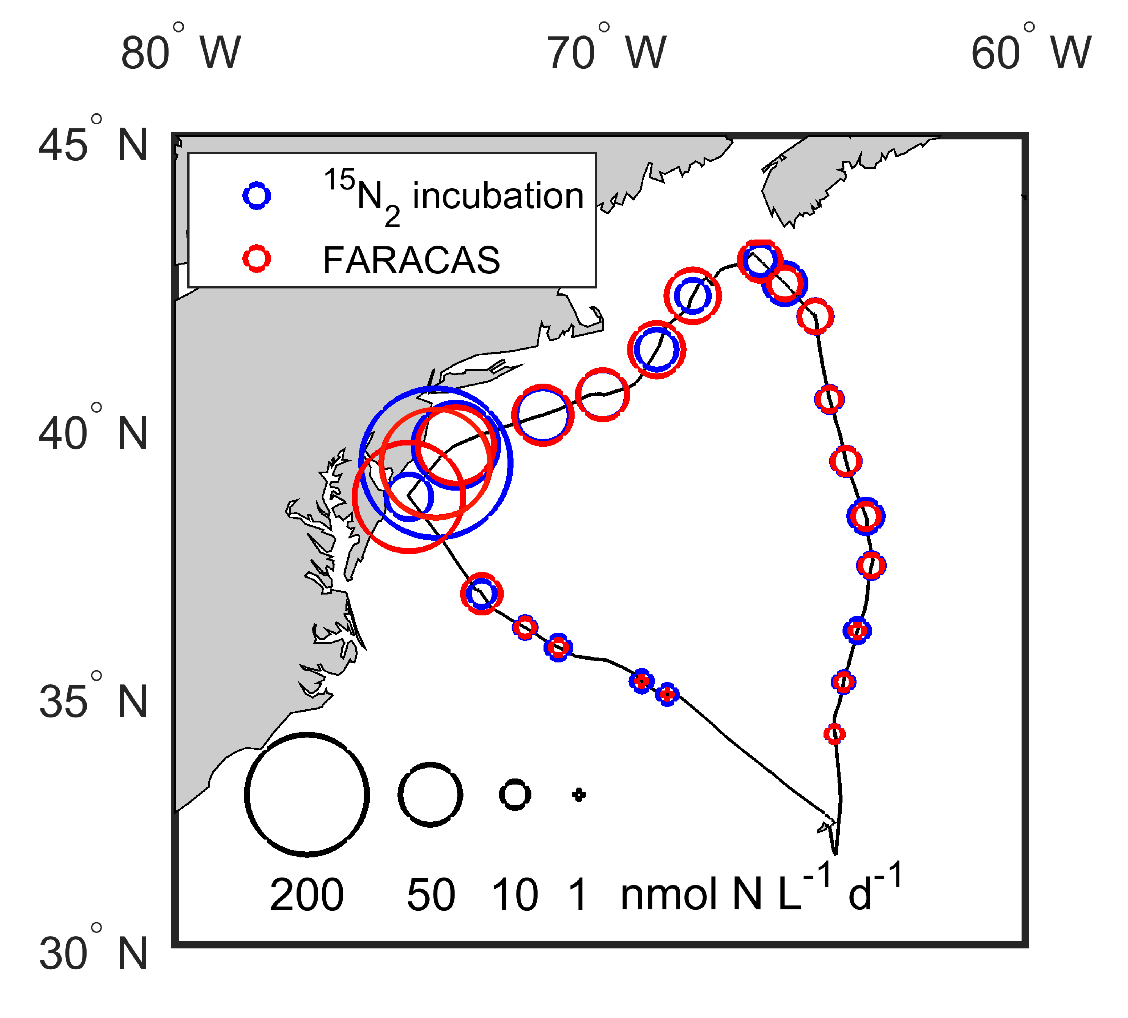
Supplementary Table 3. *nifH* gene copies from 34 RoCSI samples. BD: below detection limits.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Locations | | *Trichodesmium* (copies L-1) | | UCYN-A (copies L-1) | |
| Latitude (°N) | Longitude (°W) | *nifH* gene | standard deviation | *nifH* gene | standard deviation |
| 32.55 | -64.38 | 20820 | 5165 | 1914 | / |
| 33.85 | -64.44 | 19852 | 3327 | 261 | 142 |
| 34.19 | -64.49 | 37128 | 10673 | BD | / |
| 34.78 | -64.38 | 77889 | 5320 | 277 | 380 |
| 35.57 | -64.15 | 80182 | 18112 | BD | / |
| 36.23 | -63.93 | 229870 | 24067 | 250 | 224 |
| 36.63 | -63.77 | 47018 | 13114 | BD | / |
| 37.08 | -63.70 | 63910 | 872 | 276 | / |
| 37.44 | -63.61 | 83591 | 5319 | 797 | 276 |
| 37.84 | -63.67 | 38337 | 3326 | BD | / |
| 38.59 | -63.90 | 193602 | 13245 | 5173 | 813 |
| 39.06 | -64.12 | 153122 | 5196 | 1158 | 281706 |
| 42.92 | -66.56 | 76033 | 5196 | 1109644 | 281706 |
| 42.74 | -67.12 | BD | / | 461600 | 41947 |
| 42.52 | -67.43 | BD | / | 4976690 | 1108460 |
| 42.55 | -67.57 | 2840 | / | 3127298 | 185226 |
| 42.17 | -67.89 | BD | / | 1206545 | 581922 |
| 41.33 | -68.67 | BD | / | 254921 | 65990 |
| 40.93 | -68.99 | 15992 | 9200 | 408588 | 128904 |
| 40.62 | -69.49 | BD | / | 701648 | 112049 |
| 40.53 | -70.19 | 29719 | 2020 | 765909 | 260956 |
| 40.34 | -70.77 | 80950 | 7380 | 339202 | 68401 |
| 40.11 | -71.61 | 61977 | 7606 | 597229 | 89011 |
| 40.00 | -72.03 | 19911 | 2622 | 90756 | 31837 |
| 39.79 | -72.87 | BD | / | 4645482 | 149798 |
| 39.40 | -73.78 | BD | / | 40202944 | 2791211 |
| 39.22 | -73.95 | BD | / | 30262500 | 2103643 |
| 38.99 | -74.23 | BD | / | 10822610 | 1267333 |
| 38.61 | -74.44 | BD | / | 919038 | 97930 |
| 37.78 | -73.64 | 33178 | 0 | 44713 | 9016 |
| 37.43 | -73.34 | 22000 | 7376 | 49082 | 12979 |
| 36.57 | -72.48 | 163150 | 20025 | 122 | 151 |
| 36.32 | -71.96 | 173360 | 10267 | BD | / |

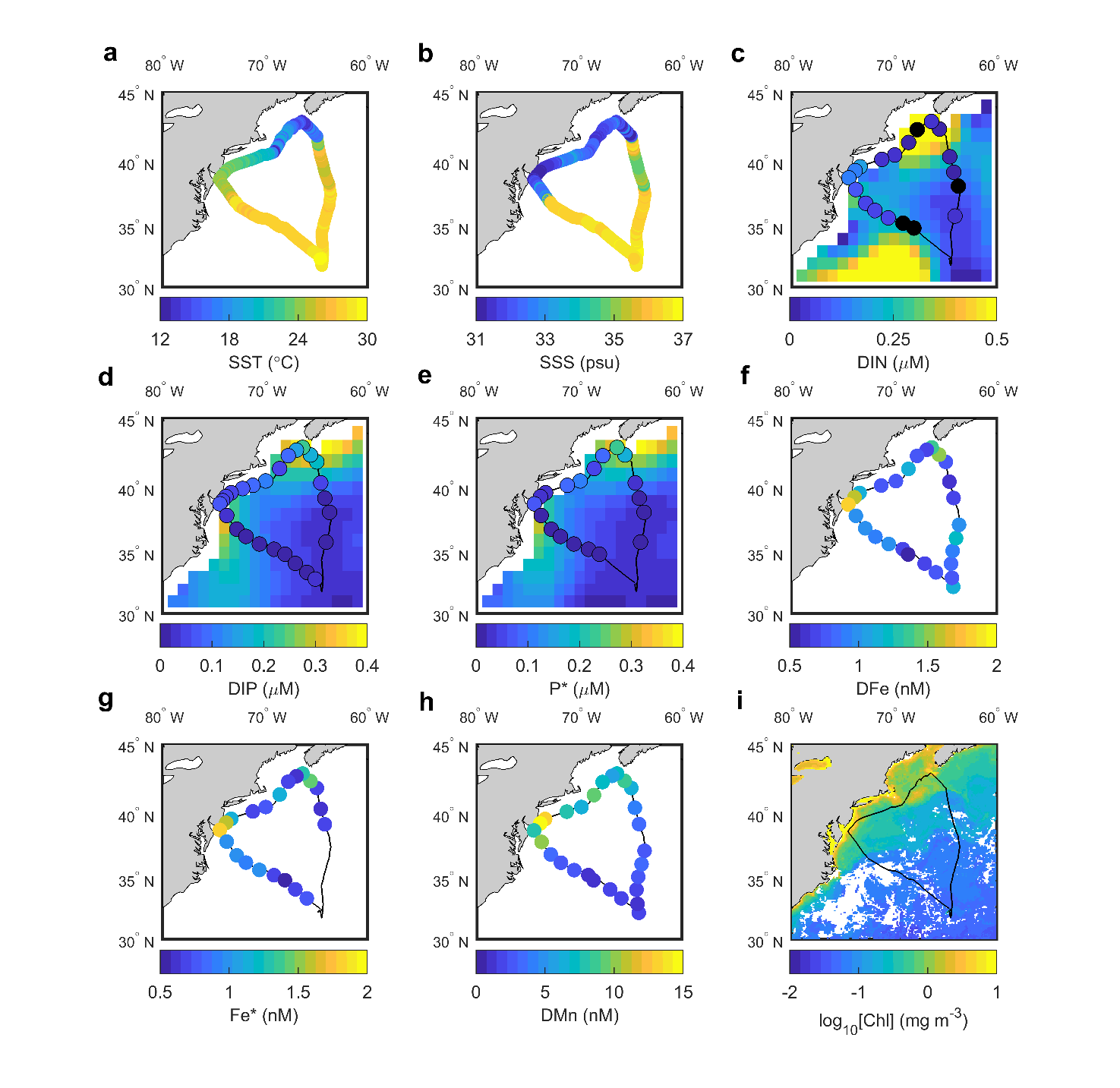
A close up of a map

Description automatically generated

Supplementary Figure 1. Comparison between *nifH* gene and *pstS* gene copy numbers of UCYN-A measured by qPCR in the same samples.



Supplementary Figure 2. Comparison of N2 fixation rates measured by FARACAS and by the dissolved 15N2 addition method.

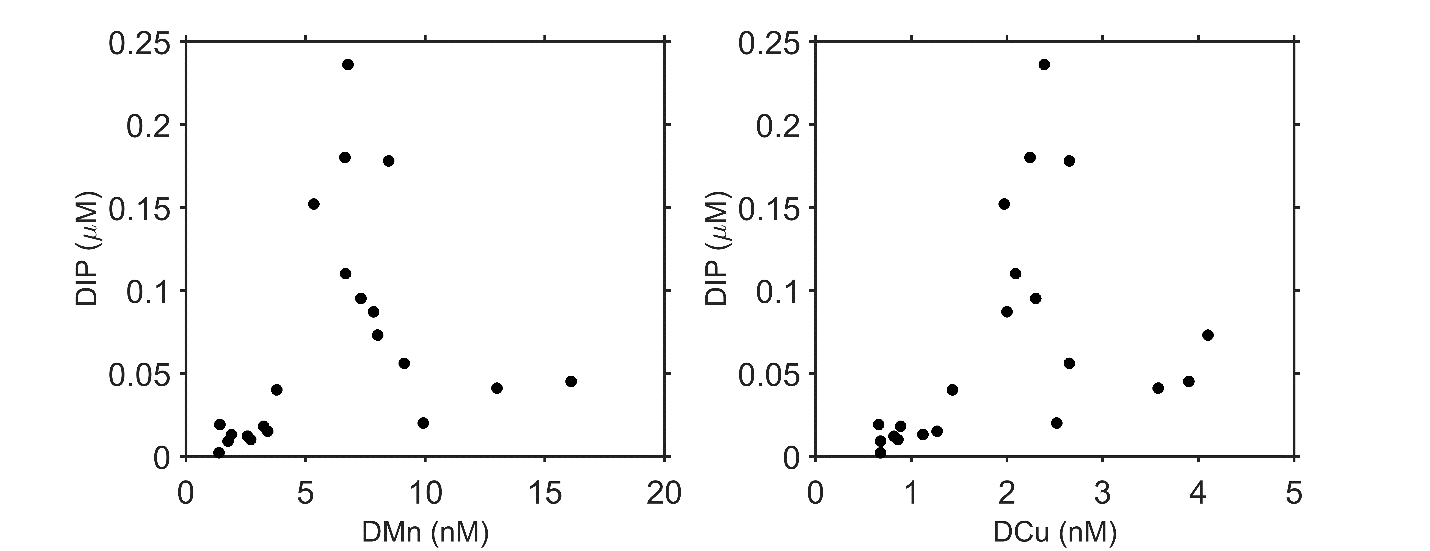


Supplementary Figure 3. Maps of surface properties. (**a**) sea surface temperature (SST), (**b**) sea surface salinity (SSS), (**c**) dissolved inorganic nitrogen concentrations (DIN = nitrate + nitrite) overlaid on climatology of DIN in August, (**d**) dissolved inorganic phosphorus concentrations (DIP) overlaid on climatology of DIP in August, (**e**) calculated excess phosphorus (P\*=DIP-DIN/16) overlaid on climatology of P\* in August, (**f**) dissolved iron concentrations (DFe), (**g**) excess DFe (; where ) calculated at stations where both DIP and DFe concentrations are available (3), (**h**) dissolved manganese concentrations (DMn), and (**i**) averaged chlorophyll-*a* concentration [Chl] observed by MODIS-Aqua during the cruise period. Samples below detection limits are shown as black circles. Nutrient climatologies were obtained from World Ocean Atlas 2013 (4).

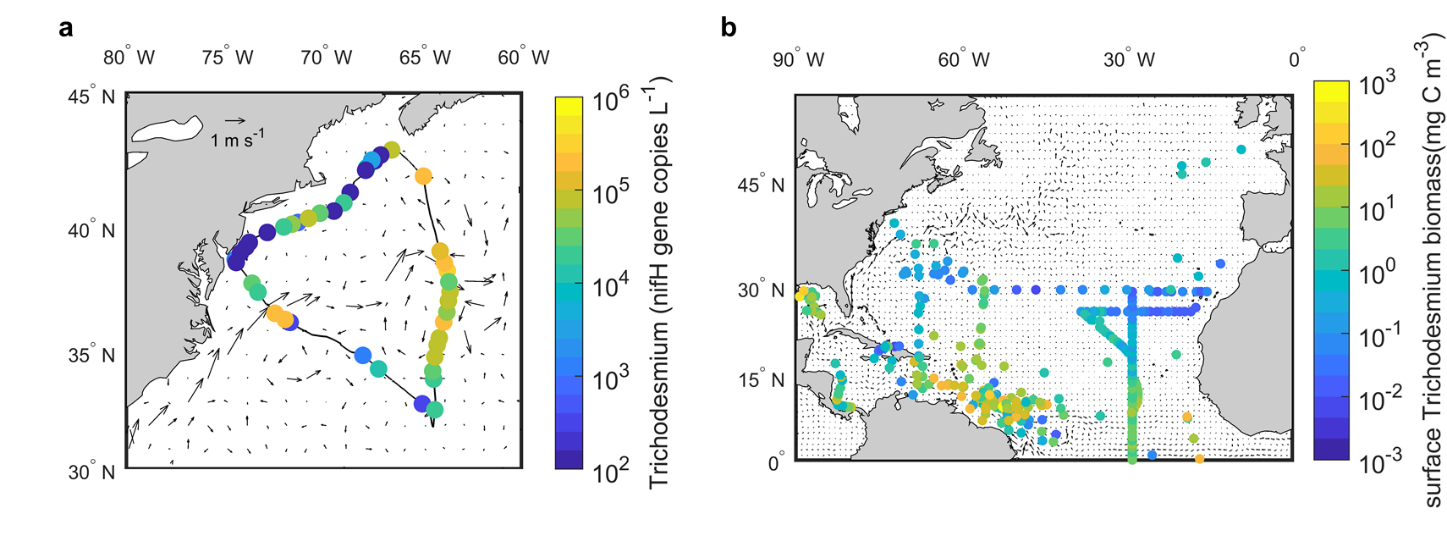
A screenshot of a cell phone

Description automatically generated

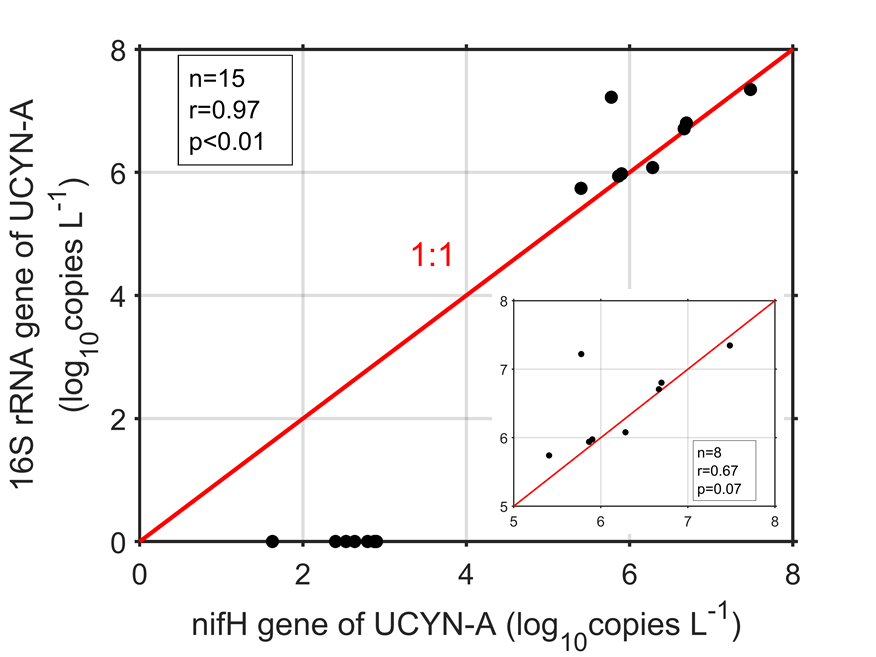
Supplementary Figure 4. Spearman’s rank correlation analyses between surface daily N2 fixation rates and environmental properties. N2 fixation rates vs (**a**) surface DIN climatology in August (r = 0.46, p<0.01), (**b**) DIN:DIP ratio (r = 0.09, p=0.73), (**c**) surface P\* (r = 0.47, p=0.04), (**d**) surface P\* climatology in August (r = 0.72, p<0.01), (**e**) surface DCu (r = 0.82, p<0.01) (**f**) satellite observed chlorophyll-*a* concentrations (r = 0.84, p<0.01).



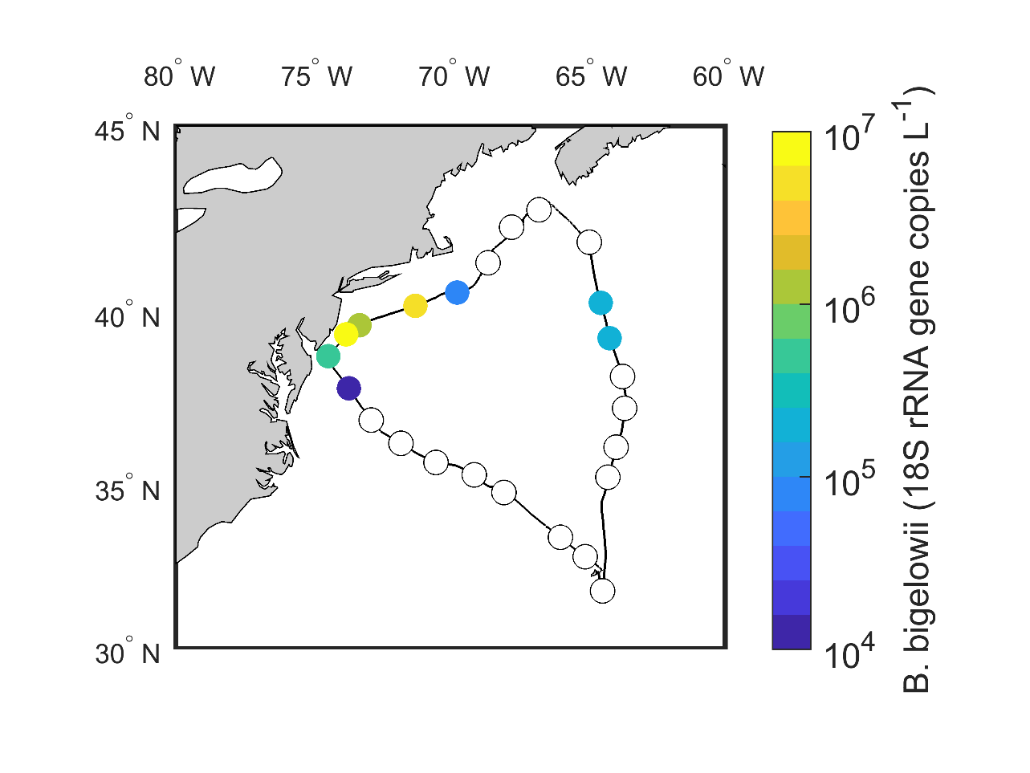
Supplementary Figure 5. Spearman’s rank correlation analyses between DMn, DCu and DIP concentrations. DIP is correlated to DMn (n=21, r=0.62, p<0.01) and DCu (n=21, r=0.67, p<0.01).



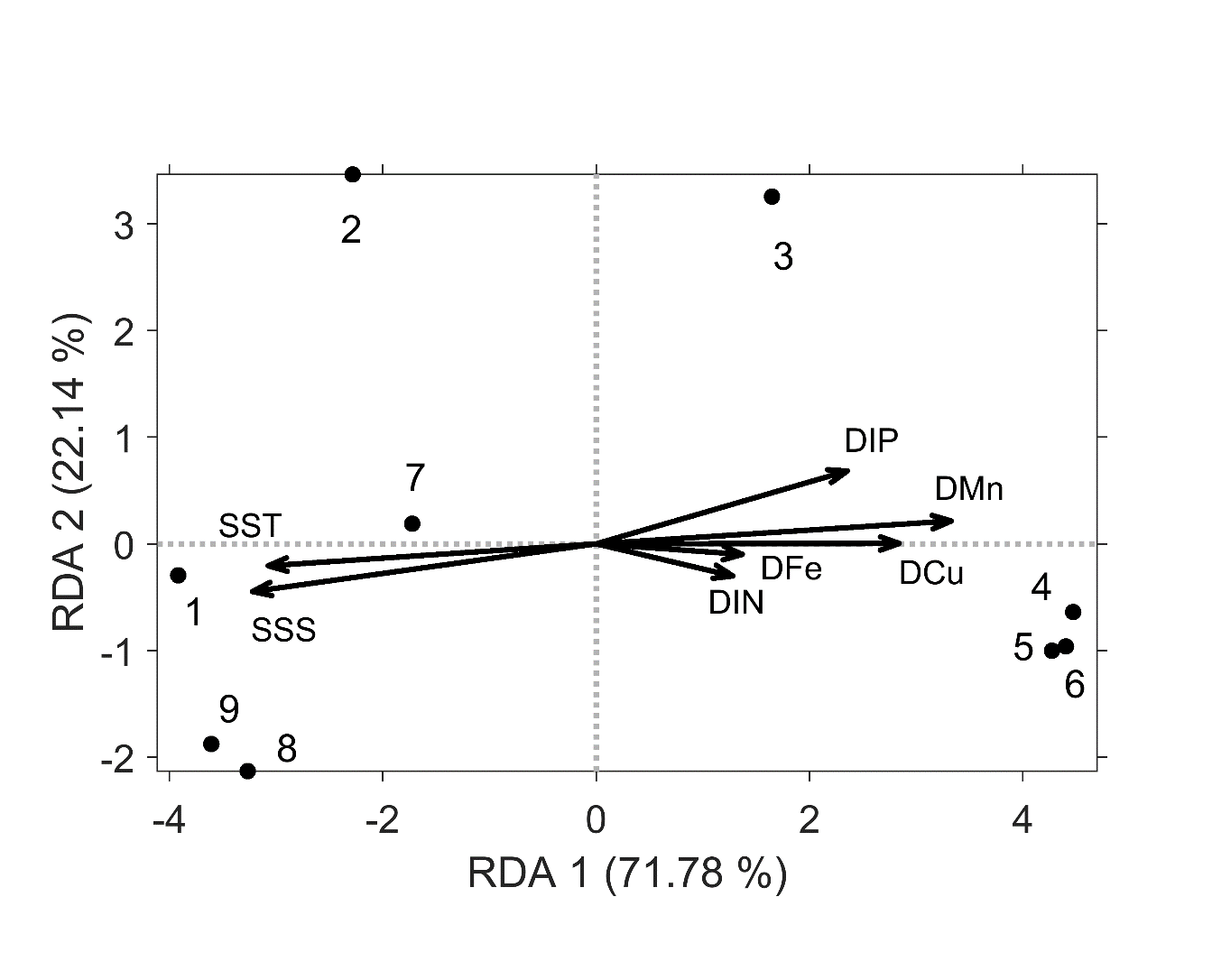
Supplementary Figure 6. (**a**) Color coded *Trichodesmium* *nifH* gene abundance measured in this study and (**b**) *Trichodesmium* biomass obtained from Luo et al., 2012 overlaid on surface currents (5).



Supplementary Figure 7. Comparison of UCYN-A abundance estimated by qPCR and quantitative 16 rRNA amplicon sequencing. UCYN-A abundance below the detection limit of quantitative 18S rRNA gene sequencing are shown as 0 on y-axis.



Supplementary Figure 8. Distribution of *Braarudosphaera bigelowii*. The white circles denote stations where *B. bigelowii* abundance wasbelow the detection limit of quantitative 18S rRNA gene sequencing (Wang et al., unpublished) (6).

****

Supplementary Figure 9. Redundancy analysis (RDA) biplot of the relationship between the *nifH* gene abundances of *Trichodesmium* and UCYN-A and environmental properties at 9 discrete daily noon stations. The numbers and points represent discrete stations.

**References:**

1. Zehr J, Turner P. Nitrogen Fixation: Nitrogenase Genes and Gene Expression. Method in Microbiology. 2001;30:271-86.

2. Church MJ, Jenkins BD, Karl DM, Zehr JP. Vertical distributions of nitrogen-fixing phylotypes at Stn ALOHA in the oligotrophic North Pacific Ocean. Aquatic Microbial Ecology. 2005;38(1):3-14.

3. Parekh P, Follows MJ, Boyle EA. Decoupling of iron and phosphate in the global ocean. Global Biogeochemical Cycles. 2005;19(2):GB2020.

4. Boyer TP, Antonov JI, Baranova OK, Coleman C, Garcia HE, Grodsky A, et al. NOAA Atlas NESDIS 72. World Ocean Database. 2013.

5. Luo YW, Doney SC, Anderson LA, Benavides M, Berman-Frank I, Bode A, et al. Database of diazotrophs in global ocean: abundance, biomass and nitrogen fixation rates. Earth System Science Data. 2012;4(1):47-73.

6. Wang S, Tang W, Delage E, Gifford S, Whitby H, Aridane G, et al. Patterns of North Atlantic microbial community structure in relation to net community production, nitrogen fixation, and nutrient availability. Unpublished.