**Appendix S1**

**This file includes:**

Supplementary Methods

Supplementary Tables S1-S2

Supplementary Figures S1-S7

Supplementary References

**Supplementary Methods**

*N2 fixation rate data compilation and calculation of basin-scale yearly rates*

We compiled volumetric (surface) and integrated N2 fixation rates from the available literature from the IO published as of October 2022, gathering a total of 269 surface station rates and 103 station integrated rates (Table S1). A total of 269 station surface rates and 103 station integrated rates are plotted in figures 2 and 3. Using interquartile range statistics (Q1 and Q3), we established the maximum (upper bound) limits for the surface and integrated rates at 23 nmol N l-1 d-1 and 361 μmol N m-2 d-1, respectively (excluding very high rates >1000 μmol N m-2 d-1 for integrated rates). Only one publication (Kromkamp et al. 1997) representing 15 station surface rates from the IO is included in the Luo et al. (2012) global database. Another global database by (Tang et al. 2019) included (Gandhi et al. 2011) and (Shiozaki et al. 2014) from the IO, corresponding to 13 surface and 13 integrated N2 fixation rates.

The spreadsheet used to calculate basin-scale yearly N2 fixation rates is available at <https://zenodo.org/record/7870536>. Briefly, depth-integrated rates were classified according to sub-basin and season, and their arithmetic mean, geometric mean and median calculated. The means and median were extrapolated to the area of each sub-basin considering the duration (days) of each season. Finally, sub-basin seasonal rates were summed for each sub-basin to obtain sub-basin yearly rates, and the latter were summed to obtain basin-scale rates.

*nifH sequence data compilation*

Currently, there are five *nif*H amplicon datasets available from the IO (Raes et al. 2018; Wu et al. 2019, 2021; Li et al. 2021; Sato et al. 2022). All the amplicon datasets from the IO were downloaded from the NCBI SRA database, European Nucleotide Archive, and DDBJ and were analyzed by the DADA2 pipeline (Callahan et al. 2016). Sequences were filtered, trimmed, and processed as per the DADA2 guidelines (Callahan et al. 2016). After merging all the cleaned reads of the forward and reverse sequences, sequence lengths of more than 300 base pairs long were considered for downstream processing. Chimeras were removed and after that filtered sequences were used to create the amplicon sequence variant (ASV) tables, ASVs abundance, and taxonomic annotations. The taxonomy of the ASVs was assigned by the recently updated nifHdada2 database <http://doi.org/10.5281/zenodo.3958370>. The output of the amplicon datasets was used to create the basin-wide heat maps and phylum percent contributions (detailed results of the analyzed dataset will be available upon request). Detailed output of the amplicon datasets and cloning-based sanger sequences used in this study are in Table S1.

We recovered 372 amplicon sequence variants (ASVs) from the BoB where 23.9% of sequences were represented by cyanobacteria and 76.1% by NCDs (Fig. S1). 964 ASVs retrieved from the EqIO by *nif*H amplicon analysis from 17 samples, where 9.02% and 90.82% sequences were represented by cyanobacteria and NCDs, respectively (Fig. S1). In the EIO, 15.3% of *nif*H sequences represent cyanobacteria, and 84.7% of the sequences are represented by NCDs (Fig. S1).

**Supplementary Tables**

**Table S1:** *nifH* amplicon sequences analyzed and cloned based *nif*H sequences used in this study.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Amplicon dataset** | **Reference** | **SRA Run ID** | **Latitude (ºN)** | **Longitude (ºE)** | **Data type** | **Total ASVs** |
| Bay of Bengal (BoB) | (Li et al. 2021) | SRR12376990 | 15.49 | 88.01 | Illumina Paired end | 372 |
| SRR12376991 | 15 | 91 |
| SRR12376992 | 15 | 88.5 |
| SRR12376994 | 15.01 | 87 |
| (Wu et al. 2019) | SRR6834173 | 9.99 | 84.49 |
| SRR6834176 | 10 | 87.99 |
| SRR6834178 | 10 | 87.99 |
| SRR6834179 | 10 | 87.99 |
| SRR6834180 | 9.99 | 84.49 |
| SRR6834181 | 9.99 | 84.49 |
| (Sato et al. 2022) | DRR393951 | 16.46 | 88.05 |
| DRR393952 | 15 | 88 |
| DRR393953 | 15 | 88 |
| DRR393954 | 9.93 | 88 |
| DRR393955 | 9.93 | 88 |
| Equatorial Indian Ocean (EqIO) | (Wu et al. 2019) | SRR6834172 | 0 | 84 | Illumina Paired end | 964 |
| SRR6834174 | 4.99 | 90.5 |
| SRR6834175 | 4.99 | 90.5 |
| SRR6834177 | 4.99 | 90.5 |
| SRR6834182 | 0 | 92 |
| SRR6834183 | 0 | 92 |
| SRR6834184 | 0 | 84 |
| SRR6834185 | 0 | 84 |
| SRR6834186 | 3.99 | 96.32 |
| SRR6834187 | 3.99 | 96.32 |
| SRR6834188 | 3.99 | 96.32 |
| SRR6834189 | 0 | 92 |
| (Sato et al. 2022) | DRR393956 | 4.99 | 88 |
| DRR393957 | 0 | 88.01 |
| DRR393958 | 0 | 88.01 |
| DRR393959 | -5 | 88.01 |
| DRR393960 | -5 | 88.01 |
| Equatorial Indian Ocean (EqIO) & Bay of Bengal (BoB) | (Wu et al. 2021) | SRR13756225 | 9.99 | 85.99 | Illumina Single end | 1270 |
| SRR13756226 | 0 | 89.01 |
| SRR13756227 | 9.99 | 87 |
| SRR13756228 | 7.03 | 89.69 |
| SRR13756229 | 5 | 81 |
| SRR13756230 | 2 | 82 |
| SRR13756231 | 5 | 96.99 |
| SRR13756232 | 2.35 | 91.84 |
| SRR13756233 | 0 | 83 |
| SRR13756234 | 10 | 84.98 |
| SRR13756235 | 2.99 | 79.99 |
| SRR13756236 | 0.99 | 94.49 |
| SRR13756237 | 5 | 90.59 |
| SRR13756238 | 0 | 91.99 |
| SRR13756239 | 0 | 88 |
| SRR13756240 | 0.5 | 80 |
| SRR13756241 | 5.05 | 90.44 |
| SRR13756242 | 1 | 80 |
| SRR13756243 | 1.97 | 89.5 |
| SRR13756244 | 10 | 88.24 |
| SRR13756245 | 10 | 99.91 |
| Eastern Indian Ocean (SIO) | (Raes et al. 2018) | SRR7038454 | -22.694 | 112.785 | Illumina Paired end | 1815 |
| SRR7038525 | -17.57 | 118.06 |
| SRR7038527 | -18.11 | 117.32 |
| SRR7038531 | -14.06 | 123.39 |
| SRR7038532 | -25.65 | 112 |
| SRR7038533 | -29.34 | 113.43 |
| SRR7038785 | -31.67 | 115.16 |
| SRR7038788 | -13.06 | 125.06 |
| SRR7038790 | -19.76 | 115.14 |
| SRR7038827 | -18.11 | 117.32 |
| SRR7038903 | -15.37 | 121.39 |
| SRR7599347 | -19.76 | 115.14 |
| SRR7599354 | -31.67 | 115.16 |
| SRR7599558 | -22.694 | 112.785 |
| SRR7599581 | -17.57 | 118.06 |
| SRR7599620 | -18.11 | 117.32 |
| SRR7599700 | -14.06 | 123.39 |
| SRR7599704 | -13.06 | 125.06 |
| SRR7599706 | -15.37 | 121.39 |
| SRR7599708 | -30.92 | 114.26 |
| SRR7599717 | -29.34 | 113.43 |
| SRR7599720 | -25.65 | 112 |
| (Sato et al. 2022) | DRR393961 | -10 | 88 |
| DRR393962 | -10 | 88 |
| DRR393963 | -15 | 88.99 |
| DRR393964 | -15 | 88.99 |
| DRR393965 | -19.99 | 88 |
| DRR393966 | -19.99 | 88 |
| DRR393967 | -19.99 | 88 |
| DRR393968 | -19.99 | 88 |
| DRR393969 | -8.99 | 105.74 |
| DRR393970 | -8.99 | 105.74 |
| DRR393971 | -12 | 107.12 |
| DRR393972 | -15 | 108.5 |
| DRR393973 | -15 | 108.5 |
| DRR393974 | -12 | 112.5 |
| DRR393975 | -12 | 112.5 |
| DRR393976 | -16 | 112 |
| DRR393977 | -16 | 112 |
| DRR393978 | -16 | 112 |
| DRR393979 | -20 | 111.5 |
| DRR393980 | -20 | 111.5 |
| DRR393981 | -22 | 111.25 |
| DRR393982 | -22 | 111.25 |
| DRR393983 | -24.01 | 111.02 |
| DRR393984 | -24.01 | 111.02 |
| DRR393985 | -23.4 | 109 |
| DRR393986 | -23.4 | 109 |
| DRR393987 | -23.4 | 109 |
| DRR393988 | -24.5 | 112.75 |
| DRR393989 | -24.5 | 112.75 |

**Table S2:** Cloning-based Sanger *nifH* sequences used in this study.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***nifH* cloning dataset** | **Reference** | **Accession No** | **Data type** | **Total Sequences** |
| Arabian Sea | (Mazard et al. 2004) | AY621706 - AY621747 | Cloned PCR products | 44 |
| AY620240 - AY620241 |
| Arabian Sea | (Bird et al. 2005) | AY800134 - AY800143 | Cloned PCR products | 10 |
| Arabian Sea | (Jayakumar et al. 2012) | JF429940 - JF429973 | Cloned PCR products | 34 |
| Arabian Sea | (Bird and Wyman 2013) | JX064466 - JX064499 | Cloned PCR products | 34 |
| Arabian Sea & Indian Ocean Subtropical Gyre | (Shiozaki et al. 2014) | AB928219 - AB928303 | Cloned PCR products | 84 |
| Bay of Bengal | (Löscher et al. 2020) | MT507504-MT507584 | Cloned PCR products | 81 |

**Supplementary Figures**

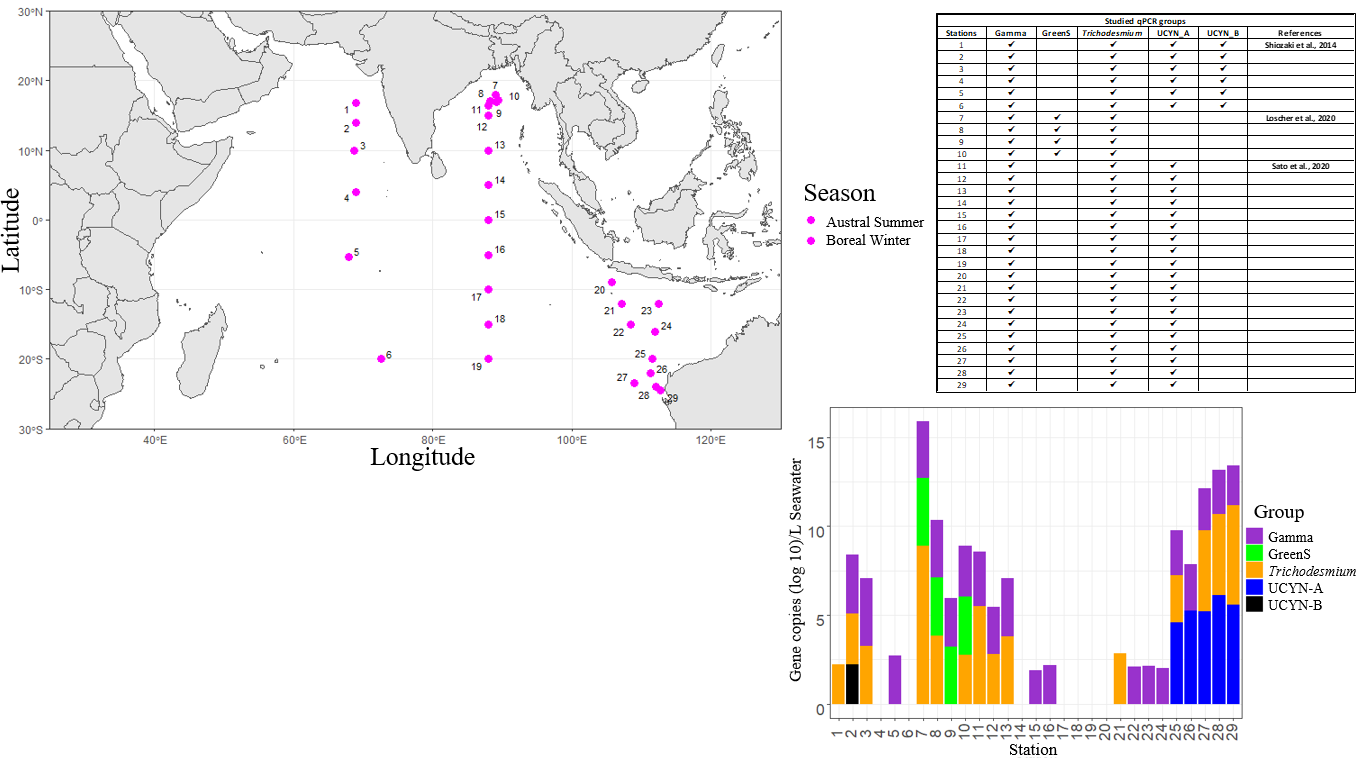
****

Fig. S1: Abundance of five diazotroph groups studied in the Indian Ocean (stations are represented in deep pink sampled during the austral summer and boreal winter)

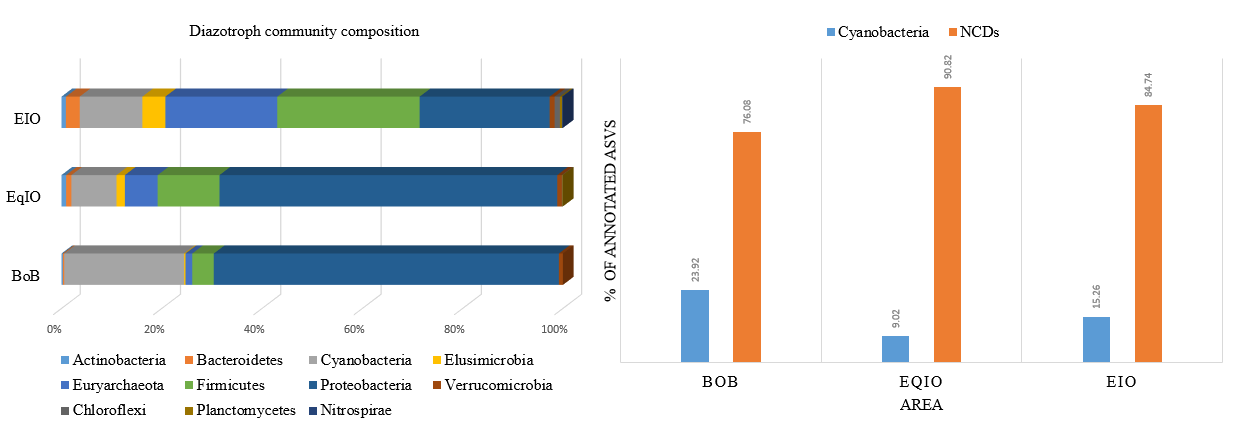


Fig. S2: Overall diazotroph community composition (A) and percentage contribution of cyanobacteria and NCDs (B) on the diazotroph community of Bay of Bengal (BoB), Equatorial Indian Ocean (EqIO) and Eastern Indian Ocean (EIO).

A picture containing timeline

Description automatically generated

Fig. S3: Heat map representing the top 15 ASVs of the diazotroph community in the Bay of Bengal (BoB) during intermonsoon (IM) boreal winter periods.

A picture containing calendar

Description automatically generated

Fig. S4: Heat map representing the top 15 ASVs of diazotroph community in the Equatorial Indian Ocean (EqIO) during the intermonsoon (IM) and boreal winter periods.

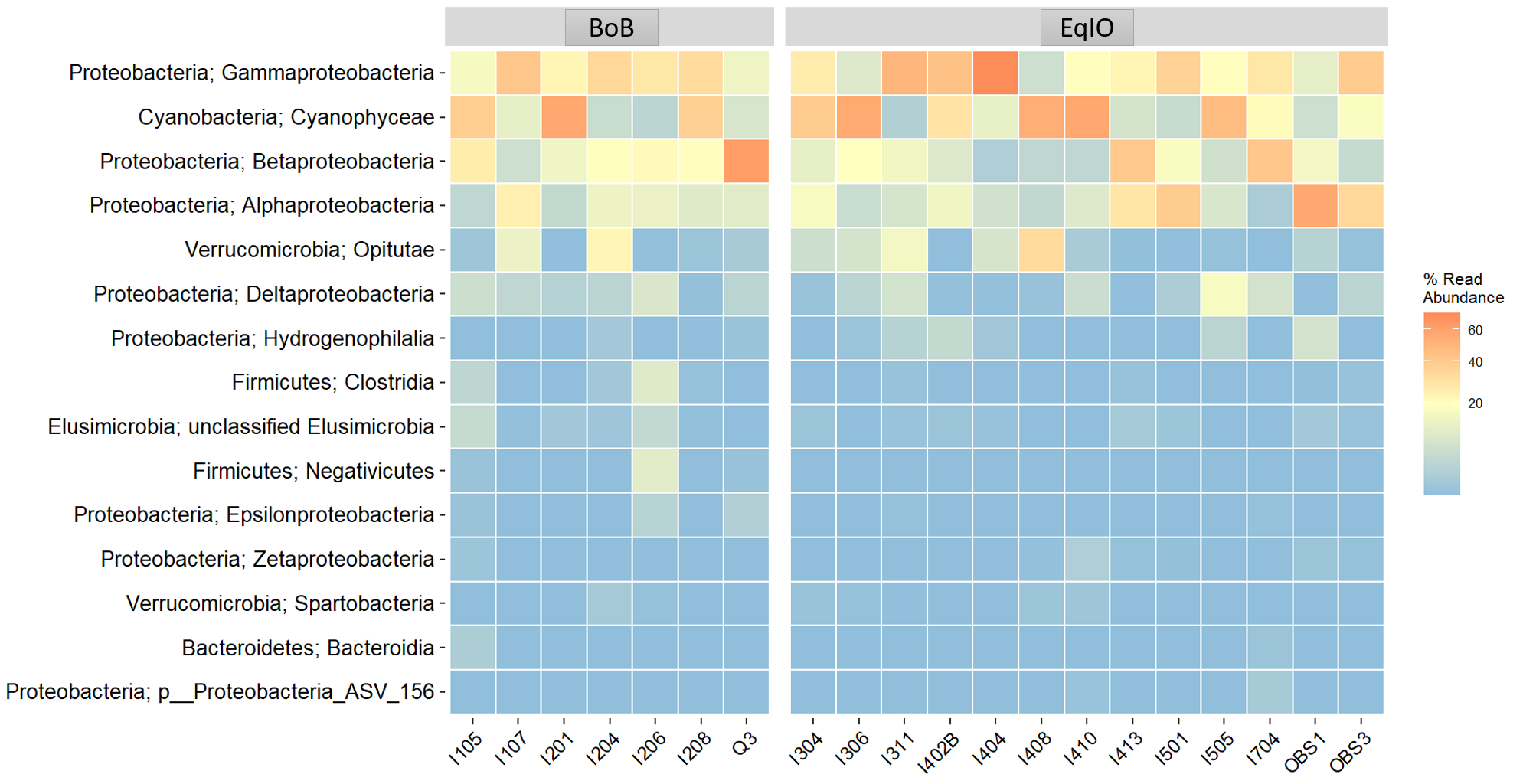


Fig. S5: Heat map representing the top 15 ASVs of diazotroph community in the Equatorial Indian Ocean (EqIO) and Bay of Bengal (BoB) during the early to late intermonsoon (IM) period.

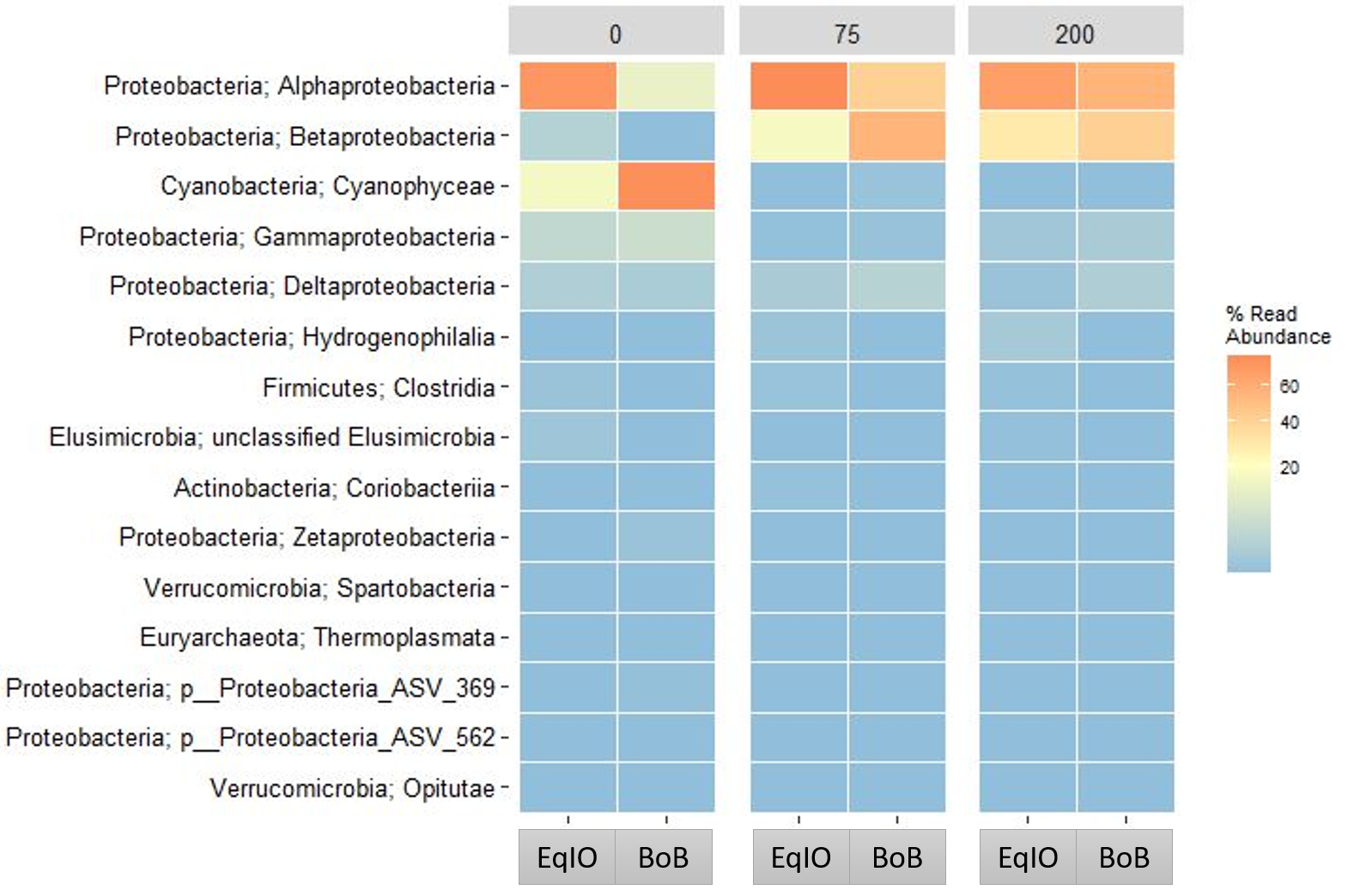


Fig. S6: Depth-wise top 15 ASVs diazotroph community distribution in the Equatorial Indian Ocean (EqIO) and Bay of Bengal (BoB) during the intermonsoon (IM) period. The upper panel represents the 0, 75, and 200 m depth.

Chart, timeline, bar chart, treemap chart

Description automatically generated

Fig. S7: Heat map representing the top 15 ASVs of the diazotrophs community in the Eastern Indian Ocean (EIO) during the austral summer and intermonsoon periods.

**Supplementary References**

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

Bird, C., J. Martinez Martinez, A. G. O’Donnell, and M. Wyman. 2005. Spatial Distribution and Transcriptional Activity of an Uncultured Clade of Planktonic Diazotrophic γ-Proteobacteria in the Arabian Sea. Appl Environ Microbiol **71**: 2079–2085. doi:10.1128/AEM.71.4.2079-2085.2005

Bird, C., and M. Wyman. 2013. Transcriptionally active heterotrophic diazotrophs are widespread in the upper water column of the Arabian Sea. FEMS Microbiol Ecol **84**: 189–200. doi:10.1111/1574-6941.12049

Brandes, J. A., A. H. Devol, T. Yoshinari, D. A. Jayakumar, and S. W. A. Naqvi. 1998. Isotopic composition of nitrate in the central Arabian Sea and eastern tropical North Pacific: A tracer for mixing and nitrogen cycles. Limnol. Oceanogr. **43**: 1680–1689. doi:10.4319/lo.1998.43.7.1680

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods **13**: 581–583. doi:10.1038/nmeth.3869

Gandhi, N., A. Singh, S. Prakash, R. Ramesh, M. Raman, M. S. Sheshshayee, and S. Shetye. 2011. First direct measurements of N 2 fixation during a *Trichodesmium* bloom in the eastern Arabian Sea: N 2 Fixation in the Arabian Sea. Global Biogeochem. Cycles **25**: 4014. doi:10.1029/2010GB003970

González, M. L., Molina, V., Florez-Leiva, L., Oriol, L., Cavagna, A. J., Dehairs, F., ... and Fernandez, C. 2014. Nitrogen fixation in the Southern Ocean: a case of study of the Fe-fertilized Kerguelen region (KEOPS II cruise). Biogeosci. Discuss. **11**: 17151-17185. doi: 10.5194/bgd-11-17151-2014

Jayakumar, A., M. M. D. Al-Rshaidat, B. B. Ward, and M. R. Mulholland. 2012. Diversity, distribution, and expression of diazotroph *nifH* genes in oxygen-deficient waters of the Arabian Sea. FEMS Microbiol Ecol **82**: 597–606. doi:10.1111/j.1574-6941.2012.01430.x

Kromkamp, J., M. De Bie, N. Goosen, J. Peene, P. Van Rijswijk, J. Sinke, and G. C. A. Duinevel. 1997. 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. doi:10.1016/S0967-0645(97)00015-5

Kumari, V. R., Ghosh, V. R. D., Rao, D. N., Krishna, M. S., and Sarma, V. V. S. S. 2022. Nitrogen fixation in the western coastal Bay of Bengal: Controlling factors and contribution to primary production. Reg. Stud. Mar. Sci. **53**: 102410. doi: 10.1016/j.rsma.2022.102410

Li, L., C. Wu, D. Huang, C. Ding, Y. Wei, and J. Sun. 2021. Integrating Stochastic and Deterministic Process in the Biogeography of N2-Fixing Cyanobacterium Candidatus Atelocyanobacterium Thalassa. Front. Microbiol. **12**: 654646. doi:10.3389/fmicb.2021.654646

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

Luo, Y.-W., and others. 2012. Database of diazotrophs in global ocean: Abundance, biomass and nitrogen fixation rates. Earth Syst. Sci. Data 4: 47–73. doi:10.5194/essd-4-47-2012

Mazard, S. L., N. J. Fuller, K. M. Orcutt, O. Bridle, and D. J. Scanlan. 2004. PCR Analysis of the Distribution of Unicellular Cyanobacterial Diazotrophs in the Arabian Sea. Appl Environ Microbiol **70**: 7355–7364. doi:10.1128/AEM.70.12.7355-7364.2004

Raes, E. J., L. Bodrossy, J. van de Kamp, A. Bissett, and A. M. Waite. 2018. Marine bacterial richness increases towards higher latitudes in the eastern Indian Ocean. Limnol Oceanogr Lett **3**: 10–19. doi:10.1002/lol2.10058

Sato, T., T. Shiozaki, F. Hashihama, M. Sato, A. Murata, K. Sasaoka, S. Umeda, and K. Takahashi. 2022. Low Nitrogen Fixation Related to Shallow Nitracline Across the Eastern Indian Ocean. JGR Biogeosciences **127**. doi:10.1029/2022JG007104

Sahoo, D., Saxena, H., Nazirahmed, S., Kumar, S., Sudheer, A. K., Bhushan, R., ... and Singh, A. 2021. Role of eddies and N 2 fixation in regulating C: N: P proportions in the Bay of Bengal. Biogeochemistry **155**: 413-429. doi: 10.1007/s10533-021-00833-4

Shiozaki, T., M. Ijichi, T. Kodama, S. Takeda, and K. Furuya. 2014. Heterotrophic bacteria as major nitrogen fixers in the euphotic zone of the Indian Ocean. Global Biogeochem. Cycles **28**: 1096–1110. doi:10.1002/2014GB004886

Tang, W., S. Wang, D. Fonseca-Batista, and others. 2019. Revisiting the distribution of oceanic N2 fixation and estimating diazotrophic contribution to marine production. Nat Commun **10**: 831. doi:10.1038/s41467-019-08640-0

Wu, C., J. Kan, H. Liu, L. Pujari, C. Guo, X. Wang, and J. Sun. 2019. Heterotrophic Bacteria Dominate the Diazotrophic Community in the Eastern Indian Ocean (EIO) during Pre-Southwest Monsoon. Microb Ecol **78**: 804–819. doi:10.1007/s00248-019-01355-1

Wu, C., J. Sun, H. Liu, W. Xu, G. Zhang, H. Lu, and Y. Guo. 2021. Evidence of the Significant Contribution of Heterotrophic Diazotrophs to Nitrogen Fixation in the Eastern Indian Ocean During Pre-Southwest Monsoon Period. Ecosystems. doi:10.1007/s10021-021-00702-z