Science Advances
AAAS

Supplementary Materials for

Nitrogen fixation in the widely distributed marine γ-proteobacterial diazotroph *Candidatus* Thalassolituus haligoni

Sonja A. Rose *et al.*

Corresponding author: Sonja A. Rose, s.rose@dal.ca; Julie LaRoche, julie.laroche@dal.ca

Sci. Adv. **10**, eadn1476 (2024)
DOI: 10.1126/sciadv.adn1476

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

Supplementary Text

Supplemental text S1: Culture Experiments under Prolonged Nitrate Concentrations and Various Carbon Sources

Various carbon substrates were tested to determine the isolate's maximum biomass and growth rate, in order to obtain sufficient material for biomolecular research and NFR measurements. The carbon substrates tested comprised of (i) a mixed carbon cocktail (MCC), (ii) individual components of the MCC, (iii) Succinate (Sigma Aldrich), (iv) Na-Pyruvate (Sigma Aldrich), and (v) L-malic acid (Sigma Aldrich). The MCC standing stock was made in a final volume of 200 mL consisting of 1 M D-Glucose (Sigma Aldrich), 1 M Na-Acetate (Sigma Aldrich), 1 M Lactate (Sigma Aldrich), and 1 M Fumarate (Sigma Aldrich). 20 mL of the MCC standing stock was then diluted into a final 1L volume of MilliQ water to create the working solution. Separate working stocks of Succinate, Na-Pyruvate, and L-Malic acid were prepared individually as 1 M solutions in a final volume of 50 mL of MilliQ. The carbon substrates were added to cultures for final concentrations of either 6.4-, or 64-mM carbon.

Cultures were prepared in 50 mL non-vented culture flasks containing artificial seawater (ASW) supplemented with modified f/2 media (6:1, N:P), resulting in a final NO_3^- concentration of 200 µM. Oxygen levels were measured within the cultures using a colorimetric resazurin dye (1:1000 final dilution), indicating oxygen concentrations as blue (fully oxic), pink (suboxic), or clear (anoxic). Cultures were prepared in biological triplicates each with respective blanks for every carbon source. Each condition was inoculated with 800 cells/µL. Cultures were maintained at 15°C under 12-hour light/dark cycles. Over 6 days, samples were collected twice daily for nutrient and cell biomass analysis. Cell density was measured on an Accuri C6 flow cytometer (BD) using a final dilution of 1:1000 SYBR Green II (Invitrogen). Nutrients were collected and stored at -20°C for later analysis on the Skalar Sann^{++} nutrient analyzer. Statistical analysis included a one-way ANOVA and Tukey post-hoc test ($\alpha = 0.05$) for growth rates and biomass.

Supplemental text S2: Preliminary N-free Culturing of *Cand.* T. haligoni

100 µL of an initial parental culture grown in ASW N-free f/2 media supplemented with 64 mM carbon of a mixed carbon cocktail (MCC) was transferred in a T-streak fashion onto a 1.2% ASW N-free f/2 agar plate + 64 mM MCC. The plate was incubated for 14 days at 15 °C on a 12-hour light/dark cycle. Once visible growth of individual colonies were observed, a single colony was selected and transferred into 50 mL of liquid culture. The liquid culture consisted of N-free f/2 ASW, supplemented with 64 mM MCC, 50 µL resazurin dye (1:1000 final dilution) and kept at 7 °C undisturbed for three weeks. After three weeks, growth within the culture was comparable to that of fixed N cultures ($\times 10^{10}$ cells/ L). Nutrients were taken at $T_{initial}$ and T_{final} and stored at -80 °C for later analysis; where $T_{initial}$ was prior to adding the colony and T_{final} was at the time of harvest.

To ensure the culture was not contaminated, 1 mL of culture was subject to the same DNA extraction protocol mentioned throughout this study and sent for *16S* rRNA sequencing. Sequences were then blasted against the *16S* rRNA sequence of *Cand.* T. haligoni with 100% identity. The remaining culture was aliquoted as follows: 1:1 glycerol stock and stored at -80 °C, 10 mL filtered onto a 0.2 µm pre-combusted glass-fibre filter for POC/PN analysis, 10 mL centrifuged down for biomolecular analysis and stored at -80 °C and the remaining used as inoculant for new cultures. The biomolecular sample was later extracted and digested for discovery proteomics using methods supplied throughout this study.

Supplemental text S3: Ecotype Detection of *Cand*. T. haligoni

Ecotype detection was carried out using samples where *nifH* amplicon yielded more than 2 reads to reduce the risk of false positive due to amplification and sequencing errors. *NifH* ASVs were aligned using MUSCLE (codon) in MEGAX and exported. Using a minimum-spanning network, PopART was then used for haplotype mapping (*123, 146*). Epsilon values were set to zero to ensure

the high stringency of the model. The Tajima's D statistic and nucleotide diversity (pi) were run to determine for any selective pressures (positive, negative or neutral); where a D statistic less than zero suggests a recent selective sweep, and a nucleotide (pi) diversity value is a measure of the difference in genetic variation. These tests were further confirmed through MEGA-X, testing for selective pressures ds/dn. Visualisation of synonymous to non-synonymous mutations was analysed in MEGA-X, and plotted in R using *ggplot2 (98, 147)*.

Supplemental text S4: Liquid Chromatography Tandem Mass Spectrometry Protein Extraction and Digestion

Proteins were extracted according to *(130)* apart from flash freezing of samples. Extract concentrations were measured using the Micro BCA Protein Assay Kit (Thermo-Scientific) on the SYNERGY H1 microplate reader (BioTek). Once concentrations were determined, protein extracts were then frozen at -80 ºC. Protein samples were thawed and aliquoted to give 50 ug of protein for each digestion. Samples were first reduced in 5 mM Dithiothreitol (DTT) [500 mM DTT in 50 mM Ammonium Bicarbonate] and incubated for 1 hour at 37 ºC at 350 rpm. Samples were then cooled and alkylated in 15 mM Iodoacetamide (IAM) [500 mM IAM in 50 mM Ammonium Bicarbonate] for 30 minutes in the dark. Another 5 mM DTT was added to samples and vortexed at room temperature, after which, 2.5 µL of 12% Phosphoric acid was added. S-trap buffer [80% aqueous methanol in 100 mM TEAB, pH 7.1] was then added to samples in a ratio of 1:7 (200 µL buffer: 20-30 µL sample). TEAB was acidified using 85% Phosphoric acid. Samples were then loaded onto S-Traps (Protifi) in 600 µL increments and washed 10 x with 600 µL S-trap buffer and 1 final rinse with 80% methanol. Columns were then moved to a clean tube and 125 µL of digestion buffer [50 mM TEAB with 1:25 wt:wt trypsin] and centrifuged at 2000 g for 1 minute with the eluent re-applied for proper saturation of the filter. Samples were then incubated for 16 hours at 37ºC. Samples were then re-dissolved in 1% formic acid, 3% acetonitrile for a final concentration of 0.5 ug/ul. Peptide concentrations were then measured using the Micro BCA Protein Assay Kit (Thermo Scientific) and measured on the SYNERGY H1 microplate reader (BioTek). Peptide extracts were then desalted using 50 mg HyperSep C18 SPE Cartridges (Thermo Scientific).

Supplemental text S5: Liquid Chromatography Tandem Mass Spectrometry Sample de-salting

Peptide extracts were then desalted using 50 mg HyperSep C18 SPE Cartridges (Thermo Scientific). Cartridges were primed with 0.5 mL of methanol then 0.5 mL of 50% acetonitrile followed by equilibration with 1 mL 0.1% TFA. Samples were diluted with 0.1% TFA to a volume of ~0.2 mL and loaded onto cartridges by positive pressure from a syringe and flow through loaded a second time. Using a vacuum manifold, the cartridges were washed 3 times with 1 mL of 0.1% TFA. Peptides were eluted with 2 x 0.2 mL 50% acetonitrile, 0.1% FA, then 0.1 mL 70% acetonitrile, 0.1% FA with positive pressure from a syringe. The eluent was then brought to dryness in a speed vac as described above.

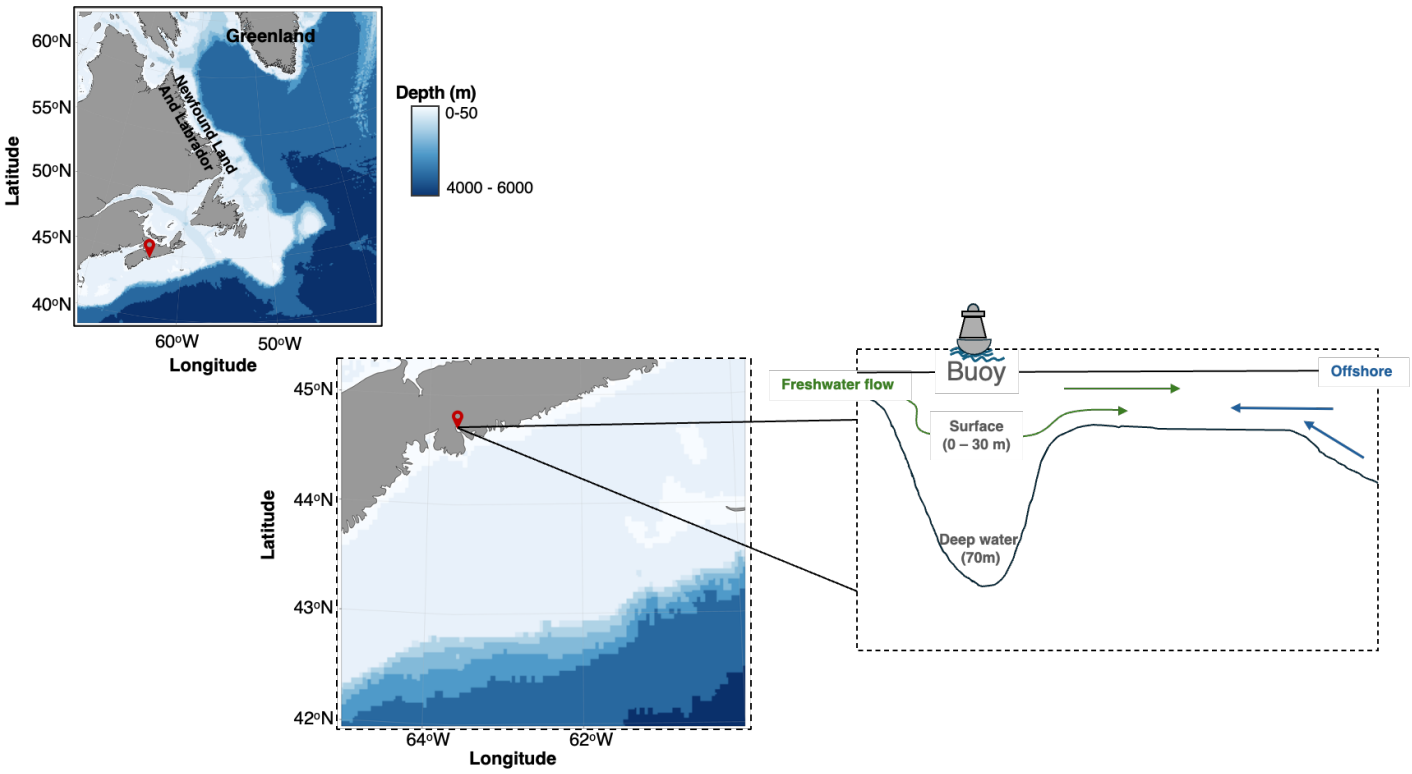

Fig. S1. Sample collection site and Campus Buoy location within the Bedford Basin, Halifax, NS, Canada of *Candidatus* Thalassolituus haligoni. Freshwater flow from the Sackville river is shown in green, and coastal water inflow shown in blue. Total depth of the sampling point within the basin is 70 m depth, and the surface waters from 0 - 30 m depth.

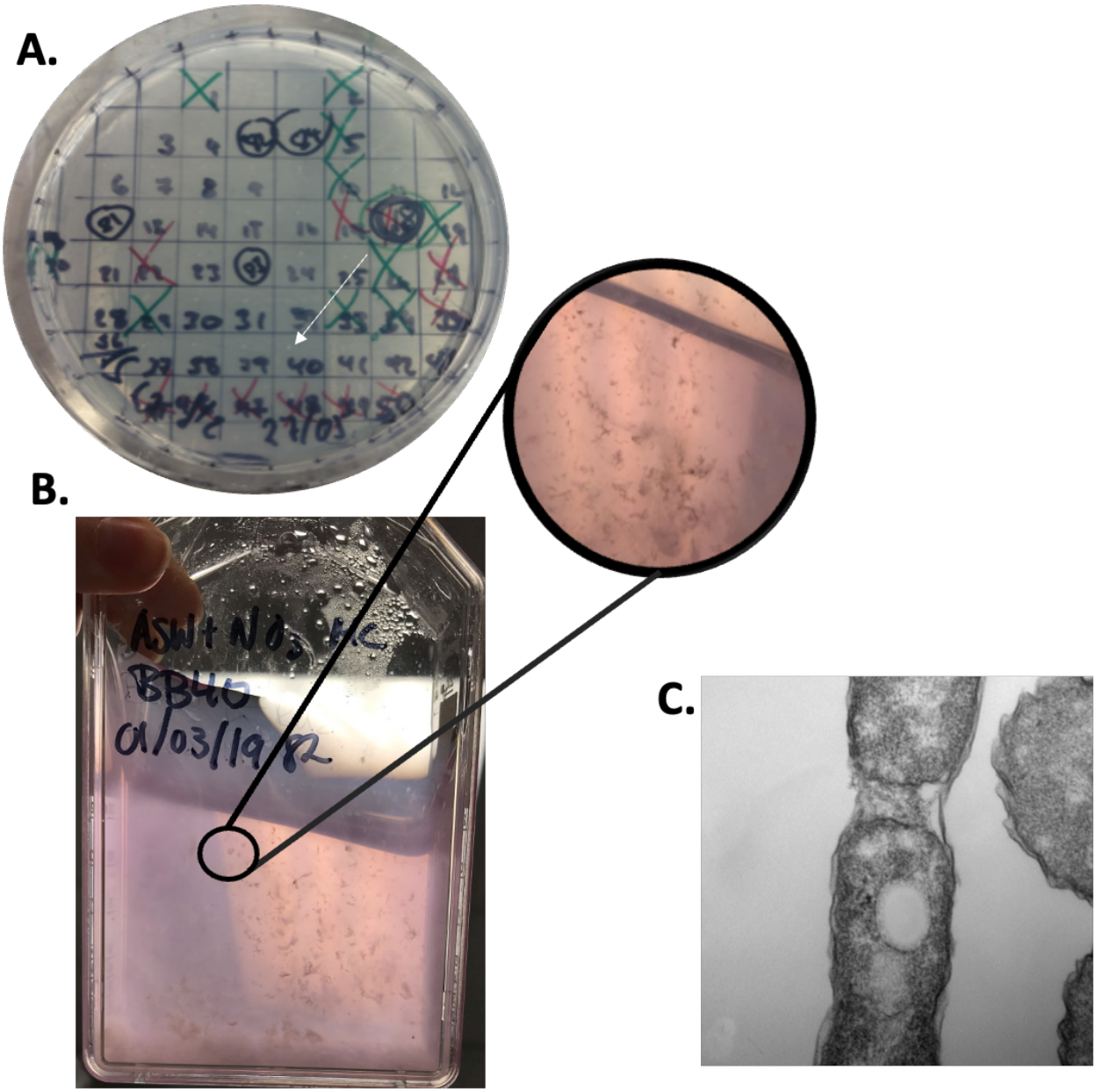

Fig. S2.

Additional observed growth characteristics of *Cand.* T. haligoni. **A)** Sorted cell colonies from enrichment treatment of positive *nifH* PCR identification. Cell number 40 is *Cand.* T. haligoni isolate and indicated by a white arrow. **B)** Self-aggregation of isolate in stationary phase. Cultures grown under suboxic N-limited conditions with a mixed carbon cocktail. **C)** Biofilm like substance secreted from *Cand.* T. haligoni under deplete NO_3^- conditions detected using TEM. Imagine obtained at $\times 10^5$ magnification.

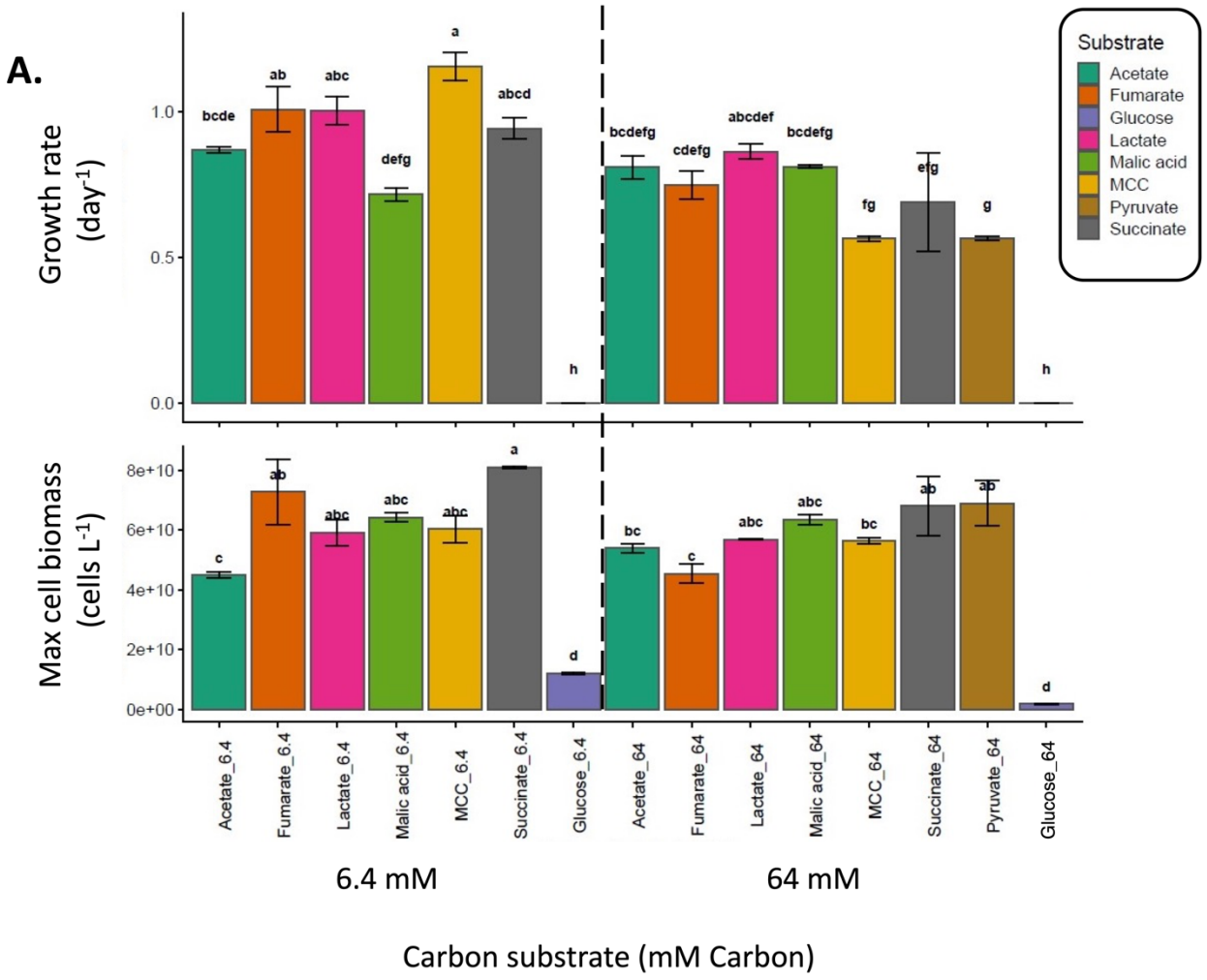

Fig. S3.

Growth of *Cand.* T. haligoni amended with diverse carbon compounds. Growth rate (day $^{-1}$) and biomass (cell L^{-1}) of isolate grown with nitrate (220 μM) and a range of carbon compounds, including a mixed carbon cocktail (MCC). All carbon compounds were supplied to triplicate culture replicates. One-way ANOVA was conducted on cell biomass and growth rate independently with Tukey post hoc analysis (letters). Carbon compounds that share the same letter indicate non-significant relationship.

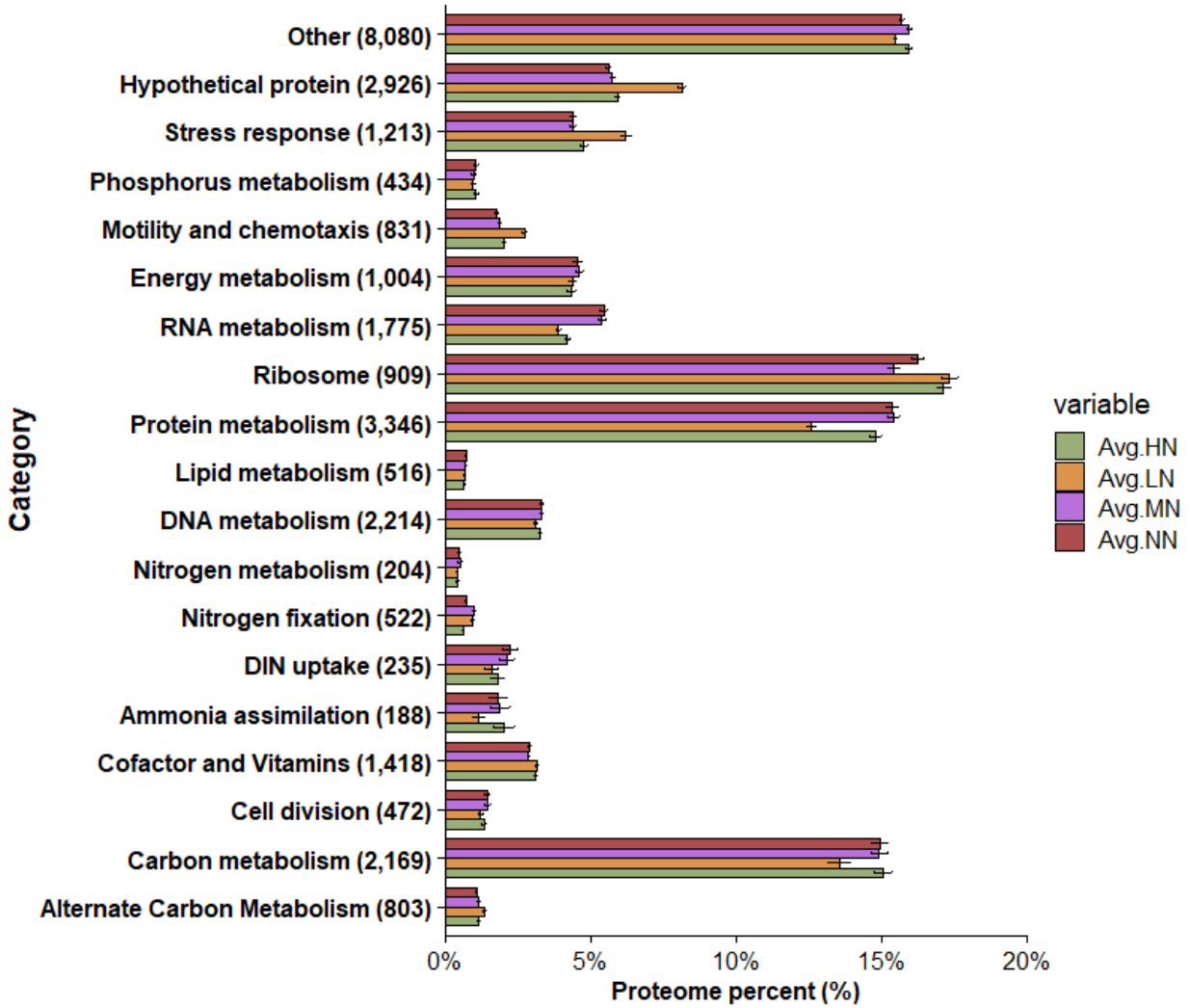

Fig. S4.

Percent of the detected proteome dedicated to specific functions in *Cand.* T. haligoni under changing NO_3^- conditions. Mass fractions calculated according to *(132)*. Numbers in brackets indicate the number of peptides associated with each function. Nitrogen metabolism category includes cyanate hydrolysis, arginine biosynthesis, and folate biosynthesis. DIN uptake includes Urea uptake, and nitrate uptake.

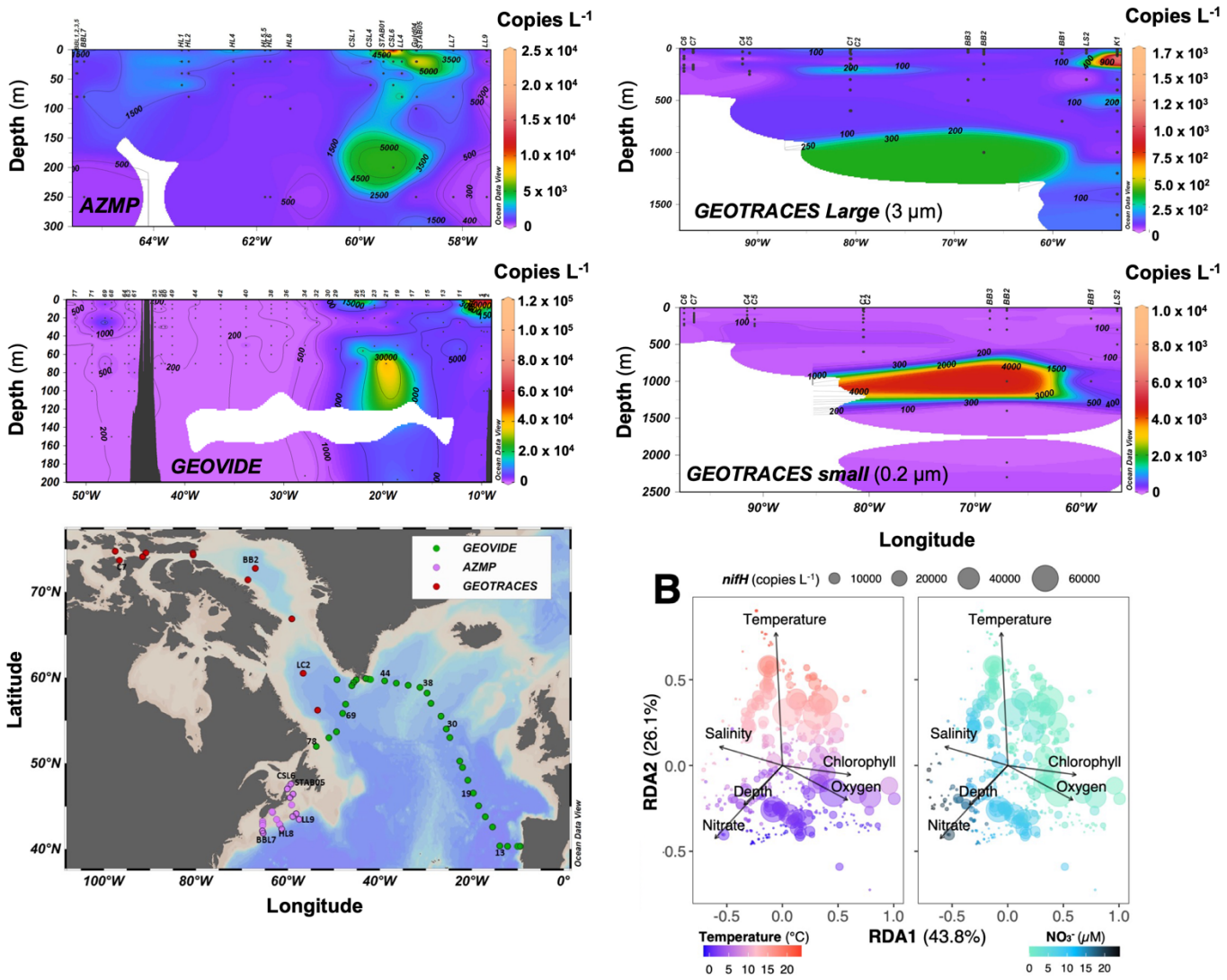

Fig. S5.

***nifH* copies per litre of seawater section plots and RDA analysis of *Cand.* T. haligoni. A)** Section plots of *Cand.* T. haligoni *nifH* qPCR copies along transects from Northern Atlantic and Canadian Arctic expeditions. Note that scales for qPCR copy numbers are not identical. The lower map displays cruise transects viewed in upper section plots. **B)** RDA analysis of *Cand.* T. haligoni *nifH* qPCR results and environmental data with results coloured by temperature (left) and nitrate concentrations (right). Data for qPCR and nutrients are provided within Data S3.

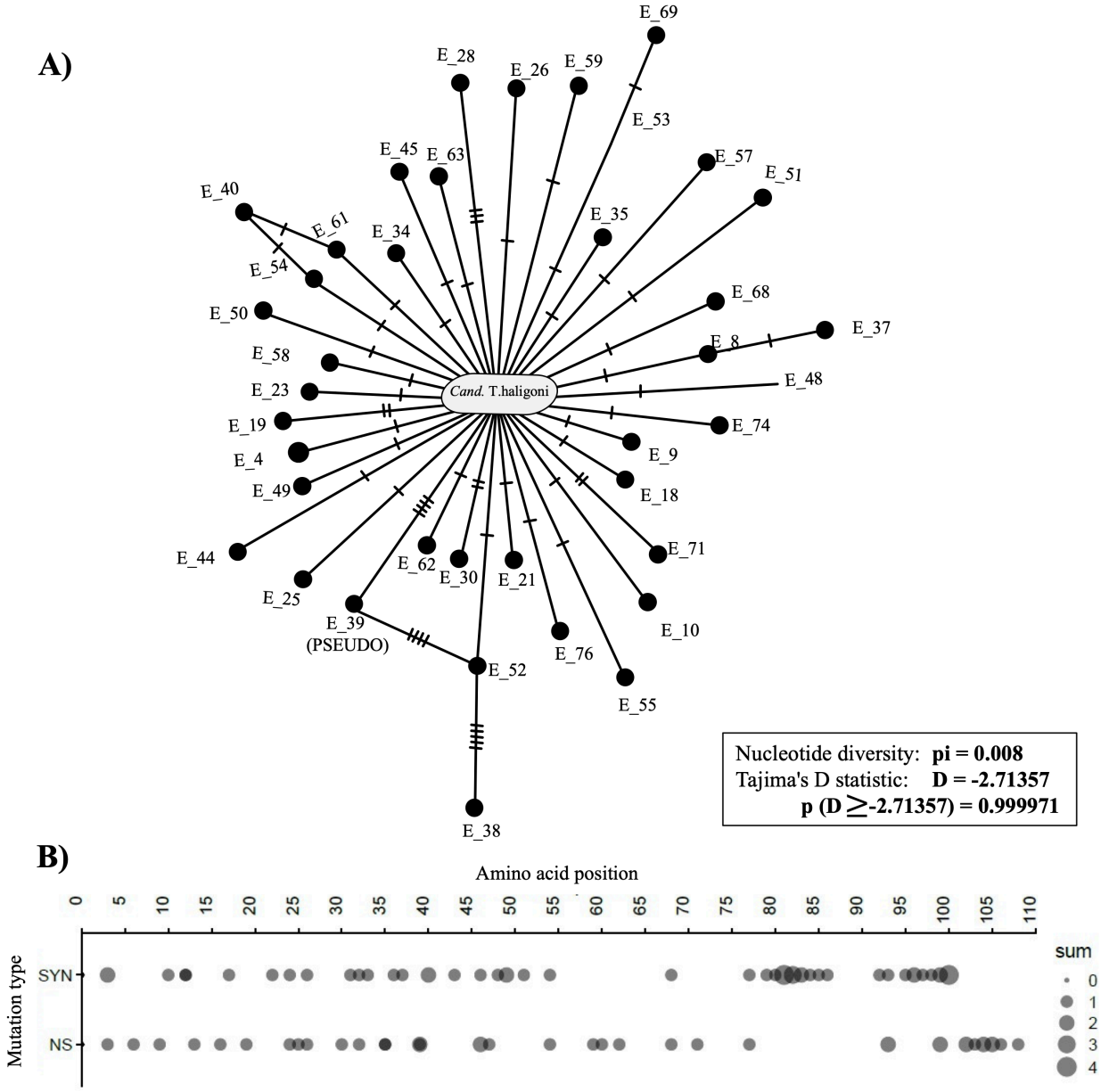

Fig. S6.

Putative ecotype detection of *Cand.* Thalassolituus haligoni. **A)** *nifH* minimum spanning network of ASVs with 98% identity to isolate's *nifH* ASV signature. Analysis limited to ASVs with >2 reads. Hash lines indicate the number of nucleotide differences between the ASV and *Cand.* T. haligoni. Minimum spanning network and statistics calculated in POPART (*145, 147*). **B)** Cumulative amino acid differences for synonymous (SYN) or non-synonymous (NS) changes between the *nifH* of *Cand.* T. haligoni and closely related ASVs in panel A. Circle sizes indicate the number of ASVs with a mutation at a given amino acid position (either synonymous or non-synonymous).

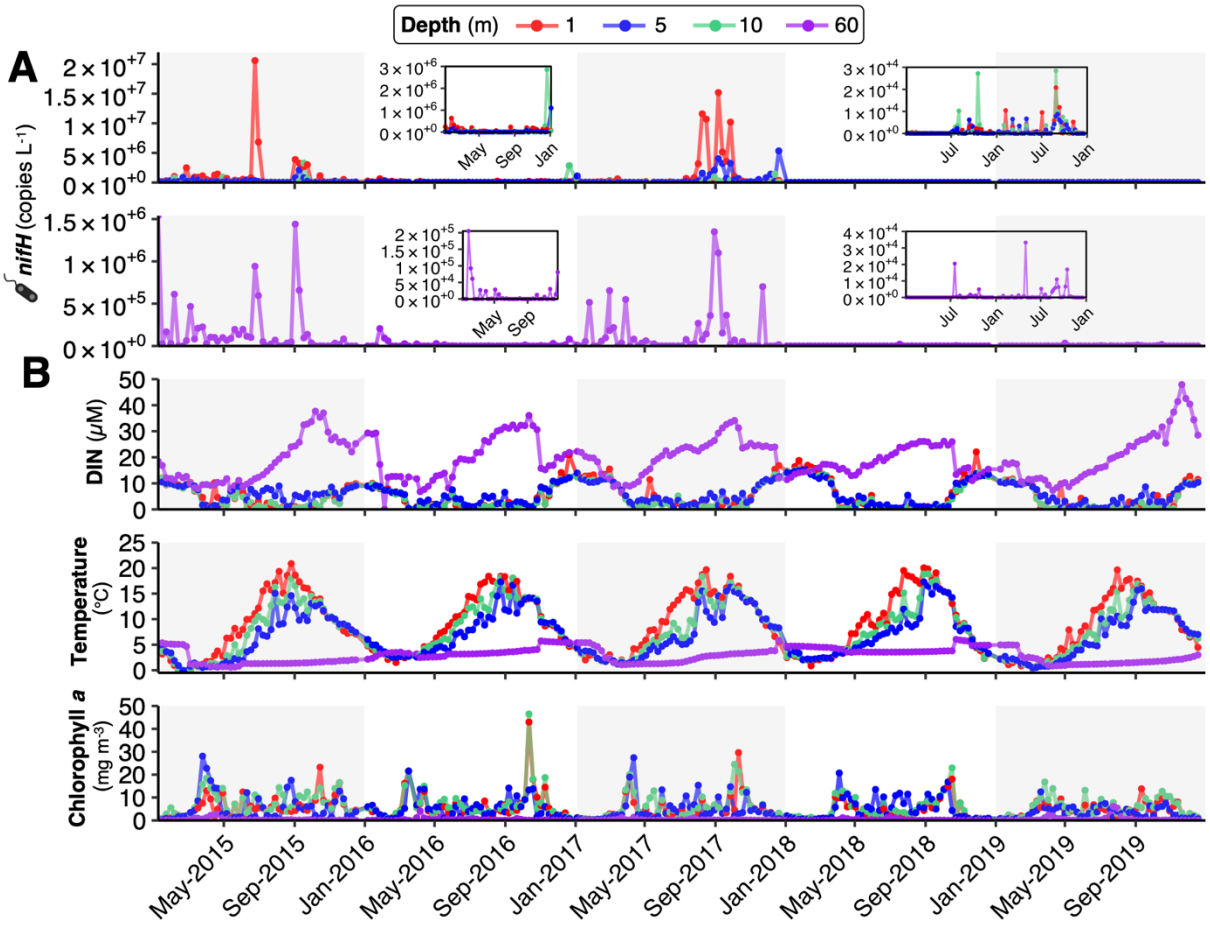

Fig. S7.
Seasonal patterns for *Cand. T. haligoni* within the coastal NWA. Water samples are from the Bedford Basin time series for the years 2015 –2019. A) *Cand.* T. haligoni *nifH* gene copies per litre for 1, 5, 10, and 60 m depths. **B**) Coordinated weekly dissolved inorganic Nitrogen (DIN), temperature, and chlorophyll *a* concentrations. White and grey highlighting shows yearly intervals. *NifH* copies per litre for 2018 and 2019 are from size fractionated data that were summed together.

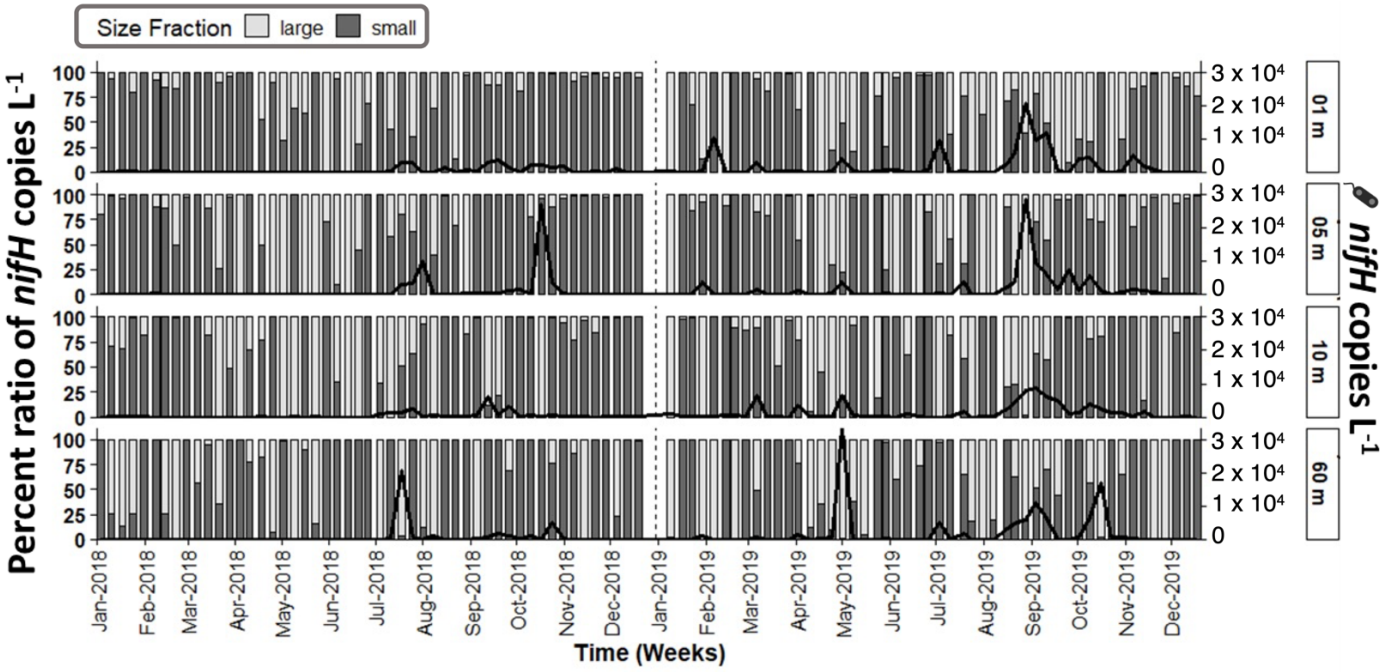

Fig. S8.
Size fraction association ratio (%) of *Cand.* T. haligoni from Bedford Basin, Halifax (NS). Data are from 2018 and 2019 at depths 1, 5, 10, and 60 m. Ratio percent calculated by taking the copies L^{-1} of small size fraction and dividing by the sum of copies L^{-1} for both size fractions. Large size fraction = 0.2 - 3 µm; small size fraction = 0.2 µm. Line represents the sum of both size fraction abundances.

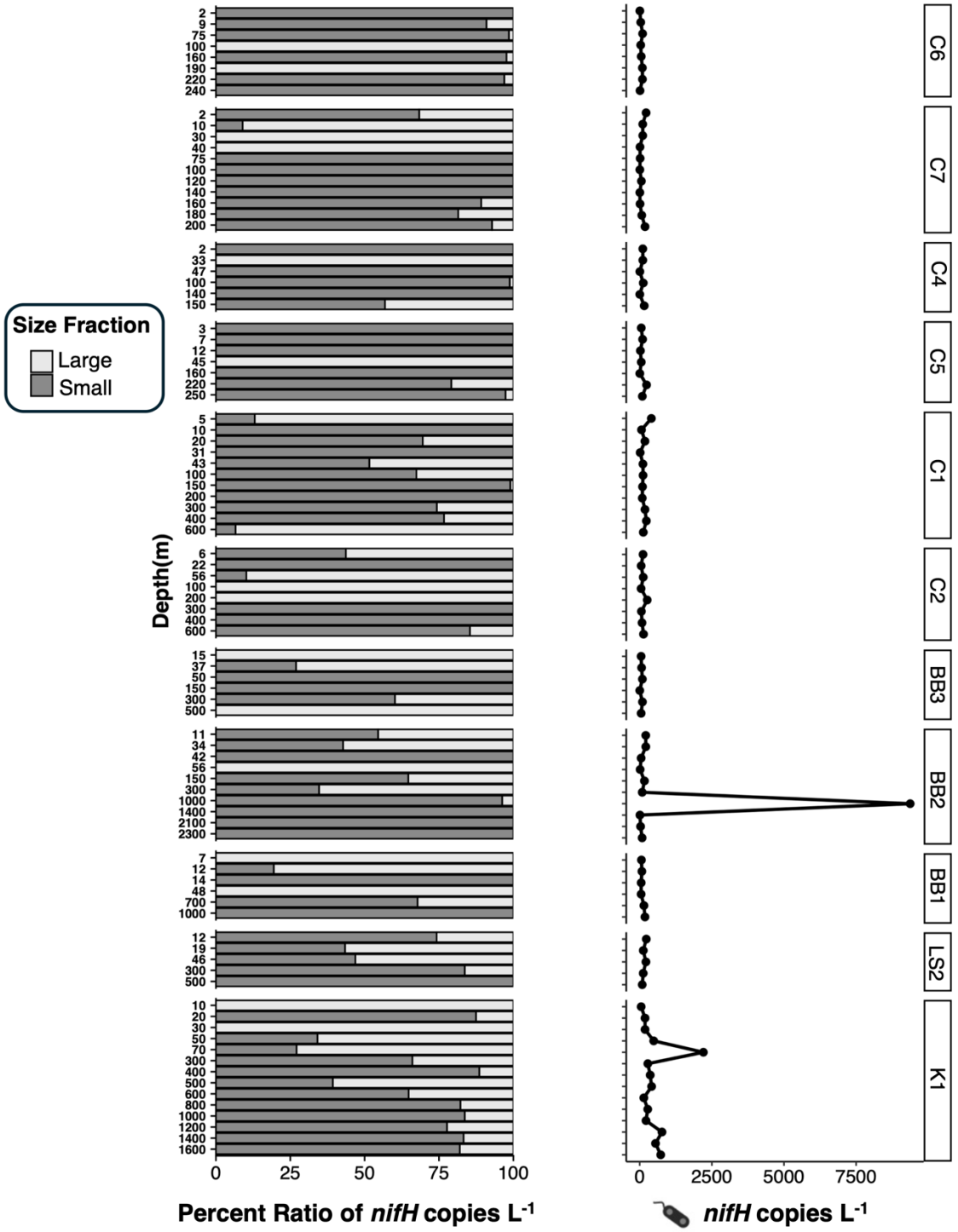

Fig. S9.

Size fraction association ratio (%) of *Cand.* T. haligoni with corresponding *nifH* copies L^{-1} from GEOTRACES Arctic cruise at indicated stations and depth. Ratio percent calculated by taking the copies L^{-1} of small size fraction and dividing by the sum of copies L^{-1} for both size fractions. Line represents the sum of the abundance in both fractions. Latitude and Longitude of each station are as follows (*116*): BB1(66.8555, -59.0573), BB2 (72.7511, -67.0000), BB3 (71.4109, -68.5960), C1 (74.5213, -80.5740), C2 (74.3143, -80.4973), C4 (74.1223, -91.5109), C5 (74.5388, -90.8024), C6 (74.7596, -97.4522), C7 (73.6729), K1 (56.175, -53.542) and LS2 (60.441, -56.535).

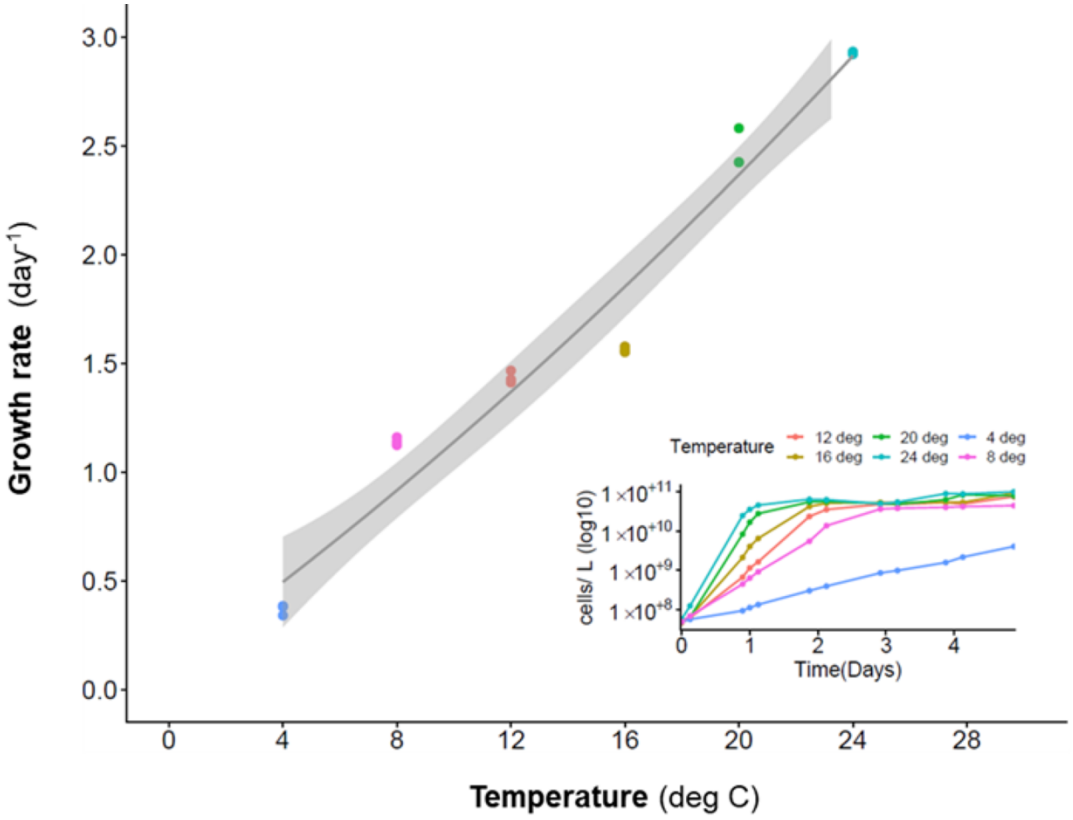

Fig. S10.
Growth rates (day^{-1}) of *Cand.* T. haligoni under various temperatures (°C). The grey line indicates the linear line of best fit with 95% confidence interval. Inset plot corresponds to the growth curves for selected temperatures. Cultures were grown using 6.4 mM C of L-malic acid with 220 µM NO_3^- conditions.

Description	HN			LN			HN		NN	
	10	11	12	25	26	27	28	30	31	33
Hemerythrin	n/d	0.38	0.31	0.30	0.51	2.30	n/d	n/d	n/d	n/d
Hemerythrin-like protein PA1673	n/d	n/d	-0.63	1.02	0.16	1.27	0.71	0.21	0.27	0.22
Hemerythrin-like protein PA1673	-0.55	-1.27	-0.73	-0.03	0.47	2.17	-0.67	-0.50	0.95	0.14
Hemerythrin-like protein PA1673	-0.63	0.18	0.30	-1.33	2.06	-1.45	0.01	0.06	0.56	0.25
Hemoglobin-like protein HbO	1.36	-0.61	-1.04	0.69	0.79	1.62	-0.80	-0.68	-0.65	-0.67
Cytochrome b	-0.90	-0.46	0.44	-1.20	-0.56	-1.07	1.67	0.65	0.13	1.29
Cytochrome b	n/d	-0.06	-0.05	-0.87	-0.84	-1.09	0.87	1.11	1.51	0.69
Alginate biosynthesis protein AlgZ/FimS	n/d	-0.43	-0.73	-0.45	-0.74	-0.69	-0.19	0.92	1.67	1.58

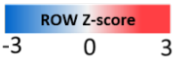

Fig. S11.

Heatmap of row z-scores of hypothesized O_2 protection mechanisms used in N_2 fixation for *Cand.* T. haligoni. Data was collected in biological replicates (indicated by numbers below treatment) across various NO_3^- treatments: High (HN), Low (LN), Re-addition of NO_3^- (NN) and 12 hours after re-addition (MN). NO_3^- concentrations and days data was collected include: HN = 100 µM (day 2), LN = 2 µM (Day 8), NN = 100 µM (Day 9) and MN = 50 µM (Day 9.5). Samples where protein was not detected are indicated by n/d.

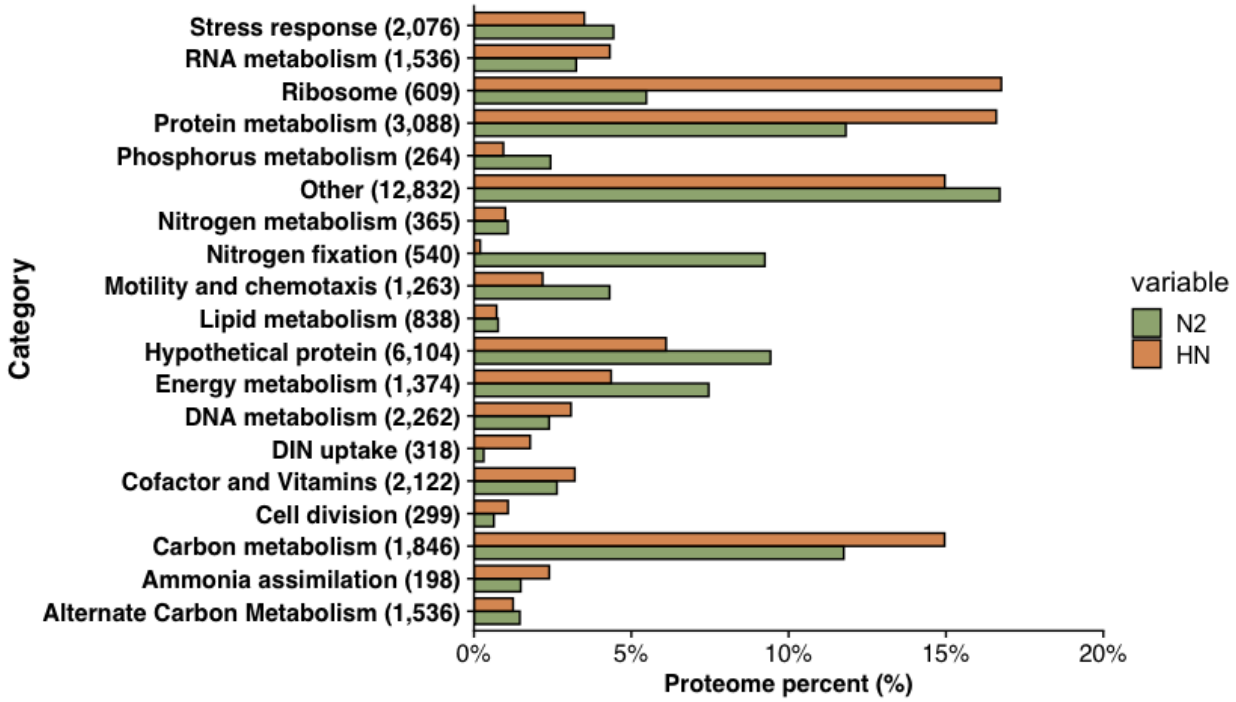

Fig. S12.

Percent of the detected proteome dedicated to specific functions in *Cand.* T. haligoni under high nitrate (HN) and N-free (N2) growth conditions. Mass fractions calculated according to *(132)*. Numbers in brackets indicate the number of peptides associated with each function. Nitrogen metabolism category includes cyanate hydrolysis, arginine biosynthesis, and folate biosynthesis. DIN uptake includes Urea uptake, and nitrate uptake. The HN condition is one of the replicates from the HN condition shown in fig S4 and fig 4. The N_2 growth condition proteomics data is from a single culture and is preliminary data. Refer to supplemental text S2 on further details of the N_2 culture.

Table S1.

Genomic features of *Cand.* T. haligoni. Genomic features of interest annotated from the RAST server (*143*).

Genome features	features
Genome size	4.37 Mbp
%GC content	53.4
Number of contigs	1
Number of subsystems	462
Number of CDS	4019
Number of RNA genes	61
	#CDS
Cell wall and capsule	**132**
Capsular and extracellular polysaccharides	44
Cell wall and capsule – no subcategory	55
Gram-negative cell wall components	32
Gram positive cell wall components	1
Motility and chemotaxis	**108**
Motility and chemotaxis – no subcategory	30
Flagellar motility in Prokaryota	78
Fatty acids, Lipids, and Isoprenoids	171
Phospholipids	27
Triacylglycerols	4
Fatty acids	74
Polyhydroxybutyrate metabolism	33
Isoprenoids	33
Nitrogen metabolism	**69**
Cyanate hydrolysis	3
Nitrogen fixation	26
Nitrosative stress	2
Nitrate and Nitrite reduction	14
Ammonia assimilation	24
Metabolism of Aromatic compounds	**127**
Salicylate ester degradation	5
Phenol hydroxylase	12
Quinate degradation	1
Biphenyl degradation	16
Benzoate degradation	11
p-hydroxybenzoate degradation	1
Catechol branch of beta-ketoadipate pathway	3
Salicylate and gentisate catabolism	16
4-hydroxyphenylacetic acid catabolic pathway	16
N-heterocyclic aromatic compound degradation	2
Central meta-cleavage pathway	25
Aromatic Amin catabolism	9
Gentisate degradation	10
Carbohydrates	**216**
Central carbohydrate metabolism	106
Di- and oligosaccharides	9
One-carbon metabolism	9
Organic acids	12
Butanol Biosynthesis	18
Acetolacetate Synthase	3
Lactate fermentation	7
Acetyl-CoA fermentation to Butyrate	30
CO_2 fixation	1
Glycerol and Glycerol-3-phosphate uptake and utilization	7
Carbon storage regulator	1
VCO266- no subcategory	1
Polysaccharides	7
Monosaccharides	5
Oxygen metabolism	**154**
Oxidative stress	53
Dioxygenases	15
Fermentation	58
Anaerobic respiratory reductases	28
Total Genes	**4019**

Table S2

ANI genome comparison to *Cand.* T. haligoni (GC content= 51.8%; isolate genome size = 4.2 Mbp). Cut off reliability for genus classification for ANI values was set to be >75%, and species classification cut off of >95 %, otherwise amino-acid identity (AAI) was used (*33, 34*).

Organism	ANI value (%)	GC content (%)	Total length (Mbp)
HBD gamma 11	74.31	51.96	3.61
HBD gamma 8	74.40	53.08	4.16
HBD gamma 4	73.57	54.97	3.51
HBD gamma 1	71.46	52.43	4.46
HBD gamma 9	71.65	52.51	4.29
HBD gamma 6	71.61	53.20	4.20
Thalassolituus HI10	70.66	48.56	5.67
Thalassolituus oliveorans	71.22	46.63	3.92
Parathalassolituus penaei gen.nov	74.70	55.43	4.44
Arc-Gamma-03 MAG	99.00	53.90	4.28

Table S3.
AAI genome comparison matrix of *Thalassolituus* sp., γ-HBD's, *P. penaei* and *Cand.* T. haligoni. Cut off acceptable AAI values for genus and species classification was set to >60% and >90% (*33, 34*).

	Genomes						
	Cand. T. haligoni	Arc-Gamma-03	*P. penaei*	HBD Gamma 08	HBD Gamma 11	HBD Gamma 04	*T. oleivorans*
Cand. T. haligoni	**100**	99	69	67	67	61	61
Arc-Gamma-03	99	**100**	69	67	67	61	62
P. penaei	69	69	**100**	65	65	60	61
HBD Gamma 08	67	67	65	**100**	80	60	61
HBD Gamma 11	67	67	65	80	**100**	60	61
HBD Gamma 04	61	61	60	60	60	**100**	65
T. oleivorans	64	63	62	61	62	65	**100**

Table S4.

Nitrogen fixation measurements from the GEOVIDE Cruise 2014. Rate measurements and dominant *nifH* contributors obtained from Fonseca-Batista et al. (*65).* NFR measurements and dominant contributor to *nifH* can be found in the supplemental material, figs 5 and 6 of Fonseca-Batista et al. *(65*) while *Cand.* T. haligoni *nifH* copies L^{-1} were used from this study at the sample site location NFR measurements were obtained from. Note "n/a' refers to depths where *nifH* sequencing was

not conducted.

Station number	depth (m)	Lat (deg N)	Lon (deg E)	NFR (nmol N L-1 day-1)	*Cand.* T. haligoni *nifH* copies L-1	Dominant contributor to *nifH* (*nifH* sequencing)
1	6	40.3	-10	4.8	1.25E+05	n/a
	11	40.3	-10	7.1	2.05E+04	n/a
	25	40.3	-10	2.5	1.02E+04	n/a
	34	40.3	-10	1.2	2.83E+03	n/a
2	11	40.3	-9.5	4.7	2.50E+03	n/a
	31	40.3	-9.5	2.8	4.50E+03	n/a
13	30	41.4	-13.9	1	3.92E+03	Verrucomicrobia and *Cand*. T. haligoni
	75	41.4	-13.9	3.9	4.11E+03	n/a
21	18	46.5	-19.7	<DL	3.04E+02	Verrucomicrobia
	25	46.5	-19.7	4.9	3.46E+02	n/a
	60	46.5	-19.7	2	5.04E+04	Gamma-proteobacteria (*Cand*. T. haligoni)

Table S5.
Enrichment treatments for diazotrophic isolation (*91*).

	Nutrient added (final concentration)				
	NH_4NO_3 (2 µM)[1]	**NaH_2PO_4 (400 nM)**	**$FeCl_3$ (4 nM)**	**D-Glucose (2 µM)**	**Thiosulfate (1 mM)**
Control	-	-	-	-	-
1.	-	+	+	-	-
2.	-	+	+	+	-
3.	-	+	+	+	+
4.	-	-	-	-	+
5.	+	+	+	-	-

1) Final concentration in sample

Table S6.

Delmont et al. (*21*) MAG assigned names versus names assigned in this study. Names in this study are based on the proposed name in the supplemental material of *(21)*.

Delmont et al. (21)	This study.
TARA_PSE-93-MAG_00027	HBD Gamma 06
TARA_PON_MAG_00023	HBD Gamma 09
TARA_PON_109_MAG_00015	HBD Gamma 04
TARA_PSE_93_MAG_00053	HBD Gamma 11
TARA_PSW_86_MAG_00031	HBD Gamma 08
TARA_IOS_50_MAG_00116	HBD Gamma 15
TARA_PSE_93_MAG_00042	HBD Gamma 07
TARA_AOS_82_MAG_0008	HBD Gamma 02
TARA_PON_109_MAG_00047	HBD Gamma 05
TARA_PON_109_MAG_00010	HBD Gamma 03
TARA_PSE_93_MAG_00126	HBD Gamma 13
TARA_AON_82_00085	HBD Gamma 12
TARA_AON_82_MAG_00185	HBD Gamma 14
TARA_AON_82_MAG_00052	HBD Gamma 10
TARA_ION_MAG_00014	HBD Gamma 01

Data S1. (separate file)

Data S1 Pangenome comparison of *Cand.* T. haligoni to γ-HBDs: A) RAST genome comparison of Arc-Gamma-03 to *Cand.* T. haligoni, **B)** core gene clusters of all compared MAGS, **C)** shared core genes between *Cand.* T. haligoni (Iso) and Arc-Gamma-03 (Arc), **D)** gene clusters only in *Cand.* T. haligoni. Note that Isolate refers to *Cand.* T. haligoni.

Data S2. (separate file)

Data S2 Proteomics data for *Cand.* T. haligoni under prolonged nitrate conditions: A) Recovery of total identified proteins in all samples, **B)** List of proteins detected (D) and not detected (ND) in *Cand.* T. haligoni proteome, **C)** Mass fraction categories, **D)** Relative protein abundance values and raw abundance data, **E)** Individual protein statistics from Fig 4A.

Data S3. (separate file)

Data S3 *NifH* quantitative PCR (qPCR) data and assay LOQ of *Cand.* T. haligoni: A*)* *nifH* qPCR(copies L^{-1}) of *Cand.* T. haligoni with Temperature, Depth, Salinity, Oxygen, Chlorophyll *a,* and Nitrate from each cruise. **B)** RDA mutations and significance from fig S3B, **C)** qPCR efficiencies calculated using LinReg, **D)** qPCR limit of quantification (LOQ).

Data S4. (separate file)

Data S4 Preliminary proteomic data of *Cand.* T. haligoni under N_2 conditions: *Cand.* T. haligoni N_2 preliminary culture data proteome. Data shows detected (D) and not detected (ND) proteins of a N-free culture (N_2) and high nitrate condition (HN; previously shown in Data S2) and protein relative abundance. HN samples were run as a control and for sample comparison with the N_2 condition.

Data S5. (separate file)

Data S5 Literature SRA references with metadata and *nifH* relative abundances of *Cand.* T. haligoni.

Data S6. (separate file)

Data S6 Media compounds and concentrations used to make modified ASW f/2 (6:1 N:P).

Data S7. (separate file)

Data S7 Nitrogen fixation rate measurement limit of quantification and detection calculations (*129)*.

REFERENCES AND NOTES

1. C. M. Moore, M. M. Mills, K. R. Arrigo, I. Berman-Frank, L. Bopp, P. W. Boyd, E. D. Galbraith, R. J. Geider, C. Guieu, S. L. Jaccard, T. D. Jickells, J. LaRoche, T. M. Lenton, N. M. Mahowald, E. Marañón, I. Marinov, J. K. Moore, T. Nakatsuka, A. Oschlies, M. A. Saito, T. F. Thingstad, A. Tsuda, O. Ulloa, Processes and patterns of oceanic nutrient limitation. *Nat. Geosci.* **6**, 701–710 (2013).

2. X. Zhang, B. B. Ward, D. M. Sigman, Global nitrogen cycle: Critical enzymes, organisms, and processes for nitrogen budgets and dynamics. *Chem. Rev.* **120**, 5308–5351 (2020).

3. K. R. Arrigo, Marine microorganisms and global nutrient cycles. *Nature* **437**, 349–355 (2005).

4. D. Karl, A. Michaels, B. Bergman, D. Capone, E. Carpenter, R. Letelier, F. Lipschultz, H. Paerl, D. Sigman, L. Stal. "Dinitrogen fixation in the world's oceans" in *The Nitrogen Cycle at Regional to Global Scales* (Springer Netherlands, 2002), pp. 47–98.

5. M. Caffin, T. Moutin, R. A. Foster, P. Bouruet-Aubertot, A. M. Doglioli, H. Berthelot, C. Guieu, O. Grosso, S. Helias-Nunige, N. Leblond, A. Gimenez, A. A. Petrenko, A. de Verneil, S. Bonnet, N2 Fixation as a dominant new N Source in the western tropical South Pacific Ocean (OUTPACE cruise). *Biogeosciences* **15**, 2565–2585 (2018).

6. T. D. Jickells, E. Buitenhuis, K. Altieri, A. R. Baker, D. Capone, R. A. Duce, F. Dentener, K. Fennel, M. Kanakidou, J. LaRoche, K. Lee, P. Liss, J. J. Middelburg, J. K. Moore, G. Okin, A. Oschlies, M. Sarin, S. Seitzinger, J. Sharples, A. Singh, P. Suntharalingam, M. Uematsu, L. M. Zamora, A reevaluation of the magnitude and impacts of anthropogenic atmospheric nitrogen inputs on the ocean. *Global Biogeochem. Cycles* **31**, 289–305 (2017).

7. T. H. Coale, V. Loconte, K. A. Turk-Kubo, B. Vanslembrouck, W. K. E. Mak, S. Cheung, A. Ekman, J.-H. Chen, K. Hagino, Y. Takano, T. Nishimura, M. Adachi, M. L. Gros, C. Larabell, J. P. Zehr, Nitrogen-fixing organelle in a marine alga. *Science* **384**, 217–222 (2024).

8. D. G. Capone, J. P. Zehr, H. W. Paerl, B. Bergman, E. J. Carpenter, *Trichodesmium*, a globally significant marine cyanobacterium. *Science* **276**, 1221–1229 (1997).

9. D. G. Capone, J. A. Burns, J. P. Montoya, A. Subramaniam, C. Mahaffey, T. Gunderson, A. F. Michaels, E. J. Carpenter, Nitrogen fixation by *Trichodesmium* spp.: An important source of new nitrogen to the tropical and subtropical North Atlantic Ocean. *Global Biogeochem. Cycles* **19**, (2005).

10. J. A. Sohm, E. A. Webb, D. G. Capone, Emerging patterns of marine nitrogen fixation. *Nat. Rev. Microbiol.* **9**, 499–508 (2011).

11. Y.-W. Luo, S. C. Doney, L. A. Anderson, M. Benavides, I. Berman-Frank, A. Bode, S. Bonnet, K. H. Boström, D. Böttjer, D. G. Capone, E. J. Carpenter, Y. L. Chen, M. J. Church, J. E. Dore, L. I. Falcón, A. Fernández, R. A. Foster, K. Furuya, F. Gómez, K. Gundersen, A. M. Hynes, D. M. Karl, S. Kitajima, R. J. Langlois, J. LaRoche, R. M. Letelier, E. Marañón, D. J. McGillicuddy, P. H. Moisander, C. M. Moore, B. Mouriño-Carballido, M. R. Mulholland, J. A. Needoba, K. M. Orcutt, A. J. Poulton, E. Rahav, P. Raimbault, A. P. Rees, L. Riemann, T. Shiozaki, A. Subramaniam, T. Tyrrell, K. A. Turk-Kubo, M. Varela, T. A. Villareal, E. A. Webb, A. E. White, J. Wu, J. P. Zehr, Database of diazotrophs in global ocean: Abundance, biomass and nitrogen fixation rates. *Earth Syst. Sci. Data* **4**, 47–73 (2012).
12. J. P. Zehr, L. A. McReynolds, Use of degenerate oligonucleotides for amplification of the nifH gene from the marine cyanobacterium *Trichodesmium thiebautii*. *Appl. Environ. Microbiol.* **55**, 2522–2526 (1989).
13. H. Farnelid, A. F. Andersson, S. Bertilsson, W. A. Al-Soud, L. H. Hansen, S. Sørensen, G. F. Steward, Å. Hagström, L. Riemann, Nitrogenase gene amplicons from global marine surface waters are dominated by genes of non-cyanobacteria. *PLOS ONE* **6**, e19223 (2011).
14. J. P. Zehr, M. T. Mellon, S. Zani, New nitrogen-fixing microorganisms detected in oligotrophic oceans by amplification of nitrogenase (*nifH*) genes. *Appl. Environ. Microbiol.* **64**, 3444–3450 (1998).
15. C. Martínez-Pérez, W. Mohr, C. R. Löscher, J. Dekaezemacker, S. Littmann, P. Yilmaz, N. Lehnen, B. M. Fuchs, G. Lavik, R. A. Schmitz, J. LaRoche, M. M. M. Kuypers, The small unicellular diazotrophic symbiont, UCYN-A, is a key player in the marine nitrogen cycle. *Nat. Microbiol.* **1**, 16163 (2016).
16. A. W. Thompson, R. A. Foster, A. Krupke, B. J. Carter, N. Musat, D. Vaulot, M. M. M. Kuypers, J. P. Zehr, Unicellular cyanobacterium symbiotic with a single-celled eukaryotic alga. *Science* **337**, 1546–1550 (2012).
17. H. Farnelid, J. Harder, M. Bentzon-Tilia, L. Riemann, Isolation of heterotrophic diazotrophic bacteria from estuarine surface waters. *Environ. Microbiol.* **16**, 3072–3082 (2014).
18. R. Langlois, T. Großkopf, M. Mills, S. Takeda, J. LaRoche, Widespread distribution and expression of gamma A (UMB), an uncultured, diazotrophic, γ-proteobacterial *nifH* phylotype. *PLOS ONE* **10**, e0128912 (2015).
19. J. P. Zehr, D. G. Capone, Changing perspectives in marine nitrogen fixation. *Science* **368**, (2020).
20. K. A. Turk-Kubo, M. R. Gradoville, S. Cheung, F. Cornejo-Castillo, K. J. Harding, M. Morando,

M. Mills, J. P. Zehr, Non-cyanobacterial diazotrophs: Global diversity, distribution, ecophysiology, and activity in marine waters. *FEMS Microbiol. Rev.* **47**, fuac046 (2022).
21. T. O. Delmont, C. Quince, A. Shaiber, Ö. C. Esen, S. T. Lee, M. S. Rappé, S. L. McLellan, S. Lücker, A. M. Eren, Nitrogen-fixing populations of planctomycetes and proteobacteria are abundant in surface ocean metagenomes. *Nat. Microbiol.* **3**, 804–813 (2018).
22. D. Bombar, R. W. Paerl, L. Riemann, Marine non-cyanobacterial diazotrophs: Moving beyond molecular detection. *Trends Microbiol.* **24**, 916–927 (2016).
23. M. Bentzon-Tilia, H. Farnelid, K. Jürgens, L. Riemann, Cultivation and isolation of N2-fixing bacteria from suboxic waters in the Baltic Sea. *FEMS Microbiol. Ecol.* **88**, 358–371 (2014).
24. D. Bombar, P. H. Moisander, J. W. Dippner, R. A. Foster, M. Voss, B. Karfeld, J. P. Zehr, Distribution of diazotrophic microorganisms and nifH gene expression in the Mekong River plume during intermonsoon. *Mar. Ecol. Prog. Ser.* **424**, 39–52 (2011).
25. W. Mohr, N. Lehnen, S. Ahmerkamp, H. K. Marchant, J. S. Graf, B. Tschitschko, P. Yilmaz, S. Littmann, H. Gruber-Vodicka, N. Leisch, M. Weber, C. Lott, C. J. Schubert, J. Milucka, M. M. M. Kuypers, Terrestrial-type nitrogen-fixing symbiosis between seagrass and a marine bacterium. *Nature* **600**, 105–109 (2021).
26. S. Cheung, J. Zehr, Gamma4: A genetically versatile Gammaproteobacterial *nifH* phylotype that is widely distributed in the North Pacific Ocean. *Environ. Microbiol.* **23**, 4246–4259 (2021).
27. C. Fernandez, L. Farías, O. Ulloa, Nitrogen fixation in denitrified marine waters. *PLOS ONE* **6**, e20539 (2011).
28. R. J. Langlois, D. Hümmer, J. LaRoche, Abundance of Gamma-A nifH genes in the Atlantic Ocean [dataset]. PANGAEA, doi:10.1594/PANGEAE.817863.
29. H. Halm, P. Lam, T. G. Ferdelman, G. Lavik, T. Dittmar, J. LaRoche, S. D'Hondt, M. M. Kuypers, Heterotrophic organisms dominate nitrogen fixation in the South Pacific Gyre. *ISME J.* **6**, 1238–1249 (2012).
30. C. R. Loescher, T. Großkopf, F. D. Desai, D. Gill, H. Schunck, P. L. Croot, C. Schlosser, S. C. Neulinger, N. Pinnow, G. Lavik, M. M. M. Kuypers, J. LaRoche, R. A. Schmitz, Facets of diazotrophy in the oxygen minimum zone waters off Peru. *ISME J.* **8**, 2180–2192 (2014).
31. L. Riemann, E. Rahav, U. Passow, H.-P. Grossart, D. de Beer, I. Klawonn, M. Eichner, M. Benavides, E. Bar-Zeev, Planktonic aggregates as hotspots for heterotrophic diazotrophy: The plot thickens. *Front. Microbiol.* **13**, 875050 (2022).

32. T. Shiozaki, Y. Nishimura, S. Yoshizawa, H. Takami, K. Hamasaki, A. Fujiwara, S. Nishino, N. Harada, Distribution and survival strategies of endemic and cosmopolitan diazotrophs in the Arctic Ocean, ISME J, 17, 1340–1350 (2023) https://doi.org/10.1038/s41396-023-01424-x.
33. M. R. Olm, A. Crits-Christoph, S. Diamond, A. Lavy, P. B. Matheus Carnevali, J. F. Banfield, Consistent metagenome-derived metrics verify and delineate bacterial species boundaries. *mSystems* **5**, e00731-19 (2020).
34. L. M. Rodriguez-R, K. T. Konstantinidis, Bypassing cultivation to identify bacterial species. *Microbe* **9**, 111–118 (2014).
35. M. Satomi, T. Fujii, "The family Oceanospirillaceae" in *The Prokaryotes: Gammaproteobacteria* (Springer, 2014), pp. 491–527.
36. M. M. Yakimov, L. Giuliano, R. Denaro, E. Crisafi, T. N. Chernikova, W.-R. Abraham, H. Luensdorf, K. N. Timmis, P. N. Golyshin, *Thalassolituus oleivorans* gen. nov., sp. nov., a novel marine bacterium that obligately utilizes hydrocarbons. *Int. J. Syst. Evol. Microbiol.* **54**, 141–148 (2004).
37. H. Chen, J. Dai, P. Yu, X. Wang, J. Wang, Y. Li, S. Wang, S. Li, D. Qiu, *Parathalassolituus penaei* gen. nov., sp. nov., a novel member of the family *Oceanospirillaceae* isolated from a coastal shrimp pond in Guangxi, PR, China. *Int. J. Syst. Evol. Microbiol.* **73**, (2023).
38. A. L. Müller, J. R. de Rezende, C. R. J. Hubert, K. U. Kjeldsen, I. Lagkouvardos, D. Berry, B. B. Jørgensen, A. Loy, Endospores of thermophilic bacteria as tracers of microbial dispersal by ocean currents. *ISME J.* **8**, 1153–1165 (2004).
39. K. D. Young, The selective value of bacterial shape. *Microbiol. Mol. Biol. Rev.* **70**, 660–703 (2006).
40. K. J. Kechris, J. C. Lin, P. J. Bickel, A. N. Glazer, Quantitative exploration of the occurrence of lateral gene transfer by using nitrogen fixation genes as a case study. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 9584–9589 (2006).
41. A. Koirala, V. S. Brözel, Phylogeny of nitrogenase structural and assembly components reveals new insights into the origin and distribution of nitrogen fixation across bacteria and archaea. *Microorganisms* **9**, 1662 (2021).
42. A. Nonaka, H. Yamamoto, N. Kamiya, H. Kotani, H. Yamakawa, R. Tsujimoto, Y. Fujita, Accessory proteins of the nitrogenase assembly, NifW, NifX/NafY, and NifZ, are essential for diazotrophic growth in the nonheterocystous cyanobacterium *Leptolyngbya boryana*. *Front. Microbiol.* **10**, 495 (2019).
43. R. Dixon, D. Kahn, Genetic regulation of biological nitrogen fixation. *Nat. Rev. Microbiol.* **2**,

621–631 (2004).
44. I. Klawonn, M. J. Eichner, S. T. Wilson, N. Moradi, B. Thamdrup, S. Kümmel, M. Gehre, A. Khalili, H.-P. Grossart, D. M. Karl, H. Ploug, Distinct nitrogen cycling and steep chemical gradients in *Trichodesmium* colonies. *ISME J* **14**, 399–412 (2019).
45. J. C. Setubal, P. Dos Santos, B. S. Goldman, H. Ertesvåg, G. Espin, L. M. Rubio, S. Valla, N. F. Almeida, D. Balasubramanian, L. Cromes, L. Curatti, Z. Du, E. Godsy, B. Goodner, K. Hellner-Burris, J. A. Hernandez, K. Houmiel, J. Imperial, C. Kennedy, T. J. Larson, P. Latreille, L. S. Ligon, J. Lu, M. Mærk, N. M. Miller, S. Norton, I. P. O'Carroll, I. Paulsen, E. C. Raulfs, R. Roemer, J. Rosser, D. Segura, S. Slater, S. L. Stricklin, D. J. Studholme, J. Sun, C. J. Viana, E. Wallin, B. Wang, C. Wheeler, H. Zhu, D. R. Dean, R. Dixon, D. Wood, Genome sequence of *Azotobacter vinelandii*, an obligate aerobe specialized to support diverse anaerobic metabolic processes. *J. Bacteriol.* **191**, 4534–4545 (2009).
46. W. Sabra, A. P. Zeng, H. Lünsdorf, W. D. Deckwer, Effect of oxygen on formation and structure of *Azotobacter vinelandii Alginate* and its role in protecting nitrogenase. *Appl. Environ. Microbiol.* **66**, 4037–4044 (2000).
47. R. K. Poole, S. Hill, Respiratory protection of nitrogenase activity in *Azotobacter vinelandii*—Roles of the terminal oxidases. *Biosci. Rep.* **17**, 303–317 (1997).
48. N. Padmaja, H. Rajaram, S. K. Apte, A novel hemerythrin DNase from the nitrogen-fixing cyanobacterium *Anabaena sp.* strain PCC 7120. *Arch. Biochem. Biophys.* **505**, 171–177 (2011).
49. C. Alvarez-Carreño, V. Alva, A. Becerra, A. Lazcano, Structure, function and evolution of the hemerythrin-like domain superfamily. *Protein Sci.* **27**, 848–860 (2018).
50. F. Mus, A. B. Alleman, N. Pence, L. C. Seefeldt, J. W. Peters, Exploring the alternatives of biological nitrogen fixation. *Metallomics* **10**, 523–538 (2018).
51. R. Navarro-González, C. P. McKay, D. N. Mvondo, A possible nitrogen crisis for Archaean life due to reduced nitrogen fixation by lightning. *Nature* **412**, 61–64 (2001).
52. E. S. Boyd, J. W. Peters, New insights into the evolutionary history of biological nitrogen fixation. *Front. Microbiol.* **4**, 201 (2013).
53. A. K. Garcia, D. F. Harris, A. J. Rivier, B. M. Carruthers, A. Pinochet-Barros, L. C. Seefeldt, B. Kaçar, Nitrogenase resurrection and the evolution of a singular enzymatic mechanism. *eLife* **12**, e85003 (2023).

54. A. Knapp, The sensitivity of marine N2 fixation to dissolved inorganic nitrogen. *Front. Microbiol.* **3**, 374 (2012).
55. M. A. Trainer, T. C. Charles, The role of PHB metabolism in the symbiosis of rhizobia with legumes. *Appl. Microbiol. Biotechnol.* **71**, 377–386 (2006).
56. D. Tec-Campos, C. Zuñiga, A. Passi, J. Del Toro, J. D. Tibocha-Bonilla, A. Zepeda, M. J. Betenbaugh, K. Zengler, Modeling of nitrogen fixation and polymer production in the heterotrophic diazotroph *Azotobacter vinelandii* DJ. *Metab. Eng. Commun.* **11**, e00132 (2020).
57. K. Mandon, F. Nazaret, D. Farajzadeh, G. Alloing, P. Frendo, Redox regulation in diazotrophic bacteria in interaction with plants. *Antioxidants* **10**, 880 (2021).
58. E. K. Heiniger, Y. Oda, S. K. Samanta, C. S. Harwood, How posttranslational modification of nitrogenase is circumvented in *Rhodopseudomonas palustris* strains that produce hydrogen gas constitutively. *Appl. Environ. Microbiol.* **78**, 1023–1032 (2012).
59. A. B. Alleman, J. W. Peters, Mechanisms for generating low potential electrons across the metabolic diversity of nitrogen-fixing bacteria. *Appl. Environ. Microbiol.* **89**, e0037823 (2023).
60. K. Mandon, N. Michel-Reydellet, S. Encarnación, P. A. Kaminski, A. Leija, M. A. Cevallos, C. Elmerich, J. Mora, Poly-β-hydroxybutyrate turnover in *Azorhizobium caulinodans* is required for growth and affects *nifA* expression. *J. Bacteriol.* **180**, 5070–5076 (1998).
61. J. M. Dubbs, F. Robert Tabita, Regulators of nonsulfur purple phototrophic bacteria and the interactive control of CO2 assimilation, nitrogen fixation, hydrogen metabolism and energy generation. *FEMS Microbiol. Rev.* **28**, 353–376 (2004).
62. C. Dingler, J. Kuhla, H. Wassink, J. Oelze, Levels and activities of nitrogenase proteins in *Azotobacter vinelandii* grown at different dissolved oxygen concentrations. *J. Bacteriol.* **170**, 2148–2152 (1988).
63. M. A. Saito, E. M. Bertrand, S. Dutkiewicz, V. V. Bulygin, D. M. Moran, F. M. Monteiro, M. J. Follows, F. W. Valois, J. B. Waterbury, Iron conservation by reduction of metalloenzyme inventories in the marine diazotroph *Crocosphaera Watsonii*. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 2184–2189 (2011).
64. F. Lv, Y. Zhan, W. Lu, X. Ke, Y. Shao, Y. Ma, J. Zheng, Z. Yang, S. Jiang, L. Shang, Y. Ma, L. Cheng, C. Elmerich, Y. Yan, M. Lin, Regulation of hierarchical carbon substrate utilization, nitrogen fixation, and root colonization by the Hfq/Crc/CrcZY genes in *Pseudomonas stutzeri*. *iScience* **25**, 105663 (2022).
65. D. Fonesca-Batista, X. Li, V. Riou, V. Michotey, F. Deman, F. Fripiat, S. Guasco, N. Brion,

N. Lemaitre, M. Tonnard, M. Gallinari, H. Planquette, F. Planchon, G. Sarthou, M. Elskens, L. Chou, F. Dehairs, Evidence of high N2 fixation rates in the temperate northeast Atlantic. *Biogeosciences* **16**, 999–1017 (2019).
66. Z. Shao, Y.-W. Luo, Controlling factors on the global distribution of a representative marine non-cyanobacterial diazotroph phylotype (Gamma A). *Biogeosciences* **19**, 2939–2952 (2022).
67. K. Turk, A. P. Rees, J. P. Zehr, N. Pereira, P. Swift, R. Shelley, M. Lohan, E. M. S. Woodward, J. Gilbert, Nitrogen fixation and nitrogenase (*nifH*) expression in tropical waters of the eastern North Atlantic. *ISME J.* **5**, 1201–1212 (2011).
68. P. H. Moisander, R. A. Beinart, I. Hewson, A. E. White, K. S. Johnson, C. A. Carlson, J. P. Montoya, J. P. Zehr, Unicellular cyanobacterial distributions broaden the oceanic N2 fixation domain. *Science* **327**, 1512–1514 (2010).
69. M. Simon, H.-P. Grossart, B. Schweitzer, H. Ploug, Microbial ecology of organic aggregates in aquatic ecosystems. *Aquat. Microb. Ecol.* **28**, 75–211 (2002).
70. E. Geisler, A. Bogler, E. Rahav, E. Bar-Zeev, Direct detection of heterotrophic diazotrophs associated with planktonic aggregates. *Sci. Rep.* **9**, 9288 (2019).
71. J. N. Pedersen, D. Bombar, R. W. Paerl, L. Riemann, Diazotrophs and N2-fixation associated with particles in coastal estuarine waters. *Front. Microbiol.* **9**, 2759 (2018).
72. H. Farnelid, K. Turk-Kubo, H. Ploug, J. E. Ossolinski, J. R. Collins, B. A. S. Van Mooy, J. P. Zehr, Diverse diazotrophs are present on sinking particles in the North Pacific Subtropical Gyre. *ISME J.* **13**, 170–182 (2019).
73. N. Lemaitre, H. Planquette, F. Planchon, G. Sarthou, S. Jacquet, M. I. García-Ibáñez, A. Gourain, M. Cheize, L. Monin, L. André, P. Laha, H. Terryn, F. Dehairs, Particulate barium tracing of significant mesopelagic carbon remineralisation in the North Atlantic. *Biogeosciences* **15**, 2289–2307 (2018).
74. F. M. Cornejo-Castillo, J. P. Zehr, Intriguing size distribution of the uncultured and globally widespread marine non-cyanobacterial diazotroph Gamma-A. *ISME J.* **15**, 124–128 (2021).
75. C. Lønborg, F. Baltar, C. Carreira, X. A. G. Morán, Dissolved organic carbon source influences tropical coastal heterotrophic bacterioplankton response to experimental warming. *Front. Microbiol.* **10**, 2807 (2019).
76. D. Li, J. Liu, R. Zhang, M. Chen, W. Yang, J. Li, Z. Fang, B. Wang, Y. Qiu, M. Zheng, N2 fixation impacted by carbon fixation via dissolved organic carbon in the changing Daya Bay, South China Sea. *Sci. Total Environ.* **674**, 592–602 (2019).

77. L. Bopp, O. Aumont, L. Kwiatkowski, C. Clerc, L. Dupont, C. Ethé, R. Séférian, A. Tagliabue, Diazotrophy as a key driver of the response of marine net primary productivity to climate change. *Biogeosciences* **19**, 4267–4285 (2022).
78. S. Bonnet, M. Benavides, F. A. C. Le Moigne, M. Camps, A. Torremocha, O. Grosso, C. Dimier, D. Spungin, I. Berman-Frank, L. Garczarek, F. M. Cornejo-Castillo, Diazotrophs are overlooked contributors to carbon and nitrogen export to the deep ocean. *ISME J.* **17**, 47–58 (2023).
79. C. Martínez-Pérez, W. Mohr, A. Schwedt, J. Dürschlag, C. M. Callbeck, H. Schunck, J. Dekaezemacker, C. R. T. Buckner, G. Lavik, B. M. Fuchs, M. M. M. Kuypers, Metabolic versatility of a novel N2-fixing alphaproteobacterium isolated from a marine oxygen minimum zone. *Environ. Microbiol.* **20**, 755–768 (2018).
80. M. Bentzon-Tilia, I. Severin, L. H. Hansen, L. Riemann, Genomics and ecophysiology of heterotrophic nitrogen-fixing bacteria isolated from estuarine surface water. *mBio* **6**, e00929 (2015).
81. K. A. Turk-Kubo, M. M. Mills, K. R. Arrigo, G. van Dijken, B. A. Henke, B. Stewart, S. T. Wilson, J. P. Zehr, UCYN-A/haptophyte symbioses dominate N2 fixation in the Southern California Current System. *ISME Commun.* **1**, 42 (2021).
82. D. Prévost, H. Antoun, L. M. Bordeleau, Effects of low temperatures on nitrogenase activity in sainfoin (*Onobrychis viciifolia*) nodulated by arctic rhizobia. *FEMS Microbiol. Ecol.* **3**, 205–210 (1987).
83. T. Shiozaki, A. Fujiwara, K. Inomura, Y. Hirose, F. Hashihama, N. Harada, Biological nitrogen fixation detected under Antarctic sea ice. *Nat. Geosci.* **13**, 729–732 (2020).
84. T. D. Niederberger, J. A. Sohm, J. Tirindelli, T. Gunderson, D. G. Capone, E. J. Carpenter, S. C. Cary, Diverse and highly active diazotrophic assemblages inhabit ephemerally wetted soils of the Antarctic Dry Valleys. *FEMS Microbiol. Ecol.* **82**, 376–390 (2012).
85. L. W. Von Friesen, L. Riemann, Nitrogen fixation in a changing arctic ocean: An overlooked source of nitrogen? *Front. Microbiol.* **11**, 596426 (2020).
86. J.-P. Bellenger, T. Wichard, Y. Xu, A. M. L. Kraepiel, Essential metals for nitrogen fixation in a free-living N2-fixing bacterium: Chelation, homeostasis and high use efficiency. *Environ. Microbiol.* **13**, 1395–1411 (2011).
87. R. Paerl, T. Hansen, N. Henriksen, A. Olesen, L. Riemann, N-fixation and related O2 constraints on model marine diazotroph Pseudomonas stutzeri BAL361. *Aquat. Microb. Ecol.* **81**, 125–136 (2018).
88. B. M. Robicheau, J. Tolman, E. M. Bertrand, J. LaRoche, Highly-resolved interannual phytoplankton community dynamics of the coastal Northwest Atlantic. *ISME Commun.* **2**, 38

(2022).
89. J. P. Zehr, J. B. Waterbury, P. J. Turner, J. P. Montoya, E. Omoregle, G. F. Steward, A. Hansen, D. M. Karl, Unicellular cyanobacteria fix N2 in the subtropical North Pacific Ocean. *Nature* **412**, 635–638 (2001).
90. R. R. L. Guillard, J. H. Ryther, Studies of marine planktonic diatoms: I. Cyclotella nana hustedt, and detonula confervacea (cleve) gran. *Can. J. Microbiol.* **8**, 229–239 (1962).
91. J.-M. Ratten, "The diversity, distribution and potential metabolism of non-cyanobacterial diazotrophs in the North Atlantic Ocean," thesis, Dalhousie University, Halifax, Nova Scotia, Canada (2017).
92. S. Koren, B. P. Walenz, K. Berlin, J. R. Miller, N. H. Bergman, A. M. Phillippy, Canu: Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* **27**, 722–736 (2017).
93. A. M. Eren, E. Kiefl, A. Shaiber, I. Veseli, S. E. Miller, M. S. Schechter, I. Fink, J. N. Pan, M. Yousef, E. C. Fogarty, F. Trigodet, A. R. Watson, Ö. C. Esen, R. M. Moore, Q. Clayssen, M. D. Lee, V. Kivenson, E. D. Graham, B. D. Merrill, A. Karkman, D. Blankenberg, J. M. Eppley, A. Sjödin, J. J. Scott, X. Vázquez-Campos, L. J. McKay, E. A. McDaniel, S. L. R. Stevens, R. E. Anderson, J. Fuessel, A. Fernandez-Guerra, L. Maignien, T. O. Delmont, A. D. Willis, Community-led, integrated, reproducible multi-omics with anvi'o. *Nat. Microbiol.* **6**, 3–6 (2021).
94. C. Rinke, P. Schwientek, A. Sczyrba, N. N. Ivanova, I. J. Anderson, J.-F. Cheng, A. Darling, S. Malfatti, B. K. Swan, E. A. Gies, J. A. Dodsworth, B. P. Hedlund, G. Tsiamis, S. M. Sievert, W.-T. Liu, J. A. Eisen, S. J. Hallam, N. C. Kyrpides, R. Stepanauskas, E. M. Rubin, P. Hugenholtz, T. Woyke, Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**, 431–437 (2013).
95. M. A. Campbell, W.-J. Chen, J. A. López, Molecular data do not provide unambiguous support for the monophyly of flatfishes (*Pleuronectiformes*): A reply to Betancur-R and Ortí. *Mol. Phylogenet. Evol.* **75**, 149–153 (2014).
96. M. D. Lee, GToTree: A user-friendly workflow for phylogenomics. *Bioinformatics* **35**, 4162–4164 (2019).
97. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLOS ONE* **5**, e9490 (2010).
98. S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).

99. R. C. Edgar, MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
100. M. Nei, S. Kumar, *Molecular Evolution and Phylogenetics* (Oxford Univ. Press, 2000).
101. I . Letunic, P. Bork, Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).
102. E . Karensti, S. G. Acinas, P. Bork, C. Bowler, C. De Varges, J. Raes, M. Sullivan, D. Arendt, F. Benzoni, J.-M. Claverie, M. Follows, G. Gorsky, P. Hingamp, D. Iudicone, O. Jaillon, S. Kandels-Lewis, U. Krzic, F. Not, H. Ogata, S. Peasant, E. G. Reynaud, C. Sardet, M. E. Sieracki, S. Speich, D. Velayoudon, J. Weissenbach, P. Wincker, A holistic approach to marine eco-systems biology. *PLOS Biol.* **9**, e1001177 (2011).
103. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie2. *Nat. Methods* **9**, 357–359 (2012).
104. D. Li, C.-M. Liu, R. Luo, K. Sadakane, T.-W. Lam, MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct *de Bruijn* graph. *Bioinformatics* **31**, 1674–1676 (2015).
105. M. Alonge, L. Lebeigle, M. Kirsche, K. Jenike, S. Ou, S. Aganezov, X. Wang, Z. B. Lippman, M. C. Schatz, S. Soyk, Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biol.* **23**, 258 (2022).
106. D. Hyatt, G.-L. Chen, P. F. LoCascio, M. L. Land, F. W. Larimer, L. J. Hauser, Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).
107. R. L. Tatusov, M. Y. Galperin, D. A. Natale, E. V. Koonin, The COG database: A tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* **28**, 33–36 (2000).
108. T. Aramaki, R. Blanc-Mathieu, H. Endo, K. Ohkubo, M. Kanehisa, S. Goto, H. Ogata, KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* **36**, 2251–2252 (2020).
109. J. Mistry, S. Chuguransky, L. Williams, M. Qureshi, G. A. Salazar, E. L. L. Sonnhammer, S. C. E. Tosatto, L. Paladin, S. Raj, L. J. Richardson, R. D. Finn, A. Bateman, Pfam: The protein families database in 2021. *Nucleic Acids Res.* **49**, D412–D419 (2021).
110. T. O. Delmont, A. M. Eren, Linking pangenomes and metagenomes: The *Prochlorococcus* metapangenome. *PeerJ* **6**, e4320 (2018).

111. D. H. Parks, M. Chuvochina, D. W. Waite, C. Rinke, A. Skarshewski, P.-A. Chaumeil, P. Hugenholtz, A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996–1004 (2018).
112. B. Buchfink, C. Xie, D. H. Huson, Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).
113. L. Pritchard, H. R. Glover, S. Humphris, J. G. Elphinstone, I. K. Toth, Genomics and taxonomy in diagnostics for food security: Soft-rotting enterobacterial plant pathogens. *Anal. Methods* **8**, 12–24 (2016).
114. J. Zorz, C. Willis, A. M. Comeau, M. G. I. Langille, C. L. Johnson, W. K. W. Li, J. LaRoche, Drivers of regional bacterial community structure and diversity in the Northwest Atlantic Ocean. *Front. Microbiol.* **10**, 281 (2019).
115. S. Haas, B. M. Robicheau, S. Rakshit, J. Tolman, C. K. Algar, J. LaRoche, D. W. R. Wallace, Physical mixing in coastal waters controls and decouples nitrification via biomass dilution. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2004877118 (2021).
116. B. M. Robicheau, J. Tolman, S. Rose, D. Desai, J. LaRoche, Marine nitrogen-fixers in the Canadian Arctic Gateway are dominated by biogeographically distinct noncyanobacterial communities. *FEMS Microbiol. Ecol.* **99**, fiad122 (2023).
117. S. Zani, M. T. Mellon, J. L. Collier, J. P. Zehr, Expression of nifH genes in natural microbial assemblages in Lake George, New York, detected by reverse transcriptase PCR. *Appl. Environ. Microbiol.* **66**, 3119–3124 (2000).
118. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, (2011).
119. A. M. Comeau, G. M. Douglas, M. G. I. Langille, Microbiome helper: A custom and streamlined workflow for microbiome research. *mSystems* **2**, e00127-16 (2017).
120. A. Amir, D. McDonald, J. A. Navas-Molina, E. Kopylova, J. T. Morton, Z. Z. Xu, E. P. Kightley, L. R. Thompson, E. R. Hyde, A. Gonzalez, R. Knight, Deblur rapidly resolves single-nucleotide community sequence patterns. *mSystems* **2**, e00191-16 (2017).
121. J. C. Gaby, D. H. Buckley, A comprehensive evaluation of PCR primers to amplify the nifH gene of nitrogenase. *PLOS ONE* **7**, e42149 (2012).
122. E. Bolyen, J. R. Rideout, M. R. Dillon, N. A. Bokulich, C. C. Abnet, G. A. Al-Ghalith, H. Alexander, E. J. Alm, M. Arumugam, F. Asnicar, Y. Bai, J. E. Bisanz, K. Bittinger, A. Brejnrod, C. J.

Brislawn, C. T. Brown, B. J. Callahan, A. M. Caraballo-Rodríguez, J. Chase, E. K. Cope, R. Da Silva, C. Diener, P. C. Dorrestein, D. M. Durall, C. Duvallet, C. F. Edwardson, M. Ernst, M. Estaki, J. Fouquier, J. M. Gauglitz, S. M. Gibbons, D. L. Gibson, A. Gonzalez, K. Gorlick, J. Guo, B. Hillmann, S. Holmes, H. Holste, C. Huttenhower, G. A. Huttley, S. Janssen, A. K. Jarmusch, L. Jiang, B. D. Kaehler, K. B. Kang, C. R. Keefe, P. Keim, S. T. Kelley, D. Knights, I. Koester, T. Kosciolek, J. Kreps, M. G. I. Langille, J. Lee, R. Ley, Y.-X. Liu, E. Loftfield, C. Lozupone, M. Maher, C. Marotz, B. D. Martin, D. McDonald, L. J. McIver, A. V. Melnik, J. L. Metcalf, S. C. Morgan, J. T. Morton, A. T. Naimey, J. A. Navas-Molina, L. F. Nothias, S. B. Orchanian, T. Pearson, S. L. Peoples, D. Petras, M. L. Preuss, E. Pruesse, L. B. Rasmussen, A. Rivers, M. S. Robeson, P. Rosenthal, N. Segata, M. Shaffer, A. Shiffer, R. Sinha, S. J. Song, J. R. Spear, A. D. Swafford, L. R. Thompson, P. J. Torres, P. Trinh, A. Tripathi, P. J. Turnbaugh, S. Ul-Hasan, J. J. J. van der Hooft, F. Vargas, Y. Vázquez-Baeza, E. Vogtmann, M. von Hippel, W. Walters, Y. Wan, M. Wang, J. Warren, K. C. Weber, C. H. D. Williamson, A. D. Willis, Z. Z. Xu, J. R. Zaneveld, Y. Zhang, Q. Zhu, R. Knight, J. G. Caporaso, Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).

123. C. Ramakers, J. M. Ruijter, R. H. L. Deprez, A. F. M. Moorman, Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci. Lett.* **339**, 62–66 (2003).

124. J. M. Ruijter, C. Ramakers, W. M. H. Hoogaars, Y. Karlen, O. Bakker, M. J. B. van den Hoff, A. F. M. Moorman, Amplification efficiency: Linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res.* **37**, e45 (2009).

125. W. Mohr, T. Großkopf, D. W. R. Wallace, J. LaRoche, Methodological underestimation of oceanic nitrogen fixation rates. *PLOS ONE* **5**, e12583 (2010).

126. T. Großkopf, W. Mohr, T. Baustian, H. Schunk, D. Gill, M. M. M. Kuypers, G. Lavik, R. A. Schmitz, D. R. Wallance, J. LaRoche, Doubling of marine dinitrogen-fixation rates based on direct measurements. *Nature* **488**, 361–364 (2012).

127. D. Fonseca-Batista, F. Dehairs, V. Riou, F. Fripiat, M. Elskens, F. Deman, N. Brion, F. Quéroué, M. Bode, H., Nitrogen fixation in the eastern Atlantic reaches similar levels in the Southern and Northern Hemisphere. *J. Geophys. Res. Oceans* **122**, 587–601 (2017).

128. R. M. Parr, S. A. Clements, “Intercomparison of enriched stable isotope reference materials for medical and biological studies” (International Atomic Energy Agency, 1991).

129. A. E. White, J. Granger, C. Selden, M. R. Gradoville, L. Potts, A. Bourbonnais, R. W. Fulweiler,

A. N. Knapp, W. Mohr, P. H. Moisander, C. R. Tobias, M. Caffin, S. T. Wilson, M. Benavides, S. Bonnet, M. R. Mulholland, B. X. Chang, A critical review of the 15N2 tracer method to measure diazotrophic production in pelagic ecosystems. *Limnol. Oceanogr. Methods* **18**, 129–147 (2020).
130. M. Wu, J. S. P. McCain, E. Rowland, R. Middag, M. Sandgren, A. E. Allen, E. M. Bertrand, Manganese and iron deficiency in Southern Ocean *Phaeocystis antarctica* populations revealed through taxon-specific protein indicators. *Nat. Commun.* **10**, 3582 (2019).
131. M. The, M. J. MacCoss, W. S. Noble, L. Käll, Fast and accurate protein false discovery rates on large-scale proteomics data sets with Percolator 3.0. *J. Am. Soc. Mass Spectrom.* **27**, 1719–1727 (2016).
132. J. S. P. McCain, A. E. Allen, E. M. Bertrand, Proteomic traits vary across taxa in a coastal Antarctic phytoplankton bloom. *ISME J.* **16**, 569–579 (2022).
133. H. Wickham, Create elegant data visualisations using the grammar of graphics; https://ggplot2.tidyverse.org/.
134. M. Morgan, M. Ramos, B. P. Maintainer, BiocVersion: Set the appropriate version of Bioconductor packages. Bioconductor version: Release (3.18), 2023. doi:10.18129/B9.bioc.BiocVersion.
135. H. Wickham, Welcome to the Tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).
136. J. Oksanen, Vegan community ecology package version 2.6-2 (2022).
137. A. Kassambara, kassambara/rstatix: Pipe-friendly framework for basic statistical tests version 0.7.2.999; https://rdrr.io/github/kassambara/rstatix/.
138. P. Massicotte, World map data from natural Earth; https://docs.ropensci.org/rnaturalearth/.
139. C. Wilke, Introduction to cowplot; https://wilkelab.org/cowplot/articles/introduction.html.
140. G. Yu, S. Xu, Scatterpie: Scatter pie plot; https://cran.r-project.org/web/packages/scatterpie/index.html.
141. D. Dunnington, B. Thorne, D. Hernangómez, Ggspatial: Spatial data framework for ggplot2; https://cran.r-project.org/web/packages/ggspatial/index.html.
142. P. Stothard, D. S. Wishart, Circular genome visualization and exploration using CGView. *Bioinformatics* **21**, 537–539 (2005).
143. R. K. Aziz, D. Bartels, A. A. Best, M. DeJongh, T. Disz, R. A. Edwards, K. Formsma, S. Gerdes, E. M. Glass, M. Kubal, F. Meyer, G. J. Olsen, R. Olson, A. L. Osterman, R. A. Overbeek, L. K. McNeil, D. Paarmann, T. Paczian, B. Parrello, G. D. Pusch, C. Reich, R. Stevens, O. Vassieva, V. Vonstein, A. Wilke, O. Zagnitko, The RAST server: Rapid annotations using subsystems technology. *BMC Genomics* **9**, 75 (2008).

144. K. E. Luxem, A. M. L. Kraepiel, L. Zhang, J. R. Waldbauer, X. Zhang, Carbon substrate re-orders relative growth of a bacterium using Mo-, V-, or Fe-nitrogenase for nitrogen fixation. *Environ. Microbiol.* **22**, 1397–1408 (2020).
145. J. W. Leigh, D. Bryant, PopART: Full-feature software for haplotype network construction. *Methods Ecol. Evol.* **6**, 1110–1116 (2015).
146. B. M. Robicheau, J. Tolman, D. Desai, J. LaRoche, Microevolutionary patterns in ecotypes of the symbiotic cyanobacterium UCYN-A revealed from a Northwest Atlantic coastal time series. *Sci. Adv.* **9**, eadh9768 (2023).
147. H. Bandelt, P. Forster, A. Röhl, Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **16**, 37–48 (1999).