Genome Biology

RESEARCH

Open Access

Non-cyanobacterial diazotrophs support the survival of marine microalgae in nitrogen-depleted environment

Udita Chandola[1], Marinna Gaudin[2,3], Camille Trottier[1], Louis-Josselin Lavier-Aydat[1], Eric Manirakiza[1], Samuel Menicot[1], Erik Jörg Fischer[4], Isabelle Louvet[5], Thomas Lacour[6], Timothée Chaumier[1], Atsuko Tanaka[7], Georg Pohnert[4], Samuel Chaffron[2,3] and Leïla Tirichine[1,8*]

*Correspondence:
tirichine-l@univ-nantes.fr

[1] UMR 6286, F-44000, Nantes Université, CNRS, Nantes US2B, France
[2] UMR 6004, Nantes Université, École Centrale Nantes, CNRS, Nantes LS2 N, France
[3] Research Federation for the Study of Global Ocean Systems Ecology and Evolution, Tara Oceans GOSEE, F-75016, Paris R2022, France
[4] Institute for Inorganic and Analytical Chemistry, Friedrich Schiller University Jena, Lessingstrasse 8, Jena 7743, Germany
[5] UMR 6230, Nantes Université, CNRS, CEISAM, Nantes 44000, France
[6] PHYTOX, PHYSALG, Rue de L'Ile d'Yeu, Nantes Cedex 03 BP2110544311, France
[7] Department of Chemistry, Biology and Marine Science, Faculty of Science, University of the Ryukyus, Okinawa 903-0213, Japan
[8] Institute for Marine and Antarctic Studies (IMAS), Ecology and Biodiversity Centre, University of Tasmania, TAS, Hobart 7004, Australia

Abstract

Background: Non-cyanobacteria diazotrophs (NCDs) are shown to dominate in surface waters shifting the long-held paradigm of cyanobacteria dominance. This raises fundamental questions on how these putative heterotrophic bacteria thrive in sunlit oceans. The absence of laboratory cultures of these bacteria significantly limits our ability to understand their behavior in natural environments and, consequently, their contribution to the marine nitrogen cycle.

Results: Here, via a multidisciplinary approach, we identify the presence of NCDs within the phycosphere of the model diatom *Phaeodactylum tricornutum* (*Pt*), which sustain the survival of *Pt* in nitrogen-depleted conditions. Through bacterial metacommunity sequencing and genome assembly, we identify multiple NCDs belonging to the Rhizobiales order, including *Bradyrhizobium*, *Mesorhizobium*, *Georhizobium*, and *Methylobacterium*. We demonstrate the nitrogen-fixing ability of *Pt*NCDs through in silico identification of nitrogen fixation genes and by other experimental assays. We show the wide occurrence of this type of interactions with the isolation of NCDs from other microalgae, their identification in the environment, and their predicted associations with photosynthetic microalgae.

Conclusions: Our study underscores the importance of microalgae interactions with NCDs to support nitrogen fixation. This work provides a unique model *Pt*-NCDs to study the ecology of this interaction, advancing our understanding of the key drivers of global marine nitrogen fixation.

Keywords: Biological nitrogen fixation, Non-cyanobacterial diazotrophs, Rhizobiales, *Phaeodactylum tricornutum*, Microalgae, Eukaryote-prokaryote interactions

Background

Nearly 80% of Earth's atmosphere is in the form of molecular nitrogen (N_2). Similarly, the main form of nitrogen in the oceans which cover 71% of Earth surface is dissolved dinitrogen. The rest is reactive nitrogen in the form of nitrate, ammonia and dissolved organic compounds, which are very scarce, rendering nitrogen one of the main limiting factors of productivity in most areas of the oceans [1]. Despite its dominance, ironically, N_2 is unavailable for use by most organisms. Biological nitrogen fixation which is an energy intensive process is specific to only some prokaryotes and archaea termed diazotrophs. It is catalyzed by the nitrogenase enzyme complex including the nitrogenase reductase encoded by the *nifH* gene and the nitrogenase composed of two pairs of different subunits, α and β, respectively encoded by the *nifD* and *nifK* genes [2]. Therefore, eukaryotes are only able to obtain fixed nitrogen through their interactions with diazotrophs [3], although recent evidence has revealed the evolutionary conversion of a cyanobacterium into a nitrogen-fixing organelle within its host, representing the first documented case of nitrogen fixation in a eukaryote [4].

In aquatic habitats, associations between diazotrophic cyanobacteria such as Richelia/Calothrix and various eukaryotes, including several diatom genera such as *Chaetoceros, Rhizosolenia* and *Hemiaulus* known as diatom-diazotroph associations are common and important in oligotrophic habitats of the ocean [5]. Because the nitrogenase complex is sensitive to oxygen, diazotrophs evolved protective nitrogen fixation alternatives that involve conditionally, temporally, or spatially separating oxidative phosphorylation or photosynthesis from nitrogen fixation [6–11]. Cyanobacteria were considered for a very long time as the main diazotrophs in the ocean. However, recent molecular analyses indicated that non-cyanobacterial diazotrophs (NCDs) are also present and active, exhibiting an important presence of diverse *nifH* genes related to non-cyanobacteria, especially to Proteobacteria and Planctomycetes [12–16].

The overwhelming dominance of NCDs raises important questions on how these presumably heterotrophic proteobacteria thrive in the photic zones of the oceans. One emerging hypothesis is that their association with aggregates and larger cellular fractions including phytoplankton, may provide the optimal conditions for nitrogen fixation [16–19]. In our work, we provide evidence for the occurrence of associations between free-living NCDs and several microalgae. More importantly, we identified this interaction in an attractive biological model, the diatom *Phaeodactylum tricornutum* (*Pt*), opening important and unique opportunities to investigate the molecular mechanisms of diazotrophy in symbiotic interactions in microalgae. We sequenced the bacterial meta-community associated with *Pt* and identified previously unknown diverse marine diazotrophs of mainly Alphaproteobacteria type, supporting the survival of host cells in the absence of any form of usable nitrogen. Further, we isolated, cultured, sequenced and fully assembled some of the *Pt*-associated NCDs. Their assembled genomes clustered phylogenetically to NCDs found in the environment and showed environmental co-occurrence patterns with microalgae including diatoms, green algae and haptophytes. Interestingly, most of the *Pt*NCDs belong to the Rhizobiales, an order known for its ability to nodulate legumes. These NCDs include *Bradyrhizobium*, *Rhizobium*, *Mesorhizobium*, *Methylobacterium*, *Phylobacterium* and *Georhizobium* raising important questions into the evolution of symbiotic nitrogen fixation and its terrestrial or marine

origins. Our findings differ from a recent study that identified an NCD residing within a diatom as an obligate symbiont with a reduced genome compared to free living NCDs [20]. Our work brings evidence for an apparently prevalent interaction between microalgae and NCDs, that might explain how both partners cope with periods of nitrogen scarcity, and how these NCDs compromise for nitrogen fixation with its high energy cost in fully oxygenated areas of the water column. These new findings will have to be considered in future nitrogen and carbon biogeochemical cycling studies.

Results

Identification of non-cyanobacterial diazotrophs interaction with diatoms

In a screening of 10 xenic environmental samples of *Pt* grown in nitrogen-depleted medium without the addition of an exogenous carbon source, one accession *P. tricornutum 15* (hereafter *Pt15*), sampled from East China sea, was found to survive in contrast to the axenic control (Fig. 1a). This suggests that diatom growth may have been supported by bacterial nitrogen fixation providing a usable source of nitrogen to diatom cells. After 3 months of nitrogen starvation, xenic *Pt15* showed similar growth to non-starved cells upon transfer to nitrate replete medium. This indicated that *Pt15* starved cells remained viable and capable of growth under favorable conditions as evidenced by the presence of brown-pigmented cells (Fig. 1b). To demonstrate nitrogen fixation, we used the widely applied Acetylene Reduction Assay (ARA) [21] which detected ethylene production in nitrogen-starved xenic *Pt15* cultures. Although the ethylene levels were low, they were notably higher than those in the *E. coli* negative control, where no ethylene production was detected. Xenic *Pt15* cultures under nitrogen-replete conditions (+ N) did not show any ethylene production. No ethylene was detected in axenic *Pt15* confirming that the measured ethylene was not endogenous to the diatom cells but originated from bacterial nitrogen fixation (Additional file 1: Table S1).

To further investigate nitrogen fixation in minus N condition (-N), we measured Particulate Organic Nitrogen quota (PON) in diatom cells and showed a stable level of PON even after 80 days of growth, supporting the maintenance of diatom cells with the amounts of usable nitrogen likely provided by *Pt15* associated bacteria (Additional file 1: Table S1; Additional file 2: Fig. S1a). No exogenous carbon source was provided to the cultures, making the diatom the sole source of carbon. Simultaneously, measured biomass of *Pt15*, indicated a stable cell population suggesting a ground state photosynthetic activity that might support the energy demand of nitrogen fixation (Additional file 2: Fig. S1b). We measured a diatom cell density of 4.3 million cells/mL in the +N condition, whereas in the -N condition, *Pt15* cell density did not exceed 360,000 cells/mL.

It is crucial to highlight that the growth conditions are extreme, with no nitrate provided to the diatom cells, likely leading to a reduction in their metabolic activity to cope with the harsh environment. Fixed nitrogen is the only nitrogen source available to both the diatom and the entire bacterial community, likely resulting in competition for this limited resource. Taken together, our results suggest bacterial nitrogen fixation activity that maintain diatom population cells stable and alive. Light microscopy performed on samples grown on agar plates showed a significant accumulation of bacteria in -N compared to +N condition. This accumulation is not an artifact but is likely a response to nitrate starvation, driven by the formation of biofilm-like clusters

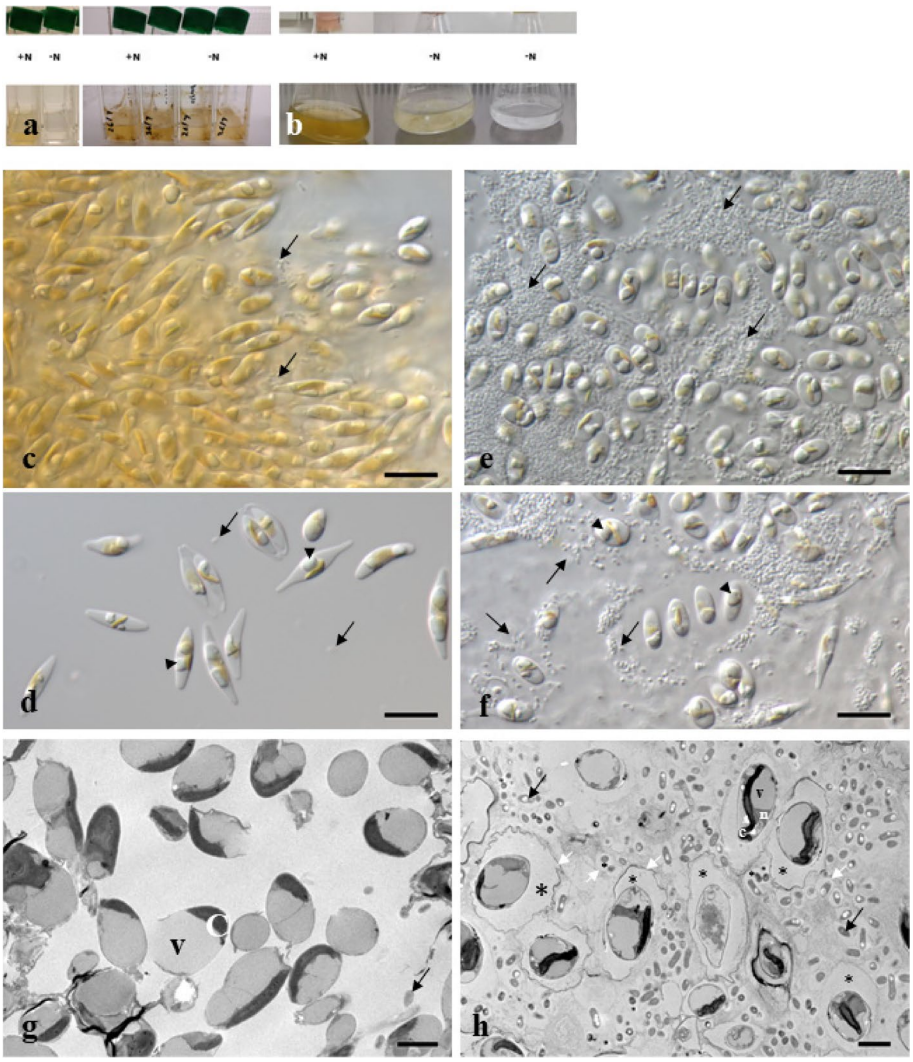

Fig. 1 Macro/microscopic phenotype of *P. tricornutum* cells under nitrogen starvation. **a**. An illustration comparing the screening results of two *P. tricornutum* accessions: xenic *Pt4* (left), which thrives in + N but does not survive in -N, and the candidate accession, xenic *Pt15* (right), shown in duplicates. **b**, from right to left: axenic *Pt15* not surviving in -N (negative control), xenic *Pt15* in -N condition and xenic *Pt15* transferred to + N medium after 3 months in -N. **c**-**h**, micrographs of *P. tricornutum* xenic cultures on agar plates. Cells were grown in -N (**e**, **f**, **h**) or + N culture plates (**c**, **d**, **g**). Living cells were observed by DIC (**c**-**f**), and fixed cells by transmission electron microscopy (**g**, **h**). **d**, Diatom cells grown in + N possessed yellow–brown color of photosynthetic pigments. Mainly in aggregated cells of fusiform and oval shapes, some bacteria were observed (arrows). **d**, Single or multiple whitish droplets, oil bodies, were developed next to chloroplasts (arrowheads). **e**, **f**, Diatom cells formed aggregates with a significant number of surrounding bacteria (arrows) in the -N culture, compared to the + N condition. **f**, Diatom cells grown without nitrate also developed oil bodies in the oval cells. **g**, **h**, Electron micrographs indicate cross-sectioned cells with or without nitrate. Organelles with high electron density are chloroplast, and major spaces were occupied by oil bodies or vacuoles, which showed low electron densities. Number of bacteria cells (arrows) were obviously larger in -N (**h**) compared to + N culture (**g**). **h**, Interestingly, both *Pt* cells and some bacteria are encased in sacs surrounded with a likely EPS made membrane (white arrows) which showed darker electron density around diatom cells compared to bacteria. Cells in 1 (**g)** and 1 (**h)** were stained with TI blue stain, which binds to negatively charged components such as anionic polysaccharides increasing electron density. A low electron density space (stars) separates *Pt* cells and bacteria from the membrane. c, chloroplast, v, vacuole, n, nucleus. Scale bars, 10 µm (**c**-**f**) and 2 µm (**g**, **h**)

of NCDs. These clusters may promote the development of low-oxygen microenvironments favorable for nitrogen fixation (Fig. 1a, d, e and f). Bacteria from agar plates were observed to surround diatom cells within the phycosphere, but not to be directly attached to or residing inside the cells. Transmission electron microscopy from agar plates samples confirmed these observations and revealed a peculiar structure with a higher electron density around diatom cells looking like a pellicle of exopolysaccharides of either diatom or bacterial origin, and found exclusively in starved cells (Fig. 1g and h). Similar structures, although less prominent, occurred around some of bacterial cells, either as a single bacterium or a group of several bacteria. The probability that these structures are composed of polysaccharides is supported by the TI blue stain, a well-known dye that has an affinity for staining polysaccharides [22].

To identify species composition of *Pt15* bacterial community in -N and +N conditions, bacterial DNA was extracted and shotgun sequenced revealing a total of 153,4 million reads. Using the GTDB reference database, we identified a total of 2 classes, 12 families, 8 orders and 90 genera. Upon comparison of the two conditions, our analysis identified 90 distinct genera and 3,224 bacterial species in the -N condition, whereas only 18 genera and 1,540 species were detected in the +N condition. At class level, Alphaproteobacteria represented 98% ±0,8 SD in +N and 99% ±0,5 SD in -N, based on relative abundance estimations, while *Gammaproteobacteria* were detected with a relative abundance below 1% (Fig. 2a). At family level, *Sphingomonadaceae* and *Rhizobiaceae*, both known to contain diazotrophic species, dominated (Fig. 2c). Attached bacteria may be underrepresented due to the filtration process used to separate diatom cells from bacteria prior to DNA extraction. However, to maximize the recovery of both attached and free-living bacteria, diatom cells were vigorously vortexed multiple times before filtration.

At genus level, among the 90 distinct taxa identified, 73 genera were only detected in -N with no reads in +N (Additional file 3: Table S2) [24–30]. At the genus level, the microbiome composition of the community showed a dominance of *Sphyngopixys* in +N, while -N showed a significant enrichment in *Mesorhizobium sp., Bradyrhizobium sp*., *Georhizobium sp*., and *Brevundimonas sp*. In addition to their identification in aquatic environments [31–35], these genera are also recognized as nitrogen-fixing taxa in terrestrial environments, either existing as free-living organisms and/or forming symbiotic relationships with legumes (Fig. 2d). Compared to +N, *Sphyngopixys* abundance decreased in -N although its dominance over the other taxa persisted under this condition. A total of 5 *nifH* genes were detected and assigned to 4 different MAGs (MAG21-*Bradyrhizobium*: 2 *nifH*, MAG8-*Bradyrhizobium*: 1 *nifH*, MAG4-Georhizobium: 1 *nifH*, and MAG5-Sphingopyxis: 1 *nifH*) from the -N bacterial community. No *nifH* genes were detected in the +N condition. Interestingly, several presumably non-nitrogen fixing species (67 out of 73) were significantly enriched in -N (Additional file 3: Table S2), suggesting a role of these species in direct or indirect interactions within the bacterial community and/or with the host. These associated non-diazotrophic bacteria, which are enriched in the absence of bioavailable nitrogen, may contribute to the interaction through various mechanisms. We hypothesize that they could modulate competition among diazotrophs interacting with the host, produce compounds that promote microalgae growth, and inhibit antagonistic

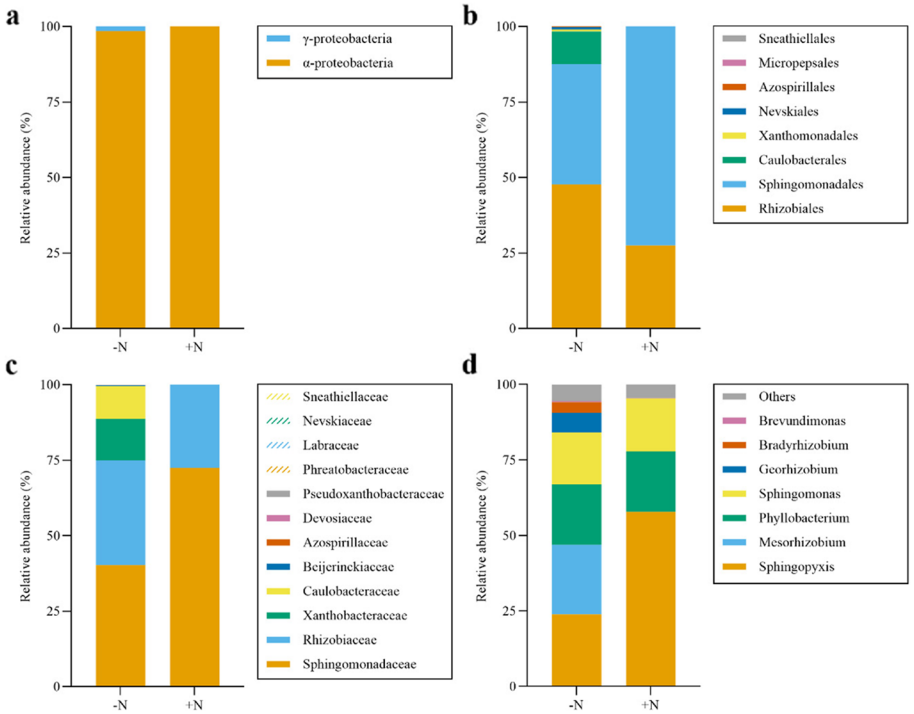

Fig. 2 Comparative relative abundance estimation of *Pt15* associated bacteria. Relative abundance of bacterial community is shown as (**a**), class, (**b**), order, (**c**), family and (**d**), genus (7 most abundant) levels between -N and + N conditions. Two replicates were performed for each condition. Taxonomic assignation of metagenomic community associated to *P. tricornutum* in nitrate-deplete and nitrate-replete conditions was done using anvio version 7.2 standard recommended pipeline using GTDB (The Genome Taxonomy Database) for assigning taxa with the in-built HMM of 22 single copy core genes [23]

organisms, thereby facilitating the interaction with the host. However, these proposed mechanisms have yet to be demonstrated.

Isolation/identification of *P. tricornutum* associated non-cyanobacterial diazotrophs

Given the unique nature of the identified interaction involving a model diatom species and NCDs, it was important to isolate the diazotrophic bacteria identified in silico from metagenomics (MetaG) data in order to further explore this interaction. We used selective media and took advantage of the species assignation provided by MetaG to target the isolation of the desired bacteria according to literature reports [36–39]. Therefore, different growth conditions (media, temperature) were tested (see methods) leading to the isolation of 9 axenic bacterial species (Table 1).

Bacterial colonies were identified using either full length 16S rRNA gene or two hypervariable regions of the 16S (V1 V2, V3 V4). The presence of *nifH* was assayed using degenerate primers (Additional file 2: Fig. S2a and b). Six bacteria were identified as nitrogen fixers based on a combination of *nifH* PCR amplification and/or *nifH* in silico detection along with observation of growth on nitrogen free medium for those that were isolated (Table 1 and additional file 4: Table S3). Among the nine isolated bacteria, some species were not detected in metagenomics sequences likely due to their low

Table 1 List of isolated bacteria from xenic *Pt15* in -N condition

Genus	MAG No.	Size (Mbp)	No. of contigs	*nifH* PCR identification	*nifH in silico* identification	Growth on NFM	Predicted plasmids
Mesorhizobium	**MAG3**	6.85	75	No	Yes	Yes	Yes
Bradyrhizobium sp1*.	**MAG21**	8.42	5	Yes	Yes	Yes	No
Sphingopyxis sp2	MAG7	5.00	23	ND	No	ND	Yes
Rhizobium sp.	NA	ND	ND	ND	ND	Yes	ND
Paracoccus sp.,	NA	ND	ND	ND	ND	Yes	ND
Methylobacterium*	NA	ND	ND	Yes	Yes	Yes	ND
Arthrobacter	MAG14	3.46	12	No	No	ND	ND
Pseudoarthrobacter	NA	ND	ND	No	ND	No	ND
Micrococcacea	NA	ND	ND	No	ND	No	ND

NCDs are indicated in bold. *NA* Not assembled,*ND* Not determined. * 15 N incorporation was used for nitrogen fixation detection

abundance, emphasizing the importance of combining sequencing and isolation/culturing approaches (Table 1). Due to the use of selective culturing media for bacterial species isolation, we identified species that were not necessarily present in the in-silico analysis of the metacommunity, which was grown under a single condition.

To provide stronger support for nitrogen fixation, we monitored $^{15}N_2$ gas-incorporation using mass spectrometry. This allowed a direct measure of nitrogen fixation in the two first isolated bacteria, *Bradyrhizobium* MAG21 and *Methylobacterium*. The labeling of various amino acids indicated substantial $^{15}N_2$ gas assimilation in both bacteria as well as the positive controls *Pseudomonas stutzeri* and *Azotobacter vinalendii*, while the negative controls, (the axenic *Pt15* and *Escherichia coli*) did not show any nitrogen fixation (Additional file 1: Table S1). Our high-resolution LC–MS approach precisely detects the label at the molecular level through isotope-resolved ions, enabling structure-resolved detection of labeled compounds. This allows for more accurate identification of ^{15} N_2 incorporation patterns, which would be harder to resolve with bulk detection methods. This detection unequivocally establishes the nitrogen-fixing capability of the *Pt15*-associated bacteria, namely *Bradyrhizobium* and *Methylobacterium*.

We could not observe $^{15}N_2$ incorporation in xenic *Pt15* cells, which can be attributed to their lower bacterial population density when compared to the lab cultured high biomass of *Bradyrhizobium* and *Methylobacterium*, as well as the competition with other numerous bacteria (~ 100 species) for fixed nitrogen. To verify that the ability of *Pt15* cells to uptake bioavailable nitrogen was not altered, we cultured xenic *Pt15* in artificial seawater containing L-valine labeled with ^{15}N, both in the presence and absence of nitrate. The analysis of endometabolomes using LC–MS distinctly revealed the incorporation of labeled valine in xenic *Pt15* when nitrate levels were at 0 μM (Additional file 1: Table S1). However, no such incorporation was observed in diatom cells when nitrate was present in the medium (545 μM). This suggests that *Pt15* is able to take up nitrogen in the form of amino acids from the associated NCDs, if available nitrogen is otherwise limited.

Diazotrophic MAGs are diverse and affiliated to environmentally sampled diazotrophs of Alphaproteobacteria type

To gain insights into *Pt15* associated diazotrophs and their phylogenetic relationship with known taxa, we assembled their genomes and recovered several MAGs with completeness exceeding 97% (Additional file 4: Table S3). To achieve a better identification of the bacterial genetic material, we sought to recover plasmids from

non-assembled reads, which are important elements for bacteria survival susceptible to contain nitrogen fixation genes. We used three different plasmid prediction tools and recovered 42 plasmids assigned to different bacterial species with an average length of 97,013 bp (Additional file 4: Table S3). Besides *nifH*, several additional *nif* genes were detected in most of the MAGs, among which two were found in plasmids, *nifU* and *nifS*.

Phylogenomic analysis of *Pt15* MAGs, MAGs of known terrestrial and marine diazotrophs from the MicroScope database [40] and MAGs from *Tara* Oceans data using anvi'o [41] was performed. The genomes from MicroScope were selected by individual phylogenetic analysis of each nine *Pt15* MAG deposited on the platform. Thus, the relatedness on the MAGs represented are to diazotrophic marine (purple), non-diazotrophic marine (white), diazotrophic terrestrial (green) and non-diazotrophic terrestrial (green) bacteria. The phylogeny was built using anvio platform and in built HMM of 71 scgs (single copy core genes), Bacteria_71. The analysis revealed two clusters of *Pt15* MAGs on the phylogeny, cluster A (blue) and cluster B (red) (Fig. 3). Cluster A (*Mesorhizobium*, *Bradyrhizobium* MAGs, *Georhizobium*, *Phylobacterium*, and *Brevundimonas*) was affiliated to diazotrophs, while, cluster B (2 *Sphingopyxis* MAGs and *Sphingobium*) was related more to denitrifiers. In both clusters, whenever the related MAGs were not identified as diazotrophs (because of the absence of *nifHDK*, likely due to incomplete assembly), they were found to belong to Rhizobiales, Parvibaculales, Caulobacterales and Sphingomonodales, orders known to contain nitrogen fixing species [29, 42, 43]. *Sphinopyxis* MAG5 was identified with *nifH* gene in silico and was found to cluster with non- or un-identified diazotrophs. For a clearer depiction of Fig. 3, an interactive phylogenetic tree is available on the following link: https://ncdstree.univ-nantes.fr. Altogether, our results including (i) survival of the xenic diatom in the absence of usable forms of nitrogen, (ii) nitrogen fixation demonstrated by ARA and ^{15}N-incorporation, (iii) detection of *nifH,* and (iv) the isolation of NCDs from the xenic *Pt15* clustering with known proteobacteria diazotrophs namely rhizobia, indicate that an interaction for nitrogen fixation likely occurs in *Pt15* xenic cultures, supporting the survival of diatoms in -N condition.

(See figure on next page.)

Fig. 3 Phylogenetic tree of *Pt15* assembled bacterial genomes amongst 1888 TARA bacterial MAGs and MAGs from MicroScope. The phylogeny of 71 single copy core genes was built using anvio platform. The upper circle represents a global view of the phylogeny built using a total of 1927 MAGs and genomes (1888 TARA MAGs, 30 terrestrial MAGs and 9 *Pt15* MAGs). The bootstrap values represented from 0–1 with replicates of 100 and complete phylogeny can be accessed as an interactive tree at: https://ncdstree.univ-nantes.fr. The lower part of the figure represents a partial zoom-in on the two clusters where the *Pt15* MAGs form groups in the tree with respect to other MAGs and genomes. Cluster A, left (blue highlighted text on the tree) consists of *Pt15* associated MAGs: MAG9, MAG4, MAG3, MAG2, MAG8 and MAG21. Cluster B, right (red highlighted text on the tree) consists of *Pt15* associated MAGs: MAG7, MAG5 and MAG6. Green highlighted text represents terrestrial bacteria, purple highlighted text represents diazotroph TARA MAGs (reported *nifH* gene in silico) and non-highlighted text/white background represents non-diazotroph TARA MAGs. Symbol legends, right lower-end denote information with bacterial genomes/MAGs on the tree. (Orange, oval): *Pt15* MAGs, (purple, oval): TARA diazotrophic genomes (green oval): MicroScope genomes, (brown, circle): terrestrial bacteria, (blue, circle): marine bacteria, (triangle): symbiotic, (box): free-living, (star): symbiotic and free-living, (inverted triangle): symbiotic and poor- or non-fixer and (circle): no diazotrophic information available. The numbers indicated on the oval symbols correspond to genome information in Additional file 7: Table S6

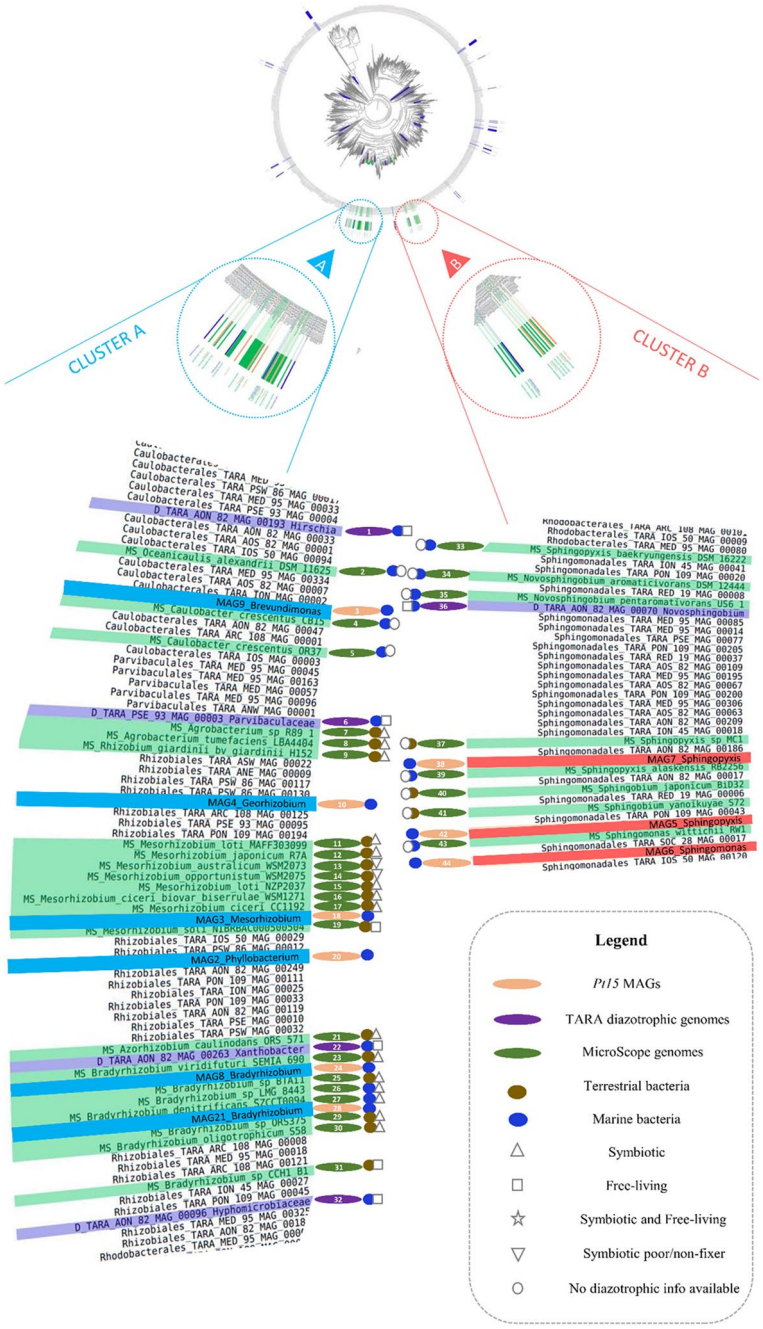

Fig. 3 (See legend on previous page.)

Functional gene annotation of *Pt15* associated NCDs in nitrate deplete versus replete conditions

To understand the biological functions of the microbial community, genes from -N and +N treatments were assigned to functional categories (Fig. 4; Additional file 5: Table S4). There was a significant enrichment of genes related to nitrogen fixation and metabolism in -N compared to +N condition. Typically, in -N, we exclusively identified *nifH* and *nifD*, two key components of the nitrogenase enzyme complex.

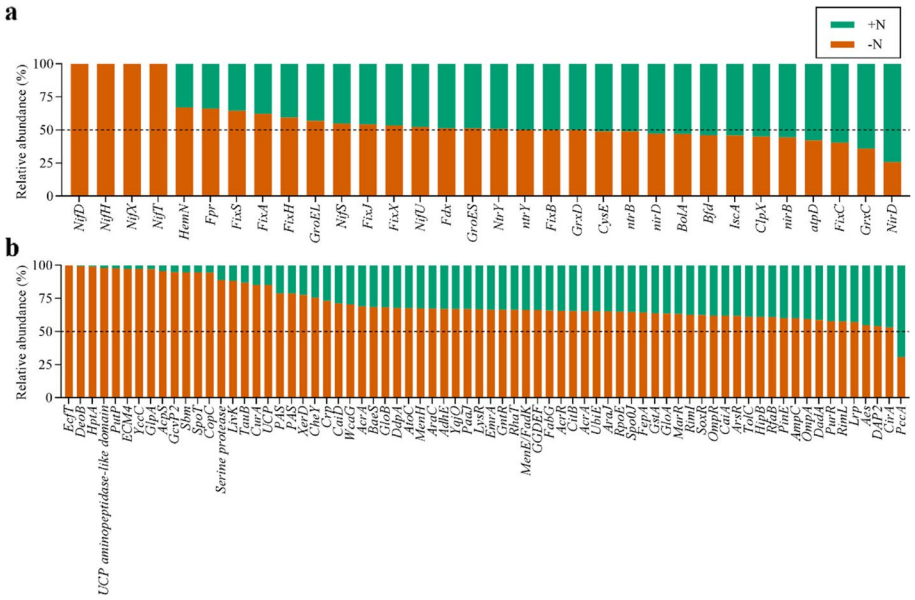

Fig. 4 Functional gene annotation of *Pt15* associated NCDs in nitrate deplete versus replete conditions. (**a**), Relative abundance of selected genes involved in nitrogen fixation and assimilation in nitrate deplete and nitrate replete conditions. (**b**), illustration representing a combined view of 50 top abundant gene in nitrate deplete and replete condition with ordering representing top relatively abundant genes in nitrate deplete conditions. Dashed line indicates the relative percentage at 50%

Likewise, *nifX* and *nifT* were only detected in -N suggesting their critical role in nitrogen fixation. Part of the *nif* operon, *NifX* was reported to be important in FeMo-co maturation and its transfer to *NifDK* proteins [44, 45]. Other genes related to nitrogen fixation metabolism emerged from the analysis, namely *HemN*, *Fpr*, *FixS*, *FixA*, *FixH*, *GroEL*, *NifS* and *FixJ* (Fig. 4), some of which are part of the Nif regulon.

The abundance of several transporters was striking in -N, suggesting an important role in energy/substrate shuttling between the host and the bacterial community. Examples included an *EcfT* gene (energy-coupling factor transporter transmembrane) known for its function in providing the energy necessary to transport a number of different substrates [46], AcpS, an acyl carrier gene, which functions as a multi-carbon units carrier into different biosynthetic pathways [47], TauB, a taurine transporter that might play a role in providing taurine, which is produced by a large number of algae and known as a potential source of carbon, nitrogen and energy [48].

Stress responsive genes were clearly more represented in -N compared to +N including SpoT, which is a general bacterial stress response gene [49] that allows bacteria to increase its persistence in stressful conditions, here the lack of nitrogen. Other genes found almost exclusively in the starved bacterial community were the PAS domain containing genes (Per-ARNT-Sim) [50], which function as signal sensors by their capacity to bind small molecules, modulators and transducers. They are known to sense oxygen to prevent degradation of oxygen sensitive enzymes suggesting a crucial and protective role in -N condition where there is an enrichment in nitrogenase containing bacteria. Overall, functional analysis of MetaG data corroborated with NCDs enrichment in -N and nitrogen fixation.

Geographical distribution and ecological prevalence of NCD-microalgae interactions

Next, we sought to assess the prevalence and importance of similar interactions in other microalgae providing further evidence that the association of NCDs with *P. tricornutum* is not an isolated case. First, we screened a collection of 14 xenic microalgae as described for *Pt15,* and identified five species of microalgae (35%) that grew in the absence of nitrogen in the growth medium suggesting the presence of diazotrophs (Additional file 6: Table S5). Similar to *Pt15*, the retained microalgae were screened using different nitrogen deprived media under various conditions for diazotrophs isolation. *NifH* degenerate and 16S rRNA gene primers were used to identify the isolated bacterial colonies, which revealed five Alpha-proteobacteria including two species of *Thalassospira sp*. (isolated from two different microalgae), *Stenotrophomonas maltophilia, Stappia* and *Rhodococcus qingshengii* suggesting that the occurrence of NCDs with microalgae may be relatively common. PCR detection using *nifH* degenerate primers was successful only in *S. maltophilia*. In all other bacteria, with the exception of *Stappia* and *S. maltophilia* which were not tested, a reduction of acetylene to ethylene was observed in the ARA assay (Additional file 6: Table S5**)**. All isolated bacterial species grew in nitrogen free medium (Additional file 6: Table S5). To provide additional support for the presence of nitrogen fixation in bacteria isolated from microalgae other than *P. tricornutum*, we carried out an $^{15}N_2$ incorporation assay followed by LC–MS analysis on *Rhodococcus.* The results clearly indicated that the bacteria assimilated labeled nitrogen, confirming its ability to fix nitrogen (Additional file 1: Table S1).

To maximize the number of species sequenced while maintaining cost-effectiveness, isolated NCDs were sequenced in pools of three bacteria using both Illumina and PacBio. Their genomes were successfully assembled, demonstrating high completeness (> 87%, with two MAGs reaching 100%) and low contamination (Additional file 4: Table S3). We then used MAGs isolated from *Pt15* and the other microalgae (Table 1; Additional file 4: Table S3) to reassess their phylogenetic relationship as described for *Pt15* isolated NCDs. The analysis revealed 3 more clusters (C, D and E) in close proximity to diazotrophic MAGs from the environment and/or MAGs belonging to diazotrophic containing orders [51–53] (Fig. S3). An interactive figure is available at https://ncdstree2.univ-nantes.fr/.

Next, we assessed the occurrence of *Pt15* MAGs in the environment in surface water and Deep Chlorophyll Maximum layer (DCM) using Tara Ocean datasets [54, 55]. Given that *Pt15* MAGs had not been previously reported, the analysis to monitor the distribution for the surface and DCM waters was performed using the two phylogenetically closest TARA MAGs. The phylogeny was built in anvio with 1888 TARA MAGs with inbuilt single copy genes HMM (Bacteria_71) and the five *Pt15* MAGs in which the *nifH* gene was identified. We used branch support and branch length as proxies to select the closely related TARA MAGs. The two TARA MAGs clustering with each individual *Pt15* MAG were selected (Additional file 4: Table S3). With the exception of two *Pt15* MAGs, *Georhizobium* and *Sphingomonas*, which show weak branch support with one of the two TARA MAGs, all other MAGs are robustly supported as being closely related to their respective TARA MAGs in the phylogeny. Among the 130 sampling stations, 92 stations showed the occurrence of at least one of the TARA affiliated MAGs reflecting a high

prevalence (Fig. 5a; Additional file 2: Fig. S4). *Pt15* clustering with TARA MAGs were more frequently occurring with 129 recorded stations. The distribution of these MAGs was extensive, spanning all latitudes from south Pacific, Indian and Atlantic Oceans to polar regions. The MAGs were found in both surface and DCM samples, mainly within the 0.8–5 µm and 0.22–1.6/3 µm size fractions (Fig. 5a; Additional file 2: Fig. S4). Interestingly, TARA MAGs affiliated to *Pt15* MAGs co-occurred in surface and DCM samples in most cases, although their dominance in surface samples was prominent (Fig. 5b).

We then asked whether the occurrence of the MAGs isolated from *Pt15* correlated with environmental factors monitored during the Tara Oceans campaign and whether they co-occur with other organisms. We considered only two *Pt15* MAGs (*Mesorhizobium* MAG3 and *Bradyrhizobium* MAG21) to assess their correlation with three environmental factors used as indicators of diazotrophy, namely nitrate, nitrate phosphate ratio and oxygen concentrations. Iron concentrations from the ECCO2-DARWIN model [56] were considered, yet no distinct relationship was identified with the distribution of NCDs. Interestingly, TARA affiliated MAGs were detected in low nitrate and almost no detection above 5 µM (Fig. 5c). For both nitrate and N/P ratio, there are no differences in their occurrence with TARA diazotrophic MAGs, while these differences are significant with non-diazotrophic MAGs only in the larger size fraction of 180–2000 µm

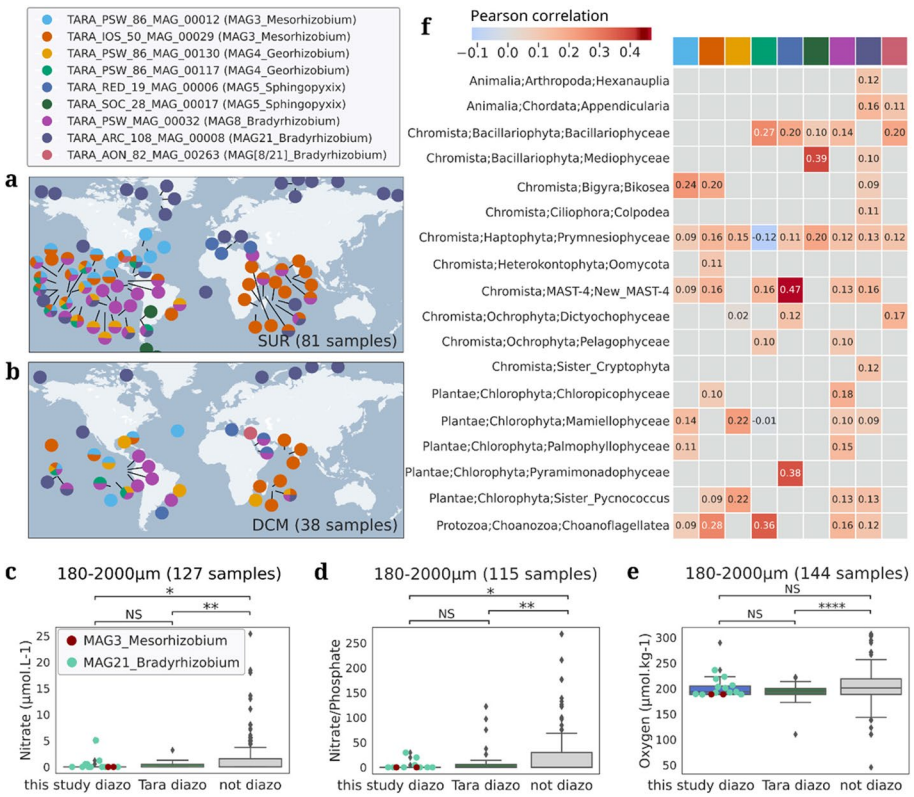

Fig. 5 Geographical distribution and ecological prevalence of isolated and assembled bacterial diazotrophs genomes correlating to TARA data. Geographical distribution of phylogenetically closest *Pt15* associated Tara Ocean MAGs in (**a**), surface and (**b**), DCM samples. (**c**-**e**), relative abundance of the top 2 most occurring MAGs of *Pt15* bacterial community in the environment in function of (c, nitrate, d, N/P ratio and e, oxygen concentration). (**f**), Co-occurrence of TARA MAGs affiliated to *Pt15* in the environment with different species of micro-eukaryotes inferred through FlashWeave, FDR < 0.01

(Fig. 5c and d) suggesting an association with particles, aggregates or other organisms. This measure is in line with nitrate limitation as a marker of diazotrophs presence [57]. In a similar way, oxygen showed a consistent pattern when comparing this study's MAGs with TARA NCDs, distinguishing these two categories from non-diazotrophic MAGs only in 0.22–1.6/3 μm, 0.8–5 μm and 180–2000 μm size fractions (Fig. 5e; Additional file 2: Fig. S5). By contrast to what is known about oxygen inhibition of nitrogenase, both MAGs seemed to thrive in relatively well oxygenated zones (200 μM) compared to less oxygenated areas or oxygen minimum zones (OMZ, 20- 50 μM O_2). This implies adaptive means by which these bacteria fix nitrogen such as creating micro-oxic environments through their interactions with other organisms.

Strikingly, a co-occurrence analysis revealed a clear association of the NCDs isolated from *Pt15* (Fig. 5h) with micro-eukaryotes. Most of the *Pt15*-associated TARA MAGs were found to co-occur primarily with Haptophytes, Bacillariophyta, and Cryptophytes. The association with Bacillariophyta is consistent with their isolation from a diatom. While we did not demonstrate nitrogen fixation in these predicted associations, this analysis provides evidence for a probable nitrogen fixing interaction between *Pt15* MAGs and various species of primarily photosynthetic micro-eukaryotes.

Discussion

Recent reports of NCDs dominance in the euphotic zones of the oceans questioned the long-held dogma that symbiotic N_2 fixation is mainly carried by cyanobacteria, but also raised fundamental questions on how these presumably heterotroph NCDs can thrive in the sunlit ocean [12–14, 58]. The nitrogenase enzyme complex responsible for nitrogen fixation is sensitive to oxygen that irreversibly inactivates the enzyme. Therefore, diazotrophs evolved several mechanisms to protect nitrogenase from oxygen degradation including hyper-respiration, uncoupling of photosynthesis in space and time in photosynthetic diazotrophs, and attachment/interaction with microalgae or debris which were found to be hotspots of nitrogen fixation [17]. Exopolysaccharides produced by microalgae can create a low-oxygen microenvironment that protects the oxygen-sensitive nitrogenase enzyme by serving as a physical barrier, thus restricting oxygen diffusion to the nitrogenase in bacteria associated with the microalgae. Using a multidisciplinary approach, the study described herein brings evidence of interactions for nitrogen fixation between microalgae, in particular the model diatom *Phaeodactylum tricornutum* and NCDs. After screening ten xenic accessions of *P. tricornutum* in the absence of a usable form of nitrogen, our study revealed that one accession, *Pt15*, interacts with nitrogen-fixing proteobacteria. Although low, this nitrogen fixation seemed to maintain the viability of the diatom and its associated bacteria. To compensate for the allocation of energy to the bacterial community, *Pt15* cells seemed to slow down growth. This diatom-NCDs interaction may reflect natural environments, where such interactions, in extreme conditions (low to no nitrogen) take place until favorable factors (e.g., nitrate repletion with upwelling or Sahara dust) permit the recovery and growth of cells, such as what we observed in our nitrate starved cultures when transferred to nitrate replete medium.

Metagenomic sequencing of the associated bacteria revealed not only one diazotrophic species but a group of nitrogen fixers predominantly Alpha-proteobacteria,

mainly represented by Rhizobiaceae, Sphingomonadaceae and Xanthobacteraceae. All the isolated bacteria identified as diazotrophs grew in nitrogen free medium. The nitrogenase gene was detected in the majority of them, through either in silico analysis or by PCR amplification, further confirming the presence of diazotrophy in these bacteria. The *nifH* gene was found in most cases in the chromosome, except for *Georhizobium* where it was predicted to be plasmidic. ARA results indicated that nitrogen fixation rates in xenic *Pt15* under -N condition were lower than those observed in the positive controls and showed variability when compared to nitrogen fixation rates reported for NCDs in natural environments. These differences likely reflect the variability in environmental conditions across study sites. For example, in the Gotland Basin, NCD nitrogen fixation rates were estimated at approximately 7.6 nmol/L/day in surface waters and 0.44 nmol/L/day in deep waters [59]. In contrast, nitrogen fixation rates of NCDs in hypoxic waters of the Southern Californian Bight averaged 0.07 nmol/L/day, which is lower than those observed in the present study [60]. However, differences in monitoring techniques and experimental conditions, particularly between in situ measurements and laboratory cultures, must be considered when drawing comparisons [59–62]. The low nitrogen fixation rates observed in xenic *Pt15* and its associated bacterial community under nitrogen-depleted conditions are likely due to two interconnected phenomena. First, competition for fixed nitrogen by non-diazotrophic bacterial species may have limited the nitrogen available to the diatoms. Second, the quiescent state of diatom cells, induced by the absence of nitrogen, likely contributed to the reduced nitrogen fixation rates in the NCDs. Our data suggest that diatom cells ceased division and maintained minimal metabolic activity, which may have further reduced the availability of photosynthates necessary to support nitrogen fixation by the associated NCDs. Quiescence is a well-documented survival strategy in certain microalgae under unfavorable environmental conditions, allowing them to save energy and sustain basic cellular functions until nutrient availability improves [63–65]. This was evident in *Pt15* cells, which resumed growth upon transfer to nitrogen-rich medium. If *Pt15* is not in an optimal physiological state to supply energy to its associated bacteria, nitrogen fixation rates will remain low, further exacerbated by the consumption of fixed nitrogen by the abundant non-diazotrophic bacteria present in the culture. Under our culture conditions, where bioavailable nitrogen is entirely absent, the presence of competing non NCDs for fixed nitrogen, combined with the quiescent state of diatom cells, which prevents nitrogen uptake due to the lack of metabolic activity, suggests that little to no nitrogen is transferred to the diatoms. This could explain our inability to detect $^{15}N_2$ incorporation in diatom cells. This aspect deserves further investigation, particularly under conditions of limited rather than complete nitrogen deprivation, coupled with the application of more sensitive methodologies to improve the detection of nitrogen transfer.

The nitrogen-fixing capability of the isolated bacteria was directly confirmed using $^{15}N_2$ incorporation monitored by isotope analysis in LC–MS, leaving no doubt about the *Pt15* associated bacteria's ability to fix nitrogen. This provides initial evidence that NCDs associated with the diatom are capable of fixing nitrogen, a process likely beneficial to other associated bacteria and potentially to the diatom, as long as it does not transition into a resting stage due to the extreme culture conditions, specifically the complete absence of bioavailable nitrogen. A recent study identified

an NCD from the Rhizobiales order residing within a diatom of the *Haslea* genus, functioning as an endosymbiont [20]. The interactions we identified differ significantly from those with *Haslea* containing NCDs. The bacteria associated with *Haslea* have a much smaller genome, typically less than 1 Mb, whereas our NCDs have genomes about 7 Mb in size [20]. This discrepancy in genome size reflects the different types of interactions: in the obligate interaction, the bacteria have lost parts of their genome and depend on the host's cellular machinery, whereas in the facultative interaction, the bacteria interact from a distance or within the phycosphere [66–69]. Our findings reflect the broad array of nitrogen fixation interactions, extending from obligatory symbiosis to epibiosis or remote interactions within the phycosphere, mirroring observations in nitrogen fixing interactions between Rhizobia and land plants. Bacteria associated with *Pt15*, identified as NCDs in this study, align with the distribution patterns of Tara NCDs in ocean regions characterized by low nitrate and nitrate phosphate ratio. Interestingly, their distribution and correlation with these environmental factors are observed only in larger size fractions. This observation is in accordance with findings from multiple studies that have reported similar diazotroph distributions under the aforementioned conditions. Indeed, diazotrophs have developed various strategies to counteract the irreversible degradation of the nitrogenase enzyme by oxygen [6, 7, 10, 70]. These strategies include forming associations with other organisms, among other mechanisms, as they seek out low-oxygen environments primarily facilitated by surrounding polymers.

Our study revealed that NCDs interaction with *Pt15* and other microalgae is likely not an isolated phenomenon and is more broadly distributed than previously anticipated. Our MAGs were detected in the environment clustering with previously identified terrestrial and marine NCDs, supporting diazotrophy and likely their ecological importance. These MAGs were identified at global scale in surface and DCM samples in size fractions enriched in microalgae, which supports the hypothesis of an interaction with them as evidenced by the co-occurrence analysis. This interaction might be taking advantage of hypo-oxic biomes microalgae and bacteria may provide with the secretion of extracellular polymers, mainly polysaccharides, that form biofilms ensuring low O_2 diffusion [71–73]. This was suggested by our TEM analysis with the formation of pellicles around *Pt* cells and bacteria. Such pellicles were reported to form around two nitrogen fixing bacteria, *Pseudomonas stutzeri* and *Azospirillum brasilense* in oxygenated surroundings implying the importance of oxygen stress in the formation of these structures [74]. The presence of these sac-like structures might have hindered diatom uptake of released ammonium leading to growth inhibition in the extreme condition tested in this study. Besides biofilm formation, this interaction may combine different strategies for both microalgae and NCDs to avoid/limit oxygen inactivation of nitrogenase such as a conformational switch that protects nitrogenase from O_2 degradation [10, 75] or hyper-respiration [76] among other protective mechanisms that need further investigations. In line with our discovery of NCDs interaction with microalgae for nitrogen fixation, a recent study [77] suggested that anaerobic processes, including heterotrophic N_2 fixation, can take place in anoxic microenvironments inside sinking particles even in fully oxygenated marine waters.

Conclusions

Our screen under nitrogen-deprived conditions of different accessions of the model diatom *P. tricornutum* revealed several previously unidentified NCDs among which some were found to be related to terrestrial nitrogen-fixing bacteria, raising questions about their evolutionary origins. The NCDs associated with *P. tricornutum* enable the diatom to survive in nitrogen-deprived environments, indicating a relationship that supports its persistence. These NCDs remain in the phycosphere and do not penetrate diatom cells, suggesting that their interactions occur at a distance. Genome analyses of *P. tricornutum*-associated bacteria showed clustering with both terrestrial and marine NCDs, and these bacteria were found to be widely distributed across latitudes, co-occurring with photosynthetic microeukaryotes, including diatoms, haptophytes, and cryptophytes. The identification of different NCDs with a model diatom opens wide perspectives to investigate the role of individual and combined interactions of bacterial species with the host and within the bacterial community. Thus, *Pt15* accession with its associated NCDs represent a unique Eukaryotic-bacteria model system and an exciting opportunity to study molecular mechanisms underlying this interaction for nitrogen fixation and beyond.

Methods

Microalgae species used and culture media

Screening of Phaeodactylum tricornutum accessions

Ten xenic *P. tricornutum* accessions were screened in presence (545 µM) and absence of nitrate in Enhanced Artificial Sea Water (EASW) [78] in order to select xenic *Pt* ecotypes growing in the absence of nitrate. The starting concentration for all accessions, measured under the microscope using Malassez counting chamber, was 100,000 cells per ml. The cells were counted and washed twice with nitrate-free EASW in order to remove all residual nitrate and resuspended in respective media for screening over a period of one month. A growth curve was made by counting the cells in periodic time intervals. The cultures were grown at 19 °C with a 12/12 h photoperiod at 30 to 50 µE/s/m2. A cocktail of antibiotics (Ampicillin 100 µg/µl, Chloramphenicol, 14 µg/µl and Streptomycin 100 µg/µl) was used to make *Pt15* axenic in liquid EASW culture medium with refreshment of antibiotic containing medium every 5 days.

Pt15 cells were regularly monitored for axenicity using peptone-rich medium and microscopy.

Additionally, axenicity was further validated by culturing the putatively axenic *Pt15* in -N medium. If the *Pt15* cells died under these conditions, it would confirm the absence of associated bacteria.

Screening of other microalgae

A total of fourteen microalgae were screened in nitrate-deplete and nitrate-replete enhanced artificial sea water for growth (Additional file 6: Table S5). About 100,000 cells were counted, resuspended in their respective media and grown as described above. Their growth was assessed for a period of two weeks at regular intervals.

Acetylene reduction assay

Acetylene reduction assay (ARA) [21] was used to quantify the conversion of acetylene (C_2H_2) to ethylene (C_2H_4) which allows the assessment of nitrogenase activity. ARA was conducted on the following samples: two positive controls, *Azotobacter vinelandii* (NCIMB 12096), *Pseudomonas stutzeri* BAL361, xenic *Pt15* grown in NFM [79, 80] or nitrogen replete medium, all incubated at three different temperatures (19, 25, 30 °C). Axenic *Pt15* culture grown in NFM medium +5 sugars (final concentration of 1% each: D-glucose, D-sucrose, D-fructose, D-galactose and D-mannose) was used in every measurement as a negative control to ensure there was no endogenous ethylene produced. All the cultures were grown at their respective growth media and conditions. After reaching an appropriate O.D. between 0.2–0.3, 1 mL of the bacterial cells were collected in an Eppendorf's tube and centrifuged at11,000 rpm for 10 min at RT to remove the medium. The cells were re-suspended in NFM with five sugars and grown in their respective medium. Sugar was not used for culturing xenic *Pt15*. When the bacterial growth reached an O.D. of 0.2–0.3, the cultures were prepared for acetylene reduction assay.

The cultures were secured with a rubber stopper in order to make them air tight. Ten percent of the air was removed from the culture flasks and supplemented with 10% acetylene gas. The cultures were incubated at their appropriate temperatures for a period of 72 h before monitoring ARA with gas chromatography. A volume of 500 µl was taken from the headspace of the samples using a Hamilton gas tight syringe (SYR 500 ul 750 RN no NDL, NDL large RN 6/pk) and injected into an Agilent gas chromatograph model 7820 equipped with a Flame Ionization Detector (GC-FID) and using Hydrogen as carrier gas at a constant flow of 7.7 mL/min. Samples were injected into a split/splitless injector, set at 160˚C in splitless mode. A GS-Alumina capillary column, 50 m × 0.53 mm × 0.25 µm (Agilent, 115–3552) was used. The programmed oven temperature was at 160˚C (isotherm program) during 2 min. The FID detector was set to 160˚C. A standard curve was made with ethylene gas using 5 volumes (100, 150, 200 µL). The raw data generated from the GC provides value of y, Area [pA*s]. This value is supplemented in the slope-intercept formula, y = mx + c, with 'm' and 'c' representing slope and constant values respectively, denoted in the calibration curve. The amounts of ethylene were calculated as described previously [81]. Acetylene reduction rates were converted to fixed N using the common 3:1 (C2H2:N2) molar conversion ratio [82, 83].

Particulate organic nitrogen measurements

One liter of 200,000 cells/mL culture of xenic *Pt15* was grown in nitrate-free EASW for over a period of 80-days at 19 °C and 12 h/12 h light/dark photo period at 70µE/s/m^2. Six time points were used for harvesting samples: 1, 4, 12, 59, 76 and 80 day(s). The cells were counted at each time point using a cell counter. Twenty mL of *Pt15* culture were harvested and injected through a sterile (treated at 450 deg/4 h) Whatman GF/F glass microfibre filter paper to collect *Pt15* cells and the filters were dried overnight in 60 °C incubator and stored at −80 °C. The 20 mL filtrate from the cultures was passed through 0.22 µm filters (Sigma ca. GVWP10050), in order to remove bacterial cells and stored at −80 °C, before being processed for particulate organic nitrogen measurement analysis.

EASW with 545 µM nitrate, 0 µM nitrate and three heat-treated Whatman GF/F glass microfibre filters were used as blanks for comparison with test samples.

Test for incorporation of the ^{15}N isotope

Stationary cultures (30 mL) of either the diatoms, *Phaeodactylum tricornutum Pt15* Xenic Plus nitrate (545 µM), *P. tricornutum Pt15* Xenic Minus nitrate (0 µM), *P. tricornutum Pt15* Axenic Plus nitrate (545 µM) or bacteria, *Bradyrhizobium MAG21*, *Methylobacterium*, *Azotobacter vinalendii*, *Pseudomonas stutzeri BAL361*, *Escherichia coli DH5a* and *Rhodococcus* were put under a $^{15}N_2$ (98 atom %, Sigma-Aldrich, Steinheim, Germany) atmosphere (20 mL). For this, a 50 mL syringe (B Braun Omnifix®) without needle was used to take up 30 mL of the culture, followed by 20 mL of $^{15}N_2$ from a lecture bottle. The syringes were immediately sealed with a plastic valve and parafilm to prevent leakage of the gas. Diatom cultures were incubated at 18 °C under a 14:10 h light: dark regime with an illumination of 60–65 µmol photons m^{-2} s^{-1}. Bacterial cultures were incubated at 100 rpm and 30 °C, except for *E. coli DH5a* which was incubated at 37 °C. After 2.5 days the diatom cultures were filtered (1.2 µm pore size, GF/C, Whatman, GE Healthcare, Little Chalfont, United Kingdom). The cells were washed off the filter using 5 mL ice-cold methanol/ethanol/chloroform 1:3:1 (methanol: HiPerSolv Chromanorm, VWR Chemicals, Dresden, Germany; ethanol: HiPerSolv Chromanorm, VWR Chemicals, Dresden, Germany; chloroform: SupraSolv, VWR Chemicals, Dresden, Germany) and the resulting extracts were treated in an ultrasonic bath for 10 min and centrifuged (16.100 × g, 20 min, 4 °C). The supernatant was stored at −20 °C until analysis. Bacterial cultures were centrifuged (3.893 × g, 20 min, 4 °C) and the cells were resuspended in 5 mL of methanol/dichlormethane/ethylacetate 10:2:3 (dichlormethane: SupraSolv, VWR Chemicals, Dresden, Germany; ethylacetate, HiPerSolv Chromanorm, VWR Chemicals, Dresden, Germany). The suspensions were sonicated (5 min, 9 cycles, 80–85%) and afterwards centrifuged (3893 × g, 20 min, 4 °C). The supernatant was stored at −20 °C until further use.

Test for uptake of ^{15}N-L-valine-by *P. tricornutum*

To stationary cultures (40 mL) of either *Phaeodactylum tricornutum Pt15* Xenic Plus nitrate (545 µM) or *P. tricornutum Pt15* Xenic Minus nitrate (0 µM), 100 µL of an aqueous ^{15}N-L-valine- solution (98 atom % ^{15}N, Merck, Saint Louis, USA, 50 µmol L^{-1} in medium) were added to reach a final concentration of 125 nm L^{-1}. The cultures were then incubated, extracted and the samples stored as described in "*Test for incorporation of the ^{15}N isotope*".

Analysis of endometabolomes with LC-MS

Cell extracts were dried in a nitrogen stream at room temperature and resuspended in either 200 µL methanol for diatom extracts or methanol/acetonitrile/water 5:9:1 (acetonitrile: HiPerSolv Chromanorm, VWR Chemicals, Dresden, Germany; water: CHEMSOLUTE, TH.GEYER, Renningen, Germany) for bacterial extracts. Metabolites were separated on a Dionex Ultimate 3000 UHPLC coupled to a Q-Exactive Plus Orbitrap mass spectrometer (Thermo Scientific, Bremen, Germany). Chromatographic separation was performed on a SeQuant ZIC-HILIC column (5 µm, 200 Å, 150 × 2.1 mm, Merck,

Germany) equipped with a SeQuant ZIC-HILIC guard column (20 × 2.1 mm, Merck, Germany) at 25 °C using water with 2% acetonitrile and 0.1% formic acid as eluent A and 90% acetonitrile with 10% water and 1 mmol L^{-1} ammonium acetate as eluent B. Full scan measurements (m/z 80–1200, resolution 70 k, automatic gain control target $3E^6$, maximum ion injection time 200 ms) were performed in positive and negative ion mode using a heated electrospray ionization source to generate molecular ions. The method duration was 14.5 min with an MS runtime from 0.5 min to 9.0 min, a flow rate of 0.6 mL min^{-1} and a gradient as follows: 85% B (0.0–4.0 min), linear gradient to 100% A (4.0–5.0 min), 100% A (5.0–9.0 min), linear gradient to 100% B (9.0–10.0 min), 100% B (10.0–12.0 min), linear gradient to 85% B (12.0–12.5 min), 85% B (12.5–14.5 min). Amino acid standards were run separately for comparison of retention time and mass spectra. Resulting raw data was analyzed with FreeStyle (Thermo Fisher Scientific).

Microscopy

Light microscopy

Living cells grown on agar plates were used directly for the preparations. Cells were taken directly from agar plates without prior treatment, transferred onto glass slides, and covered with a cover slip before observation. Samples were observed with differential interference contrast (DIC) microscopy by BX53 microscope (Olympus, Japan) and Axiocam 705 color (Carl Zeiss, Germany). The agar medium is composed of EASW [78], supplemented with 15 g/L of agar.

Transmission electron microscopy

Samples were prepared as described previously [84]. Embedded samples were cut for making pale yellow ultra-thin Sects. (70–80 nm thickness) with a microtome EM UC6 (Leica Microsystems, Germany). Sections were stained with TI blue (Nisshin EM, Japan) at a 1:10 dilution in distilled water followed by lead citrate treatment. After staining, the sections were air dried before inspection with a JEM-1011 KM II transmission electron microscope (JEOL, Japan).

Bacterial isolation and culture media used

A total of eight media recipes were used for bacteria isolation and culture under different conditions, all summarized in (Additional file 8: File S1). All isolated bacteria were checked for purity by repeated streaking and 16S sequencing. Bacterial species were maintained in 20% glycerol at −80 °C for long-term storage.

DNA extraction and PCR amplification of marker genes

For initial identification of bacteria, the isolated colonies were picked under sterile conditions using filter tips, resuspended and mixed thoroughly in 50 µL nuclease-free water (NFW). The suspension was heat/cold treated at 95 °C following 4 °C, 10 min each and twice for cell lysis. Post-treatment, an additional 50 µL NFW was added to the suspension for a total volume of 100 µL. Five µL of the colony DNA was used as template for PCR.

Wizard® Promega Genomic DNA extraction kit was used for extracting high-quality genomic DNA from bacterial cells. The bacteria were grown in their appropriate media

until the culture optical density (OD) reached 0.3. The cells were harvested and washed in 1X PBS twice. The cell pellets were resuspended in 480 µL 50 mM EDTA and treated with 120 µL of 20 mg/ml lyzozyme (VWR cat no. 0663-5G), incubated at 37 °C for 60 min in order to lyze possible gram-positive bacteria. Post incubation, the cells were centrifuged and the supernatant was removed. Nuclear Lysis solution, 600 µL (Promega cat no. A7941), was used to re-suspend and gently mix the pellet for incubation at 80 °C for 5 min. The suspension was allowed to cool at room temperature (RT) for 10 min before adding 1 µL RNAse at 10 mg/mL (VWR cat no. 0675-250MG) and incubation at 37 °C for 60 min. After incubation, the suspension was cooled at RT for 10 min. Protein lysis solution, 200 µL (cat no. VWR 0663-5G) was added, vortexed and incubated on ice (4 °C for 5 min). The mixture was centrifuged and the clear supernatant was transferred to a clean Eppendorf with 600 µL RT isopropanol and the tube was mixed gently before being centrifuged. The cell pellet was washed with freshly prepared 70% ethanol and re-centrifuged. The ethanol was carefully removed without disturbing the pellet and air dried before resuspension in appropriate amount of NFW. The quality of the DNA was checked using Nanodrop 260/280 and 230/260 ratios and on 1% agarose gel. The identity of the bacteria was re-confirmed using 16S rRNA gene amplification and Sanger sequencing.

For the *nifH* PCR amplification, the reaction mixture included 2.5 µL of DreamTaq Green Buffer (1X), 0.1 µL of BSA (0.08 µg/µL), 2 µL of each primer (10 µM), 0.2 µL of DreamTaq Green Polymerase (5 U/µL), 0.5 µL of dNTP mix (10 mM), 3 µL of DNA template (normalized to 10 ng/µL), and H2O adjusted to a total volume of 25 µL. The thermocycling protocol started with an initial denaturation step at 95 °C for 5 min, followed by 40 cycles of denaturation at 95 °C for 30 s, annealing at 53 °C for 30 s, and extension at 72 °C for 30 s, and a final extension step at 72 °C for 7 min. Following the reaction, amplification results were assessed through 1% agarose gel electrophoresis, and DNA bands of interest were isolated using the NucleoSpin® Gel & PCR Cleanup Kit from Macherey–Nagel. The extracted DNA underwent Sanger sequencing at Eurofins Scientific. Sequences were analyzed using Geneious. For species identification, 16S PCR fd1-rd1, variable regions PCR V1 V2 V3 V4 and V5 V7 were used. The reaction was performed with 35 cycles of 95 °C (1 min), 55 °C (1 min) and 72 °C (1 min). All primer sequence details are reported in (Additional file 9: Table S7).

Sample preparation and DNA extraction for Illumina and PacBio sequencing

Duplicate with a starting concentration of 200,000 cells/mL of 3-L culture volumes of xenic *Pt15* in nitrate-replete and nitrate-deplete enhanced artificial sea water were grown for a period of 2 months (60-days) at 19 °C and 12 h/12 h light/dark photoperiod at 70µE/s/m^2 before collection for sequencing. The xenic *Pt15* cells were filtered through a 3 µm filter (MF-millipore hydrophobic nitrocellulose, diameter 47 mm (cat no. SSWP04700) to retain the cells and allow passage of the filtrate containing the bacterial community. The filtrate was filtered twice in order to avoid contamination of the bacterial cells with residual *Pt15* cells. The double filtered bacterial community was collected in 0.22-micron filter MF-millipore hydrophobic nitrocellulose, diameter 47 mm (cat no. GSWP04700) and resuspended in 1X PBS. DNA was extracted for metagenome sequencing using Quick-DNA fungal/bacterial mini prep kit (cat no. DG005) and

according to manufacturer instructions. The bacterial DNA quality was checked using 1% agarose gel to assess intact DNA and quantified using Qubit broad range DNA kit, before being sequenced. Additionally, 16S rRNA sequencing was performed to assess the presence of bacterial DNA and 18S rRNA PCR was run as a negative control to inspect absence of *P. tricornutum* DNA. When bacterial isolates were available, they were sequenced either as a single species or in a pool of 3 bacteria using Illumina and PacBio.

Data processing

Metagenomic analysis

Two different strategies were used to analyze the metagenomic reads: (i) a "Single assemblies" strategy, where assembly and binning was performed for each sample individually. This strategy is known to produce less fragmented assemblies with a limitation of chimeric contigs formation [85]. (ii) a co-assembly strategy was also performed to complete the results with additional likely lower abundant MAGs. The idea is to pool all the available samples reads and build a unique and common assembly for all samples. Six samples from 2 experiments were used: one with sequences obtained at 2 g/L salinity (in + N and -N conditions, no replicate) and the other one at 20 g/L salinity with duplicates for each of the two conditions. Both datasets are similar.

Single assembly strategy

DNA raw reads were analyzed using a dedicated metagenomic pipeline ATLAS v2.7 [85]. This tool includes four major steps of shotgun analysis: quality control of the raw reads, assembly, binning, taxonomic and functional annotations. Briefly, quality control of raw sequences is performed using utilities in the BBTools suite (https://sourceforge.net/projects/bbmap/): BBDuk for trimming and removing adapters, BBSplit for decontamination. In our case *Phaeodactylum tricornutum* reference genome (accession GCA_000150955.2), including mitochondria and chloroplast sequences, was used for this specific step. Quality-control (QC) reads assemblies were performed using metaSPADES [86] with ATLAS default parameters. The assembled contigs shorter than 500 bases length and without mapped reads were removed to keep high-quality contigs.

Metagenome-assembled genomes (MAGs) were obtained using two distinct binners: maxbin2 [87] and metabat2 [88]. The bins obtained by the different binning tools were combined using DASTool [89] and dereplicated with dRep [90] using default parameters. Completeness and contamination were computed using CheckM [91]. Only bins with a completeness above 50% and a contamination rate under 10% were kept. Prediction of open reading frames (ORFs) was performed, at the assembly level (not only at the MAGs level) using Prodigal [92]. Linclust [93] was used to cluster the translated gene product obtained to generate gene and protein catalogs common to all samples.

Co-assembly strategy

QC reads and co-assembly were analyzed using a shotgun dedicated sequential pipeline MetaWRAP v1.3.2 [89]. Briefly, QC reads are pulled together and addressed to the chosen assembler metaSPAdes [86] using MetaWRAP default parameters. Metagenome-assembled genomes (MAGs) were obtained using two distinct binners: maxbin2 and

metabat2 and the bins obtained were combined using an internal MetaWRAP dedicated tool. Completeness and contamination were computed using CheckM [91]. Only bins with a completeness above 50% and a contamination rate under 10% were kept. MetaWRAP also includes a re-assemble bins module that collect reads belonging to each bin, and then reassemble them independently with a"permissive"and a"strict"algorithm. Only improved bins are kept in the final set. Using this strategy, we were able to add an additional MAG.

MAGs relative abundances

The QC reads were mapped back to the contigs, and bam files were generated to organize downstream analysis. Median coverage of the genomes was computed to evaluate MAGs relative abundance estimations.

Taxonomic assignation of metagenomic data

We utilized anvio version 7.2 [41] for taxonomic profiling of the *P. tricornutum* associated metagenomes in nitrate deplete and replete condition. The workflow uses The Genome Taxonomy Database (GTDB) [23] to determine the taxonomy based on 22 single copy core genes (scgs, anvio/anvio/data/misc/SCG_TAXONOMY). Sequence alignment search was done using DIAMOND [94].

Hybrid assembly of MAGs

Assembly

Metagenomic assembly of Illumina technology reads does not yield complete. Genomes due to the difficulty of assembling repetitive regions [95]. Whenever Illumina and PacBio reads were available, we used a strategy to improve MAGs assembly by combining both type of reads. Metagenomic assembly of Illumina paired-end and PacBio reads post quality trimming were assembled using metaSPAdes [86] or SPAdes (v3.15.5) with option flag (–meta). To validate the newly assembled genomes, we verified that the long reads covered regions that were not supported by the mapped Illumina sequences only and that we get high quality bins. We obtained more contiguous genomes in which some regions had not been supported by the assembly of Illumina reads only.

Binning

The contigs were grouped and assigned to the individual genome. To do this, we used MaxBin2 and MetaBAT2 which are clustering methods based on compositional features or on alignment (similarity), or both [96].

Refinement

Bins from MaxBin2 and MetaBAT2 were consolidated into one more solid set of bins. To do this we used Binning_refiner which improves genomic bins by combining different binning programs available at: https://github.com/songweizhi/Binning_refiner [97].

Taxonomic assignation

GTDB-Tk has already been independently and positively evaluated for the classification of MAGs [98]. After consolidating the bins in the refinement step, we determined

the taxonomy of each bin. The gtdbtk classify_wf module was used in this work using a GTDB database https://github.com/ecogenomics/gtdbtk [99].

Gene annotation of metagenomic data and MAGs

Gene annotation of metagenomics data in two conditions was done using anvi'o version 7.2. Three databases were used for a robust gene observations: COG20 [100], KEGG [101] and Prokka [102]. DIAMOND was used to search NCBI's Database. In built, anvi'o program annotates a contig database with HMM hits from KOfam, a database of KEGG Orthologs (KOs, https://www.genome.jp/kegg/ko.html). For Prokka, the annotation was done externally and the gene calls were imported to anvi'o platform into the individual metagenome contigs databases.

Phylogenomic analysis of MAGs

The phylogenomic tree was built using anvi'o −7.1 with database of 1888 TARA ocean MAGs (contig-level FASTA files) submitted to Genoscope (https://www.genoscope.cns.fr/tara/). Additional closest terrestrial and marine genomes (Additional file 7: Table S6) to the individual MAGs were selected using MicroScope database (https://mage.genoscope.cns.fr/microscope/) using genome clustering, utilizing MASH [103] for ANI (Average Nucleotide Identity) for calculating pairwise genomic distancing, neighbor-joining (https://www.npmjs.com/package/neighbor-joining) for tree-construction and computing the clustering using Louvain Community Detection (https://github.com/taynaud/python-louvain). anvi'o utilizes Prodigal [92] to identify open reading frames for the contig database. The HMM profiling for 71 single core-copy bacterial genes (SCGs) is done using in-built anvi'o database, Bacteria_71 and utilizes MUSCLE [104] for multiple sequence alignment. The bootstrap values are represented from 0–1 with replicates of 100.

Plasmid prediction and annotation

Plasmids were predicted from MetaG data using the following plasmid assembly tools: SCAPP [105], metaplasmidSPAdes [106] and PlaScope [107]. Predicted plasmids were annotated using Dfast [108], eggNOG-mapper v 2 [109] and Prokka 1.14.6 [102].

Tara Oceans MAGs biogeography and environmental preferences

Previously reconstructed MAGs from Tara Oceans metagenomes and respective abundance count tables and fasta files were downloaded from open-source released Tara Oceans databases at https://www.genoscope.cns.fr/tara/ [13, 110]. The data included 2,601 MAGs, comprising 1,888 prokaryotic and 713 eukaryotic MAGs, collected from 81 surface samples and 38 DCM samples across five size fractions (0.22–3 µm, 3–20 µm, 0.8–5 µm, 20–200 µm, and 180–2,000 µm) [111].

The biogeographical occurrences of TARA MAGs phylogenetically close to *Pt15* were projected onto global maps for both surface and DCM depths, either by grouping the 5 size fractions (Fig. 5), or separately by size fraction (Additional file 2: Fig. S4).

Tara Oceans physico-chemical parameters were downloaded from Pangaea [54]. MAGs TSS relative abundances were projected and analyzed in relationship with

ammonium (µmol/L), chlorophyll A (mg/m3), iron (µmol/L), nitrate (µmol/L), nitrate/phosphate and oxygen (µmol/kg).

Co-occurrence network inference

Co-occurrence network inference was performed using FlashWeave v0.18.0 [112] with Julia v1.5.3, for eukaryotic and prokaryotic Tara MAGs linked to *Pt15* MAGs. To account for the compositional nature of MAGs abundance data, a multiplicative replacement method [113] was applied to replace zeros, and a Centered Log Ratio transformation [114] was applied separately on both prokaryotic and eukaryotic abundance matrices. FlashWeave parameters used were max_k = 1, n_obs_min = 10, heterogeneous = False and normalize = False.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13059-025-03597-4.

Additional file 1. Table S1. Combined table representing amount of ethylene, N15 incorporation assay and PON detection in xenic Pt15 in nitrate deplete condition

Additional file 2. Supplementary Figs. 1 to 5

Additional file 3. Table S2. Details of taxonomy for bacterial metagenomics data in nitrate deplete and nitrate replete conditions

Additional file 4. Table S3. MAG features, species-specific plasmid detection from Pt15 meta-genomic data and TARA MAGs corresponding to Pt15 MAGs

Additional file 5. Table S4. Gene annotation using anvio_COG20, anvio_KEGG and anvio_Prokka of the metagenomics data in nitrate deplete and replete conditions

Additional file 6. Table S5. Screening of culture collection microalgae under nitrate-depleted conditions and isolation of associated bacteria

Additional file 7. Table S6. Features of the genomes selected from MircoScope for phylogenetic studies

Additional file 8. File S1. Media recipes used for bacterial isolation

Additional file 9. Table S7. List of primers used in the study

Acknowledgements
We acknowledge Clara Guillouche and Irene Romero Rodriguez for their technical assistance with PCR screening and Romain le Balch for his help with GC runs.

Peer review information
Wenjing She was the primary editor of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team. The peer-review history is available in the online version of this article.

Authors'contributions
LT conceived and designed the study. UC performed most of the experiments. SM contributed to the ARA assay and the screening. LJLA assisted with the screening and performed bacterial isolation. TC, CT, EM, UC performed bioinformatics analysis. MG performed bioinformatics analysis and generated Fig. 5. TL conducted the PON assay. AT performed the EM study. IL assisted with the ARA assay. EJF performed 15 N assay. GP supervised EJF and provided guidance on the 15 N assay. SC supervised MG. UC, TC, CT, MG, SC, EJF, GP, AT, TL and LT analysed and interpreted the data. LT wrote the manuscript with methodological input from UC, MG, EF, AT, TL, EM and TC. LT supervised and coordinated the study. All authors read, edited and approved the manuscript.

Funding
LT acknowledges support from the region of Pays de la Loire (ConnecTalent EPIALG project), Epicycle ANR project (ANR-19-CE20- 0028–02) and µAlgaNIF France-Japan International Research Project. UC was supported by grant 998UMR6286 Connect Talent EpiAlg from Région Pays de la Loire to LT. SC acknowledges support from the CNRS MITI through the interdisciplinary program ***Modélisation du Vivant*** (GOBITMAP grant), and the H2020 European Commission project AtlantECO (award number 862923). GP acknowledges support from Germany's Excellence Strategy—EXC 2051—Project-ID 390713860 (Balance of the Microverse). We are grateful to the Genomics Core Facility GenoA, member of Biogenouest and France Genomique and to the Bioinformatics Core Facility BiRD, member of Biogenouest and Institut Français de Bioinformatique (IFB) (ANR-11-INBS-0013) for the use of their resources and their technical support.

Data availability
All sequencing data and generated assemblies are available at NCBI under BioProject PRJNA923279, https://www.ncbi.nlm.nih.gov/bioproject/?term = PRJNA923279 [115]. No other scripts and software were used other than those mentioned in the Methods section.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 2 September 2024 Accepted: 28 April 2025
Published online: 28 May 2025

References

1. Moore CM, Mills, M.M., Arrigo, K.R. I. Berman-Frank, L. Bopp, P. W. Boyd, E. D. Galbraith, , R. J. Geider CG, S. L. Jaccard, T. D. Jickells, J. La Roche, T. M. Lenton, N. M. Mahowald, , E. Marañón IM, J. K. Moore, T. Nakatsuka, A. Oschlies, M. A. Saito, T. F. Thingstad, , Ulllo ATO: Processes and patterns of oceanic nutrient limitation. Nature Geoscience 2013, 6:701–710.
2. Einsle O, Rees DC. Structural Enzymology of Nitrogenase Enzymes. Chem Rev. 2020;120:4969–5004.
3. Kneip C, Lockhart P, Voss C, Maier UG. Nitrogen fixation in eukaryotes–new models for symbiosis. BMC Evol Biol. 2007;7:55.
4. Coale TH, Loconte V, Turk-Kubo KA, Vanslembrouck B, Mak WKE, Cheung S, Ekman A, Chen JH, Hagino K, Takano Y, et al. Nitrogen-fixing organelle in a marine alga. Science. 2024;384:217–22.
5. Fiore CL, Jarett JK, Olson ND, Lesser MP. Nitrogen fixation and nitrogen transformations in marine symbioses. Trends Microbiol. 2010;18:455–63.
6. Cornejo-Castillo FM, Zehr JP. Hopanoid lipids may facilitate aerobic nitrogen fixation in the ocean. Proc Natl Acad Sci U S A. 2019;116:18269–71.
7. Inomura K, Deutsch C, Wilson ST, Masuda T, Lawrenz E, Lenka B, Sobotka R, Gauglitz JM, Saito MA, Prasil O, Follows MJ. quantifying oxygen management and temperature and light dependencies of nitrogen fixation by crocosphaera watsonii. mSphere. 2019;4(6):e00531–19. https://doi.org/10.1128/mSphere.00531-19.
8. Inomura K, Wilson ST, Deutsch C. Mechanistic model for the coexistence of nitrogen fixation and photosynthesis in marine trichodesmium. mSystems. 2019;4(4):10.1128/msystems.0021019. https://doi.org/10.1128/msystems.00210-19.
9. Munoz-Marin MDC, Shilova IN, Shi T, Farnelid H, Cabello AM, Zehr JP. The transcriptional cycle is suited to daytime n2 fixation in the unicellular cyanobacterium "candidatus atelocyanobacterium thalassa" (UCYN-A). Mbio. 2019;10(1):e02495–18. https://doi.org/10.1128/mbio.02495-18.
10. Schlesier J, Rohde M, Gerhardt S, Einsle O. A Conformational Switch Triggers Nitrogenase Protection from Oxygen Damage by Shethna Protein II (FeSII). J Am Chem Soc. 2016;138:239–47.
11. Bothe H, Tripp HJ, Zehr JP. Unicellular cyanobacteria with a new mode of life: the lack of photosynthetic oxygen evolution allows nitrogen fixation to proceed. Arch Microbiol. 2010;192:783–90.
12. Delmont TO, Quince C, Shaiber A, Esen OC, Lee ST, Rappe MS, McLellan SL, Lucker S, Eren AM. Nitrogen-fixing populations of Planctomycetes and Proteobacteria are abundant in surface ocean metagenomes. Nat Microbiol. 2018;3:804–13.
13. Delmont TOPK, J.S. Veseli, I., Fuessel, J. Murat Eren, A., Foster, R.A., Bowler, C., Wincker, P., Pelletier, E.: Heterotrophic bacterial diazotrophs are more abundant than their cyanobacterial counterparts in metagenomes covering most of the sunlit ocean.https://doi.org/10.1038/s41396-021-01135-1 2021.
14. Farnelid H, Andersson AF, Bertilsson S, Al-Soud WA, Hansen LH, Sorensen S, Steward GF, Hagstrom A, Riemann L. Nitrogenase gene amplicons from global marine surface waters are dominated by genes of non-cyanobacteria. PLoS ONE. 2011;6: e19223.
15. Moisander PH, Benavides M, Bonnet S, Berman-Frank I, White AE, Riemann L. Chasing after Non-cyanobacterial Nitrogen Fixation in Marine Pelagic Environments. Front Microbiol. 2017;8:1736.
16. Harding KJ, Turk-Kubo KA, Mak EWK, Weber PK, Mayali X, Zehr JP. Cell-specific measurements show nitrogen fixation by particle-attached putative non-cyanobacterial diazotrophs in the North Pacific Subtropical Gyre. Nat Commun. 2022;13:6979.
17. Geisler E, Bogler A, Rahav E, Bar-Zeev E. Direct Detection of Heterotrophic Diazotrophs Associated with Planktonic Aggregates. Sci Rep. 2019;9:9288.
18. Ababou F-ELMF, Cornet-Barthaux V, Taillandier V, Bonnet S. Composition of the sinking particle flux in a hot spot of dinitrogen fixation revealed through polyacrylamide gel traps. Front Mar Sci. 2024;10:1290625.
19. Geisler E, Bogler A, Bar-Zeev E, Rahav E. Heterotrophic Nitrogen Fixation at the Hyper-Eutrophic Qishon River and Estuary System. Front Microbiol. 2020;11:1370.
20. Tschitschko B, Esti M, Philippi M, Kidane AT, Littmann S, Kitzinger K, Speth DR, Li S, Kraberg A, Tienken D, et al. Rhizobia-diatom symbiosis fixes missing nitrogen in the ocean. Nature. 2024;630:899–904.

21. Sullivan BW, Smith WK, Townsend AR, Nasto MK, Reed SC, Chazdon RL, Cleveland CC. Spatially robust estimates of biological nitrogen (N) fixation imply substantial human alteration of the tropical N cycle. Proc Natl Acad Sci U S A. 2014;111:8101–6.
22. Nagasato C, Terauchi, M., Tanaka, A. and Motomura, T.: Development and function of plasmodesmata in zygotes of Fucus distichus. Botanica Marina 2015, 58:229–238.
23. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, Hugenholtz P. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol. 2018;36:996–1004.
24. Rojas A, Holguin G, Glick BR, Bashan Y: Synergism between Phyllobacterium sp. (N(2)-fixer) and Bacillus licheniformis (P-solubilizer), both from a semiarid mangrove rhizosphere. FEMS Microbiol Ecol 2001, 35:181–187.
25. Klepa MS, diCenzo GC, Hungria M. Comparative genomic analysis of Bradyrhizobium strains with natural variability in the efficiency of nitrogen fixation, competitiveness, and adaptation to stressful edaphoclimatic conditions. Microbiol Spectr. 2024;12: e0026024.
26. Laranjo M, Alexandre A, Oliveira S. Legume growth-promoting rhizobia: an overview on the Mesorhizobium genus. Microbiol Res. 2014;169:2–17.
27. Kabdullayeva T, Crosbie DB, Marin M. Mesorhizobium norvegicum sp. nov., a rhizobium isolated from a Lotus corniculatus root nodule in Norway. Int J Syst Evol Microbiol. 2020;70:388–96.
28. Naqqash T, Imran A, Hameed S, Shahid M, Majeed A, Iqbal J, Hanif MK, Ejaz S, Malik KA: First report of diazotrophic Brevundimonas spp. as growth enhancer and root colonizer of potato. Sci Rep 2020, 10:12893.
29. Videira SS, de Araujo JL, Rodrigues Lda S, Baldani VL, Baldani JI. Occurrence and diversity of nitrogen-fixing Sphingomonas bacteria associated with rice plants grown in Brazil. FEMS Microbiol Lett. 2009;293:11–9.
30. Xi H, Ryder M, Searle IR. Complete genome sequence of allorhizobium vitis strain K306, the causal agent of grapevine crown gall. Microbiol Resour Announc. 2020;9(29):e00565–20. https://doi.org/10.1128/MRA.00565-20.
31. Wu C, Sun J, Liu H, Xu W, Guicheng Z, Lu H, Guo Y. Evidence of the significant contribution of heterotrophic diazotrophs to nitrogen fixation in the Eastern Indian ocean during pre-Southwest Monsoon period. Ecosystems. 24(6). https://doi.org/10.1007/s10021-021-00702-z.
32. Tang W, Wang S, Fonseca-Batista D, Dehairs F, Gifford S, Gonzalez AG, Gallinari M, Planquette H, Sarthou G, Cassar N. Revisiting the distribution of oceanic N(2) fixation and estimating diazotrophic contribution to marine production. Nat Commun. 2019;10:831.
33. Tao J, Wang S, Liao T, Luo H. Evolutionary origin and ecological implication of a unique nif island in free-living Bradyrhizobium lineages. ISME J. 2021;15:3195–206.
34. Krick A, Kehraus S, Eberl L, Riedel K, Anke H, Kaesler I, Graeber I, Szewzyk U, Konig GM: A marine Mesorhizobium sp. produces structurally novel long-chain N-acyl-L-homoserine lactones. Appl Environ Microbiol 2007, 73:3587–3594.
35. Fu GY, Yu XY, Zhang CY, Zhao Z, Wu D, Su Y, Wang RJ, Han SB, Wu M, Sun C: Mesorhizobium oceanicum sp. nov., isolated from deep seawater. Int J Syst Evol Microbiol 2017, 67:2739–2745.
36. de Lajudie P, Willems A, Nick G, Moreira F, Molouba F, Hoste B, Torck U, Neyra M, Collins MD, Lindstrom K, et al: Characterization of tropical tree rhizobia and description of Mesorhizobium plurifarium sp. nov. Int J Syst Bacteriol 1998, 48 Pt 2:369–382.
37. Yang X, Jiang Z, Zhang J, Zhou X, Zhang X, Wang L, Yu T, Wang Z, Bei J, Dong B, et al: Mesorhizobium alexandrii sp. nov., isolated from phycosphere microbiota of PSTs-producing marine dinoflagellate Alexandrium minutum amtk4. Antonie Van Leeuwenhoek 2020, 113:907–917.
38. Zhang X, Tong, J., Dong, M., Akhtar, K. and He, B.: Isolation, identification and characterization of nitrogen fixing endophytic bacteria and their effects on cassava production. PeerJ 10: e12677 2022, 10.
39. Bostrom KH, Riemann L, Kuhl M, Hagstrom A. Isolation and gene quantification of heterotrophic N2-fixing bacterioplankton in the Baltic Sea. Environ Microbiol. 2007;9:152–64.
40. Medigue C, Calteau A, Cruveiller S, Gachet M, Gautreau G, Josso A, Lajus A, Langlois J, Pereira H, Planel R, et al. MicroScope-an integrated resource for community expertise of gene functions and comparative analysis of microbial genomic and metabolic data. Brief Bioinform. 2019;20:1071–84.
41. Eren AM, Kiefl E, Shaiber A, Veseli I, Miller SE, Schechter MS, Fink I, Pan JN, Yousef M, Fogarty EC, et al. Community-led, integrated, reproducible multi-omics with anvi'o. Nat Microbiol. 2021;6:3–6.
42. Lesser MP, Morrow KM, Pankey SM, Noonan SHC. Diazotroph diversity and nitrogen fixation in the coral Stylophora pistillata from the Great Barrier Reef. ISME J. 2018;12:813–24.
43. Skorupska A, Janczarek M, Marczak M, Mazur A, Krol J. Rhizobial exopolysaccharides: genetic control and symbiotic functions. Microb Cell Fact. 2006;5:7.
44. Hernandez JA, Igarashi RY, Soboh B, Curatti L, Dean DR, Ludden PW, Rubio LM. NifX and NifEN exchange NifB cofactor and the VK-cluster, a newly isolated intermediate of the iron-molybdenum cofactor biosynthetic pathway. Mol Microbiol. 2007;63:177–92.
45. Jasniewski AJ, Lee CC, Ribbe MW, Hu Y. Reactivity, Mechanism, and Assembly of the Alternative Nitrogenases. Chem Rev. 2020;120:5107–57.
46. Wang T, Fu G, Pan X, Wu J, Gong X, Wang J, Shi Y. Structure of a bacterial energy-coupling factor transporter. Nature. 2013;497:272–6.
47. Geiger O, Lopez-Lara IM. Rhizobial acyl carrier proteins and their roles in the formation of bacterial cell-surface components that are required for the development of nitrogen-fixing root nodules on legume hosts. FEMS Microbiol Lett. 2002;208:153–62.
48. Cook AM, Denger K. Metabolism of taurine in microorganisms: a primer in molecular biodiversity? Adv Exp Med Biol. 2006;583:3–13.
49. Boutte CC, Crosson S. Bacterial lifestyle shapes stringent response activation. Trends Microbiol. 2013;21:174–80.
50. Stuffle EC, Johnson MS, Watts KJ. PAS domains in bacterial signal transduction. Curr Opin Microbiol. 2021;61:8–15.
51. Madigan M, Cox SS, Stegeman RA. Nitrogen fixation and nitrogenase activities in members of the family Rhodospirillaceae. J Bacteriol. 1984;157:73–8.

52. Singh RK, Singh P, Sharma A, Guo D-J, Upadhyay SK, Song Q-Q, Verma KK, Li D-P, Malviya MK, Song X-P, et al. Unraveling Nitrogen Fixing Potential of Endophytic Diazotrophs of Different Saccharum Species for Sustainable Sugarcane Growth. Int J Mol Sci. 2022;23:6242.
53. Sellstedt A, Richau KH. Aspects of nitrogen-fixing Actinobacteria, in particular free-living and symbiotic Frankia. FEMS Microbiol Lett. 2013;342:179–86.
54. Pesant S, Not F, Picheral M, Kandels-Lewis S, Le Bescot N, Gorsky G, Iudicone D, Karsenti E, Speich S, Trouble R, et al. Open science resources for the discovery and analysis of Tara Oceans data. Sci Data. 2015;2: 150023.
55. Karsenti E, Acinas SG, Bork P, Bowler C, De Vargas C, Raes J, Sullivan M, Arendt D, Benzoni F, Claverie JM, et al. A holistic approach to marine eco-systems biology. PLoS Biol. 2011;9: e1001177.
56. Menemenlis BD, Campin, J., Heimbach, P., Hill, C., & Lee, T., 31(October), 13–21. : ECCO2: High resolution global ocean and sea ice data synthesis. Mercator Ocean Quarterly Newsletter 2008, 31:13–21.
57. Ward BADS, Moore CM, Follows MJ. Iron, phosphorus, and nitrogen supply ratios define the biogeography of nitrogen fixation. Limnol Oceanogr. 2013;58:2059–75.
58. Bombar D, Paerl RW, Riemann L. Marine Non-Cyanobacterial Diazotrophs: Moving beyond Molecular Detection. Trends Microbiol. 2016;24:916–27.
59. Mohr W, Grosskopf T, Wallace DW, LaRoche J. Methodological underestimation of oceanic nitrogen fixation rates. PLoS ONE. 2010;5: e12583.
60. Hamersley RTK, Leinweber A, Gruber N, Zehr JP, et al. Nitrogen fixation within the water column associated with two hypoxic basins within the Southern California Bight. Aquat Microb Ecol. 2011;63:193–205.
61. Farnelid H, Bentzon-Tilia M, Andersson AF, Bertilsson S, Jost G, Labrenz M, Jurgens K, Riemann L. Active nitrogen-fixing heterotrophic bacteria at and below the chemocline of the central Baltic Sea. ISME J. 2013;7:1413–23.
62. Fernandez C, Farias L, Ulloa O. Nitrogen fixation in denitrified marine waters. PLoS ONE. 2011;6: e20539.
63. Pelusi A, Ambrosino L, Miralto M, Chiusano ML, Rogato A, Ferrante MI, Montresor M. Gene expression during the formation of resting spores induced by nitrogen starvation in the marine diatom Chaetoceros socialis. BMC Genomics. 2023;24:106.
64. Takeuchi T, Benning C. Nitrogen-dependent coordination of cell cycle, quiescence and TAG accumulation in Chlamydomonas. Biotechnol Biofuels. 2019;12:292.
65. Wang G, Huang L, Zhuang S, Han F, Huang Q, Hao M, Lin G, Chen L, Shen B, Li F, et al. Resting cell formation in the marine diatom Thalassiosira pseudonana. New Phytol. 2024;243:1347–60.
66. Turk-Kubo KA, Mills MM, Arrigo KR, van Dijken G, Henke BA, Stewart B, Wilson ST, Zehr JP. UCYN-A/haptophyte symbioses dominate N(2) fixation in the Southern California Current System. ISME Commun. 2021;1:42.
67. Carpenter EJ, Janson S. Intracellular Cyanobacterial Symbionts in the Marine Diatom Climacodium Frauenfeldianum (Bacillariophyceae). J Phycol. 2000;36:540–4.
68. Janson S, Wouters J, Bergman B, Carpenter EJ. Host specificity in the Richelia-diatom symbiosis revealed by hetR gene sequence analysis. Environ Microbiol. 1999;1:431–8.
69. Prechtl J, Kneip C, Lockhart P, Wenderoth K, Maier UG. Intracellular spheroid bodies of Rhopalodia gibba have nitrogen-fixing apparatus of cyanobacterial origin. Mol Biol Evol. 2004;21:1477–81.
70. Sabra W, Zeng AP, Lunsdorf H, Deckwer WD. Effect of oxygen on formation and structure of Azotobacter vinelandii alginate and its role in protecting nitrogenase. Appl Environ Microbiol. 2000;66:4037–44.
71. Limoli DH, Jones CJ, Wozniak DJ. Bacterial extracellular polysaccharides in biofilm Formation and function. Microbiol Spectr . 2015;3(3):10.1128/microbiolspec.MB-0011-2014. https://doi.org/10.1128/microbiolspec.MB-0011-2014.
72. Paerl HW, Prufert LE. Oxygen-poor microzones as potential sites of microbial n(2) fixation in nitrogen-depleted aerobic marine waters. Appl Environ Microbiol. 1987;53:1078–87.
73. Riemann L, Rahav E, Passow U, Grossart HP, de Beer D, Klawonn I, Eichner M, Benavides M, Bar-Zeev E. Planktonic Aggregates as Hotspots for Heterotrophic Diazotrophy: The Plot Thickens. Front Microbiol. 2022;13: 875050.
74. Wang D, Xu A, Elmerich C, Ma LZ. Biofilm formation enables free-living nitrogen-fixing rhizobacteria to fix nitrogen under aerobic conditions. ISME J. 2017;11:1602–13.
75. Maier RJ, Moshiri F. Role of the Azotobacter vinelandii nitrogenase-protective shethna protein in preventing oxygen-mediated cell death. J Bacteriol. 2000;182:3854–7.
76. Brian ACJ, C. W. : The respiratory system of Azotobacter vinelandii. Eur J Biochem 1971, 20:29–35.
77. Chakraborty S, Andersen KH, Visser AW, Inomura K, Follows MJ, Riemann L. Quantifying nitrogen fixation by heterotrophic bacteria in sinking marine particles. Nat Commun. 2021;12:4085.
78. Vartanian M, Descles J, Quinet M, Douady S, Lopez PJ. Plasticity and robustness of pattern formation in the model diatom Phaeodactylum tricornutum. New Phytol. 2009;182:429–42.
79. Jorgensen JH, Pfaller MA, Carroll KC, Funke G, Landry ML, Richter SS, Warnock DW. Manual of Clinical Microbiology. 11th ed. Washington (DC): ASM Press; 2015.
80. Isenberg HD. Clinical microbiology procedures handbook. 2nd ed. Washington, DC: American Society for Microbiology; 1992.
81. Paudel D, Liu F, Wang L, Crook M, Maya S, Peng Z, Kelley K, Ane JM, Wang J. Isolation, Characterization, and Complete Genome Sequence of a Bradyrhizobium Strain Lb8 From Nodules of Peanut Utilizing Crack Entry Infection. Front Microbiol. 2020;11:93.
82. Seitzinger SP, Garber JH. Nitrogen-fixation and N-15(2) calibration of the acetylene-reduction assay in coastal marine-sediments. Mar Ecol Progr Ser. 1987;65–73:65–73.
83. Hardy RWFHR, Jackson EK, Burns RC. The Acetylene-Ethylene Assay for N2 fixation: laboratory and field evaluation. Plant Phys. 1968;43:1185–207.
84. Tanaka ADMA, Amato A, Montsant A, Mathieu B, Rostaing P, Tirichine L, Bowler C. Ultrastructure and Membrane Traffic During Cell Division in the Marine Pennate Diatom Phaeodactylum tricornutum. Protist. 2015;166:506–21.
85. Kieser S, Brown J, Zdobnov EM, Trajkovski M, McCue LA. ATLAS: a Snakemake workflow for assembly, annotation, and genomic binning of metagenome sequence data. BMC Bioinformatics. 2020;21:257.

86. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. Genome Res. 2017;27:824–34.
87. Wu YW, Simmons BA, Singer SW: MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 2016, 32:605–607.
88. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ. 2019;7: e7359.
89. Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, Banfield JF. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. Nat Microbiol. 2018;3:836–43.
90. Olm MR, Brown CT, Brooks B, Banfield JF. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. ISME J. 2017;11:2864–8.
91. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015;25:1043–55.
92. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010;11:119.
93. Steinegger M, Soding J. Clustering huge protein sequence sets in linear time. Nat Commun. 2018;9:2542.
94. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015;12:59–60.
95. Giguere DJ, Bahcheli, A.T., Joris, B.R., Paulssen, J.M., Gieg, L.M., Flatley, M.W., Gloor, G.B.: Complete and validated genomes from a metagenome.https://doi.org/10.1101/2020.04.08.032540 2020.
96. Meyer F, Lesker TR, Koslicki D, Fritz A, Gurevich A, Darling AE, Sczyrba A, Bremges A, McHardy AC. Tutorial: assessing metagenomics software with the CAMI benchmarking toolkit. Nat Protoc. 2021;16:1785–801.
97. Song WZ, Thomas T. Binning_refiner: improving genome bins through the combination of different binning programs. Bioinformatics. 2017;33:1873–5.
98. Coil DA, Jospin G, Darling AE, Wallis C, Davis IJ, Harris S, Eisen JA, Holcombe LJ, O'Flynn C. Genomes from bacteria associated with the canine oral cavity: A test case for automated genome-based taxonomic assignment. PLoS ONE. 2019;14: e0214354.
99. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. Bioinformatics. 2020;36(6):1925–27. https://doi.org/10.1093/bioinformatics/btz848.
100. Galperin MY, Makarova KS, Wolf YI, Koonin EV. Expanded microbial genome coverage and improved protein family annotation in the COG database. Nucleic Acids Res. 2015;43:D261-269.
101. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28:27–30.
102. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068–9.
103. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, Phillippy AM. Mash: fast genome and metagenome distance estimation using MinHash. Genome Biol. 2016;17:132.
104. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32:1792–7.
105. Pellow D, Zorea A, Probst M, Furman O, Segal A, Mizrahi I, Shamir R. SCAPP: an algorithm for improved plasmid assembly in metagenomes. Microbiome. 2021;9:144.
106. Antipov D, Raiko M, Lapidus A, Pevzner PA. Plasmid detection and assembly in genomic and metagenomic data sets. Genome Res. 2019;29:961–8.
107. Royer G, Decousser JW, Branger C, Dubois M, Medigue C, Denamur E, Vallenet D. PlaScope: a targeted approach to assess the plasmidome from genome assemblies at the species level. Microb Genom . 2018;4(9):e000211. https://doi.org/10.1099/mgen.0.000211.
108. Tanizawa Y, Fujisawa T, Kaminuma E, Nakamura Y, Arita M. DFAST and DAGA: web-based integrated genome annotation tools and resources. Biosci Microbiota Food Health. 2016;35:173–84.
109. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Mol Biol Evol. 2021;38(12):5825–9. https://doi.org/10.1093/molbev/msab293. PMID: 34597405; PMCID: PMC8662613.
110. Delmont T, Gaia M, Hinsinger DD, et al. Functional repertoire convergence of distantly related eukaryotic plankton lineages revealed by genome-resolved metagenomics. Cell Genomics. 2022;2(5):100123. https://doi.org/10.1016/j.xgen.2022.100123.
111. Delmont TO, Pierella Karlusich JJ, Veseli I, Fuessel J, Eren AM, Foster RA, Bowler C, Wincker P, Pelletier E. Heterotrophic bacterial diazotrophs are more abundant than their cyanobacterial counterparts in metagenomes covering most of the sunlit ocean. ISME J. 2022;16:927–36.
112. Tackmann J, Matias Rodrigues JF, von Mering C. Rapid Inference of Direct Interactions in Large-Scale Ecological Networks from Heterogeneous Microbial Sequencing Data. Cell Syst. 2019;9(286–296): e288.
113. Fernández JAM, Vidal, C.B.& Glahn, V.P. : Dealing with Zeros and Missing Values in Compositional Data Sets Using Nonparametric Imputation. Mathematical Geology 2003, 35: pages253–278.
114. Quinn TP, Erb I, Gloor G, Notredame C, Richardson MF, Crowley TM. A field guide for the compositional analysis of any-omics data. GigaScience. 2019;8(9):giz107. https://doi.org/10.1093/gigascience/giz107.
115. Chandola U, Gaudin, M., Trottier, C., Lavier Aydat, LJ., Manirakiza, E., Menicot, S., Jörg Fischer, E., Louvet, I., Lacour, T., Chaumier, T., Tanaka, A., Pohnert, P., Chaffron, S., and Tirichine, L. : Non-cyanobacterial diazotrophs support the survival of marine microalgae in nitrogen-depleted environments. 2025. https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA923279

Publisher's Note