

ARTICLE OPEN



Characterization of a bloom-associated alphaproteobacterial lineage, 'Candidatus Phycosocius': insights into freshwater algal-bacterial interactions

Yuuhiiko Tanabe ^{1,2}✉, Haruyo Yamaguchi ¹, Masaki Yoshida², Atsushi Kai² and Yusuke Okazaki³

© The Author(s) 2023, corrected publication 2023

Marine bacterial lineages associated with algal blooms, such as the *Roseobacter* clade, have been well characterized in ecological and genomic contexts, yet such lineages have rarely been explored in freshwater blooms. This study performed phenotypic and genomic analyses of an alphaproteobacterial lineage 'Candidatus Phycosocius' (denoted the CaP clade), one of the few lineages ubiquitously associated with freshwater algal blooms, and described a novel species: 'Ca. Phycosocius spiralis.' Phylogenomic analyses indicated that the CaP clade is a deeply branching lineage in the *Caulobacterales*. Pangenome analyses revealed characteristic features of the CaP clade: aerobic anoxygenic photosynthesis and essential vitamin B auxotrophy. Genome size varies widely among members of the CaP clade (2.5–3.7 Mb), likely a result of independent genome reductions at each lineage. This includes a loss of tight adherence pilus genes (*tad*) in 'Ca. P. spiralis' that may reflect its adoption of a unique spiral cell shape and corkscrew-like burrowing activity at the algal surface. Notably, quorum sensing (QS) proteins showed incongruent phylogenies, suggesting that horizontal transfers of QS genes and QS-involved interactions with specific algal partners might drive CaP clade diversification. This study elucidates the ecophysiology and evolution of proteobacteria associated with freshwater algal blooms.

ISME Communications; <https://doi.org/10.1038/s43705-023-00228-6>

INTRODUCTION

In marine environments, algal blooms often occur in early spring to early summer in the stratified euphotic zone, where nutrients are replenished annually through winter mixing. As such, marine algal blooms constitute part of marine biogeochemical cycles [1]. Each algal cell or colony in a marine bloom develops a phycosphere. This phycosphere attracts many heterotrophic bacteria that utilize dissolved organic matter, such as polysaccharides, that are rarely available in the open oligotrophic ocean [2]. Among the marine phycosphere bacteria, the *Roseobacter* clade is one of the most abundant bacterial groups in the pelagic environment and plays a significant role in global carbon and sulfur cycling [1]. Members of the *Roseobacter* clade are associated with a broad taxonomic range of algae, including dinoflagellates, diatoms, and haptophytes [1, 3]. Aside from its affinity to phytoplankton, a prominent characteristic of the *Roseobacter* clade is its metabolic versatility. These adaptations include anoxygenic photosynthesis, sulfur and carbon monoxide oxidations as sources of energy (chemolithotrophy), and algae-derived dimethylsulfoniopropionate (DMSP) assimilation [3]. In addition, members of the *Roseobacter* clade show chemotaxis toward algae or algae-derived dissolved organic matter and become sessile on the nutrient-rich algal cell surface. This is often called a 'swim-or-stick' lifestyle [4] and is facilitated by flagellar assembly and type IV pili expression [5]. These behaviors are controlled by a suite of phosphorelays centered on a global

transcriptional regulator, CtrA, with quorum-sensing (QS) also involved in its regulation [6].

In freshwater systems, algal blooms often occur in response to increasing nitrogen and phosphorus inputs resulting from human activity, although the occurrence of blooms predates human history [7]. Climate change is also thought to accelerate freshwater blooms [7]. Freshwater bloom-forming algae include *Cyanobacteria* (e.g., *Microcystis*, *Dolichospermum*, and *Aphanizomenon*), diatoms [8], and on rare occasions, green algae (e.g., *Botryococcus braunii*) [9]. Cyanobacterial blooms are known to produce cyanotoxins that are health hazards for both humans and wildlife [7]. Many freshwater bloom-forming algae form carbohydrate-rich colonies that contribute to a variety of phycospheres. A growing number of studies have revealed that *Alphaproteobacteria*, *Gammaproteobacteria*, and *Bacteroidetes* are major constituents of freshwater phycosphere bacterial communities [10–14]. However, ecophysiological interactions between freshwater bloom-forming algae and their phycosphere bacteria have been poorly described. One study suggested the presence of metabolic complementation between *Microcystis aeruginosa* and *Roseomonas* spp. [13]. Another study suggested that *B. braunii* supplemented biotin to *Brevundimonas* sp., which in return protected *B. braunii* from colonization by other harmful bacteria [14].

A previous study isolated a novel alphaproteobacterium, 'Candidatus Phycosocius bacilliformis,' from a bloom of *B. braunii*, a colony-forming green alga (*Chlorophyta*) [15]. *B. braunii* is unique

¹Biodiversity Division, National Institute for Environmental Studies, Ibaraki 305-8506, Japan. ²Algae Biomass and Energy System R&D Center, University of Tsukuba, Ibaraki 305-8572, Japan. ³Institute for Chemical Research, Kyoto University, Kyoto 611-0011, Japan. ✉email: ytanabehiko@gmail.com

Received: 18 October 2022 Revised: 23 February 2023 Accepted: 1 March 2023

Published online: 11 March 2023

in that it accumulates hydrocarbons within the colony and has thus attracted attention as a feedstock for biofuel production [15]. ‘*Ca. P. bacilliformis*’ 16S rDNA clustered with two bacterial sequences recovered from *Microcystis* cultures, and similar sequences (showing >99.4% DNA similarity) were frequently detected in *Botryococcus* and *Microcystis* blooms collected from lakes in Asia and Africa [15]. A subsequent metagenomic study of *Microcystis*-associated bacteria in 14 US lakes revealed that relatives of ‘*Ca. P. bacilliformis*’ were detected in all 46 *Microcystis* colonies and accounted for up to 19% of bacterial abundance in the *Microcystis* phycosphere [16]. These results suggest that ‘*Ca. P. bacilliformis*’ and its relatives represent a ubiquitous lineage closely associated with freshwater phytoplankton populations. Interestingly, ‘*Ca. P. bacilliformis*’ possesses genes for bacteriochlorophyll *a* (Bchl *a*), suggesting that this species is a photosynthetic bacterium [17]. However, knowledge of this lineage is limited due to the poor availability of genome sequences.

Previous research observed another bacterium, morphologically distinct from ‘*Ca. P. bacilliformis*,’ co-occurring with ‘*Ca. P. bacilliformis*’ in a nonaxenic culture of *B. braunii* [15]. ‘*Ca. P. bacilliformis*’ was rod-shaped, whereas the novel bacterium was spiral in shape (Fig. S1). In this study, the spiral bacterium was cultivated, characterized, and designated BOTRYCO-1. Analysis of 16S sequences revealed that BOTRYCO-1 was closely related to ‘*Ca. P. bacilliformis*,’ motivating further characterization of ‘*Ca. P. Phycosocius*.’ The BOTRYCO-1 genome sequence was obtained and used in a pangenome analysis of four ‘*Ca. Phycosocius*’ relatives. Together, these bacteria represented a unique alpha-proteobacterial lineage characterized by photoheterotrophy and close association with freshwater phytoplankton blooms.

MATERIALS AND METHODS

Cultures

B. braunii Ba10, a nonaxenic isolate recovered from a *B. braunii* bloom in a reservoir in Myanmar [15], was used to cultivate a spiral bacterium, designated BOTRYCO-1. A two-membered culture of a bacterium and *B. braunii* Ba10 was established by inoculating algal-free bacterial suspension (obtained by filtration of Ba10 using a Minisart® syringe filter, pore size 1.2 µm, Sartorius, Göttingen, Germany) into a bacteria-free culture of Ba10 (ref. 15). The culture was maintained in liquid AF-6 medium [18] at 25 °C under continuous white light (50 µmol photon m⁻²s⁻¹). The detailed cultivation protocol was as described previously [15].

Light and electron microscopic studies

Light microscopy was performed using an IX71 microscope with a DP72 camera (Olympus, Tokyo, Japan). Catalyzed reporter deposition-fluorescence *in situ* hybridization (CARD-FISH) was performed using a BAG645 probe specific to the *CaP* clade [15], as previously described [15]. Scanning electron microscopy (SEM) was performed using a JSM-6330F field emission SEM (JEOL, Ltd., Tokyo, Japan). Transmission electron microscopy (TEM) was performed using a TEM H-7650 (Hitachi, Tokyo, Japan). Sample preparation was as previously described for SEM [19] and TEM [15].

Whole genome analysis

BOTRYCO-1 cells were collected by Minisart® filtration of 80 mL of a two-membered liquid culture comprising BOTRYCO-1 and *B. braunii* Ba10. Genomic DNA was extracted using NucleoBond AXG columns with buffer set III (Macherey-Nagel, Düren, Germany). Whole-genome sequencing of BOTRYCO-1 was performed using the MiSeq platform (Illumina, San Diego, CA, USA) as previously described [17]. De novo assembly of short reads was performed using SPAdes ver. 3.14.1 (ref. 20), and assembled scaffolds were polished using Pilon ver 1.24 (ref. 21). After removing experimental contaminants (based on lower K-mer coverage and BLAST analysis), the remaining scaffolds were annotated using dFAST [22]. Genome completeness was assessed using CheckM [23]. Average nucleotide (ANI) and amino acid identities (AAI) were calculated using online ANI and AAI calculators [24].

Phylogenetic analyses

Genome sequences related to BOTRYCO-1 were retrieved from GenBank, and a phylogenomic tree was constructed using PhyloPhlAn 3.0 (ref. 25) using amino acid sequences of 400 universal marker genes (PhyloPhlAn database) [26]. The statistical confidence of each tree branch was evaluated using the bootstrap method with 100 replicates. The PufL and PufM sequences of type strains of the *Caulobacterales* related to BOTRYCO-1 were retrieved from GenBank and aligned using Clustal W [27]. Maximum likelihood (ML) phylogenetic analyses were performed using RaxML-NG ver. 1.0.1 (ref. 28). Modeltest-NG ver. 0.1.6 (ref. 29) was used to select the amino acid substitution model for ML phylogenetic reconstruction.

Pangenome analyses and functional assessment of coding sequences

Pangenome analyses were performed using the GET_HOMOLOGUES pipeline [30], and the result was visualized using ‘ggplot2’ in R ver. 4.0.5. Initial functional annotation of proteins was performed using eggNOG mapper v.2 (ref. 31). Using the outputs of the eggNOG mapper, KEGG pathway reconstruction [32] was used to infer metabolic pathways present for each genome. Enzyme presence was further inspected by individual BLAST searches.

RESULTS AND DISCUSSION

Phenotypic and phylogenetic characterization of a novel species of ‘*Ca. Phycosocius*’

TEM microscopy of *B. braunii* Ba10 indicated the presence of spiral bacteria inside the extracellular matrix (ECM; Fig. S2). This differed from ‘*Ca. P. bacilliformis*,’ which is not seen inside the ECM [15]. The spiral bacterium, BOTRYCO-1, showed no growth in a low nutrient medium in which ‘*Ca. P. bacilliformis*’ had exhibited transient growth (Table S1). A culture exclusively containing *B. braunii* and BOTRYCO-1 was established, and BOTRYCO-1 growth was abundant in cocultivation (Fig. 1A), reaching > 10⁴ cells/µL during the early stationary growth phase of *B. braunii*. BOTRYCO-1 was observed actively swimming in a spiral manner in the liquid medium (Movie S1) and was aggressively motile toward *B. braunii*.

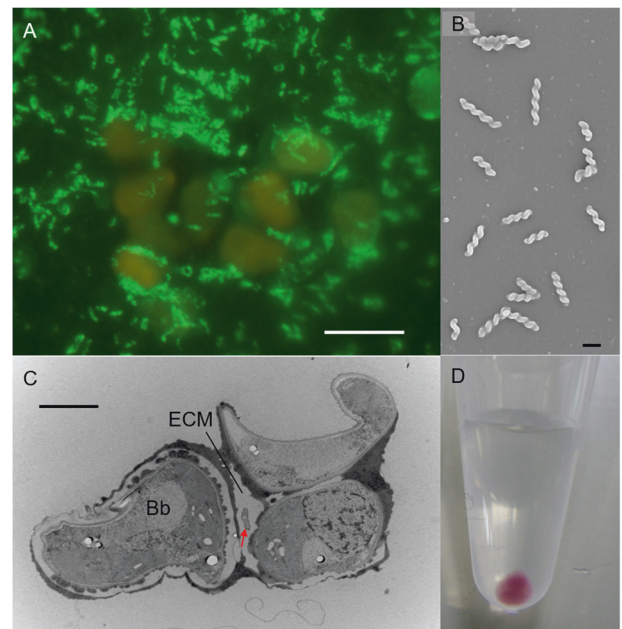


Fig. 1 Micrographs of BOTRYCO-1. **A** CARD-FISH image of the two-membered culture of *B. braunii* (orange) and BOTRYCO-1 (yellow-green). Scale bar, 10 µm. **B** SEM photo of BOTRYCO-1 showing a spiral shape. Scale bar, 1 µm. **C** TEM photo showing a BOTRYCO-1 cell inside the extracellular matrix (ECM) of *B. braunii*. Arrow indicates BOTRYCO-1. Bb, *B. braunii*. Scale bar, 2 µm. **D** Collected cells of BOTRYCO-1 in a 1.5 mL microtube.

BOTRYCO-1 burrowed through the ECM in a back-and-forth spiral motion near the surface of the *B. braunii* colony (Movie S2). Numerous nonmotile spiral cells were present deep inside the ECM, suggesting that, after successful colonization, the bacterium stopped moving and presumably lost its motility, reminiscent of the 'swim-or-stick' lifestyle of the *Roseobacter* clade [4]. SEM microscopy indicated that BOTRYCO-1 was 1.2–2.6 μm long (Fig. 1B), and TEM examination of the coculture confirmed that BOTRYCO-1 was present inside the ECM (Fig. 1C). Collected BOTRYCO-1 cells were burgundy-red (Fig. 1D), similar to those of '*Ca. P. bacilliformis*' [15].

BLAST analysis indicated that the BOTRYCO-1 16S rDNA sequence had 98.9–99.3% sequence similarity to about a dozen bacterial sequences, including those of '*Ca. P. bacilliformis*', the recently described species *Aquidulcibacter paucihalophilus* TH1-2^T [33], and a metagenome assembled genome UKL13-1 recovered from an *Aphanizomenon flos-aquae* cyanobacterial bloom [34]. Given the morphological dissimilarities (Table S1), the close relationships observed between BOTRYCO-1 and these two species were unexpected. Phylogenetic analysis of 16S rDNA indicated that these sequences, including the BOTRYCO-1 sequence, formed a monophyletic clade, hereafter denoted the *CaP* (*Candidatus* Phycosocius) clade (Fig. S3A, B). Notably, almost all sequences of this clade were recovered from isolates or colonies of freshwater microalgae (Table S1, Fig. S3B). Our literature survey identified *CaP* bacteria in freshwater phycosphere microbiomes worldwide (Table S2). Furthermore, analyses of global *Microcystis* microbiome data [12] revealed that *CaP* bacteria were distributed worldwide and were one of the dominant groups of bacteria (>1% relative abundance) in 4 of 12 *Microcystis* blooms examined (Table S3). These results confirm the hypothesis that the *CaP* clade is ubiquitously associated with freshwater blooms of broad taxonomic groups of algae [15, 16]. The *CaP* clade and two monospecific genera, *Vitreimonas* [35] and *Terricaulis* [36], constituted a deeply branching lineage of the *Caulobacterales* both in phylogenomic and 16S rDNA trees, although the branching order within the *CaP* clade was slightly different (Fig. 2A, Fig. S3A). '*Ca. P. bacilliformis*' 16S rDNA had 100% similarity to *A. paucihalophilus*; however, ANI and AAI values between the four genomes in the *CaP* clade were <95% (Fig. 2A), which is a proposed cutoff value used for bacterial species delineation [37, 38]. This suggests that each member of the *CaP* clade represents a different species. ANI values between *A. paucihalophilus* and '*Ca. P. bacilliformis*' (81.4%), and *A. paucihalophilus* and BOTRYCO-1 (79.6%), were higher than the mean genus cutoff value (73.9%) [39], indicating that '*Ca. P. bacilliformis*' and BOTRYCO-1 may be affiliated with the genus *Aquidulcibacter*. However, the colony colors of BOTRYCO-1 and '*Ca. P. bacilliformis*' were purplish, different from the yellowish color of *A. paucihalophilus* (Table S1) and from those of other known *Caulobacteraceae* species [36]. In addition, species affiliated with the *Caulobacterales* have various cell shapes, including rod, vibrioid, and fusiform, but, other than BOTRYCO-1, spiral-shaped cells have not been described [40]. On the basis of its distinct cell morphology and its genetic and phenotypic similarity to '*Ca. P. bacilliformis*', we tentatively propose '*Candidatus* Phycosocius spiralis' for BOTRYCO-1. Hereafter, '*Ca. P. spiralis*' is used for BOTRYCO-1.

Aerobic anoxygenic photosynthesis (AAnP) in the *CaP* clade

Whole-genome assembly of '*Ca. P. spiralis*' produced 46 contigs with an N50 value of 149 808 bp, with the longest contig being 351 315 bp and with total contig length of 2.5 Mb. CheckM analysis indicated 100% completeness and 0% contamination, suggesting that the estimated genome size approximated the true genome size of '*Ca. P. spiralis*', and that the recovered scaffolds covered almost all genes. The estimated genome size differed from other members of the *CaP* clade, with '*Ca. P. spiralis*' having the smallest genome and *A. paucihalophilus* having the largest

(3.7 Mb; Table S1). To infer the genomic basis underlying the phenotypic characteristics of *CaP* bacteria, pangenome analyses, including the two most closely related species of the *CaP* clade, *Terricaulis silvestris* and *Vitreimonas flagellata*, were performed. Results indicated that each *CaP* bacterium possessed a substantial number of proteins that were not shared with other *CaP* bacteria (Fig. 2B). The differences in genome size and protein contents among the species in the *CaP* clade (Fig. 2B) were explained either by genome reduction or by multiple large-scale horizontal gene transfers (HGT). Accumulating evidence favors the former because the increasing association between bacteria and other organisms has often driven genome reduction in bacterial associates [41, 42]. Regardless, the genome content heterogeneity among *CaP* bacteria likely results from independent adaptive evolution in response to different environments such as diverse bacteria-host algae interactions.

Genome annotation of '*Ca. P. spiralis*' revealed the presence of a complete set of genes for BChl *a*-based photosynthesis. These include genes for BChl *a* (*bch*) and carotenoid syntheses (*crt*), and light-harvesting and reaction center proteins (*puf* and *puh*). These genes were also found in the genome of '*Ca. P. bacilliformis*' and other genomes of the *CaP* clade, while being completely absent in *T. silvestris* and *V. flagellata* (Table S4). The photosynthesis genes in the *CaP* clade were located in the photosynthetic gene cluster (PGC) and were divided between at least two subclusters separated by 20–40 kbp (Fig. 3A). Gene contents within the PGCs resembled those of other phototrophic alphaproteobacteria [43]. A characteristic feature of the PGC in the *CaP* clade is the lack of *crtA*. This encodes spheroidene monooxygenase, which catalyzes the synthesis of spheroidenone, a photosynthetic carotenoid. By contrast, the PGCs of the *CaP* clade contain the complete set of genes for synthesis of another carotenoid spirilloxanthin [44]. Consistent with the presence of gene sets for two photosynthetic pigments, absorption peaks indicative of BChl *a* and spirilloxanthin were detected in the spectral data of '*Ca. P. spiralis*' (Fig. 3B). The purplish colony colors of '*Ca. P. spiralis*' (Fig. 1D) and '*Ca. P. bacilliformis*' [15] are typical for photosynthetic alphaproteobacteria [45] and are consistent with their photosynthetic potential. The yellow colony color of *A. paucihalophilus* (Table S1) suggests that yellow pigmented carotenoids are expressed in this species [46].

Comparison of PGCs among members of the *CaP* clade indicated that each of the PGC genes were highly similar, and gene organization was highly conserved (Fig. 3A), suggesting common ancestry for the PGCs of the *CaP* clade. One exceptional difference was the lack of *pufB* in the UKL-13 genome (Fig. 3A). The absence of either *pufA* or *pufB* in the PGC of alphaproteobacteria was reported previously [47]. *pufA* and *pufB* encode light-harvesting protein B-875 α and β chains, respectively [48]. Given the structural similarities between *pufA* and *pufB*, the *pufA* product likely compensates for the function of the *pufB* product in UKL-13. Notably, all the *CaP* genomes contained an aerobic oxidative cyclase gene, *acsF*, within the PGC as well as an oxygen-dependent coproporphyrinogen-III oxidase gene, *hemF*, both of which are hallmarks of AAnP [49, 50]. Their anaerobic counterparts, *bchE* and *hemN*, were not found in *CaP* genomes (Table S4). Calvin-Benson cycle enzymes (e.g., RuBisCO) were also absent in *CaP* genomes, whereas proteins involved in anaplerotic CO₂ assimilation, phosphoenolpyruvate carboxylase, and malate dehydrogenase were present (Fig. 4 and Table S4). The latter two enzymes catalyze the formation of oxaloacetate and malate by incorporating the carboxyl group from bicarbonate into phosphoenolpyruvate and pyruvate, respectively [51]. A gene encoding a sodium-dependent bicarbonate transporter (*Sbt*) was found in the genomes of the *CaP* clade (Fig. 4 and Table S4), and this was likely involved in bicarbonate uptake for anaplerotic CO₂ assimilation. Genes for lithoautotrophy were not found in the genomes of the *CaP* clade (Table S4). Together, these results

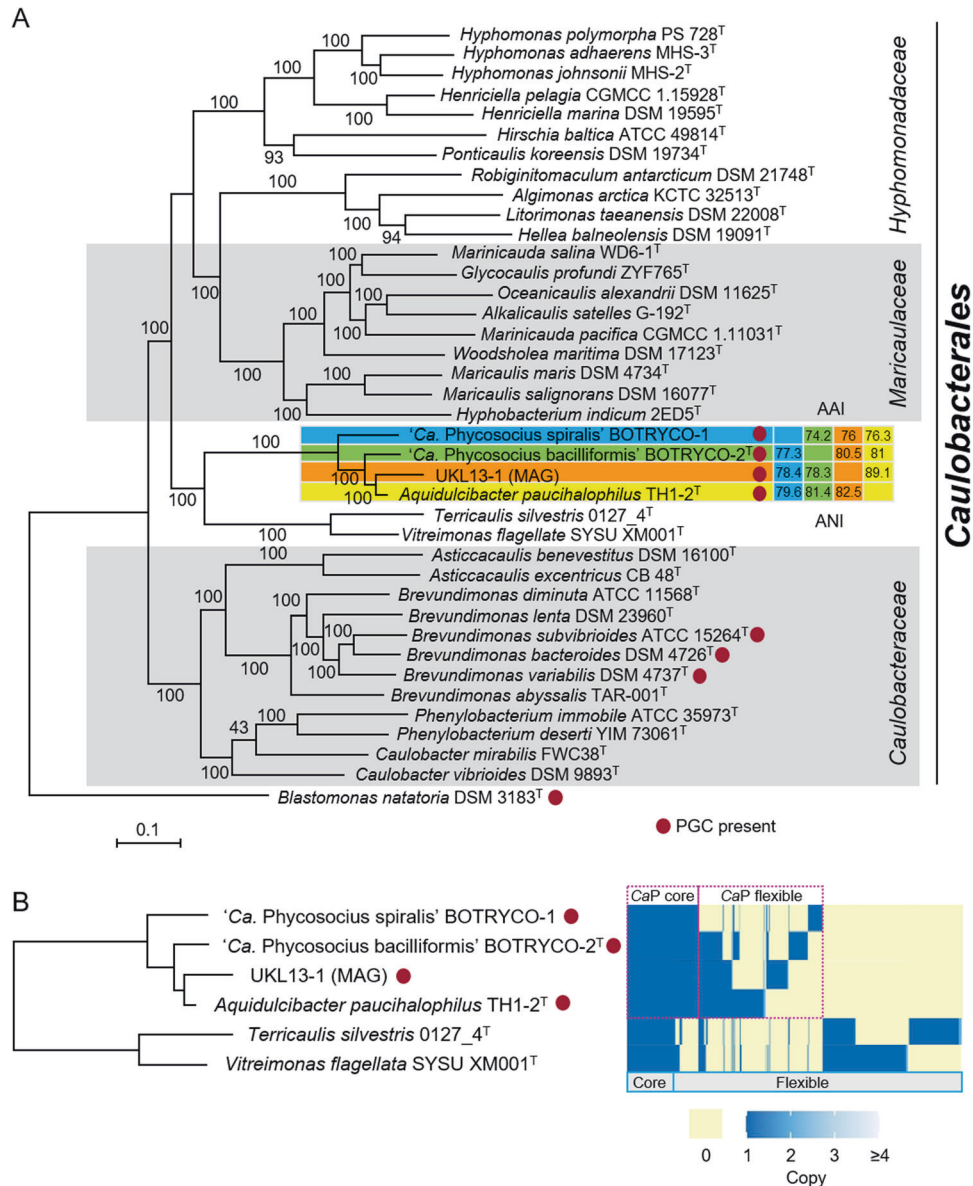


Fig. 2 Pangenome analysis of 'Ca. Phycosocius' and relatives. A ML phylogenomic tree rooted with *Blastomonas natoria*. ML bootstrap values (>80%) are indicated at each branch. Pairwise ANI and AAI values were indicated at the cluster, including 'Ca. P. spiralis'. **B** Distribution of orthologous gene clusters in the CaP clade. A heatmap indicates the presence and absence of orthologs. Core and flexible genomes indicate orthologs present in all six genomes and less than six genomes of CaP bacteria, respectively. CaP core and flexible indicate orthologs in all and less than four CaP genomes, respectively.

suggest that CaP members are aerobic anoxygenic photoheterotrophic bacteria.

Comprehensive genome mining suggested that the PGC is rare in the *Caulobacterales*; only three type strains of *Brevundimonas* species harbor a PGC (Fig. 2A). This sporadicity can be explained either by HGT and/or by repeated loss of the gene cluster across the *Caulobacterales*. The phylogenomic tree (Fig. 2A) and *pufLM* genealogy of the *Caulobacterales* (Fig. 3C) were concordant. Recent phylogenetic analysis of *Brevundimonas* showed that *Brevundimonas bacterioides* was most closely related to *B. variabilis*, whereas *B. subvibrioides* was more closely related to other *Brevundimonas* species that lacked a PGC [52]. This favors the HGT of PGC between the CaP clade and *Brevundimonas*, although repeated losses of PGC in the *Caulobacterales* cannot be ruled out. The HGT hypothesis is further supported by frequent HGT of PGC in AAnP bacteria [53].

Metabolic features of 'Ca. Phycosocius' and relatives

Central carbon metabolism. All members of the CaP clade share the same nonoxidative pentose phosphate and glycolysis (Entner-Doudoroff) pathways, but the Embden-Meyerhof pathway is incomplete due to the absence of phosphofructokinase (Fig. 4). This type of glycolysis is often found in Gram-negative bacteria, in which the Embden-Meyerhof pathway is adopted for gluconeogenesis [54, 55] and is appropriate for CaP bacteria because photosynthetic ATP production can reduce the need for glycolysis [56].

Differences in culturability in synthetic media (Table S1) suggest that CaP bacteria have different nutritional requirements. Consistent with this, genes involved in galactose oligomer metabolism and transport (*ganPQ*, *cyb*, and *msmX*) [57] are present in 'Ca. P. bacilliformis' and *A. paucihalophilus* but are absent in 'Ca. P. spiralis' (Fig. 4, Table S4 and Fig. S4). The presence of these

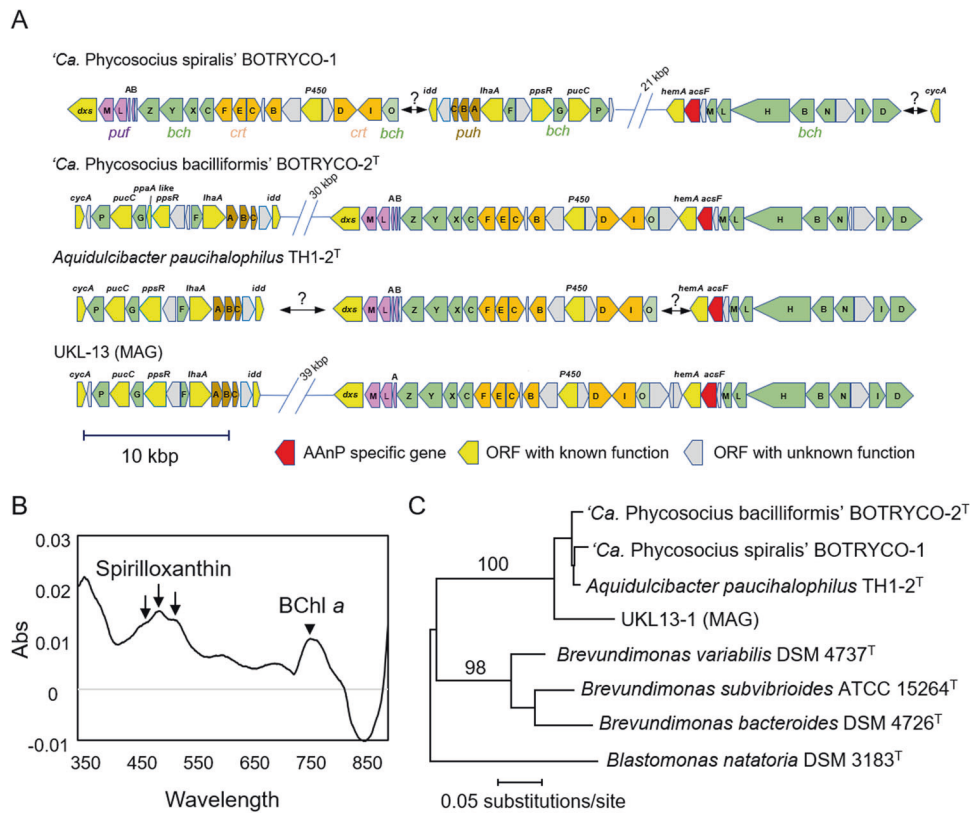


Fig. 3 Evidence for AANP in 'Ca. Phycosocius' and relatives. A Organization of PGC in the *CaP* clade. **B** Spectral analysis of methanol extract of photosynthesis pigments in 'Ca. *P. spiralis*.' Arrows and an arrowhead indicated putative absorbance peaks of spirilloxanthin [44] and bacteriochlorophyll *a* (BChl *a*), respectively. The protocol is described in Supplementary Methods. **C** ML phylogenetic tree of concatenated amino acid sequences of PufL and PufM. The tree was generated based on CPREV + G4 + F amino acid substitution model using a gap-free alignment (578 amino acids). ML bootstrap values (>80%) based on 1000 replicates are indicated at each branch.

genes in 'Ca. *P. bacilliformis*' corresponds with galactose as the major component of the ECM in *B. braunii* [58]. 'Ca. *P. bacilliformis*' and *A. paucihalophilus* likely utilize galactose oligomers as one of their carbon sources, whereas 'Ca. *P. spiralis*' utilizes different carbon sources for growth.

Metabolism related to algal-bacterial interactions. Thiamine and biotin are essential B vitamins (VBs) for virtually all living organisms and are needed by enzymes involved in aldehyde and carboxyl group translocation, respectively. VB intercomplementation is one of the metabolic highlights of algal-bacterial mutualism [59]. *CaP* bacteria lack genes essential for thiamine and biotin synthesis. Growth experiments indicated that axenic *B. braunii* could grow without thiamine and biotin [14], suggesting that *B. braunii* can synthesize both these VBs *de novo*. Putative transporters involved in thiamine uptake were detected in the genomes of *CaP* bacteria (Table S4), but biotin transporters are yet to be identified. *CaP* bacteria obtain VBs from associated algae (Fig. 4), other coexisting microbes, or both. Consistent with the former hypothesis, biotin uptake from *B. braunii* by phycosphere bacteria has been observed previously [14], although the transporters involved in VB export in *B. braunii* remain to be identified.

Cobalamin functions as a cofactor for methyl-group transfer in several metabolic reactions. A hallmark of the cobalamin requirement is its role as a cofactor in methionine synthesis [60]. *CaP* bacteria encode a cobalamin-dependent methionine synthase (MetH) but do not encode cobalamin-independent methionine synthase (MetE). *CaP* bacteria also possess a set of genes for cobalamin transport, including TonB-dependent cobalamin transporter (BtuB), a periplasmic cobalamin-binding protein (BtuF), and two putative ABC transporters for cobalamin (BtuC and BtuD; Fig. 4

and Table S4). However, cobalamin synthesis genes were absent in *CaP* bacteria. These observations suggest that *CaP* bacteria utilize cobalamin produced by bacteria coexisting in the *B. braunii* phycosphere.

Members of the *Roseobacter* clade are known to synthesize bioactive secondary metabolites, such as tropodithietic acid and roseobacticide, to protect or kill associated algae [61]. Genes for tropodithietic acid and roseobacticide productions (*tda*) were not found in *CaP* bacteria (Table S4). In addition, polyketide synthase and nonribosomal peptide synthetases, which are responsible for the production of many bioactive compounds in bacteria, were not found in *CaP* bacteria. Nonribosomal peptide synthetases were found in several genomes of the *Roseobacter* clade [62, 63]. One of their products, indigoidine, was shown to act against several competing algal pathogens, including *Vibrio* spp. [64]. Interactions between marine algae and *Roseobacter* bacteria via these bioactive compounds may govern the rise and fall of marine algal blooms [61]. In this context, the lack of bioactive compounds, either beneficial or harmful to algae, in *CaP* bacteria is unexpected. This is indicative of a different type of interaction between *CaP* bacteria and competing microbes and associated algae compared to bacteria of the *Roseobacter* clade. Unlike the *Roseobacter* clade, genes for DMSP assimilation were not found in *CaP* bacteria (e.g., *DmdA*, *DmdD*, and DMSP lyases) [65] (Table S4). Consistent with this, substantial DMSP production was observed only rarely in freshwater algae [66]. DMSP metabolisms involving marine algal blooms and the *Roseobacter* clade were proposed to have an essential role in marine carbon and sulfur cycles [1]. As sulfur concentrations in freshwater are much lower than in the ocean [67], freshwater algae and associated *CaP* bacteria are less likely to have developed similar interactions via sulfur compounds.

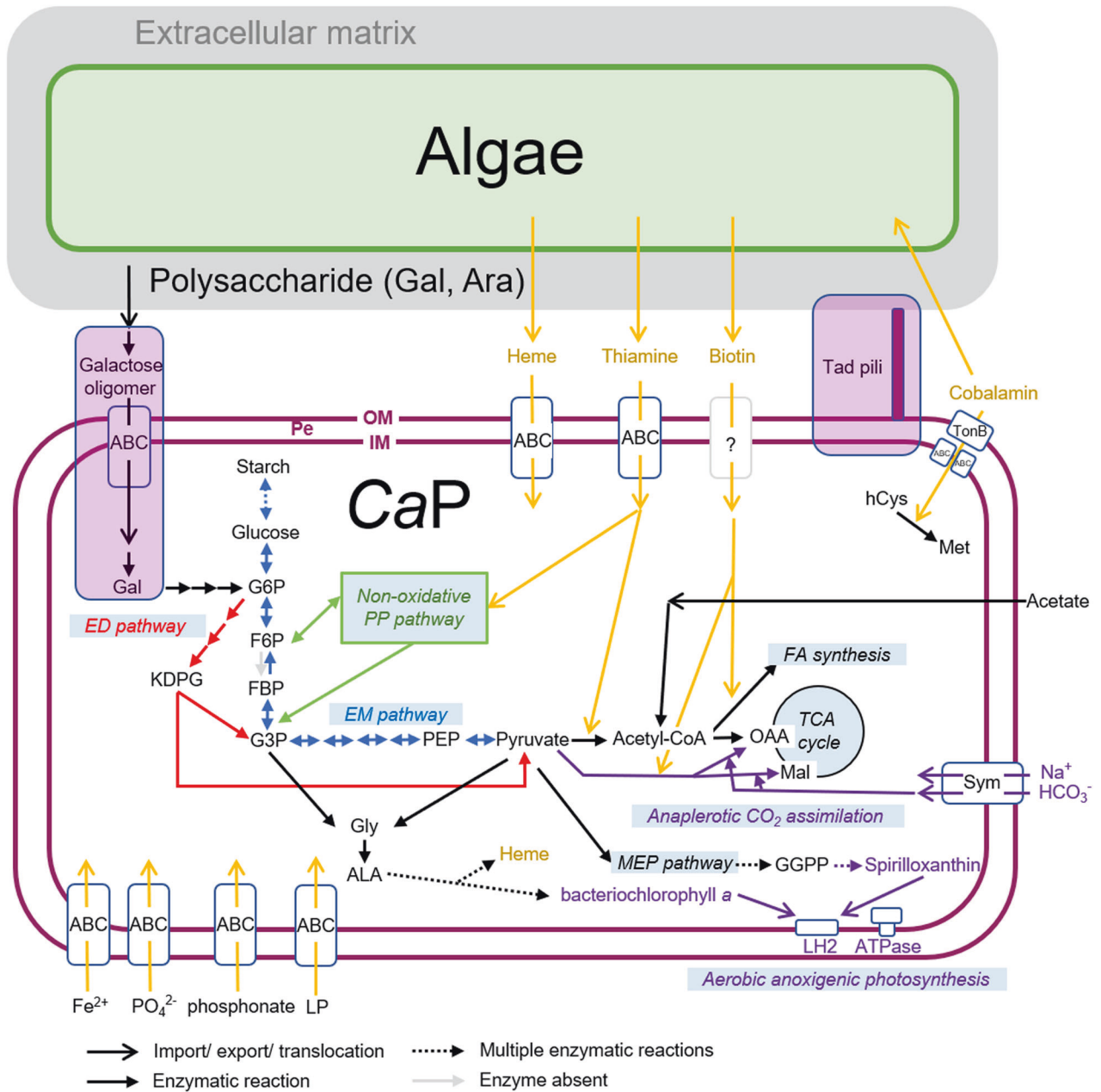


Fig. 4 Metabolic modeling of CaP clade bacteria. Features not shared in all CaP bacteria are indicated in the lavender box. Note that phosphofruktokinase is absent in CaP bacteria (indicated by a gray arrow). PE periplasmic space, OM outer membrane, IM inner membrane, Gal galactose, Ara arabinose, G6P glucose-6-phosphate, F6P fructose 6-phosphate, FBP fructose-1,6-bisphosphate, G3P glyceraldehyde-3-phosphate, PEP phosphoenolpyruvate, OAA oxaloacetate, Mal malate, FA fatty acid, hCys homocysteine, Met methionine, Gly glycine, Ala alanine, GGPP geranylgeranyl diphosphate, Sym symporter, LP lipoprotein.

QS and cell motility. Genes for QS were identified in all CaP bacterial genomes, as would be expected from the frequent observation of QS in bacterial pathogens and symbionts of eukaryotes [68]. Autoinducer synthase (LuxI) and response regulator (LuxR), a pair of proteins involved in QS [69], are present in '*Ca. P. spiralis*' and '*Ca. P. bacilliformis*', both of which have a green algal partner (Table S4). *A. paucihalophilus* and UKL-13, recovered from cyanobacteria, lack LuxI and contain only one LuxR. Such orphan LuxRs likely respond to exogenous QS molecules produced by coexisting microbes [68]. QS systems not only are involved in bacterial colonization of the algal surface [5] but also induce algal responses such as growth promotion or inhibition [70]. It is therefore possible that different modes of QS signaling might help to specify the algal partners of CaP bacteria,

and this can be elucidated by examination of host specificities for the different CaP bacteria. Phylogenetic analyses revealed discordant and unresolved phylogenies of LuxI and LuxR (Fig. S5A, B), suggesting complex evolutionary histories of QS genes, including within the CaP clade. The phylogenetic complexity can most easily be explained by HGT of QS genes. This is consistent with previous research suggesting that HGT events involving QS genes are frequent in proteobacteria [68]. Therefore, HGT of QS genes might have altered CaP bacterial preferences for specific algae, which in turn may have led to the divergence of the CaP clade.

Host-bacterial colonization is governed by chemotaxis [71], and it is therefore reasonable to assume that the 'aggressive' motility of '*Ca. P. spiralis*' toward *B. braunii* (Movie S2) occurs in response to

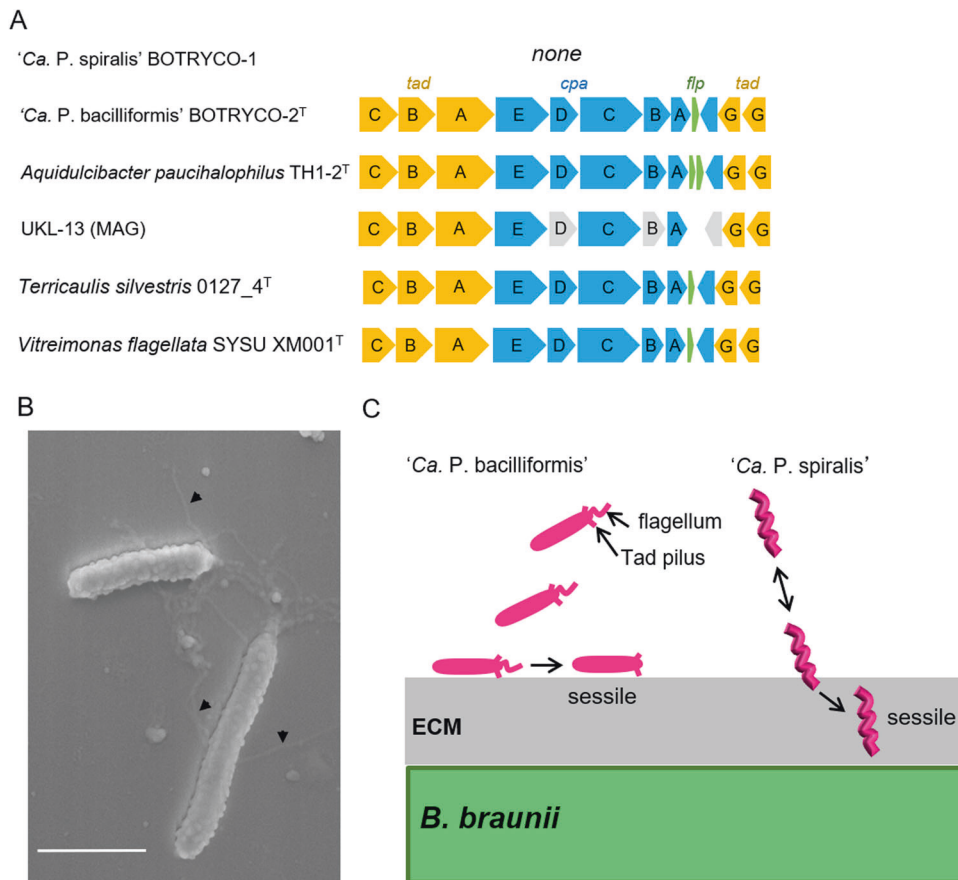


Fig. 5 Tad pili and their possible role in algal surface attachment. **A** Tad gene cluster in the *CaP* clade. Gray boxes indicate putative pseudogenes. **B** SEM image shows putative Tad pili in 'Ca. P. bacilliformis'. Scale bar, 1 μ m. **C** Schematic of different modes of attachment of two 'Ca. Phycosocius' species onto/into the algal surface.

molecules produced by the alga. This is supported by the presence of genes for chemotaxis in all *CaP* bacteria, namely, *cheA*, *cheB*, *cheD*, *cheR*, *cheW*, *cheX*, *cheY*, and *cheZ* in a *che* operon, and 2–4 chemotaxis receptor proteins (methyl-accepting chemotaxis proteins, MCP; Table S4). Different MCPs bind to different compounds, thereby directing chemotaxis [72]. Molecules implicated in bacterial chemotaxis toward algae include DMSP, amino acids, sugars, and carbohydrates [71]. At present, it is difficult to predict a chemical ligand for the *CaP* MCPs due to their low sequence similarities to characterized MCPs.

Genes involved in flagellum expression were found in all *CaP* bacteria (Table S4), although flagella were not observed in 'Ca. P. spiralis'. It is possible that flagella were expressed in hidden life stages, or that technical limitations hampered their observation. In the *Caulobacterales*, flagellar assembly is synchronized with the cell cycle, which is controlled by >20 regulatory proteins, including CtrA, PleC/D, and DivJ/K [73]. All *CaP* bacteria possess most of these regulatory genes (Table S4), suggesting that a similar regulatory pathway governs flagellar formation and the cell cycle.

Implications for algal-*CaP* bacterial interactions and their ecological impacts

Previous research showed that 'Ca. P. bacilliformis' promoted the growth of the host alga *B. braunii* during its growth, indicative of a mutualistic relationship [15]. However, preliminary growth experiments indicated that 'Ca. P. spiralis' had unstable effects (sometimes positive, sometimes neutral) on the growth of *B. braunii* (Fig. S6). It is possible that bacterial abundance relative to *B. braunii* abundance in the culture determines the fate of the algal-bacterial interaction [74]. The phycosphere of *B. braunii* in natural

environments harbors many microbes in addition to 'Ca. P. spiralis' and likely has multiple alga-microbe interactions. For example, a nonaxenic culture of *B. braunii* Ba10 harbored at least four species in addition to 'Ca. Phycosocius' (Fig. S7). The impact of 'Ca. P. spiralis' on the growth of *B. braunii* might therefore vary in a microbiome-dependent manner. One suggestion is that the *Roseobacter* clade and flavobacteria synergistically remineralize phytoplankton-derived organic matter in marine environments [1], and similar interactions might occur between *CaP* and other bacteria.

Given the smaller genome size of 'Ca. P. spiralis' compared to other *CaP* bacteria, our initial hypothesis was that 'Ca. P. bacilliformis' was ancestral to 'Ca. P. spiralis'. However, pangenome analyses suggest independent evolution for each lineage. Increasing the nutritional and metabolic dependence of 'Ca. P. spiralis' on associated algae may have accelerated its genome reduction, transforming 'Ca. P. spiralis' from a mutualist into a commensal. Both theoretical and empirical data suggest that symbiotic relationships are not stable and oscillate along the continuum between mutualism and parasitism [41]. In this context, *CaP* bacteria fluctuate between mutualism and commensalism.

One characteristic gene set that is absent in 'Ca. P. spiralis' but present in other members of the *CaP* clade is the Tad gene cluster, which encodes tight adherence (Tad) pili (a class of type IV pili) and associated regulatory proteins (Fig. 5A) [75]. Tad pili are widespread in bacteria, including *Caulobacter*, a distant relative of the *CaP* clade [76], and are involved in surface attachment and biofilm formation [75]. Swimming and sessile cells of 'Ca. P. spiralis' were observed off and on host algae, respectively

(Movie S2). This shift in cell motility was reminiscent of the 'swim-or-stick' lifestyle of algal-associated members of the *Roseobacter* clade [4], which use Tad pili to attach to algal partners [77]. SEM imaging captured a putative Tad pilus-like structure in '*Ca. P. bacilliformis*' in coculture with *B. braunii* (Fig. 5B). This suggests that all *CaP* bacteria except '*Ca. P. spiralis*' use Tad pili, allowing for a *Roseobacter*-like 'swim-or-stick' lifestyle. The loss of a Tad gene cluster in '*Ca. P. spiralis*' likely coincides with its acquisition of spiral motility and ECM burrowing, which renders attachment to the host surface no longer necessary (Fig. 5C). '*Ca. P. spiralis*' therefore adopts a unique strategy for its 'swim-or-stick' lifestyle that does not employ Tad pili for algal surface attachment. The Tad gene cluster in UKL-13 may be incomplete due to the pseudogenization of three genes, including *cpeE* and *cpeD* (Fig. 5A), and Tad pili might thus be absent in UKL-13. If so, UKL-13 might have a spiral-shaped morphology similar to that of '*Ca. P. spiralis*'.

Several features of the *CaP* clade, such as association with algal colonies and the metabolic potential of AANP, point to the clade having partial ecophysiological similarity to the marine *Roseobacter* clade. Of these features, the shared presence of AANP is of particular interest. A close correlation between marine AANP bacteria and occurrence of phytoplankton blooms has been reported. The reason for this correlation is not yet fully understood [1, 46]; however, it has been suggested that freshwater bloom-associated AANP bacteria play a role in maintaining redox balance in blooms via thiosulfate reduction and dissimilatory sulfate reduction pathways [12, 78]. *CaP* bacteria are unlikely to have a similar role because the genes involved in these processes (*sox*) are absent in *CaP* bacteria (Table S4). Field surveys indicate that AANP bacterial abundance is positively correlated to temporal phytoplankton bloom dynamics [79], and *CaP* members likely comprise some of these AANP bacteria. AANP bacteria show higher growth rates than nonphotosynthetic heterotrophs when optimal light is available due to their capacity to reduce respiration in order to produce ATP [80]. AANP bacteria are thus preferentially selected in organic matter-rich and sunlit environments like algal blooms in both marine and freshwater environments. This suggests that sunlight exposure owing to the association with blooms is one of the driving forces for the evolution of AANP in the *CaP* clade. A recent report indicated that lineages in *Gemmatimonadota* associated with phytoplankton are photoheterotrophs [81]. Many species of the alphaproteobacterium genus *Methylobacterium*, which are often found on plant leaves, also possess genes for AANP [47]. Photoheterotrophy might thus be one of the derived features in bacteria most closely associated with phototrophic organisms such as algae and plants. Conversely, close association with algae exposes associated bacteria to higher levels of sunlight-derived UV exposure. To cope with the DNA damage from UV exposure, bacteriorhodopsin-containing bacteria in marine environments, for example, possess more photolyase genes than bacteria armed with light-screening pigments [82]. With the exception of '*Ca. P. spiralis*', which has a reduced genome, this is also the case with *CaP* bacteria, which have three photolyase genes. Nonphotosynthetic *CaP* relatives have one or no photolyase genes (Table S4).

In conclusion, this study highlights the ecological and physiological characteristics of members of the *CaP* clade, which have adopted photoheterotrophic and VB auxotrophic growth in close association with algal blooms in freshwater. At least two species of the *CaP* clade, '*Ca. P. bacilliformis*' [15] and '*Ca. P. spiralis*', stick to the algal colony and subsequently become sessile. With the exception of VB auxotrophy, these ecophysiological features resemble those of the marine *Roseobacter* clade. In this context, the *CaP* clade is a freshwater counterpart of the *Roseobacter* clade. However, the *CaP* clade is genetically far less diverse (99% 16S rDNA sequence divergence) than the marine *Roseobacter* clade, which includes >50 genera encompassing

~89% 16S rDNA sequence diversity [1, 83]. This suggests that the *CaP* clade may have derived much later than the *Roseobacter* clade, although the possibility cannot be excluded that the diversity of the *CaP* clade was underestimated in this study owing to the poor availability of *CaP* genomes. Future studies include (1) unraveling the genomic and metabolic diversity of the *CaP* clade by including more strains; (2) exploring molecular interactions between *CaP* bacteria and their algal partners, including QS signaling; (3) investigating detailed population dynamics of *CaP* bacteria in relation to algal bloom dynamics; and (4) elucidating the metabolic regulation and balance between heterotrophic and phototrophic energy acquisitions of *CaP* bacteria. These studies will shed light on the ecological impact of the *CaP* clade on carbon cycling during algal blooms in freshwater environments. These studies will also allow examination of the hypothesis that evolution of photoheterotrophy in bacteria was driven by close association with phototrophic organisms, including microalgae.

DATA AVAILABILITY

Genome sequence data of '*Ca. P. spiralis*' BOTRYCO-1 are available at GenBank under the accession number BPFZ01000000.

REFERENCES

- Buchan A, LeCleir GR, Gulvik CA, González JM. Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nat Rev Microbiol.* 2014;12:686–98.
- Mühlenbruch M, Grossart HP, Eigemann F, Voss M. Mini-review: phytoplankton-derived polysaccharides in the marine environment and their interactions with heterotrophic bacteria. *Environ Microbiol.* 2018;20:2671–85.
- Luo H, Moran MA. Evolutionary ecology of the marine *Roseobacter* clade. *Microbiol Mol Biol Rev.* 2014;78:573–87.
- Belas R, Horikawa E, Aizawa SI, Suvanasthiti R. Genetic determinants of *Silicibacter* sp. TM1040 motility. *J Bacteriol.* 2009;191:4502–12.
- Fei C, Ochsenkühn MA, Shibl AA, Isaac A, Wang C, Amin SA. Quorum sensing regulates 'swim-or-stick' lifestyle in the phycosphere. *Environ Microbiol.* 2020;22:4761–78.
- Koppenhöfer S, Wang H, Scharfe M, Kaever V, Wagner-Döbler I, Tomasch J. Integrated transcriptional regulatory network of quorum sensing, replication control, and SOS response in *Dinoroseobacter shibae*. *Front Microbiol.* 2019;10:803.
- Huisman J, Codd GA, Paerl HW, Ibelings BW, Verspagen JMH, Visser PM. Cyanobacterial blooms. *Nat Rev Microbiol.* 2018;16:471–83.
- Reavie ED, Barbiero RP, Allinger LE, Warren GJ. Phytoplankton trends in the Great Lakes, 2001–2011. *J Gt Lakes Res.* 2014;40:618–39.
- Teng ST, Law K, Hanifah AH, Bojo OB, Idrus FA, Mannaf AH, et al. Bloom of a freshwater green alga *Botryococcus braunii* (*Botryococcaceae*, *Trebouxiophyceae*) and the associated mass fish mortality in a man-made lake, Sarawak, Malaysia. *Plankton Biosci Res.* 2021;16:59–68.
- Smith DJ, Tan JY, Powers MA, Lin XN, Davis TW, Dick GJ. Individual *Microcystis* colonies harbour distinct bacterial communities that differ by *Microcystis* oligo-type and with time. *Environ Microbiol.* 2021;23:5652–7.
- Jankowiak JG, Gobler CJ. The composition and function of microbiomes within *Microcystis* colonies are significantly different than native bacterial assemblages in two North American lakes. *Front Microbiol.* 2020;11:1016.
- Cook KV, Li C, Cai H, Krumholz LR, Hambright KD, Paerl HW, et al. The global *Microcystis* interactome. *Limnol Oceanogr.* 2020;65:5194–5207.
- Pérez-Carrascal OM, Tromas N, Terrat Y, Moreno E, Giani A, Corrêa Braga Marques L, et al. Single-colony sequencing reveals microbe-by-microbiome phylosymbiosis between the cyanobacterium *Microcystis* and its associated bacteria. *Microbiome.* 2021;9:1–21.
- Blifernez-Klassen O, Klassen V, Wibberg D, Cebeci E, Henke C, Rückert C, et al. Phytoplankton consortia as a blueprint for mutually beneficial eukaryote-bacteria ecosystems based on the biocoenosis of *Botryococcus* consortia. *Sci Rep.* 2021;11:1726.
- Tanabe Y, Okazaki Y, Yoshida M, Matsuura H, Kai A, Shiratori T, et al. A novel alphaproteobacterial ectosymbiont promotes the growth of the hydrocarbon-rich green alga *Botryococcus braunii*. *Sci Rep.* 2015;5:10467.
- Jackrel SL, White JD, Evans JT, Buffin K, Hayden K, Sarnelle O, et al. Genome evolution and host-microbiome shifts correspond with intraspecific niche divergence within harmful algal bloom-forming *Microcystis aeruginosa*. *Mol Ecol.* 2019;28:3994–4011.

17. Tanabe Y, Yamaguchi H, Watanabe MM. Draft genome sequence of "Candidatus Phycosocius bacilliformis," an alphaproteobacterial ectosymbiont of the hydrocarbon-producing green alga *Botryococcus braunii*. *Genome Announc*. 2018;6:e00396–18.
18. Kasai F, Kawachi M, Erata M, Watanabe MM. NIES-Collection, List of strains, microalgae and protozoa. 7th ed. National Institute for Environmental Studies, Tsukuba, Japan, 2004.
19. Takahashi Y, Yoshida M, Inouye I, Watanabe MM. *Fibrophrys columna* gen. nov., sp. nov.: A member of the family *Amphifiliidae*. *Eur J Protistol*. 2016;56:41–50.
20. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19:455–77.
21. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sarkthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One*. 2014;9:e112963.
22. Tanizawa Y, Fujisawa T, Nakamura Y. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics*. 2018;34:1037–9.
23. 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.
24. Rodriguez-R LM, Konstantinidis KT. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ Preprints*. 2016;4:e1900v1.
25. Asnicar F, Thomas AM, Beghini F, Mengoni C, Manara S, Manghi P, et al. Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. *Nat Commun*. 2020;11:1–10.
26. Segata N, Börnigen D, Morgan XC, Huttenhower C. PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. *Nat Commun*. 2013;4:1–11.
27. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007;23:2947–8.
28. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*. 2019;35:4453–5.
29. Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T. ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Mol Biol Evol*. 2020;37:291–4.
30. Contreras-Moreira B, Vinuesa P. GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Appl Environ Microbiol*. 2013;79:7696–701.
31. Cantalapiedra CP, Hernández-Plaza A, Letunic J, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol*. 2021;38:5825–9.
32. Kanehisa M, Sato Y. KEGG Mapper for inferring cellular functions from protein sequences. *Protein Sci*. 2020;29:28–35.
33. Cai H, Shi Y, Wang Y, Cui H, Jiang H. *Aquidulcibacter paucihalophilus* gen. nov., sp. nov., a novel member of family *Caulobacteraceae* isolated from cyanobacterial aggregates in a eutrophic lake. *Antonie Van Leeuwenhoek*. 2017;110:1169–77.
34. Driscoll CB, Otten TG, Brown NM, Dreher TW. Towards long-read metagenomics: complete assembly of three novel genomes from bacteria dependent on a diazotrophic cyanobacterium in a freshwater lake co-culture. *Stand Genom. Sci*. 2017;12:1–6.
35. Asem MD, Salam N, Zheng W, Liao LH, Zhang XT, Han MX, et al. *Vitreimonas flagellata* gen. nov., sp. nov., a novel member of the family *Hyphomonadaceae* isolated from an activated sludge sample. *Int J Syst Evol Microbiol*. 2020;70:2632–9.
36. Vieira S, Pascual J, Boedeker C, Geppert A, Riedel T, Rohde M, et al. *Terricaulis silvestris* gen. nov., sp. nov., a novel prothecate, budding member of the family *Caulobacteraceae* isolated from forest soil. *Int J Syst Evol Microbiol*. 2020;70:4966–77.
37. Thompson CC, Chimento L, Edwards RA, Swings J, Stackebrandt E, Thompson FL. Microbial genomic taxonomy. *BMC Genomics*. 2013. 2013;14:913.
38. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun*. 2018;9:1–8.
39. Barco RA, Garrity GM, Scott JJ, Amend JP, Neelson KH, Emerson D. A genus definition for Bacteria and Archaea based on a standard genome relatedness index. *mBio*. 2020;11:e02475–19.
40. Poindexter JS *Caulobacter*. In: Trujillo ME, Dedysch S, DeVos P, Hedlund B, Kämpfer P, Rainey FA, et al. (eds.) *Bergey's Manual of Systematics of Archaea and Bacteria*. 2015; <https://doi.org/10.1002/9781118960608.gbm00792>.
41. Drew GC, Stevens EJ, King KC. Microbial evolution and transitions along the parasite–mutualist continuum. *Nat Rev Microbiol*. 2021;19:623–38.
42. Foster RA, Zehr JP. Diversity, genomics, and distribution of phytoplankton-cyanobacterium single-cell symbiotic associations. *Annu Rev Microbiol*. 2019;73:435–56.
43. Kopejtká K, Tomasch J, Zeng Y, Tichý M, Sorokin DY, Koblížek M. Genomic analysis of the evolution of phototrophy among haloalkaliphilic *Rhodobacterales*. *Genome Biol Evol*. 2017;9:1950–62.
44. Kovács AT, Rákhely G, Kovács KL. Genes involved in the biosynthesis of photosynthetic pigments in the purple sulfur photosynthetic bacterium *Thiocapsa roseopersicina*. *Appl Environ Microbiol*. 2003;69:3093–102.
45. Imhoff JF, Rahn T, Künzel S, Neulinger SC. Photosynthesis is widely distributed among *Proteobacteria* as demonstrated by the phylogeny of PufLM reaction center proteins. *Front Microbiol*. 2018;8:2679.
46. Koblížek M. Ecology of aerobic anoxygenic phototrophs in aquatic environments. *FEMS Microbiol Rev*. 2015;39:854–70.
47. Zervas A, Zeng Y, Madsen AM, Hansen LH. Genomics of aerobic photoheterotrophs in wheat phyllosphere reveals divergent evolutionary patterns of photosynthetic genes in *Methylobacterium* spp. *Genome Biol. Evol*. 2019;11:2895–908.
48. Kiley PJ, Donohue TJ, Havelka WA, Kaplan S. DNA sequence and in vitro expression of the B875 light-harvesting polypeptides of *Rhodobacter sphaeroides*. *J Bacteriol*. 1987;169:742–50.
49. Pinta V, Picard M, Reiss-Husson F, Astier C. *Rubrivivax gelatinosus acsF* (previously *orf358*) codes for a conserved, putative binuclear-iron-cluster-containing protein involved in aerobic oxidative cyclization of Mg-protoporphyrin IX monomethylester. *J Bacteriol*. 2002;184:746–53.
50. Zeilstra-Ryalls JH, Schornberg KL. Analysis of *hemF* gene function and expression in *Rhodobacter sphaeroides* 2.4. 1. *J. Bacteriol*. 2006;188:801–4.
51. Tang KH, Tang YJ, Blankenship RE. Carbon metabolic pathways in phototrophic bacteria and their broader evolutionary implications. *Front Microbiol*. 2011;2:165.
52. Lee YW, Lee KH, Lee SY, Im WT. *Brevundimonas fluminis* sp. nov., isolated from a river. *Int J Syst Evol Microbiol*. 2020;70:204–10.
53. Brinkmann H, Göker M, Koblížek M, Wagner-Döbler I, Petersen J. Horizontal operon transfer, plasmids, and the evolution of photosynthesis in *Rhodobacteraceae*. *ISME J*. 2018;2:1994–2010.
54. Klingner A, Bartsch A, Dogs M, Wagner-Döbler I, Jahn D, Simon M, et al. Large-Scale ¹³C flux profiling reveals conservation of the Entner-Doudoroff pathway as a glycolytic strategy among marine bacteria that use glucose. *Appl Environ Microbiol*. 2015;81:2408–22.
55. Jyoti P, Shree M, Joshi C, Prakash T, Ray SK, Satapathy SS, et al. The Entner-Doudoroff and nonoxidative pentose phosphate pathways bypass glycolysis and the oxidative pentose phosphate pathway in *Ralstonia solanacearum*. *mSystems*. 2020;5:e00091–20.
56. Hauruseu D, Koblížek M. Influence of light on carbon utilization in aerobic anoxygenic phototrophs. *Appl Environ Microbiol*. 2012;78:7414–9.
57. Watzlawick H, Morabbi Heravi K, Altenbuchner J. Role of the *ganSPQAB* operon in degradation of galactan by *Bacillus subtilis*. *J. Bacteriol*. 2016;198:2887–96.
58. Weiss TL, Roth R, Goodson C, Vitha S, Black I, Azadi P, et al. Colony organization in the green alga *Botryococcus braunii* (Race B) is specified by a complex extracellular matrix. *Eukaryot Cell*. 2012;11:1424–40.
59. Cooper MB, Kazamia E, Helliwell KE, Kudahl UJ, Sayer A, Wheeler GL, et al. Cross-exchange of B-vitamins underpins a mutualistic interaction between *Ostreococcus tauri* and *Dinoroseobacter shibae*. *ISME J*. 2019;13:334–45.
60. Croft MT, Lawrence AD, Raux-Deery E, Warren MJ, Smith AG. Algae acquire vitamin B12 through a symbiotic relationship with bacteria. *Nature*. 2005;438:90–3.
61. Wang R, Gallant É, Seyedsayamdst MR. Investigation of the genetics and biochemistry of roseobactin production in the *Roseobacter* clade bacterium *Phaeobacter inhibens*. *mBio*. 2016;7:e02118–15.
62. Martens T, Gram L, Grossart HP, Kessler D, Müller R, Simon M, et al. Bacteria of the *Roseobacter* clade show potential for secondary metabolite production. *Microb Ecol*. 2007;54:31–42.
63. Wagner-Döbler I, Ballhausen B, Berger M, Brinkhoff T, Buchholz I, Bunk B, et al. The complete genome sequence of the algal symbiont *Dinoroseobacter shibae*: a hitchhiker's guide to life in the sea. *ISME J*. 2010;4:61–77.
64. Cude WN, Mooney J, Tavanaei AA, Hadden MK, Frank AM, Gulvik CA, et al. Production of the antimicrobial secondary metabolite indigoidine contributes to competitive surface colonization by the marine roseobacter *Phaeobacter* sp. strain Y4I. *Appl Environ Microbiol*. 2012;78:4771–80.
65. Bullock HA, Luo H, Whitman WB. Evolution of dimethylsulfoniopropionate metabolism in marine phytoplankton and bacteria. *Front Microbiol*. 2017;8:637.
66. Ginzburg B, Chalifa I, Gun J, Dor I, Hadas O, Lev O. DMS formation by dimethylsulfoniopropionate route in freshwater. *Environ Sci Technol*. 1998;2:2130–6.
67. Giordano M, Norici A, Hell R. Sulfur and phytoplankton: acquisition, metabolism and impact on the environment. *New Phytol*. 2005;166:371–82.

68. Case R, Labbate M, Kjelleberg S. AHL-driven quorum-sensing circuits: their frequency and function among the *Proteobacteria*. *ISME J*. 2008;2:345–9.
69. Cude WN, Buchan A. Acyl-homoserine lactone-based quorum sensing in the *Roseobacter* clade: complex cell-to-cell communication controls multiple physiologies. *Front Microbiol*. 2013;4:336.
70. Stock F, Bilcke G, De Decker S, Osuna-Cruz CM, Van den Berge K, Vancaester E, et al. Distinctive growth and transcriptional changes of the diatom *Seminavis robusta* in response to quorum sensing related compounds. *Front Microbiol*. 2020;11:1240.
71. Raina JB, Fernandez V, Lambert B, Stocker R, Seymour JR. The role of microbial motility and chemotaxis in symbiosis. *Nat Rev Microbiol*. 2019;17:284–94.
72. Webre DJ, Wolanin PM, Stock JB. Bacterial chemotaxis. *Curr Biol*. 2003;13:R47–49.
73. van Teeseling MCF, Thanbichler M. Generating asymmetry in a changing environment: cell cycle regulation in dimorphic alphaproteobacteria. *Biol Chem*. 2020;401:1349–63.
74. Morris MM, Kimbrel JA, Geng H, Tran-Gyamfi MB, Yu ET, Sale KL, et al. Bacterial community assembly, succession, and metabolic function during outdoor cultivation of *Microchloropsis salina*. *mSphere*. 2022;22:e00231–22.
75. Tomich M, Planet PJ, Figurski DH. The tad locus: postcards from the widespread colonization island. *Nat. Rev. Microbiol*. 2007;5:363–75.
76. Sangermani M, Hug I, Sauter N, Pfohl T, Jenal U. Tad pili play a dynamic role in *Caulobacter crescentus* surface colonization. *mBio*. 2019;10:e01237–19.
77. Isaac A, Francis B, Amann RI, Amin SA. Tight adherence (Tad) pilus genes indicate putative niche differentiation in phytoplankton bloom associated *Rhodobacteriales*. *Front Microbiol*. 2021;2:718297.
78. Li Q, Lin F, Yang C, Wang J, Lin Y, Shen M, et al. A large-scale comparative metagenomic study reveals the functional interactions in six bloom-forming *Microcystis*-epibiont communities. *Front Microbiol*. 2018;9:746.
79. Cepáková Z, Hrouzek P, Žiškova E, Nuyanzina-Boldareva E, Šorf M, Kozlíková-Zapomělová E, et al. High turnover rates of aerobic anoxygenic phototrophs in European freshwater lakes. *Environ Microbiol*. 2016;18:5063–71.
80. Piwosz K, Vrdoljak A, Frenken T, González-Olalla JM, Šantić D, McKay R, et al. Light and primary production shape bacterial activity and community composition of aerobic anoxygenic phototrophic bacteria in a microcosm experiment. *mSphere*. 2020;5:e00354–20.
81. Mujakić I, Andrei AŞ, Shabarova T, Fecskeová LK, Salcher MM, Piwosz K, et al. Common presence of phototrophic *Gemmatimonadota* in temperate freshwater lakes. *mSystems*. 2021;6:e01241–20.
82. Kumagai Y, Yoshizawa S, Nakajima Y, Watanabe M, Fukunaga T, Ogura Y, et al. Solar-panel and parasol strategies shape the proteorhodopsin distribution pattern in marine *Flavobacteriia*. *ISME J*. 2018;12:1329–43.
83. Buchan A, González JM, Moran MA. Overview of the marine roseobacter lineage. *Appl Environ Microbiol*. 2005;71:5665–77.

ACKNOWLEDGEMENTS

We thank Nanda Kyaw Thu for the isolation of BOTRYCO-1, Mina Nomizu for the guidance of TEM sample preparation, Masahiro Koide for taking video images, and

Iwane Suzuki and Makoto M. Watanabe for the use of their laboratory facilities. We also thank Katherine Cook and David Hambright for providing the raw data of their work on worldwide *Microcystis* blooms. This work is partly supported by Next-generation Energies for Tohoku recovery (NET) from Reconstruction Agency.

AUTHOR CONTRIBUTIONS

YT designed the research, YT performed isolation, culture, light microscopy, pigment, whole genome, and bioinformatics analyses, HY performed NGS, MY and AK performed electron microscopy, YO performed CARD-FISH and 16S amplicon analyses, YT and YO wrote the paper. All authors read and approved the final manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s43705-023-00228-6>.

Correspondence and requests for materials should be addressed to Yuuhiko Tanabe.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023, corrected publication 2023