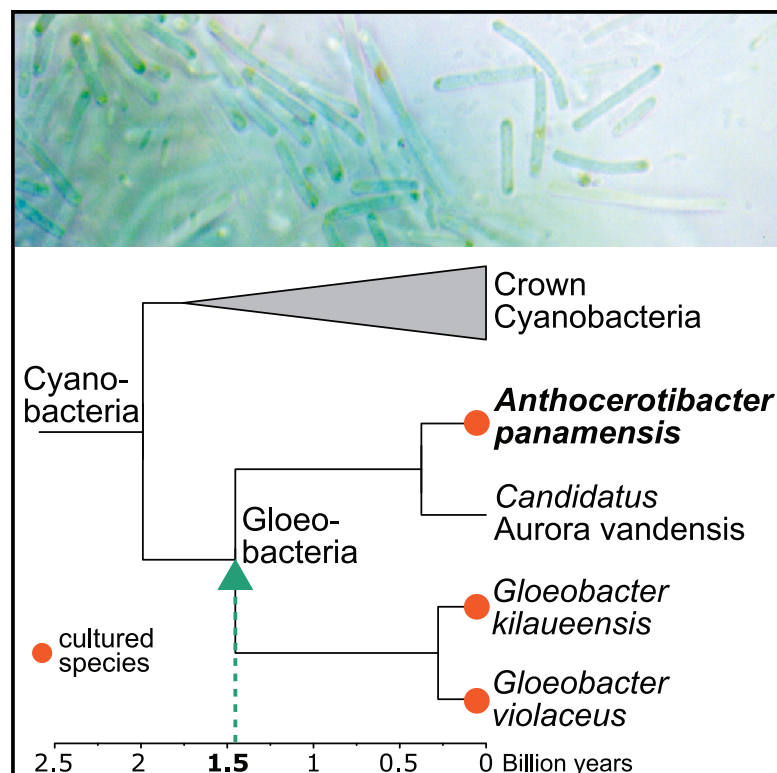


Current Biology

A novel thylakoid-less isolate fills a billion-year gap in the evolution of Cyanobacteria

Graphical abstract



Authors

Nasim Rahmatpour, Duncan A. Hauser, Jessica M. Nelson, Pa Yu Chen, Juan Carlos Villarreal A., Ming-Yang Ho, Fay-Wei Li

Correspondence

mingyang@ntu.edu.tw (M.-Y.H.), fl329@cornell.edu (F.-W.L.)

In brief

Very few cultured species exist in Gloeobacteria, an enigmatic lineage that is sister to the crown Cyanobacteria. Rahmatpour et al. discover a new and deeply diverged member of Gloeobacteria. This new species exhibits a suite of unique morphological, physiological, and genomic features, thus shedding new light on the evolution of Cyanobacteria.

Highlights

- A new thylakoid-less cyanobacterium cultured from a tropical hornwort
- This isolate is sister to a clade of environmental samples from the polar regions
- It diverged >1.4 Ga ago from the closest cultured taxa, *Gloeobacter* spp.
- Unique gene repertoires of phycobilisome and photosystems were uncovered



Article

A novel thylakoid-less isolate fills a billion-year gap in the evolution of Cyanobacteria

Nasim Rahmatpour,¹ Duncan A. Hauser,¹ Jessica M. Nelson,¹ Pa Yu Chen,² Juan Carlos Villarreal A.,^{3,4} Ming-Yang Ho,^{2,5,*} and Fay-Wei Li^{1,6,7,*}

¹Boyce Thompson Institute, Ithaca, NY, USA

²Department of Life Science, National Taiwan University, Taipei, Taiwan

³Department of Biology, Laval University, Quebec City, QC, Canada

⁴Smithsonian Tropical Research Institute, Panama City, Panama

⁵Institute of Plant Biology, National Taiwan University, Taipei, Taiwan

⁶Plant Biology Section, Cornell University, Ithaca, NY, USA

⁷Lead contact

*Correspondence: mingyang@ntu.edu.tw (M.-Y.H.), fl329@cornell.edu (F.-W.L.)

<https://doi.org/10.1016/j.cub.2021.04.042>

SUMMARY

Cyanobacteria have played pivotal roles in Earth's geological history, especially during the rise of atmospheric oxygen. However, our ability to infer the early transitions in Cyanobacteria evolution has been limited by their extremely lopsided tree of life—the vast majority of extant diversity belongs to Phycobacteria (or “crown Cyanobacteria”), while its sister lineage, Gloeobacteria, is depauperate and contains only two closely related species of *Gloeobacter* and a metagenome-assembled genome. Here, we describe a new cultured member of Gloeobacteria, *Anthocerotibacter panamensis*, isolated from a tropical hornwort. *Anthocerotibacter* diverged from *Gloeobacter* over 1.4 Ga ago and has low 16S rDNA identities with environmental samples. Our ultrastructural, physiological, and genomic analyses revealed that this species possesses a unique combination of traits that are exclusively shared with either Gloeobacteria or Phycobacteria. For example, similar to *Gloeobacter*, it lacks thylakoids and circadian clock genes, but the carotenoid biosynthesis pathway is typical of Phycobacteria. Furthermore, *Anthocerotibacter* has one of the most reduced gene sets for photosystems and phycobilisomes among Cyanobacteria. Despite this, *Anthocerotibacter* is capable of oxygenic photosynthesis under a wide range of light intensities, albeit with much less efficiency. Given its key phylogenetic position, distinct trait combination, and availability as a culture, *Anthocerotibacter* opens a new window to further illuminate the dawn of oxygenic photosynthesis.

INTRODUCTION

The rise of atmospheric oxygen is undoubtedly one of the most transformative events in Earth's history. During the Paleoproterozoic era, oxygenic photosynthesis carried out by Cyanobacteria fueled the Great Oxygenation Event, which altered the biogeochemical cycles and fundamentally shifted the evolutionary trajectories of life on Earth.^{1,2} None of the closest living relatives of Cyanobacteria, Vampirovibronia and Sericytochromatia,³ are capable of oxygenic photosynthesis or carry any intermediate photosynthetic machinery. Therefore, to understand how photosynthesis evolved over time, it is imperative to identify and examine the lineages that diverged the deepest within the Cyanobacteria tree of life.

The phylum Cyanobacteria (defined by Garcia-Pichel et al.⁴) is composed of two extant groups: Gloeobacteria and Phycobacteria that diverged around 2 Ga ago (bya).^{5,6} Phycobacteria encompasses >99.9% of the known cyanobacterial diversity and is sometimes also referred to as the “crown Cyanobacteria.”

Gloeobacteria, on the other hand, is rather enigmatic and has only two species described thus far: *Gloeobacter violaceus* and *G. kilaueensis*. Little is known about the ecology and distribution of these two species; *G. kilaueensis* is only known from a lava cave in Hawaii,⁷ and *G. violaceus* was originally isolated from a limestone rock in the Swiss alps⁸ and subsequently reported on waterfall walls in Europe and Mexico.^{9,10} The diversity of Gloeobacteria is likely much greater than these two species, as evidenced by a few distinct 16S rDNA clades from environmental samples¹¹ and a recent metagenome-assembled genome (MAG), *Candidatus Aurora vandensis*, from Lake Vanda, Antarctica.¹²

Gloeobacteria is a pivotal lineage to reconstruct the early evolution of Cyanobacteria, given their many distinct and presumably pleisiomorphic traits. In contrast to Phycobacteria, *Gloeobacter* lacks thylakoid membranes as well as a circadian clock.^{8,13} Due to the absence of thylakoids, the photosynthesis and respiratory apparatus are both located on the cytoplasmic membrane.¹⁴ The *Gloeobacter* photosystems are also atypical,^{15–19} and their phycobilisomes are bundle shaped instead



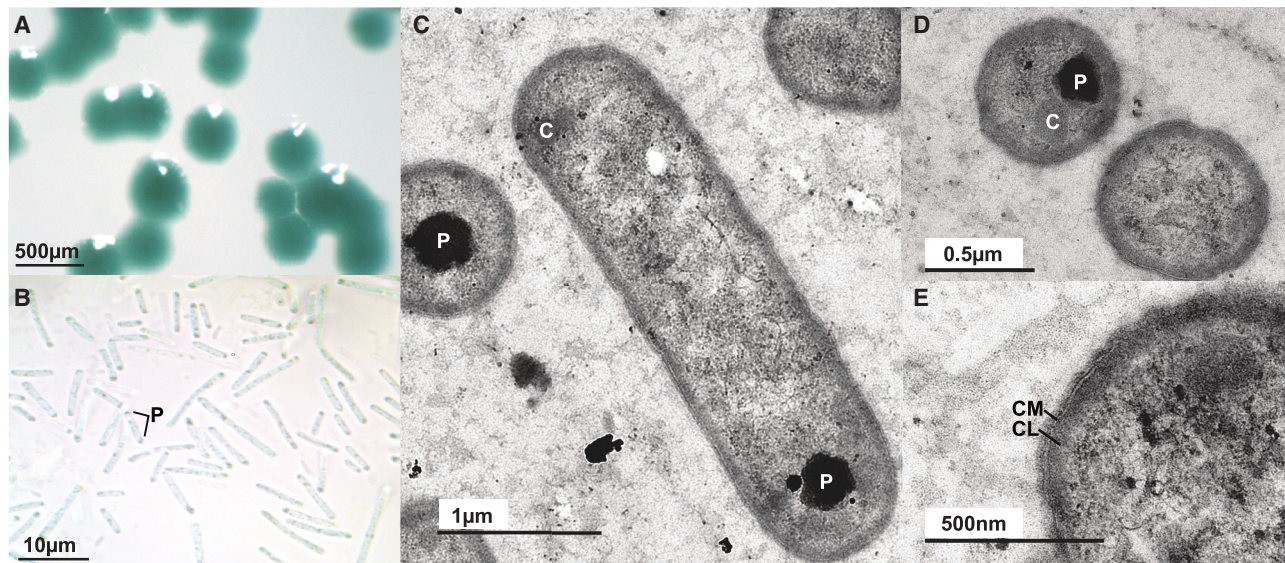


Figure 1. Morphology of *Anthocerotibacter panamensis*

(A) Blue-green colonies on a solid agar plate.

(B) Rod-shaped cells under light microscope with polyphosphate granules (P) often visible toward the cell poles.

(C–E) TEM images of *Anthocerotibacter* in longitudinal (C) and transverse (D and E) sections. No thylakoid is present. C, carboxysome; CL, electron-dense cytoplasmic layer; CM, cytoplasmic membrane.

of hemidiscoidal, as found in *Phycobacteria*.^{20,21} In addition, the carotenoid biosynthesis pathway in *Gloeobacter* is of “bacteria-type,”^{22,23} which is again different from other photosynthetic organisms. To date, a large body of research has been done on *Gloeobacter* and has significantly advanced our knowledge on Cyanobacteria biology and evolution.

However, the depauperate nature of *Gloeobacteria* makes it difficult to ascertain whether the idiosyncrasies of *Gloeobacter* can properly reflect the ancestral states or are instead a product of lineage-specific reductive evolution. The recent discovery of *Candidatus Aurora* MAG is an important step toward capturing the broader genomic diversity of *Gloeobacteria*,¹² although the fragmented nature of MAG, as well as its absence as a cultured organism, prohibits further in-depth investigations.

Here, we report a new cultured *Gloeobacteria* species that is distantly related to *Gloeobacter*. Its complete circular genome, ultrastructure, and basic physiological properties provided new insights into the early evolution of Cyanobacteria and associated oxygenic photosynthesis.

RESULTS

Isolation of *Anthocerotibacter panamensis* from a tropical hornwort

A new cyanobacterial culture was obtained from a surface-sterilized thallus of the hornwort (Bryophyta) *Leiosporoceros dussii* from Panama (Figure 1). Our 16S rDNA phylogeny places this isolate (named *Anthocerotibacter panamensis*) sister to a clade of uncultured environmental samples from the Arctic and Antarctic regions (e.g., Canadian tundra, Iceland, Antarctic lakes, and Patagonia; Figure 2A). This “polar clade” also includes *Candidatus Aurora* vandensis MAGs from Antarctica. The 16S rDNA sequence identity between *Anthocerotibacter* is around 96%,

88%, and 70% when compared to polar clade members, *Gloeobacter*, and *Synechocystis* sp. PCC 6803, respectively (Figure 2D). The equatorial origin of *Anthocerotibacter* is in stark contrast to its sister polar clade, which speaks to the uniqueness of this new isolate.

Because all hornwort species have symbiotic associations with filamentous nitrogen-fixing Cyanobacteria,^{25,26} it raises the question of whether *Anthocerotibacter* could be a major symbiont. We suggest this is unlikely to be the case as *Anthocerotibacter* is not filamentous (Figure 1B), and its genome lacks nitrogenase genes although nitrogen fixation is the selective advantage in plant-Cyanobacteria symbiosis. Recently, the microbiome of *L. dussii* was profiled using 16S rDNA amplicon sequencing,²⁷ but no trace of *Anthocerotibacter* could be found (though interestingly, a unique clade of *Gloeobacter* was uncovered therein: ASV172; ASV719; and ASV661 in Figure 2A). We therefore speculate that *Anthocerotibacter* could be a soil bacterium that happened to survive our surface sterilization, an endophyte that inhabits hornwort intercellular space, or a freeloader inside symbiotic cavities with other functional nitrogen-fixing cyanobionts. Future studies are needed to clarify the ecology and natural history of *Anthocerotibacter*.

Anthocerotibacter represents a novel and deeply branched cyanobacterial lineage

To corroborate the phylogenetic placement of *Anthocerotibacter* and to estimate the timing of divergence, we compiled a 12-gene matrix based on the dataset of Shih et al.⁶ These genes are slow evolving and are shared between plastid and mitochondria genomes, thereby allowing for cross-validations to reuse and link the same calibration across multiple nodes. As in the 16S phylogeny, we recovered the same branching order among *Anthocerotibacter*, *Candidatus Aurora*, *Gloeobacter*, and

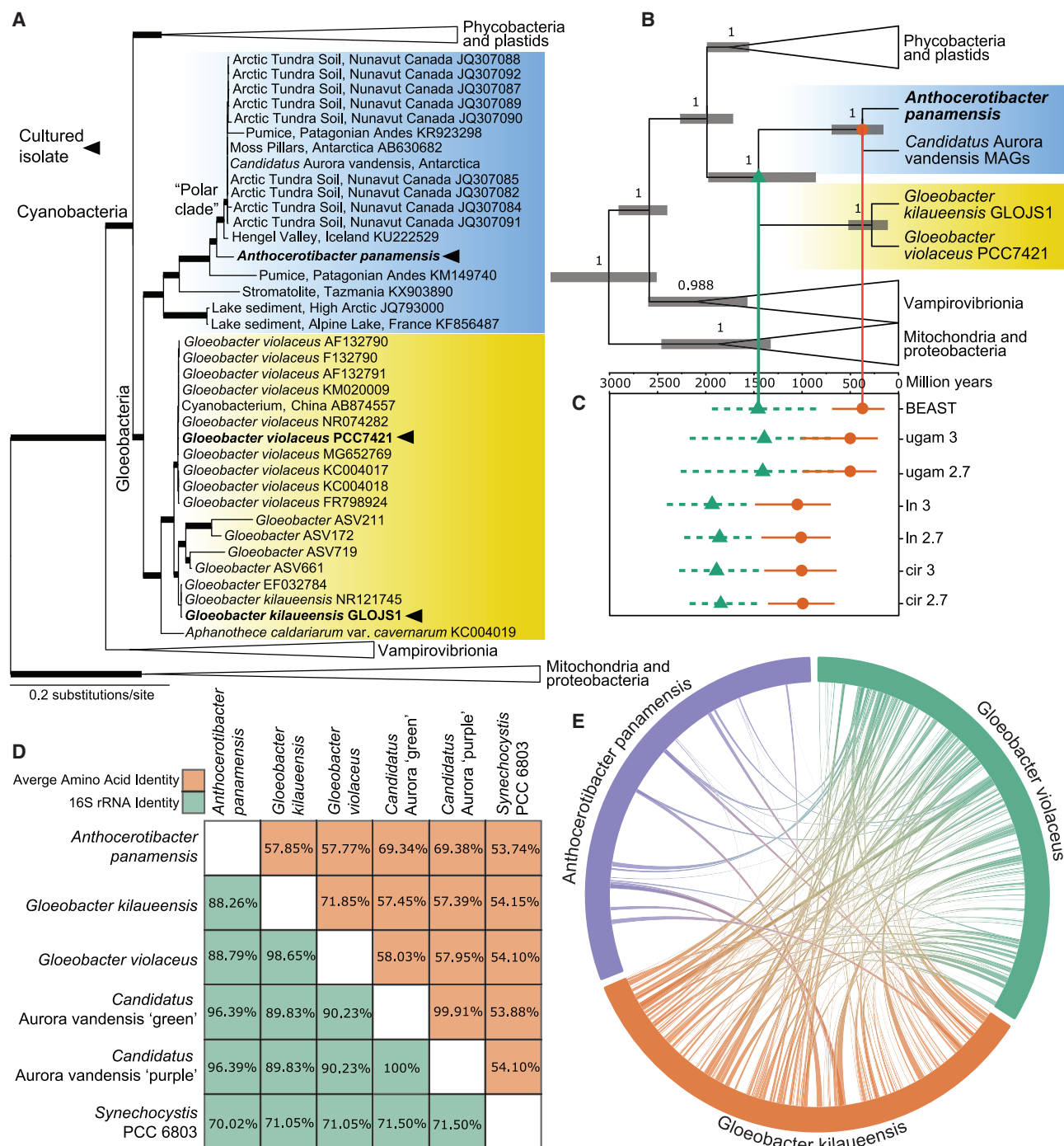


Figure 2. Relationship of *Anthocerotibacter panamensis* and other Cyanobacteria

(A) *Anthocerotibacter* is sister to a clade of environmental samples from the polar regions, based on 16S rDNA sequences. Thickened branches received bootstrap support >80. Arrowhead marks the strains that are available as cultures.

(B) A time-calibrated phylogeny revealed that *Anthocerotibacter* represents a deeply diverged lineage within Cyanobacteria. Values above branches are posterior probabilities. The size of collapsed wedges does not correspond to taxa number.

(C) Comparison of divergence time estimates from (B) ("BEAST") and from PhyloBayes following the calibration strategies of Baracaldo et al.²⁴ A total of six PhyloBayes runs were carried out, using a combination of different clock models: log normal (ln); Cox-Ingersoll-Ross model (cir); and uncorrelated gamma

(legend continued on next page)

Phycobacteria with strong support (Figure 2B). It should be noted that our topology is different from the 16S tree of Grettenberger et al.,¹² which resolved *Candidatus Aurora* sister to *Gloeobacter* + Phycobacteria, but is consistent with their 37-gene phylogeny. Given the high branch support from both of our datasets, we believe our result (((*Anthocerotibacter*, *Candidatus Aurora*), *Gloeobacter*), Phycobacteria) represents the most plausible relationship.

The chronogram from BEAST suggested that *Anthocerotibacter* diverged from *Candidatus Aurora* at around 374.74 mya (95% highest posterior density [HPD]: 155.86–694.79 mya) and from *Gloeobacter* around 1.45 bya (95% HPD: 857.64–1,975.35 mya; Figure 2B). Because molecular dating analyses can be strongly influenced by calibrations, we further explored the robustness of our time estimates by adopting a suite of different calibration schemes, models, and data matrices made by Baracaldo et al.²⁴ Comparable but generally deeper divergences were obtained, and the results are summarized in Figure 2C. Taken together, our phylogenetic and molecular clock analyses indicate that *Anthocerotibacter* is currently the sole cultured representative of a novel cyanobacterial lineage, which split from *Gloeobacter* over 1.4 bya.

Anthocerotibacter genome

Nanopore sequencing resulted in one circular genome assembly of 4,187,797 bp and a GC content of 55.51%. The average amino-acid identity (AAI) is around 57% when compared to *Gloeobacter* and 69% to *Candidatus Aurora* (Figure 2D). A total of 4,166 protein-coding genes were predicted, 1,805 of which can be functionally annotated. Protein family circumscription based on 103 cyanobacterial species identified a total of 18,419 orthogroups. Of these, 2,818 gene families are present in *Anthocerotibacter*, 35 are species specific, and 140 are unique to *Anthocerotibacter* + *Candidatus Aurora* (Figure S1; Data S1). A total of 454 predicted genes in *Anthocerotibacter* were singletons that could not be assigned to any orthogroup. Furthermore, we found very little synteny between *Anthocerotibacter* and *Gloeobacter* spp., consistent with the deep divergence between the two genera (Figure 2E).

Anthocerotibacter lacks thylakoids

One of the most distinctive features of *Gloeobacter* is its lack of thylakoids.^{8,9} Although the ultrastructure of *Candidatus Aurora* is not available, analysis of its metagenome found a lack of the *VIPP1* (vesicle-inducing protein in plastid 1) gene,¹² which is essential for biogenesis of thylakoid membrane.²⁸ The availability of *Anthocerotibacter* as a culture allowed us to directly examine it using transmission electron microscopy (TEM). We found no trace of thylakoids (Figures 1C–1E) or a *VIPP1* homolog in the genome. This result reaffirmed that the lack of thylakoid is likely a general feature of *Gloeobacteria* and thylakoids were an innovation of Phycobacteria.

Spectroscopy reveals unique features of Anthocerotibacter

In a stark contrast to the purple to brown color of *Gloeobacter*,^{7,9} *Anthocerotibacter* is blue-green and has no apparent red hue (Figure 1A). To better understand the pigment composition, we characterized and compared *Anthocerotibacter*'s absorption and fluorescence emission spectra with those of *Gloeobacter violaceus* and *Synechocystis* sp. PCC 6803 (a model unicellular species from Phycobacteria; hereafter Syn6803). In the whole-cell absorption spectra (Figure 3A), *Anthocerotibacter* lacks clear peaks representing phycoerythrin (PE), which is different from *G. violaceus*²⁹ but consistent with the absence of PE genes in the *Anthocerotibacter* genome (see below). Compared to Syn6803, which shows predominant Soret and Q_y bands of chlorophyll (Chl) *a* (440 nm and 682 nm), the absorption at 682 nm in *Anthocerotibacter* is lower, not showing a pronounced peak (Figure 3A). The peak at 635 nm from phycocyanin (PC) is higher in *Anthocerotibacter* than in Syn6803 at 629 nm, though it is common that the absorption peaks of the same phycobiliproteins differ from species to species.^{21,30,31} The absorption feature at 495 nm is from carotenoids. Because *Anthocerotibacter* has no PE or phycoerythrocyanin (PEC) identified in the genome, the 582 nm feature, which is absent in Syn6803, is from an unknown source, possibly from carotenoid-binding proteins. In *G. violaceus*, this wavelength region is overlaid with the absorption of PE in the whole cell absorption spectra and was therefore not discussed previously.²⁹

The absorption spectra from methanol-extracted pigments (Figure 3B) show that *Anthocerotibacter* contain only Chl *a* (665 nm and 436 nm) and no other chlorophylls, and the Q_y band (665 nm) matches perfectly with the Q_y band from Chl *a* in Syn6803. However, although both strains show features of carotenoids at 475 nm, *Anthocerotibacter* has stronger absorption at 436 nm and an additional absorption feature around 530 nm, indicating it has one or more carotenoids with long-wavelength absorption. A long-wavelength absorbing carotenoid was also identified in *G. violaceus*, and its absorption maximum is similar to oscilloxanthin.^{14,29} At high-light condition (100 μmol photons m⁻²s⁻¹), *Anthocerotibacter* produced more carotenoids (absorption peaks at wavelengths 443 nm, 474 nm, and 530 nm) per unit of Chl *a* (Figures 3C and 5B). Further analyses are required to characterize the species and quantities of carotenoids in *Anthocerotibacter*.

Finally, 77K fluorescence spectroscopy was used to measure fluorescence emission from chromophore-bound protein complexes, such as phycobilisomes and photosystem I (PSI) and II (PSII). In *Anthocerotibacter*, phycobiliproteins are the predominant light-harvesting complexes, resulting fluorescence emission largely coming from phycobiliproteins, even though the excitation wavelength (at 440 nm) is targeted for chlorophylls. Three emission peaks are observed in *Anthocerotibacter* and assigned as PC (646 nm), allophycocyanin (AP) (664 nm), and PSII (684 nm; Figure 3D). The peak at 684 nm is asymmetric,

multiplier (ugam) and two different root ages: 2.7 and 3 bya. The orange and green lines are the 95% HPD of age estimates for *Anthocerotibacter*-*Candidatus Aurora* and *Anthocerotibacter*-*Gloeobacter* splits, respectively.

(D) *Anthocerotibacter* has low pairwise 16S rDNA identities and low AAI with other cyanobacterial species. *Candidatus Aurora* vandensis “green” and “purple” are MAGs from separate sediment layers.

(E) Very little genomic synteny exists between *Anthocerotibacter* and *Gloeobacter* spp.

See also Figure S1.

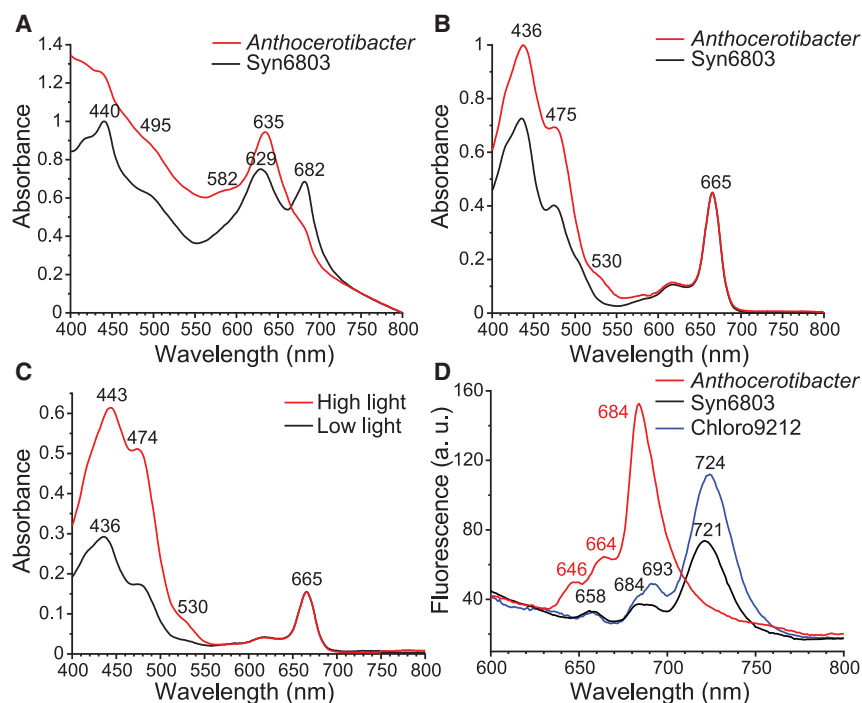


Figure 3. Spectral analyses of *Anthocerotibacter* in comparison with other Cyanobacteria

(A–C) Absorption spectra of (A) whole cells, (B) pigment extraction in methanol, and (C) pigment extraction in methanol of *Anthocerotibacter* in high and low light (100 and 10 $\mu\text{mol photons m}^{-2}\text{s}^{-1}$). (D) Fluorescence spectroscopy was taken at 77K with excitation wavelength at 440 nm using whole cells of *Anthocerotibacter*, Syn6803, and *Chlorogloeopsis fritschii* PCC 9212 (Chloro9212). The absorbance spectra were normalized at 750 nm in (A) and at 665 nm in (B) and (C). Specific wavelengths were labeled above the peaks.

indicating that the second peak of PSII is likely present, but not strong enough, forming a separate peak, a pattern also found in *G. violaceus*.^{15,29} The long-wavelength emission from PSI is absent in *Anthocerotibacter*, which is different from Syn6803 (and other Phycobacteria) but similar to *G. violaceus* (Figure 3D).^{15,29} Previous studies in *G. violaceus* concluded that the absence of long-wavelength emission beyond 700 nm is due to the lack of long-wavelength Chl *a* in PSI,²⁹ which is possibly the situation in *Anthocerotibacter* as well. These measurements indicate that the spectral properties of *Anthocerotibacter* are distinct from Syn6803 but largely similar to *G. violaceus*, with the exception that *Anthocerotibacter* lacks PE.

Anthocerotibacter lacks cyanobacterial circadian clock genes

Circadian clocks enable organisms to adjust their physiological activities according to extrinsic daily changes, such as light, temperature, and humidity. In Cyanobacteria, environmental signals are transmitted through redox sensitive proteins to core oscillators (KaiA, KaiB, and KaiC), followed by output pathways to regulate physiological activities.³² We found that the *Anthocerotibacter* genome is missing the core oscillator *KaiABC* genes, similar to what was reported in *Gloeobacter*.¹³ This implies that the cyanobacterial circadian clock most likely evolved in the common ancestor of Phycobacteria.

The carotenoid biosynthesis pathways of *Anthocerotibacter* and *Gloeobacter* differ

Carotenoids are an integral part of light-harvesting complexes and play key roles in photoprotection in conjunction with orange carotenoid proteins (OCP). Although *Anthocerotibacter* OCP is in the OCPX clade (defined by Bao et al.³³) like *Gloeobacter* (Figure S2; Data S2), the pathway through which carotenoids are

made differs between the two lineages. The first step of carotenoid biosynthesis is the condensation of two geranylgeranyl pyrophosphates into phytoene-by-phytoene synthase (CrtB). Following this, phytoene is converted to lycopene through two different routes, depending on the organisms. The “plant-type” pathway, which can be found in all Cyanobacteria (except *Gloeobacter*) and plants, relies on three separate enzymes: phytoene desaturase (CrtP); ζ -carotene desaturase (CrtQ); and *cis*-carotene isomerase (CrtH).³⁴ By contrast, the “bacteria type,” found in *Gloeobacter*, anoxygenic photosynthetic bacteria, and fungi, uses only a single phytoene desaturase (CrtI).^{22,23} Interestingly, we found that both *Anthocerotibacter* and *Candidatus Aurora* have the genetic chassis for the plant-type carotenoid biosynthesis, instead of the bacteria-type reported in *Gloeobacter* (Data S2). This implies that the origin of the plant-type carotenoid biosynthesis pathway should be pushed back to the ancestor of all Cyanobacteria, with *Gloeobacter* later substituting it with the bacterial pathway. Although we cannot entirely rule out that *Anthocerotibacter* horizontally acquired the plant-type pathway from Phycobacteria, the phylogenetic positions of *Anthocerotibacter* *crtP*, *crtQ*, and *crtH* are consistent with the species phylogeny (Figure S3).

Anthocerotibacter has a unique phycobilisome composition

Cyanobacteria use phycobilisomes as the light-harvesting antennae to capture and transfer light energy to photosystems. Typically, phycobilisomes are fan shaped and composed of the AP core and several peripheral proteins: PC and PE or PEC.^{21,30} One notable exception is *Gloeobacter*, whose phycobilisomes are bundle shaped and consist of AP, PC, and PE.^{20,21} Compared to *Gloeobacter*, *Anthocerotibacter* further lacks homologs for PE (and PEC), a situation also reported from *Candidatus Aurora* MAG (Figure 4).¹² On the other hand, most of the AP and PC subunits are present in *Anthocerotibacter*, which is consistent with our absorption and fluorescence emission spectral data (Figures 3A and 3D). The missing AP and PC subunits are ApcD, ApcF, and CpcD, although the mutations of which did not appear to affect the phycobilisome function in *Synechococcus* sp. PCC 7002^{35,36} or in Syn6803.^{37,38}

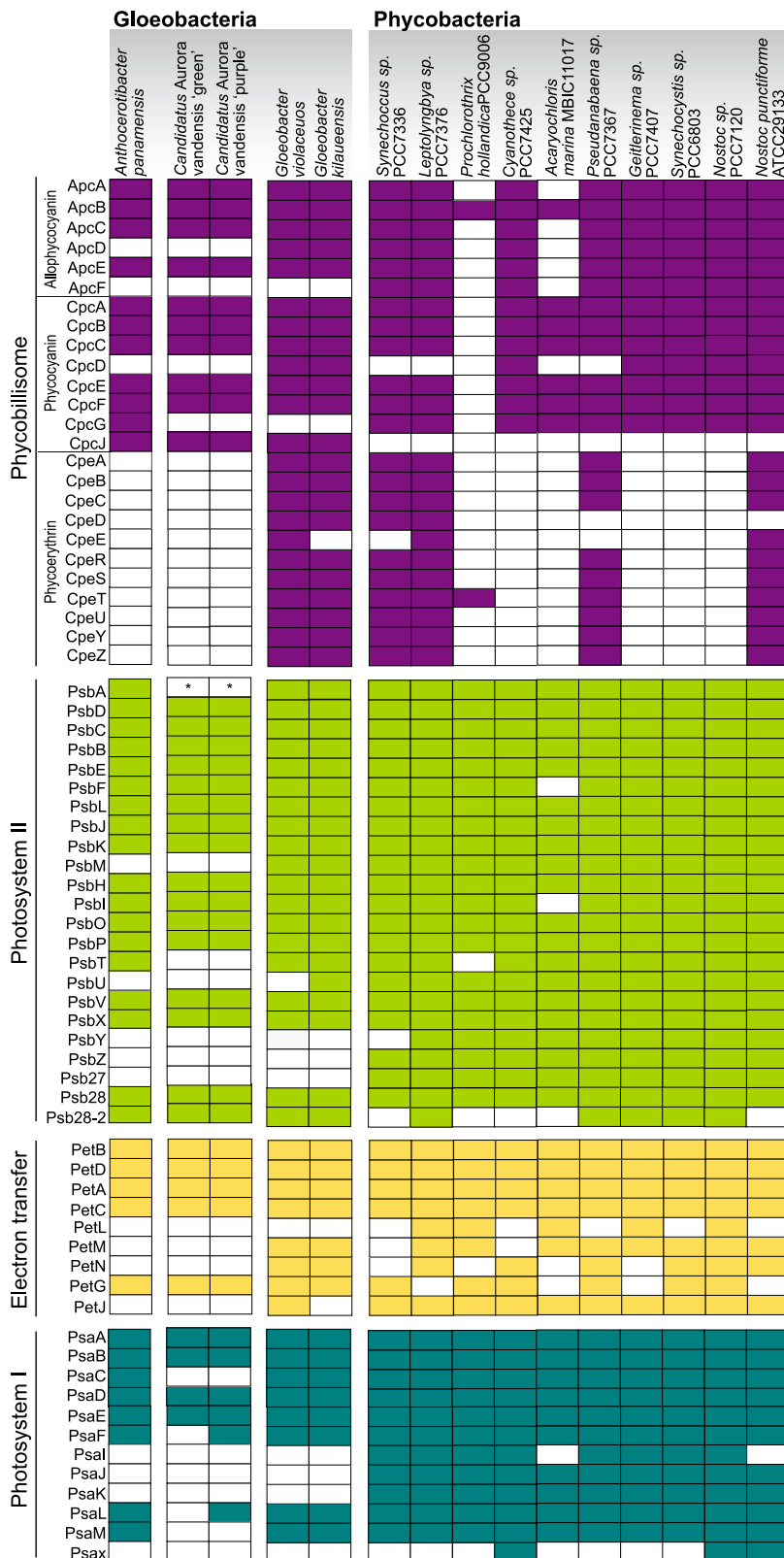


Figure 4. Comparison of phycobilisome and photosystem subunits among *Anthocerotibacter*, *Candidatus Aurora*, *Gloeobacter*, and other Cyanobacteria

Filled and empty boxes indicate gene presence and absence, respectively. *Candidatus Aurora* "green" and "purple" are MAG from separate sediment layers. **psbA* was not found in Aurora MAG but on separate contigs that could not be properly binned. See also [Data S1](#) and [S2](#) and [Figures S4](#) and [S5](#).

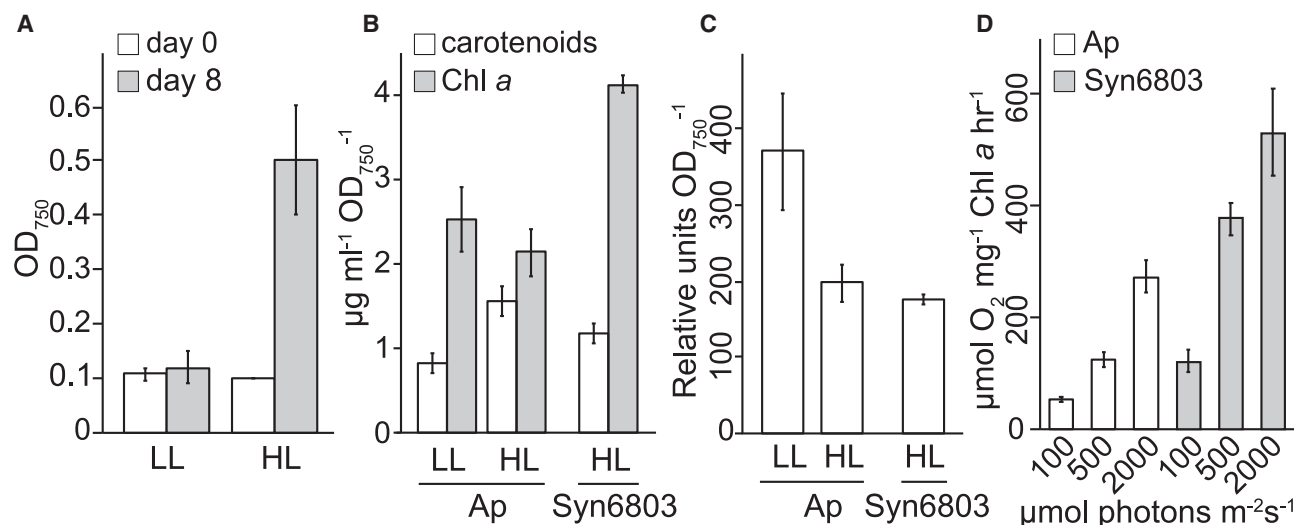


Figure 5. Measurements of growth, pigment analyses, and oxygen evolution

Anthocerotibacter was grown in high light (HL) (100 µmol photon m⁻² s⁻¹) and low light (LL) (10 µmol photon m⁻² s⁻¹) conditions for 8 days. Three biological replicates were used for all the measurements. Standard deviations were shown as error bars.

(A) The optical density 750 (OD₇₅₀) values at the beginning of inoculation (day 0) and the end of cultivation (day 8) were recorded for estimation of doubling time. (B and C) Quantifications of (B) Chl a and carotenoids and (C) phycobiliproteins were performed using *Anthocerotibacter* (Ap) cells grown in LL and HL on day 8. *Synechocystis* sp. PCC6803 (Syn6803) grown under HL was used as a comparison.

(D) Rates of oxygen evolution (y axis) over different light intensities (µmol photons m⁻² s⁻¹).

Another peculiar feature of *Anthocerotibacter* phycobilisomes is the presence of *cpcG*, which encodes for a rod-core linker protein. This protein is critical in phycobilisome assembly and energy transfer from PC to AP but is absent in *Gloeobacter* and *Candidatus Aurora*. In *Gloeobacter*, a different linker protein, CpcJ, was used instead of CpcG.³⁹ Interestingly, this *Gloeobacter*-specific *cpcJ* gene is also present in *Anthocerotibacter*, suggesting that *Anthocerotibacter* has two different types of rod-core linker (CpcG and CpcJ) for connecting PC and AP (Data S2). In other words, *Anthocerotibacter* may have a considerably different phycobilisome from that of *Gloeobacter*, which is already distinct among Cyanobacteria. Furthermore, our phylogenetic analysis of phycobilisome linker proteins placed *Anthocerotibacter cpcG* sister to the rest of cyanobacterial homologs (Figure S4), suggesting that the ancestor of Cyanobacteria very likely had a CpcG linker.

Anthocerotibacter has a reduced set of photosystem subunits

With respect to photosystems, *Anthocerotibacter* is similar to *Gloeobacter* in that they both lack several highly conserved genes encoding PsbY, PsbZ, and Psb27 at PSII and PsaI, PsaJ, PsaK, and PsaX at PSI (Figure 4). These subunits are involved in electron transfer (e.g., PsbZ)⁴⁰ or the formation and stabilization of photosystems (e.g., Psb27, PsaI, and PsaJ).^{41–45} *Anthocerotibacter* also lacks the known pathway genes for synthesizing phyloquinone/menaquinone (KEGG M00116), which is a secondary electron acceptor of PSI; such absence is also shared with *Gloeobacter*.¹⁸

Interestingly, we found even more genes encoding photosystem subunits are missing in *Anthocerotibacter* (and to some extent *Candidatus Aurora* as well) compared to *Gloeobacter*, including PsbM, PsbU, PetM, and PetN. PsbM is localized in

the PSII center and plays a key role in PSII dimer formation and stability.⁴⁶ PsbU is needed for making a thermally stable oxygen-evolving complex and aids electron transport and energy transfer from phycobilisome to PSII.^{47,48} At the electron transport chain, PetN is necessary for the proper function and assembly of the cytochrome b₆f complex at least in Syn6803 and tobacco.⁴⁹ Although PetM mutations in Syn6803 did not impair the cytochrome b₆f function, the amounts of phycobilisomes and PSI decreased.⁵⁰ In all, *Anthocerotibacter* has one of the most reduced sets of photosynthesis subunits among Cyanobacteria, which prompted us to further investigate the physiological properties of *Anthocerotibacter* and compare them with existing cyanobacterial strains.

Anthocerotibacter is able to grow photoautotrophically and evolve oxygen but slowly

The growth of *Anthocerotibacter* was measured under low-light (LL) and high-light (HL) conditions (10 and 100 µmol photons m⁻² s⁻¹, respectively) in the medium without supplement of carbon sources. *Anthocerotibacter* can grow photoautotrophically under both light intensities, with a faster rate under HL (Figure 5A). The pigment analyses show that, under HL, the Chl a:phycobiliprotein ratio is about half in *Anthocerotibacter* as in Syn6803 (Figures 5B and 5C). *Anthocerotibacter* produces less Chl a and phycobiliproteins but more carotenoids in HL than in LL, probably because photoprotection is turned on. A similar change from lower to higher light intensity was also observed in *Gloeobacter*.¹⁵

Compared to other model Cyanobacteria, *Anthocerotibacter* is very slow growing; its doubling time is estimated around 67 h under HL and over 8 days under LL (Figure 5A). By comparison, Syn6803 doubles in around 8 h under the same HL condition. The reported doubling time for *Gloeobacter* ranges from

73 h to 17 days,^{14,51} depending on the culturing conditions. The slow growth of *Anthocerotibacter* likely results from its inefficient photosynthetic apparatus and/or the inability to tolerate high light intensities, which could be caused by the fact that they lack accessory subunits of PSI and PSII (Figure 4).

Finally, we demonstrated that *Anthocerotibacter* does produce oxygen, and the level of which increased with light intensity (from 100, 500, to 2,000 $\mu\text{mol photons m}^{-2}\text{s}^{-1}$; Figure 5D). In general, *Anthocerotibacter* has about half of the oxygen evolution rates compared to Syn6803 under the same conditions. Although *Gloeobacter* was not included in our experiments, we could cross-reference the results with those of Koyama et al.,⁵² who reported a similar difference in oxygen evolution rates between Syn6803 and *G. violaceus*. In other words, despite having even fewer photosynthesis subunits than *Gloeobacter*, *Anthocerotibacter* can nevertheless carry out comparable oxygenic photosynthesis. These results in *Anthocerotibacter* may contribute to determination of the minimal and essential photosynthetic subunits required for photoautotrophic growth in nature.

DISCUSSION

Anthocerotibacter described here represents a novel and ancient lineage of Cyanobacteria that diverged from *Gloeobacter* over 1.4 bya. Very recently, *Candidatus* Aurora was described from a metagenome assembly,¹¹ which, based on our phylogenetic analysis, belongs to the same lineage as *Anthocerotibacter* (Figure 2). There are, however, several key differences between the two. First, the assembly of *Candidatus* Aurora MAG is fragmented and incomplete, making it difficult to confidently conclude gene presence and absence and discuss the broader implications on cyanobacterial evolution. For example, *psbA* (which encodes PSII reaction center D1 protein) was not present in Aurora MAG but instead found on a separate small contig that was not binned properly. Here, with a complete circular genome of *Anthocerotibacter*, we are able to fully describe the gene content. A case in point is the phycobilisome linker protein CpcG. Because it is missing in *Gloeobacter* genomes as well as Aurora MAG, *cpcG* gene would be inferred to originate in Phycobacteria. However, we found a clear *cpcG* homolog in *Anthocerotibacter* that is phylogenetically sister to the rest of cyanobacterial *cpcG* (Figure S4). This suggests that *cpcG* is probably ancestral to all Cyanobacteria (but subsequently lost in *Gloeobacter*) and raises the possibility that its absence in Aurora MAG might be due to incomplete assembly. Second, the existence of *Candidatus* Aurora had been hinted by earlier environmental samples from Arctic and Antarctic regions, which share highly similar 16S rDNA sequences (>99% identity; Figure 2).¹¹ *Anthocerotibacter*, on the hand, was not known from any environmental sample that we are aware of. In addition, as an isolate from a tropical hornwort, *Anthocerotibacter* shows a contrasting distribution pattern from its sister polar clade (Figure 2). Most importantly, the fact that *Anthocerotibacter* is the only cultured organism in this lineage allowed us to characterize the morphology and physiology of this representing strain in detail. The lack of thylakoid in *Anthocerotibacter* was observed by TEM (Figure 1), which is much more convincing than basing on the absence of *VIPP1* homolog alone in *Candidatus*

Aurora,¹¹ especially considering gene absence could also be due to incomplete MAG assembly. Our spectral and growth experiments also provided important insights into the biology of *Anthocerotibacter* that are otherwise impossible to obtain for a metagenomic sample.

The deep divergence and availability as a culture enable us to use *Anthocerotibacter* to revisit the early evolution of Cyanobacteria. One key question we were able to address is whether the many unique features in *Gloeobacter* are indeed characteristic of the entire *Gloeobacteria* lineage. On one hand, we found that, similar to *Gloeobacter*, *Anthocerotibacter* lacks thylakoids, circadian clock oscillators, nitrogenase, canonical pathway for phyloquinone synthesis, long-wavelength PSI fluorescence, as well as many conserved photosynthesis subunits (summarized in Figure S5). The most parsimonious interpretation is that the origins of these components took place in the most recent common ancestor (MRCA) of Phycobacteria and the absence of which likely defines *Gloeobacteria*. Conversely, the carotenoid biosynthesis pathway in *Anthocerotibacter* is not the bacterial type as found in *Gloeobacter* but typical of the rest of Cyanobacteria, implying that the plant-type pathway likely evolved in the MRCA of Cyanobacteria and not of Phycobacteria (Figure S5). *Anthocerotibacter* also possesses several unique features, such as its phycobilisome composition (with both CpcJ and CpcG linker proteins) and an even further reduced collection of photosystem subunits. Nevertheless, we were able to demonstrate that *Anthocerotibacter* can indeed grow photoautotrophically and evolve oxygen across a range of light intensities, although less efficiently compared to other cyanobacterial strains. In summary, *Anthocerotibacter* is an important addition to the depauperate *Gloeobacteria* and will facilitate future investigations into the origin of Cyanobacteria and oxygenic photosynthesis.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Cyanobacteria cultures
 - Taxonomic treatments
- METHOD DETAILS
 - DNA extraction
 - Nanopore and Illumina sequencing
 - Genome assembly and annotation
 - Genome comparisons
 - Phylogenetic analyses and divergence time estimates
 - Phylogeny of orange carotenoid proteins (OCP)
 - Phylogeny of phycobilisome linker proteins
 - Microscopy
 - Absorption and fluorescence spectral analyses
 - Measurements of pigment composition and oxygen evolution
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cub.2021.04.042>.

ACKNOWLEDGMENTS

We thank Dr. Noris Salazar Allen for providing fieldwork support to J.C.V.A., Dr. Hsiu-An Chu and Keng-Min Lin for assisting oxygen evolution measurement, Dr. Patrick Shih for sharing alignment matrices and early discussions, Dr. Aharon Oren for consulting on nomenclature, and Dr. Kathleen Pryer, Dr. John Meeks, and two anonymous reviewers for comments. This work was funded by National Science Foundation (NSF) grant no. DEB1831428 to F.-W.L., Ministry of Science and Technology (Taiwan) Young Scholar Fellowship Einstein Program (108C5368 and 109C5369) and Ministry of Education Yushan Young Scholar Program (108V1102 and 109V1102) to M.-Y.H., and Chaires de recherche du Canada (950-232698), CRNSG (RGPIN-2016-05967), and Earl S. Tupper Fellowship (STRI) to J.C.V.A.

AUTHOR CONTRIBUTIONS

Conceptualization, F.-W.L.; formal analysis, N.R., M.-Y.H., and F.-W.L.; investigation, N.R., D.A.H., J.M.N., P.Y.C., J.C.V.A., M.-Y.H., and F.-W.L.; resources, J.C.V.A.; data curation, N.R.; writing, N.R., M.-Y.H., and F.-W.L.; visualization, N.R., M.-Y.H., and F.-W.L.; funding acquisition, M.-Y.H. and F.-W.L.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 25, 2021

Revised: April 2, 2021

Accepted: April 16, 2021

Published: May 13, 2021

REFERENCES

1. Soo, R.M., Hemp, J., Parks, D.H., Fischer, W.W., and Hugenholtz, P. (2017). On the origins of oxygenic photosynthesis and aerobic respiration in Cyanobacteria. *Science* 355, 1436–1440.
2. Planavsky, N.J., Reinhard, C.T., Wang, X., Thomson, D., McGoldrick, P., Rainbird, R.H., Johnson, T., Fischer, W.W., and Lyons, T.W. (2014). Earth history. Low mid-Proterozoic atmospheric oxygen levels and the delayed rise of animals. *Science* 346, 635–638.
3. Soo, R.M., Hemp, J., and Hugenholtz, P. (2019). Evolution of photosynthesis and aerobic respiration in the cyanobacteria. *Free Radic. Biol. Med.* 140, 200–205.
4. Garcia-Pichel, F., Zehr, J.P., Bhattacharya, D., and Pakrasi, H.B. (2020). What's in a name? The case of cyanobacteria. *J. Phycol.* 56, 1–5.
5. Cavalier-Smith, T. (2002). The neomuran origin of archaeobacteria, the neigibacterial root of the universal tree and bacterial megaclassification. *Int. J. Syst. Evol. Microbiol.* 52, 7–76.
6. Shih, P.M., Hemp, J., Ward, L.M., Matzke, N.J., and Fischer, W.W. (2017). Crown group Oxyphotobacteria postdate the rise of oxygen. *Geobiology* 15, 19–29.
7. Saw, J.H.W., Schatz, M., Brown, M.V., Kunkel, D.D., Foster, J.S., Shick, H., Christensen, S., Hou, S., Wan, X., and Donachie, S.P. (2013). Cultivation and complete genome sequencing of *Gloeobacter kilaeensis* sp. nov., from a lava cave in Kilauea Caldera, Hawai'i. *PLoS ONE* 8, e76376.
8. Rippka, R., Waterbury, J., and Cohen-Bazire, G. (1974). A cyanobacterium which lacks thylakoids. *Arch. Microbiol.* 100, 419–436.
9. Mareš, J., Hrouzek, P., Kaňa, R., Ventura, S., Strunecký, O., and Komárek, J. (2013). The primitive thylakoid-less cyanobacterium *Gloeobacter* is a common rock-dwelling organism. *PLoS ONE* 8, e66323.
10. Montejano, G., Becerra-Absalón, I., Gold-Morgan, M., and Osorio-Santos, K. (2018). *Gloeobacter violaceus*: primitive reproductive scheme and its significance. *Plant Syst. Evol.* 304, 1221–1229.
11. Lynch, M.D.J., Bartram, A.K., and Neufeld, J.D. (2012). Targeted recovery of novel phylogenetic diversity from next-generation sequence data. *ISME J.* 6, 2067–2077.
12. Grottenberger, C.L., Sumner, D.Y., Wall, K., Brown, C.T., Eisen, J.A., Mackey, T.J., Hawes, I., Jospin, G., and Jungblut, A.D. (2020). A phylogenetically novel cyanobacterium most closely related to *Gloeobacter*. *ISME J.* 14, 2142–2152.
13. Nakamura, Y., Kaneko, T., Sato, S., Mimuro, M., Miyashita, H., Tsuchiya, T., Sasamoto, S., Watanabe, A., Kawashima, K., Kishida, Y., et al. (2003). Complete genome structure of *Gloeobacter violaceus* PCC 7421, a cyanobacterium that lacks thylakoids. *DNA Res.* 10, 137–145.
14. Rexroth, S., Mullineaux, C.W., Ellinger, D., Sendtko, E., Rögner, M., and Koenig, F. (2011). The plasma membrane of the cyanobacterium *Gloeobacter violaceus* contains segregated bioenergetic domains. *Plant Cell* 23, 2379–2390.
15. Koenig, F., and Schmidt, M. (1995). *Gloeobacter violaceus*—investigation of an unusual photosynthetic apparatus. Absence of the long wavelength emission of photosystem I in 77 K fluorescence spectra. *Physiol. Plant.* 94, 621–628.
16. Mangels, D., Krup, J., Berry, S., Rögner, M., Boekema, E.J., and Koenig, F. (2002). Photosystem I from the unusual cyanobacterium *Gloeobacter violaceus*. *Photosynth. Res.* 72, 307–319.
17. Inoue, H., Tsuchiya, T., Satoh, S., Miyashita, H., Kaneko, T., Tabata, S., Tanaka, A., and Mimuro, M. (2004). Unique constitution of photosystem I with a novel subunit in the cyanobacterium *Gloeobacter violaceus* PCC 7421. *FEBS Lett.* 578, 275–279.
18. Mimuro, M., Tsuchiya, T., Inoue, H., Sakuragi, Y., Itoh, Y., Gotoh, T., Miyashita, H., Bryant, D.A., and Kobayashi, M. (2005). The secondary electron acceptor of photosystem I in *Gloeobacter violaceus* PCC 7421 is menaquinone-4 that is synthesized by a unique but unknown pathway. *FEBS Lett.* 579, 3493–3496.
19. Dreher, C., Hielscher, R., Prodöhl, A., Hellwig, P., and Schneider, D. (2010). Characterization of two cytochrome *b₆* proteins from the cyanobacterium *Gloeobacter violaceus* PCC 7421. *J. Bioenerg. Biomembr.* 42, 517–526.
20. Guglielmi, G., Cohen-Bazire, G., and Bryant, D.A. (1981). The structure of *Gloeobacter violaceus* and its phycobilisomes. *Arch. Microbiol.* 129, 181–189.
21. Bryant, D.A., and Canniffe, D.P. (2018). How nature designs light-harvesting antenna systems: design principles and functional realization in chlorophototrophic prokaryotes. *J. Phys. B* 51, 033001.
22. Tsuchiya, T., Takaichi, S., Misawa, N., Maoka, T., Miyashita, H., and Mimuro, M. (2005). The cyanobacterium *Gloeobacter violaceus* PCC 7421 uses bacterial-type phytoene desaturase in carotenoid biosynthesis. *FEBS Lett.* 579, 2125–2129.
23. Steiger, S., Jackisch, Y., and Sandmann, G. (2005). Carotenoid biosynthesis in *Gloeobacter violaceus* PCC7421 involves a single crtI-type phytoene desaturase instead of typical cyanobacterial enzymes. *Arch. Microbiol.* 184, 207–214.
24. Sánchez-Baracaldo, P., Raven, J.A., Pisani, D., and Knoll, A.H. (2017). Early photosynthetic eukaryotes inhabited low-salinity habitats. *Proc. Natl. Acad. Sci. USA* 114, E7737–E7745.
25. Meeks, J.C. (1998). Symbiosis between nitrogen-fixing cyanobacteria and plants. *BioScience* 48, 266–276.
26. Duckett, J.G., Prasad, A.K.S.K., Davis, D.A., and Walker, S. (1977). A cytological analysis of the *Nostoc*-Bryophyte relationship. *New Phytol.* 79, 349–362.
27. Bouchard, R., Peñaloza-Bojacá, G., Toupin, S., Guadalupe, Y., Gudiño, J., Allen, N.S., Li, F.W., and Villarreal, A.J.C. (2020). Contrasting bacteriome of the hornwort *Leiosporoceros dussii* in two nearby sites with emphasis on the hornwort-cyanobacterial symbiosis. *Symbiosis* 81, 39–52.

28. Kroll, D., Meierhoff, K., Bechtold, N., Kinoshita, M., Westphal, S., Voithknecht, U.C., Soll, J., and Westhoff, P. (2001). *VIPP1*, a nuclear gene of *Arabidopsis thaliana* essential for thylakoid membrane formation. *Proc. Natl. Acad. Sci. USA* 98, 4238–4242.
29. Mimuro, M., Ookubo, T., Takahashi, D., Sakawa, T., Akimoto, S., Yamazaki, I., and Miyashita, H. (2002). Unique fluorescence properties of a cyanobacterium *Gloeobacter violaceus* PCC 7421: reasons for absence of the long-wavelength PSI Chl a fluorescence at -196°C . *Plant Cell Physiol.* 43, 587–594.
30. Bryant, D.A., Guglielmi, G., Tandeau de Marsac, N., Castets, A.-M., and Cohen-Bazire, G. (1979). The structure of cyanobacterial phycobilisomes: a model. *Arch. Microbiol.* 123, 113–127.
31. Bryant, D.A. (1982). Phycoerythrocyanin and phycoerythrin: properties and occurrence in cyanobacteria. *J. Gen. Microbiol.* 128, 835–844.
32. Cohen, S.E., and Golden, S.S. (2015). Circadian rhythms in Cyanobacteria. *Microbiol. Mol. Biol. Rev.* 79, 373–385.
33. Bao, H., Melnicki, M.R., Pawlowski, E.G., Sutter, M., Agostoni, M., Lechno-Yossef, S., Cai, F., Montgomery, B.L., and Kerfeld, C.A. (2017). Additional families of orange carotenoid proteins in the photoprotective system of cyanobacteria. *Nat. Plants* 3, 17089.
34. Takaichi, S. (2011). Carotenoids in algae: distributions, biosyntheses and functions. *Mar. Drugs* 9, 1101–1118.
35. Ashby, M.K., and Mullineaux, C.W. (1999). The role of ApcD and ApcF in energy transfer from phycobilisomes to PS I and PS II in a cyanobacterium. *Photosynth. Res.* 61, 169–179.
36. Gindt, Y.M., Zhou, J., Bryant, D.A., and Sauer, K. (1992). Core mutations of *Synechococcus* sp. PCC 7002 phycobilisomes: a spectroscopic study. *J. Photochem. Photobiol. B* 15, 75–89.
37. Ughy, B., and Ajlani, G. (2004). Phycobilisome rod mutants in *Synechocystis* sp. strain PCC6803. *Microbiology (Reading)* 150, 4147–4156.
38. Watanabe, M., and Ikeuchi, M. (2013). Phycobilisome: architecture of a light-harvesting supercomplex. *Photosynth. Res.* 116, 265–276.
39. Koyama, K., Tsuchiya, T., Akimoto, S., Yokono, M., Miyashita, H., and Mimuro, M. (2006). New linker proteins in phycobilisomes isolated from the cyanobacterium *Gloeobacter violaceus* PCC 7421. *FEBS Lett.* 580, 3457–3461.
40. Bishop, C.L., Ulas, S., Baena-Gonzalez, E., Aro, E.-M., Purton, S., Nugent, J.H.A., and Mäenpää, P. (2007). The PsbZ subunit of Photosystem II in *Synechocystis* sp. PCC 6803 modulates electron flow through the photosynthetic electron transfer chain. *Photosynth. Res.* 93, 139–147.
41. Shi, L.-X., and Schröder, W.P. (2004). The low molecular mass subunits of the photosynthetic supracomplex, photosystem II. *Biochim. Biophys. Acta* 1608, 75–96.
42. Roose, J.L., and Pakrasi, H.B. (2008). The Psb27 protein facilitates manganese cluster assembly in photosystem II. *J. Biol. Chem.* 283, 4044–4050.
43. Bečková, M., Gardian, Z., Yu, J., Konik, P., Nixon, P.J., and Komenda, J. (2017). Association of Psb28 and Psb27 proteins with PSII-PSI supercomplexes upon exposure of *Synechocystis* sp. PCC 6803 to high light. *Mol. Plant* 10, 62–72.
44. Xu, Q., Odom, W.R., Guikema, J.A., Chitnis, V.P., and Chitnis, P.R. (1994). Targeted deletion of *psaJ* from the cyanobacterium *Synechocystis* sp. PCC 6803 indicates structural interactions between the Psaj and Psaf subunits of photosystem I. *Plant Mol. Biol.* 26, 291–302.
45. Xu, Q., Yu, L., Chitnis, V.P., and Chitnis, P.R. (1994). Function and organization of photosystem I in a cyanobacterial mutant strain that lacks Psaf and Psaj subunits. *J. Biol. Chem.* 269, 3205–3211.
46. Kawakami, K., Umena, Y., Iwai, M., Kawabata, Y., Ikeuchi, M., Kamiya, N., and Shen, J.-R. (2011). Roles of Psbl and Psbm in photosystem II dimer formation and stability studied by deletion mutagenesis and X-ray crystallography. *Biochim. Biophys. Acta* 1807, 319–325.
47. Inoue-Kashino, N., Kashino, Y., Satoh, K., Terashima, I., and Pakrasi, H.B. (2005). PsbU provides a stable architecture for the oxygen-evolving system in cyanobacterial photosystem II. *Biochemistry* 44, 12214–12228.
48. Veerman, J., Bentley, F.K., Eaton-Rye, J.J., Mullineaux, C.W., Vasil'ev, S., and Bruce, D. (2005). The PsbU subunit of photosystem II stabilizes energy transfer and primary photochemistry in the phycobilisome-photosystem II assembly of *Synechocystis* sp. PCC 6803. *Biochemistry* 44, 16939–16948.
49. Schwenkert, S., Legen, J., Takami, T., Shikanai, T., Herrmann, R.G., and Meurer, J. (2007). Role of the low-molecular-weight subunits PetL, PetG, and PetN in assembly, stability, and dimerization of the cytochrome b_6f complex in tobacco. *Plant Physiol.* 144, 1924–1935.
50. Schneider, D., Berry, S., Rich, P., Seidler, A., and Rögner, M. (2001). A regulatory role of the PetM subunit in a cyanobacterial cytochrome b_6f complex. *J. Biol. Chem.* 276, 16780–16785.
51. Castenholz, R.W. (2001). Oxygenic photosynthetic bacteria. In *Bergey's Manual of Systematics Bacteriology*, D.R. Boone, R.W. Castenholz, and G.M. Garrity, eds. (Springer-Verlag), pp. 474–487.
52. Koyama, K., Suzuki, H., Noguchi, T., Akimoto, S., Tsuchiya, T., and Mimuro, M. (2008). Oxygen evolution in the thylakoid-lacking cyanobacterium *Gloeobacter violaceus* PCC 7421. *Biochim. Biophys. Acta* 1777, 369–378.
53. Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P.A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* 37, 540–546.
54. Davis, J.J., Wattam, A.R., Aziz, R.K., Brettin, T., Butler, R., Butler, R.M., Chlenski, P., Conrad, N., Dickerman, A., Dietrich, E.M., et al. (2020). The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. *Nucleic Acids Res.* 48 (D1), D606–D612.
55. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.
56. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760.
57. Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K., and Earl, A.M. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* 9, e112963.
58. Kanehisa, M., Sato, Y., and Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG Tools for functional characterization of genome and metagenome sequences. *J. Mol. Biol.* 428, 726–731.
59. Charif, D., and Lobry, J.R. (2007). SeqinR 1.0-2: a contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. In *Structural Approaches to Sequence Evolution: Molecules, Networks, Populations*, U. Bastolla, M. Porto, H.E. Roman, and M. Vendruscolo, eds. (Springer-Verlag), pp. 207–232.
60. Emms, D.M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20, 238.
61. Conway, J.R., Lex, A., and Gehlenborg, N. (2017). UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* 33, 2938–2940.
62. Wang, Y., Tang, H., DeBarry, J.D., Tan, X., Li, J., Wang, X., Lee, T.H., Jin, H., Marler, B., Guo, H., et al. (2012). MScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40, e49.
63. Mirarab, S., Nguyen, N., Guo, S., Wang, L.-S., Kim, J., and Warnow, T. (2015). PASTA: ultra-large multiple sequence alignment for nucleotide and amino-acid sequences. *J. Comput. Biol.* 22, 377–386.
64. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534.
65. Drummond, A.J., and Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7, 214–221.

66. Miller, M.A., Pfeiffer, W., and Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In 2010 Gateway Computing Environments Workshop (GCE) (IEEE), pp. 1–8.
67. Lartillot, N., Lepage, T., and Blanquart, S. (2009). PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 25, 2286–2288.
68. Nelson, J.M., Hauser, D.A., Gudiño, J.A., Guadalupe, Y.A., Meeks, J.C., Salazar Allen, N., Villarreal, J.C., and Li, F.-W. (2019). Complete genomes of symbiotic cyanobacteria clarify the evolution of Vanadium-nitrogenase. *Genome Biol. Evol.* 11, 1959–1964.
69. Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M., and Stanier, R.Y. (1979). Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *Microbiology* 111, 1–61.
70. Dubbs, J.M., and Bryant, D.A. (1991). Molecular cloning and transcriptional analysis of the *cpeBA* operon of the cyanobacterium *Pseudanabaena* species PCC7409. *Mol. Microbiol.* 5, 3073–3085.
71. Buchfink, B., Xie, C., and Huson, D.H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60.
72. Muzzopappa, F., and Kirilovsky, D. (2020). Changing color for photoprotection: The orange carotenoid protein. *Trends Plant Sci.* 25, 92–104.
73. Villarreal A, J.C., and Renzaglia, K.S. (2006). Structure and development of *Nostoc* strands in *Leiosporoceros dussii* (Anthocerotophyta): a novel symbiosis in land plants. *Am. J. Bot.* 93, 693–705.
74. Ho, M.-Y., Shen, G., Canniffe, D.P., Zhao, C., and Bryant, D.A. (2016). Light-dependent chlorophyll *f* synthase is a highly divergent paralog of *PsbA* of photosystem II. *Science* 353, aaf9178.
75. Sakamoto, T., and Bryant, D.A. (1998). Growth at low temperature causes nitrogen limitation in the cyanobacterium *Synechococcus* sp. PCC 7002. *Arch. Microbiol.* 169, 10–19.
76. Huang, J.-Y., Chiu, Y.-F., Ortega, J.M., Wang, H.-T., Tseng, T.-S., Ke, S.-C., Roncel, M., and Chu, H.-A. (2016). Mutations of cytochrome *b₅₅₉* and *PsbJ* on and near the *Q_C* Site in Photosystem II influence the regulation of short-term light response and photosynthetic growth of the cyanobacterium *Synechocystis* sp. PCC 6803. *Biochemistry* 55, 2214–2226.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Anthocerotibacter panamensis	This paper; UTEX Culture Collection of Algae	Strain ID: UTEX 3164
Synechocystis sp	Institute of Plant and Microbial Biology, Academia Sinica, Taiwan	Strain ID: PCC 6803
Chlorogloeopsis fritschii	Pasteur Culture Collection	Strain ID: PCC 9212
Critical commercial assays		
MinION R9 flowcell	Nanopore	Cat#FLO-MIN106D
Ligation Sequencing kit	Nanopore	Cat#SQK-LSK109
Deposited data		
Nanopore and Illumina raw reads	This paper	NCBI SRA: SRR12713672, SRR12713673
Genome	This paper	NCBI BioProject: PRJNA665722
Software and algorithms		
Flye v2.4.1	Kolmogorov et al. ⁵³	https://github.com/fenderglass/Flye/
Medaka	N/A	https://github.com/nanoporetech/medaka
PATRIC	Davis et al. ⁵⁴	https://www.patricbrc.org
TRIMMOMATIC	Bolger et al. ⁵⁵	https://github.com/usadellab/Trimmomatic
BWA	Li and Durbin ⁵⁶	https://github.com/lh3/bwa
PILON v3.6.1	Walker et al. ⁵⁷	https://github.com/broadinstitute/pilon
KEGG (BlastKOALA)	Kanehisa et al. ⁵⁸	https://www.kegg.jp/blastkoala/
Seqinr v.3.6.1	Charif and Lobry ⁵⁹	http://seqinr.r-forge.r-project.org/src/mainmatter/getseqflat.pdf
CompareM v0.1.1	N/A	https://github.com/dparks1134/CompareM
Orthofinder v2.3.12	Emms and Kelly ⁶⁰	https://github.com/davidemms/OrthoFinder
UpsetR	Conway et al. ⁶¹	https://www.rdocumentation.org/packages/UpSetR/versions/1.4.0
MCSanX	Wang et al. ⁶²	https://github.com/wyp1125/MCSanX
AccuSyn	N/A	https://accusyn.usask.ca
PASTA v3	Mirarab et al. ⁶³	https://github.com/smirarab/pasta
IQTREE v1.6.11, v.2.0.3	Minh et al. ⁶⁴	https://github.com/iqtree/iqtree2
BEAST v1.8.3	Drummond and Rambaut ⁶⁵	http://beast.community
CIPRES Science Gateway v3.3	Miller et al. ⁶⁶	https://www.phylo.org

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
LogCombiner	Drummond and Rambaut ⁶⁵	http://beast.community
TreeAnnotator	Drummond and Rambaut ⁶⁵	http://beast.community
PhyloBayes	Lartillot et al. ⁶⁷	https://hpc.nih.gov/apps/PhyloBayes.html
Other		
Sequence alignments, genome and annotation files, input and output files for BEAST, PhyloBayes, and Orthofinder	This paper	https://gitlab.com/NasimR/anthocerotibacter-panamensis
MAPADA	Shanghai Mapada	Cat#UV1800
Spectrophotometer UV1800	Instruments	
F-4500 fluorescence spectrophotometer	Hitachi	Cat#F-4500
Oxytherm System oxygen electrode	Hansatech Instruments	Cat#V-560

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Fay-Wei Li (fi329@cornell.edu).

Materials availability

The type strain of *Anthocerotibacter panamensis* C109 was deposited at Culture Collection of Algae at the University of Texas at Austin (accession: UTEX 3164) and the permanent holotype specimen was archived at Bailey Hortorium (accession: BH 283079) at Cornell University.

Data and code availability

The raw nanopore and Illumina reads generated during this study are available at NCBI SRA:SRR12713672 and SRR12713673. The genome assembly was deposited at NCBI BioProject: PRJNA665722. Phylogenetic and orthogroup datasets are available at <https://gitlab.com/NasimR/anthocerotibacter-panamensis>.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cyanobacteria cultures

The hornwort *Leiosporoceros dussii* (Steph.) Hässel was collected in Coclé, Panama (8°37'34.6" N, 80°08'13.7" W) under the collection permit No. SE/P-33-16. Sterilization of the plant thalli and initial culturing of *Anthocerotibacter panamensis* isolate (type strain C109) followed the protocol of Nelson et al.⁶⁸ *Synechocystis* sp. PCC 6803 was a gift from Dr. Hsiu-An Chu at Institute of Plant and Microbial Biology, Academia Sinica, Taiwan, and *Chlorogloeopsis fritschii* PCC 9212 was obtained from the Pasteur Culture Collection.⁶⁹ *Anthocerotibacter*, *Chlorogloeopsis fritschii* PCC 9212, and *Synechocystis* sp. PCC 6803 were grown in the B-HEPES growth medium,⁷⁰ a modified BG11 medium containing 1.1 g L⁻¹ HEPES (final concentration) with the pH adjusted to 8.0 with 2.0 M KOH. Cool white LED light provided continuous illumination at either 10 or 100 μmol photons m⁻² s⁻¹ in a 30°C growth chamber supplemented with 1% (v/v) CO₂ in air. The doubling time is estimated based on the change of OD₇₅₀ absorption over time.

Taxonomic treatments

Anthocerotibacter F.-W. Li gen. nov. (Figure 1)

Differ from genera of Chroococcales, Chroococcidiopsidales, Nostocales, Oscillatoriales, Pleurocapsales, Spirulinales, and Synechococcales in lacking thylakoid membranes. Most similar to *Gloeobacter* in Gloeobacterales but lacks phycoerythrin.

Etymology. An.tho.ce.ro.ti.bac'ter. N.L. masc. n. *Anthoceros* a genus of hornworts; N.L. masc. n. *bacter* a rod; N.L. masc. n. *Anthocerotibacter* a rod from hornworts.

Type species. *Anthocerotibacter panamensis* F.-W. Li sp. nov.

***Anthocerotibacter panamensis* F.-W. Li sp. nov. (Figure 1)**

The cells are rod-shaped, 0.5–1 μm wide and 2–10 μm long. Thylakoid membranes are absent. Carboxysomes are present, and polyphosphate granules often appear toward the cell poles. The major pigments are allophycocyanin, phycocyanin, and chlorophyll *a*. No phycoerythrin nor phycoerythrocyanin is present. Colonies are generally light to dark green in both liquid and solid BG11 medium, and never purple, orange, or brown like in *Gloeobacter*. Growth can occur under 10–100 $\mu\text{mol photons m}^{-2}\text{s}^{-1}$ light intensities. No nitrogenase genes are present and cells do not grow without supplementing combined nitrogen. The type strain (C109) was originally isolated from a surface-sterilized hornwort *Leiosporoceros dussii* from Coclé, Panama (8°37'34.6" N, 80°08'13.7" W). GC content of the type strain is 55.51%.

Etymology. pa.na.men'sis. N.L. masc./fem. adj. panamensis, of or pertaining to Panama where the type strain was isolated.

Holotype. A permanent slide of the type strain C109 was archived at Bailey Hortorium (accession: BH 283079) at Cornell University.

Type culture. The type strain C109 was deposited at Culture Collection of Algae at the University of Texas at Austin (accession: UTEX 3164).

METHOD DETAILS**DNA extraction**

Cyanobacteria colonies from either liquid cultures or solid BG11 plates were centrifuged at 17,000 $\times$ g for 2 mins to pellet the cells. The cells were ground with a pestle in liquid nitrogen and resuspended in a prewarmed 2X CTAB solution with 1% beta-mercaptoethanol. The mixture was incubated at 65°C for 1 h, with gentle mixing every 15 min. Equal volume of 24:1 chloroform:isoamyl alcohol was added to the samples twice, each time with 5 min mixing on a Labnet Mini Labroller and 5 min centrifugation at 17,000 $\times$ g. The supernatant was transferred to a new tube by wide bore pipette tips to maintain DNA integrity. DNA was precipitated with an equal volume of isopropanol and pelleted for 30 min at 17,000 $\times$ g and 4°C. The DNA pellets were washed twice with 70% ethanol, then air-dried in a sterile air hood and resuspended in Tris–HCl pH 7.5, before treated with RNase for 1 h at 37°C.

Nanopore and Illumina sequencing

Genomic DNA was sequenced on both Oxford Nanopore MinION and Illumina platforms. Nanopore libraries were prepared using the Ligation Sequencing kit (SQK-LSK109) and sequenced on two MinION R9 flowcells (FLO-MIN106D) for 60 hours. In one MinION run, the focal sample was multiplexed with another cyanobacteria sample and barcoded by Native Barcoding Expansion Kit 1–12, while the other run contains only the focal sample. Basecalling was carried out using Guppy v3.1.5 with the high accuracy flip-flop mode. The Illumina library was prepared and sequenced on NextSeq500 by MiGS (Microbial Genome Sequencing Center).

Genome assembly and annotation

The basecalled nanopore reads were first filtered to remove reads shorter than 10kb, and then *de novo* assembled using Flye v2.4.1⁵³ with the “–plasmids” flag on, followed by polishing by medaka (<https://github.com/nanoporetech/medaka>). The resulting assembly contained three circular contigs. Based on the PATRIC taxonomic assignments,⁵⁴ one has the closest affinity to *Gloeobacter*, while the other two belong to *Pseudomonas fluorescens*, indicating contaminations were introduced during liquid culture (which was later confirmed by light microscopy). We therefore started a new culture for generating Illumina data. After trimming by TRIMMOMATIC (min score = 25, min length = 36),⁵⁵ the resulting Illumina reads had 91.74% mapping rate against the cyanobacterial contig (based on the bwa mem aligner⁵⁶), suggesting most (if not all) contaminations were removed. Genome polishing with Illumina reads was done by pilon⁵⁷ for 4 iterations. Structural and functional annotation was performed using PATRIC,⁵⁴ with additional KEGG annotation by BlastKOALA⁵⁸ (Data S2).

Genome comparisons

Pairwise 16S rDNA sequence divergence and Average Amino-acid Identity (AAI) were calculated by seqinr v3.6–1⁵⁹ and CompareM v0.1.1 (<https://github.com/dparks1134/CompareM>), respectively. We used Orthofinder v2.3.12⁶⁰ to classify families of orthologous genes in the genomes of *Anthocerotibacter* and 102 other cyanobacterial species (Data S1; Figure S1). UpSetR⁶¹ was used to summarize the shared and unique orthogroups found in *Anthocerotibacter*. To infer synteny between *Anthocerotibacter* and the two *Gloeobacter* species, MCScanX⁶² was used with Diamond⁷¹ as the search engine. The syntenic relationship was then visualized in AccuSyn (<https://accusyn.usask.ca>).

Phylogenetic analyses and divergence time estimates

We first compiled a 16S data matrix that includes major Cyanobacteria lineages, outgroups,⁶ and available environmental sequences with high sequence similarities to *Gloeobacter* and *Anthocerotibacter*.^{11,12} Alignment was done by PASTA v3,⁶³ and maximum-likelihood phylogeny was inferred using IQTREE v1.6.11 with 1,000 ultrafast bootstrap replicates.⁶⁴

Next, we integrated our data with previous molecular clock analyses on Cyanobacteria: Shih et al.⁶ and Baracaldo et al.²⁴ The dataset from Shih et al.⁶ is composed of 11 concatenated protein sequences (*atpA*, *atpB*, *atpE*, *atpF*, *atpH*, *atpI*, *rpl2*, *rpl16*, *rps3*, *rps12*, and elongation factor Tu) and 16S nucleotide sequences. We used the same age calibrations and substitution models (CpREV for proteins and GTR+Gamma for 16S) as in Shih et al.⁶ Phylogeny and divergence time were inferred using BEAST v.1.8.3⁶⁵ on CIPRES Science Gateway (v3.3).⁶⁶ Five MCMC were run in total, and for each chain samples were taken every 10,000 generations. The

posterior distributions were inspected in Tracer v.1.7.1 to ensure proper mixing and convergence. The resulting trees from five runs were combined by LogCombiner and summarized by TreeAnnotator.⁶⁵

The dataset from Baracaldo et al.²⁴ is composed of 8 protein-coding genes (*atpA*, *atpB*, *petB*, *psaC*, *psbA*, *psbD*, *rbcl*, and *S12*). Following the protocol of Baracaldo et al.,²⁴ PhyloBayes⁶⁷ was used to estimate divergence time, with CAT-GTR+Gamma substitution model, three different relaxed molecular clock models: log normal (ln), Cox–Ingersoll–Ross model (cir) and uncorrelated gamma multiplier (ugam), and two different root priors: (1) 95% of the prior distribution falls between 2.32 and 2.7 bya and (2) between 2.3 and 3 bya.

Phylogenetic inferences were also done for selected orthogroups using IQTREE v1.6.11 with 1,000 ultrafast bootstrap replicates and automatic model selection (Figure S3).⁶⁴

Phylogeny of orange carotenoid proteins (OCP)

OCP is composed of two domains, N-terminal domain (NTD) and C-terminal domain (CTD), connected by a flexible loop linker.⁷² Both domains have their own paralogs with either stand-alone NTD or CTD; the former is called helical carotenoid protein (HCP) and the latter CTD-like proteins (CTDH).⁷² *Anthocerotibacter* has one OCP and one NTD, but no CTDH (Data S2). While Grettenberger et al.¹² did report on the presence of OCP in *Candidatus Aurora*, they did not distinguish OCP, HCP, and CTDH. Our reanalysis showed that *Candidatus Aurora* has one OCP, one HCP, and one putative CTDH. We then incorporated the OCP sequences from *Anthocerotibacter* and *Candidatus Aurora* into the dataset compiled by Bao et al.,³³ which was used to define the major OCP clades.³³ Sequence alignment was done by PASTA v3⁶³ followed by phylogenetic inference using IQTREE v2.0.3 with automatic model selection and 1,000 ultrafast bootstrap replicates⁶⁴ (Figure S2).

Phylogeny of phycobilisome linker proteins

The phycobilisome linkers from *Anthocerotibacter* (Data S2) and *Gloeobacter violaceus* were added to the protein sequences compiled by Watanabe and Ikeuchi.³⁸ Sequence alignment was done by PASTA v3,⁶³ and phylogeny was reconstructed using IQTREE v2.0.3 with automatic model selection and 1,000 ultrafast bootstrap replicates⁶⁴ (Figure S4).

Microscopy

Pure cultures were processed for transmission electron microscopy as described previously⁷³ with slight modifications. Cultures were fixed in 3% glutaraldehyde, 1% fresh formaldehyde, and 0.75% tannic acid in 0.05 M Na-cacodylate buffer, pH 7, for 3 h at room temperature. After several rinses in 0.1M buffer, the samples were post-fixed in buffered (0.1M, pH 6.8) 1% osmium tetroxide overnight at 4°C, dehydrated in an ethanol series and embedded in Spurr's resin via ethanol. Thin sections were cut with a diamond histo-knife, stained with methanolic uranyl acetate for 15 min and in Reynolds' lead citrate for 10 min, and observed with a Hitachi H-7100 transmission electron microscope at the Imaging-Microscopy Platform of the IBIS, Université Laval. Images of colonies on an agar plate were taken by a SMZ-171TLED stereomicroscope (Bio Pioneer Tech Co, Taiwan) equipped with a TrueChrome AF digital camera (TUCSEN, China). Light microscopy images were taken under a Leica DMR microscope (Leica, Germany) equipped with a Leica MC 170HD digital camera (Leica, Germany).

Absorption and fluorescence spectral analyses

The spectroscopy measurements were performed as previously described with modifications.⁷⁴ The OD₇₅₀ values of the cyanobacterial cultures were adjusted to 0.25 in 1 mL of B-HEPES medium for absorption spectroscopy, and 0.25 in 1 mL of PEG buffer (500 µl of B-HEPES medium and 500 µl of 30% (w/v) PEG4000) for fluorescence spectroscopy. The measurement of absorption spectra (wavelengths from 400 to 800 nm) was performed by MAPADA UV1800 with an equipped PC software (Datech Instruments Ltd., Taiwan). Fluorescence spectroscopy was performed by the F-4500 fluorescence spectrophotometer (Hitachi, Japan). Fluorescence emission was recorded from 600 to 800 nm by excitation at 440 nm for chlorophyll *a*.

Measurements of pigment composition and oxygen evolution

The quantifications of chlorophyll *a*, carotenoids, and phycobiliproteins were performed as previously described.⁷⁵ Steady-state rates of oxygen evolution were measured as previously described with minor modifications.⁷⁶ An Oxytherm System oxygen electrode (Model V-560, Hansatech Instruments Ltd, UK) was used for the measurement. Cyanobacterial cells were diluted in B-HEPES medium to the final concentration as 3.5 µg (Chl) ml⁻¹ in a stirred, water-jacketed cell chamber at 25°C. Potassium ferricyanide and 2,6-dichloro-p-benzoquinone (DCBQ) were added sequentially into the cell chamber as artificial electron acceptors to the final concentration as 2 mM. Different strengths of light intensities were provided from both sides of the cell chamber by two fiber-optic illuminators (Dolan-Jenner model MI 150, USA). Data recording and processing were performed with the manufacture software O₂ view (Hansatech Instruments Ltd, UK).

QUANTIFICATION AND STATISTICAL ANALYSIS

For gene phylogenies, the best-fitting models were selected by the ModelFinder function implemented in IQTREE according to the Bayesian Information Criterion. We inferred the branch supports using IQTREE's ultrafast bootstrapping method. To estimate the divergence time and 95% HPD, BEAST and PhyloBayes analyses were carried out as described above. Three biological replicates were done for quantifying growth rate, pigments, and oxygen evolution.