

Horizontal Gene Transfers Underpin Ribose Heterotrophy and Central Carbon Metabolism Remodeling in *Gloeobacteraceae*

Edi Sudianto ¹, Denis Baurain ¹, Luc Cornet ^{1,*}

¹InBioS–PhytoSYSTEMS, Eukaryotic Phylogenomics, University of Liège, Liège 4000, Belgium

*Corresponding author: E-mail: luc.cornet@uliege.be.

Accepted: May 15, 2026

Abstract

Gloeobacterales occupy a key phylogenetic position among cyanobacteria and are distinguished by the absence of thylakoid membranes. Using comparative genomics and phylogenetic analyses, we show that horizontal gene transfer has played a major role in shaping the central carbon metabolism of this lineage. In *Gloeobacteraceae*—one of the two families within the order—we identify a complete ribose ATP-synthase binding cassette importer and associated metabolic enzymes that enable ribose uptake and assimilation into central carbon metabolism alongside photosynthesis, indicative of a photomixotrophic lifestyle. Beyond ribose utilization, their central carbon metabolism exhibits a mosaic architecture shaped by the integration of foreign genes into the Calvin–Benson–Bassham cycle, the pentose phosphate pathway, and the Embden–Meyerhof–Parnas pathway. Uniquely, these genes appear to have been acquired through multiple independent transfer events, as reflected by their dispersed genomic locations and diverse bacterial donors, including other cyanobacteria and *Pseudomonadota*. These findings highlight *Gloeobacterales* as a dynamic lineage that continues to adapt and evolve through metabolic innovation and the assimilation of foreign genes into its genomes.

Key words: *Gloeobacterales*, cyanobacteria, photomixotrophy, Calvin–Benson–Bassham cycle, pentose phosphate pathway, Embden–Meyerhof–Parnas pathway.

Significance

Gloeobacterales are unusual among cyanobacteria because they lack thylakoids, a feature often interpreted as ancestral, yet the evolution of their central carbon metabolism remains poorly understood. Here, we show that *Gloeobacterales* have repeatedly acquired genes from other bacteria, including ribose importer genes in *Gloeobacteraceae* and additional genes involved in central carbon metabolism. These genes likely enable *Gloeobacteraceae* to use external carbon sources in addition to photosynthesis and suggest a mosaic origin for parts of their central carbon metabolism, with pathway components derived from different bacterial lineages and integrated into the native metabolic framework. Our findings indicate that gene exchange has contributed to metabolic flexibility in this lineage and provide new insight into the evolution of *Gloeobacterales*.

Introduction

Cyanobacteria are the only prokaryotes capable of oxygenic photosynthesis. This process relies on a linear electron transfer chain to convert light energy and CO₂ into carbohydrates and release O₂ as a byproduct (Johnson 2016). Photosynthesis is estimated to have originated ~3.4 billion years ago in the last common ancestor of the cyanobacteria (Fournier et al. 2021). Owing to this ancient innovation, cyanobacteria are traditionally regarded as strictly photosynthetic organisms (Smith et al. 1967). However, both experimental and environmental evidence have increasingly demonstrated that photomixotrophy—the ability to perform photosynthesis while assimilating additional carbon sources—is more widespread than previously assumed within the phylum (Muñoz-Marín et al. 2024). The first large-scale study by Rippka et al. (1979) in the 1970s found that diverse cyanobacteria are capable of photomixotrophy on a range of carbon sources, including glucose, fructose, sucrose, ribose, and glycerol. For instance, *Nostoc punctiforme* PCC 73102 can metabolize ribose, glucose, fructose, and sucrose (Rippka et al. 1979; Stebbegg et al. 2023). In addition, picocyanobacteria have been shown to use amino acids, dimethylsulfoniopropionate, and phosphonate as organic carbon sources for photomixotrophy (Muñoz-Marín et al. 2020; Druce et al. 2025). These substrates are typically funneled into central carbon metabolic pathways, such as the pentose phosphate pathway (PPP), the Entner–Doudoroff (ED) pathway, the Embden–Meyerhof–Parnas (EMP) pathway, and the Calvin–Benson–Bassham (CBB) cycle (Muñoz-Marín et al. 2020; Stebbegg et al. 2023; Lucius and Hagemann 2024).

In most cyanobacteria, photosynthesis occurs within specialized intracellular membranes known as thylakoids. Gloeobacterales, however, lack these membranes and instead localize photosynthetic complexes to the cytoplasmic membrane (Raven and Sánchez-Baracaldo 2021; Cornet 2025). This distinctive feature, together with their phylogenetic position within cyanobacteria, makes Gloeobacterales an important group for studying the diversification of cyanobacterial photosynthesis and thylakoid emergence (Cornet 2025). Despite their ecological diversity, Gloeobacterales remain poorly characterized due to their recalcitrance to cultivation; only four strains have been successfully isolated since their discovery in the 1980s (Grettenberger 2021; Raven and Sánchez-Baracaldo 2021; Saw et al. 2021). Here, we report the first genomic evidence of photomixotrophy within Gloeobacteraceae, expanding the known metabolic repertoire of this nonthylakoid-bearing cyanobacterial lineage. We further show that the central carbon metabolism of Gloeobacterales has been shaped by horizontal gene transfer (HGT) events that likely occurred

after the diversification of Gloeobacteraceae and Anthocerotibacteraceae. These observations reveal that Gloeobacterales, although retaining ancestral characteristics, such as lack of thylakoids (Guglielmi et al. 1981; Rexroth et al. 2011), presence of a divergent form of the D1 subunit of photosystem II (Cardona et al. 2015), as well as bacteriorhodopsin-like protein (Bernát et al. 2012), can acquire new genes and metabolic traits, demonstrating substantial adaptation since their divergence.

Results

Phylogeny and Hierarchical Orthologous Group Data Summary

Here we compiled a dataset of 343 cyanobacterial and closely related outgroup genomes, spanning both photosynthetic and nonphotosynthetic lineages. Using 251 conserved single-copy genes (SCGs), we reconstructed a phylogenomic species tree and inferred hierarchical orthologous groups (HOGs) from the corresponding predicted proteomes for downstream evolutionary analyses. Out of the 251 SCGs in the cyanobacterial reference profile hidden Markov models (pHMMs) of GToTree, our taxa contained 98 to 250 of the SCGs (Table S1). Nonphotosynthetic taxa possessed fewer SCGs than the photosynthetic ones (Table S1); specifically, “*Candidatus* Margulisbacteria” (104 to 142), “*Candidatus* Sericytochromatia” (112 to 149), and Vampirovibrionophyceae (98 to 150) contained fewer SCGs than photosynthetic cyanobacteria (159 to 250). The phylogeny reconstructed from these SCGs showed clear delineation of the major cyanobacterial orders with strong bootstrap supports (Fig. S1). We reconstructed and inferred 41,823 HOGs at the root node (N0 in Fig. S1). For downstream analyses, we identified 34 HOGs of interest corresponding to 37 selected genes primarily involved in ribose uptake and metabolism and central carbon metabolism (Table S2).

Ribose Uptake in Gloeobacteraceae

Our HOG analysis revealed that all Gloeobacteraceae taxa included in this study possess three genes associated with an ATP-synthase binding cassette (ABC) transporter system. These three genes are clustered in three distinct HOGs with genes annotated as fructose importers (*frtA*, *frtB*, and *frtC*) in other cyanobacteria (see Table S3) and code for an ATP-binding protein, a substrate-binding protein, and a permease protein, respectively. Interestingly, no Anthocerotibacteraceae taxa have any genes being clustered in these three HOGs. These genes are also absent from nonphotosynthetic cyanobacterial relatives (e.g. Vampirovibrionophyceae and “*C. Margulisbacteria*”; Table S3).

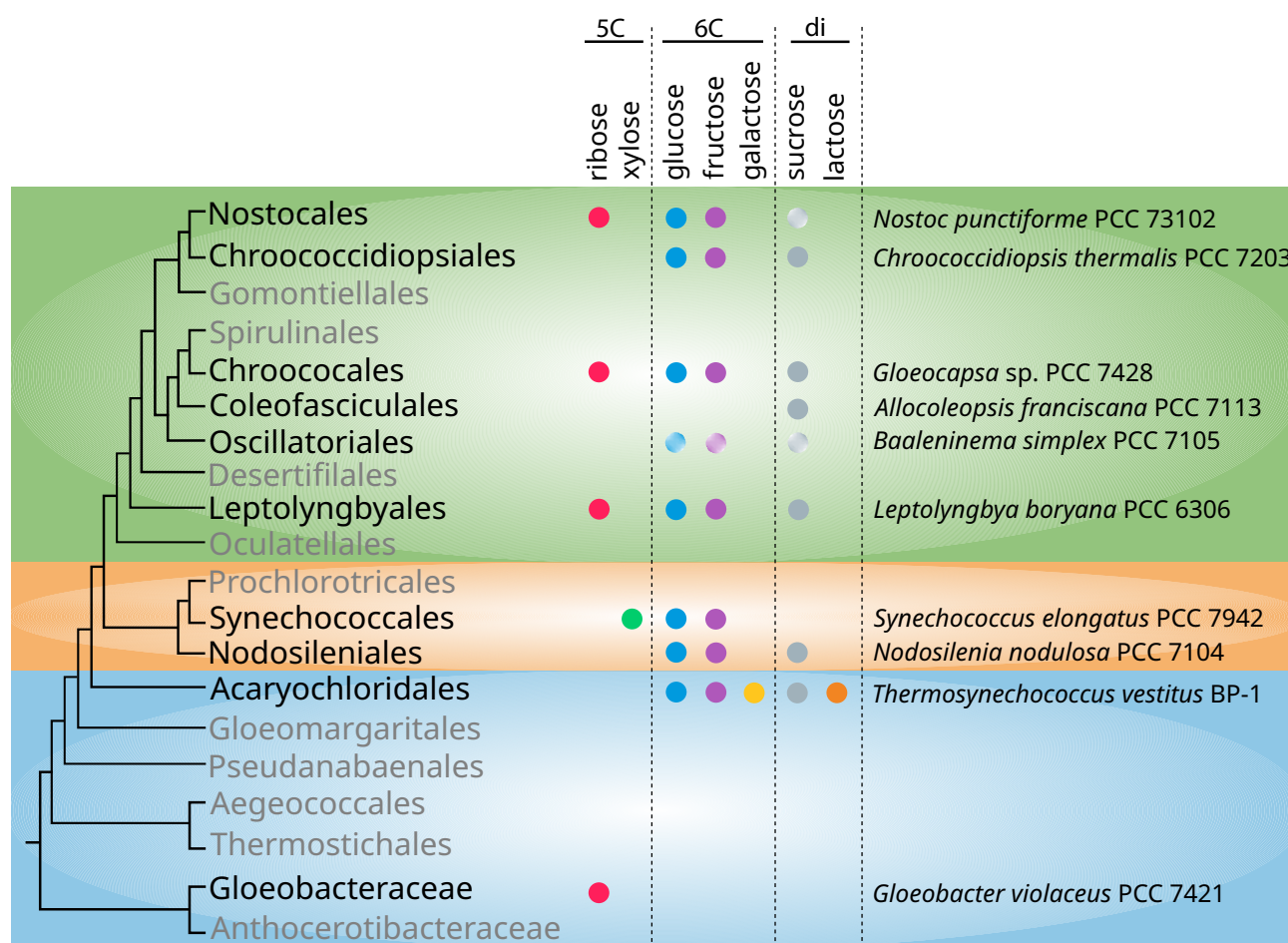


Fig. 1. Overview of heterotrophic capability among 19 cyanobacterial orders. The listed taxa and their heterotrophic capability (except for ribose heterotrophy in Gloeobacteraceae) were derived from Stebbins et al. (2023). The phylogeny backbone follows the classification of Strunecký et al. (2023). The Gloeobacterales order is divided into two families, Gloeobacteraceae and Anthocerotibacteraceae, to highlight differences in heterotrophic capability. Orders without known heterotrophic taxa are marked as gray. Circles with a white gradient indicate “weak” ability to utilize the listed sugars. 5C, five carbon monosaccharides (pentoses); 6C, six carbon monosaccharides (hexoses); di, disaccharides. The classification of cyanobacterial orders follows Gisriel et al. (2023).

Based on structural alignment analysis (Table S4) comparing the substrate-binding protein to reference proteins, we determined that this protein is likely a ribose importer (Template modeling score [TM-score]: 0.94; root mean square deviation [RMSD]: 1.5) rather than a fructose importer (TM-score: 0.75 to 0.77; RMSD: 2.09 to 2.38). This interpretation is also supported by the Kyoto Encyclopedia of Genes and Genomes (KEGG; based on sequence homology and the presence of specific motifs), which annotates the three genes as ribose transporter components (e.g. *rbsA* in *Gloeobacter violaceus*: <https://www.kegg.jp/entry/gvi:glr0808>).

These results indicate that all examined Gloeobacteraceae taxa possess the complete genetic repertoire required to assemble a functional ribose importer protein. Thus, the seven Gloeobacteraceae taxa included in this study are likely capable of utilizing

ribose as an external carbon source in addition to performing oxygenic photosynthesis. In contrast, the absence of these genes in the Anthocerotibacteraceae and their scattered distribution in other cyanobacteria suggest that the ribose heterotrophy is restricted to the Gloeobacteraceae and was potentially acquired only after it diverged from its sister family (Fig. 1). This discovery provides the first genomic evidence of heterotrophic capability in Gloeobacteraceae and suggests that photomixotrophy in this group is a derived metabolic innovation rather than an ancestral trait.

Gene Gain and Loss Mosaics in the Central Carbon Metabolism of Gloeobacterales

We found clear signatures of gene gain and loss events across the central carbon metabolism (Fig. 2). The three

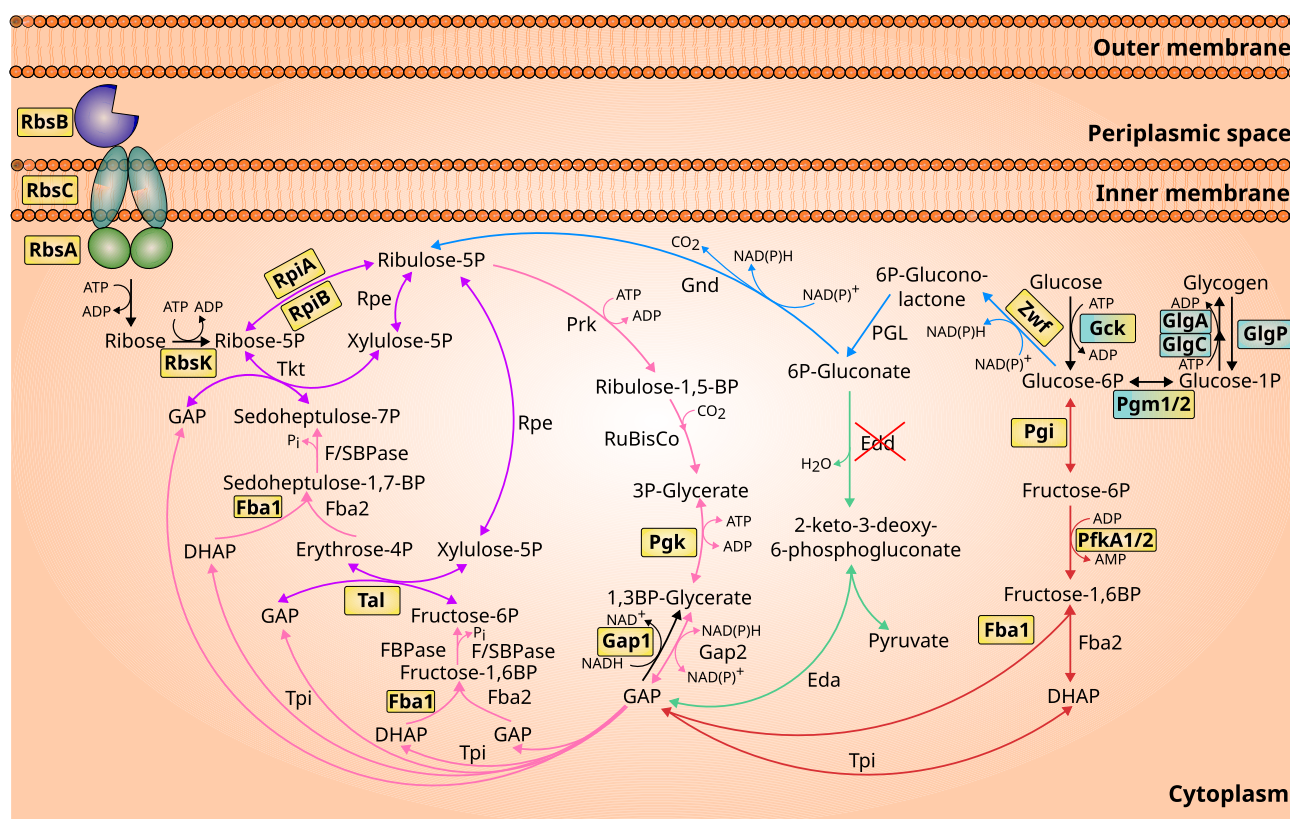


Fig. 2. Summary of the central carbon metabolism pathways involved in ribose import and utilization in Gloeobacteriales. Arrow colors denote distinct pathways: the CBB cycle is pink, the PPP is blue, shared pathways between CBB and PPP is purple, the ED pathway is green, and the EMP pathway is red. Key enzymes potentially acquired through HGT are highlighted in yellow boxes, while enzymes lost in some taxa are highlighted in cyan boxes. Enzymes exhibiting signatures of horizontally acquired genes before being lost in some taxa are shown in yellow–cyan boxes. An enzyme absent in all Gloeobacteriales taxa is marked with a red cross. Illustration based on metabolic pathway information from Bowyer and Leegood (1997) and Caliebe et al. (2025). RbsA, ribose import ATP-binding protein; RbsB, ribose import substrate-binding protein; RbsC, ribose import permease protein; RbsK, ribokinase; RpiA, ribose-5-phosphate isomerase A; RpiB, ribose-5-phosphate isomerase B; Rpe, ribose-5-phosphate epimerase; Prk, phosphoribulokinase; RuBisCo, ribulose-1,5-bis-p carboxylase/oxygenase; Pgk, phosphoglycerate kinase; Gap, glyceraldehyde 3-phosphate dehydrogenase; Tpi, triose-phosphate isomerase; Fba, fructose-bisphosphate aldolase; FBpase, fructose-1,6-bisphosphatase; F/SBPase, bifunctional fructose-1,6/sedoheptulose-1,7-bisphosphatase; Tal, transaldolase; Tkt, transketolase; Gnd, 6-phosphogluconate dehydrogenase; Pgl, 6-phosphogluconolactonase; Zwf, glucose-6-phosphate 1-dehydrogenase; Edd, 6-phosphogluconate dehydratase; Eda, 2-keto-3-deoxygluconate-6-phosphate (KDPG) aldolase; Pgi, glucose-6-phosphate isomerase; PfkA, phosphofructokinase A; Gck, glucokinase; Pgm1/2, phosphoglucomutase; GlgA, glycogen synthetase; GlgC, glucose-1-phosphate adenyltransferase catalytic site; GlgP, glycogen phosphorylase.

genes responsible for the ribose importer are likely products of HGT into an ancestor of Gloeobacteraceae (Figs. 3 and 4), as they are absent from the HOGs of Anthocerotibacteraceae and nonphotosynthetic close relatives (Table S3). Similarly, within the Gloeobacteriales, the ribokinase (*rbsK*) and ribose-5-phosphate isomerase B (*rpiB*) HOGs are exclusive to the *Gloeobacter* genus and Gloeobacteraceae, respectively, suggesting independent HGT events (Table S3). Although all Gloeobacteriales taxa possess a single *rpiA* homolog, phylogenetic analysis indicates that the Gloeobacteraceae copy is noncyanobacterial (Fig. 5). Interestingly, nonphotosynthetic cyanobacterial relatives lack both *rbsK* and *rpiA*, but encode the *rpiB* gene (Table S3).

In Gloeobacteriales, additional HGT-derived genes were identified across the CBB cycle, the PPP, and the EMP pathways, including *pgk*, *gap1*, *fba1*, *tal*, *zwf*, *pgi*, and *pfkA* (Fig. S2). In contrast, several genes (*gck*, *pgm1/2*, *glgA*, *glgC*, and *glgP*) are absent in subsets of Anthocerotibacteraceae (Fig. 2; Table S3). HGT analysis (Fig. S2) indicates that *gck* and *pgm1/2* show signs of horizontally acquired genes before being lost in some members of Anthocerotibacteraceae. Nevertheless, our HOG data reveal that the majority of cyanobacterial taxa, including Gloeobacteraceae, still retain these genes (Table S3). The *edd* gene was absent from all cyanobacterial genomes, while *pfkA* was detected only in three Anthocerotibacteraceae genomes, potentially indicating a recent acquisition event. Overall, our

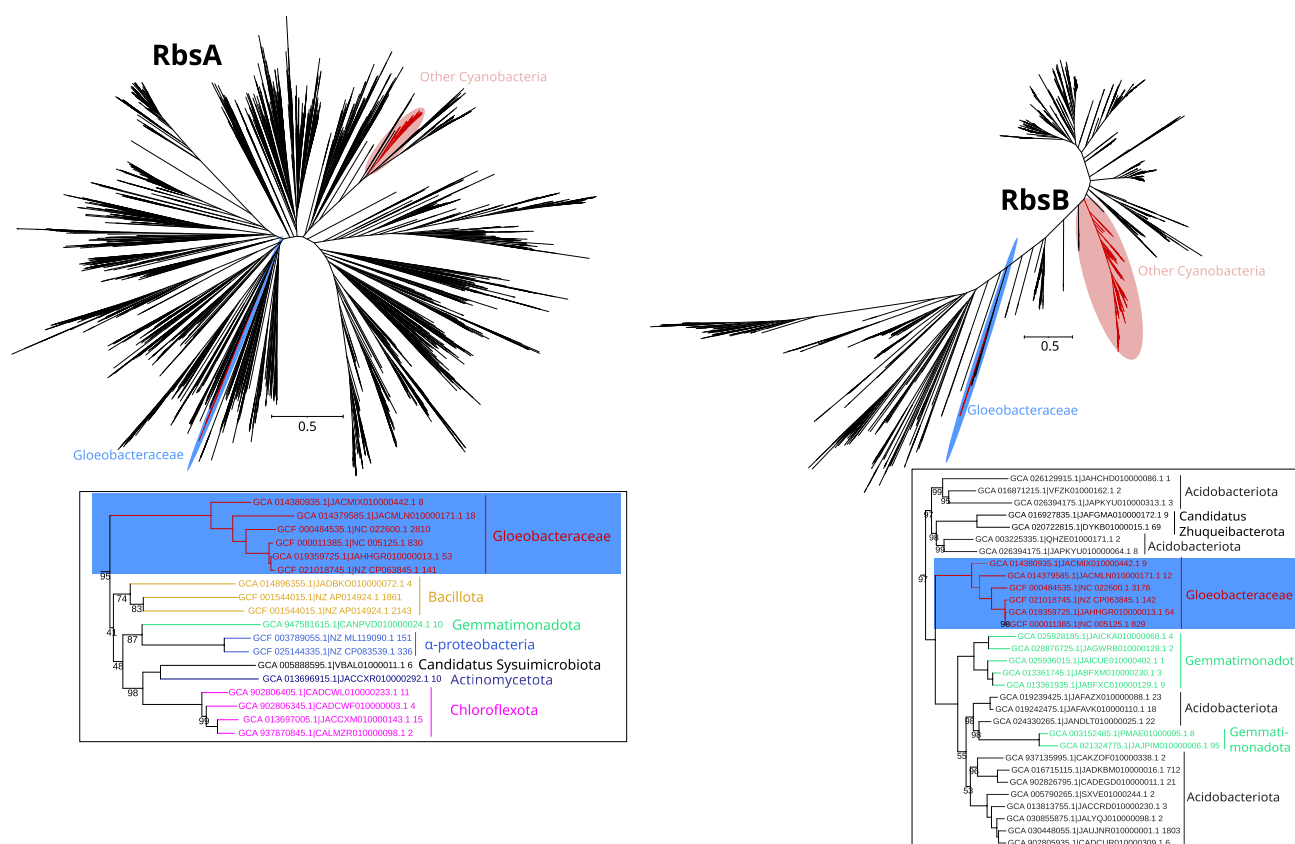


Fig. 3. Gene trees of ribose import ATP-binding protein (*RbsA*) and ribose import substrate-binding protein (*RbsB*). The red, blue, and green ellipses in the full gene trees show the position of other cyanobacteria, Gloeobacteraceae, and Anthocerotibacteraceae taxa, respectively. The tree snippets beside each gene tree display the potential donor of HGT to Gloeobacteraceae. Font colors in the snippets represent different bacterial groups and are independent of the ellipsis color scheme used in the full trees. Only nodes with bootstrap supports of <100% are shown. These maximum likelihood trees were reconstructed using LG+G substitution models and 1,000 ultrafast bootstrap replicates in the following alignment dimensions: *RbsA* (3,738 sequences × 1,329 sites); *RbsB* (1,942 sequences × 4,108 sites).

findings suggest that the central carbon metabolism in Gloeobacterales has been shaped by HGT and differential gene loss, resulting in a mosaic metabolic architecture.

Ribose Metabolism Genes in Gloeobacteraceae Are Mostly Acquired Via HGT

Phylogenetic analyses of the six ribose import and metabolism genes (including *rbsA*, *rbsB*, *rbsC*, *rbsK*, *rpiA*, and *rpiB*) demonstrate that all Gloeobacteraceae sequences were acquired through HGT (Figs. 3 to 5). Anthocerotibacteraceae lack these ribose-related genes entirely, except for *rpiA* (Figs. 3 to 5).

Surprisingly, the inferred donors are diverse, with only *rbsB* and *rbsC* likely originating from the same source, possibly Acidobacteriota or Gemmatimonadota. The gene trees indeed show that *rbsB* and *rbsC* share the same donor lineage, suggesting cotransfer as a part of a gene cluster. *RbsA*, however, likely has a distinct evolutionary origin.

There is no clear identification of the *rbsA* donor as its gene tree places the Gloeobacteraceae sequences as sister to a group composed of diverse bacteria, including Bacillota, Gemmatimonadota, Alphaproteobacteria, “*C. Syssymbiobacteria*,” Actinomycetota, and Chloroflexota. Thus, it can be inferred that the donor of *rbsA* is distinct from those of *rbsB* and *rbsC*.

Furthermore, our gene trees clearly indicate Alphaproteobacteria as the donor of *rbsK* and *rpiA* and SAR324/Desulfobacterota as the donor of *rpiB* in Gloeobacteraceae. *RpiA* represents an especially interesting case, where the HGT copy appears to have replaced the endogenous cyanobacterial copy still present in Anthocerotibacteraceae (Fig. 5). Topology tests (approximately unbiased [AU], Shimodaira–Hasegawa (SH), and Kishino–Hasegawa (KH); $P < 0.05$; Table S5) further support these inferences, rejecting the alternative topology enforcing cyanobacterial monophyly for *rbsA*, *rbsB*, *rbsC*, and *rpiA*. The remaining genes (*rbsK* and *rpiB*) were not included in topology

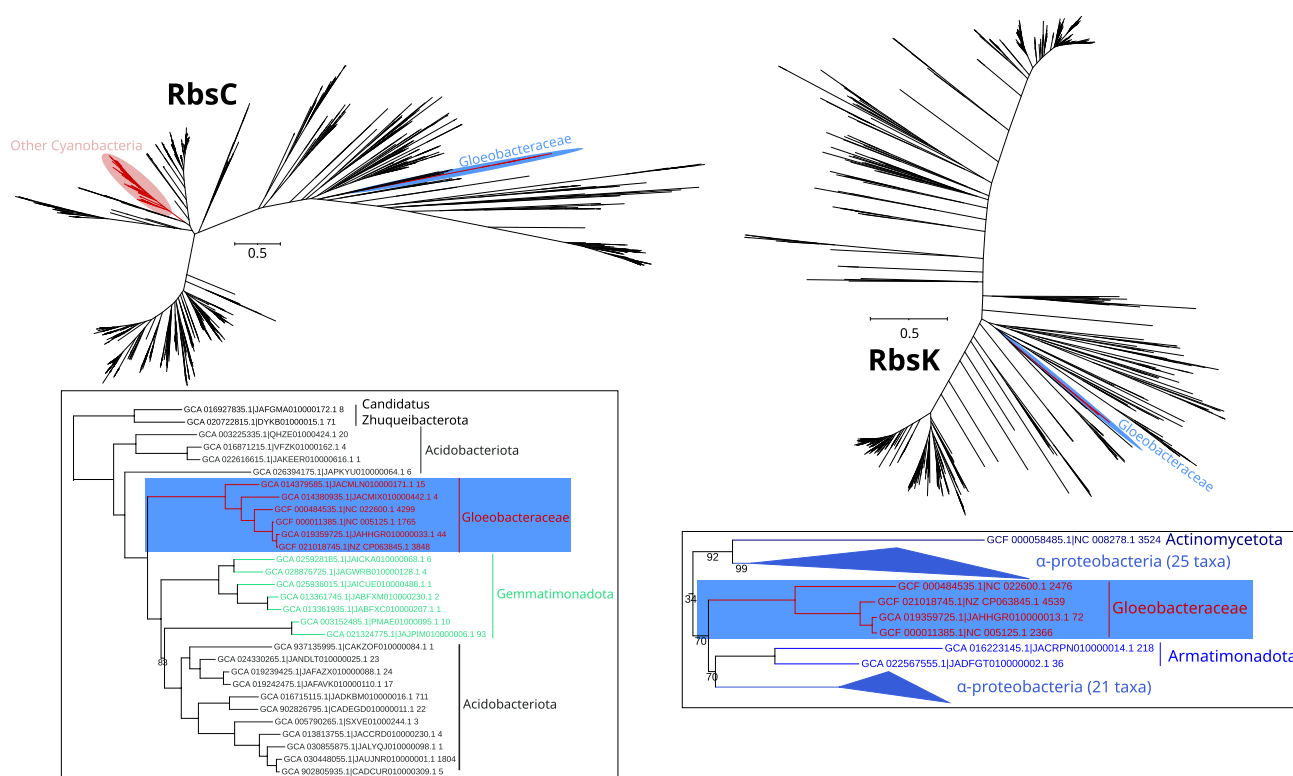


Fig. 4. Gene trees of ribose import permease protein (RbsC) and ribokinase (RbsK). The red, blue, and green ellipses in the full gene trees show the position of other cyanobacteria, Gloeobacteraceae, and Anthocerotibacteraceae taxa, respectively. The tree snippets beside each gene tree display the potential donor of HGT to Gloeobacteraceae. Font colors in the snippets represent different bacterial groups and are independent of the ellipsis color scheme used in the full trees. Only nodes with bootstrap supports of <100% are shown. These maximum likelihood trees were reconstructed using LG+F+G substitution models and 1,000 ultrafast bootstrap replicates in the following alignment dimensions: RbsC (1,797 sequences × 2,895 sites); RbsK (784 sequences × 398 sites).

testing because other phylogenetically informative cyanobacterial homologs were absent from the gene trees, despite their detection in the corresponding HOGs (Table S3). Together, these findings strongly imply that all genes involved in ribose import and metabolism in Gloeobacteraceae were independently acquired via HGT from multiple bacterial sources, underscoring the evolutionary plasticity of this cyanobacterial lineage.

Contrasting the Genetic Repertoire of Gloeobacteraceae and Anthocerotibacteraceae

To elucidate the evolutionary dynamics of genes involved in ribose import and metabolism and central carbon metabolic pathways across Gloeobacterales, we inferred patterns of gene acquisition and loss, along with their putative donor lineages (Fig. 6). This analysis revealed marked contrast in the evolutionary trajectories of Gloeobacteraceae and Anthocerotibacteraceae.

In Gloeobacteraceae, numerous horizontally transferred genes were inferred to have been acquired at

the base of the family. These include *rbsA*–*C*, *rpiAB*, *pgk_2*, *zwf_2*, and *fba1*, accompanied by the loss of *pgk_1*. These genes likely originated from phylogenetically diverse donors (Figs. 6 and S2), such as Acidobacteriota/Gemmatimonadota (*rbsBC*), Pseudomonadota (*rpiA* and *zwf_2*), and SAR324/Desulfobacterota (*rpiB*). The donors of *rbsA*, *pgk_2*, and *fba1* remain unidentified. Subsequent acquisitions include *rbsK* (putatively from Pseudomonadota) and *gap1* (other cyanobacteria) in the common ancestor of the four *Gloeobacter* species. Additional transfers, such as *tal_2* (cyanobacterial origin) and *pgi_2* (unknown origin), occurred later within the *G. violaceus*–*Gloeobacter morelensis* lineage.

In contrast, the Anthocerotibacteraceae is characterized by a single inferred HGT event (*tal_1*) and multiple gene losses (*pgm1/2*, *glgC*, and *glgA*) in the clade containing “*C. Cyanaurora vandensis*” and *Anthocerotibacter panamensis*. These taxa appear to have replaced their native transaldolase (*tal_1*) gene with an unclassified HGT-derived copy. Additional losses include *glgP*, which is absent from four species

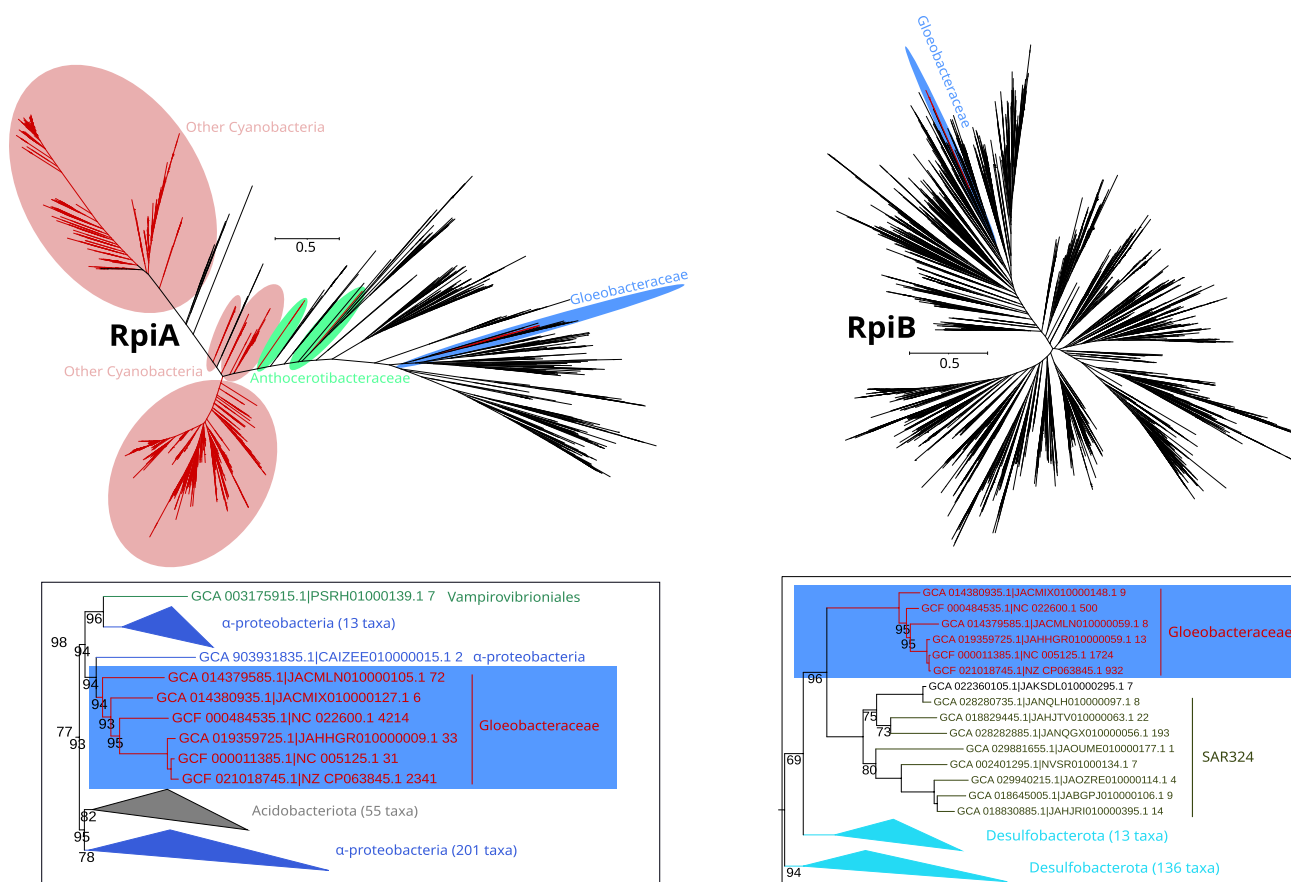


Fig. 5. Gene trees of ribose-5-phosphate isomerase a (RpiA) and ribose-5-phosphate isomerase B (RpiB). The red, blue, and green ellipses in the full gene trees show the position of other cyanobacteria, Gloeobacteraceae, and Anthocerotibacteraceae taxa, respectively. The tree snippets beside each gene tree display the potential donor of HGT to Gloeobacteraceae. Font colors in the snippets represent different bacterial groups and are independent of the ellipsis color scheme used in the full trees. Only nodes with bootstrap supports of <100% are shown. These maximum likelihood trees were reconstructed using LG+F+G substitution models and 1,000 ultrafast bootstrap replicates in the following alignment dimensions: RpiA (1,834 sequences × 434 sites); RpiB (3,740 sequences × 386 sites).

in the family but retained in one Gloeobacterales genome (GCA_031398925.1). Three taxa, including two “*C. Cyanaurora vandensis*” and a Gloeobacterales cyanobacterium (GCA_038226175.1), acquired *pgi*₂ from the Pseudomonadota. Note that despite being annotated as Gloeobacterales cyanobacterium or Gloeobacteraceae cyanobacterium (GCA_038226175.1, GCA_038245785.1, and GCA_031398925.1), these genomes are confidently placed within Anthocerotibacteraceae according to our species tree (Fig. S1). Furthermore, the “*C. Sivonenia alaskensis*” clade uniquely carries a *pfkA* gene of another cyanobacterial origin, while this gene is absent from all other Gloeobacterales taxa examined. These patterns suggest that Gloeobacteraceae underwent frequent HGT and metabolic innovation within the context of ribose uptake and central carbon metabolism. In contrast, Anthocerotibacteraceae followed a conservative

trajectory with selective gene loss and limited HGT in genes related to central carbon metabolism, reflecting divergent trends in metabolic evolution within the order. Whether this contrast extends beyond these metabolic functions to the broader genome remains to be determined.

Frequent Gene Rearrangements and Genome Plasticity in Gloeobacterales

Most horizontally transferred genes in Gloeobacterales are dispersed rather than organized in operons or conserved gene clusters. Synteny comparisons among the four complete genome assemblies of Gloeobacterales, including *G. violaceus*, *G. morelensis*, *Gloeobacter kilaueensis*, and *A. panamensis*, revealed limited conservation of local gene neighborhoods (Fig. 7).

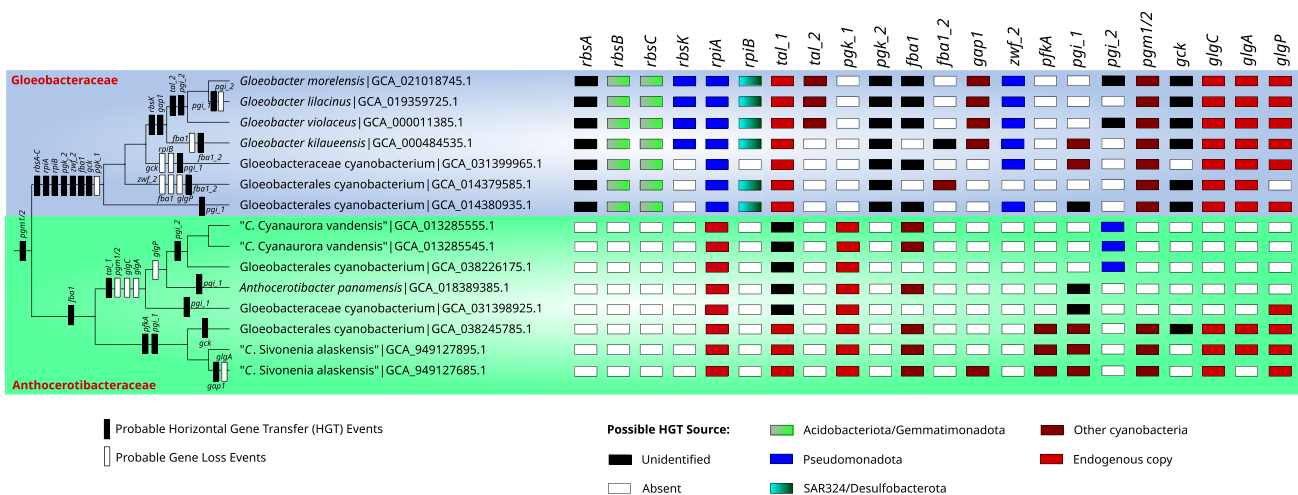


Fig. 6. Proposed model of HGT and gene loss events for genes involved in central carbon metabolism among the two Gloeobacterales families, as indicated in fig. 2. The Gloeobacterales phylogeny is based on the cyanobacterial phylogeny in Fig. S1. Colored boxes on the right denote the gene presence/absence and their potential HGT donor. It is noteworthy that three genomes are named either Gloeobacterales cyanobacterium (GCA_038226175.1, GCA_038245785.1) or Gloeobacteraceae cyanobacterium (GCA_031398925.1) while they clearly belong to Anthocerotobacteraceae in our species tree (Fig. S1).

The ribose importer genes *rbxA* and *rbxB* were adjacent only in *G. violaceus* and *G. morelensis*, but not in *G. kilaueensis*. Even native cyanobacterial gene arrangements, such as *prk-tkt*, *glgC*, and *eda*, differed markedly among the *Gloeobacter* genomes (Fig. 7a). Similarly, the *pgi_2* gene was located near the *rpiB* gene in *G. violaceus* but not in *G. morelensis* (Fig. 7b).

Genes acquired from the same lineages also displayed variable genomic positions. For example, *rpiA* and *zwf_2*, both originating from Pseudomonadota, are located close to each other in *G. kilaueensis* but are situated in distant genomic regions in *G. violaceus* and *G. morelensis* (Fig. 7c). In contrast, some ancestral operons, such as the *rbxL-rbxS*, have retained synteny across the family (Fig. 7d). Collectively, these results suggest that horizontally transferred genes are integrated in the Gloeobacterales genomes and subject to frequent rearrangements, reflecting postacquisition gene reshuffling rather than stable incorporation into conserved metabolic loci.

Discussion

Ribose Is Likely an Important Carbon Source for Gloeobacteraceae

Previous studies (White and Shilo 1975; Rippka et al. 1979; Stebbins et al. 2023) have identified some cyanobacteria capable of heterotrophic ribose utilization, including members of the Nostocales, Chroococcales, Leptolyngbyales, and Geitlerinemaales (Fig. 1). Here, we provide the first genomic evidence of a similar

capability in the Gloeobacteraceae, inferred from the presence of ribose ABC importer genes in all examined genomes (Table S3). This finding suggests that the capacity to import and metabolize ribose emerged after the divergence from Anthocerotobacteraceae but before the diversification of Gloeobacteraceae (ca. <~1.4 billion years ago—>300 Ma; Rahmatpour et al. (2021)). Phylogenetic analyses (Figs. 3 and 4) indicate that these genes were likely horizontally transferred from at least two distinct donors: (i) either Acidobacteriota/Gemmatimonadota for the ribose-binding and permease components and (ii) an unresolved bacterial taxon for the ATP-binding protein. The noncontiguous arrangement of ribose importer genes in Gloeobacteraceae (Fig. 7) further supports multiple, independent acquisition events. Such a chimeric ABC transporter is unusual, as they are typically transferred as an intact unit, e.g. maltose transporters in Thermococcales and Thermotogales (Noll et al. 2008) and phosphate transporters in Firmicutes (Moreno-Letelier et al. 2011). The gene fragmentation pattern, however, is consistent with the extensive genome rearrangements typical of cyanobacteria, which often lack conserved synteny even in stable operons such as the *dcw* cluster (Léonard et al. 2022), secondary metabolite genes (Entfellner et al. 2022), and exopolysaccharide synthesis genes (Kim et al. 2024).

The functional implications of this ribose uptake system are particularly intriguing in light of the respiratory characteristics of *G. violaceus*. Koyama et al. (2008) reported exceptionally high respiratory activity—ca. 20-fold higher than *Synechocystis* PCC 6803—suggesting significant reliance on respiration. However, the

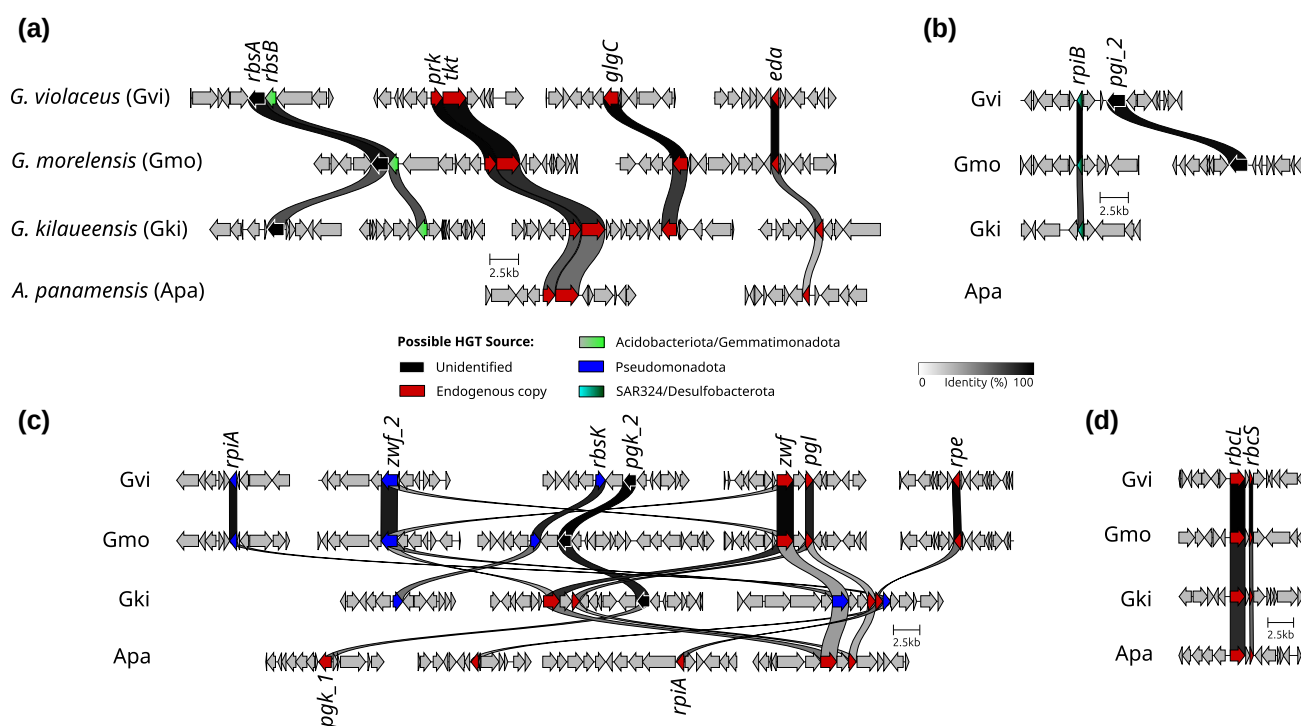


Fig. 7. Synteny of selected central carbon metabolism genes across the four complete genomes of Gloeobacteriales that are currently available. The illustrated regions contain both endogenous (red) or horizontally transferred genes (color corresponding to donor lineage), including a) *rbsA-rbsB*, b) *rpiB-pgi_2*, c) *rpiA*, *rbsK*, *zwf_2*, and d) the endogenous *rbcL-rbcS*. Only regions with two or more neighboring genes in at least one of the four genomes are shown. Gvi, *G. violaceus*; Gmo, *G. morelensis*; Gki, *G. kilaueensis*; Apa, *A. panamensis*.

substrate driving this oxygen consumption has remained enigmatic (Koyama et al. 2008). Our findings offer a possible explanation for this pattern. The presence of a complete ribose uptake and catabolic system within the Gloeobacteraceae suggests a capacity for photomixotrophy, in which ribose could serve as an auxiliary carbon source. Imported ribose may subsequently be oxidized via the PPP to generate reducing power for respiration, thereby sustaining energy metabolism when oxygenic photosynthetic electron flow is limited. This interpretation is consistent with the elevated oxygen uptake and may help explain the slow but sustained growth of *G. violaceus* in laboratory culture. Nevertheless, this inferred capacity of Gloeobacteraceae to import and metabolize ribose is currently supported only by comparative genomics and warrants future experimental validation.

The presence of both *rpiA* and *rpiB* genes in Gloeobacteraceae further highlights the importance of ribose metabolism. These enzymes catalyze the inter-conversion between ribose-5P (R5P) and ribulose-5P (Ru-5P), a key metabolic step linking ribose utilization with the central carbon network. Both proteins share largely the same function but are structurally distinct (Chen et al. 2020). The *rpiA* gene is nearly universal

among organisms, but *rpiB* occurs mainly in Actinobacteria and diverse eukaryotes (Tang et al. 2021; Santana-Molina et al. 2025). It was proposed that, while *rpiA* primarily catalyzes the conversion of Ru-5P to R5P, the reverse conversion is preferentially performed by *rpiB* under high R5P concentrations (Tang et al. 2021). Thus, the presence of *rpiB* in Gloeobacteraceae might be linked to elevated intracellular R5P levels resulting from active ribose import.

Collectively, the presence of all key genes required for ribose import (*rbsA-C*) and its assimilation (*rbsK*, *rpiA*, and *rpiB*) into the central carbon metabolism suggests that heterotrophic ribose utilization is an integral metabolic feature of Gloeobacteraceae. These genes were likely acquired through multiple independent HGT events rather than a single operon transfer. The capacity for ribose heterotrophy may contribute to the elevated respiratory activity reported for *G. violaceus*, providing a plausible physiological explanation for the acquisition and retention of ribose uptake and metabolism genes. These findings reveal that Gloeobacteraceae have evolved a more versatile and dynamic carbon metabolism through HGT than previously assumed, combining an ancient photosynthetic machinery with a more recently acquired heterotrophic capability.

HGT and Gene Loss Reshape Central Carbon Metabolism in Gloeobacterales

Several studies have reported HGT in the “core” central carbon metabolism of some cyanobacteria, including the acquisition of *rbcL-rbcS* from α -proteobacteria in marine Synechococcaceae (Cabello-Yeves et al. 2022), *fba1* from red algae in marine Synechococcaceae (Rogers et al. 2007; Godde et al. 2018), and *gap1* from *Anabaena* in *G. violaceus* (Figge et al. 1999). Our analyses further expand the list of potential HGTs in the central carbon metabolism of Gloeobacterales, identifying *tal*, *pgk*, *fba1*, *pgi*, *pfkA*, and *zwf* as additional candidates (Figs. 2 and 6). The lack of *edd* in Gloeobacterales verified findings by Evans et al. (2024), who suggested the absence of a complete ED pathway in cyanobacteria and plants.

Phylogenetic reconstructions (Fig. S2) confirm that *gap1* was horizontally transferred from other cyanobacteria in all four examined *Gloeobacter* species, consistent with Figge et al. (1999). Additional transfers from other cyanobacterial donors are clear for *tal_2*, *fba1*, *fba1_2*, *pfkA*, *pgi_1*, and *pgm1/2* (Fig. 6). These transfers are relatively “recent,” postdating the diversification of donor lineages. This finding underscores other cyanobacteria as major contributors of metabolic genes to Gloeobacterales. Members of Pseudomonadota also appear as frequent donors, contributing *rbsK*, *rpiA*, a second copy of *zwf* (in Gloeobacteraceae), and *pgi_2* (in Anthocerotibacteraceae). Conversely, even with extensive searches across more than 100,000 representative genomes in the Genome Taxonomy Database (GTDB), some HGT donors remained unidentifiable, possibly due to limited phylogenetic resolution, extinction of donor lineages, or incomplete taxonomic representation in the current database.

An alternative explanation for the inferred HGT events is vertical inheritance from the last common ancestor of cyanobacteria, followed by differential gene loss in Anthocerotibacteraceae and other cyanobacterial lineages, which may produce patchy phylogenetic patterns resembling HGT. However, in four of the six genes involved in ribose import and metabolism (*rbsA–C* and *rpiA*), other cyanobacterial lineages are present (Figs. 3 to 5; Table S3) and form distinct clades separate from Gloeobacteraceae, a pattern further supported by topology tests (Table S5). For the remaining two genes (*rbsK* and *rpiB*), homologs from other cyanobacteria are also detected within the corresponding HOGs (Table S3), but their sequences are highly divergent and were therefore not included in the phylogenetic analyses used to identify potential donor lineages. Thus, a scenario of vertical inheritance followed by differential gene loss appears less likely. Taken together, these results favor HGT as the most

likely explanation for the evolutionary history of these genes. Nevertheless, differential gene loss cannot be fully excluded and may contribute to the observed patchy distribution in some cases.

The ecological context of these transfers is consistent with the lifestyle of Gloeobacterales. Cyanobacteria often form dense biofilms in microbial mats (Bhaya et al. 2025; Bozan et al. 2025), which promote DNA exchanges through close physical contact with other bacteria. Indeed, Gloeobacterales microbial mats typically harbor various cyanobacterial taxa, including Nostocales, Leptolyngbyales, Synechococcales, and Chroococcales (Cuzman et al. 2010; Pessi et al. 2023), as well as abundant Pseudomonadota (Pessi et al. 2023). The dominance of these two phyla within Gloeobacterales-associated mats parallels their role as principal HGT donors. Other putative donors, such as Acidobacteriota and Gemmatimonadota, are also commonly found in microbial mats containing Gloeobacterales taxa (Zeng et al. 2020; Powell et al. 2024).

The absence of genes involved in glycogen metabolism in certain members of the Anthocerotibacteraceae (Fig. 6) represents an intriguing metabolic feature. Glycogen serves as the primary storage molecule for photosynthetically fixed carbon in cyanobacteria, thereby buffering against fluctuations in energy and carbon (Xu et al. 2013; Shinde et al. 2020). Interestingly, the lack of glycogen metabolism has been frequently reported in parasitic, symbiotic, or fastidious bacteria (Henrissat et al. 2002). Mutants of *glgC* and *glgA* in *Synechocystis* and *Synechococcus* exhibit severe defects in glycogen accumulation (Xu et al. 2013; Carrieri et al. 2017). The loss of these genes can impair growth recovery, alter sugar excretion, and increase stress susceptibility (Gründel et al. 2012; Xu et al. 2013; Shinde et al. 2020). Alternatively, members of Anthocerotibacteraceae lacking canonical *glg* genes may possess alternative carbon storage routes, such as those mediated by Rv3032-, GlgE-, or glucosylglycerol-based pathways observed in other bacteria (Chandra et al. 2011; Ortega-Martínez et al. 2025).

These findings demonstrate that HGT and gene loss have been major drivers of metabolic reorganization in Gloeobacterales. The repeated recruitment of cyanobacterial and proteobacterial genes suggests that Gloeobacterales did not simply conserve ancestral metabolic traits but instead actively restructured their physiology through foreign gene assimilation.

Evolutionary and Ecological Drivers of Metabolic Innovation in Gloeobacterales

The absence of thylakoid membranes inherently limits their photosynthetic efficiency (Rahmatpour et al.

2021), which explains both their slow growth and the difficulty of cultivating them in the laboratory. When the photosynthetic electron transport chain first emerged in the cytoplasmic membrane of ancestral organisms—possibly resembling the organization of modern *Gloeobacter*—this innovation would have conferred a significant evolutionary advantage (Cornet 2025). However, as thylakoid-bearing cyanobacteria evolved and diversified within the same microbial mats (Cuzman et al. 2010; Pessi et al. 2023), Gloeobacterales may have faced increasing ecological competition, and the lack of thylakoid membranes became a liability over time.

The evidence presented here suggests that Gloeobacteraceae subsequently evolved toward photo-mixotrophy, secondarily acquiring the capacity to exploit exogenous organic carbon sources such as ribose to compensate for their lack of competitiveness relative to other cyanobacteria. This transition was facilitated by multiple horizontal gene transfers, producing a patchwork of imported and rearranged genes even within core metabolic pathways. The genomic mosaicism observed in their central metabolism, spanning sugar importers, the CBB cycle enzymes, and the PPP components, underscores the evolutionary plasticity of these nonthylakoid-bearing cyanobacteria.

Conclusion

Our findings show that HGT has profoundly shaped the evolution of Gloeobacteraceae, enabling them to persist in low-light, nutrient-variable habitats by coupling inefficient photosynthesis with heterotrophic ribose utilization. HGT has also remodeled the Gloeobacterales' central carbon metabolism by combining endogenous genes with horizontally acquired genes, resulting in mosaic metabolic pathways. The microbial mats inhabited by Gloeobacterales likely serve as active gene-exchange networks, with other cyanobacteria and Pseudomonadota acting as the primary donors. This reinterpretation positions Gloeobacterales as metabolically versatile survivors whose genomes record a long history of innovation that reflects both ancient photosynthetic features and more recent adaptation through gene acquisition.

Materials and Methods

Taxa Selection and Phylogenetic Tree Reconstruction

We curated a dataset comprising 343 cyanobacteria and related nonphotosynthetic taxa encompassing the full diversity of cyanobacterial orders described by Strunecký et al. (2023). In brief, we downloaded the cyanobacterial genome list from GTDB release R220 (Parks et al. 2022) and selected one to two

representatives for each genus, prioritizing those with a high CheckM (Parks et al. 2015) completeness score (>80%) and a low contamination level (<10%). We included all available genomes from NCBI for the orders Gloeobacterales, Thermostichales, Pseudanabaenales, and Gloeomargaritales. The Pseudanabaenales dataset was further dereplicated using ToRQuEMaDA v0.2.1 (Léonard et al. 2021) with the following parameters: “ $P=200$, $l=200$, $t=0.80$, $j=12$, $a=100$, $r=10$, $d=1$, $\text{min}=1$, $\text{type}=\text{taxonomic}$, $\text{alg}=\text{Jl}$, $\text{egn}=\text{jellyfish}$.” Additionally, the taxon sampling was enriched with strains reported to exhibit heterotrophic capacity. In total, our data comprises 104 nonphotosynthetic taxa (including “*C. Margulisbacteria*,” “*C. Sericytochromatia*,” and Vampirovibrionophyceae) and 239 (photosynthetic) cyanobacteria (Table S1).

We used GToTree v1.8.8 (Lee 2019) and the cyanobacterial SCG pHMMs available in the tool to reconstruct the SCG supermatrix based on 251 individual alignments for the 343 selected taxa. The pipeline uses MUSCLE v5.1 (Edgar 2022) and TrimAl v1.4.rev15 (Capella-Gutiérrez et al. 2009) for sequence alignment and trimming, respectively. We slightly modified the parameters of GToTree by adding the following options: “ $-c\ 1$ and $-G\ 0$.” The supermatrix resulting from concatenated SCG alignments was used to infer a phylogenomic tree using IQ-TREE v2.3.6 (Minh et al. 2020) with the automatic best substitution model ($-m\ \text{MFP}$; Kalyanamoorthy et al. 2017) and 1,000 ultrafast bootstrap (Hoang et al. 2018) replicates. The final tree is shown in Fig. S1.

Hierarchical Orthologous Group Inference and Ortholog Search

The 343 selected genomes were first annotated using the Prodigal v2.6.3 (Hyatt et al. 2010) function implemented in the GEN-ERA toolbox v3.0.0 (Cornet et al. 2023). The orthologous gene inference was subsequently performed using OrthoFinder v2.5.4 (Emms and Kelly 2019) together with the previously reconstructed phylogeny (Fig. S1) as the species tree. The inferred HOGs were used to identify orthologs of genes of interest, including those related to heterotrophy and the central carbon metabolic pathways (Table S2). Reference genes were obtained from the literature (e.g. *Nostoc* fructose importer [Ekman et al. 2013] and *Synechocystis* fructose-1,6-biphosphatase [Caliebe et al. 2025]) or the KEGG database (Kanehisa et al. 2023). The presence or absence of each gene across HOGs is summarized in Table S3.

Ribose Structural Alignment Analyses

We used US-align v20241108 (Zhang et al. 2022) to assess the substrate specificity of the transporter gene identified in this study. We first downloaded the

predicted protein structure of *G. violaceus*' substrate-binding transporter (AlphaFold Protein DB: Q7NMF9) from the AlphaFold Protein Structure Database (Varadi et al. 2022). Reference proteins with known substrate specificities, e.g. for fructose, glucose, galactose, ribose, and maltodextrin (Table S4), were also retrieved from either the RSCB Protein Data Bank (Berman et al. 2000) or AlphaFold Protein Structure Database (Varadi et al. 2022). Each reference protein was aligned to the *G. violaceus* substrate-binding protein using US-align. Both the TM-score and RMSD were recorded. A perfectly aligned protein yields a TM-score of 1 (Zhang and Skolnick 2004), while an RMSD <2 Å indicates high structural similarity (Shamsian et al. 2024).

Identification of HGT Events

We first extracted the cyanobacterial sequences within the 34 identified HOGs of interest in this study (Table S3). Next, these protein sequences were aligned using MAFFT v7.5.2 (Katoh et al. 2002). The resulting alignments were converted into pHMMs using the hmmbuild function in HMMER v3.4 (<http://hmm.org/>). These pHMMs were then searched against the GTDB R220 bacteria representative genome dataset (107,235 species) using the hmmsearch function. To reduce redundancy, we filtered the hits using OmpaPa v0.252040 (<https://metacpan.org/dist/Bio-MUST-Apps-OmpaPa>), retaining approximately 704 to 5,206 sequences for downstream analyses. These sequences were subsequently realigned using MAFFT and trimmed with Clipkit v2.4.1 in the default (smart-gap) parameter (Steenwyk et al. 2020) before gene tree reconstruction using IQ-TREE v2.3.6. We set the IQ-TREE parameters to 1,000 ultrafast bootstrap replicates and the LG+G substitution model (for computational efficiency). The resulting trees were color-formatted using the format-tree.pl function in Bio-MUST-Core v0.252040 (<https://metacpan.org/dist/Bio-MUST-Core>) and visualized using iTOL v7 (Letunic and Bork 2024). We further performed topology tests in IQ-TREE by comparing the unconstrained maximum likelihood gene tree to a topology in which cyanobacterial sequences (including lineages such as Gloeobacterales) were constrained to be monophyletic. Constrained trees were inferred using the -g option under the same substitution model as the unconstrained trees. Candidate topologies were then evaluated using the -z option with the AU, SH, and KH tests, based on 10,000 resampling estimated log-likelihoods (RELL) bootstrap replicates without additional tree search (-n 0 -zb 10000 -au). Statistical significance was assessed at $P < 0.05$.

Because the initial HMM profiles were constructed from all cyanobacterial sequences within each HOG,

some putative HGTs into Gloeobacterales may have been masked. To mitigate this effect and improve donor identification, we repeated the above workflow with a modification: the pHMMs were constructed using only the Gloeobacterales sequences available in the HOGs. The remaining steps, such as GTDB search, OmpaPa filtering, and gene tree reconstruction, followed the same workflow described above.

Synteny Analyses

To investigate the genomic context of the horizontally transferred genes, we examined the synteny of those HGT-acquired genes using clinker v0.0.31 (Gilchrist and Chooi 2021). Four Gloeobacterales taxa with complete genome assemblies were selected for comparison, including *G. violaceus*, *G. morelensis*, *G. kilaeensis*, and *A. panamensis*. For each taxon, genomic regions encompassing the gene of interest (typically ± 5 kb flanking regions) were extracted using an in-house script (<https://doi.org/10.6084/m9.figshare.30422521>). These regions were compared and visualized using clinker, which identifies homologous genes based on translated sequence similarity and aligns genomic loci accordingly (Gilchrist and Chooi 2021).

Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online.

Acknowledgments

We thank Raphaël Léonard for the help in running the ToRQuEMaDA software and Louise Hambücken for assistance in identifying key primary literature relevant to this study.

Author Contributions

E.S.: Data curation; formal analysis; investigation; visualization; writing—original draft; writing—review & editing. D.B.: Conceptualization; funding acquisition; software; writing—review & editing. L.C.: Conceptualization; funding acquisition; supervision; writing—review & editing.

Funding

This study was funded by the Belgian National Fund for Scientific Research (F.R.S.-FNRS) through a research grant (PDR T.0018.24 OR-OX-PHOT-IN-CYN) awarded to D.B. L.C. is supported by a fellowship from F.R.S.-FNRS. Computational resources were provided by the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the F.R.S.-FNRS under Grant No. 2.5020.11 and by the Walloon Region.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability

All genomic data, generated data, and custom scripts used in this study are deposited in the FigShare repository (<https://doi.org/10.6084/m9.figshare.30422521>) under the CC BY 4.0 license. The GTDB R220 bacteria representative genome dataset can be downloaded from the GTDB website (<https://gtdb.ecogenomic.org/downloads>).

Literature Cited

- Berman HM, et al. The Protein Data Bank. *Nucleic Acids Res.* 2000;28:235–242. <https://doi.org/10.1093/nar/28.1.235>.
- Bernát G, et al. Unique properties vs. common themes: the atypical cyanobacterium *Gloeobacter violaceus* PCC 7421 is capable of state transitions and blue-light-induced fluorescence quenching. *Plant Cell Physiol.* 2012;53:528–542. <https://doi.org/10.1093/pcp/pcs009>.
- Bhaya D, Birzu G, Rocha EPC. Horizontal gene transfer and recombination in Cyanobacteriota. *Annu Rev Microbiol.* 2025;79:685–711. <https://doi.org/10.1146/annurev-micro-041522-100420>.
- Bowyer JR, Leegood RC. Photosynthesis. In: Dey PM, Harborne JB, editors. *Plant biochemistry*. Academic Press; 1997. p. 49–104.
- Bozan M, Berreth H, Lindberg P, Bühler K. Cyanobacterial biofilms: from natural systems to applications. *Trends Biotechnol.* 2025;43:318–332. <https://doi.org/10.1016/j.tibtech.2024.08.005>.
- Cabello-Yeves PJ, et al. α -cyanobacteria possessing form IA RuBisCO globally dominate aquatic habitats. *ISME J.* 2022;16:2421–2432. <https://doi.org/10.1038/s41396-022-01282-z>.
- Caliebe F, et al. Fructose-1,6-bisphosphatase is involved in heterotrophic growth and glycogen metabolism in cyanobacteria. *J Exp Bot.* 2025;76:6911–6929. <https://doi.org/10.1093/jxb/eraf347>.
- Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. Trimal: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics.* 2009;25:1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>.
- Cardona T, Murray JW, Rutherford AW. Origin and evolution of water oxidation before the last common ancestor of the cyanobacteria. *Mol Biol Evol.* 2015;32:1310–1328. <https://doi.org/10.1093/molbev/msv024>.
- Carrier D, et al. Transcriptome and proteome analysis of nitrogen starvation responses in *Synechocystis* 6803 Δ glgC. *Algal Res.* 2017;21:64–75. <https://doi.org/10.1016/j.algal.2016.11.003>.
- Chandra G, Chater KF, Bornemann S. Unexpected and widespread connections between bacterial glycogen and trehalose metabolism. *Microbiology.* 2011;157:1565–1572. <https://doi.org/10.1099/mic.0.044263-0>.
- Chen J, Wu H, Zhang W, Mu W. Ribose-5-phosphate isomerases: characteristics, structural features, and applications. *Appl Microbiol Biotechnol.* 2020;104:6429–6441. <https://doi.org/10.1007/s00253-020-10735-4>.
- Cornet L, et al. The GEN-ERA toolbox: unified and reproducible workflows for research in microbial genomics. *GigaScience.* 2023;12:giad022. <https://doi.org/10.1093/gigascience/giad022>.
- Cornet L. On the non-oxygenic origins of thylakoids. *Commun Biol.* 2025;8:1697. <https://doi.org/10.1038/s42003-025-09100-w>.
- Cuzman OA, et al. Biodiversity of phototrophic biofilms dwelling on monumental fountains. *Microb Ecol.* 2010;60:81–95. <https://doi.org/10.1007/s00248-010-9672-z>.
- Druce E, Maberly SC, Sánchez-Baracaldo P. Wide-ranging organic nitrogen diets of freshwater picocyanobacteria. *ISME J.* 2025;19:wrae236. <https://doi.org/10.1093/ismejo/wrae236>.
- Edgar RC. Muscle5: high-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *Nat Commun.* 2022;13:6968. <https://doi.org/10.1038/s41467-022-34630-w>.
- Ekman M, Picossi S, Campbell EL, Meeks JC, Flores E. A *Nostoc punctiforme* sugar transporter necessary to establish a cyanobacterium-plant symbiosis. *Plant Physiol.* 2013;161:1984–1992. <https://doi.org/10.1104/pp.112.213116>.
- Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 2019;20:238. <https://doi.org/10.1186/s13059-019-1832-y>.
- Entfellner E, et al. Toxic/bioactive peptide synthesis genes rearranged by insertion sequence elements among the bloom-forming cyanobacteria Planktothrix. *Front Microbiol.* 2022;13:901762. <https://doi.org/10.3389/fmicb.2022.901762>.
- Evans SE, et al. Plastid ancestors lacked a complete Entner-Doudoroff pathway, limiting plants to glycolysis and the pentose phosphate pathway. *Nat Commun.* 2024;15:1102. <https://doi.org/10.1038/s41467-024-45384-y>.
- Figge RM, et al. Glyceraldehyde-3-phosphate dehydrogenase gene diversity in eubacteria and eukaryotes: evidence for intra- and inter-kingdom gene transfer. *Mol Biol Evol.* 1999;16:429–440. <https://doi.org/10.1093/oxfordjournals.molbev.a026125>.
- Fournier GP, et al. The Archean origin of oxygenic photosynthesis and extant cyanobacterial lineages. *Proc R Soc B Biol Sci.* 2021;288:20210675. <https://doi.org/10.1098/rspb.2021.0675>.
- Gilchrist CLM, Chooi YH. Clinker & clustermap.js: automatic generation of gene cluster comparison figures. *Bioinformatics.* 2021;37:2473–2475. <https://doi.org/10.1093/bioinformatics/btab007>.
- Gisriel CJ, Bryant DA, Brudvig GW, Cardona T. Molecular diversity and evolution of far-red light-acclimated photosystem I. *Front Plant Sci.* 2023;14:1289199. <https://doi.org/10.3389/fpls.2023.1289199>.
- Godde JS, et al. Comparison of genomic islands in cyanobacteria: evidence of bacteriophage-mediated horizontal gene transfer from eukaryotes. *Microbiol Res.* 2018;211:31–46. <https://doi.org/10.1016/j.micres.2018.03.005>.
- Grettenberger CL. Novel *Gloeobacteriales* spp. from diverse environments across the globe. *mSphere.* 2021;6:e0006121. <https://doi.org/10.1128/mSphere.00061-21>.
- Gründel M, et al. Impaired glycogen synthesis causes metabolic overflow reactions and affects stress responses in the cyanobacterium *Synechocystis* sp. PCC 6803. *Microbiology.* 2012;158:3032–3043. <https://doi.org/10.1099/mic.0.062950-0>.
- Guglielmi G, Cohen-Bazire G, Bryant DA. The structure of *Gloeobacter violaceus* and its phycobilisomes. *Arch Microbiol.* 1981;129:181–189. <https://doi.org/10.1007/BF00425248>.
- Henrissat B, Deleury E, Coutinho PM. Glycogen metabolism loss: a common marker of parasitic behaviour in bacteria? *Trends Genet.* 2002;18:437–440. [https://doi.org/10.1016/S0168-9525\(02\)02734-8](https://doi.org/10.1016/S0168-9525(02)02734-8).
- Hoang DT, et al. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol.* 2018;35:518–522. <https://doi.org/10.1093/molbev/msx281>.

- Hyatt D, et al. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*. 2010;11:119. <https://doi.org/10.1186/1471-2105-11-119>.
- Johnson MP. Photosynthesis. *Essays Biochem*. 2016;60:255–273. <https://doi.org/10.1042/EBC20160016>.
- Kalyaanamoorthy S, et al. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. 2017;14:587–589. <https://doi.org/10.1038/nmeth.4285>.
- Kanehisa M, et al. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res*. 2023;51:D587–D592. <https://doi.org/10.1093/nar/gkac963>.
- Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*. 2002;30:3059–3066. <https://doi.org/10.1093/nar/gkf436>.
- Kim M, et al. Extensive genomic rearrangement of catalase-less cyanobloom-forming *Microcystis aeruginosa* in freshwater ecosystems. *J Microbiol*. 2024;62:933–950. <https://doi.org/10.1007/s12275-024-00172-7>.
- Koyama K, et al. Oxygen evolution in the thylakoid-lacking cyanobacterium *Gloeobacter violaceus* PCC 7421. *Biochim Biophys Acta Bioenerg*. 2008;1777:369–378. <https://doi.org/10.1016/j.bbabi.2008.01.009>.
- Lee MD. GToTree: a user-friendly workflow for phylogenomics. *Bioinformatics*. 2019;35:4162–4164. <https://doi.org/10.1093/bioinformatics/btz188>.
- Léonard RR, et al. ToRQuEMaDA: tool for retrieving queried Eubacteria, metadata and dereplicating assemblies. *PeerJ*. 2021;9:e11348. <https://doi.org/10.7717/peerj.11348>.
- Léonard RR, et al. Was the last bacterial common ancestor a monoderm after all? *Genes (Basel)*. 2022;13:376. <https://doi.org/10.3390/genes13020376>.
- Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res*. 2024;52:W78–W82. <https://doi.org/10.1093/nar/gkae268>.
- Lucius S, Hagemann M. The primary carbon metabolism in cyanobacteria and its regulation. *Front Plant Sci*. 2024;15:1417680. <https://doi.org/10.3389/fpls.2024.1417680>.
- Minh BQ, et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol*. 2020;37:1530–1534. <https://doi.org/10.1093/molbev/msaa015>.
- Moreno-Letelier A, et al. Parallel evolution and horizontal gene transfer of the pst operon in Firmicutes from oligotrophic environments. *Int J Evol Biol*. 2011;2011:781642. <https://doi.org/10.4061/2011/781642>.
- Muñoz-Marín MC, et al. Mixotrophy in marine picocyanobacteria: use of organic compounds by *Prochlorococcus* and *Synechococcus*. *ISME J*. 2020;14:1065–1073. <https://doi.org/10.1038/s41396-020-0603-9>.
- Muñoz-Marín MC, et al. Mixotrophy in cyanobacteria. *Curr Opin Microbiol*. 2024;78:102432. <https://doi.org/10.1016/j.mib.2024.102432>.
- Noll KM, et al. Evolution of mal ABC transporter operons in the Thermococcales and Thermotogales. *BMC Evol Biol*. 2008;8:7. <https://doi.org/10.1186/1471-2148-8-7>.
- Ortega-Martínez P, et al. Identification of an alternative glycogen synthesis pathway independent of glycogen synthases in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiol Biochem*. 2025;229:110519. <https://doi.org/10.1016/j.plaphy.2025.110519>.
- Parks DH, et al. Checkm: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*. 2015;25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
- Parks DH, et al. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and genome-based taxonomy. *Nucleic Acids Res*. 2022;50:D785–D794. <https://doi.org/10.1093/nar/gkab776>.
- Pessi IS, et al. Novel diversity of polar cyanobacteria revealed by genome-resolved metagenomics. *Microb Genomics*. 2023;9:001056. <https://doi.org/10.1099/mgen.0.001056>.
- Powell T, et al. Metagenome-assembled bacterial genomes from benthic microbial mats in ice-covered Lake Vanda, Antarctica. *Microbiol Resour Announc*. 2024;13:e0125023. <https://doi.org/10.1128/mra.01250-23>.
- Rahmatpour N, et al. A novel thylakoid-less isolate fills a billion-year gap in the evolution of cyanobacteria. *Curr Biol*. 2021;31:2857–2867.e4. <https://doi.org/10.1016/j.cub.2021.04.042>.
- Raven JA, Sánchez-Baracaldo P. Gloeobacter and the implications of a freshwater origin of cyanobacteria. *Phycologia*. 2021;60:402–418. <https://doi.org/10.1080/00318884.2021.1881729>.
- Rexroth S, et al. The plasma membrane of the cyanobacterium *Gloeobacter violaceus* contains segregated bioenergetic domains. *Plant Cell*. 2011;23:2379–2390. <https://doi.org/10.1105/tpc.111.085779>.
- Rippka R, et al. Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *Microbiology*. 1979;111:1–61. <https://doi.org/10.1099/00221287-111-1-1>.
- Rogers MB, Patron NJ, Keeling PJ. Horizontal transfer of a eukaryotic plastid-targeted protein gene to cyanobacteria. *BMC Biol*. 2007;5:26. <https://doi.org/10.1186/1741-7007-5-26>.
- Santana-Molina C, et al. Chimeric origins and dynamic evolution of central carbon metabolism in eukaryotes. *Nat Ecol Evol*. 2025;9:613–627. <https://doi.org/10.1038/s41559-025-02648-0>.
- Saw JH, Cardona T, Montejano G. Complete genome sequencing of a novel *Gloeobacter* species from a waterfall cave in Mexico. *Genome Biol Evol*. 2021;13:evab264. <https://doi.org/10.1093/gbe/evab264>.
- Shamsian S, et al. Benchmarking different docking protocols for predicting ligand binding poses. *BiolImpacts*. 2024;14:29955. <https://doi.org/10.34172/bi.2023.29955>.
- Shinde S, et al. Glycogen metabolism supports photosynthesis start through the oxidative pentose phosphate pathway in cyanobacteria. *Plant Physiol*. 2020;182:507–517. <https://doi.org/10.1104/pp.19.01184>.
- Smith AJ, London J, Stanier RY. Biochemical basis of obligate autotrophy in blue-green algae and thiobacilli. *J Bacteriol*. 1967;94:972–983. <https://doi.org/10.1128/jb.94.4.972-983.1967>.
- Stebegg R, Schmetterer G, Rempel A. Heterotrophy among cyanobacteria. *ACS Omega*. 2023;8:33098–33114. <https://doi.org/10.1021/acsomega.3c02205>.
- Steenwyk JL, et al. ClipKIT: a multiple sequence alignment trimming software for accurate phylogenomic inference. *PLoS Biol*. 2020;18:e3001007. <https://doi.org/10.1371/journal.pbio.3001007>.
- Strunecký O, et al. An updated classification of cyanobacterial orders and families based on phylogenomic analysis. *J Phycol*. 2023;59:12–51. <https://doi.org/10.1111/jpy.13304>.
- Tang H, et al. Engineering ribose-5-phosphate isomerase B into a promising sugar biocatalyst. *Appl Microbiol Biotechnol*. 2021;105:509–523. <https://doi.org/10.1007/s00253-020-11075-z>.
- Varadi M, et al. AlphaFold protein structure database: expanding structural coverage of protein-sequence space. *Nucleic Acids Res*. 2022;50:D439–D444. <https://doi.org/10.1093/nar/gkab1061>.

- White AW, Shilo M. Heterotrophic growth of the filamentous blue-green alga *Plectonema boryanum*. Arch Microbiol. 1975;102:123–127. <https://doi.org/10.1007/BF00428356>.
- Xu Y, et al. Altered carbohydrate metabolism in glycogen synthase mutants of *Synechococcus* sp. Strain PCC 7002: cell factories for soluble sugars. Metab Eng. 2013;16:56–67. <https://doi.org/10.1016/j.ymben.2012.12.002>.
- Zeng Y, et al. Potential rhodopsin- and bacteriochlorophyll-based dual phototrophy in a high Arctic glacier. mBio. 2020;11:e02641–e02620. <https://doi.org/10.1128/mBio.02641-20>.
- Zhang C, Shine M, Pyle AM, Zhang Y. US-align: universal structure alignments of proteins, nucleic acids, and macromolecular complexes. Nat Methods. 2022;19:1109–1115. <https://doi.org/10.1038/s41592-022-01585-1>.
- Zhang Y, Skolnick J. Scoring function for automated assessment of protein structure template quality. Proteins. 2004;57:702–710. <https://doi.org/10.1002/prot.20264>.

Associate editor: Jennifer Wisecaver