From Cytoplasmic Membrane to Thylakoids: Evolution of Membrane Biogenesis and Photosystem II assembly in early-diverging Cyanobacteria

Louise Hambücken[1], Denis Baurain[1], Luc Cornet[1*]

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

* Author to whom correspondence should be addressed.
luc.cornet@uliege.be

Abstract

Thylakoid membranes (TM) in cyanobacteria and chloroplasts host the light-dependent reactions of oxygenic photosynthesis, which involve a linear electron transfer (LET) chain composed of multi-subunit complexes, including notably Photosystem II (PSII). *Gloeobacterales*, the earliest-diverging cyanobacterial lineage, lack TM and perform photosynthesis within specialized regions of the cytoplasmic membrane (CM), thereby representing an ancestral state with respect to other cyanobacteria, all equipped with TM and known as Phycobacteria. The emergence of TM, which increased the membrane surface available for oxygenic photosynthesis, was a key innovation that likely contributed to the Great Oxidation Event. This evolutionary transition involved the formation of a distinct membrane compartment, followed by the relocation of LET components from the CM to TM. Here, we present a phylogenomic analysis identifying three candidate proteins associated with membrane trafficking that may contribute to TM biogenesis, including the SPFH family member Slr1106, which we show was acquired via lateral gene transfer. Moreover, evolutionary analysis of 36 PSII assembly factors indicates key modifications in late-stage PSII assembly, notably in manganese homeostasis. We further highlight structural changes in the early-acting YidC translocase that may have facilitated the relocation of LET components from the CM to TM. Altogether, our phylogenetic and functional prediction analyses of proteins involved in membrane dynamics and PSII assembly factors bring new insights into the molecular innovations that led to the emergence of TM.

Main

Introduction

Thylakoids are specialized membranes present in both cyanobacteria and chloroplasts that carry out the light-dependent reactions of oxygenic photosynthesis. They generate NADPH and ATP, for subsequent CO_2 fixation in the Calvin–Benson cycle, through a linear electron transport chain (LET) composed of multi-subunit complexes, notably Photosystem II (PSII) and Photosystem I (PSI), together with associated electron carriers. The emergence of this intracellular compartment provided cyanobacteria with an expanded membrane surface for light energy conversion, which is hypothesized to have triggered the Great Oxidation Event (GOE)[1] approximately 2.4 billion years ago, profoundly transforming Earth's biosphere[2].

Among extant cyanobacteria, members of the order *Gloeobacterales*, the earliest-diverging lineage, lack thylakoid membranes (TM)[3,4]. Instead, they carry out oxygenic photosynthesis within specialized regions of their cytoplasmic membrane (CM)[5]. The absence of TM in this basal group suggests that these structures arose after the divergence of *Gloeobacterales* but before the emergence of *Thermostichales*, the second earliest-diverging cyanobacterial lineage, which already possesses TM.

Consequently, *Gloeobacterales* retain the ancestral evolutionary state preceding TM formation and provide a unique model for exploring the transition from a LET located in the CM to a LET integrated within TM[6]. The proteins mediating the assembly of photosynthetic complexes within the TM may retain evolutionary signatures of this ancestral relocation, offering valuable insights into how cyanobacteria with TM, known as Phycobacteria, have emerged[6].

The mechanisms underlying TM biogenesis remain poorly understood. Four main hypotheses have been proposed: fusion between CM and TM, vesicular transport, direct lipid transfer via soluble lipid carriers, or self-assembly driven by lipid phase transitions[1,7]. VIPP1, an homolog of the eukaryotic ESCRT-III system implicated in membrane deformation and scission, was suggested to play an essential role in the biogenesis of TM, thereby fitting the first scenario[8–10]. This conclusion was supported by the fact that VIPP1 was only found in oxygenic photosynthetic organisms and that its presence is essential for thylakoid formation[10]. VIPP1 evolved from a duplication of the stress response protein PspA at the root of Phycobacteria and differs from PspA by an additional C-terminal extension[9]. However, a recent study showed that the *Gloeobacterales pspA* gene could complement a *vipp1-null* mutant in *Arabidopsis*[11]. Furthermore, proteomic data from TM biogenesis indicate that VIPP1 accumulates at later stages of TM formation[12]. Overall, these recent findings question the previously assumed central role of VIPP1 in the process.

Although the mechanisms underlying the assembly of most of the protein complexes forming the LET remain poorly studied, *de novo* assembly of PSII was shown to be a highly synchronized spatial and temporal process in *Synechocystis*, mediated by at least 36 proteins known as assembly factors[13–15]. During the early steps of PSII biogenesis, the precursor of D1 (pD1) is translocated across the CM by the Sec translocon and the YidC/Alb3 insertase, followed by chlorophyll incorporation and D1 maturation. The D1–D2 heterodimer then forms within a specialized region of the CM enriched in assembly factors, known as the PratA-defined membrane (PDM)[16]. During the later steps of PSII biogenesis, the inner antenna subunits CP47 and CP43 are sequentially incorporated, and the Mn_4CaO_5 cluster is assembled and activated[17]. These later steps likely occur within the developing TM.

Here, we present a phylogenomic analysis of thirteen proteins involved in membrane trafficking that identify one member of the SPFH family and one member of the DedA family as potential contributors to TM biogenesis, providing an alternative to the VIPP1-based model. We further report the first evolutionary analysis of the 36 PSII assembly factors known so far, suggesting important modifications in the later stages of PSII assembly, notably in manganese homeostasis, after the divergence of *Gloeobacterales*. In parallel, we pinpoint structural changes in the translocase YidC and the Sec machinery that may have facilitated the relocation of the LET components from CM to TM **(Figure 1)**.

Materials and methods

Genome selection and phylogenomic species tree

45,316 representative species genomes of Terrabacteria were retrieved from GTDB release r220[18] **(Table S1).** To limit computational burden during orthologous gene inference, a subset of 950 genomes were selected based on genome quality and phylogenetic position. Briefly, only the 43,827 genomes that passed the GUNC[19] contamination filter were retained. Within each GTDB order, genomes were then ranked by the difference between their CheckM2[20] completeness and contamination scores. For orders represented by a single genus, the two highest-quality genomes were kept; for all other orders, the best genome of each genus was selected, yielding 1,452 genomes in total, excluding the Cyanobacteriota phylum **(Table S2)**. Sampling within cyanobacteria was more exhaustive, resulting in 432 genomes, i.e., one or two representative genomes were chosen per genus, prioritizing those with >80% completeness and <10% contamination, as estimated by CheckM2[20]

(Table S3). To ensure coverage of early-diverging cyanobacterial lineages, all available genomes from the orders *Gloeobacterales*, *Thermostichales* and *Gloeomargaritales* were incorporated, whereas the more abundant *Pseudanabaenales* were dereplicated using ToRQuEMaDA v0.2.1, with default parameters, resulting in 343 genomes[21].

Protein-coding sequences were obtained directly from NCBI proteomes when available or predicted using Prodigal v2.6.3[22] **(Table S4)**. A phylogenomic tree based on ribosomal proteins was reconstructed using the ORPER workflow with default parameters from the GEN-ERA toolbox v1[23], with the exception that MAFFT v7.471 was used for the alignment of the ribosomal proteins[24] and that the ribosomal tree was inferred using IQ-TREE v2.4.0[25] with the best-fit substitution model (Q.YEAST+I+R10) and 1,000 ultrafast bootstrap replicates **(Figure S2)**. The resulting phylogeny was pruned using PARNAS v0.1.6[26] to retain a final set of 950 representative species **(Table S5)**, encompassing approximately 90% of the total phylogenetic diversity, as estimated by diversity score metrics, and comprising 30 cyanobacterial species **(Figure S1)**.

Orthologous group inference

The orthologous groups were inferred with OrthoFinder v2.5.5[27] from the 950 genomes selected using default parameters. An approach based on Hierarchical Orthologous Groups (HOGs) was first explored to account for multigene families stemming from duplications. However, this strategy, proved to be less accurate than the one based on standard Orthologous Groups (OGs), detailed below.

Identification of the orthologous groups of interest

The OGs corresponding to known assembly factors and membrane dynamics proteins were identified by performing BLASTp searches (v2.9.0+)[28] of every protein from the 950 proteomes against a database containing reference sequences of known PSII assembly factors and proteins involved in membrane dynamics (**Tables S6,7**). The e-value threshold was fixed at 10^{-3} and only the first hit was kept. The hits were then plotted based on their bitscore distributions for each OG **(Figures S5-S50)**. For each assembly factor and membrane dynamics protein, the OG with the highest median bitscore for cyanobacterial hits was selected **(Table S8)**. To check for optimal OG delineation, a reciprocal BLASTp of every reference sequence against a database containing the 950 proteomes was also performed. The best hits were then filtered based on their e-value, sequence length, number of copies per organism (first best hit per organism for the best score) and the fraction of the sequence covered by the alignment with the reference protein using OmpaPa (https://metacpan.org/dist/Bio-MUST-Apps-OmpaPa; D. Baurain). We then cross-validated the OGs derived from the filtered sequences against those obtained from the bitscore distribution plots.

Orthologous gene tree inference

The gene trees of each OG of interest were generated using IQ-TREE v3.0.0[29] with the best-fit substitution model and 1,000 ultrafast bootstrap replicates **(Table S8, Figures S96-S136)**. The sequences corresponding to the assembly factors and membrane dynamics proteins were then pinpointed in the gene trees based on the bitscores. When subsequent analyses were needed, either for taxonomic distribution determination or tree topology resolution, a HMM profile (pHMM) was built with the cyanobacterial sequences and searched against a database comprising the inferred proteomes of 343 Cyanobacteriota species (the dereplicated cyanobacterial dataset of 432 genomes, hereafter referred to as the cyanobacterial dataset) using HMMER v3.4[30]. The best hits were then filtered based on their e-value, sequence length, number of copies per organism (the three best hits per organism were kept) and the fraction of the sequence covered by the pHMM alignment (minimum 0.7) using OmpaPa. As the selected parameters varied depending on the protein under study, the filtering thresholds are detailed in figshare https://doi.org/10.6084/m9.figshare.31135354. Minimal lengths

ranged from 8 to 556 amino acids, with a median of 154, and minimal –log10 e-values ranged from 0 to 281, with a median of 33.

The selected sequences were then aligned using MAFFT v7.505[24]. Alignment sites were conserved if they had at most 70% gaps and the sequences with a minimum of 30% length (with respect to the longest unaligned sequence) were exported using ali2phylip.pl (https://metacpan.org/release/Bio-MUST-Core; D. Baurain). Finally, a gene tree was inferred from the filtered alignment using IQ-TREE v3.0.0[29] with the best-fit substitution model and 1,000 ultrafast bootstrap replicates **(Figures S51-S95)**. The presence or absence of the proteins was mapped onto a species tree of the cyanobacterial dataset[31] using iTOL[32].

Functional protein and 3D protein structure prediction

Functional protein domains were annotated with InterPro v5.48-83.0[33]. When the experimental structures were not available, the 3D protein structures were modeled using the AlphaFold 3 server[34]. The structures were then visualized and aligned using PyMOL v2.5.5[35].

Results

Membrane trafficking proteins

SPFH proteins (implicated in vesicular lipid transport), DedA family proteins (associated with lipid rearrangement), and dynamin-like proteins (involved in membrane fission and fusion) have been proposed by Siebenaller and Schneider (2023) as potential contributors to lipid trafficking in cyanobacteria[36]. To further explore their evolutionary history among Terrabacteria, we conducted a separate phylogenetic analysis for thirteen of these proteins (**Figure 2**). Their taxonomic distribution revealed that two are absent from *Gloeobacterales* but present in most Phycobacteria included in this study: Slr1106 and Slr0305. The prohibitin Slr1106 is a member of the SPFH family. Its phylogeny within Terrabacteria **(Figure 3)**, which also includes another SPFH protein, Slr1768, and suggests for both an origin through lateral gene transfer, possibly from Patescibacteria. The prohibitin Slr1106 is a member of the SPFH family. Its phylogeny within Terrabacteria (**Figure 3**) suggests for both an origin through lateral gene transfer, possibly from Patescibacteria. Slr1106 appears to have been acquired at the base of the Phycobacteria, whereas Slr1768 was acquired later, prior to the emergence of the *Pseudanabaenales* but after that of the *Thermostichales* (**Figure 2**). Our analyses showed that the tree of Slr1106 and Slr1768 also contains two groups of basal Cyanobacteria, including *Gloeobacterales*, which lack orthologues in other cyanobacterial orders (**Figure 3**). Moreover, structural predictions revealed an additional β-sheet positioned outside of the ring structure formed during oligomerization in the proteins of the basal group I (**Figure S137**). The DedA protein Slr0305 was not detected in *Gloeobacterales* but was present from the root of Phycobacteria (**Figure 4**). Another DedA protein, Slr0232, is for its part conserved across all cyanobacteria. However, its phylogeny exhibits a long branch at the base of Phycobacteria (**Figure 5**), albeit without discernible structural or sequence divergence.

Assembly factors of PSII

Late-stage assembly factors

The largest differences between *Gloeobacterales* and Phycobacteria were observed at the level of late-stage assembly factors (Figure 1). Hence, Psb27 and Psb34 were not identified in *Gloeobacterales* in our analyses (**Figure 2**). Psb34 is thought to bind CP47 and compete with HliA/B to enable the final stage of PSII assembly[37], while Psb27 binds both unassembled and assembled CP43 in the PSII complex[38]. Sll0606, a protein of unknown function but leading to a massive reduction of D1 and CP43 in PSII assembly upon deletion[39], was not detected either in *Gloeobacterales* (**Figure 2**). Domain

analysis of Sll0606 suggests a potential role in the transport of hydrophobic components, and distant homologues in *Gloeobacterales* have been proposed[39]. In this work, we also identified such a homolog in *Gloeobacterales*: BAC90294.1. Its sequence is not orthologous to the *Synechocystis* PCC6803 reference sequence, but InterPro domain analysis annotates three domains capable of binding hydrophobic components at the N-terminal region **(Figure S138).** Hlips family phylogenies are not well resolved in our analyses. Indeed, HliA, HliB, HliC, and HliD were scattered over several OGs **(Figure S9)**, and a combined phylogenetic tree did not reveal a clear separation among these four genes **(Figure S55)**.

Except for the accession GCA_038245785.1 (*Gloeobacterales* cyanobacterium), *Gloeobacterales* were also shown to lack Slr0144, a protein containing a 4-vinyl reductase (V4R) domain that binds small hydrophobic molecules and is inferred to interact with chlorophyll precursors in *Rhodobacter capsulatus*[40,41]. In *Synechocystis* sp. PCC6803, Slr0144 is encoded within the *slr0144–slr0152* operon[41], which is involved in PSII stabilization[42]. Slr0144 and Slr1471, the only two proteins of *Synechocystis* sp. PCC6803 containing a V4R domain, are the result of a gene duplication (**Figure 6**). Slr0144 and Slr1471 differ notably by Slr0144 having an N-terminal extension, along with changes in its C-terminal region (**Figure S139**).

Early-stage assembly factors

The early stages of PSII biogenesis do not show any differences in terms of presence or absence of assembly factors between *Gloeobacterales* and Phycobacteria. Indeed, whether considering the processing and incorporation of the D1 precursor (pD1), D2 folding, or the formation of the reaction center, all corresponding assembly factors are present in *Gloeobacterales* (**Figure 2**). However, a structural difference exists in the periplasmic domain of the translocase YidC and SecD between *Gloeobacterales* and Phycobacteria, with the latter harboring a specific features in its periplasmic domain.

YidC is part of the Oxa1/Alb3/YidC family of proteins, which are involved in the insertion and assembly of membrane proteins and have homologs across all three domains of life[43]. Notably, bacterial YidC shares homology with mitochondrial inner membrane Oxa1 and chloroplast thylakoid Alb3 proteins. YidC has been shown to participate in PSII biogenesis alongside the SecYEG machinery by mediating the membrane integration of the D1 precursor (pD1) in *Synechocystis* PCC6803, a process homologous to the one described in chloroplasts[44]. The periplasmic domain of YidC is only present in diderm bacteria, such as cyanobacteria, as its size is markedly reduced in other bacterial and chloroplastic homologs[43]. Interestingly, our results indicate that the periplasmic domain of cyanobacterial YidC is largely conserved across the cyanobacterial phylum, except in *Gloeobacterales.* Sequence alignment reveals well-conserved N-terminal and C-terminal regions; however, the periplasmic domain of *Gloeobacterales* YidC differs, with missing amino acids in the sequence (**Figure S140**). Besides, 3D structure prediction of the periplasmic domain highlights these differences, showing that Phycobacteria YidC possesses a P1 domain with structural changes and two additional helix loops compared to the *Gloeobacterales* homolog (**Figure 8**). In order to investigate how these changes affect YidC interaction with the Sec machinery, we performed the same analyses for SecD and SecF, the only two components of the Sec machinery that harbor periplasmic domains[61]. Our data showed that SecD contained one additional helix loop in *Thermostichales*, with this loop being larger in *Synechocystis* **(Figure S146)**. In contrast, no structural differences were observed in the periplasmic domain of *Thermostichales* SecF, whereas such differences were identified in *Synechocystis* **(Figure S147)**.

PratA, a periplasmic tetratricopeptide repeat (TPR) protein strongly associated with the PDM region[17], has been shown to facilitate efficient Mn delivery to pD1[45]. Our taxonomic distribution analyses showed the absence of PratA in most *Gloeobacterales*, at the exception of two accessions: GCA_018389385.1 (*Anthocerotibacter panamensis*) and GCA_038245785.1 (*Gloeobacterales* cyanobacterium) (**Figure 2**). CyanoP, the cyanobacterial homolog of the PsbP protein known in land

plants, is a Mn-binding protein that plays a role either in the binding association of CP47 and CP43 or in the photoactivation of the oxygen-evolving complex[46]. While our analyses show the presence of CyanoP in *Gloeobacterales*, its sequence has diverged substantially compared to Phycobacteria (**Figure S141**). APBS electrostatics calculations suggest that *Gloeobacterales* CyanoP either binds divalent ions differently compared to its Phycobacteria counterpart or does not bind ions at all (**Figure 7**). PAM71 is a Mn transporter involved in manganese homeostasis, sequestering excess cytoplasmic manganese into the periplasmic and lumenal compartments[47]. Its role in Mn uptake is considered to support PSII oxygen-evolving complex performance, suggesting a PratA-independent Mn delivery pathway[48]. Our results indicate its absence in *Gloeobacterales* (**Figure 2**). MncA plays an important role in manganese binding within the periplasmic space[49]. However, its OG could not be determined, as we could not extract an OG that contained most cyanobacterial sequences with high sequence similarity **(Figure S9)**.

Discussion

The emergence of TM is necessarily linked to a biological innovation involving lipid environments[1,7]. VIPP1 has long been considered a key factor in TM formation, a hypothesis supported by Tan *et al.* (2024), who demonstrated that VIPP1 originated from a duplication and subsequent C-terminal extension of the stress response protein PspA at the root of Phycobacteria[9]. However, the functional role of VIPP1 remains ambiguous. In the two model cyanobacteria *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002, VIPP1 does not appear to participate in TM biogenesis but rather in the assembly of photosynthetic complexes[50,51]. This observation is consistent with the proteomic analysis of TM by Huang *et al.* (2023), which detected VIPP1 only during late stages of TM development[12]. Moreover, a recent study identified a PspA homolog in *Gloeobacterales* capable of complementing a *vipp1-null* mutant in *Arabidopsis*[11]. Although its function is still undefined, these results collectively indicate that VIPP1 is unlikely to be directly involved in TM formation.

Our results show that four additional proteins, from prohibitin and DedA families, involved in membrane trafficking, might constitute an alternative to the VIPP1-based model for the emergence of TM. Slr1106 is absent in *Gloeobacterales* but present in Phycobacteria, including *Thermostichales*, although its distribution within the latter group is limited. Slr1106 was acquired through lateral gene transfer (LGT) to the ancestor of Phycobacteria, but the donor could not be unambiguously determined (**Figure S134**), possibly due to the ancien nature of the transfer. Nevertheless, some Patescibacteria sequences appear to be closely-related to Slr1106 in the phylogeny **(Figure 3)**. Slr1106 belongs to the prohibitin family, within the SPFH protein superfamily, typically associated with the inner mitochondrial membrane or the nucleus in eukaryotic cells, where they contribute to structural stabilization, apoptosis regulation, and cell signaling[54]. Members of the SPFH family are characterized by four α-helices and a twisted β-sheet structure, and they assemble into large oligomeric complexes[55]. Slr1106 is localized in both CM and TM[37], consistent with a potential role in membrane organization and the formation of lipid microdomains. Besides Slr1106, the OG corresponding to the prohibitin Slr1106 includes proteins that are restricted to basal cyanobacteria, including *Gloeobacterales*, suggesting that these ancestral variants may have participated in the formation of primitive photosynthetic regions in CM, as observed in *Gloeobacterales*[57]. Another member of this family, Slr1768, was also identified within the same orthologous group and has already been shown to play a role in TM maintenance under high light conditions[56]. Slr1768 appears to be absent in *Gloeobacterales*, *Thermostichales* and *Gloeomargaritales* and only partially conserved across other cyanobacterial lineages **(Figure 3)**. In addition to these prohibitins, our analyses identified two DedA proteins that might have played a role in TM emergence: Slr0305 **(Figure 4)**, acquired at the root of Phycobacteria, and Slr0232 **(Figure 5)**, which exhibits an unusually long phycobacterial branch in its phylogeny. The DedA family contains three members in *Synechocystis* (Slr0232, Slr0305, and Slr0509) and is known to be involved in membrane trafficking in eukaryotic cells, with Slr0305 most closely related to the putative lipid scramblase of the yeast Tvp38[58]. This would therefore suggest a vesicular TM biogenesis system, in contrast to the CM–TM fusion model proposed with the ESCRT-III homolog VIPP1.

Cornet (2025) hypothesized that the complex stepwise relocation of the LET components from CM to TM involved intermediate evolutionary stages between *Gloeobacterales* and Phycobacteria[6]. One possible scenario is that the earliest of these stages, prior to the relocation of the LET chain to TM, acquired one or several of the three genes mentionned above (Slr1106, Slr0305, Slr0232). For example, the presence of Slr0305 at the base of Phycobacteria, despite its absence in *Gloeomargaritales*, may have contributed to the emergence of TM as new cellular compartment as Slr0305 shares high sequence similarity with a protein predicted from the plastid of *Arabidopsis thaliana* (AT5G19070.1) that is annotated as a SNARE-associated Golgi protein[58]. Similarly, another possible scenario is that this ancestral stage was tightly associated with a Patescibacterium, subsequently leading to the lateral gene transfer of Slr1106 and to the emergence of TM, although this would imply secondary gene loss in *Thermostichales*. This hypothesis is plausible given that Patescibacteria, which are characterized by reduced genomes and limited metabolic capabilities, are generally regarded as parasitic or symbiotic bacteria[52,53].

The differences observed in PSII biogenesis are consistent with the membrane architecture of *Gloeobacterales*. On one hand, the absence (or distant *Gloeobacterales* homologs) of assembly factors related to the late stages of PSII biogenesis, namely Psb27, Psb34, Sll0606, may indicate an increased reliance on late-stage assembly steps during PSII repair in TM[59]. On the other hand, the absence of TM facilitates the biogenesis of the LET components, as it occurs within a single cellular compartment, the CM. This is particularly clear in the case of Mn utilization. Mn is stored in the periplasmic space[60]. In *Gloeobacterales*, the Mn cluster of PSII is located at the periplasmic side, and therefore does not require Mn transport across membranes for PSII assembly, unlike in Phycobacteria. The absence of PratA and PAM71 can thus be explained by the lack of TM. Interestingly, the mutation of PAM71 in *Synechocystis* sp. PCC6803 leads to a massive up-regulation of the Mn-binding CyanoP, which suggests a role of CyanoP both in PSII assembly and Mn homeostasis[47]. Our results show that *Gloeobacterales* CyanoP exhibits a less negatively charged surface compared with its orthologs in Phycobacteria **(Figure 7)**, which would suggest a change in Mn homeostasis from *Gloeobacterales* to Phycobacteria. Nevertheless, the Mn-binding protein Mn cupin A (MncA), which is associated with Mn storage in the periplasmic space[49], might also be expected in *Gloeobacterales*, although we were unable to confirm its occurrence. Little is known about the function of the V4R domain of Slr0144 and Slr1471 in *Synechocystis*. Their absence in the majority of *Gloeobacterales*, along with duplication events at the base of Phycobacteria and within Thermostichales **(Figure 6)**, underscores unresolved questions regarding their roles in PSII assembly as part of the photosynthetically-related *slr0144–slr0152* operon.

The presence of all assembly factors involved in early stages of PSII biogenesis is somewhat unexpected, as one might have anticipated the emergence of new assembly factors at the root of Phycobacteria. However, the presence of additional helix loops in the periplasmic domains of the phycobacterial YidC and SecD from *Thermostichales* might explain the differential localization of PSII between *Gloeobacterales* and Phycobacteria **(Figure 8).** In *Escherichia coli*, YidC directly interacts with Sec components, particularly SecDF, which contain periplasmic domains[61,62]. Interestingly, our data indicate that SecD of Phycobacteria harbors two additional helix–loop insertions in its periplasmic domain **(Figure S146A,B).** In contrast, phycobacterial SecF differed only by its cytoplasmic domain **(Figure S147A)** in *Thermostichales*, whereas an additional helix–loop region was observed in *Synechocystis* sp. Nevertheless, interaction with YidC cannot be resolved, as the periplasmic domains of both YidC and SecD differ from those characterized in *E. coli* **(Figure S146C)**. Moreover, SecD localization within TM has not yet been established. Deletion of 90% of the periplasmic domain of *E. coli* YidC does not impair inner membrane protein biogenesis or cell viability[62], and functional complementation experiments indicate that this domain is not strictly required for protein translocation, as Oxa1, despite lacking this domain, can complement YidC function in *E. coli*[63]. Altogether, these observations suggest that the periplasmic domain of YidC has a yet-undefined role in the cell, despite its conservation across diderm bacteria[64]. Cyanobacteria possess only a single set of Sec genes that operates in both the cytoplasmic membrane (CM) and the thylakoid membrane (TM), raising questions about how translocated proteins are selectively targeted to these two compartments[65,66]. In contrast, the YidC translocase is specifically localized to the PDM and TM.

Since the lumenal space of Phycobacteria is homologous to the periplasmic space of *Gloeobacterales*[67], the additional helix loops and structural modifications observed in the periplasmic domain of phycobacterial YidC could potentially underlie changes in its functional properties and contribute to YidC's affinity for TM.

Our results on PSII, the photosynthetic complex whose biogenesis is the best characterized , highlight the value of evolutionary analyses of assembly factors for understanding the emergence of TM and the relocation of photosynthetic complexes from the ancestral CM-localized state—represented today by *Gloeobacterales*[6]—to modern TM. Extending this approach to other complexes (PSI, the cytochrome *b6f*, ATP synthase, NDH-1) would provide a framework for elucidating how thylakoid-associated photosynthetic complexes are assembled and how this assembly process evolved throughout the radiation of Cyanobacteria. Nevertheless, the biogenesis of these other complexes remains poorly characterized, with only a limited number of assembly factors identified to date[6]. This underscores the need to search for additional genes involved in the biogenesis of photosynthetic complexes if one wants to fully reconstruct the evolutionary emergence and diversification of the cyanobacterial photosynthetic apparatus.

Conclusion

Our comparative analyses of membrane trafficking proteins and PSII assembly factors highlight key molecular innovations associated with the emergence of TM in Phycobacteria. The acquisition of prohibitins and DedA-family proteins, such as Slr1106, via lateral gene transfer (possibly from Patescibacteria), and DedA-family proteins like Slr0305, probably facilitated the organization of lipid microdomains and the development of membrane compartments, providing a structural basis for thylakoids. In contrast, *Gloeobacterales* retain a simplified system, lacking V4R-domain proteins, late-stage PSII assembly factors (Psb27, Psb34, Sll0606), and manganese transporters, consistent with PSII biogenesis occurring entirely within the cytoplasmic membrane. Early-stage assembly factors are surprisingly conserved, but structural modifications in periplasmic domains of YidC and SecD in Phycobacteria may contribute to selective targeting of photosynthetic complexes to TM. Overall, thylakoid evolution may have required gene acquisitions for compartment formation and structural innovations enabling PSII migration.

Contributions

LH performed all analyses and drew figures. LH, DB, and LC wrote the manuscript. All authors read and approved the final version of the manuscript.

Acknowledgments

We thank Edi Sudianto for kindly providing an additional dataset of Cyanobacteriota genomes and the corresponding species tree. This work was supported by a research grant (PDR T.0018.24 OR-OX-PHOT-IN-CYN) from the Belgian National Fund for Scientific Research (F.R.S.-FNRS) to DB. LC is supported by a mandate from the Belgian National Fund for Scientific Research (F.R.S.-FNRS). LH is a FRIA grantee of the Fonds de la Recherche Scientifique – 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 competing interest.

Data Availability

Multiple sequence alignments, phylogenetic trees, 3D structure predictions, and analysis scripts are available on Figshare (DOI: doi.org/10.6084/m9.figshare.31135354). Supplementary figures (including phylogenetic trees of the genes of interest, the ribosomal species tree, and OG selection), tables, and additional analyses are provided in the Supplementary Information.

References

1. Guéguen, N. & Maréchal, E. Origin of cyanobacterial thylakoids via a non-vesicular glycolipid phase transition and their impact on the Great Oxygenation Event. *Journal of Experimental Botany* **73**, 2721–2734 (2022).
2. Bekker, A. *et al.* Dating the rise of atmospheric oxygen. *Nature* **427**, 117–120 (2004).
3. Rahmatpour, N. *et al.* A novel thylakoid-less isolate fills a billion-year gap in the evolution of Cyanobacteria. *Current Biology* **31**, 2857-2867.e4 (2021).
4. Rippka, R., Waterbury, J. & Cohen-Bazire, G. A cyanobacterium which lacks thylakoids. *Arch. Microbiol.* **100**, 419–436 (1974).
5. Rexroth, S. *et al.* Functional Characterization of the Small Regulatory Subunit PetP from the Cytochrome $b_6 f$ Complex in *Thermosynechococcus elongatus*. *Plant Cell* **26**, 3435–3448 (2014).
6. Luc Cornet. On the Non-Oxygenic Origins of Thylakoids. *Communications Biology (in press)* (2025).
7. Rast, A., Heinz, S. & Nickelsen, J. Biogenesis of thylakoid membranes. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1847**, 821–830 (2015).
8. Heidrich, J., Thurotte, A. & Schneider, D. Specific interaction of IM30/Vipp1 with cyanobacterial and chloroplast membranes results in membrane remodeling and eventually in membrane fusion. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1859**, 537–549 (2017).
9. Tan, S. Exploring the Origins and Evolution of Oxygenic and Anoxygenic Photosynthesis in Deeply Branched Cyanobacteriota. *Molecular Biology and Evolution* **41**, (2024).
10. Westphal, S., Heins, L., Soll, J. & Vothknecht, U. C. *Vipp1* deletion mutant of *Synechocystis* : A connection between bacterial phage shock and thylakoid biogenesis? *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4243–4248 (2001).
11. Ma, L. *et al.* VESICLE-INDUCING PROTEIN IN PLASTIDS 1 from thylakoid-lacking Gloeobacter promotes thylakoid formation in Arabidopsis. *Plant Physiology* **199**, kiaf359 (2025).

12. Huang, F. *et al.* Dynamic Changes in the Thylakoid Proteome of Cyanobacteria during Light-Regulated Thylakoid Membrane Development. *Plants* **12**, 3967 (2023).
13. Heinz, S., Liauw, P., Nickelsen, J. & Nowaczyk, M. Analysis of photosystem II biogenesis in cyanobacteria. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1857**, 274–287 (2016).
14. Rast, A. *et al.* Biogenic regions of cyanobacterial thylakoids form contact sites with the plasma membrane. *Nat. Plants* **5**, 436–446 (2019).
15. Van De Meene, A. M. L., Hohmann-Marriott, M. F., Vermaas, W. F. J. & Roberson, R. W. The three-dimensional structure of the cyanobacterium Synechocystis sp. PCC 6803. *Arch Microbiol* **184**, 259–270 (2006).
16. Stengel, A. *et al.* Initial Steps of Photosystem II de Novo Assembly and Preloading with Manganese Take Place in Biogenesis Centers in *Synechocystis*. *Plant Cell* **24**, 660–675 (2012).
17. Rengstl, B., Oster, U., Stengel, A. & Nickelsen, J. An Intermediate Membrane Subfraction in Cyanobacteria Is Involved in an Assembly Network for Photosystem II Biogenesis. *Journal of Biological Chemistry* **286**, 21944–21951 (2011).
18. Parks, D. H. *et al.* GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Research* **50**, D785–D794 (2022).
19. Orakov, A. *et al.* GUNC: detection of chimerism and contamination in prokaryotic genomes. *Genome Biol* **22**, 178 (2021).
20. Chklovski, A., Parks, D. H., Woodcroft, B. J. & Tyson, G. W. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. *Nature Methods* **20**, 1203–1212 (2023).
21. Léonard, R. R., Leleu, M., Van Vlierberghe, M., Kerff, F. & Baurain, D. ToRQuEMaDA: Tool for Retrieving Queried Eubacteria, Metadata and Dereplicating Assemblies. Preprint at https://doi.org/10.1101/2020.11.15.363259 (2020).
22. Hyatt, D. *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).

23. Cornet, L. *et al.* The GEN-ERA toolbox: unified and reproducible workflows for research in microbial genomics. *GigaScience* **12**, (2023).
24. Katoh, K. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* **30**, 3059–3066 (2002).
25. Minh, B. Q. *et al.* IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution* **37**, 1530–1534 (2020).
26. Markin, A. *et al.* PARNAS: Objectively Selecting the Most Representative Taxa on a Phylogeny. *Systematic Biology* **72**, 1052–1063 (2023).
27. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* **20**, 238 (2019).
28. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).
29. Wong, T. K. F. *et al.* IQ-TREE 3: Phylogenomic Inference Software using Complex Evolutionary Models. *EcoEvoRxiv* https://doi.org/10.32942/X2P62N (2025) doi:https://doi.org/10.32942/X2P62N.
30. Finn, R. D., Clements, J. & Eddy, S. R. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Research* **39**, W29–W37 (2011).
31. Edi Seduanto, Denis Baurain, & Luc Cornet. Horizontal Gene Transfers Underpin Ribose Heterotrophy and Central Carbon Metabolism Remodeling in Gloeobacteraceae. https://doi.org/10.1101/2025.11.10.686926 (2025) doi:https://doi.org/10.1101/2025.11.10.686926.
32. Letunic, I. & Bork, P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research* **52**, W78–W82 (2024).
33. Jones, P. *et al.* InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).
34. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
35. Schrödinger, LLC. The PyMOL molecular graphics system, version 1.8. (2015).
36. Siebenaller, C. & Schneider, D. Cyanobacterial membrane dynamics in the light of eukaryotic principles. *Bioscience Reports* **43**, BSR20221269 (2023).

37. Rahimzadeh-Karvansara, P., Pascual-Aznar, G., Bečková, M. & Komenda, J. Psb34 protein modulates binding of high-light-inducible proteins to CP47-containing photosystem II assembly intermediates in the cyanobacterium Synechocystis sp. PCC 6803. *Photosynth Res* **152**, 333–346 (2022).
38. Komenda, J. *et al.* The Psb27 Assembly Factor Binds to the CP43 Complex of Photosystem II in the Cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiology* **158**, 476–486 (2012).
39. Zhang, S., Frankel, L. K. & Bricker, T. M. The Sll0606 Protein Is Required for Photosystem II Assembly/Stability in the Cyanobacterium Synechocystis sp. PCC 6803. *Journal of Biological Chemistry* **285**, 32047–32054 (2010).
40. Anantharaman, V., Koonin, E. V. & Aravind, L. Regulatory potential, phyletic distribution and evolution of ancient, intracellular small-molecule-binding domains11Edited by F. Cohen. *Journal of Molecular Biology* **307**, 1271–1292 (2001).
41. Singh, A. K., Li, H. & Sherman, L. A. Microarray analysis and redox control of gene expression in the cyanobacterium *Synechocystis* sp. PCC 6803. *Physiologia Plantarum* **120**, 27–35 (2004).
42. Yang, H. *et al.* Slr0151 in *Synechocystis* sp. PCC 6803 is required for efficient repair of photosystem II under high-light condition. *JIPB* **56**, 1136–1150 (2014).
43. Hennon, S. W., Soman, R., Zhu, L. & Dalbey, R. E. YidC/Alb3/Oxa1 Family of Insertases. *Journal of Biological Chemistry* **290**, 14866–14874 (2015).
44. Gathmann, Sven, Eva Rupprecht, Uwe Kahmann, & Dirk Schneider. A Conserved Structure and Function of the YidC Homologous Protein Slr1471 from Synechocystis sp. PCC 6803. *Journal of microbiology and biotechnology* **18**, 1090–1094 (2008).
45. Klinkert, B. *et al.* PratA, a Periplasmic Tetratricopeptide Repeat Protein Involved in Biogenesis of Photosystem II in Synechocystis sp. PCC 6803. *Journal of Biological Chemistry* **279**, 44639–44644 (2004).
46. Cormann, K. U., Bartsch, M., RÃ¶gner, M. & Nowaczyk, M. M. Localization of the CyanoP binding site on photosystem II by surface plasmon resonance spectroscopy. *Front. Plant Sci.* **5**, (2014).

47. Gandini, C., Schmidt, S. B., Husted, S., Schneider, A. & Leister, D. The transporter Syn PAM 71 is located in the plasma membrane and thylakoids, and mediates manganese tolerance in *Synechocystis* PCC 6803. *New Phytologist* **215**, 256–268 (2017).
48. Eisenhut, M. Manganese Homeostasis in Cyanobacteria. *Plants* **9**, 18 (2019).
49. Tottey, S. *et al.* Protein-folding location can regulate manganese-binding versus copper- or zinc-binding. *Nature* **455**, 1138–1142 (2008).
50. Gao, H. & Xu, X. Depletion of Vipp1 in *Synechocystis* sp. PCC 6803 affects photosynthetic activity before the loss of thylakoid membranes. *FEMS Microbiology Letters* **292**, 63–70 (2009).
51. Zhang, S., Shen, G., Li, Z., Golbeck, J. H. & Bryant, D. A. Vipp1 Is Essential for the Biogenesis of Photosystem I but Not Thylakoid Membranes in Synechococcus sp. PCC 7002. *Journal of Biological Chemistry* **289**, 15904–15914 (2014).
52. Ji, Y. *et al.* Widespread but Poorly Understood Bacteria: Candidate Phyla Radiation. *Microorganisms* **10**, 2232 (2022).
53. Kuroda, K. *et al.* Symbiosis between *Candidatus* Patescibacteria and Archaea Discovered in Wastewater-Treating Bioreactors. *mBio* **13**, e01711-22 (2022).
54. Browman, D. T., Hoegg, M. B. & Robbins, S. M. The SPFH domain-containing proteins: more than lipid raft markers. *Trends in Cell Biology* **17**, 394–402 (2007).
55. Takekawa, N. *et al.* Structure of *Vibrio* FliL, a New Stomatin-like Protein That Assists the Bacterial Flagellar Motor Function. *mBio* **10**, e00292-19 (2019).
56. Bryan, S. J. *et al.* Loss of the SPHF Homologue Slr1768 Leads to a Catastrophic Failure in the Maintenance of Thylakoid Membranes in Synechocystis sp. PCC 6803. *PLoS ONE* **6**, e19625 (2011).
57. Rexroth, S. *et al.* The Plasma Membrane of the Cyanobacterium *Gloeobacter violaceus* Contains Segregated Bioenergetic Domains. *The Plant Cell* **23**, 2379–2390 (2011).
58. Keller, R. & Schneider, D. Homologs of the yeast Tvp38 vesicle-associated protein are conserved in chloroplasts and cyanobacteria. *Front. Plant Sci.* **4**, (2013).
59. Komenda, J., Sobotka, R. & Nixon, P. J. The biogenesis and maintenance of PSII: Recent advances and current challenges. *The Plant Cell* **36**, 3997–4013 (2024).

60. Keren, N., Kidd, M. J., Penner-Hahn, J. E. & Pakrasi, H. B. A Light-Dependent Mechanism for Massive Accumulation of Manganese in the Photosynthetic Bacterium *Synechocystis* sp. PCC 6803. *Biochemistry* **41**, 15085–15092 (2002).
61. Nouwen, N. & Driessen, A. J. M. SecDFyajC forms a heterotetrameric complex with YidC. *Molecular Microbiology* **44**, 1397–1405 (2002).
62. Xie, K., Kiefer, D., Nagler, G., Dalbey, R. E. & Kuhn, A. Different Regions of the Nonconserved Large Periplasmic Domain of *Escherichia coli* YidC Are Involved in the SecF Interaction and Membrane Insertase Activity. *Biochemistry* **45**, 13401–13408 (2006).
63. Van Bloois, E. *et al.* The Sec-independent Function of Escherichia coli YidC Is Evolutionary-conserved and Essential. *Journal of Biological Chemistry* **280**, 12996–13003 (2005).
64. Oliver, D. C. & Paetzel, M. Crystal Structure of the Major Periplasmic Domain of the Bacterial Membrane Protein Assembly Facilitator YidC. *Journal of Biological Chemistry* **283**, 5208–5216 (2008).
65. Masato Nakai, Daigo Sugita, Tatsuo Omata, & Toshiya Endo. Sec-Y protein is localized in both the cytoplasmic and thylakoid membranes in the cyanobacterium Synchococcus PCC9742. *Biochemical and Biophysical Research Communications* **193**, 228–234 (1993).
66. Srivastava, R., Pisareva, T. & Norling, B. Proteomic studies of the thylakoid membrane of ***Synechocystis*** sp. PCC 6803. *Proteomics* **5**, 4905–4916 (2005).
67. Herrmann, J. M., Kauff, F. & Neuhaus, H. E. Thiol oxidation in bacteria, mitochondria and chloroplasts: Common principles but three unrelated machineries? *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1793**, 71–77 (2009).

Table and Figures

Figure 1. Overview of the results. Phylogenetic analyses of 36 assembly factors if PSII and 13 proteins involved in membranes dynamics summarized in (36) were performed to investigate the taxonomic distributions and structural differences between the thylakoid-less *Gloeobacterales* and the Phycobacteria.

Figure 2. Taxonomic distribution of the PSII assembly factors and proteins involved in membrane dynamics in *Synechocystis*. A black dot indicates the presence of a protein, whereas a white dot indicates its absence. Highlighted proteins are discussed in the manuscript. Proteins marked with an asterisk indicate cases in which the *Gloeobacterales* homolog may not share the same evolutionary origin as the corresponding phycobacterial protein.

Figure 3. Phylogenetic tree of the cyanobacterial orthologous sequences of Slr1106 and Slr1768 (SPFH family) in the Terrabacteria dataset. The phylogenetic tree was constructed from 273 sequences across 510 unambiguously aligned positions using IQ-TREE (Q.PLANT+F+I+R6 model) with 1000 ultrafast bootstrap replicates. Bootstrap values ≥70% are shown. The tree annotation corresponds to the region showing the highest similarity to the reference *Synechocystis* sequences, as determined by BLAST.

Figure 4. Phylogenetic tree of the cyanobacterial orthologous sequences of Slr0305 (DedA family) in the Terrabacteria dataset. (A) The phylogenetic tree was constructed from 347 sequences across 504 unambiguously aligned positions using IQ-TREE (LG+F+R7 model) with 1000 ultrafast bootstrap replicates. Bootstrap values ≥70% are shown. The region labeled Slr0305 corresponds to the region showing the highest similarity to the reference *Synechocystis* sequence, as identified by BLAST. (B) The region labeled as Slr0305 in (A) was zoomed.

Figure 5. Phylogenetic tree of the cyanobacterial orthologous sequences of Slr0232 (DedA family) in the Terrabacteria dataset. (A) The phylogenetic tree was constructed from 987 sequences across 522 unambiguously aligned positions using IQ-TREE (LG+F+R10 model) with 1000 ultrafast bootstrap replicates. Bootstrap values ≥70% are shown. The region labeled Slr0232 corresponds to the region showing the highest similarity to the reference *Synechocystis* sequence, as identified by BLAST. (B) The region labeled as Slr0232 in (A) was zoomed. The long branch at the basis of Phycobacteria is indicated by an asterisk.

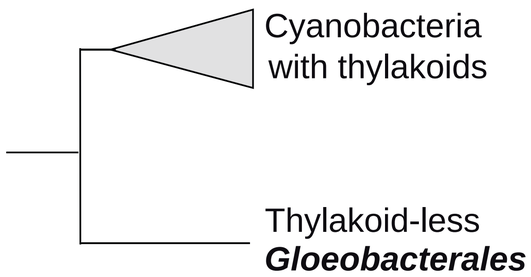
Figure 6. Phylogenetic tree of the cyanobacterial orthologues sequences of Slr0144 and Slr0147 in the cyanobacterial dataset. The phylogenetic tree was constructed from 311 sequences across 234 unambiguously aligned positions using IQ-TREE (Q.PFAM+R6) with 1000 ultrafast bootstrap replicates. The tree annotation corresponds to the region showing the highest similarity to the reference *Synechocystis* sequences, as determined by BLAST. Bootstrap values ≥70% are shown.

Figure 7. Prediction of the 3D structure with APBS electrostatics estimation of *Gloeobacterales* and PDB 2XB3 CyanoP structure from *Synechocystis* sp. PCC 6803. The APBS electrostatics was estimated based on the PyMOL plugin. The *Gloeobacterales* sequence was aligned to the PDB structure 2XB3 in order to find the ion-binding sites. (A, B) APBS electrostatics estimation of *Gloeobacterales* structure with two different views. (C, D) APBS electrostatics estimation of 2LNJ structure with two different views. (E) Alignment of *Gloeobacterales* CyanoP and 2LNJ structures.

Figure 8. Prediction of the 3D structure of Cyanobacterial YidC. (A) Structure alignment the P1 domain of of *Gloeobacterales* (cyan), *Thermostichales* (olive green) and *Synechocystis* (green). (B) Extra helices found in the P1 domain in Thermostichales (olive green) and *Synechocystis* (green) are highlighted in salmon.

Taxonomic distributions of 36 PSII assembly factors across six assembly steps and 13 membrane dynamics proteins in Cyanobacteria

Absent in *Gloeobacterales* | **Possible distinct origins in *Gloeobacterales*** | **Structural differences in *Gloeobacterales***



1 pD1 insertion

YidC

Sec machinery (SecYEGADF)

2 pD1 pre-processing

Chlorophyll incoporation:
ChlG, Ycf39, HliC, HliD, Pitt

Chlorophyll-binding molecule: Slr0144

Mn preloading: PratA

Cleavage of the C-ter of pD1: CtpA

3 RC complex assembly

D2 folding: Slr0286, Slr2013

Complex formation:
Ycf48, CtpA Sll0408, RubA

4 RC47 intermediate assembly

Psb28, Sll0933, HliA, HliB, Psb35

5 CP43 incorporation

Sll0933, Psb27, Sll0606, Psb34

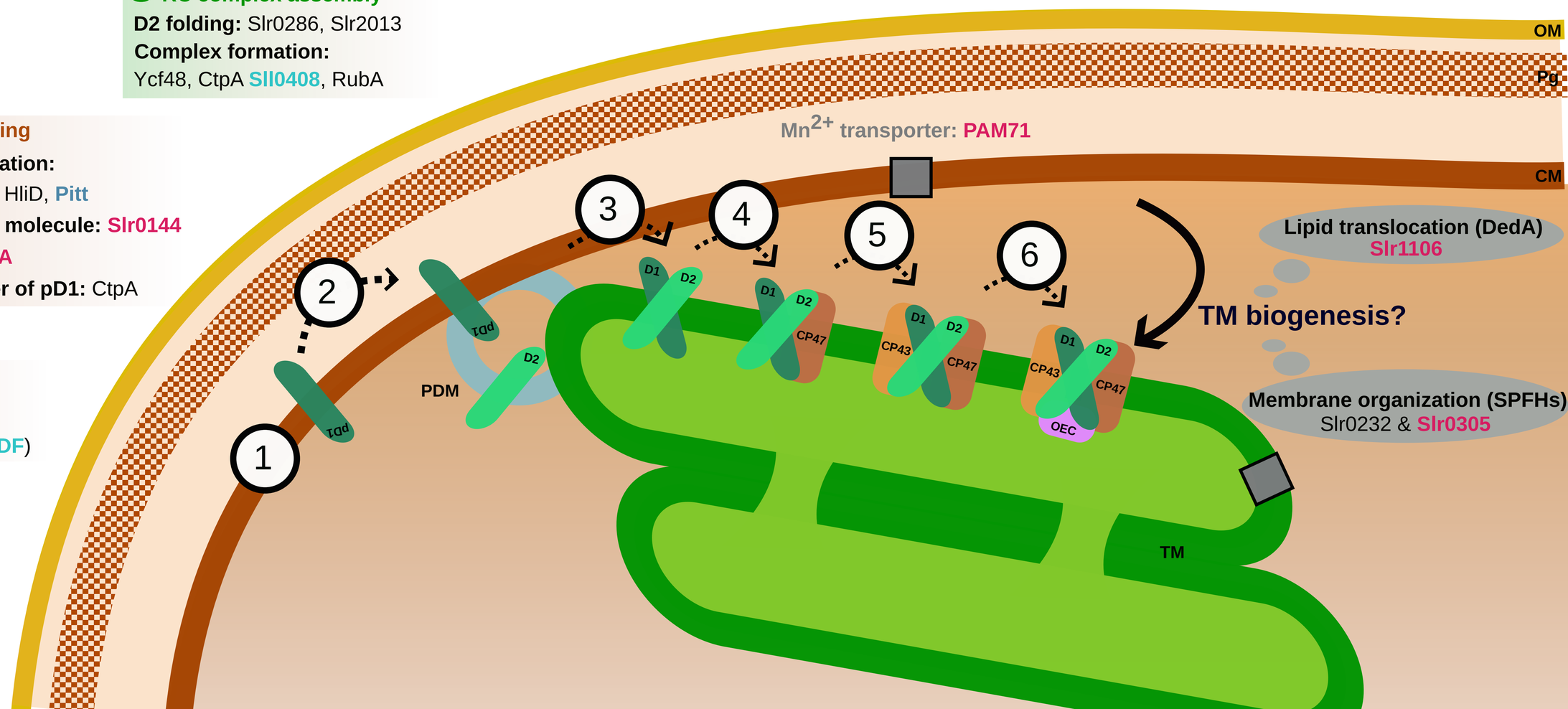
6 Functional PSII assembly

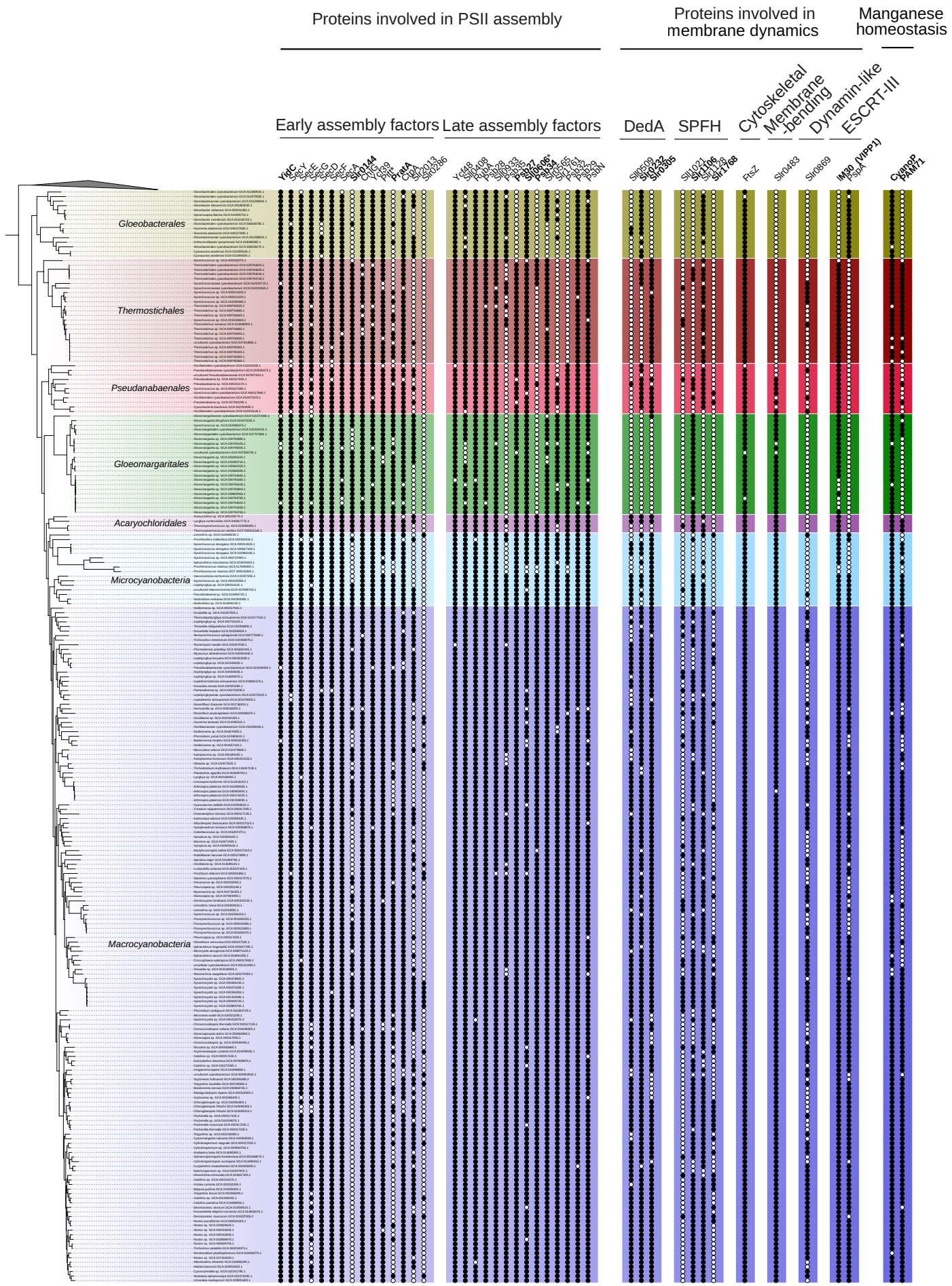
Photoactivation of the OEC:
CyanoP, Slr0565

Supexcomplex assembly: Slr1761

Photodamage protection: Psb32

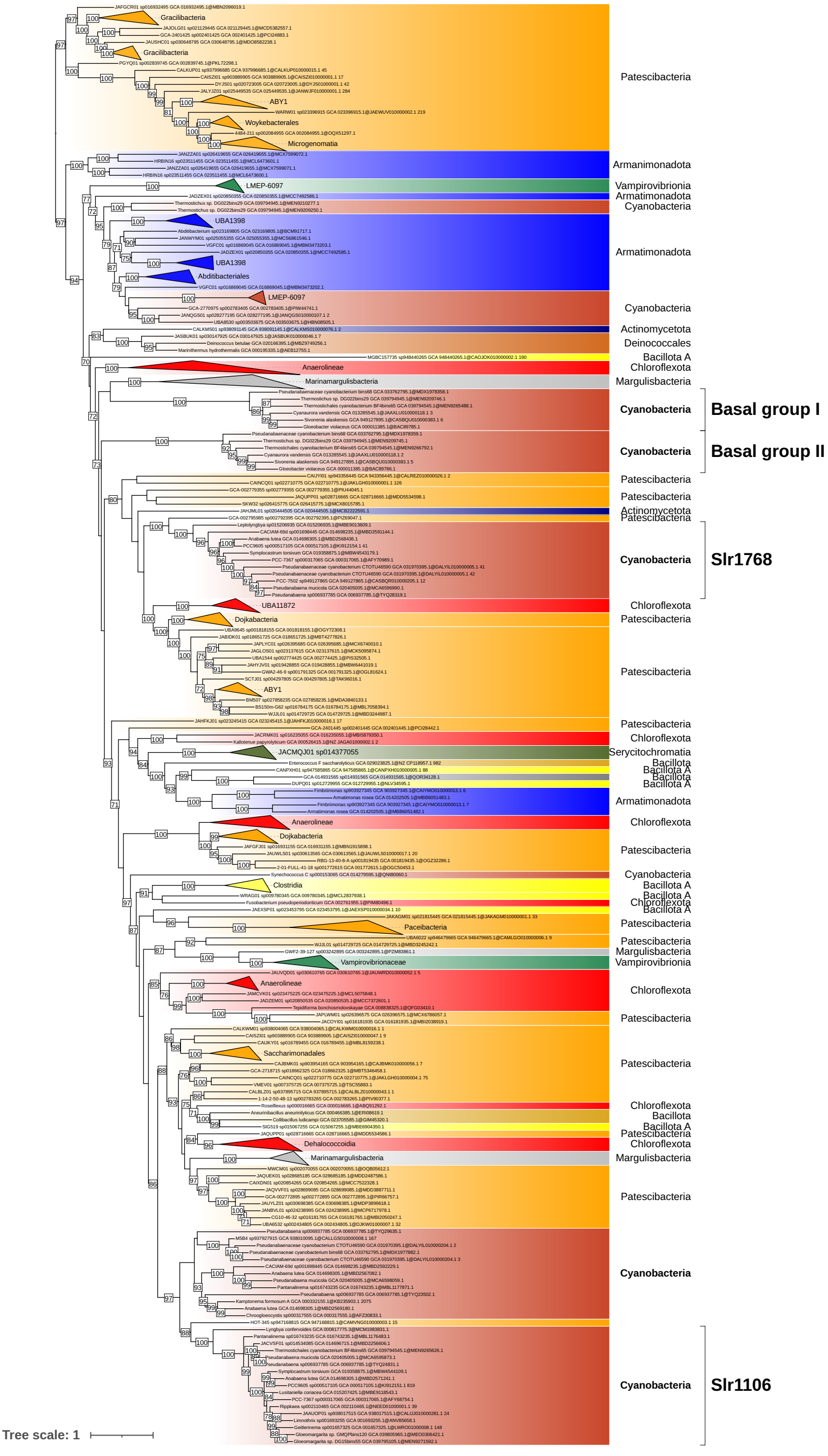
PSII repair: Psb29, Slr0151, PsbN



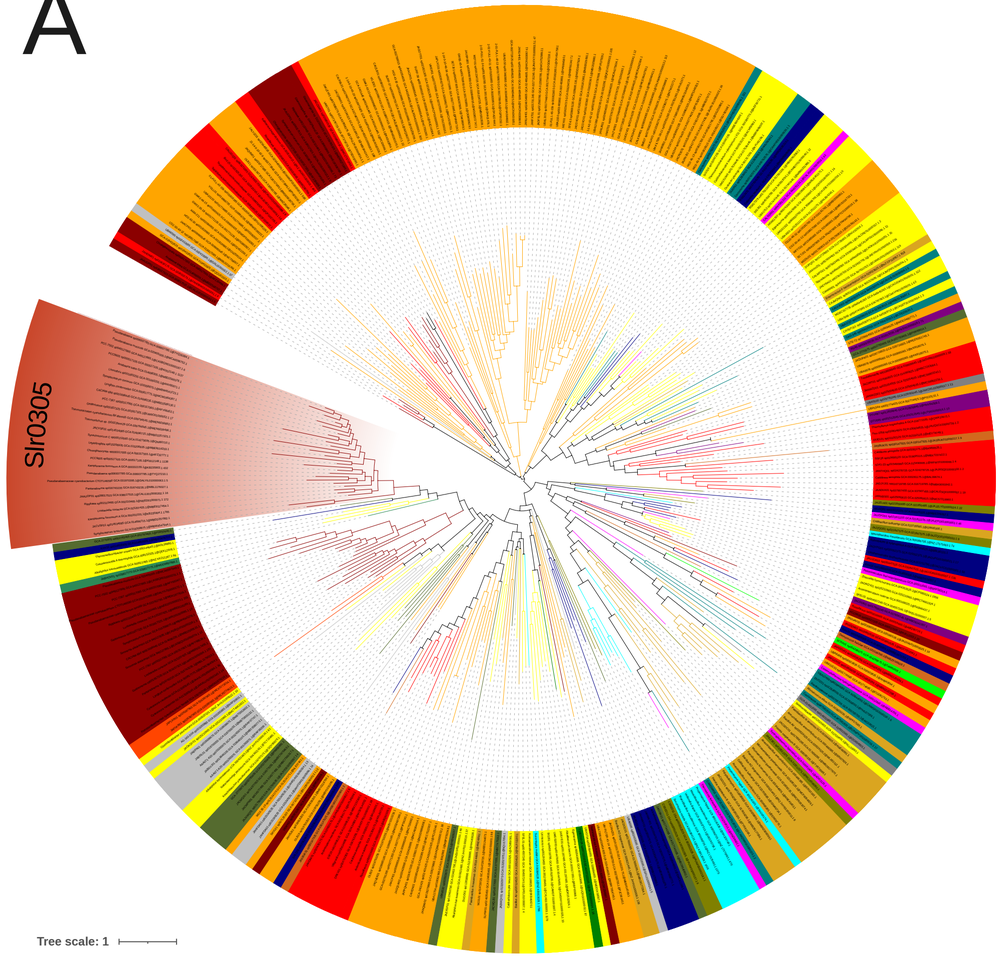


Proteins involved in PSII assembly
Proteins involved in membrane dynamics
Manganese homeostasis
Early assembly factors
Late assembly factors
DedA
SPFH
Cytoskeletal
Membrane -bending
Dynamin-like
ESCRT-III
YidC
SecY
SecE
SecG
SecD
SecF
SecA
Sir0144
CtnG
Ctg9
PitK
PratA
CtpA
Sir2013
Sir0286
Ycf48
Sll0408
NuBA
Sll0928
Sll0933
Psb35
Psb27
Sll0606*
Psb34
Sll0565
Sll1762
Psb32
Psb29
PsbN
Sll0509
Sir0232
Sir0305
Sll1021
Sll1106
Sll1128
Sir1768
FtsZ
Slr0483
Slr0869
IM30 (VIPP1)
PspA
CyanoP
PAM71
Gloeobacterales
Thermostichales
Pseudanabaenales
Gloeomargaritales
Acaryochloridales
Microcyanobacteria
Macrocyanobacteria

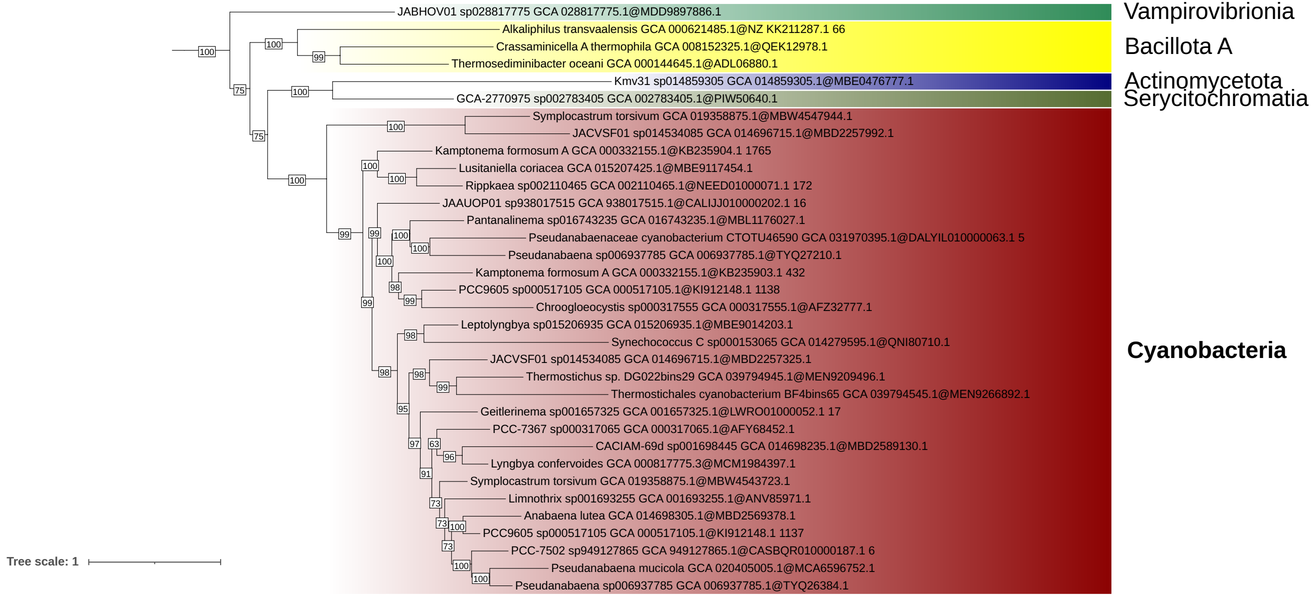
Tree scale: 1

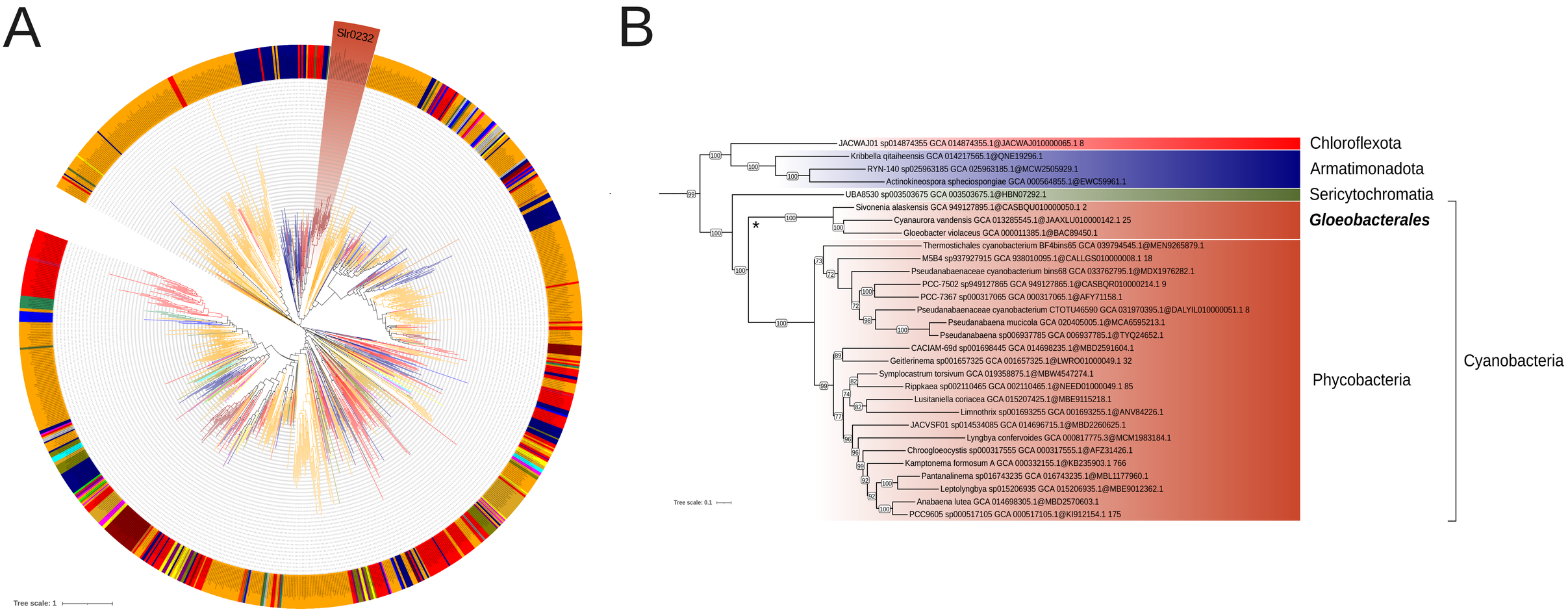

Gracilibacteria
Gracilibacteria
Patescibacteria
ABY1
Woykebacterales
Microgenomatia
Armanimonadota
Vampirovibrionia
Armatimonadota
Cyanobacteria
UBA1398
Armatimonadota
UBA1398
Abditibacteriales
LMEP-6097
Cyanobacteria
Actinomycetota
Deinococcales
Bacillota A
Anaerolineae
Chloroflexota
Marinamargulisbacteria
Margulisbacteria
Cyanobacteria
Basal group I
Cyanobacteria
Basal group II
Patescibacteria
Patescibacteria
Actinomycetota
Patescibacteria
Cyanobacteria
Slr1768
UBA11872
Chloroflexota
Dojkabacteria
Patescibacteria
Patescibacteria
Patescibacteria
Chloroflexota
JACMQJ01 sp014377055
Serycitochromatia
Bacillota A
Bacillota A
Armatimonadota
Anaerolineae
Chloroflexota
Dojkabacteria
Patescibacteria
Cyanobacteria
Clostridia
Bacillota A
Bacillota A
Chloroflexota
Bacillota A
Paceibacteria
Patescibacteria
Patescibacteria
Margulisbacteria
Vampirovibrionaceae
Vampirovibrionia
Anaerolineae
Chloroflexota
Patescibacteria
Saccharimonadales
Patescibacteria
Chloroflexota
Bacillota
Bacillota A
Patescibacteria
Dehalococcoidia
Chloroflexota
Marinamargulisbacteria
Margulisbacteria
Patescibacteria
Cyanobacteria
Cyanobacteria
Slr1106
Tree scale: 1

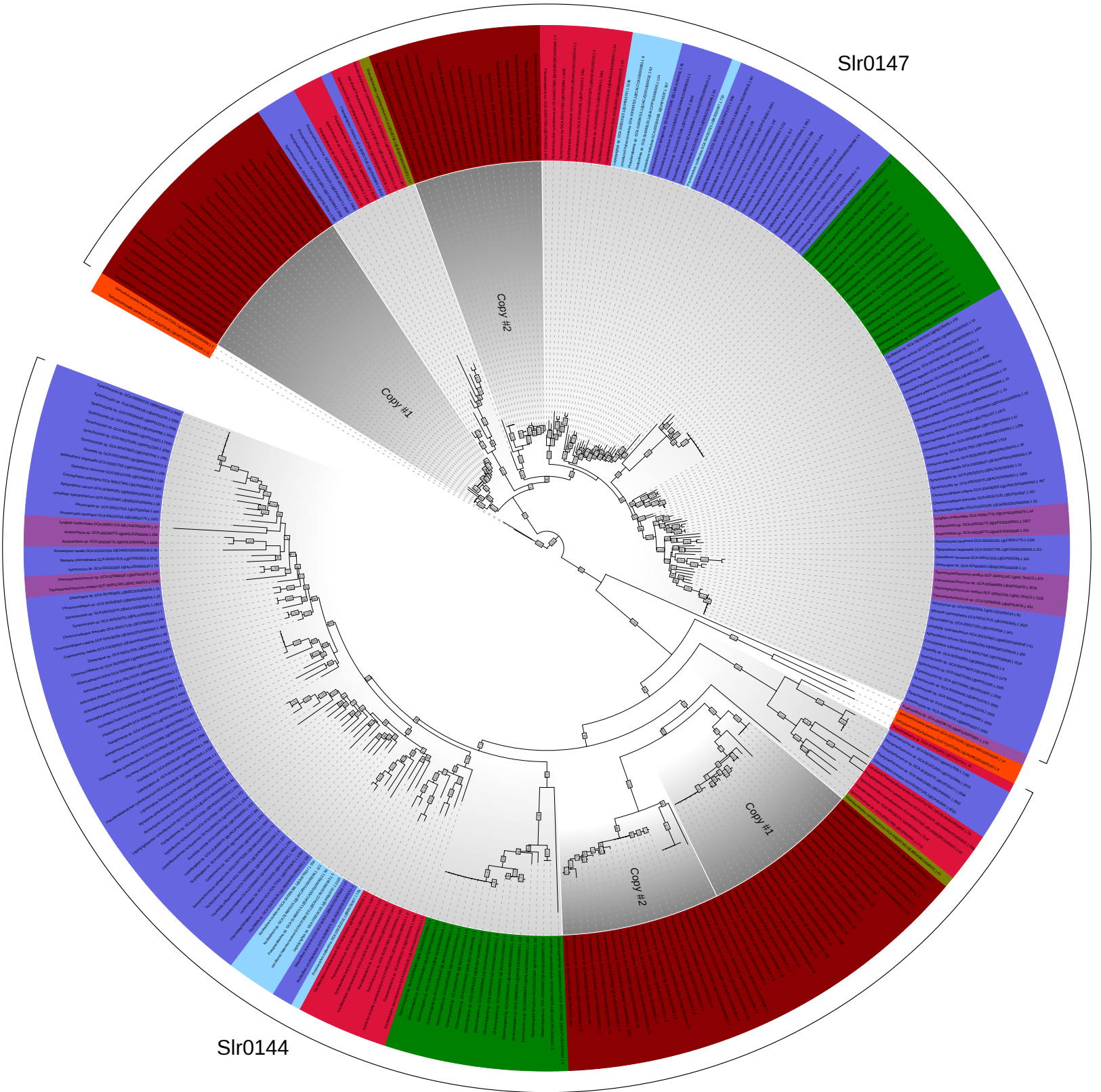
A

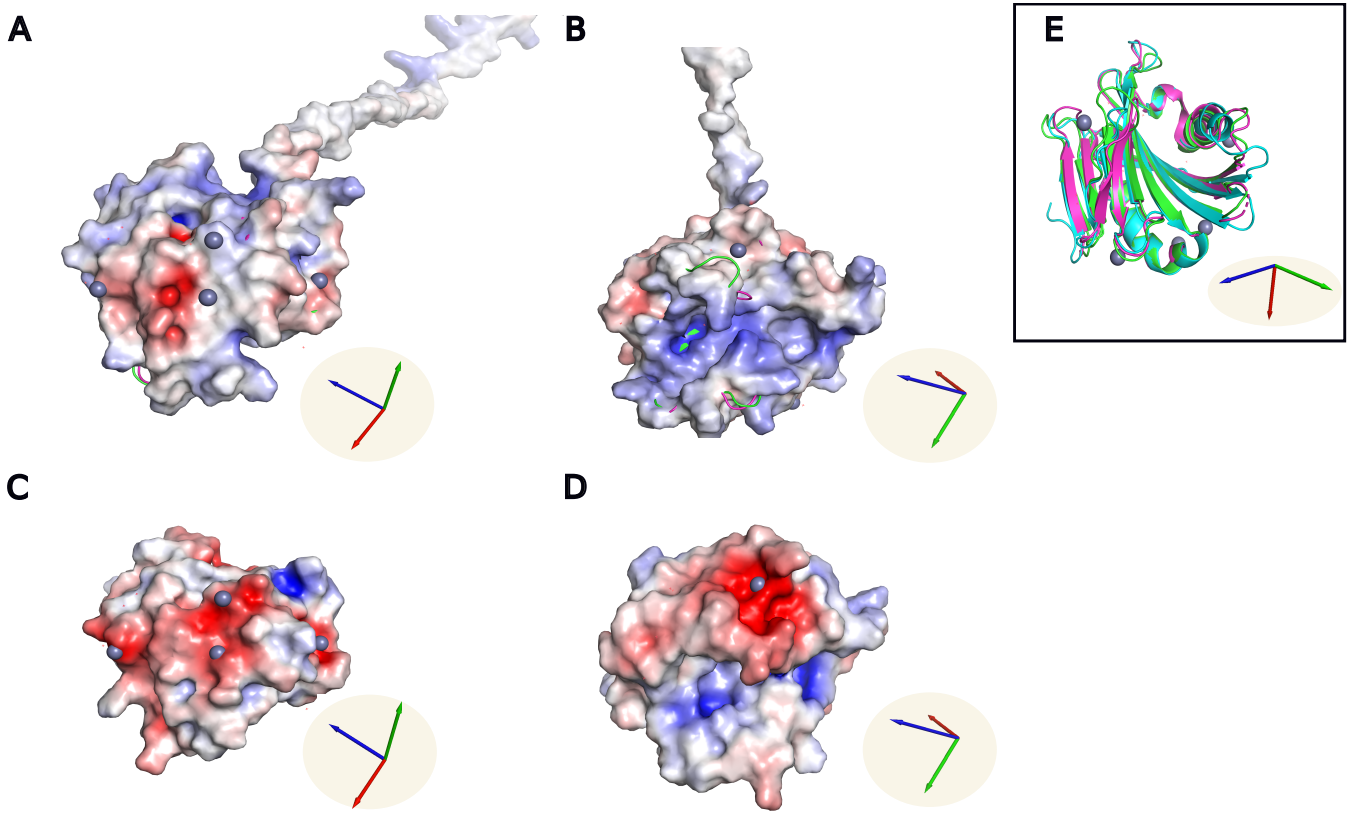

Slr0305
Tree scale: 1

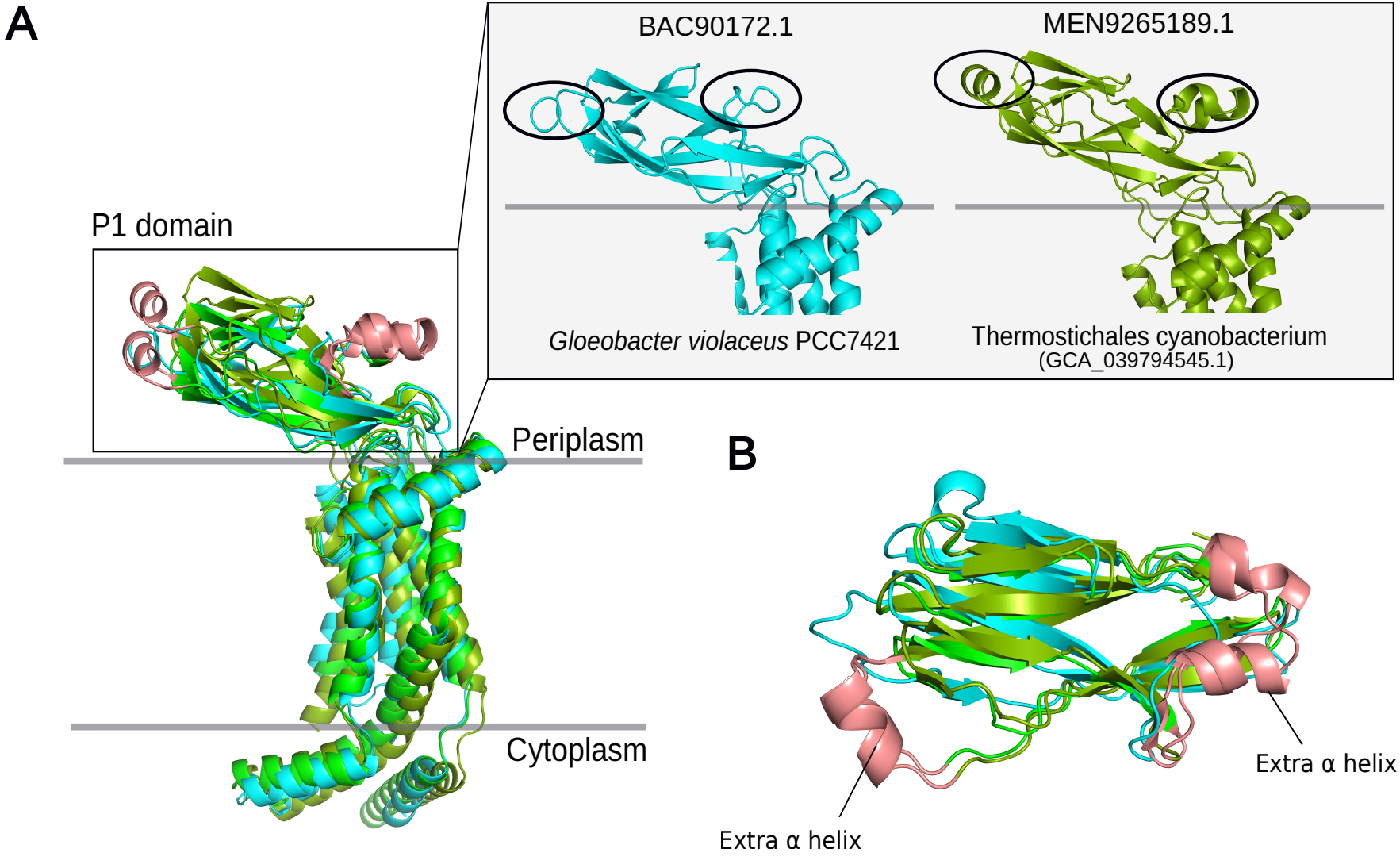
B


JABHOV01 sp028817775 GCA 028817775.1@MDD9897886.1
Alkaliphilus transvaalensis GCA 000621485.1@NZ KK211287.1 66
Crassaminicella A thermophila GCA 008152325.1@QEK12978.1
Thermosediminibacter oceani GCA 000144645.1@ADL06880.1
Kmv31 sp014859305 GCA 014859305.1@MBE0476777.1
GCA-2770975 sp002783405 GCA 002783405.1@PIW50640.1
Symplocastrum torsivum GCA 019358875.1@MBW4547944.1
JACVSF01 sp014534085 GCA 014696715.1@MBD2257992.1
Kamptonema formosum A GCA 000332155.1@KB235904.1 1765
Lusitaniella coriacea GCA 015207425.1@MBE9117454.1
Rippkaea sp002110465 GCA 002110465.1@NEED01000071.1 172
JAAUOP01 sp938017515 GCA 938017515.1@CALIJJ010000202.1 16
Pantanalinema sp016743235 GCA 016743235.1@MBL1176027.1
Pseudanabaenaceae cyanobacterium CTOTU46590 GCA 031970395.1@DALYIL010000063.1 5
Pseudanabaena sp006937785 GCA 006937785.1@TYQ27210.1
Kamptonema formosum A GCA 000332155.1@KB235903.1 432
PCC9605 sp000517105 GCA 000517105.1@KI912148.1 1138
Chroogloeocystis sp000317555 GCA 000317555.1@AFZ32777.1
Leptolyngbya sp015206935 GCA 015206935.1@MBE9014203.1
Synechococcus C sp000153065 GCA 014279595.1@QNI80710.1
JACVSF01 sp014534085 GCA 014696715.1@MBD2257325.1
Thermostichus sp. DG022bins29 GCA 039794945.1@MEN9209496.1
Thermostichales cyanobacterium BF4bins65 GCA 039794545.1@MEN9266892.1
Geitlerinema sp001657325 GCA 001657325.1@LWRO01000052.1 17
PCC-7367 sp000317065 GCA 000317065.1@AFY68452.1
CACIAM-69d sp001698445 GCA 014698235.1@MBD2589130.1
Lyngbya confervoides GCA 000817775.3@MCM1984397.1
Symplocastrum torsivum GCA 019358875.1@MBW4543723.1
Limnothrix sp001693255 GCA 001693255.1@ANV85971.1
Anabaena lutea GCA 014698305.1@MBD2569378.1
PCC9605 sp000517105 GCA 000517105.1@KI912148.1 1137
PCC-7502 sp949127865 GCA 949127865.1@CASBQR010000187.1 6
Pseudanabaena mucicola GCA 020405005.1@MCA6596752.1
Pseudanabaena sp006937785 GCA 006937785.1@TYQ26384.1
Vampirovibrionia
Bacillota A
Actinomycetota
Serycitochromatia
Cyanobacteria
Tree scale: 1


A
Slr0232
Tree scale: 1
B
JACWAJ01 sp014874355 GCA 014874355.1@JACWAJ010000065.1 8
Kribbella qitaiheensis GCA 014217565.1@QNE19296.1
RYN-140 sp025963185 GCA 025963185.1@MCW2505929.1
Actinokineospora sphecospongiae GCA 000564855.1@EWC59961.1
UBA8530 sp003503675 GCA 003503675.1@HBN07292.1
Sivonenia alaskensis GCA 949127895.1@CASBQU010000050.1 2
Cyanaurora vandensis GCA 013285545.1@JAAXLU010000142.1 25
Gloeobacter violaceus GCA 000011385.1@BAC89450.1
Thermostichales cyanobacterium BF4bins65 GCA 039794545.1@MEN9265879.1
M5B4 sp937927915 GCA 938010095.1@CALLGS010000008.1 18
Pseudanabaenaceae cyanobacterium bins68 GCA 033762795.1@MDX1976282.1
PCC-7502 sp949127865 GCA 949127865.1@CASBQR010000214.1 9
PCC-7367 sp000317065 GCA 000317065.1@AFY71158.1
Pseudanabaenaceae cyanobacterium CTOTU46590 GCA 031970395.1@DALYIL010000051.1 8
Pseudanabaena mucicola GCA 020405005.1@MCA6595213.1
Pseudanabaena sp006937785 GCA 006937785.1@TYQ24652.1
CACIAM-69d sp001698445 GCA 014698235.1@MBD2591604.1
Geitlerinema sp001657325 GCA 001657325.1@LWRO01000049.1 32
Symplocastrum torsivum GCA 019358875.1@MBW4547274.1
Rippkaea sp002110465 GCA 002110465.1@NEED01000049.1 85
Lusitaniella coriacea GCA 015207425.1@MBE9115218.1
Limnothrix sp001693255 GCA 001693255.1@ANV84226.1
JACVSF01 sp014534085 GCA 014696715.1@MBD2260625.1
Lyngbya confervoides GCA 000817775.3@MCM1983184.1
Chroogloeocystis sp000317555 GCA 000317555.1@AFZ31426.1
Kamptonema formosum A GCA 000332155.1@KB235903.1 766
Pantanalinema sp016743235 GCA 016743235.1@MBL1177960.1
Leptolyngbya sp015206935 GCA 015206935.1@MBE9012362.1
Anabaena lutea GCA 014698305.1@MBD2570603.1
PCC9605 sp000517105 GCA 000517105.1@KI912154.1 175
Tree scale: 0.1
Chloroflexota
Armatimonadota
Sericytochromatia
Gloeobacterales
Phycobacteria
Cyanobacteria


Slr0147
Copy #2
Copy #1
Copy #1
Copy #2
Slr0144
Tree scale: 1


A
B
C
D
E


A
BAC90172.1
MEN9265189.1
P1 domain
Gloeobacter violaceus PCC7421
Thermostichales cyanobacterium
(GCA_039794545.1)
Periplasm
Cytoplasm
B
Extra α helix
Extra α helix