



Supplemental Figure 9. Maximum likelihood phylogeny of GlgA2 (**A**) and 16S RNA (**B**) of the same strains of Cyanobacteria. All the sequence for each was taken from the NCBI database, aligned with MUSCLE and then blocs were selected with BMGE (Criscuolo and Gribaldo 2010) and checked manually. The maximum likelihood phylogenies were built using IQTREE (Nguyen et al., 2015), and either LG4X matrix (Le et al. 2012) (**A**) and GTR+I+G matrix (Lanave et al. 1984, Yang 1994) (**B**). The bootstrap values are indicated on each branch, based on 1000 bootstrap replicas. (**A**) The outgroup was chosen among the closest sequences of cyanobacteria GlgA2 present in the NCBI database and that belongs to GlgA2 family. (**B**) The outgroup is composed of *Gloeobacter* 16S RNA sequences, that is known to be an early branching cyanobacterial group. The groups of organisms found in both trees are highlighted in different colors.

Relevant supplemental references

12 **Criscuolo A, Gribaldo S.** BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions
13 from multiple sequence alignments. *BMC Evol Biol.* 2010 Jul 13;10:210.

14 **Lanave C, Preparata G, Saccone C, Serio G.** A new method for calculating evolutionary substitution rates. *J Mol Evol.* 1984;20:86-93.

15 **Le SQ, Dang CC, Gascuel O.** Modeling protein evolution with several amino acid replacement matrices depending on site rates. *Mol Biol Evol.*
16 2012 Oct;29(10):2921-36. **Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ.** IQ-TREE: a fast and effective stochastic algorithm for estimating
17 maximum-likelihood phylogenies. *Mol Biol Evol.* 2015 Jan;32(1):268-74.

18 **Yang Z.** Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J Mol Evol.*
19 1994;39:306-314.

20