Supplementary Information for

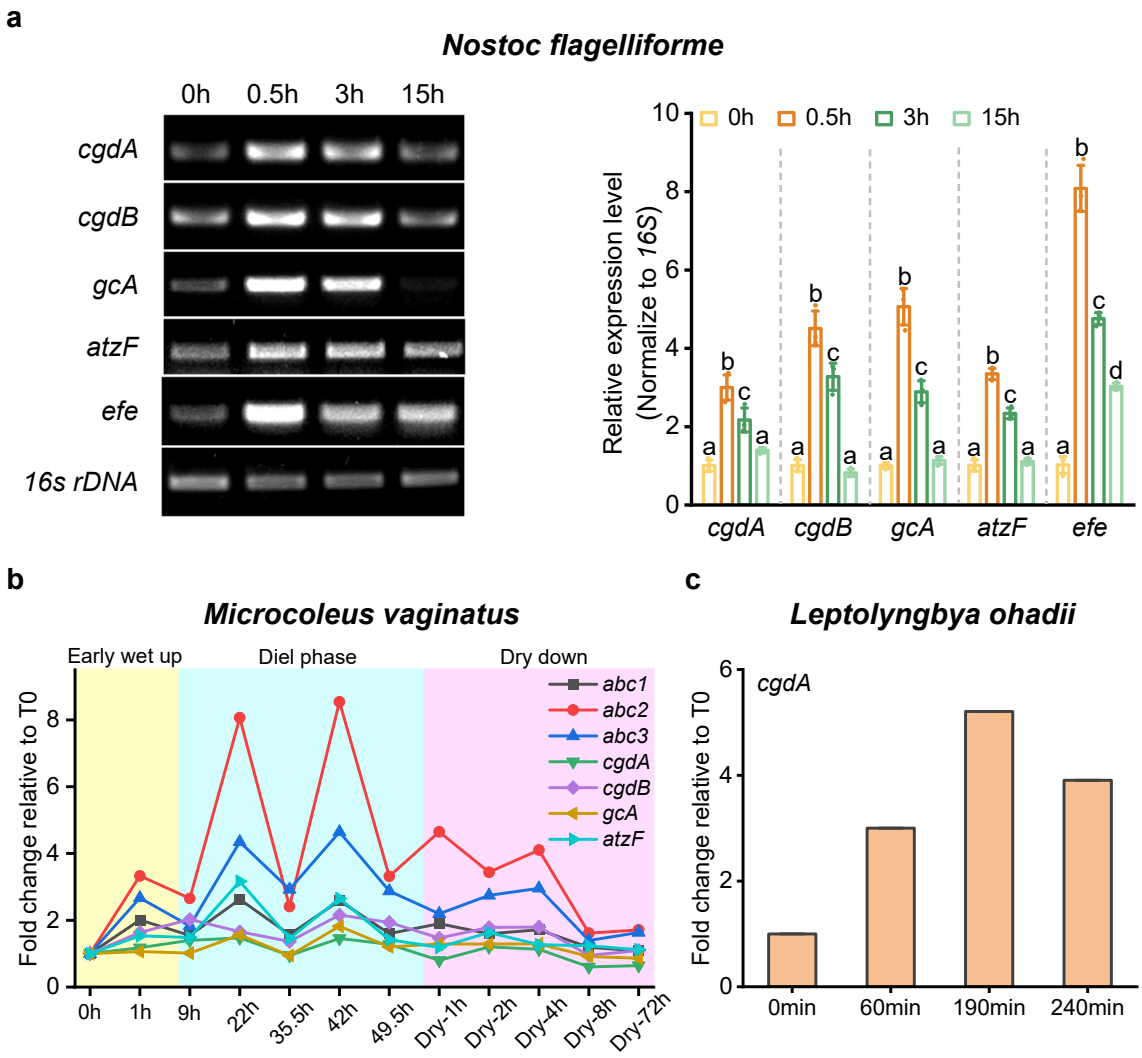
Guanidine sparks rapid resurrection of desert cyanobacteria

Jin-Long Shang, Guo-Wei Qiu, Lei Zhao, Hai-Feng Xu, Yao Cheng, Yong Li, Zhong-Chun Zhang, Guo-Zheng Dai, Shengwei Hou, Chen Yang, Wolfgang R. Hess & Bao-Sheng Qiu

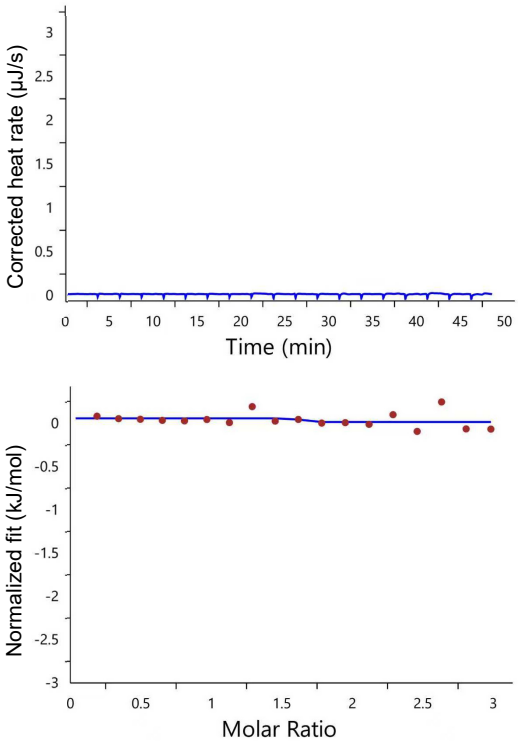
Corresponding author: Bao-Sheng Qiu (email: bsqiu@mail.ccnu.edu.cn).

This file includes

Supplementary Figures 1 to **15**.

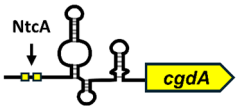


Supplementary Fig. 1 Transcriptional responses of guanidine metabolism genes to dehydration and rehydration in representative desert cyanobacteria. a Expression of *Nostoc flagelliforme* genes involved in guanidine production (*efe*) and carboxylation (*cgdA*, *cgdB*, *gcA*, *atzF*) during rehydration was analyzed by RT-PCR and RT-qPCR. Data represent mean ± SD of three independent biological replicates, normalized to transcript levels under desiccation. Letters above bars indicate statistical significance ($P < 0.05$; one-way ANOVA across all treatments). **b, c** Transcriptional changes of guanidine metabolism genes in *Microcoleus vaginatus* (**b**) during simulated dehydration-rehydration cycles, and in *Leptolyngbya ohadii* (**c**) during dehydration. Data were re-analyzed and plotted from Rajeev *et al.*[1] and Oren *et al.*[2]. *Leptolyngbya ohadii* shows reduced water content after 120 minutes of treatment.



Supplementary Fig. 2 Binding affinity of the *cdgA* riboswitch for urea. The upper panel displays the original titration data, and the lower panel shows the integrated heat measurements. Binding isotherms derived from a one-site binding model characterize the interaction between the guanidine-I riboswitch and urea. Data are from a representative experiment of three independent biological replicates.

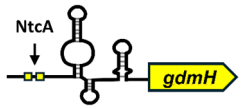
a



NtcA binding site −10 box TSS

GTAGTAAGTGATACaaattgctttgggcatccttgctaaatttatcatGagcaaacattgtgttaa
aattattaatataaataaaaacactctgttgacaatattgtagaagtttcgcaaattcaaattgcaaactaa
ataaggatgagcttctagggttccgatcgcgcctttgtgtgtgatggctggtccaagagaagcag
ccgactttgagagtcggtcacacggcgggaaaaatgcccgggagagccagacagtatgactgt
attggtaatctcctgggctttgctatttgagtcagtcgtttcttcaattccaaaggttgagcctgtggaaattcct
aaaccgatcgtccaaccgtaacaggagaatttttATG…

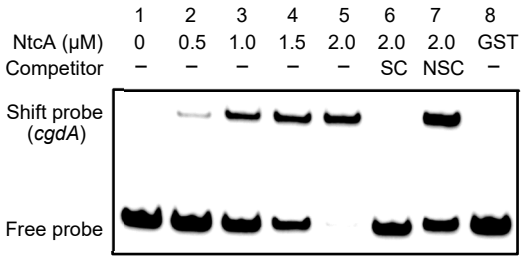
b



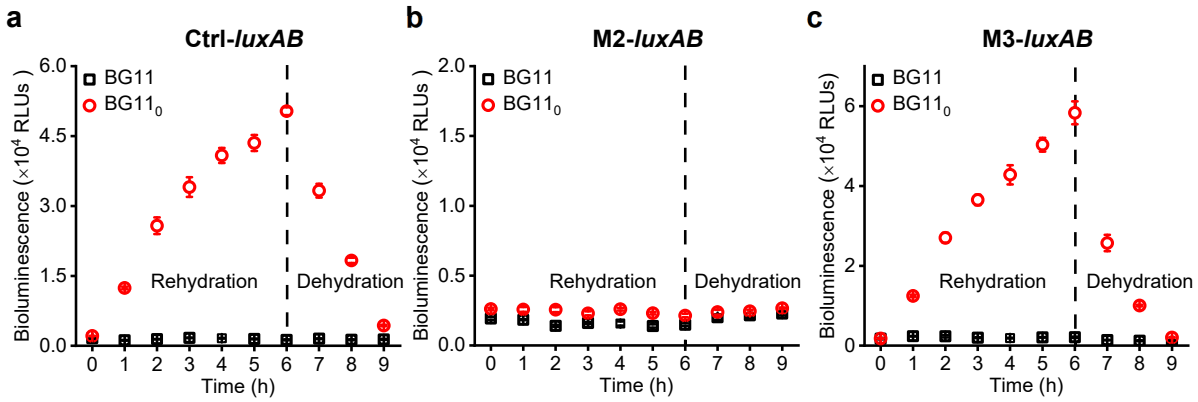
NtcA binding site −10 box TSS

GTAGCACAGAACACagttaacttaatcatccctagctagggtaggggAtattgtttctagggtt
ccgattcctggtgatgtggggaatggctggtccgagagaaacaggctaggctttgtaaagatttt
tgccactagctacacggcgggacaaaagcccgggagacttaatgaccagtgcttaggtgttgtca
ttaataactcctgggcagattttgttgtttagctagtaactacctATG…

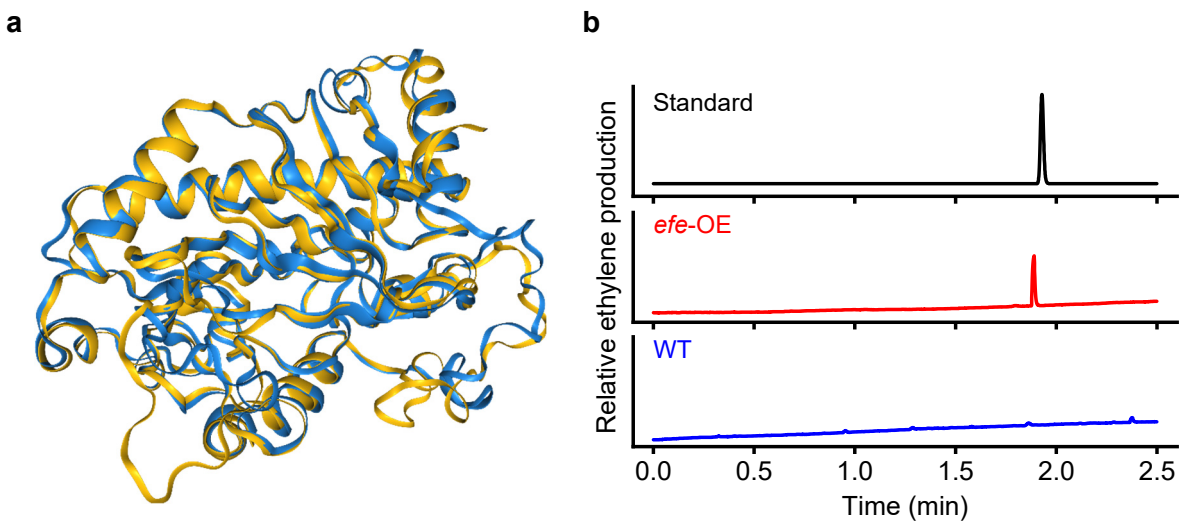
c



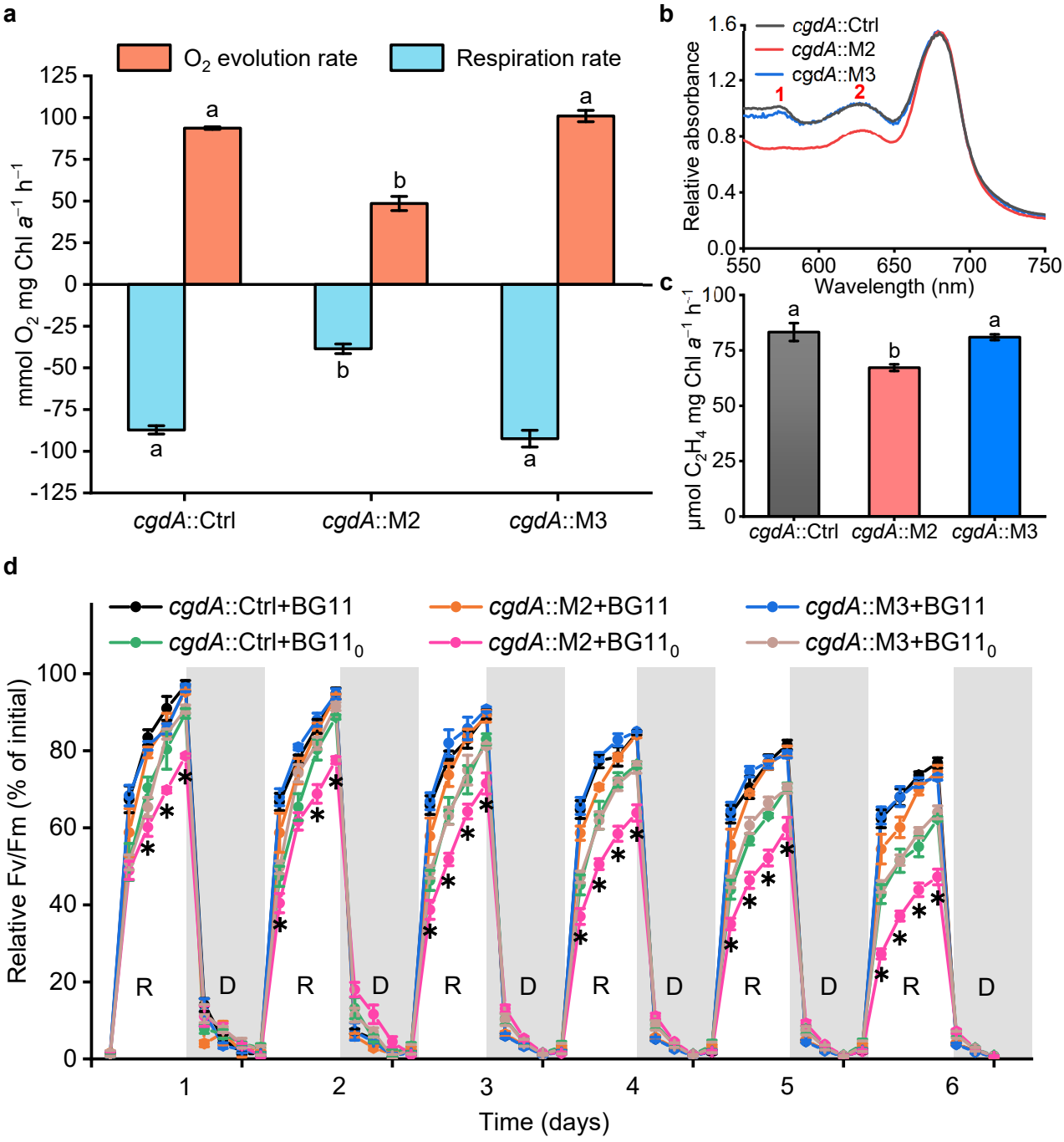
Supplementary Fig. 3 Examination of the interaction between NtcA and the *cgdA* promoter. The schematic diagram illustrates the location of potential NtcA-binding sites within the promoters of key genes in the guanidine carboxylation (*cgdA* in *Nostoc flagelliforme* CCNUN1, **a**) and hydrolysis (*gdmH* in *Synechocystis* sp. PCC 6803, **b**) pathways. The corresponding DNA sequences of these *cis*-acting elements are shown on the right, with the nucleotide associated with the guanidine-I riboswitch highlighted in blue. The interaction between NtcA and the *cgdA* promoter was further confirmed by EMSA (**c**), with the NtcA concentration in each reaction mixture indicated. Specificity of the binding was assessed by competition with unlabeled specific competitor (SC) DNA and non-specific poly (dI-dC) (NSC). Data are representative of three independent experiments.



Supplementary Fig. 4 Promoter activity of *cgdA* over an extended rehydration time. Exponentially growing *N. flagelliforme* strains harboring LuxAB reporter fusions, with the wild-type (Ctrl-*luxAB*, **a**), riboswitch mutant (M2-*luxAB*, **b**), or riboswitch-complemented (M3-*luxAB*, **c**) promoter, were first incubated in liquid BG11 or $BG11_0$ (nitrate-deficient BG11) medium for 3 days. The cells were then collected on 0.45 μm nitrocellulose membranes and subjected to alternating dehydration-rehydration cycles (21 h : 3 h) for 2 days. After this, a time-course bioluminescence analysis was performed during an extended 6-hour rehydration period on the third day. Promoter activity was quantified as relative luminescence units (RLU). Data are presented as mean ± SD of three independent biological replicates.

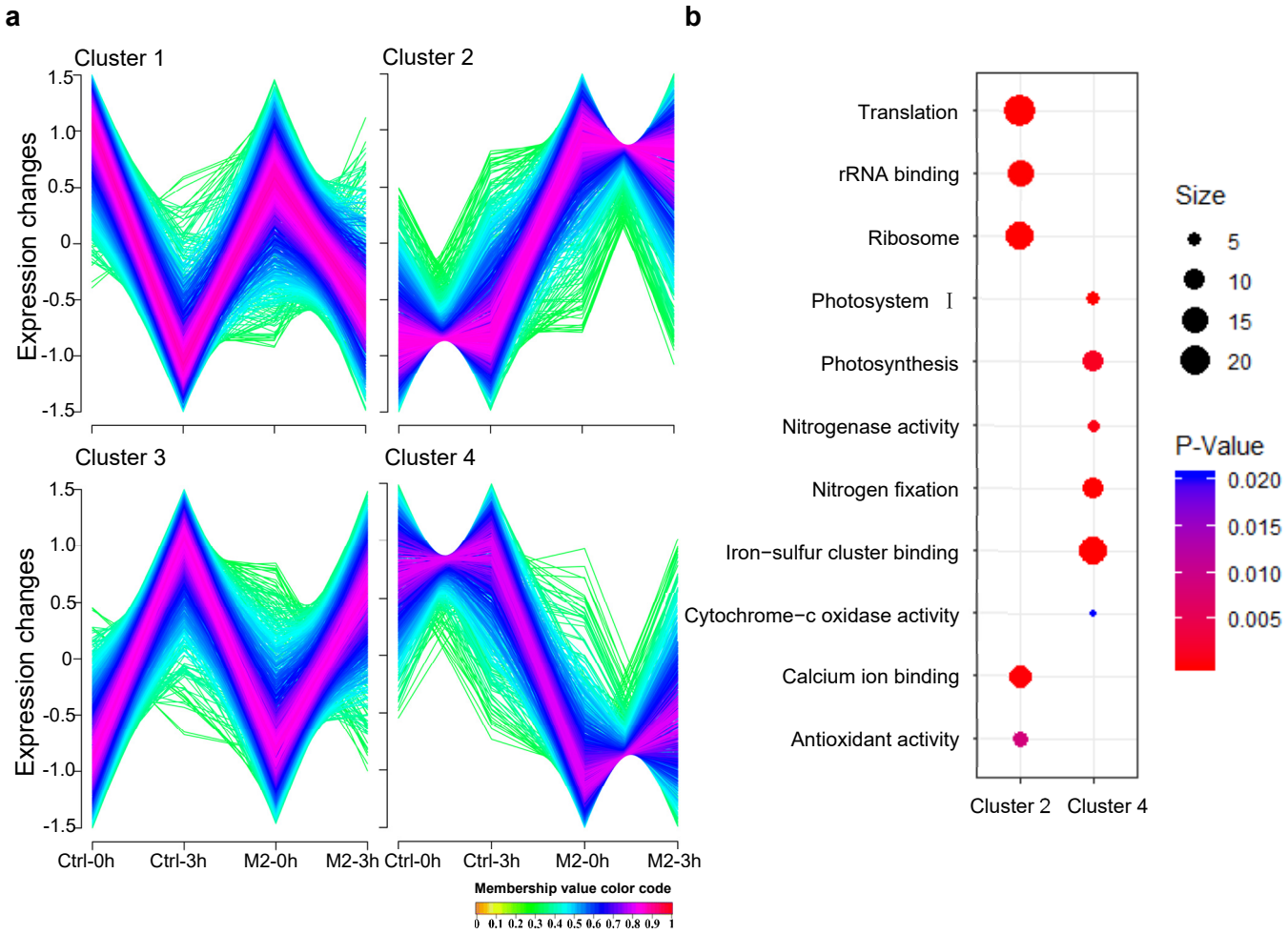


Supplementary Fig. 5 Identification of a putative guanidine-producing enzyme in *N. flagelliforme*. a The predicted structure of *N. flagelliforme* COO91_01204 exhibits high structural similarity to the 2-oxoglutarate-dependent ethylene/succinate-forming enzyme (Efe) from *Pseudomonas aeruginosa* PAO1 (TM-score = 0.92199), as identified by Foldseek[3]. The aligned structures of COO91_01204 (blue) and Efe (yellow) are shown. **b** Analysis of ethylene production by heterologous expression of COO91_01204 in *Synechocystis* sp. PCC 6803 via gas chromatography. Ethylene is produced according to the stoichiometric formula[4]: 3 2-oxoglutarate + 3 O_2 + L-arginine → 2 ethylene + succinate + 7 CO_2 + 3 H_2O + guanidine + pyrroline-5-carboxylate. Ethylene gas and wild-type *Synechocystis* served as positive and negative controls, respectively. Data are from one representative experiment of three biological replicates.

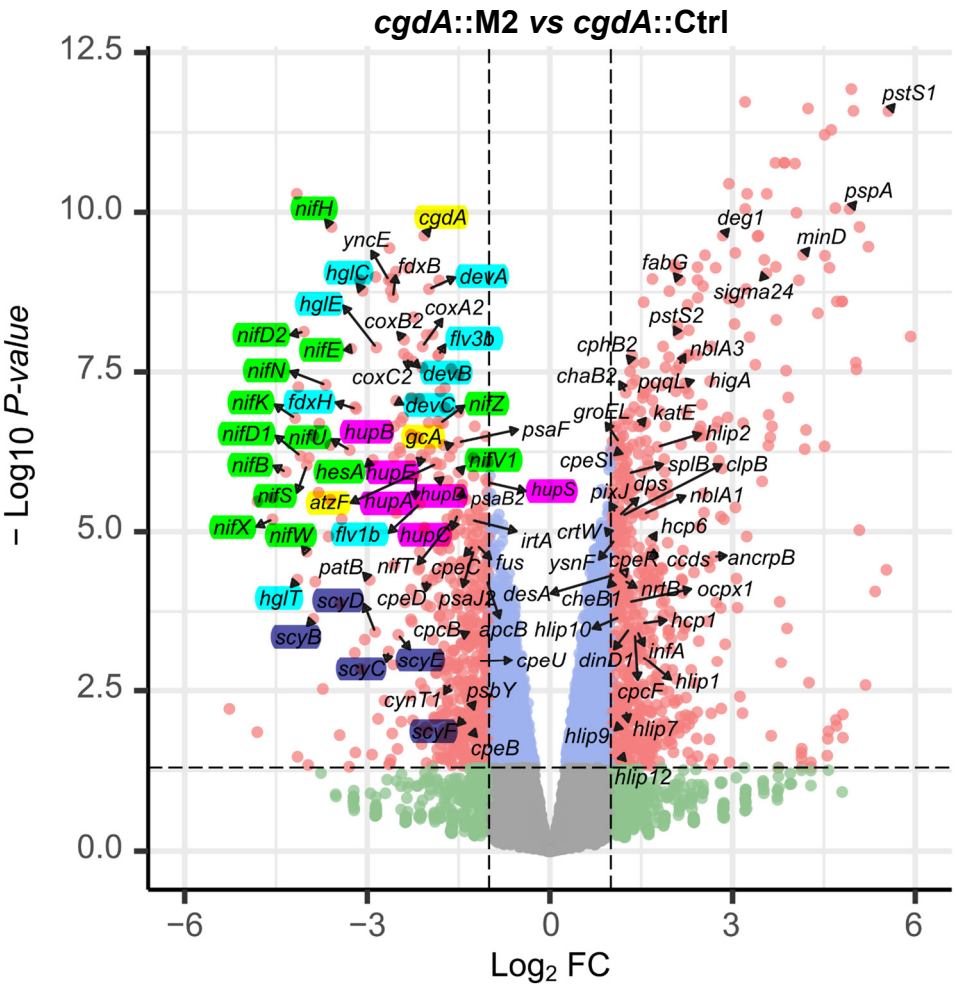


Supplementary Fig. 6 Characterization of the *N. flagelliforme cgdA*::Ctrl, *cgdA*::M2, and *cgdA*::M3 strains. a-c Phenotype of *cgdA*::Ctrl, *cgdA*::M2, and *cgdA*::M3 strains after 3 hours of rewetting. Cells pre-adapted in nitrogen-deplete medium for three days (30 mL, OD_{750} = 0.5) were collected on 0.45 μm nitrocellulose membrane filters and subjected to three dehydration-rehydration cycles (21 h : 3 h). Parameters were measured after 3 hours of rehydration following the final cycle. Net oxygen evolution and respiration rates (**a**); Representative whole-cell absorption spectra (one of three independent biological replicates), with peaks corresponding to phycoerythrin (1) and phycocyanin (2) indicated (**b**); Nitrogenase activity (**c**). Data are mean ± SD of three independent biological replicates for panels **a** and **c**. Different letters above bars indicate statistically significant differences ($P < 0.05$, one-way ANOVA across all treatments). **d** Dynamics of photosynthetic efficiency (Fv/Fm) in *cgdA*::Ctrl, *cgdA*::M2, and *cgdA*::M3 strains during dehydration (D) and rehydration (R) cycles. Data are mean ± SD of three independent biological replicates. The asterisk

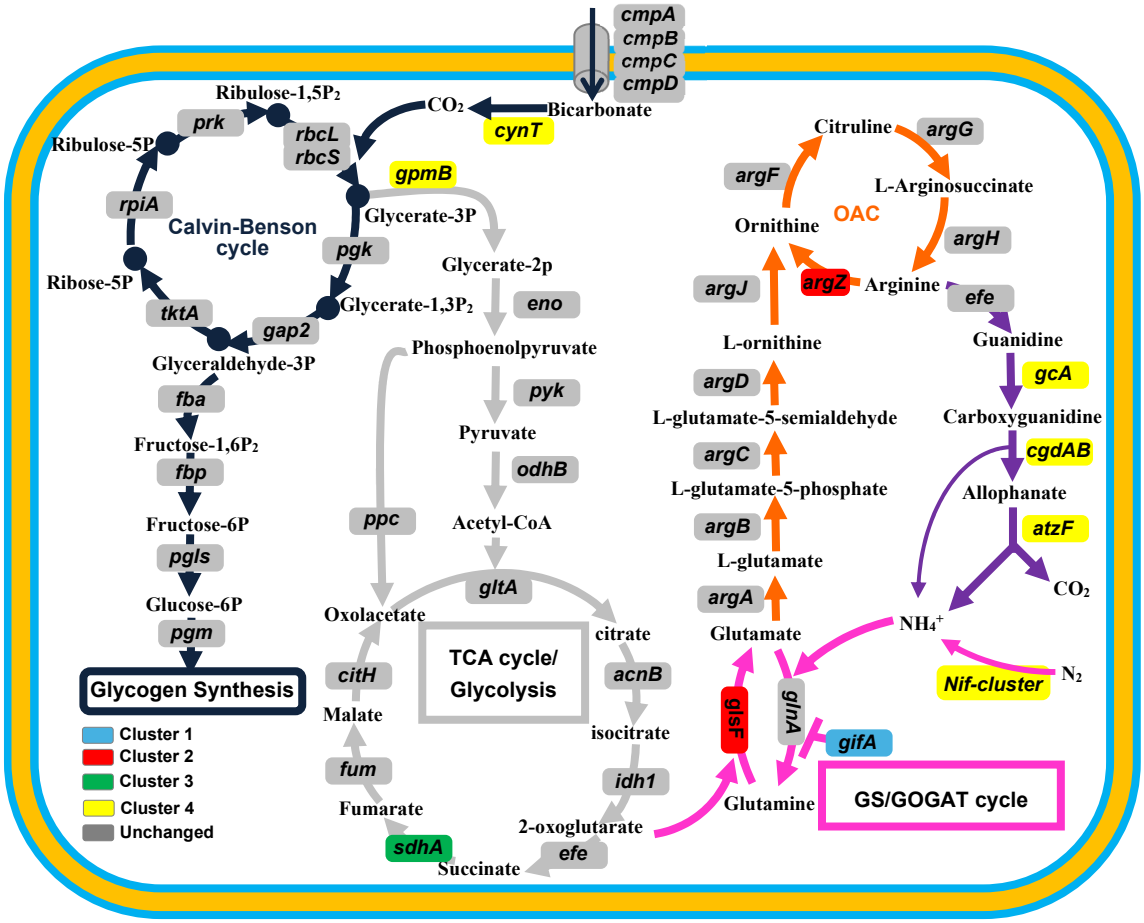
denotes a significant difference between nitrogen-starved *cgdA*::Ctrl (*cgdA*::Ctrl+BG11$_0$) and *cgdA*::M2 (*cgdA*::M2+BG11$_0$) strains as determined by a two-sided Student's *t*-test ($P < 0.05$).



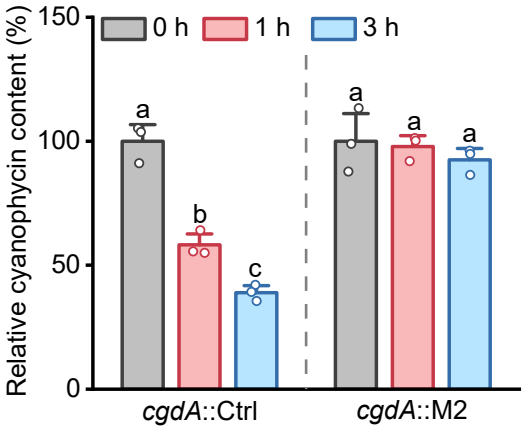
Supplementary Fig. 7 Transcriptional patterns of *N. flagelliforme cgdA*::Ctrl and *cgdA*::M2 strains during rewetting after dehydration and rehydration cycles. a Cluster analysis of gene expression profiles at 0 and 3 hours of rehydration, performed using the Fuzzy C-Means (FCM) algorithm. **b** Gene ontology enrichment analysis for differentially expressed genes identified in Cluster 2 and Cluster 4.



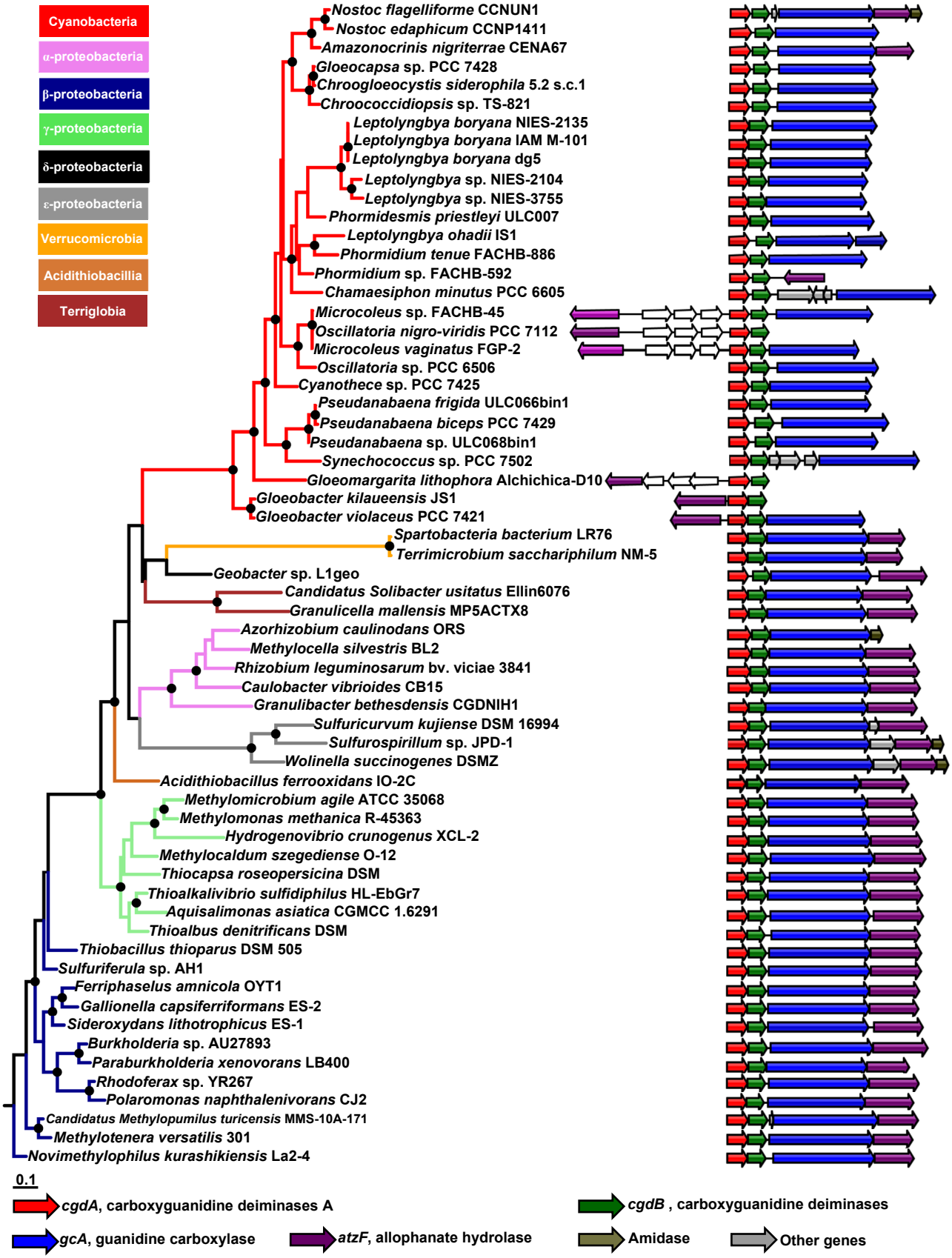
Supplementary Fig. 8 Global expression analysis of *cgdA*::M2 vs. *cgdA*::Ctrl strains after 3 hours of rewetting from dehydration and rehydration cycles. Dashed lines indicate the thresholds for statistical significance ($P < 0.05$) and fold change (FC = ±1). Functional groups are distinguished by color, and genes meeting both significance criteria (differentially expressed genes, DEGs) are highlighted in red. The corresponding numerical data are provided in Supplementary Data 1.



Supplementary Fig. 9 Alterations in carbon and nitrogen metabolism genes in *cgdA*::M2 vs. *cgdA*::Ctrl strains after 3 h of rewetting from dehydration and rehydration cycles. Genes are color-coded based on their expression patterns. The corresponding numerical values for the differentially expressed genes are available in Supplementary Data 1.

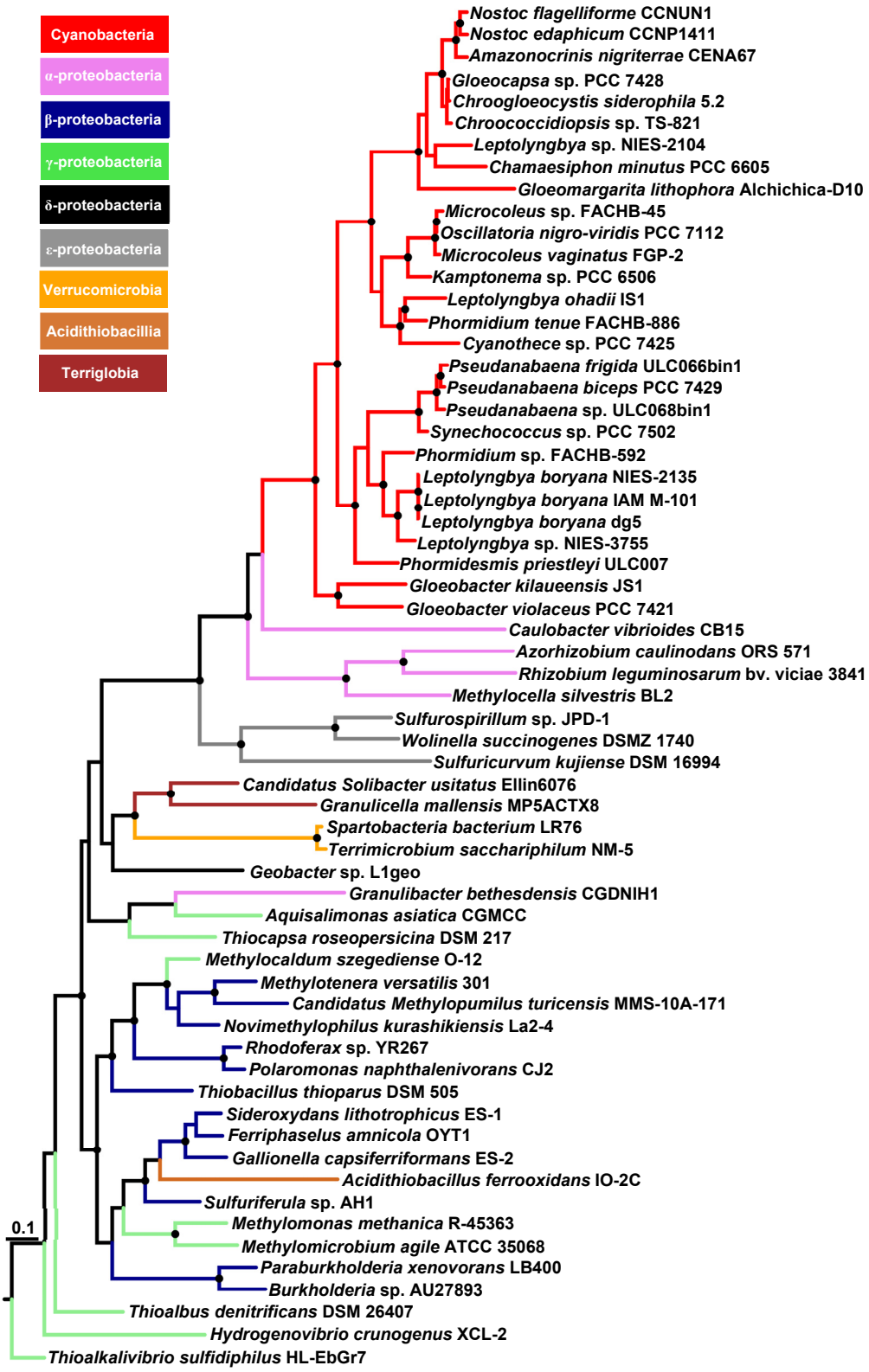


Supplementary Fig. 10 Intracellular cyanophycin levels in *cgdA*::Ctrl and *cgdA*::M2 strains after 3 hours of rewetting from dehydration and rehydration cycles. Data are shown as mean ± SD of three independent biological replicates, with data normalized to the pre-rewetting (0 h) levels. Different lowercase letters above the bars indicate statistically significant differences ($P < 0.05$) as determined by one-way ANOVA.

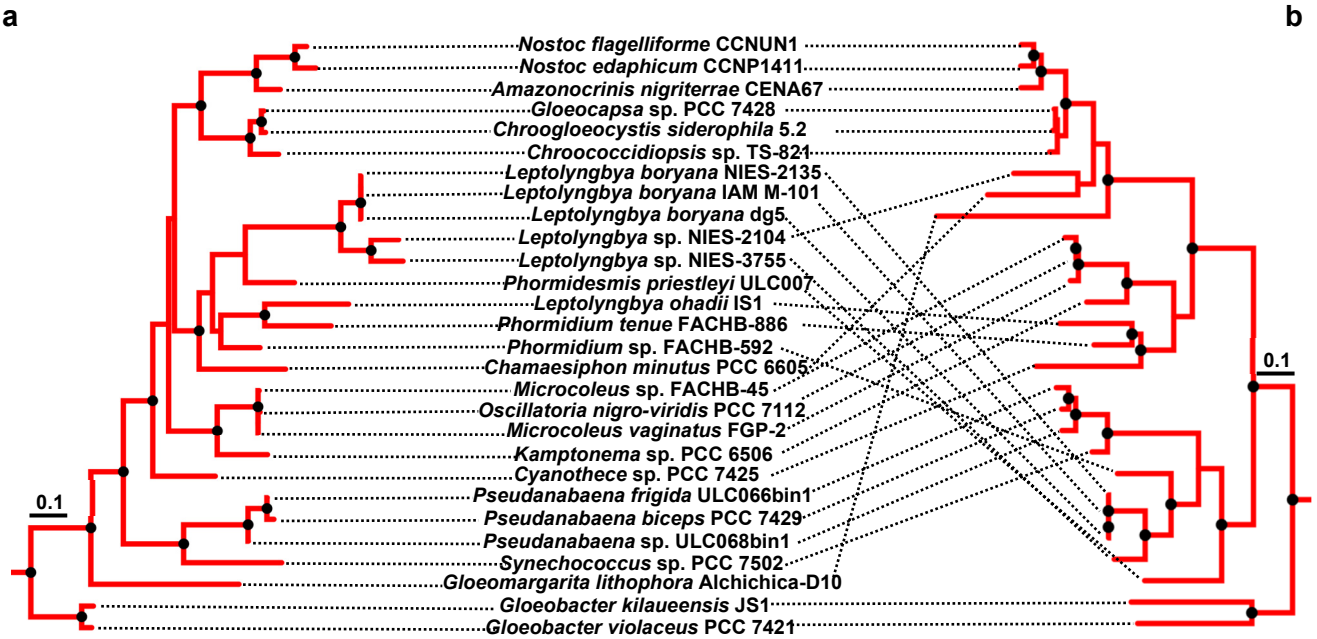


Supplementary Fig. 11 Phylogenetic analysis of representative cyanobacteria and other bacterial species based on 16S rRNA gene sequences. A phylogenetic tree was constructed using the maximum-

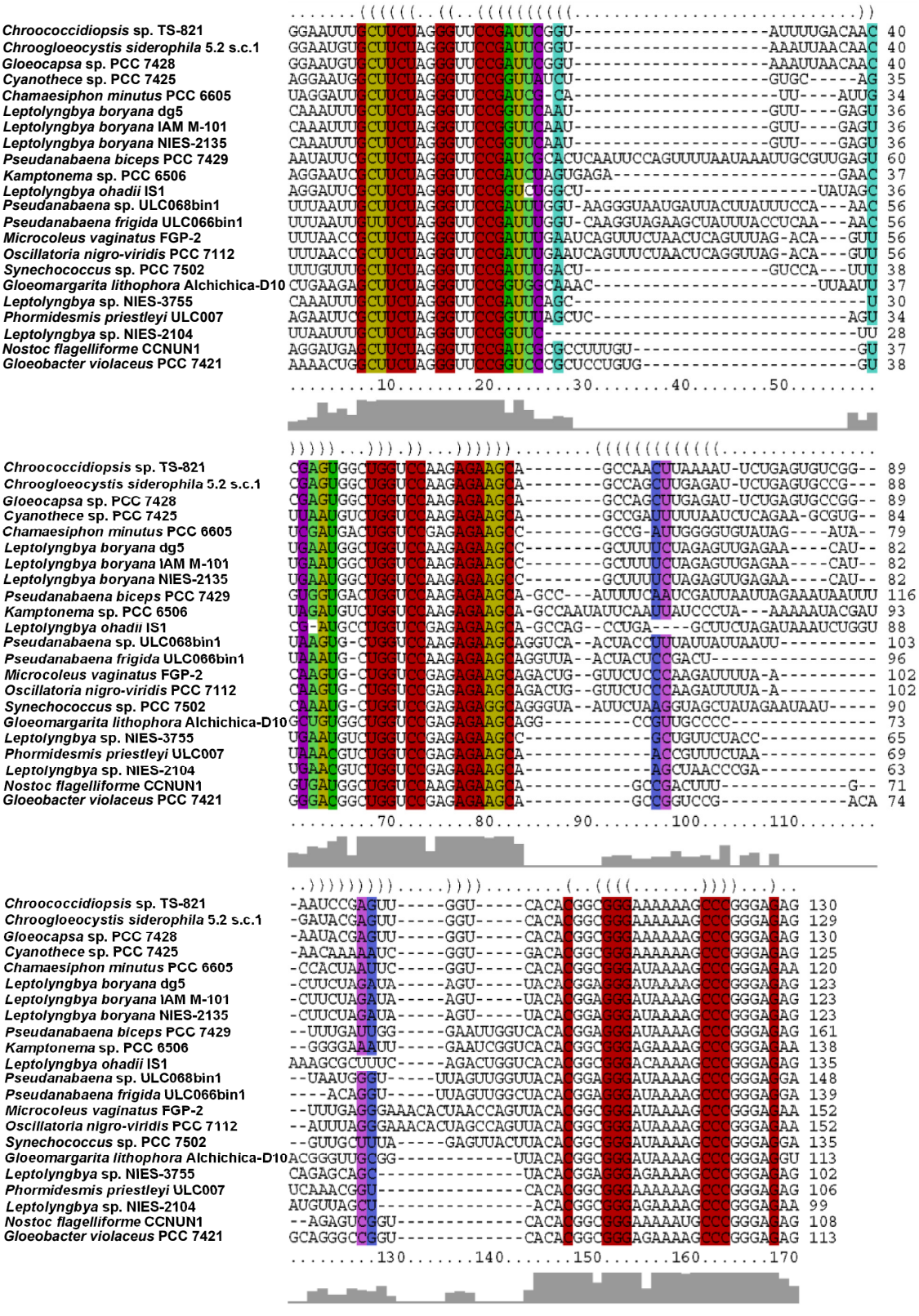
likelihood method based on 16S rRNA sequences, encompassing representative cyanobacterial and other bacterial species harboring the carboxylation pathway gene clusters. Branches are color-coded by taxonomic group as shown in the legend. *Novimethylophilus kurashikiensis* La2-4 was used as the outgroup. Nodes with bootstrap support values ≥ 70% are marked with black dots.



Supplementary Fig. 12 Phylogenetic analysis based on comparative analysis of guanidine carboxylase proteins from representative cyanobacteria and other bacteria. A maximum-likelihood phylogenetic tree was constructed based on concatenated CgdA and CgdB protein sequences, encompassing cyanobacterial and other bacterial species that harbor the carboxylation pathway gene clusters. Branch colors indicate the respective taxonomic groups, as shown in the legend. *Thioalkalivibrio sulfidiphilus* HL-EbGr7 was set as the outgroup. Black dots denote nodes with bootstrap support ≥ 70%.

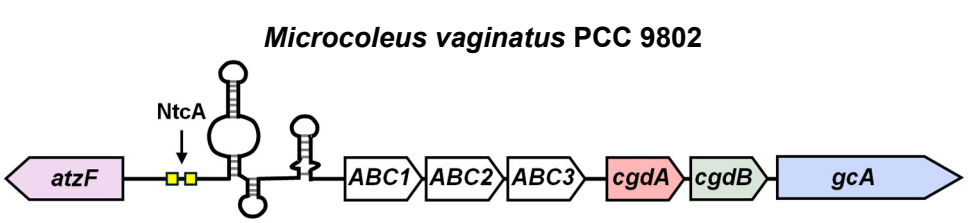


Supplementary Fig. 13 Congruence analysis between the 16S rRNA and guanidine-degradation gene datasets. Maximum-likelihood phylogenetic trees based on the 16S rRNA gene (**a**) and the concatenated sequences of CgdA and CgdB (**b**). Black dots denote nodes with bootstrap support ≥ 70%.



Supplementary Fig. 14 Sequence comparison of the guanidine-I riboswitch across cyanobacteria. The multiple sequence alignment was conducted using the LocARNA webserver[5].

a

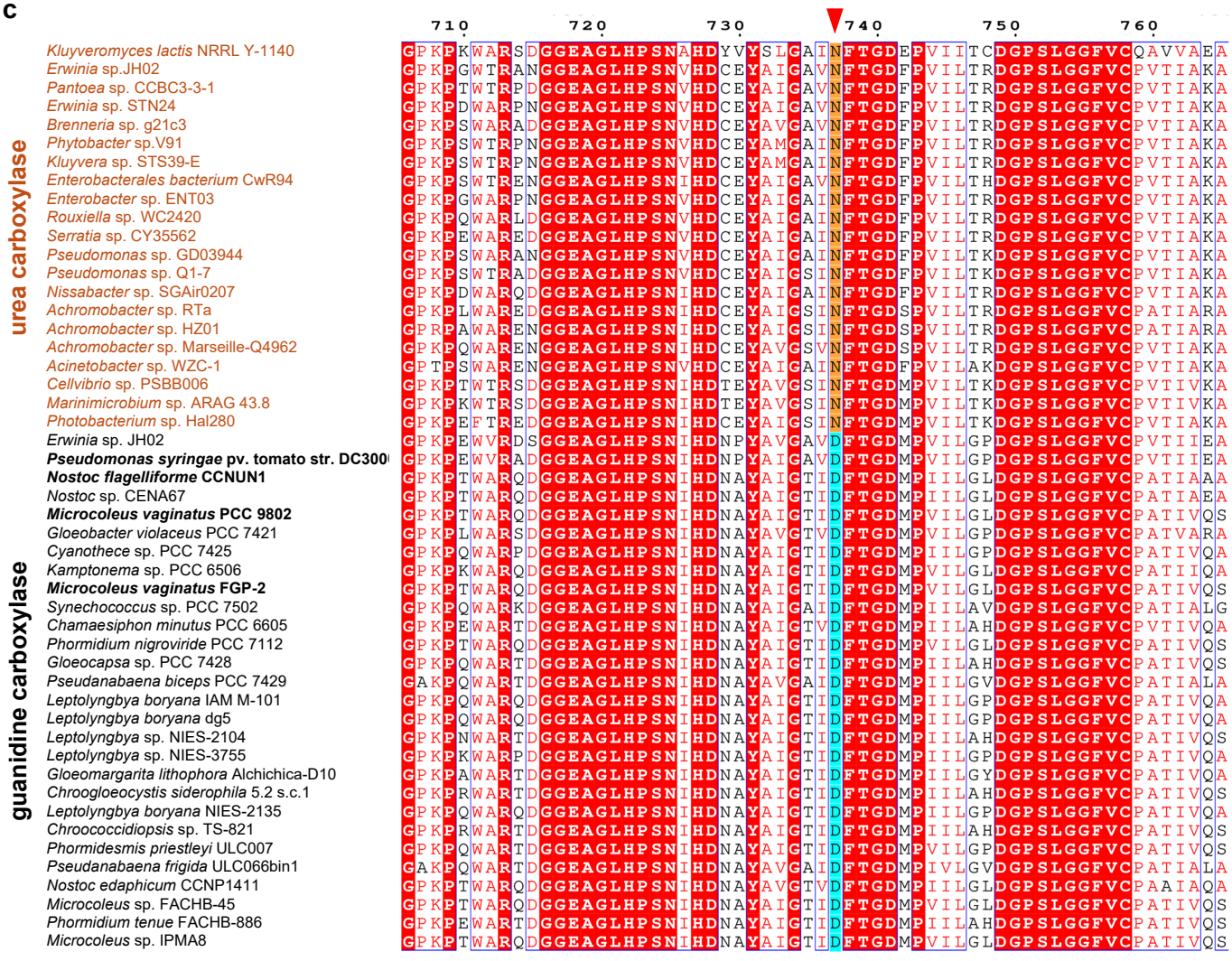


b

NtcA binding site **−10 box** **TSS**

GTAACATTTTACACacattttcgcatcaagtttgctgttaccgtttgaaacgtagcaatctagtaagattgcaaattacagtAtgcttacctttcaa
aaagctaagtgcattgctgttactgacagtctcgaataactccttgggttgcaaacaggacttatacagacaaaaaccgaaaaaattttaaccgcttc
tagggttccgatttgaatcagtttctaactcagtttagacagttcaagtgctggtccgagagaagcagactggttctcccaagattttaatttg
agggaaacactaaccagttacacggcgggataaaagcccgggagaaacgcgatcggtcatcaagactcggttcgttagctctcgggctttat
ccttgtatatcaggggatttgcggaagctaaggaagccaaagaaggtgcaaaagctcgatcgacatattcacctctaaccagcaaccagaatctg
ctATG...

c



Supplementary Fig. 15 Guanidine-I riboswitch-regulated guanidine carboxylase pathway in *Microcoleus vaginatus* PCC 9802. a Schematic of the genomic region encoding CgdA, CgdB, GcA, AtzF, and ABC-type guanidine importer in *Microcoleus vaginatus* PCC 9802. A potential NtcA-binding site is indicated. **b** Sequence of the promoter *cis*-acting elements, with the guanidine-I riboswitch nucleotide highlighted in blue. **c** Sequence alignment highlighting an amino acid substitution in the carboxyltransferase active site between urea and guanidine carboxylases. Alignments were generated with MAFFT and visualized using ESPript. Strains discussed in this study are shown in bold.

References

1. Rajeev, L. *et al.* Dynamic cyanobacterial response to hydration and dehydration in a desert biological soil crust. *ISME J.* **7**, 2178–2191 (2013).
2. Oren, N., Raanan, H., Murik, O., Keren, N. & Kaplan, A. Dawn illumination prepares desert cyanobacteria for dehydration. *Curr. Biol.* **27**, R1056–R1057 (2017).
3. van Kempen, M. *et al.* Fast and accurate protein structure search with Foldseek. *Nat. Biotechnol.* **42**, 243–246 (2024).
4. Fukuda, H. *et al.* Two reactions are simultaneously catalyzed by a single enzyme: The arginine-dependent simultaneous formation of two products, ethylene and succinate, from 2-oxoglutarate by an enzyme from *Pseudomonas syringae*. *Biochem. Biophys. Res. Commun.* **188**, 483–489 (1992).
5. Smith, C., Heyne, S., Richter, A. S., Will, S. & Backofen, R. Freiburg RNA Tools: a web server integrating INTARNA, EXPARNA and LOCARNA. *Nucleic Acids Res.* **38**, 373–377 (2010).