

Supplementary material for Blaise Li, João S Lopes, Peter G Foster, T Martin Embley, and Cymon J Cox, 2014 "Compositional biases among synonymous substitutions cause conflict between gene and protein trees for plastid origins"

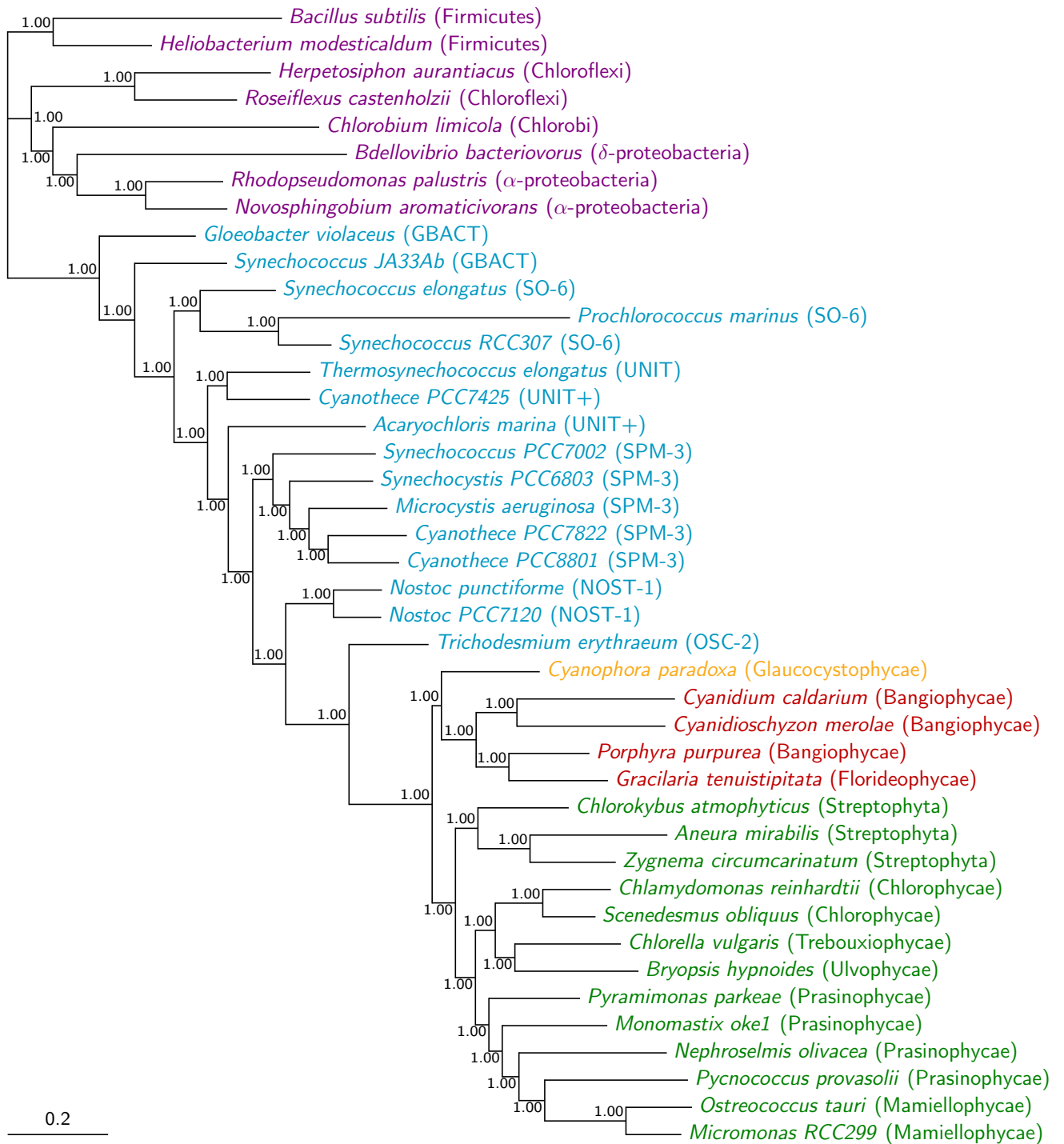


Figure S1: MCMC Bayesian analysis of the protein-coding gene dataset ‘cg75_stat’ assuming homogeneous composition. 50% majority-rule consensus tree of the ‘post-burnin’ MCMC (composition homogeneous GTR+I+ Γ) sampled trees. The proportion of samples to discard as ‘burnin’ was estimated by eye by looking for a plateauing of the likelihoods of the trees sampled by the Markov chain. Numbers at the nodes are their posterior probabilities, and branch lengths represent the number of nucleotide substitutions per site. Colours indicate taxonomic group (see legend Figure 1).

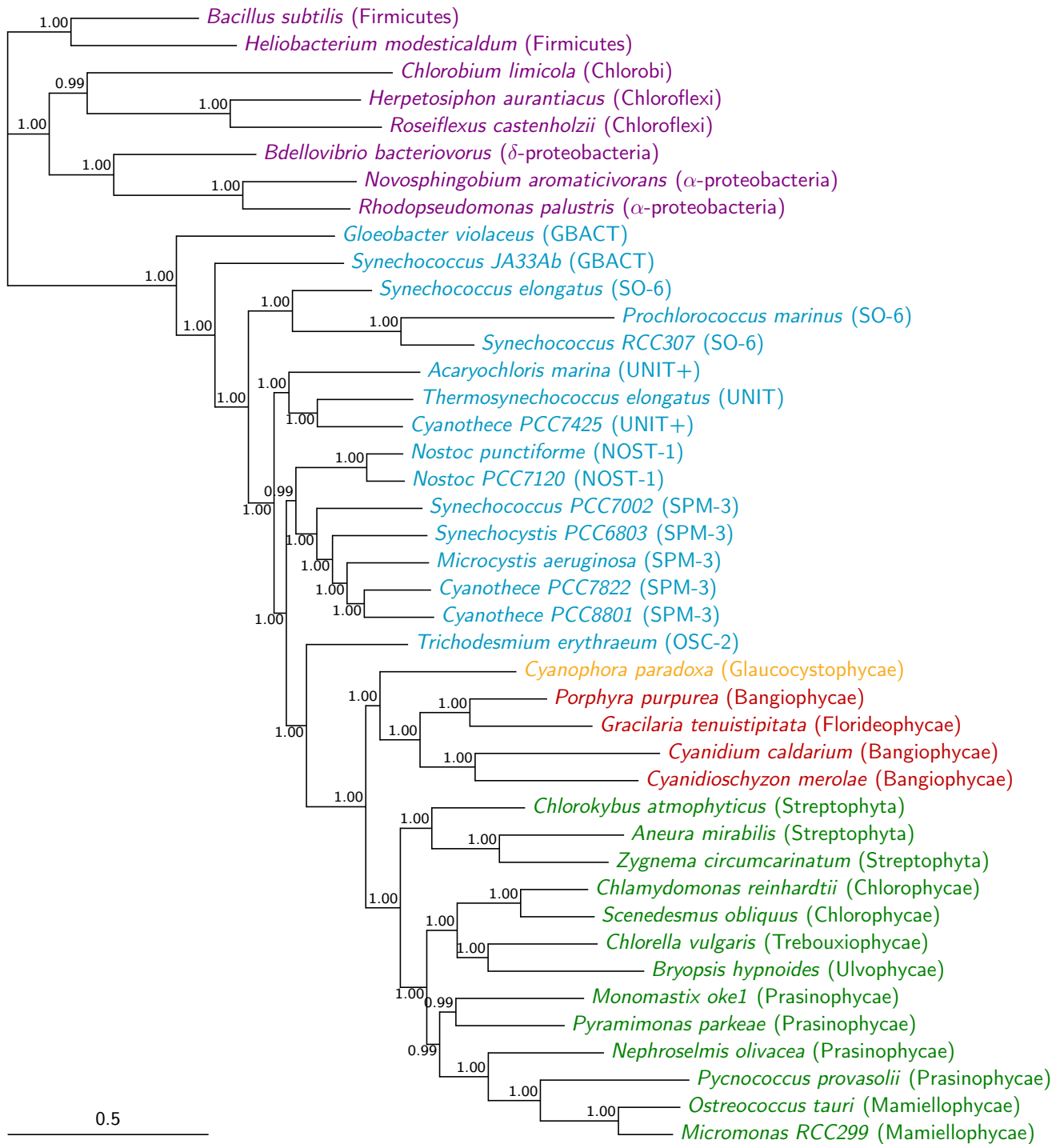


Figure S2: MCMC Bayesian analysis of the protein-coding gene dataset ‘cg75_NDCH’ under the NDCH model. 50% majority-rule consensus tree of the ‘post-burnin’ MCMC (node-heterogeneous GTR+I+ Γ +2CV) sampled trees. The proportion of samples to discard as ‘burnin’ was estimated by eye by looking for a plateauing of the likelihoods of the trees sampled by the Markov chain. Numbers at the nodes are their posterior probabilities, and branch lengths represent the number of nucleotide substitutions per site. Colours indicate taxonomic group (see legend Figure 1).

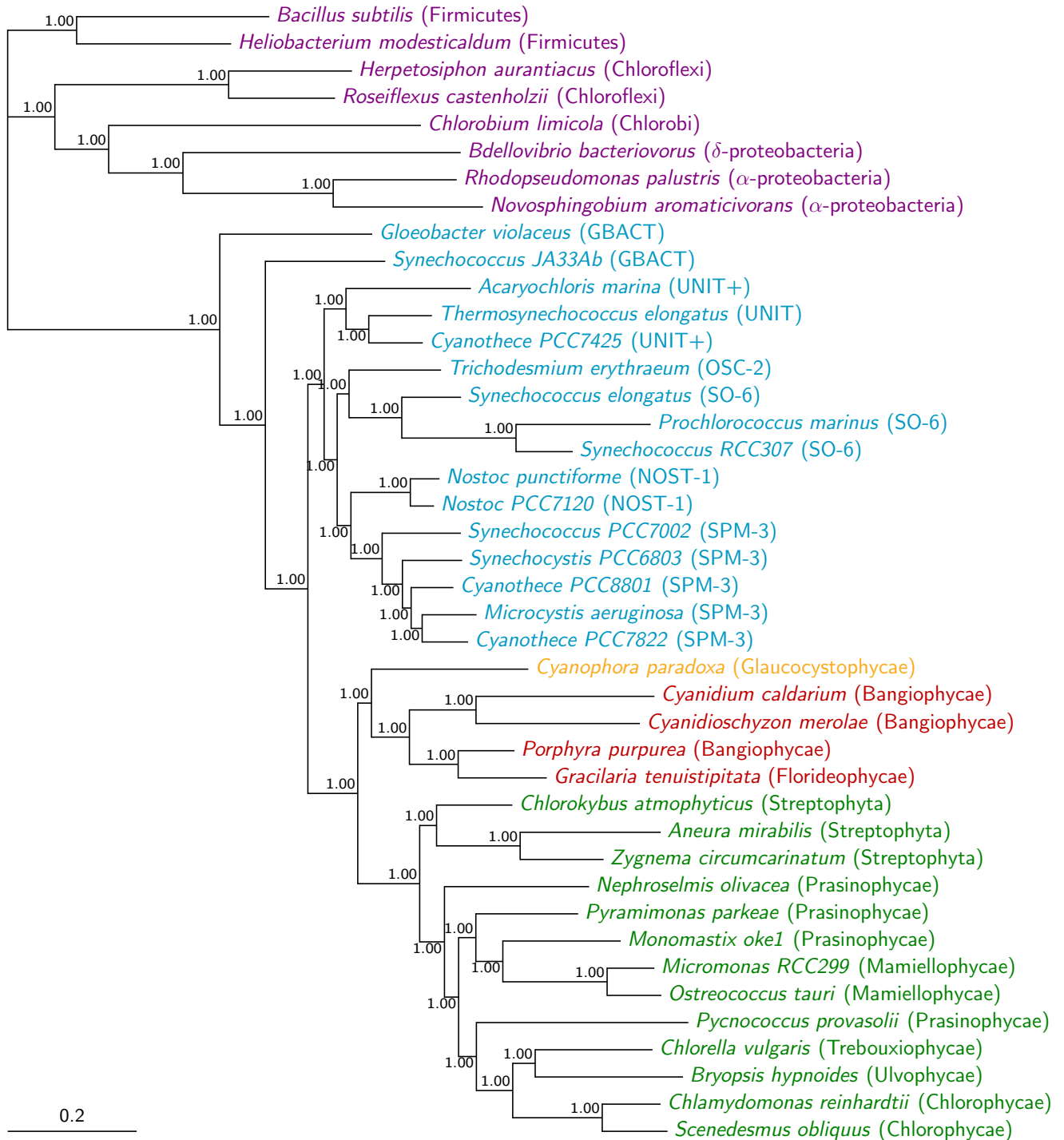


Figure S3: MCMC Bayesian analysis of the protein dataset ‘cp75_stat’ assuming homogeneous composition. 50% majority-rule consensus tree of the ‘post-burnin’ MCMC (composition homogeneous LG+I+ Γ) sampled trees. The proportion of samples to discard as ‘burnin’ was estimated by eye by looking for a plateauing of the likelihoods of the trees sampled by the Markov chain. Values above branches are posterior probabilities, and branch lengths represent the number of amino-acid substitutions per site. Colours indicate taxonomic group (see legend Figure 1).

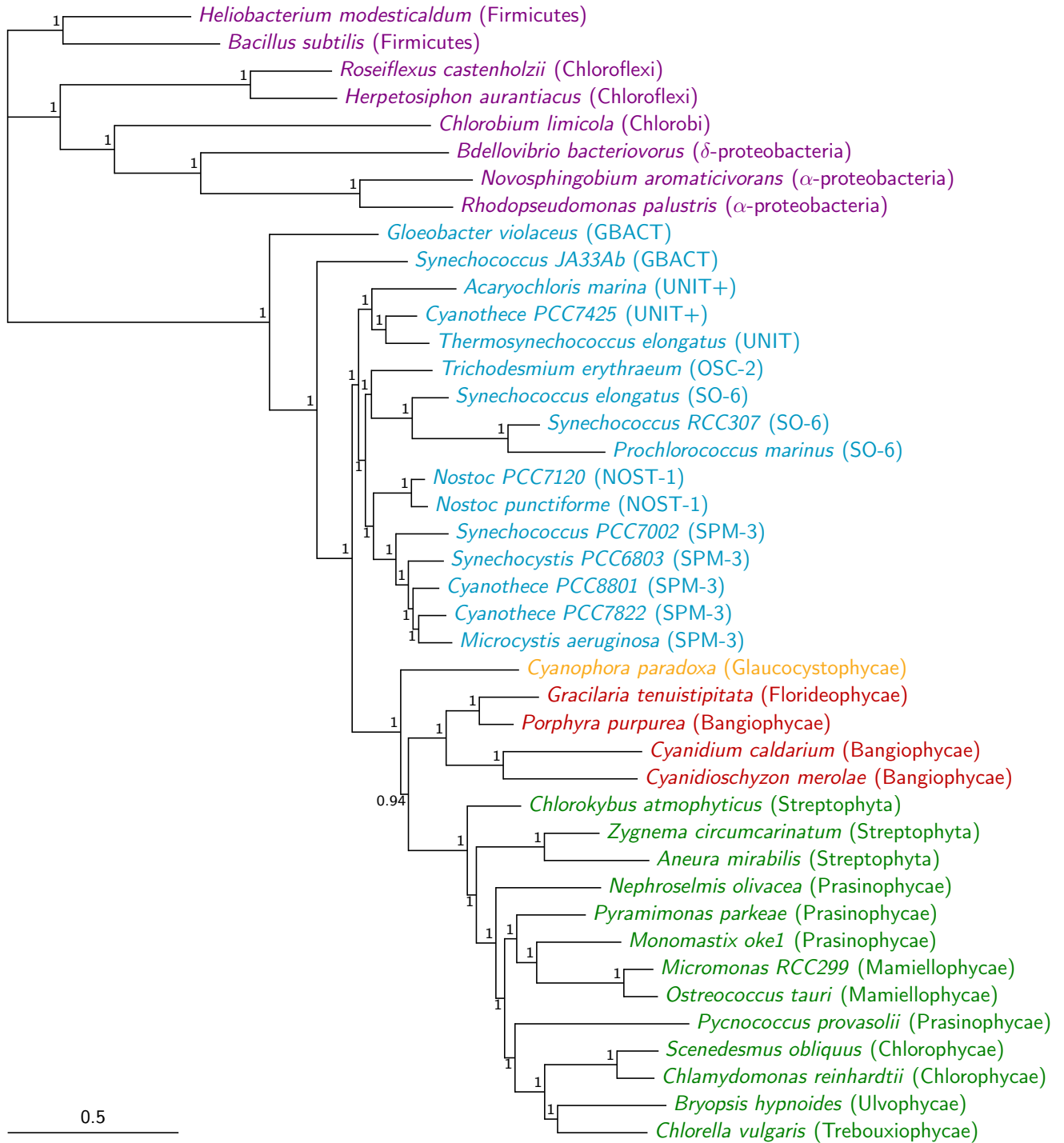


Figure S4: MCMC Bayesian analysis of the protein dataset ‘cp75_CAT’ under the CAT model
 50% majority-rule consensus tree of the ‘post-burnin’ MCMC (site-heterogeneous CAT+LG+ Γ) sampled trees. The automatic stopping criterion implemented in Phylobayes considered a ‘burnin’ of $\frac{1}{5}$ of the samples, so this was the proportion of trees discarded as ‘burnin’. Numbers at the nodes are their posterior probabilities, and branch lengths represent the number of amino-acid substitutions per site. Colours indicate taxonomic group (see legend Figure 1).

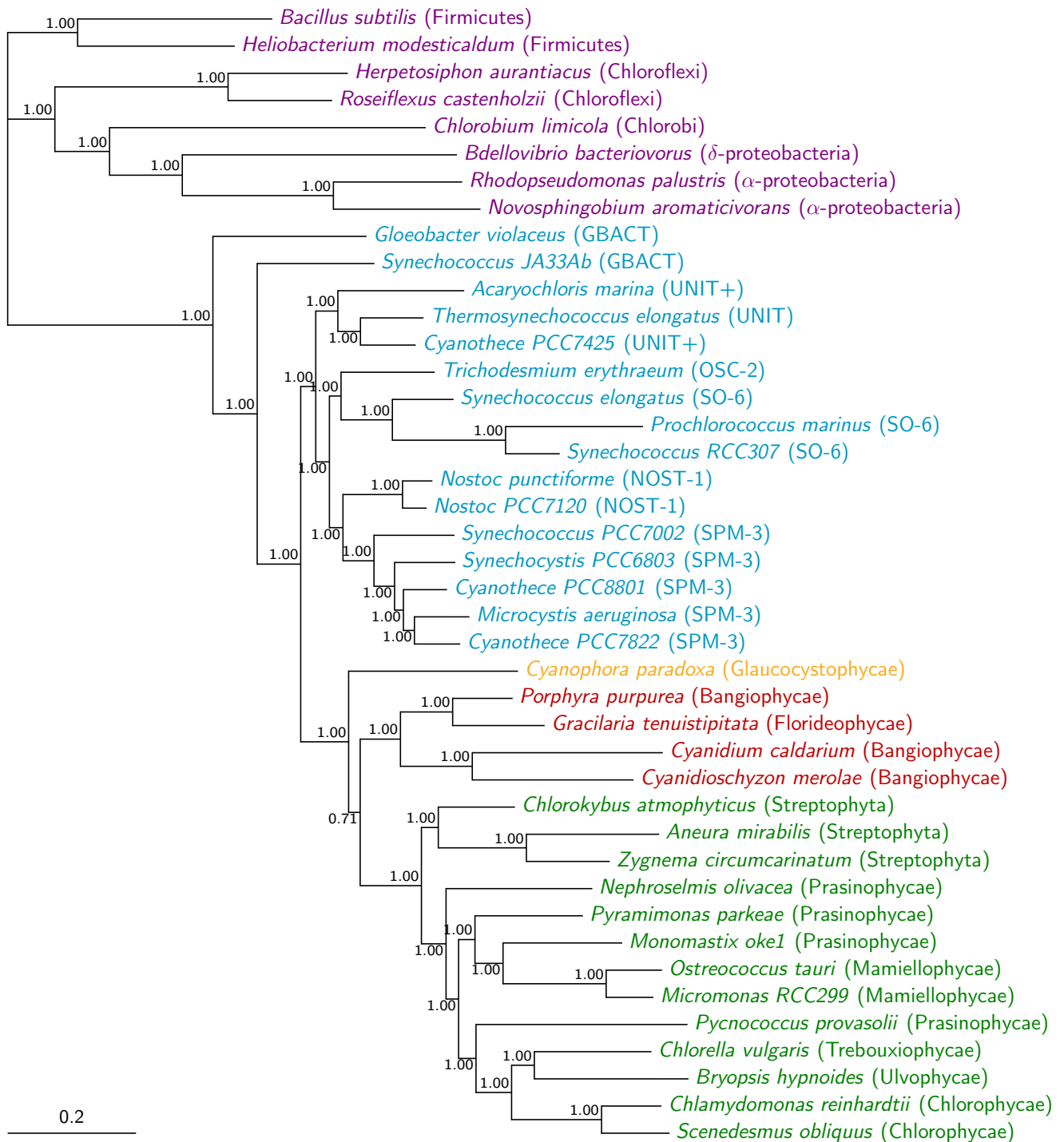


Figure S5: MCMC Bayesian analysis of the protein dataset ‘cp75_NDCH’ under the NDCH model. 50% majority-rule consensus tree of the ‘post-burnin’ MCMC (node-heterogeneous LG+I+Γ+30CV) sampled trees. The proportion of samples to discard as ‘burnin’ was estimated by eye by looking for a plateauing of the likelihoods of the trees sampled by the Markov chain. It should be noted that the 30 composition vectors were not enough to pass the χ^2 test and that another MCMC analysis under the same conditions yielded a tree with a slightly lower likelihood that differed from the present tree by the grouping of red algae with glaucocystophytes. But all topologies obtained displayed the sister-group relationship between core-cyanobacteria and plastids. Numbers at the nodes are their posterior probabilities, and branch lengths represent the number of nucleotide substitutions per site. Colours indicate taxonomic group (see legend Figure 1).

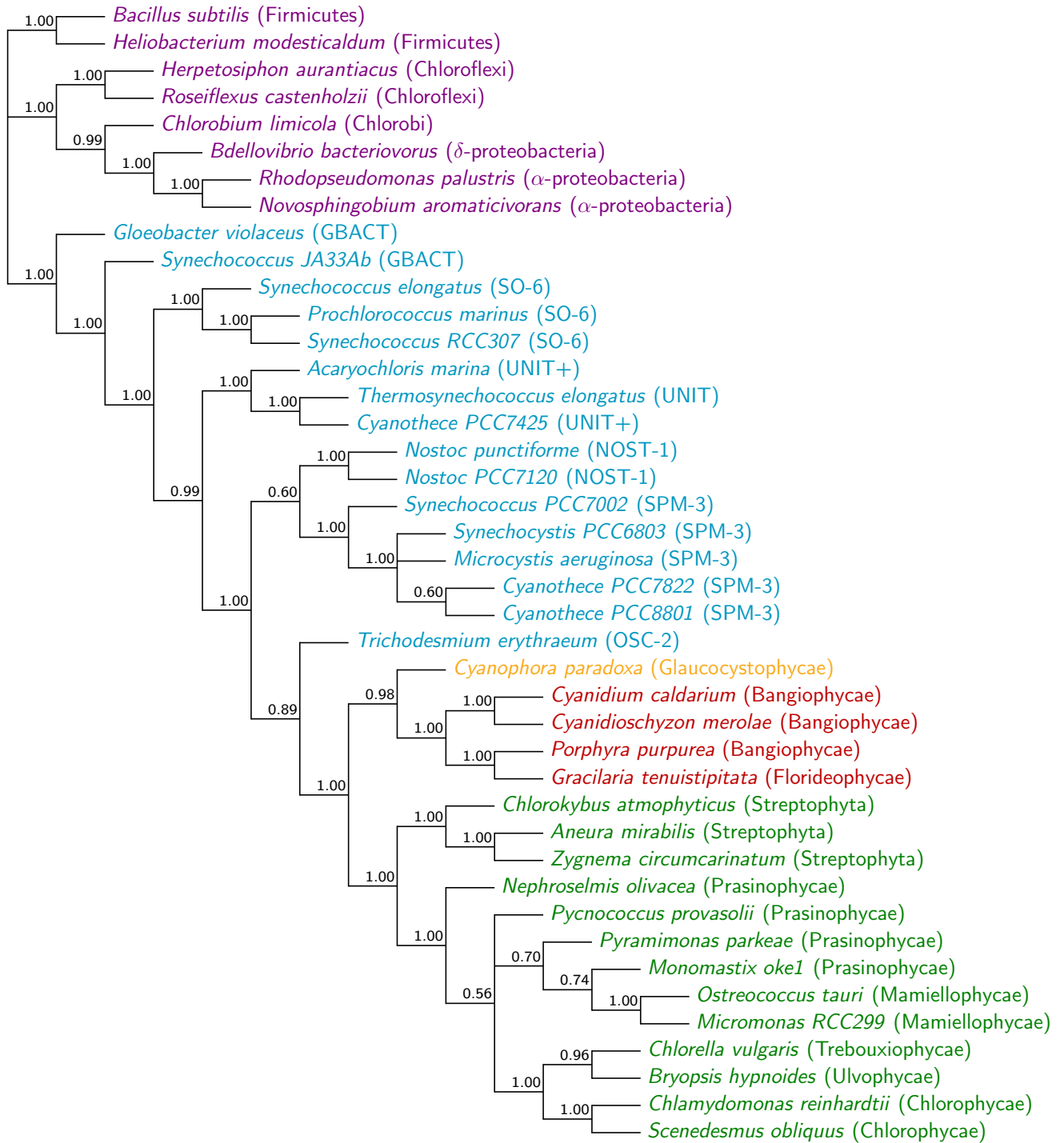


Figure S6: ML bootstrap analysis of the recoded protein-coding gene dataset ‘cg75_degen3’. 50% majority-rule consensus tree of 200 ML (GTR+I+Γ) bootstrap trees. Values above branches are bootstrap proportions. Dataset characteristics are indicated in Table S3. The values above the branches represent their bootstrap proportions. Colours indicate taxonomic group (see legend Figure 1).

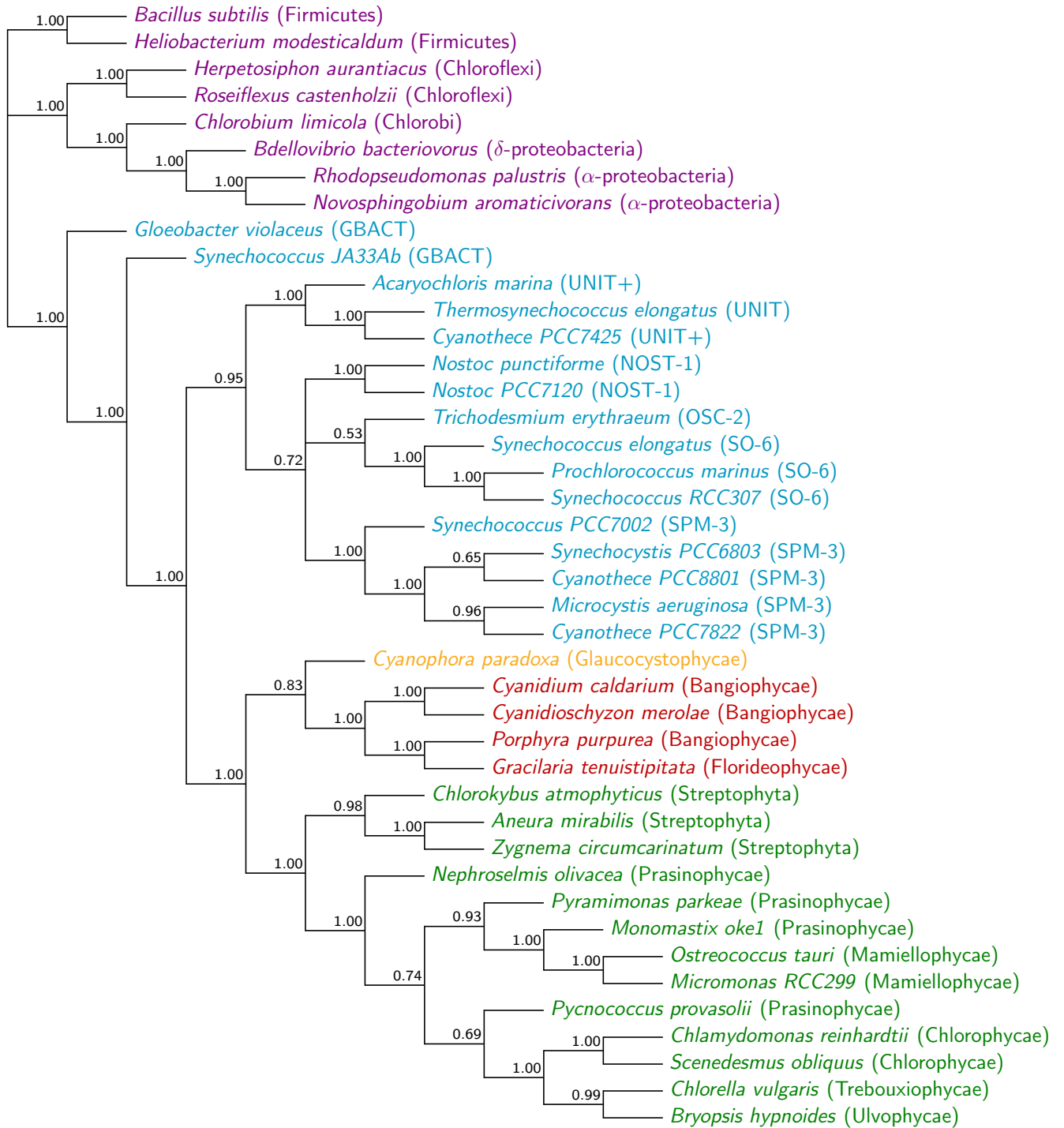


Figure S7: ML bootstrap analysis of the recoded protein-coding gene dataset 'cg75_degenLR3'. 50% majority-rule consensus tree of 200 ML (GTR+I+Γ) bootstrap trees. Values above branches are bootstrap proportions. Dataset characteristics are indicated in Table S3. The values above the branches represent their bootstrap proportions. Colours indicate taxonomic group (see legend Figure 1).

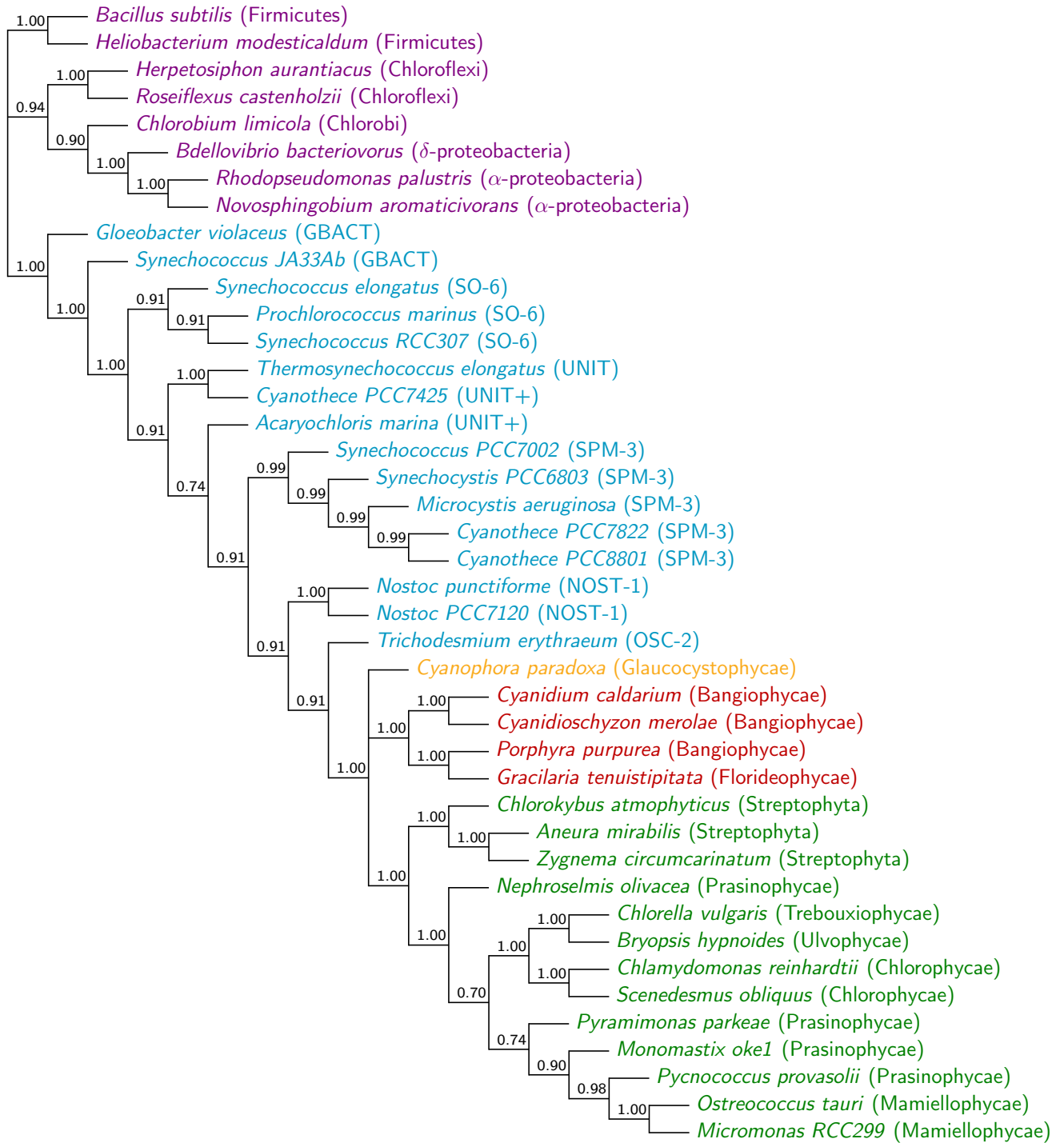


Figure S8: ML bootstrap analysis of the recoded protein-coding gene dataset 'cg75_degen1LR'. 50% majority-rule consensus tree of 200 ML (GTR+I+Γ) bootstrap trees. Values above branches are bootstrap proportions. Dataset characteristics are indicated in Table S3. The values above the branches represent their bootstrap proportions. Colours indicate taxonomic group (see legend Figure 1).

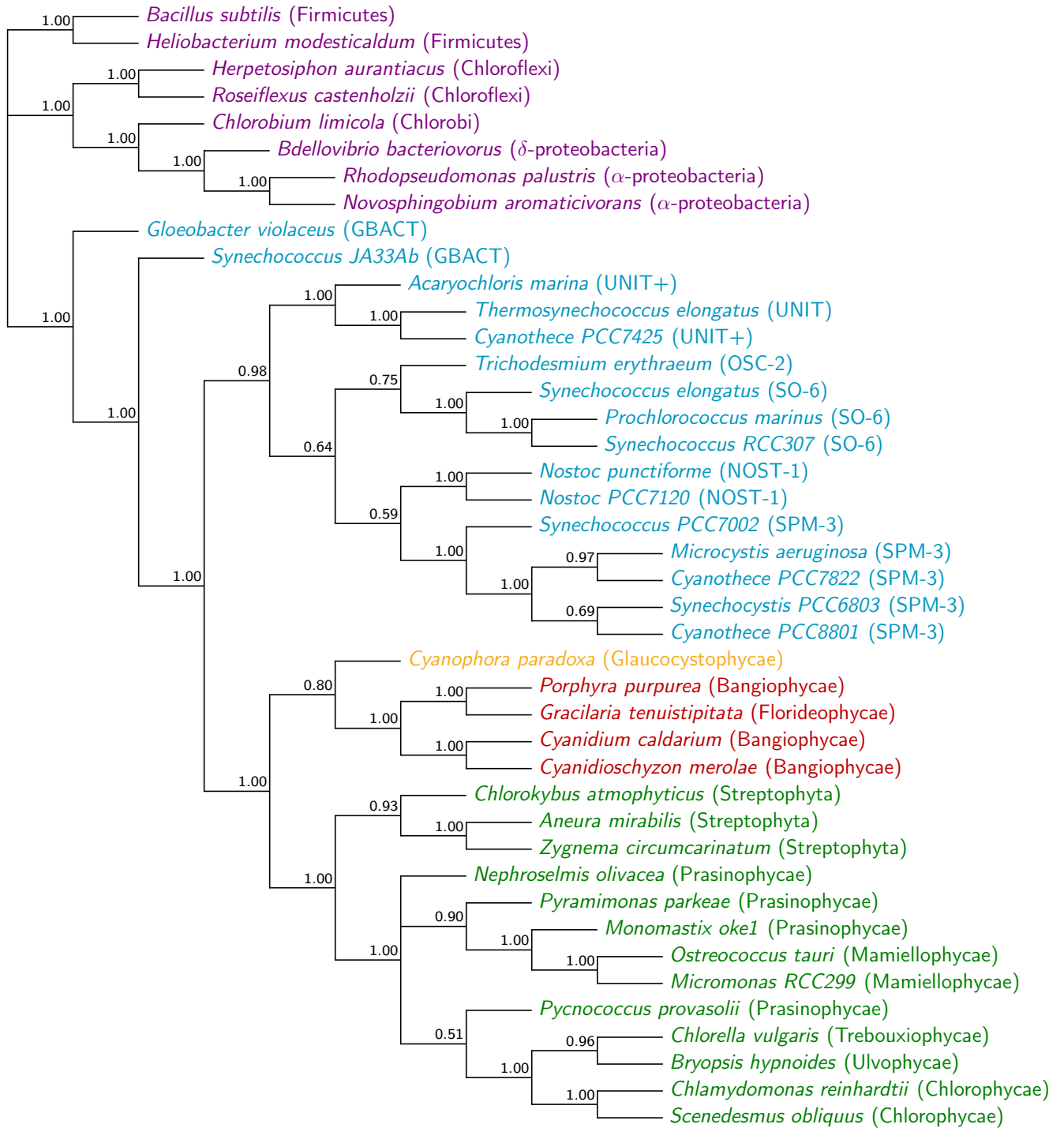


Figure S9: ML bootstrap analysis of the recoded protein-coding gene dataset ‘cg75_degen’. 50% majority-rule consensus tree of 200 ML (GTR+I+ Γ) bootstrap trees. Values above branches are bootstrap proportions. Dataset characteristics are indicated in Table S3. The values above the branches represent their bootstrap proportions. Colours indicate taxonomic group (see legend Figure 1).

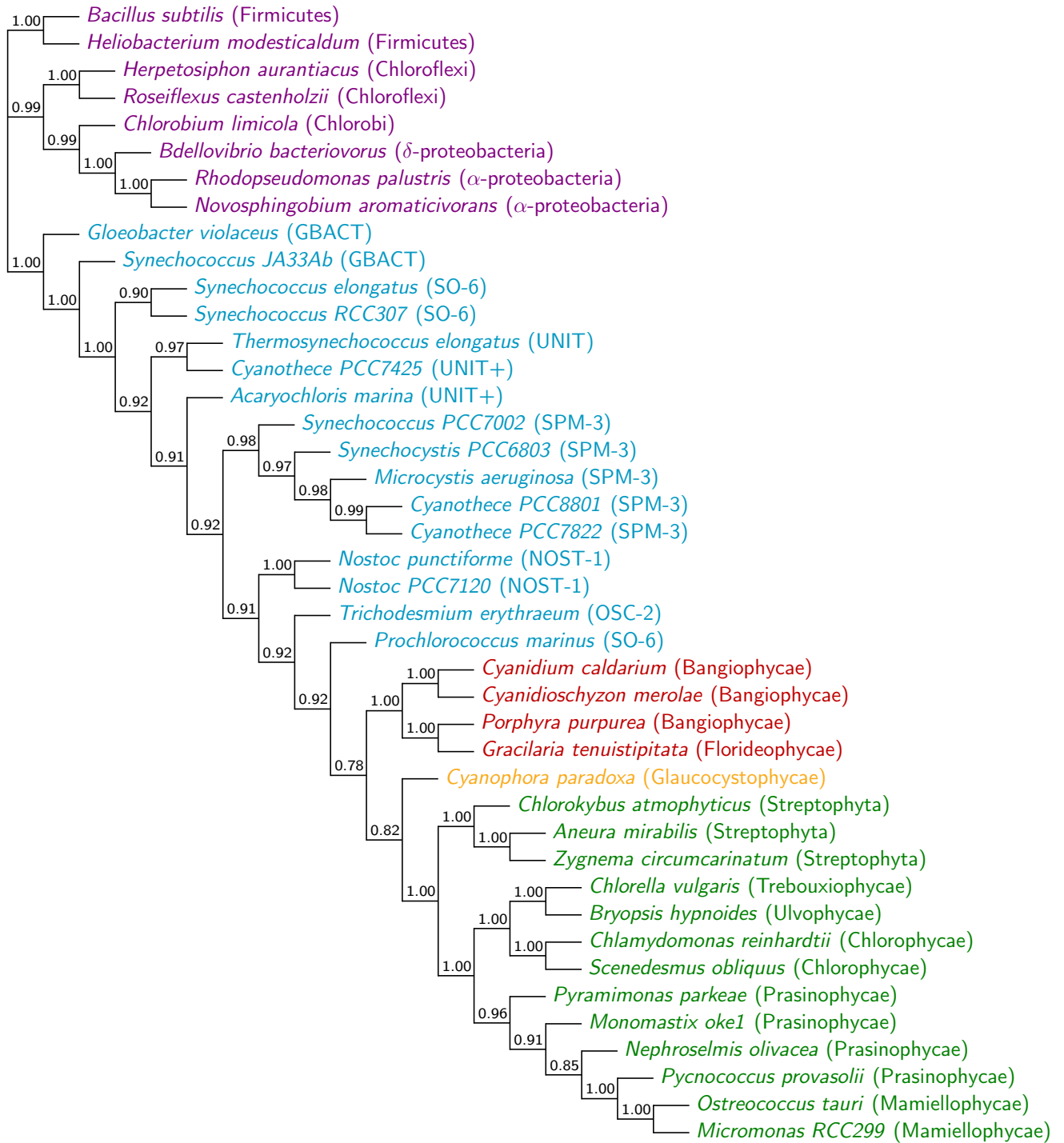


Figure S10: ML bootstrap analysis of the recoded protein-coding gene dataset ‘cg75_degen12S’. 50% majority-rule consensus tree of 200 ML (GTR+I+Γ) bootstrap trees. Values above branches are bootstrap proportions. Dataset characteristics are indicated in Table S3. The values above the branches represent their bootstrap proportions. Colours indicate taxonomic group (see legend Figure 1).

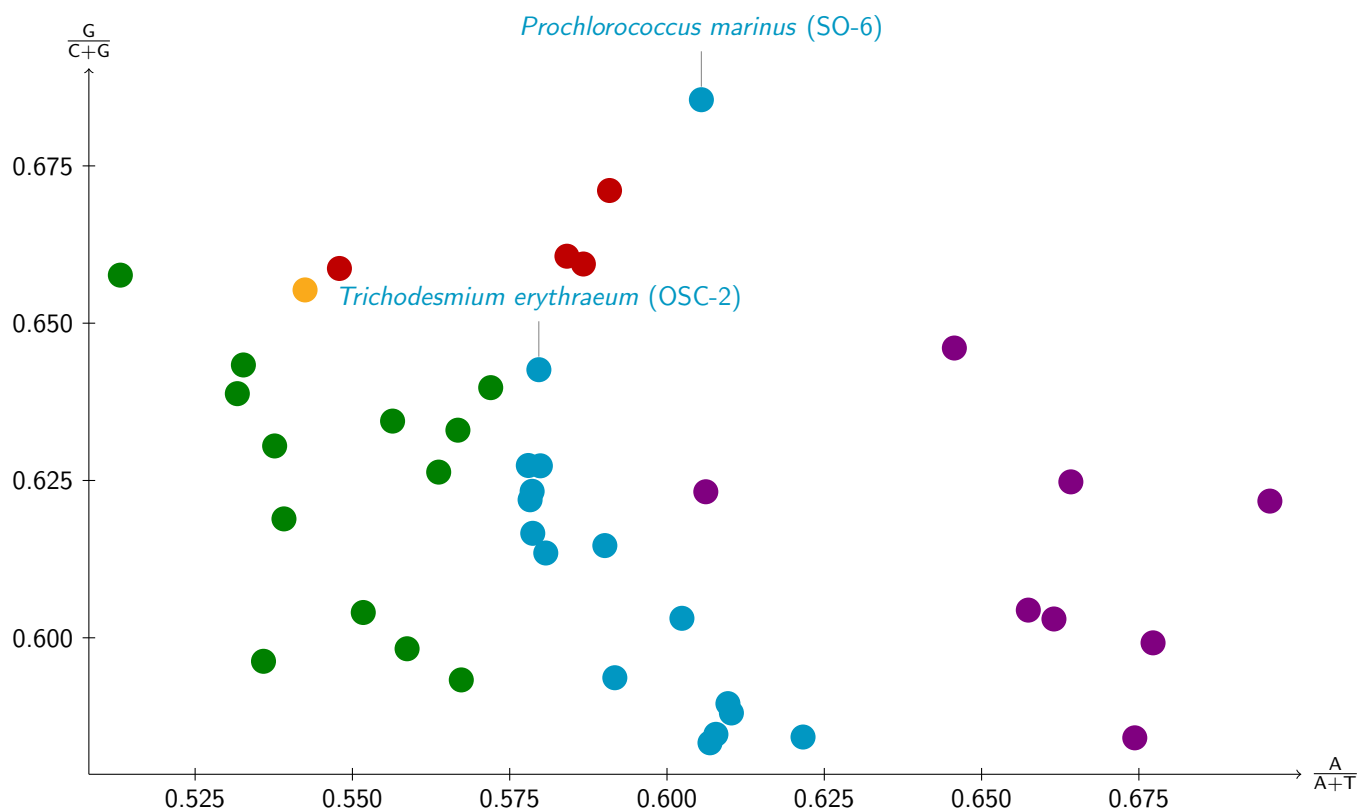


Figure S11: Nucleotide composition at first codon positions. Composition at first codon positions in the concatenated nucleotide dataset ‘cg75’. Each concatenated sequence is represented by its $\frac{A}{A+T}$ and $\frac{G}{G+C}$ ratios at first codon positions. The colours used for the dots indicate the groups to which the sequences belong (see Table S1). Only *Prochlorococcus marinus* and *Trichodesmium erythraeum* are labelled. G+C proportions are displayed on Figure 3.

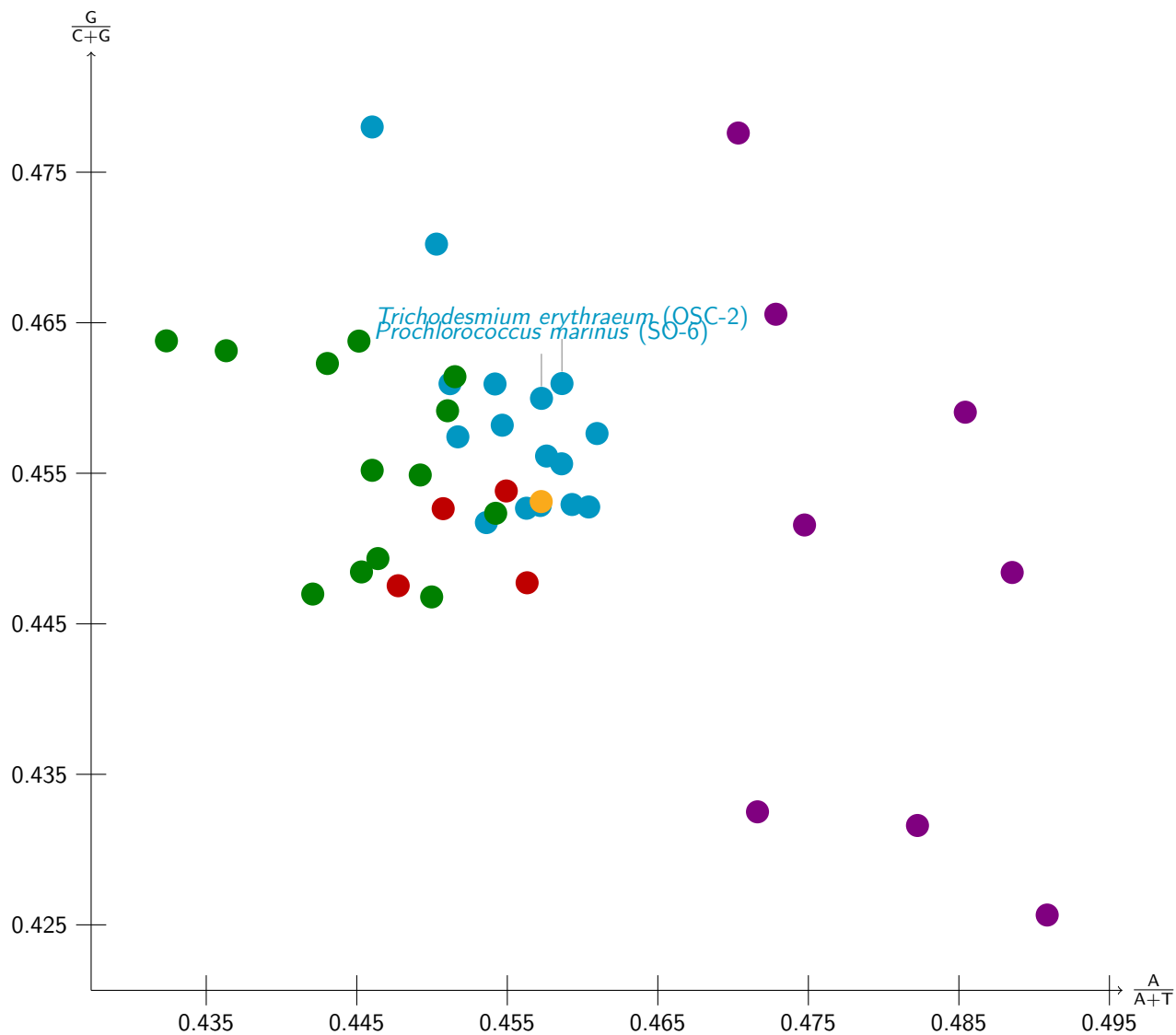


Figure S12: Nucleotide composition at second codon positions. Composition at second codon positions in the concatenated nucleotide dataset ‘cg75’. Each concatenated sequence is represented by its $\frac{A}{A+T}$ and $\frac{G}{G+C}$ ratios at second codon positions. The colours used for the dots indicate the groups to which the sequences belong (see Table S1). Only *Prochlorococcus marinus* and *Trichodesmium erythraeum* are labelled. G+C proportions are displayed on Figure 3.

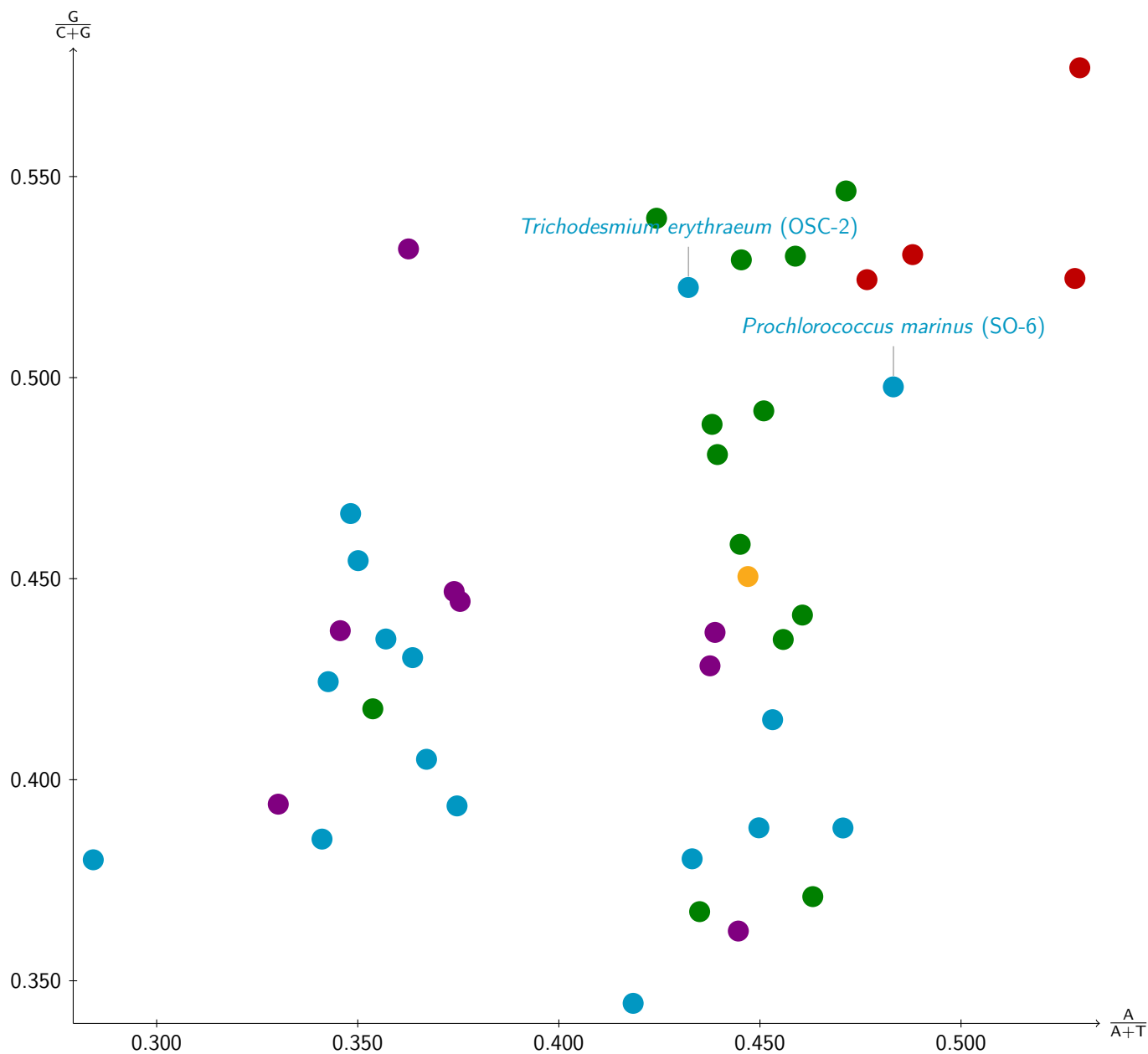


Figure S13: Nucleotide composition at third codon positions. Composition at third codon positions in the concatenated nucleotide dataset 'cg75'. Each concatenated sequence is represented by its $\frac{A}{A+T}$ and $\frac{G}{G+C}$ ratios at third codon positions. The colours used for the dots indicate the groups to which the sequences belong (see Table S1. Only *Prochlorococcus marinus* and *Trichodesmium erythraeum* are labelled. G+C proportions are displayed on Figure 3.

Table S1: Data matrix properties with respect to potential phylogenetic signal from synonymous codons. Phylogenetic signal resulting from codon synonymy as properties of the different constructed data matrices. Present ("yes") or absent ("no"). '1LR' designates the signal associated with first codon position synonymy among leucine and arginine codons. '12S' designates the signal associated with first and second codon position synonymy among serine codons. '3' designates the signal associated with third codon position synonymy among all codons families. The 'degen*' analyses are performed under standard maximum likelihood, but some of the signal associated with codon synonymy is suppressed by recoding some of the codon positions where codon degeneracy exist. The amino-acid dataset ('cp75') does not contain any of these signals because synonymous codons are represented by a common amino-acid. 'CAT' designates the site-heterogeneous model implemented in Phylobayes. 'NDCH' designates the non-stationary composition model implemented in P4.

Analysis	Codon-synonymy signal		
	1LR	12S	3
cg75_mlboot	yes	yes	yes
cp75_mlboot	no	no	no
cp75_stat	no	no	no
cp75_CAT	no	no	no
cp75_NDCH	no	no	no
cg75_stat	yes	yes	yes
cg75_NDCH	yes	yes	yes
cg75_degen3	yes	yes	no
cg75_degen	no	no	no
cg75_degenLR3	no	yes	no
cg75_degen1LR	no	yes	yes
cg75_degen12S	yes	no	yes

Table S2: Amino acid and codon-family variant exchange rates. The data were recoded so that each codon family of Arg, Leu, and Ser was individually distinguished: AGR: Arg_A, CGN: Arg_C, TTR: Leu_T, CTN: Leu_C, AGY: Ser_{AG}, TCN: Ser_{TC}. Instantaneous substitution rates in a 23 character state amino-acid model were estimated from a P4 MCMC chain using a GTR+I+ Γ model with fixed topology, α parameter of the Γ distribution, and proportion of invariable sites, the values of which were taken from the optimal ML results obtained by RAxML under a CPREV+I+ Γ model on the standard 20 state amino-acid data.

exchange			rate
arginineA	↔	arginineC	20061.9
leucineC	↔	leucineT	15275.5
serineAG	↔	serineTC	4393.5
lysine	↔	arginineA	3136.8
asparagine	↔	serineAG	2783.9
threonine	↔	serineAG	2262.4
isoleucine	↔	valine	2015.7
aspartate	↔	glutamate	1804.1
glutamate	↔	glutamine	1786.7
aspartate	↔	asparagine	1505.3
lysine	↔	glutamine	1500.6
alanine	↔	serineTC	1473.2
histidine	↔	asparagine	1395.8
histidine	↔	glutamine	1363.5
phenylalanine	↔	tyrosine	1268.1
histidine	↔	tyrosine	1249.0
leucineC	↔	methionine	1047.8
threonine	↔	serineTC	955.8
cysteine	↔	serineTC	923.6
leucineC	↔	isoleucine	857.7
asparagine	↔	glutamine	842.3
lysine	↔	arginineC	832.5
isoleucine	↔	methionine	824.7
leucineT	↔	methionine	767.6
cysteine	↔	serineAG	729.8
glycine	↔	serineAG	666.5
asparagine	↔	threonine	654.7
arginineA	↔	glutamine	644.7
glutamine	↔	arginineC	611.7
lysine	↔	asparagine	574.8
isoleucine	↔	leucineT	564.8
cysteine	↔	valine	549.8
aspartate	↔	serineAG	522.0
glutamine	↔	serineAG	516.0
cysteine	↔	threonine	490.5
alanine	↔	cysteine	485.5
methionine	↔	glutamine	473.1
phenylalanine	↔	leucineT	462.8
threonine	↔	valine	451.7
proline	↔	serineTC	450.8
histidine	↔	arginineC	450.4
arginineA	↔	serineAG	450.2

continued on next page

continued from previous page

exchange		rate
alanine	↔ threonine	420.8
methionine	↔ threonine	377.2
phenylalanine	↔ tryptophan	373.8
histidine	↔ serineAG	373.0
glutamine	↔ threonine	355.9
leucineC	↔ phenylalanine	325.6
lysine	↔ threonine	310.0
glutamate	↔ serineAG	306.8
alanine	↔ serineAG	306.0
asparagine	↔ arginineA	296.5
glutamate	↔ lysine	293.2
alanine	↔ glycine	288.8
glycine	↔ asparagine	286.9
asparagine	↔ serineTC	286.1
alanine	↔ valine	286.0
cysteine	↔ tyrosine	280.4
lysine	↔ serineAG	280.0
methionine	↔ valine	278.4
isoleucine	↔ threonine	272.8
histidine	↔ arginineA	271.2
glutamine	↔ serineTC	268.4
phenylalanine	↔ methionine	267.3
alanine	↔ proline	250.6
tryptophan	↔ tyrosine	244.9
leucineT	↔ valine	236.2
alanine	↔ glutamine	233.2
cysteine	↔ phenylalanine	229.4
cysteine	↔ arginineA	225.5
leucineC	↔ valine	222.0
histidine	↔ lysine	219.0
arginineA	↔ threonine	216.8
alanine	↔ glutamate	214.2
aspartate	↔ glutamine	210.7
cysteine	↔ asparagine	209.6
aspartate	↔ histidine	207.5
cysteine	↔ leucineT	202.5
cysteine	↔ methionine	200.6
cysteine	↔ arginineC	199.1
arginineC	↔ serineAG	195.8
histidine	↔ serineTC	190.2
glutamate	↔ asparagine	173.5
glutamate	↔ threonine	172.4
phenylalanine	↔ isoleucine	162.8
cysteine	↔ histidine	162.1
proline	↔ threonine	161.5
leucineC	↔ glutamine	161.0
phenylalanine	↔ histidine	158.2
leucineT	↔ glutamine	153.7
histidine	↔ threonine	152.2
cysteine	↔ glycine	145.0
		<i>continued on next page</i>

continued from previous page

exchange		rate
arginineA	↔ tyrosine	142.5
asparagine	↔ tyrosine	142.1
leucineC	↔ cysteine	140.4
proline	↔ glutamine	134.1
leucineT	↔ serineTC	129.3
glutamate	↔ histidine	125.4
glutamine	↔ tyrosine	123.8
glutamate	↔ serineTC	119.5
histidine	↔ tryptophan	115.6
phenylalanine	↔ serineTC	112.1
leucineC	↔ histidine	111.4
cysteine	↔ proline	111.0
aspartate	↔ glycine	110.7
lysine	↔ serineTC	103.8
phenylalanine	↔ valine	101.3
alanine	↔ methionine	97.5
tyrosine	↔ serineTC	97.2
glycine	↔ arginineA	95.2
glycine	↔ serineTC	94.8
aspartate	↔ proline	93.7
glutamine	↔ valine	92.6
aspartate	↔ serineTC	92.6
lysine	↔ methionine	91.3
glutamate	↔ proline	90.9
cysteine	↔ tryptophan	90.7
lysine	↔ proline	87.7
arginineA	↔ tryptophan	85.1
leucineC	↔ arginineA	84.8
alanine	↔ asparagine	84.3
leucineT	↔ threonine	82.3
leucineT	↔ tryptophan	79.8
tryptophan	↔ serineTC	77.3
proline	↔ serineAG	77.1
alanine	↔ lysine	76.9
leucineC	↔ tryptophan	76.5
asparagine	↔ arginineC	76.2
histidine	↔ proline	75.7
serineAG	↔ tyrosine	74.6
leucineC	↔ threonine	70.3
methionine	↔ serineAG	67.6
leucineC	↔ proline	67.1
asparagine	↔ proline	66.8
glycine	↔ glutamine	66.8
aspartate	↔ threonine	65.9
alanine	↔ aspartate	65.1
methionine	↔ tryptophan	64.5
threonine	↔ tyrosine	60.4
glutamate	↔ arginineA	60.4
methionine	↔ tyrosine	60.3
alanine	↔ histidine	59.4

continued on next page

continued from previous page

exchange			rate
proline	↔	valine	58.8
glutamine	↔	tryptophan	58.1
leucineT	↔	arginineA	57.2
methionine	↔	serineTC	56.7
isoleucine	↔	arginineA	54.3
alanine	↔	leucineT	53.9
leucineC	↔	lysine	51.8
glutamate	↔	glycine	50.1
arginineC	↔	serineTC	49.9
glutamate	↔	valine	49.3
methionine	↔	asparagine	49.0
proline	↔	arginineC	48.7
threonine	↔	arginineC	48.4
leucineT	↔	tyrosine	48.3
methionine	↔	arginineC	48.2
leucineC	↔	tyrosine	48.0
isoleucine	↔	serineAG	47.3
methionine	↔	arginineA	47.1
glycine	↔	lysine	47.1
aspartate	↔	lysine	46.8
cysteine	↔	glutamine	45.7
histidine	↔	methionine	44.2
isoleucine	↔	asparagine	43.9
cysteine	↔	lysine	43.0
lysine	↔	valine	42.7
isoleucine	↔	tyrosine	41.3
leucineC	↔	arginineC	40.9
phenylalanine	↔	threonine	40.1
arginineA	↔	serineTC	39.9
arginineA	↔	proline	39.0
isoleucine	↔	glutamine	38.8
glycine	↔	proline	38.2
valine	↔	serineAG	37.8
arginineC	↔	tryptophan	37.6
alanine	↔	isoleucine	37.2
phenylalanine	↔	proline	36.4
methionine	↔	proline	36.1
alanine	↔	leucineC	35.6
leucineT	↔	proline	35.5
valine	↔	serineTC	35.4
alanine	↔	phenylalanine	34.3
arginineA	↔	valine	34.2
phenylalanine	↔	glutamine	34.2
cysteine	↔	isoleucine	34.0
phenylalanine	↔	arginineA	32.5
cysteine	↔	aspartate	32.5
histidine	↔	valine	32.3
phenylalanine	↔	asparagine	32.3
histidine	↔	leucineT	31.3
arginineC	↔	tyrosine	30.8

continued on next page

continued from previous page

exchange			rate
isoleucine	↔	proline	29.4
valine	↔	tyrosine	28.3
aspartate	↔	arginineA	27.8
aspartate	↔	tyrosine	27.5
leucineT	↔	asparagine	27.4
isoleucine	↔	lysine	27.2
glycine	↔	histidine	27.2
glutamate	↔	tyrosine	27.0
proline	↔	tyrosine	26.8
lysine	↔	tyrosine	23.7
glutamate	↔	arginineC	23.3
alanine	↔	tyrosine	23.0
tryptophan	↔	serineAG	22.8
alanine	↔	arginineC	22.0
alanine	↔	arginineA	22.0
glutamate	↔	leucineT	21.8
asparagine	↔	valine	21.3
leucineC	↔	serineAG	20.3
leucineT	↔	serineAG	19.7
cysteine	↔	glutamate	19.4
phenylalanine	↔	serineAG	19.1
glycine	↔	tryptophan	18.7
glycine	↔	methionine	18.2
glycine	↔	threonine	17.4
isoleucine	↔	tryptophan	17.2
glycine	↔	leucineT	16.7
leucineC	↔	asparagine	16.7
isoleucine	↔	arginineC	16.5
glutamate	↔	methionine	16.1
histidine	↔	isoleucine	15.6
valine	↔	tryptophan	15.2
glycine	↔	valine	14.4
glycine	↔	arginineC	13.2
phenylalanine	↔	glycine	12.9
leucineT	↔	lysine	12.0
proline	↔	tryptophan	11.9
leucineC	↔	glutamate	11.4
asparagine	↔	tryptophan	11.3
aspartate	↔	methionine	10.7
alanine	↔	tryptophan	9.6
aspartate	↔	arginineC	9.4
leucineC	↔	serineTC	9.0
aspartate	↔	tryptophan	8.4
glycine	↔	tyrosine	8.1
lysine	↔	tryptophan	7.9
phenylalanine	↔	arginineC	7.4
threonine	↔	tryptophan	6.4
glutamate	↔	isoleucine	6.3
leucineT	↔	arginineC	6.2
arginineC	↔	valine	5.2

continued on next page

continued from previous page

exchange			rate
phenylalanine	↔	lysine	5.1
glutamate	↔	tryptophan	5.0
aspartate	↔	phenylalanine	5.0
glutamate	↔	phenylalanine	4.6
aspartate	↔	valine	4.4
leucineC	↔	glycine	4.2
aspartate	↔	leucineT	4.1
aspartate	↔	isoleucine	2.9
glycine	↔	isoleucine	1.9
isoleucine	↔	serineTC	1.8
leucineC	↔	aspartate	1.7

Table S3: Taxonomic sampling. Taxa included in the study, and the NCBI GenBank accession number of their genome (or plastidial genome). The taxonomic groups to which they belong are indicated and represented in the same colour that was used in the figures.

Taxon	NCBI accession	Taxonomic group	Sub-group
<i>Bacillus subtilis</i>	NC_000964	Bacteria	Firmicutes
<i>Heliobacterium modesticaldum</i>	NC_010337	Bacteria	Firmicutes
<i>Herpetosiphon aurantiacus</i>	NC_009972	Bacteria	Chloroflexi
<i>Roseiflexus castenholzii</i>	NC_009767	Bacteria	Chloroflexi
<i>Chlorobium limicola</i>	NC_010803	Bacteria	Chlorobi
<i>Bdellovibrio bacteriovorus</i>	NC_005363	Bacteria	δ -proteobacteria
<i>Rhodospseudomonas palustris</i>	NC_008435	Bacteria	α -proteobacteria
<i>Novosphingobium aromaticivorans</i>	NC_007794	Bacteria	α -proteobacteria
<i>Gloeobacter violaceus</i>	NC_005125	Cyanobacteria	GBACT
<i>Synechococcus JA33Ab</i>	NC_007775	Cyanobacteria	GBACT
<i>Acaryochloris marina</i>	NC_009925	Cyanobacteria	UNIT+
<i>Thermosynechococcus elongatus</i>	NC_004113	Cyanobacteria	UNIT
<i>Cyanothece PCC7425</i>	NC_011884	Cyanobacteria	UNIT+
<i>Trichodesmium erythraeum</i>	NC_008312	Cyanobacteria	OSC-2
<i>Synechococcus elongatus</i>	NC_006576	Cyanobacteria	SO-6
<i>Synechococcus RCC307</i>	NC_009482	Cyanobacteria	SO-6
<i>Prochlorococcus marinus</i>	NC_009091	Cyanobacteria	SO-6
<i>Nostoc punctiforme</i>	NC_010628	Cyanobacteria	NOST-1
<i>Nostoc PCC7120</i>	NC_003272	Cyanobacteria	NOST-1
<i>Synechococcus PCC7002</i>	NC_010475	Cyanobacteria	SPM-3
<i>Synechocystis PCC6803</i>	NC_000911	Cyanobacteria	SPM-3
<i>Microcystis aeruginosa</i>	NC_010296	Cyanobacteria	SPM-3
<i>Cyanothece PCC7822</i>	NC_014501	Cyanobacteria	SPM-3
<i>Cyanothece PCC8801</i>	NC_011726	Cyanobacteria	SPM-3
<i>Cyanophora paradoxa</i>	NC_001675	Glaucophyta	Glaucocystophyceae
<i>Cyanidium caldarium</i>	NC_001840	Rhodoophyta	Bangiophyceae
<i>Cyanidioschyzon merolae</i>	NC_004799	Rhodoophyta	Bangiophyceae
<i>Porphyra purpurea</i>	NC_000925	Rhodoophyta	Bangiophyceae
<i>Gracilaria tenuistipitata</i>	NC_006137	Rhodoophyta	Florideophyceae
<i>Chlorokybus atmophyticus</i>	NC_008822	Viridiplantae	Streptophyta
<i>Aneura mirabilis</i>	NC_010359	Viridiplantae	Streptophyta
<i>Zygnema circumcarinatum</i>	NC_008117	Viridiplantae	Streptophyta
<i>Ostreococcus tauri</i>	NC_008289	Viridiplantae	Mamiellophycae
<i>Micromonas RCC299</i>	NC_012575	Viridiplantae	Mamiellophycae
<i>Nephroselmis olivacea</i>	NC_000927	Viridiplantae	Prasinophyceae
<i>Pyramimonas parkeae</i>	NC_012099	Viridiplantae	Prasinophyceae
<i>Monomastix oke1</i>	NC_012101	Viridiplantae	Prasinophyceae
<i>Pycnococcus provasolii</i>	NC_012097	Viridiplantae	Prasinophyceae
<i>Chlorella vulgaris</i>	NC_001865	Viridiplantae	Trebouxiophyceae
<i>Bryopsis hypnoides</i>	NC_013359	Viridiplantae	Ulvophyceae
<i>Chlamydomonas reinhardtii</i>	NC_005353	Viridiplantae	Chlorophyceae
<i>Scenedesmus obliquus</i>	NC_008101	Viridiplantae	Chlorophyceae

Table S4: Sequence provenance by taxon: "X" indicates that a sequence was excluded after consideration. Black coloured entries were identified by their annotation in the GenBank record; blue coloured entries were identified using HMMER gene models on bacterial and plastid genome data of the taxa; red coloured entries were identified using NCBI BLAST on bacterial and plastid genome data of the taxa; green coloured entries were identified by NCBI BLAST on GenBank "nt" nucleotide database.

	tps2	tps3	tps4	tps7	tps8	tps9	tps11	tps12	tps14	tps18	tps19	tps12	tps15	tps112	tps114	tps116	tps119	tps120	tps122	tps123	tps132	tps133	tps136	petA	petB
<i>Helicobacterium modesticaldum</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Bacillus subtilis</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Bdellovibrio bacteriovorus</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Novosphingobium aromaticivorans</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Rhodospseudomonas palustris</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Chlorobium limicola</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Herpetosiphon aurantiacus</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Roseiflexus castenholzii</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Gloeobacter violaceus</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Synechococcus PCC7002</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Synechococcus JA39Ab</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Acyrochloris marina</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Synechococcus elongatus</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Synechococcus RCC307</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Prochlorococcus marinus</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Nostoc punctiforme</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Nostoc PCC7120</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Trichodesmium erythraeum</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Thermosynechococcus elongatus</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Cyanothece PCC7425</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Cyanothece PCC6803</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Synechocystis aeruginosa</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Microcystis aeruginosa</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Cyanothece PCC7822</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Cyanothece PCC8801</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Cyanophora paradoxa</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Cyanidium caldarium</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Cyanidioschyzon merolae</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Porphyra purpurea</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Gracilaria tenuistipitata</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Ostreococcus tauri</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Pyramimonas parkeae</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Pyrococcus provasolii</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Monomastix okei</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Macromonas RCC899</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Chlamydomonas reinhardtii</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Scenedesmus obliquus</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Chlorella vulgaris</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Bryopsis hypnoides</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Nephroselmis olivacea</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Chlorokybus atmophyllicus</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Zygnema circumcarinatum</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Aneura mirabilis</i>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Table S4: (continued)

[illegible]

Table S4: (continued)

	<i>ribL</i>	<i>ycf3</i>	<i>ycf4</i>	<i>accD</i>	<i>atpA</i>	<i>atpB</i>	<i>atpE</i>	<i>atpF</i>	<i>atpH</i>	<i>atpI</i>	<i>csaA</i>	<i>cemA</i>	<i>chlB</i>	<i>chlI</i>	<i>chlL</i>	<i>chlN</i>	<i>clpP</i>	<i>cysT</i>	<i>fsH</i>	<i>infA</i>	<i>rpoA</i>	<i>rpoB</i>	<i>rpoC1</i>	<i>rpoC2</i>	<i>tufA</i>	<i>ycf12</i>
<i>Helicobacterium modesticaldum</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Bacillus subtilis</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Bdellovibrio bacteriovorus</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Novosphingobium aromaticivorans</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Rhodopseudomonas palustris</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Chlorobium limicola</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Herpetosiphon aurantiacus</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Roseiflexus castenholzii</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Gloeobacter violaceus</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Synechococcus PCC7002</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Synechococcus JA33Ab</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Acaryochloris marina</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Synechococcus elongatus</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Synechococcus RCC307</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Prochlorococcus marinus</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Nostoc punctiforme</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Nostoc PCC7120</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Trichodesmium erythraeum</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Thermosynechococcus elongatus</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Cyanothece PCC7425</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Synechocystis PCC6803</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Microcystis aeruginosa</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Cyanothece PCC7822</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Cyanothece PCC8801</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Cyanophora paradoxa</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Cyanidium caldarium</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Cyanidioschyzon merolae</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Porphyra purpurea</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Gracilaria tenuistipitata</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Ostreococcus tauri</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Pyramimonas parkae</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Pycnococcus provasolii</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Monomastix okei</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Monomonas RCC299</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Micromonas reinhardtii</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Scenedesmus obliquus</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Chlorella vulgaris</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Bryopsis hypnoides</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Nephroselmis olivacea</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Chlorokybus atmophyticus</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Zygnema circumcarinatum</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Aneura mirabilis</i>					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Table S5: Summary statistics of the data matrices of the 75 chosen loci. Numbers of taxa, total alignment length (length), and numbers of sites included (sites included) for each of the loci that formed the combined 75 loci multi-gene data matrix ('cp75').

locus	number of taxa	length	sites included
<i>rps2</i>	42	3447	579
<i>rps3</i>	42	2556	411
<i>rps4</i>	43	939	363
<i>rps7</i>	42	585	444
<i>rps8</i>	42	510	297
<i>rps9</i>	38	777	321
<i>rps11</i>	42	420	333
<i>rps12</i>	42	459	360
<i>rps14</i>	43	354	300
<i>rps18</i>	43	663	147
<i>rps19</i>	42	357	228
<i>rpl2</i>	42	969	672
<i>rpl5</i>	41	639	522
<i>rpl12</i>	34	456	297
<i>rpl14</i>	42	402	348
<i>rpl16</i>	42	486	387
<i>rpl19</i>	35	606	276
<i>rpl20</i>	42	411	306
<i>rpl22</i>	33	459	297
<i>rpl23</i>	41	384	216
<i>rpl32</i>	38	294	54
<i>rpl33</i>	33	210	99
<i>rpl36</i>	41	255	114
<i>petA</i>	34	1116	777
<i>petB</i>	40	2496	618
<i>petD</i>	34	507	468
<i>petG</i>	34	123	93
<i>petL</i>	20	111	72
<i>petN</i>	28	117	81
<i>psaA</i>	33	2523	2049
<i>psaB</i>	37	2781	2052
<i>psaC</i>	33	249	243
<i>psaI</i>	30	183	75
<i>psaJ</i>	32	195	87
<i>psaM</i>	29	135	78
<i>psbA</i>	54	1149	1044
<i>psbB</i>	34	1686	1482
<i>psbC</i>	34	1560	1251
<i>psbD</i>	36	1107	1020
<i>psbE</i>	35	297	207
<i>psbF</i>	34	354	105
<i>psbH</i>	34	306	165
<i>psbI</i>	33	183	96
<i>psbJ</i>	34	228	96
<i>psbK</i>	34	198	108
<i>psbL</i>	34	163	99
<i>psbM</i>	26	159	87
<i>psbN</i>	35	213	117
<i>psbT</i>	34	156	87
<i>psbZ</i>	33	345	180
<i>rbcL</i>	35	1473	1320
<i>ycf3</i>	33	552	483
<i>ycf4</i>	32	783	459
<i>accD</i>	35	2031	726
<i>atpA</i>	42	1671	1350
<i>atpB</i>	42	1575	1296
<i>atpE</i>	42	489	171
<i>atpF</i>	42	666	369
<i>atpH</i>	42	396	195
<i>atpI</i>	40	1296	447
<i>ccsA</i>	35	2523	525
<i>cemA</i>	32	2154	537
<i>chlB</i>	31	2463	1164
<i>chlI</i>	39	1326	882
<i>chlL</i>	32	1140	756
<i>chlN</i>	32	1788	396
<i>clpP</i>	83	1890	483
<i>ftsH</i>	83	2664	1437
<i>infA</i>	40	378	201
<i>rpoA</i>	41	2118	402
<i>rpoB</i>	39	6219	2145
<i>rpoC1</i>	41	11382	945
<i>rpoC2</i>	42	16206	933
<i>tufA</i>	40	1278	972
<i>ycf12</i>	29	315	96

Table S6: NCBI GenBank Accession Numbers of all sequences included in the combined 76 proteins. Locus *cystT* was eventually eliminated due to lack of data and the remaining 75 proteins combined to for the multi-gene protein ('cp75') data set

Taxon	accD	atpA	atpB	atpE	atpF	atpH	atpI	ccsA
<i>Helicobacterium modesticaldum</i>	YP_001678717.1	YP_001679483.1	YP_001679481.1	YP_001679480.1	YP_001679485.1	YP_001679486.1	YP_001679487.1	YP_001680531.1
<i>Bacillus subtilis</i>	NP_390799.2	NP_391562.1	NP_391562.1	NP_391561.1	NP_391566.1	NP_391567.1	NP_391568.1	NP_390194.2
<i>Deinococcus radiodurans</i>	missing	NP_970600.1	NP_970598.1	NP_970597.1	NP_970602.1	NP_970601.1	NP_967039.1	NP_968490.1
<i>Neosporobolus bacteriivorans</i>	YP_496578.1	YP_497670.1	YP_497672.1	YP_497673.1	YP_496597.1	YP_496595.1	YP_496594.1	missing
<i>Rhodospirillum rubrum</i>	YP_779395.1	YP_779221.1	YP_779219.1	YP_779218.1	YP_783677.1	YP_783677.1	YP_783678.1	missing
<i>Chlorobium limicola</i>	YP_001942475.1	YP_001944334.1	YP_001942116.1	YP_001942115.1	YP_001944377.1	YP_001944488.1	YP_001944488.1	YP_001942471.1
<i>Herpetosiphon aurantiacus</i>	YP_001546301.1	YP_001546830.1	YP_001546832.1	YP_001546833.1	YP_001546828.1	YP_001546827.1	YP_001546826.1	missing
<i>Roseiflexus castellanii</i>	YP_001431267.1	YP_001431380.1	YP_001431380.1	YP_001431379.1	YP_001431384.1	YP_001431385.1	YP_001431386.1	missing
<i>Gloeobacter violaceus</i>	NP_924551.1	NP_925851.1	NP_925516.1	NP_925514.1	NP_925853.1	NP_925855.1	NP_925856.1	NP_924968.1
<i>Synechococcus PCC7002</i>	YP_001733331.1	YP_001733995.1	YP_001734010.1	YP_001734011.1	YP_001733997.1	YP_001733999.1	YP_001734000.1	YP_001735164.1
<i>Synechococcus JA83Ab</i>	YP_475507.1	YP_475516.1	YP_473902.1	YP_474970.1	YP_475518.1	YP_475520.1	YP_475521.1	YP_474686.1
<i>Acaryochloris marina</i>	YP_001518184.1	YP_001512521.1	YP_001519605.1	YP_001519604.1	YP_001515250.1	YP_001515248.1	YP_001515247.1	YP_001516053.1
<i>Synechococcus elongatus</i>	YP_172849.1	YP_171887.1	YP_172497.1	YP_172496.1	YP_171889.1	YP_171891.1	YP_171892.1	YP_171719.1
<i>Synechococcus RCG807</i>	YP_001227249.1	YP_001228134.1	YP_001228117.1	YP_001228118.1	YP_001228136.1	YP_001228138.1	YP_001228139.1	YP_001227516.1
<i>Prochlorococcus marinus</i>	YP_001091069.1	YP_001091855.1	YP_001091852.1	YP_001091853.1	YP_001091867.1	YP_001091869.1	YP_001091870.1	YP_001091049.1
<i>Nostoc punctiforme</i>	YP_001866275.1	YP_001868153.1	YP_001867730.1	YP_001867729.1	YP_001868151.1	YP_001868149.1	YP_001868148.1	YP_001869189.1
<i>Nostoc PCC7120</i>	NP_486404.1	NP_484049.1	NP_489079.1	NP_489078.1	NP_484051.1	NP_484053.1	NP_484054.1	NP_484079.1
<i>Thermosynechococcus erythraeus</i>	YP_722346.1	YP_721901.1	YP_722959.1	YP_722960.1	YP_721903.1	NP_721905.1	NP_721906.1	NP_720548.1
<i>Thermosynechococcus elongatus</i>	NP_682433.1	NP_681225.1	NP_681315.1	NP_681316.1	NP_681223.1	NP_681221.1	NP_681220.1	NP_682405.1
<i>Cyanospora PCC7725</i>	YP_002484544.1	YP_002482223.1	YP_002485676.1	YP_002485675.1	YP_002482221.1	YP_002482219.1	YP_002482218.1	YP_002485134.1
<i>Synechocystis PCC6803</i>	NP_442022.1	NP_440055.1	NP_441407.1	NP_441408.2	NP_440057.1	NP_440059.1	NP_440060.1	NP_440317.1
<i>Microcystis aeruginosa</i>	YP_00160767.1	YP_001660030.1	YP_001655106.1	YP_001655107.1	YP_001660028.1	YP_001660026.1	YP_001660025.1	YP_001655204.1
<i>Cyanospora PCC7892</i>	YP_003888484.1	YP_003890651.1	YP_003889227.1	YP_003889226.1	YP_003890653.1	YP_003890651.1	YP_003890656.1	YP_003888654.1
<i>Cyanospora PCC8801</i>	YP_002370462.1	YP_002372877.1	YP_002373775.1	YP_002373774.1	YP_002372875.1	YP_002372873.1	YP_002372872.1	YP_002372479.1
<i>Cyanospora parvula</i>	missing	NP_043222.1	NP_043241.1	NP_043242.1	NP_043224.1	NP_043226.1	missing	NP_043267.1
<i>Cyanospora caroliniana</i>	NP_045155.1	NP_045041.1	NP_045025.1	NP_045026.1	NP_045039.1	NP_045037.1	NP_045036.1	NP_045073.1
<i>Cyanospora merolae</i>	NP_849109.1	NP_849125.1	NP_849101.1	NP_849102.1	NP_849123.1	NP_849121.1	NP_849120.1	NP_849076.1
<i>Paraphysa purpurea</i>	NP_053808.1	NP_053852.1	NP_053869.1	NP_053870.1	NP_053854.1	NP_053856.1	NP_053857.1	NP_053979.1
<i>Gyrodactylus salicetorum</i>	YP_063517.1	YP_063652.1	YP_063636.1	YP_063635.1	YP_063650.1	YP_063648.1	YP_063647.1	YP_063685.2
<i>Ostreococcus tauri</i>	missing	YP_717250.1	YP_717233.1	YP_717212.1	YP_717249.1	YP_717248.1	YP_717247.1	YP_717246.1
<i>Pyraminococcus porphyreus</i>	missing	YP_002600941.1	YP_002600926.1	YP_002600928.1	YP_002600942.1	YP_002600943.1	YP_002600944.1	YP_002600969.1
<i>Pyraminococcus porphyreus</i>	missing	YP_002600829.1	YP_002600811.1	YP_002600828.1	YP_002600828.1	YP_002600827.1	YP_002600826.1	YP_002600835.1
<i>Microcystis Plectonon</i>	missing	YP_002601009.1	YP_002601016.1	YP_002601015.1	YP_002601010.1	YP_002601012.1	YP_002601013.1	YP_002601035.1
<i>Microcystis PCC999</i>	missing	YP_002808620.1	NP_002808620.1	NP_002808656.1	YP_002808621.1	NP_002808622.1	NP_002808623.1	XP_M_002505571.1
<i>Chlamydomonas reinhardtii</i>	XP_M_001703135.1	NP_958406.1	NP_958414.1	NP_958379.1	NP_958401.1	NP_958409.1	NP_958418.1	NP_958384.1
<i>Synechococcus obliquus</i>	NP_045833.1	NP_045781.1	NP_045906.1	NP_045905.1	NP_045783.1	NP_045784.1	NP_045785.1	NP_045898.1
<i>Chlorella vulgaris</i>	YP_003227083.1	YP_003227033.1	YP_003227061.1	YP_003227075.1	YP_003227034.1	YP_003227035.1	YP_003227036.1	NP_003227085.1
<i>Bradyrhizobium elaeagnis</i>	NP_050834.1	NP_050829.1	NP_050811.1	NP_050812.1	NP_050830.1	NP_050831.1	NP_050832.1	NP_050913.1
<i>Chlorogobius amphiplus</i>	YP_001019128.1	YP_001019097.1	YP_001019150.1	YP_001019149.1	YP_001019098.1	YP_001019099.1	YP_001019100.1	YP_001019174.1
<i>Zygnema crenulicarmatum</i>	YP_636510.1	YP_636515.1	YP_636546.1	YP_636545.1	YP_636516.1	YP_636517.1	YP_636518.1	YP_636518.2
<i>Artemia maritima</i>	YP_001687225.1	YP_001687211.1	YP_001687223.1	YP_001687222.1	YP_001687210.1	YP_001687209.1	YP_001687208.1	missing

Table S6: (continued)

Taxon	cernA	chlB	chlI	chlL	chlN	chlP	cyoT	ftsH
<i>Heliobacterium modesticaldum</i>	missing	YP_001679878.1	YP_001679879.1	YP_001679876.1	YP_001679877.1	YP_001679876.1	YP_001679876.1	YP_001679876.1
<i>Bacillus subtilis</i>	missing	missing	missing	missing	missing	NP_391334.1	missing	NP_387950.1
<i>Bdellovibrio bacteriovorus</i>	missing	missing	missing	missing	missing	NP_970462.1	missing	NP_969465.1
<i>Nonesphondonium aromatioclorans</i>	missing	missing	missing	missing	missing	YP_498324.1	missing	YP_495359.1
<i>Rhodospseudomonas palustris</i>	missing	YP_780280.1	YP_780282.1	YP_780282.1	YP_780279.1	YP_781428.1	YP_780697.1	YP_783598.1
<i>Chlorobium limicola</i>	missing	YP_001944196.1	YP_001944195.1	YP_001944195.1	YP_001944197.1	YP_001942477.1	missing	YP_001942430.1
<i>Herpetosiphon aurantiacus</i>	missing	missing	missing	missing	missing	YP_001945234.1	missing	YP_001545795.1
<i>Roseiflexus castellanii</i>	missing	NP_923161.1	NP_923161.1	NP_923161.1	NP_923151.1	YP_001431588.1	missing	YP_001430856.1
<i>Gloeobacter violaceus</i>	missing	YP_001734898.1	NP_924660.1	NP_924660.1	NP_925316.1	NP_926713.1	NP_925019.1	NP_924863.1
<i>Synechococcus PCC7002</i>	YP_001736023.1	YP_001734898.1	YP_001735490.1	YP_001735490.1	YP_001735580.1	YP_001735091.1	YP_001734057.1	YP_001734431.1
<i>Synechococcus JA83Ab</i>	YP_475942.1	YP_474361.1	YP_475345.1	YP_475345.1	YP_475773.1	YP_474133.1	YP_473950.1	YP_473500.1
<i>Acaryochloris marina</i>	YP_001518773.1	YP_171263.1	YP_001519196.1	YP_001519196.1	YP_001515787.1	YP_001516654.1	YP_173118.1	YP_001518660.1
<i>Synechococcus RCC807</i>	YP_001227489.1	YP_172966.1	YP_171650.1	YP_170847.1	YP_170846.1	YP_172821.1	YP_173118.1	YP_170949.1
<i>Prochlorococcus marinus</i>	missing	YP_001090794.1	YP_001091385.1	YP_001091385.1	YP_001090795.1	YP_001091723.1	missing	YP_001227658.1
<i>Prochlorococcus marinus</i>	YP_001867567.1	YP_001868220.1	YP_001869529.1	YP_001869529.1	YP_001867733.1	YP_001869446.1	missing	YP_001091238.1
<i>Nostoc PCC7120</i>	NP_485713.1	NP_487481.1	NP_484196.1	NP_484196.1	NP_489118.1	NP_487723.1	NP_484365.1	NP_488816.1
<i>Thermosynechococcus erythraeus</i>	YP_723002.1	YP_723330.1	YP_723165.1	YP_721286.1	NP_721284.1	NP_721067.1	missing	YP_722840.1
<i>Thermosynechococcus elongatus</i>	NP_681537.1	NP_683182.1	NP_682301.1	NP_683137.1	NP_683135.1	NP_681862.1	NP_681066.1	NP_680922.1
<i>Cyanothecis PCC7795</i>	YP_002482256.1	YP_002484347.1	YP_002481717.1	YP_002484347.1	YP_002484618.1	YP_002482249.1	NP_002484370.1	NP_002484929.1
<i>Synechocystis PCC6803</i>	NP_440313.1	NP_442044.1	NP_440486.1	NP_442933.1	NP_442934.1	NP_442796.1	NP_442841.1	NP_442160.1
<i>Microcystis aeruginosa</i>	YP_001658626.1	YP_001660400.1	YP_001659002.1	YP_001656637.1	YP_001656639.1	YP_001660733.1	YP_001685167.1	YP_001660152.1
<i>Cyanothecis PCC782</i>	YP_003889894.1	YP_003885707.1	YP_003885890.1	YP_003885051.1	YP_003890521.1	YP_003887883.1	YP_003885875.1	YP_003888550.1
<i>Cyanothecis PCC8801</i>	YP_002373813.1	YP_002373046.1	YP_002370706.1	YP_002371029.1	YP_002371027.1	YP_002373656.1	YP_002372905.1	YP_002370534.1
<i>Cyanophora paradoxa</i>	missing	NP_043149.1	NP_043214.1	NP_043285.1	NP_043286.1	NP_043261.1	missing	missing
<i>Cyanidium caldarium</i>	NP_045051.1	NP_045141.1	NP_045141.1	NP_045141.1	missing	missing	missing	NP_045094.1
<i>Cyanidochytrium merolae</i>	NP_849136.1	missing	NP_848052.1	missing	missing	AP006494	NP_848065.1	NP_849040.1
<i>Porphyra purpurea</i>	NP_053842.1	NP_053888.1	NP_054004.1	NP_053797.1	NP_053798.1	missing	missing	NP_053937.1
<i>Gelidium coulteri</i>	YP_063662.1	missing	YP_063709.1	missing	missing	missing	missing	YP_063571.1
<i>Ostreococcus tauri</i>	missing	missing	missing	missing	missing	YP_717211.1	missing	missing
<i>Pyraminococcus porfereae</i>	YP_002600931.1	YP_002600931.1	XP_003074999	YP_002600945.1	YP_002600976.1	YP_002600892.1	missing	missing
<i>Pyrococcus furiosus</i>	missing	missing	YP_002600814.1	YP_002600814.1	YP_002600836.1	YP_002600868.1	missing	missing
<i>Micromonas Ples</i>	missing	missing	YP_002601047.1	missing	missing	YP_002601006.1	missing	missing
<i>Micromonas PCC999</i>	missing	missing	XP_002607588	missing	missing	XP_002808657	missing	missing
<i>Chlamydomonas reinhardtii</i>	NP_958408.1	NP_958360.1	NP_958366.1	NP_958366.1	NP_958412.2	NP_958364.1	XP_001692407	XP_002507540
<i>Saccharomyces cerevisiae</i>	YP_635982.1	YP_635962.1	YP_635962.1	YP_635962.1	NP_635948.1	NP_635990.1	missing	missing
<i>Chlorella vulgaris</i>	NP_045761.1	NP_045840.1	NP_045840.1	NP_045840.1	NP_045840.1	NP_045839.1	NP_045890.1	missing
<i>Bradyrhizobium japonicum</i>	YP_003227074.1	YP_003227057.1	NP_0183350	YP_003227064.1	YP_003227063.1	YP_003227081.1	NP_045890.1	missing
<i>Nepenthes alata</i>	NP_050853.1	NP_050960.1	NP_050837.1	NP_050935.1	NP_050936.1	NP_050936.1	NP_050928.1	missing
<i>Chlorogobius atmophilus</i>	YP_001019143.1	YP_001019110.1	YP_001019148.1	YP_001019148.1	YP_001019127.1	YP_001019072.1	YP_001019170.1	missing
<i>Zygnema carnicum</i>	YP_636559.1	YP_636477.1	YP_636541.1	YP_636519.1	YP_636520.1	YP_636469.1	YP_636569.1	missing
<i>Aurea mirabilis</i>	YP_001687228.1	YP_001687215.1	missing	YP_001687261.1	YP_001687260.1	YP_001687238.1	missing	missing

Table S6: (continued)

Taxon	infA	petA	petB	petD	petG	petL	petN	pasA
<i>Halobacterium modesticum</i>	YP_001679900.1	missing	YP_001679860.1	YP_001679866.1	missing	missing	missing	missing
<i>Bacillus subtilis</i>	NP_388020.1	missing	NP_390136.1	missing	missing	missing	missing	missing
<i>Deinobacterio bacteriovorus</i>	YP_967899.1	missing	missing	missing	missing	missing	missing	missing
<i>Novosphingobium aromaticivorans</i>	YP_496337.1	missing	YP_495958.1	missing	missing	missing	missing	missing
<i>Rhodopseudomonas palustris</i>	YP_782652.1	missing	YP_779885.1	missing	missing	missing	missing	missing
<i>Chlorobium limicola</i>	YP_001944210.1	missing	YP_001942434.1	missing	missing	missing	missing	missing
<i>Herpetosiphon aurantiacus</i>	YP_001547699.1	missing	YP_001547012.1	missing	missing	missing	missing	missing
<i>Roseiflexus castenholzii</i>	YP_001434055.1	missing	missing	missing	missing	missing	missing	missing
<i>Gloeobacter violaceus</i>	NP_925985.1	NP_925985.1	NP_924865.1	NP_924864.1	NP_923457.1	NP_923457.1	NP_926046.1	NP_926384.1
<i>Synechococcus PCC7002</i>	YP_001735153.1	YP_001735153.1	YP_001734102.1	YP_001734101.1	YP_001733640.1	YP_001733640.1	NC_010475.1	YP_001735204.1
<i>Synechococcus JAR33Ab</i>	YP_474837.1	YP_474837.1	YP_475212.1	YP_475213.1	YP_476039.1	YP_476141.1	YP_476141.1	YP_475904.1
<i>Acaryochloris marina</i>	YP_001519006.1	YP_001518743.1	YP_001518958.1	YP_001518957.1	YP_001516381.1	YP_001516381.1	YP_001516381.1	YP_001516779.1
<i>Synechococcus elongatus</i>	YP_172306.1	YP_171029.1	YP_172481.1	YP_172480.1	NC_006576.1	YP_172101.1	YP_171754.1	YP_172754.1
<i>Synechococcus RCC307</i>	YP_001227940.1	YP_001226959.1	YP_001226701.1	YP_001226702.1	YP_001226699.1	YP_001226702.1	YP_001227663.1	YP_001227663.1
<i>Synechococcus RCC907</i>	YP_001097106.1	YP_001090575.1	YP_001090575.1	YP_001090576.1	YP_001091388.1	NC_009091.1	YP_001091024.1	YP_001091940.1
<i>Prochlorococcus marinus</i>	YP_001869768.1	YP_001863874.1	YP_001864038.1	YP_001864039.1	NC_010628.1	missing	YP_001865227.1	YP_001867145.1
<i>Nostoc punctiforme</i>	NP_486492.1	NP_486492.1	NP_487461.1	NP_487462.1	NP_485409.1	NP_485962.1	NP_488303.1	NP_489194.1
<i>Nostoc PCC71120</i>	NP_482825.1	NP_482825.1	YP_720972.1	YP_720971.1	NP_721848.1	NP_721848.1	NC_008312.1	YP_724110.1
<i>Trichodesmium erythraeum</i>	YP_681575.1	YP_681575.1	YP_681585.1	NP_681586.1	NP_681607.1	YP_654185.1	NP_681607.1	YP_681520.1
<i>Thermosynechococcus elongatus</i>	YP_002482050.1	YP_002482152.1	YP_002483066.1	YP_002483067.1	NP_002484620.1	NP_001035872.1	YP_002484350.1	NP_002485338.1
<i>Cyanospora PCC7425</i>	NP_002482050.1	NP_002482152.1	YP_002483066.1	YP_002483067.1	NP_002484620.1	NP_001035872.1	YP_002484350.1	NP_002485338.1
<i>Microcystis PCC6803</i>	NC_440947.1	NP_440947.1	YP_0024709.1	YP_0024709.1	NP_441556.1	NP_440044.1	NP_440044.1	YP_440757.1
<i>Microcystis aeruginosa</i>	NC_010296.1	YP_001656937.1	YP_001658371.1	NP_001658370.1	NP_001659648.1	YP_001659770.1	YP_001659770.1	NP_001659770.1
<i>Synechococcus PCC7422</i>	YP_003890352.1	YP_003885561.1	YP_003890352.1	YP_003890353.1	YP_003890922.1	NC_014501.1	YP_003890154.1	YP_003890154.1
<i>Cyanospora PCC8801</i>	YP_002370993.1	YP_002370993.1	YP_002373659.1	YP_002373658.1	YP_002370675.1	YP_002371562.1	YP_002372562.1	YP_002374191.1
<i>Cyanophora paradoxa</i>	missing	NP_043244.1	NP_043175.1	NP_043174.1	NP_043220.1	NP_043237.1	NP_043216.1	NP_043150.1
<i>Cyanidium caldarium</i>	missing	NP_045119.1	NP_045120.1	NP_045119.1	NP_045119.1	NP_045210.1	NP_045203.1	NP_045214.1
<i>Cyanidium</i>	missing	NP_849061.1	NP_849013.1	NP_849014.1	NP_849048.1	NP_849134.1	NP_849054.1	NP_849044.1
<i>Cyanodictyon merolae</i>	NP_053875.1	NP_053875.1	NP_053951.1	NP_053950.1	NP_053928.1	NP_053886.1	NP_053886.1	NP_053894.1
<i>Porphyra purpurea</i>	missing	YP_063630.1	YP_063557.1	NP_063556.1	YP_063611.1	NP_063620.1	NP_063620.1	NP_063614.2
<i>Gracilaria tenuistipitata</i>	missing	YP_063630.1	YP_063557.1	NP_063556.1	YP_063611.1	NP_063620.1	NP_063620.1	NP_063614.2
<i>Ostreococcus tenuis</i>	YP_717243.1	YP_717252.1	YP_717215.1	XM_003080111.4	YP_717263.1	YP_717254.1	YP_717254.1	YP_717254.1
<i>Pyramimonas parkae</i>	YP_002600915.1	YP_002600940.1	YP_002600887.1	missing	YP_002600893.1	missing	YP_002600934.1	YP_002600932.1
<i>Pyramimonas provaschii</i>	YP_002600847.1	YP_002600847.1	YP_002600831.1	YP_002600830.1	YP_002600825.1	missing	YP_002600825.1	YP_002600861.1
<i>Monomastix okei</i>	YP_002601062.1	YP_002601034.1	YP_002601048.1	YP_002601049.1	YP_002601008.1	missing	YP_002600998.1	YP_002600993.1
<i>Micromonas RCC299</i>	YP_002808627	YP_002808618	YP_002808654	XM_002149473	YP_002808617	missing	missing	missing
<i>Chlamydomonas reinhardtii</i>	XM_001692203	NP_958358.1	NP_958365.1	NP_958359.1	NP_958401.1	NP_958424.2	NP_958424.2	NP_958424.2
<i>Scenedesmus obliquus</i>	YP_635967.1	YP_636003.1	YP_635992.1	YP_636001.1	NP_635953.1	NP_635964.1	missing	missing
<i>Chlorella vulgaris</i>	YP_045911.2	NP_045911.2	NP_045838.1	NP_045836.1	NP_045908.1	NP_045909.1	missing	missing
<i>Bryopsis hypnoides</i>	YP_003227054.1	YP_003227054.1	YP_003227078.1	YP_003227079.1	NP_003227032.1	YP_003227056.1	missing	missing
<i>Nephrolepis olivacea</i>	NP_050820.1	NP_050820.1	NP_050814.1	NP_050813.1	NP_050857.1	NP_050856.1	NP_050844.1	NP_050883.1
<i>Chloroklebsmia olivacea</i>	YP_001019084.1	YP_001019144.1	YP_001019078.1	YP_001019079.1	NP_001019067.1	NP_00101946.1	YP_001019105.1	YP_001019111.1
<i>Zugmura circumcarinata</i>	YP_636495.1	NP_636533.1	YP_636562.1	NP_636563.1	YP_636550.1	YP_636549.1	YP_636549.1	YP_636558.1
<i>Aurea mirabilis</i>	YP_001687246.1	NC_010359.1	NP_010359.1	YP_001687242.1	YP_001687233.1	YP_001687232.1	YP_001687203.1	YP_001687203.1

Table S6: (continued)

Taxon	psaB	psaC	psaI	psaJ	psaM	psaA	psaB	psaC
<i>Heliobacterium modesticaldum</i>	missing	missing	missing	missing	missing	missing	missing	missing
<i>Bacillus subtilis</i>	missing	missing	missing	missing	missing	missing	missing	missing
<i>Bdellovibrio bacteriovorus</i>	missing	missing	missing	missing	missing	missing	missing	missing
<i>Nonosphingobium aromaticivorans</i>	missing	missing	missing	missing	missing	missing	missing	missing
<i>Rhodospseudomonas palustris</i>	missing	missing	missing	missing	missing	missing	missing	missing
<i>Chlorobium limicola</i>	missing	missing	missing	missing	missing	missing	missing	missing
<i>Herpetosiphon aurantiacus</i>	missing	missing	missing	missing	missing	missing	missing	missing
<i>Roseiflexus castenholzii</i>	missing	missing	missing	missing	missing	missing	missing	missing
<i>Gloeobacter violaceus</i>	NP_926385.1	NP_926233.1	NP_001735853.1	NP_001734268.1	NP_925347.1	NP_924652.1	NP_925945.1	NP_925270.1
<i>Synechococcus PCC7002</i>	YP_001735205.1	YP_001734836.1	YP_001735853.1	YP_001734268.1	YP_001735400.1	YP_001735400.1	YP_001735004.1	YP_001734806.1
<i>Synechococcus JA83Ab</i>	YP_475905.1	YP_473973.1	YP_473531.1	YP_473781.1	YP_475419.1	YP_475166.1	YP_474121.1	YP_475748.1
<i>Acaryochloris marina</i>	YP_001516780.1	YP_001515995.1	YP_172470.1	YP_001517781.1	YP_001518034.1	YP_001514818.1	YP_001516355.1	YP_001515437.1
<i>Synechococcus elongatus</i>	YP_172755.1	YP_171696.1	YP_001228409.1	YP_001226965.1	YP_172892.1	YP_171803.1	YP_171543.1	YP_171582.1
<i>Synechococcus RCC807</i>	YP_001228413.1	YP_001226389.1	YP_001228409.1	YP_001226965.1	YP_001227830.1	YP_001227697.1	YP_001226674.1	YP_001227952.1
<i>Prochlorococcus marinus</i>	YP_001091939.1	YP_001092024.1	YP_001091936.1	YP_001090717.1	YP_001090790.1	YP_001090469.1	YP_001090564.1	YP_001091574.1
<i>Nostoc punctiforme</i>	YP_001863941.1	YP_001868486.1	NC_010628.1	YP_001867190.1	YP_001868026.1	YP_001868705.1	YP_001865975.1	YP_001866969.1
<i>Nostoc PCC7120</i>	NP_489195.1	NP_487503.1	NP_487889.1	NP_484152.1	NP_488697.1	NP_488906.1	NP_484182.1	NP_488331.1
<i>Thermosynechococcus erythraeus</i>	YP_724109.1	YP_720386.1	YP_721034.1	YP_721382.1	NC_008312.1	NP_720146.1	NP_724107.1	YP_720441.1
<i>Thermosynechococcus elongatus</i>	NP_681521.1	NP_681803.1	NP_683195.1	NP_683202.1	NP_680988.1	NP_682267.1	NP_682320.1	NP_682421.1
<i>Cyanothecis PCC7795</i>	NP_002483452.1	NP_002482471.1	YP_002482500.1	YP_002484676.1	YP_002483058.1	NP_002483684.1	YP_002481608.1	YP_002480974.1
<i>Synechocystis PCC6803</i>	NP_440758.1	NP_441966.1	NP_442962.1	NP_441427.1	NP_440325.1	NP_439906.1	NP_442388.1	NP_441119.1
<i>Microcystis aeruginosa</i>	YP_001659771.1	YP_001660937.1	YP_001659384.1	YP_001659744.1	YP_001655963.1	YP_001656036.1	YP_001685313.1	YP_001659129.1
<i>Cyanosphaera PCC782</i>	YP_003890153.1	YP_003885668.1	YP_003888541.1	YP_003888378.1	YP_003890329.1	YP_003888964.1	YP_003886535.1	YP_003886960.1
<i>Cyanosphaera PCC8801</i>	YP_002374192.1	YP_002371499.1	YP_002374078.1	YP_002373525.1	YP_002371085.1	YP_002373956.1	YP_002373206.1	YP_002370318.1
<i>Cyanosphaera paradoxa</i>	NP_043151.1	NP_043270.1	NP_043182.1	NP_043152.1	NP_043246.1	NP_043238.1	NP_043167.1	NP_043248.3
<i>Cyanidium caldarium</i>	NP_045213.1	NP_045270.1	NP_045043.1	NP_045153.1	NP_045143.1	NP_045067.1	NP_045100.1	NP_045020.2
<i>Cyanidioschyzon merolae</i>	NP_849045.1	NP_848982.1	NP_849127.1	NP_849105.1	NP_848961.1	NP_848970.1	NP_849032.1	NP_849003.2
<i>Porphyra purpurea</i>	NP_053895.1	NP_053984.1	NP_053997.1	NP_053804.1	NP_054005.1	NP_053822.1	NP_053932.1	NP_053966.2
<i>Gyrodinium aureolum</i>	YP_063613.1	YP_063690.1	YP_063702.1	YP_063513.1	YP_063710.1	YP_063531.1	YP_063576.1	YP_063543.2
<i>Ostreococcus tauri</i>	YP_717253.1	YP_717214.1	YP_717255.1	YP_717261.1	YP_717257.1	YP_717233.1	YP_717207.1	YP_717204.2
<i>Pyraminococcus porfereae</i>	YP_002600933.1	YP_002600959.1	YP_002600909.1	YP_002600910.1	YP_002600900.1	YP_002600883.1	YP_002600891.1	YP_002600929.1
<i>Pyrococcus furiosus</i>	YP_002600862.1	YP_002600805.1	YP_002600869.1	missing	missing	YP_002600858.1	YP_002600857.1	YP_002600851.1
<i>Micromonas Ples1</i>	YP_002600994.1	missing	YP_002601017.1	YP_002601037.1	YP_002601036.1	YP_002601066.1	YP_002601044.1	YP_002601085.1
<i>Micromonas PCC999</i>	YP_002808607	YP_002808655	YP_002808609	YP_002808615	YP_002808611	YP_002808637	YP_002808602	YP_002808605
<i>Chlamydomonas reinhardtii</i>	NP_958404.1	NP_958423.1	NP_958417.1	NP_958417.1	missing	NP_958377.1	NP_958388.1	NP_958422.1
<i>Saccharina obliqua</i>	NP_635999.1	YP_635963.1	NP_635955.1	NP_635955.1	NP_635976.1	NP_635976.1	NP_635968.1	NP_635942.1
<i>Chlorella vulgaris</i>	NP_045852.1	NP_045771.1	NP_045808.1	NP_045808.1	NP_045767.1	NP_045767.1	NP_045800.2	NP_045800.2
<i>Bradyrhizobium elaeagnae</i>	YP_003227059.1	YP_003227055.1	NP_013359.1	YP_003227088.1	YP_003227030.1	NP_003227030.1	YP_003227067.1	YP_003227062.1
<i>Nepenthes bicalcarata</i>	NP_050884.1	NP_050917.1	NP_050833.1	NP_050847.1	NP_050810.1	NP_050810.1	NP_050881.1	NP_050846.2
<i>Chlorogobius atmophilus</i>	YP_001019121.1	YP_001019157.1	YP_001019127.1	YP_001019136.1	YP_001019071.1	YP_001019070.1	YP_001019073.1	YP_001019066.1
<i>Zygnema carolinianum</i>	YP_636557.1	YP_636506.1	YP_636509.1	YP_636514.1	YP_636514.1	YP_636509.1	YP_636470.1	YP_636500.1
<i>Aurea mirabilis</i>	NC_010359.1	YP_001687237.1	YP_001687226.1	YP_001687234.1	YP_001687213.1	YP_001687217.1	NC_010359.1	NC_010359.1

Table S6: (continued)

Taxon	psbD	psbE	psbF	psbH	psbI	psbJ	psbK	psbL
<i>Heliobacterium modesticaldum</i>	missing	missing	missing	missing	missing	missing	missing	missing
<i>Bacillus subtilis</i>	missing	missing	missing	missing	missing	missing	missing	missing
<i>Bdellovibrio bacteriovorus</i>	missing	missing	missing	missing	missing	missing	missing	missing
<i>Nonosphingobium aromatiaciorans</i>	missing	missing	missing	missing	missing	missing	missing	missing
<i>Rhodospseudomonas palustris</i>	missing	missing	missing	missing	missing	missing	missing	missing
<i>Chlorobium limicola</i>	missing	missing	missing	missing	missing	missing	missing	missing
<i>Herpetosiphon aurantiacus</i>	missing	missing	missing	missing	missing	missing	missing	missing
<i>Roseiflexus castenholzii</i>	missing	missing	missing	missing	missing	missing	missing	missing
<i>Gloeobacter violaceus</i>	NP_925269.1	NP_923802.1	NP_923803.1	NP_925948.1	NP_926580.1	NP_923805.1	NP_925753.1	NP_923804.1
<i>Synechococcus PCC7002</i>	YP_001734807.1	YP_001733498.1	YP_001734499.1	YP_001734068.1	YP_001733501.1	YP_001733501.1	YP_001736008.1	YP_001733500.1
<i>Synechococcus JA93Ab</i>	YP_475747.1	YP_475815.1	YP_475814.1	YP_474766.1	YP_476131.1	YP_475812.1	YP_475342.1	YP_475813.1
<i>Acaryochloris marina</i>	YP_001515436.1	YP_001516952.1	YP_001515482.1	YP_001519785.1	YP_001518103.1	YP_001516950.1	YP_001518153.1	YP_001516951.1
<i>Synechococcus elongatus</i>	YP_171583.1	YP_171083.1	YP_171084.1	YP_171998.1	YP_173096.1	YP_171086.1	YP_171772.1	YP_171085.1
<i>Synechococcus RCC807</i>	YP_001226515.1	YP_001226466.1	YP_001226465.1	YP_001228594.1	YP_001228591.1	YP_001226463.1	YP_001228566.1	YP_001226464.1
<i>Prochlorococcus marinus</i>	YP_001091573.1	YP_001090545.1	YP_001090546.1	YP_001090498.1	YP_001090501.1	YP_001090548.1	YP_001090519.1	YP_001090547.1
<i>Nostoc punctiforme</i>	YP_001866970.1	YP_001868799.1	YP_001868800.1	YP_001867633.1	YP_001868395.1	YP_001868802.1	YP_001864004.1	YP_001868801.1
<i>Nostoc PCC7120</i>	NP_488330.1	NP_487885.1	NP_487886.1	NP_484889.1	NP_485320.1	NP_487888.1	NP_484928.1	NP_487887.1
<i>Thermosynechococcus erythraeus</i>	YP_721054.1	YP_723067.1	YP_723068.1	YP_722514.1	YP_722791.1	NP_723070.1	NP_723401.1	YP_723069.1
<i>Thermosynechococcus elongatus</i>	NP_681245.1	NP_682331.1	NP_682332.1	NP_682176.1	NP_681865.1	NP_682334.1	NP_680967.1	NP_682333.1
<i>Cyanothecis PCC7795</i>	YP_002480975.1	YP_002482950.1	YP_002482951.1	YP_002484197.1	YP_002482810.1	YP_002482952.1	YP_002484503.1	NP_011884.1
<i>Synechocystis PCC6803</i>	NP_441120.1	NP_440412.1	NP_440413.1	NP_440949.1	NP_442015.1	NP_440415.1	NP_440395.1	NP_440414.1
<i>Microcystis aeruginosa</i>	YP_001658812.1	YP_001658317.1	YP_001658316.1	YP_001660414.1	YP_001660137.1	YP_001658314.1	YP_001656204.1	YP_001658315.1
<i>Cyanothecis PCC782</i>	YP_003886961.1	YP_003886924.1	YP_003886925.1	YP_003888779.1	YP_003887576.1	YP_003886927.1	YP_003886628.1	YP_003886926.1
<i>Cyanothecis PCC8801</i>	YP_002370319.1	YP_002374235.1	YP_002374234.1	YP_002371852.1	YP_002373406.1	YP_002374232.1	YP_002370780.1	YP_002374233.1
<i>Cyanophora paradoxa</i>	NP_043247.1	NP_043178.1	NP_043179.1	NP_043169.1	NP_043236.1	NP_043181.1	NP_043249.1	NP_043180.1
<i>Cyanidium caldarium</i>	NP_045019.1	NP_045047.1	NP_045046.1	NP_045097.1	NP_045054.1	NP_045044.1	NP_045209.1	NP_045045.1
<i>Cyanidochytrion merolae</i>	NP_849002.1	NP_849131.1	NP_849130.1	NP_849038.1	NP_849139.1	NP_849128.1	NP_849049.1	NP_849129.1
<i>Porphyra purpurea</i>	NP_053967.1	NP_054001.1	NP_054000.1	NP_053935.1	NP_053846.1	NP_053998.1	NP_053927.1	NP_053999.1
<i>Gracilaria tikvahiae</i>	YP_063542.1	YP_063706.1	YP_063705.1	YP_063573.1	YP_063658.1	YP_063703.1	YP_063610.1	YP_063704.1
<i>Ostreococcus tauri</i>	YP_717205.1	YP_717219.1	YP_717218.1	YP_717221.1	YP_717280.1	YP_717216.1	YP_717259.1	YP_717217.1
<i>Pyraminococcus porfereae</i>	YP_002600930.1	YP_002600891.1	YP_002600897.1	YP_002600888.1	YP_002600899.1	YP_002600895.1	YP_002600902.1	YP_002600896.1
<i>Pyraminococcus bleasii</i>	YP_002600850.1	YP_002600815.1	YP_002600816.1	YP_002600854.1	YP_002600804.1	YP_002600818.1	YP_002600822.1	YP_002600817.1
<i>Microcystis Plect</i>	YP_002600990.1	YP_002601025.1	YP_002601026.1	YP_002601038.1	YP_002601081.1	YP_002601028.1	YP_002601042.1	YP_002601027.1
<i>Micromonas PCC999</i>	YP_002808604.1	YP_002808650.1	YP_002808651.1	YP_002808649.1	YP_002808614.1	YP_002808653.1	YP_002808613.1	YP_002808652.1
<i>Chlamydomonas reinhardtii</i>	NP_958420.1	NP_958396.1	NP_958399.1	NP_958385.1	NP_958407.1	NP_958419.1	NP_958361.1	NP_958400.2
<i>Saccharina obliqua</i>	NP_635958.1	NP_636009.1	NP_635951.1	NP_635971.1	NP_635983.1	NP_635957.1	NP_635972.1	NP_635952.1
<i>Chlorella vulgaris</i>	NP_045801.1	NP_045829.1	NP_045827.1	NP_045845.1	NP_045788.1	NP_045825.1	NP_045866.2	NP_045826.1
<i>Bracteopsis lagynoides</i>	YP_003227061.1	YP_003227084.1	YP_003227085.1	YP_003227069.1	YP_003227070.1	YP_003227087.1	YP_003227072.1	YP_003227086.1
<i>Nepenthes alpinica</i>	NP_050845.1	NP_050864.1	NP_050863.1	NP_050878.1	NP_050892.1	NP_050861.1	NP_050858.1	NP_050862.1
<i>Chlorogobius atmophiticus</i>	YP_001019065.1	YP_001019123.1	YP_001019124.1	YP_001019076.1	YP_001019109.1	YP_001019126.1	YP_001019108.1	YP_001019125.1
<i>Zygnema caribaeum</i>	YP_636501.1	YP_636531.1	YP_636530.1	YP_636473.1	YP_636484.1	YP_636478.1	YP_636483.1	YP_636529.1
<i>Aurea mirabilis</i>	NC_010359.1	NC_010359.1	YP_001687231.1	YP_001687241.1	YP_001687244.1	YP_001687229.1	NC_010359.1	YP_001687230.1

Table S6: (continued)

Taxon	psbM	psbN	psbT	psbZ	rbcL	rpl12	rpl14	rpl16
<i>Heliobacterium modesticaldum</i>	missing	missing	missing	missing	missing	YP_001679566.1	YP_001679566.1	YP_001679566.1
<i>Bacillus subtilis</i>	missing	missing	missing	missing	missing	NP_387986.1	NP_388007.1	NP_388004.1
<i>Bdellovibrio bacteriovorus</i>	missing	missing	missing	missing	missing	NP_969764.1	NP_969745.1	NP_969748.1
<i>Nonosphingobium aromatioclorans</i>	missing	missing	missing	missing	missing	YP_495313.1	YP_496537.1	YP_496534.1
<i>Rhodospseudomonas palustris</i>	missing	missing	missing	missing	YP_780293.1	YP_782508.1	YP_782486.1	YP_782489.1
<i>Chlorobium limicola</i>	missing	missing	missing	missing	missing	YP_001944222.1	YP_001944222.1	YP_001944225.1
<i>Herpetosiphon aurantiacus</i>	missing	missing	missing	missing	missing	YP_001542921.1	YP_001547682.1	YP_001547680.1
<i>Roseiflexus castellanii</i>	missing	missing	missing	missing	missing	YP_001434359.1	YP_001434067.1	YP_001434070.1
<i>Gloeobacter violaceus</i>	NP_925943.1	NP_925947.1	NP_925946.1	missing	NP_925102.1	NP_924548.1	NP_926863.1	NP_926866.1
<i>Synechococcus PCC7002</i>	YP_001735387.1	YP_001734069.1	YP_001734216.1	YP_00173406.1	YP_001735042.1	YP_001734285.1	YP_001734313.1	YP_001734316.1
<i>Synechococcus JA83Ab</i>	YP_474119.1	YP_474122.1	YP_474121.1	YP_473583.1	YP_474643.1	YP_475200.1	YP_474619.1	YP_474622.1
<i>Acaryochloris marina</i>	YP_001516353.1	YP_001519784.1	YP_001519335.1	YP_001515365.1	YP_001516201.1	YP_001517503.1	YP_001518996.1	YP_001518993.1
<i>Synechococcus elongatus</i>	YP_171541.1	YP_171909.1	YP_171544.1	YP_172563.1	YP_170840.1	YP_171604.1	YP_172585.1	YP_172582.1
<i>Synechococcus RCC807</i>	YP_001226876.1	YP_001228593.1	YP_001226673.1	YP_001226835.1	YP_001227075.1	YP_001228611.1	YP_001228382.1	YP_001228379.1
<i>Prochlorococcus marinus</i>	YP_001090566.1	YP_001090500.1	YP_001090563.1	YP_001092058.1	YP_001090800.1	YP_001090445.1	YP_001091963.1	YP_001091966.1
<i>Nostoc punctiforme</i>	YP_001864006.1	YP_001867632.1	YP_001865974.1	YP_001864012.1	YP_001867513.1	YP_001869095.1	YP_001867696.1	YP_001867699.1
<i>Nostoc PCC7120</i>	NP_484926.1	NP_484890.1	NP_484181.1	NP_488032.1	NP_485564.1	NP_489343.1	NP_488245.1	NP_488248.1
<i>Thermosynechococcus erythraeus</i>	NP_008312.1	YP_722513.2	NP_724106.1	NP_721203.1	NP_723871.1	NP_720218.1	NP_722630.1	NP_722633.1
<i>Thermosynechococcus elongatus</i>	NP_682842.1	NP_682177.1	NP_682321.1	NP_682757.1	NP_682296.1	NP_681089.1	NP_680882.1	NP_680879.1
<i>Cyanospora PCC7795</i>	YP_002481609.1	YP_002484196.1	YP_002481609.1	YP_002483284.1	YP_002484107.1	YP_002484685.1	YP_002482040.1	YP_002482037.1
<i>Synechocystis PCC6803</i>	NP_440206.1	NP_440950.1	NP_442063.1	NP_440888.1	NP_442120.1	NP_440736.1	NP_440659.1	NP_440662.1
<i>Microcystis aeruginosa</i>	YP_001660340.1	YP_001658669.1	YP_001659403.1	YP_001658235.1	YP_001658803.1	NC_010296.1	YP_001660747.1	YP_001660750.1
<i>Cyanospora PCC782</i>	YP_003887441.1	YP_003888778.1	YP_003886543.1	YP_003890643.1	YP_003888132.1	YP_003887960.1	YP_003890216.1	YP_003890219.1
<i>Cyanospora PCC8801</i>	YP_002373016.1	YP_002371851.1	YP_002373207.1	YP_002372638.1	YP_002371810.1	YP_002371940.1	YP_002370501.1	YP_002370504.1
<i>Cyanospora parvula</i>	NP_043250.1	NP_043165.1	NP_043166.1	NP_043204.1	NP_043240.1	NP_043173.1	NP_043193.1	NP_043195.1
<i>Cyanospora salinarum</i>	missing	NP_045098.1	NP_045099.1	NP_045208.1	NP_045134.1	NP_045118.1	NP_045177.1	NP_045180.1
<i>Cyanospora merulae</i>	missing	NP_849037.1	NP_849036.1	NP_849050.1	NP_848945.1	NP_849019.1	NP_849079.1	NP_849076.1
<i>Paraphysa purpurea</i>	missing	NP_053934.1	NP_053933.1	NP_053926.1	NP_053836.1	NP_053949.1	NP_053917.1	NP_053917.1
<i>Gyrodinium aureolum</i>	missing	YP_063574.1	YP_063575.1	YP_063609.1	YP_063575.1	YP_063555.1	YP_063597.1	YP_063600.1
<i>Ostreococcus tauri</i>	missing	YP_717222.1	YP_717206.1	YP_717234.1	YP_717262.1	missing	YP_717240.1	YP_717239.1
<i>Pyraminococcus porfere</i>	YP_002600890.1	YP_002600890.1	YP_002600890.1	YP_002600952.1	YP_002600925.1	YP_002600904.1	YP_002600918.1	YP_002600919.1
<i>Pyraminococcus bleekeri</i>	missing	YP_002600855.1	YP_002600856.1	YP_002600807.1	YP_002600803.1	missing	YP_002600844.1	YP_002600843.1
<i>Microcystis Plect</i>	YP_002601046.1	YP_002601046.1	YP_002601045.1	YP_002601041.1	YP_002600999.1	missing	YP_002601058.1	YP_002601056.1
<i>Microcystis PCC99</i>	missing	YP_002808648.1	YP_002808603.1	YP_002808636.1	YP_002808616.1	missing	YP_002808630.1	YP_002808631.1
<i>Chlamydomonas reinhardtii</i>	NP_958382.1	NP_958386.1	NP_958387.1	NP_958387.1	NP_958405.1	missing	NP_958372.1	NP_958371.1
<i>Saccharina obliqua</i>	YP_635996.1	YP_635970.1	YP_635969.1	NP_635997.1	YP_635971.1	YP_635981.1	YP_636013.1	YP_636014.1
<i>Chlorella vulgaris</i>	NP_045843.1	NP_045847.1	NP_045848.1	NP_045822.1	NP_045897.1	NP_045915.1	NP_045928.1	NP_045929.1
<i>Bracteopsis lagynoides</i>	YP_003227026.1	YP_003227071.1	YP_003227068.1	YP_003227060.1	YP_003227031.1	YP_003227040.1	YP_003227048.1	YP_003227049.1
<i>Nepenthes alba</i>	NP_050808.1	NP_050808.1	NP_050808.1	NP_050873.1	NP_050889.1	NP_050815.1	NP_050823.1	NP_050824.1
<i>Chlorogobius atmophycus</i>	YP_007019130.1	YP_007019075.1	YP_007019074.1	YP_007019121.1	YP_007019152.1	missing	YP_007019087.1	YP_007019088.1
<i>Zygnema carinatum</i>	YP_636481.1	YP_636472.1	YP_636471.1	YP_636508.1	YP_636475.1	missing	YP_636492.1	YP_636491.1
<i>Aurea mirabilis</i>	YP_001687202.1	YP_001687240.1	YP_001687239.1	YP_001687219.1	YP_001687244.1	missing	YP_001687248.1	YP_001687249.1

Table S6: (continued)

Taxon	rp/19	rp/12	rp/20	rp/22	rp/23	rp/32	rp/33	rp/36
<i>Heliobacterium modesticaldum</i>	YP_001680755.1	YP_001679969.1	YP_001680367.1	YP_001679971.1	YP_001679981.1	YP_001680726.1	YP_001679950.1	YP_001679991.1
<i>Bacillus subtilis</i>	NP_389486.1	NP_388000.2	NP_390763.1	NP_388002.1	NP_387999.2	NP_389391.1	NP_387980.1	NP_388021.1
<i>Bdellovibrio bacteriovorus</i>	NP_968965.1	NP_969752.1	NP_968508.1	NP_969750.1	NP_969752.1	NP_968870.1	NP_976253.1	NP_976254.1
<i>Neorhombium aromatiatorans</i>	YP_496886.1	YP_496530.1	YP_495967.1	YP_496532.1	YP_496529.1	YP_497638.1	YP_498457.1	YP_498542.1
<i>Rhodospirillum rubrum</i>	YP_779270.1	YP_782493.1	YP_779436.1	YP_782494.1	YP_782494.1	YP_779114.1	YP_782295.1	YP_783021.1
<i>Chlorobium limicola</i>	YP_001943434.1	YP_001944229.1	YP_001942231.1	YP_001944227.1	YP_001944230.1	YP_001942249.1	YP_001943582.1	YP_001944209.1
<i>Herpetosiphon aurantiacus</i>	YP_001543006.1	YP_001547676.1	YP_001547597.1	YP_001547678.1	YP_001547678.1	YP_001543249.1	YP_001542899.1	NC_009972.1
<i>Roseiflexus castellanii</i>	YP_001432657.1	YP_001434074.1	YP_001431549.1	YP_001434072.1	YP_001434075.1	YP_001431914.1	YP_001433402.1	YP_001434054.1
<i>Gloeobacter violaceus</i>	NP_923774.1	NP_923849.1	NP_923730.1	NP_926868.1	NP_923833.1	NP_923738.1	NP_926006.1	NP_926521.1
<i>Synechococcus PCC7002</i>	YP_001734291.1	YP_001734320.1	YP_001735416.1	YP_001734318.1	YP_001734321.1	YP_001735138.1	YP_001738835.1	NC_010475.1
<i>Synechococcus JA33Ab</i>	YP_475042.1	YP_474624.1	YP_476177.1	YP_474624.1	YP_474627.1	YP_474586.1	YP_475071.1	NC_010485.1
<i>Acaryochloris marina</i>	YP_001517177.1	YP_001518989.1	YP_001516452.1	YP_001518989.1	YP_001518987.1	YP_001517257.1	YP_001519102.1	NC_009925.1
<i>Synechococcus elongatus</i>	YP_172279.1	YP_172578.1	YP_170985.1	YP_172580.1	YP_172577.1	YP_171250.1	YP_171137.1	NC_009925.1
<i>Synechococcus RCG807</i>	YP_001227920.1	YP_001228375.1	YP_001226315.1	YP_001228377.1	YP_001228374.1	YP_001227744.1	YP_001227480.1	YP_001228392.1
<i>Prochlorococcus marinus</i>	YP_001090724.1	YP_001091970.1	YP_001092076.1	YP_001091968.1	YP_001091971.1	YP_001091239.1	YP_001091214.1	NC_009091.1
<i>Nostoc punctiforme</i>	YP_001869089.1	YP_001867703.1	YP_001868225.1	YP_001867701.1	YP_001867704.1	YP_001864523.1	YP_001868229.1	YP_001867685.1
<i>Nostoc PCC7120</i>	NP_489337.1	NP_488252.1	NP_487468.1	NP_488250.1	NP_488253.1	NP_487714.1	NP_488492.1	NP_488234.1
<i>Thermosynechococcus erythraeus</i>	YP_720224.1	YP_722637.1	YP_721611.1	YP_722635.1	YP_722638.1	NP_720634.1	NP_723257.1	YP_722619.1
<i>Thermosynechococcus elongatus</i>	NP_682093.1	NP_680875.1	NP_682948.1	NP_680877.1	NP_680874.1	NP_682705.1	NP_682850.1	NP_680893.1
<i>Cyanospora PCC7795</i>	YP_002481192.1	YP_002482033.1	YP_002482209.1	YP_002482035.1	YP_002482032.1	YP_002484734.1	YP_002481650.1	YP_002482051.1
<i>Synechocystis PCC6803</i>	NP_440742.1	NP_440666.1	NP_442051.1	NP_440664.1	NP_440667.1	NP_440354.1	NP_442438.1	NP_440648.1
<i>Microcystis aeruginosa</i>	YP_001658673.1	YP_001660754.1	YP_001659348.1	YP_001660752.1	YP_001660755.1	NC_010296.1	YP_001659181.1	NC_010296.1
<i>Cyanospora PCC782</i>	YP_003887966.1	YP_003890223.1	YP_003889179.1	YP_003890221.1	YP_003890224.1	YP_003886492.1	YP_003887484.1	YP_003890205.1
<i>Cyanospora PCC8801</i>	YP_002371946.1	YP_002370508.1	YP_002372766.1	YP_002370506.1	YP_002370509.1	YP_002372851.1	YP_002371877.1	YP_002370490.1
<i>Cyanospora paradoxa</i>	NP_043156.1	NP_043199.1	NP_043162.1	NP_043197.1	NP_043197.1	missing	NP_043202.1	NP_043259.1
<i>Cyanidium caldarium</i>	NP_045111.1	NP_045184.1	NP_045199.1	NP_045182.1	NP_045185.1	NP_045066.1	NP_045028.1	NP_045169.1
<i>Cyanidiosclerotium mirabile</i>	NP_849022.1	NP_849072.1	NP_849057.1	NP_849074.1	NP_849071.1	NP_848969.1	NP_849114.1	NP_849087.1
<i>Paraphysa purpurea</i>	NP_053941.1	NP_053921.1	NP_053879.1	NP_053919.1	NP_053922.2	NP_053817.1	NP_053865.1	NP_053906.1
<i>Gyrodinium aureolum</i>	NP_063567.1	NP_063604.1	NP_063626.1	NP_063602.1	NP_063605.1	NP_063526.1	NP_063640.1	NP_063589.1
<i>Ostreococcus tauri</i>	missing	YP_717236.1	YP_717258.1	XM_003084057	YP_717235.1	YP_717210.1	missing	YP_717208.1
<i>Pyraminococcus porfere</i>	YP_002600923.1	YP_002600923.1	YP_002600935.1	YP_002600921.1	YP_002600924.1	YP_002600962.1	YP_002600914.1	YP_002600914.1
<i>Pyrococcus furiosus</i>	YP_002600852.1	YP_002600840.1	YP_002600864.1	missing	YP_002600839.1	missing	YP_002608411.1	YP_002608411.1
<i>Microcystis Plect</i>	missing	YP_002601052.1	YP_002600996.1	missing	YP_002601050.1	YP_002601064.1	missing	YP_002601005.1
<i>Micromonas PCC99</i>	XM_002504578	YP_002504578	YP_002504612	missing	YP_002508635	YP_002508658	missing	missing
<i>Chlamydomonas reinhardtii</i>	missing	NP_958369.1	NP_958363.1	missing	NP_958368.1	NP_958368.1	missing	NP_958367.1
<i>Saccharomyces cerevisiae</i>	missing	NP_635937.1	NP_635975.1	missing	NP_635988.2	NP_045889.1	missing	NP_636004.1
<i>Chlorella vulgaris</i>	NP_045812.1	NP_045933.1	NP_045805.1	NP_045975.1	NP_045984.1	NP_045889.1	missing	NP_045923.1
<i>Bradyrhodospira</i>	NP_013359.1	YP_003227052.1	YP_003227025.1	missing	YP_003227038.1	YP_003227038.1	missing	NP_003227044.1
<i>Neptostichus obsoletus</i>	NP_050851.1	NP_050827.1	NP_050871.1	missing	NP_050828.1	NP_050927.1	missing	NP_050819.1
<i>Chlorogobius amphiplus</i>	YP_001019140.1	YP_001019092.1	YP_001019133.1	YP_001019090.1	YP_001019093.1	YP_001019169.1	YP_001019135.1	YP_001019083.1
<i>Zygnema caricatum</i>	missing	YP_636488.1	YP_636467.1	missing	YP_636487.1	missing	YP_636453.1	YP_636496.1
<i>Aurea anabitis</i>	missing	YP_001687233.1	YP_001687237.1	YP_001687251.1	YP_001687254.1	YP_001687256.1	YP_001687235.1	YP_001687245.1

Table S6: (continued)

Table S6: (continued)

Taxon	tps18	tps19	tps2	tps3	tps4	tps7	tps8	tps9
<i>Halobacterium modesticaldum</i>	YP_001680277.1	YP_001679970.1	YP_001680823.1	YP_001679972.1	YP_001679941.1	YP_001679962.1	YP_001679980.1	YP_001680002.1
<i>Bacillus subtilis</i>	NP_391969.2	NP_388001.1	NP_389531.1	NP_388003.2	NP_390844.1	NP_387992.2	NP_388011.2	NP_388031.1
<i>Bdellovibrio bacteriovorus</i>	NP_967887.1	NP_969751.1	NP_970487.1	NP_969749.1	NP_969733.1	NP_969758.1	NP_969743.1	NP_967480.1
<i>Novosphingobium aromaticivorans</i>	YP_496693.1	YP_496531.1	YP_496647.1	YP_496533.1	YP_497516.1	YP_496523.1	YP_496541.1	YP_498554.1
<i>Rhodospseudomonas palustris</i>	YP_782223.1	YP_782492.1	YP_781471.1	YP_782490.1	YP_780560.1	YP_782500.1	YP_782482.1	YP_781775.1
<i>Chlorobium limicola</i>	YP_001942227.1	YP_001944228.1	YP_001942496.1	YP_001944226.1	YP_001944206.1	YP_001944236.1	YP_001944218.1	YP_001942495.1
<i>Herpetosiphon aurantiacus</i>	YP_001547722.1	YP_001547677.1	YP_001547719.1	YP_001547697.1	YP_001547697.1	YP_001546604.1	YP_001547686.1	YP_001547702.1
<i>Roseiflexus castellanii</i>	YP_001431639.1	YP_001431407.1	YP_001433976.1	YP_001434071.1	YP_001434051.1	YP_001434081.1	YP_001434063.1	YP_001434046.1
<i>Gloeobacter violaceus</i>	NP_926005.1	NP_923850.1	NP_924776.1	NP_926867.1	NP_926518.1	NP_926872.1	NP_926859.1	NP_927365.1
<i>Synechococcus PCC7002</i>	YP_001733836.1	YP_001734319.1	YP_001735202.1	YP_001734317.1	YP_001735324.1	YP_001735305.1	YP_001734310.1	YP_001734296.1
<i>Synechococcus JA33Ab</i>	YP_475070.1	YP_474625.1	YP_475059.1	YP_474623.1	YP_474947.1	YP_474744.1	YP_474616.1	YP_474878.1
<i>Acaryochloris marina</i>	YP_001519101.1	YP_001518990.1	YP_001517738.1	YP_001518992.1	YP_001517738.1	YP_001516131.1	YP_001518999.1	YP_001515600.1
<i>Synechococcus elongatus</i>	YP_171136.1	YP_172579.1	YP_172290.1	YP_172581.1	YP_173183.1	YP_171364.1	YP_172588.1	YP_172602.1
<i>Synechococcus RCC907</i>	YP_001227481.1	YP_001228376.1	YP_001227530.1	YP_001228378.1	YP_001226877.1	YP_001228424.1	YP_001228385.1	YP_001228399.1
<i>Prochlorococcus marinus</i>	YP_001091215.1	YP_001091969.1	YP_001091037.1	YP_001091967.1	YP_001090653.1	YP_001091925.1	YP_001091960.1	YP_001091947.1
<i>Nostoc punctiforme</i>	YP_001868228.1	YP_001867702.1	YP_00186818.1	YP_001867700.1	NP_001865332.1	YP_001867209.1	YP_001867693.1	YP_001867678.1
<i>Nostoc PCC7120</i>	NP_488491.1	NP_488251.1	NP_488832.1	NP_488249.1	NP_486777.1	NP_488379.1	NP_488242.1	NP_488227.1
<i>Trichodesmium erythraeum</i>	YP_723258.1	YP_722636.1	YP_720995.1	YP_722634.1	YP_720272.1	YP_720403.1	YP_722627.1	YP_722612.1
<i>Thermosynechococcus elongatus</i>	NP_682851.1	NP_680876.1	NP_682478.1	NP_680878.1	NP_680938.1	NP_682538.1	NP_680885.1	NP_680900.1
<i>Cyanospora PCC71425</i>	YP_002481649.1	YP_002482034.1	YP_002484910.1	YP_002482036.1	YP_002481751.1	YP_002481458.1	YP_002482043.1	YP_002482058.1
<i>Synechocystis PCC6803</i>	NP_442439.1	NP_440665.1	NP_441467.1	NP_440663.1	NP_442284.1	NP_441643.1	NP_440656.1	NP_440641.1
<i>Microcystis aeruginosa</i>	YP_001659819.1	YP_001660753.1	YP_001659507.1	YP_001660751.1	YP_001658257.1	YP_001659292.1	YP_001660744.1	YP_001660264.1
<i>Cyanospora PCC7822</i>	YP_003887483.1	YP_003890222.1	YP_003886505.1	YP_003890220.1	YP_003889427.1	YP_003887390.1	YP_003890213.1	YP_003890198.1
<i>Cyanospora PCC8801</i>	YP_002371876.1	YP_002370507.1	YP_002372740.1	YP_002370505.1	YP_002374484.1	YP_002371747.1	YP_002370498.1	YP_002370483.1
<i>Cyanophora paradoxa</i>	NP_043203.1	NP_043198.1	NP_043227.1	NP_043196.1	NP_043212.1	NP_043208.1	NP_043191.1	NP_043255.1
<i>Cyanidium caldarium</i>	NP_045027.1	NP_045183.1	NP_045034.1	NP_045181.1	NP_045130.1	NP_045161.1	NP_045174.1	NP_045164.1
<i>Cyanidioschyzon merolae</i>	NP_849113.1	NP_849073.1	NP_849118.1	NP_849075.1	NP_848941.1	NP_849095.2	NP_849082.2	NP_849092.1
<i>Porphyra purpurea</i>	NP_053866.1	NP_053920.1	NP_053859.1	NP_053918.1	NP_053829.1	NP_053898.1	NP_053911.1	NP_053901.1
<i>Gracilaria tenuistipitata</i>	YP_063639.1	YP_063603.1	YP_063645.1	YP_063601.1	NP_063676.1	YP_063581.1	YP_063594.1	NP_063584.1
<i>Ostreococcus tauri</i>	YP_717225.1	YP_717237.1	YP_717232.1	YP_717238.1	YP_717223.1	YP_717227.1	YP_717242.1	YP_717246.1
<i>Pyramimonas parkae</i>	YP_002600903.1	YP_002600922.1	YP_002600951.1	YP_002600920.1	YP_002600908.1	YP_002600937.1	YP_002600916.1	YP_002600905.1
<i>Pycnococcus provasolii</i>	YP_002600865.1	YP_002600841.1	YP_002600867.1	YP_002600842.1	YP_002600860.1	YP_002600824.1	YP_002600846.1	missing
<i>Monomastix okei</i>	YP_002601063.1	YP_002601053.1	YP_002601019.1	YP_002601055.1	YP_002601030.1	YP_002600991.1	YP_002601061.1	missing
<i>Micromonas RCC299</i>	YP_002808645	YP_002808633	YP_002808638	YP_002808632	YP_002808647	YP_002808643	YP_002808628	YP_002808624
<i>Chlamydomonas reinhardtii</i>	NP_958392.1	NP_958370.1	NP_958392.1	NP_958402.1	NP_958376.1	NP_958380.1	NP_958374.2	NP_958395.1
<i>Scenedesmus obliquus</i>	NP_635993.1	NP_635986.1	NP_635981.1	NP_635947.1	NP_635989.1	NP_635960.1	NP_636011.1	NP_636008.1
<i>Chlorella vulgaris</i>	NP_045806.1	NP_045801.1	NP_045931.1	NP_045930.1	NP_045904.1	NP_045810.1	NP_045925.1	NP_045919.1
<i>Bryopsis hypnoides</i>	YP_003227024.1	YP_003227031.1	YP_003227037.1	YP_003227050.1	NP_003227080.1	YP_003227027.1	YP_003227046.1	YP_003227039.1
<i>Nephroselmis olivacea</i>	NP_050867.1	NP_050826.1	NP_050842.1	NP_050825.1	NP_050870.1	NP_050849.1	NP_050821.1	NP_050816.1
<i>Chlorokibbutia amphiphycus</i>	YP_001019134.1	YP_001019091.1	YP_001019101.1	YP_001019089.1	YP_001019095.1	YP_001019138.1	YP_001019085.1	YP_001019080.1
<i>Zygnema circumcinctatum</i>	YP_636554.1	YP_636489.1	YP_636476.1	YP_636490.1	YP_636511.1	YP_636479.1	YP_636494.1	missing
<i>Aurea mirabilis</i>	YP_001687236.1	YP_001687252.1	YP_001687207.1	YP_001687207.1	YP_001687221.1	YP_001687201.1	YP_001687247.1	missing
a. Join(complement(NC_005353.1..88116..89828), complement(NC_005353.1..87098..87724)), b. Join(NC_008101.1..54809..55948)								

Table S6: (continued)

Taxon	tufA	ycf12	ycf3	ycf4
<i>Heliobacterium modesticaldum</i>	YP_001679964.1	missing	missing	missing
<i>Bacillus subtilis</i>	NP_387994.1	missing	missing	missing
<i>Bdellovibrio bacteriovorus</i>	NP_969770.1	missing	missing	missing
<i>Nonophthodinium aromatiiciorans</i>	YP_496525.1	missing	missing	missing
<i>Rhodospirillum rubrum</i>	YP_782498.1	missing	missing	missing
<i>Chlorobium limicola</i>	YP_001944234.1	missing	missing	missing
<i>Herpetosiphon aurantiacus</i>	YP_001542898.1	missing	missing	missing
<i>Roseiflexus castellanii</i>	YP_001434079.1	missing	missing	missing
<i>Gloeobacter violaceus</i>	NP_926874.1	missing	missing	missing
<i>Synechococcus PCC7002</i>	YP_001735303.1	YP_001734240.1	NP_923661.1	NP_923025.1
<i>Synechococcus JAS3Ab</i>	YP_474746.1	YP_474024.1	YP_001734813.1	YP_001734351.1
<i>Acaryochloris marina</i>	YP_001516129.1	missing	YP_474781.1	YP_474915.1
<i>Synechococcus elongatus</i>	YP_171366.1	missing	YP_172491.1	YP_171584.1
<i>Synechococcus RCG807</i>	YP_001228426.1	missing	YP_001226508.1	YP_001227950.1
<i>Prochlorococcus marinus</i>	YP_001091923.1	missing	YP_001090375.1	YP_001091572.1
<i>Nostoc punctiforme</i>	YP_001867211.1	missing	YP_001866872.1	YP_001866971.1
<i>Nostoc PCC7120</i>	NP_488377.1	NP_489168.1	NP_487233.1	NP_488329.1
<i>Thermosynechococcus erythraeus</i>	YP_720405.1	NP_722001.1	YP_724029.1	YP_720440.1
<i>Thermosynechococcus elongatus</i>	NP_682540.1	NP_682032.1	NP_682315.1	NP_682178.1
<i>Cyanospora PCC7795</i>	YP_002481456.1	YP_002485210.1	YP_002481304.1	YP_002481804.1
<i>Synechocystis PCC6803</i>	NP_441641.1	NP_442709.1	NP_441459.1	NP_440032.1
<i>Microcystis aeruginosa</i>	YP_001659290.1	YP_001655272.1	YP_001659918.1	YP_001659495.1
<i>Cyanospora PCC782</i>	YP_003887393.1	YP_003888325.1	YP_003889979.1	YP_003886962.1
<i>Cyanospora PCC8801</i>	YP_002371745.1	YP_002373709.1	YP_002372338.1	YP_002370789.1
<i>Cyanospora paradoxa</i>	NP_043207.1	NP_043213.1	NP_043154.1	NP_043147.1
<i>Cyanidium caldarium</i>	NP_045160.1	NP_045148.1	NP_045024.1	NP_045048.1
<i>Cyanidium ruber</i>	NP_849096.1	NP_849037.1	NP_849102.1	NP_849132.1
<i>Porphyra purpurea</i>	NP_063897.1	NP_053995.1	NP_053868.1	NP_053830.1
<i>Gyrodinium aureolum</i>	NP_063580.1	NP_063700.1	NP_063637.1	NP_063665.1
<i>Ostreococcus tauri</i>	YP_717228.1	YP_717256.1	YP_717244.1	XM_603079929
<i>Pyraminococcus porfereae</i>	YP_002600938.1	YP_002600901.1	YP_002600911.1	YP_002600939.1
<i>Pyraminococcus pascuensis</i>	YP_002600853.1	YP_002600806.1	YP_002600859.1	YP_002600837.1
<i>Microcystis Pect1</i>	YP_002600997.1	YP_002601043.1	YP_002601000.1	missing
<i>Microcystis PCC999</i>	YP_002808642.1	YP_002808610.1	YP_002808646.1	missing
<i>Chlamydomonas reinhardtii</i>	NP_958362.1	NP_958378.1	NP_958393.1	NP_958394.1
<i>Saccharomyces cerevisiae</i>	YP_635995.1	YP_635961.1	YP_636006.1	YP_636007.1
<i>Chlorella vulgaris</i>	NP_045811.1	NP_045864.1	NP_045789.1	NP_045813.1
<i>Bracteolaxys pygmaea</i>	YP_003227082.1	YP_003227073.1	YP_003227023.1	YP_003227029.1
<i>Nepenthes albertiana</i>	NP_050850.1	NP_050859.1	NP_050843.1	NP_050852.1
<i>Chlorogobius amphipterus</i>	YP_001019139.1	YP_001019106.1	YP_001019153.1	YP_001019142.1
<i>Zigana curvicaudata</i>	missing	YP_636485.1	YP_636512.2	YP_636560.1
<i>Aurea mobilis</i>	missing	YP_001687212.1	missing	YP_001687227.1

Table S7: Selected loci The 75 markers retained in this study. For each marker, the table gives the model used in the RAxML and P4 NDCH analyses, for nucleotide ('nuc') and amino-acid ('aa') datasets. 'I' indicates that the model included a proportion of variable sites. 'Γ' indicates that the model included four gamma-distributed discrete evolutionary rate categories. In the case of amino-acid evolution models, 'F' indicates that the amino-acid frequencies are estimated from the data instead of taking the frequencies implied by the empirical transition matrix. In the case of P4 NDCH analyses, the optimal number of composition vectors is given ('CV').

Marker	RAxML model (nuc)	RAxML model (aa)	P4 model (nuc)	P4 model (aa)
<i>accD</i>	GTR+I+Γ	WAG+Γ	GTR+I+Γ+4CV	LG+Γ+3CV
<i>atpA</i>	GTR+I+Γ	WAG+I+Γ	GTR+I+Γ+2CV	LG+I+Γ+1CV
<i>atpB</i>	GTR+I+Γ	WAG+I+Γ	GTR+I+Γ+2CV	LG+I+Γ+1CV
<i>atpE</i>	GTR+I+Γ	CPREV+I+Γ	SYM+I+Γ+3CV	LG+I+Γ+2CV
<i>atpF</i>	GTR+I+Γ	RTREV+Γ	GTR+I+Γ+6CV	LG+Γ+F+6CV
<i>atpH</i>	GTR+Γ	RTREV+Γ	GTR+Γ+2CV	LG+Γ+F+1CV
<i>atpI</i>	GTR+Γ	RTREV+Γ	GTR+Γ+2CV	LG+Γ+1CV
<i>ccsA</i>	GTR+I+Γ	CPREV+I+Γ	GTR+I+Γ+4CV	CPREV+I+Γ+F+2CV
<i>cemA</i>	GTR+I+Γ	CPREV+I+Γ	GTR+I+Γ+3CV	CPREV+I+Γ+3CV
<i>chlB</i>	GTR+I+Γ	CPREV+Γ	GTR+I+Γ+2CV	LG+Γ+1CV
<i>chlI</i>	GTR+I+Γ	WAG+I+Γ	GTR+I+Γ+4CV	LG+I+Γ+1CV
<i>chlL</i>	GTR+I+Γ	RTREV+Γ	GTR+I+Γ+2CV	LG+Γ+1CV
<i>chlN</i>	GTR+I+Γ	WAG+Γ	GTR+I+Γ+2CV	LG+Γ+1CV
<i>clpP</i>	GTR+I+Γ	RTREV+Γ	GTR+I+Γ+3CV	LG+Γ+1CV
<i>ftsH</i>	GTR+I+Γ	RTREV+I+Γ	GTR+I+Γ+4CV	LG+I+Γ+4CV
<i>infA</i>	GTR+I+Γ	RTREV+Γ	SYM+I+Γ+2CV	LG+Γ+1CV
<i>petA</i>	GTR+I+Γ	WAG+I+Γ	GTR+I+Γ+3CV	LG+I+Γ+2CV
<i>petB</i>	GTR+Γ	RTREV+Γ	GTR+Γ+2CV	LG+Γ+1CV
<i>petD</i>	GTR+I+Γ	CPREV+Γ	GTR+I+Γ+2CV	LG+Γ+1CV
<i>petG</i>	GTR+I+Γ	CPREV+Γ	GTR+I+Γ+2CV	CPREV+Γ+1CV
<i>petL</i>	GTR+Γ	JTT+Γ	HKY+Γ+2CV	MTART+Γ+F+1CV
<i>petN</i>	GTR+I+Γ	BLOSUM62+Γ	GTR+I+Γ+2CV	MTART+Γ+1CV
<i>psaA</i>	GTR+I+Γ	CPREV+I+Γ	GTR+I+Γ+3CV	LG+I+Γ+F+1CV
<i>psaB</i>	GTR+I+Γ	CPREV+I+Γ	GTR+I+Γ+3CV	LG+I+Γ+1CV
<i>psaC</i>	GTR+I+Γ	DAYHOFF+I+Γ	GTR+I+Γ+2CV	D78+I+Γ+1CV
<i>psaI</i>	GTR+Γ	CPREV+Γ	GTR+Γ+2CV	MTART+Γ+1CV
<i>psaJ</i>	GTR+I+Γ	RTREV+Γ	GTR+I+Γ+2CV	MTART+Γ+1CV
<i>psaM</i>	GTR+Γ	CPREV+Γ	GTR+Γ+4CV	LG+Γ+1CV
<i>psbA</i>	GTR+I+Γ	CPREV+I+Γ	GTR+I+Γ+3CV	LG+I+Γ+1CV
<i>psbB</i>	GTR+I+Γ	CPREV+Γ	GTR+I+Γ+3CV	LG+Γ+1CV
<i>psbC</i>	GTR+I+Γ	CPREV+Γ	GTR+I+Γ+3CV	LG+Γ+1CV
<i>psbD</i>	GTR+I+Γ	CPREV+I+Γ	GTR+I+Γ+3CV	CPREV+I+Γ+1CV
<i>psbE</i>	GTR+I+Γ	WAG+Γ	GTR+I+Γ+2CV	LG+Γ+1CV
<i>psbF</i>	GTR+I+Γ	CPREV+Γ	GTR+I+Γ+3CV	CPREV+Γ+1CV
<i>psbH</i>	GTR+I+Γ	CPREV+Γ	GTR+I+Γ+3CV	CPREV+Γ+1CV
<i>psbI</i>	GTR+I+Γ	MTREV+Γ	GTR+I+Γ+2CV	MTART+Γ+1CV
<i>psbJ</i>	GTR+I+Γ	CPREV+Γ	GTR+I+Γ+2CV	LG+Γ+1CV
<i>psbK</i>	GTR+Γ	CPREV+Γ	GTR+Γ+2CV	CPREV+Γ+1CV
<i>psbL</i>	GTR+I+Γ	CPREV+Γ	GTR+I+Γ+2CV	LG+Γ+1CV
<i>psbM</i>	GTR+Γ	CPREV+Γ	HKY+Γ+2CV	CPREV+Γ+1CV
<i>psbN</i>	GTR+I+Γ	CPREV+I+Γ	HKY+I+Γ+2CV	LG+I+Γ+1CV
<i>psbT</i>	GTR+Γ	CPREV+Γ	HKY+Γ+2CV	LG+Γ+1CV
<i>psbZ</i>	GTR+I+Γ	CPREV+Γ	GTR+I+Γ+2CV	LG+Γ+F+1CV
<i>rbcL</i>	GTR+I+Γ	RTREV+I+Γ	GTR+I+Γ+3CV	LG+I+Γ+1CV
<i>rpl2</i>	GTR+I+Γ	RTREV+I+Γ	GTR+I+Γ+4CV	LG+I+Γ+1CV
<i>rpl5</i>	GTR+I+Γ	RTREV+Γ	GTR+I+Γ+3CV	LG+Γ+2CV

Marker	RAxML model (nuc)	RAxML model (aa)	P4 model (nuc)	P4 model (aa)
<i>rpl12</i>	GTR+I+Γ	RTREV+Γ	GTR+I+Γ+3CV	LG+Γ+F+2CV
<i>rpl14</i>	GTR+I+Γ	RTREV+I+Γ	GTR+I+Γ+4CV	LG+I+Γ+1CV
<i>rpl16</i>	GTR+I+Γ	RTREV+Γ	GTR+I+Γ+2CV	LG+Γ+1CV
<i>rpl19</i>	GTR+I+Γ	CPREV+I+Γ	GTR+I+Γ+4CV	LG+I+Γ+2CV
<i>rpl20</i>	GTR+I+Γ	JTT+Γ	GTR+I+Γ+2CV	LG+Γ+2CV
<i>rpl22</i>	GTR+I+Γ	WAG+Γ	GTR+I+Γ+3CV	LG+Γ+2CV
<i>rpl23</i>	GTR+I+Γ	WAG+Γ	GTR+I+Γ+3CV	LG+Γ+4CV
<i>rpl32</i>	GTR+Γ	RTREV+Γ	HKY+Γ+2CV	LG+Γ+6CV
<i>rpl33</i>	GTR+I+Γ	JTT+Γ	GTR+I+Γ+2CV	LG+Γ+3CV
<i>rpl36</i>	GTR+Γ	CPREV+Γ	GTR+Γ+2CV	LG+Γ+1CV
<i>rpoA</i>	GTR+I+Γ	RTREV+I+Γ	GTR+I+Γ+2CV	LG+I+Γ+2CV
<i>rpoB</i>	GTR+I+Γ	CPREV+I+Γ	GTR+I+Γ+3CV	LG+I+Γ+4CV
<i>rpoC1</i>	GTR+I+Γ	CPREV+I+Γ	GTR+I+Γ+3CV	LG+I+Γ+1CV
<i>rpoC2</i>	GTR+I+Γ	CPREV+I+Γ	GTR+I+Γ+2CV	LG+I+Γ+2CV
<i>rps2</i>	GTR+I+Γ	CPREV+I+Γ	GTR+I+Γ+3CV	LG+I+Γ+2CV
<i>rps3</i>	GTR+I+Γ	RTREV+I+Γ	GTR+I+Γ+2CV	LG+I+Γ+1CV
<i>rps4</i>	GTR+I+Γ	RTREV+I+Γ	GTR+I+Γ+3CV	LG+I+Γ+2CV
<i>rps7</i>	GTR+I+Γ	CPREV+Γ	GTR+I+Γ+3CV	LG+Γ+1CV
<i>rps8</i>	GTR+I+Γ	CPREV+I+Γ	GTR+I+Γ+4CV	LG+I+Γ+1CV
<i>rps9</i>	GTR+I+Γ	RTREV+I+Γ	GTR+I+Γ+3CV	LG+I+Γ+4CV
<i>rps11</i>	GTR+I+Γ	CPREV+Γ	GTR+I+Γ+2CV	LG+Γ+1CV
<i>rps12</i>	GTR+I+Γ	JTT+I+Γ	GTR+I+Γ+4CV	LG+I+Γ+1CV
<i>rps14</i>	GTR+I+Γ	WAG+I+Γ	GTR+I+Γ+3CV	LG+Γ+1CV
<i>rps18</i>	GTR+I+Γ	CPREV+Γ	GTR+I+Γ+5CV	LG+Γ+3CV
<i>rps19</i>	GTR+I+Γ	WAG+Γ	GTR+I+Γ+2CV	LG+Γ+1CV
<i>tufA</i>	GTR+I+Γ	WAG+I+Γ	GTR+I+Γ+3CV	LG+I+Γ+1CV
<i>ycf3</i>	GTR+I+Γ	CPREV+I+Γ	GTR+I+Γ+2CV	LG+I+Γ+1CV
<i>ycf4</i>	GTR+I+Γ	CPREV+I+Γ	GTR+I+Γ+8CV	CPREV+I+Γ+5CV
<i>ycf12</i>	GTR+I+Γ	CPREV+Γ	GTR+I+Γ+2CV	LG+Γ+1CV