Supplementary Table 1. Pairwise Average Nucleotide Identity (ANI). Values above 80% are highlighted in pink. The genome names are color-coded as indicated in Figure 1A. See Materials and Methods for details on ANI values calculation.

	CCMP1375	**CCMP1986**	**MIT9313**	**MIT9303**	**MIT9312**	**MIT9515**	**NATL1A**	**NATL2A**	**CC9311**	**CC9605**	**CC9902**	**WH8102**	**MIT9211**	**MIT9215**	**RCC307**	**WH7803**	**MIT9301**	**PCC7002**
P. marinus AS9601	68.25	78.81	66.20	66.25	90.33	78.31	68.53	68.51	65.92	64.79	65.60	64.77	68.18	91.62	64.64	65.23	93.76	63.91
P. marinus CCMP1375	-	68.09	66.84	66.98	68.28	68.16	69.27	69.32	66.36	65.66	66.29	65.70	71.33	68.24	65.23	65.85	68.24	64.23
P. marinus CCMP1986	-	-	65.98	66.04	79.21	86.45	68.47	68.45	65.83	64.48	65.47	64.66	68.23	78.77	64.47	64.89	78.99	64.00
P. marinus MIT9313	-	-	-	96.06	66.30	66.00	66.74	66.84	69.59	70.01	69.49	70.18	67.10	66.15	69.04	70.03	66.23	65.19
P. marinus MIT9303	-	-	-	-	66.30	65.88	66.83	66.82	69.68	70.03	69.46	70.14	67.16	66.20	69.06	69.97	66.12	65.16
P. marinus MIT9312	-	-	-	-	-	78.61	68.38	68.37	65.87	64.74	65.58	64.72	68.14	89.59	64.47	65.23	90.22	64.04
P. marinus MIT9515	-	-	-	-	-	-	68.34	68.40	65.77	64.59	65.40	64.76	68.13	78.45	64.46	65.04	78.28	64.19
P. marinus NATL1A	-	-	-	-	-	-	-	97.01	66.32	65.70	66.47	65.75	69.17	68.43	65.35	65.83	68.52	64.19
P. marinus NATL2A	-	-	-	-	-	-	-	-	66.30	65.76	66.38	65.59	69.22	68.52	65.33	66.03	68.50	64.08
Synechococcus CC9311	-	-	-	-	-	-	-	-	-	70.73	70.53	70.87	66.58	65.92	69.63	71.73	65.96	65.46
Synechococcus CC9605	-	-	-	-	-	-	-	-	-	-	73.44	76.62	65.90	64.77	71.47	73.04	64.85	65.69
Synechococcus CC9902	-	-	-	-	-	-	-	-	-	-	-	73.14	66.38	65.50	69.79	70.88	65.74	65.31
Synechococcus WH8102	-	-	-	-	-	-	-	-	-	-	-	-	65.75	64.86	71.35	73.22	64.88	65.59
P. marinus MIT9211	-	-	-	-	-	-	-	-	-	-	-	-	-	68.13	65.40	65.97	68.08	64.15
P. marinus MIT9215	-	-	-	-	-	-	-	-	-	-	-	-	-	-	64.59	64.86	91.30	63.79
Synechococcus RCC307	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	71.29	64.58	65.68
Synechococcus WH7803	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	65.06	65.63
P. marinus MIT9301	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	63.80
Synechococcus PCC7002	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Supplementary Table 2. Pairwise Average Amino acid Identity (AAI). Values above 80% are highlighted in pink. The genome names are color-coded as indicated in Figure 1A. See Materials and Methods on AAI values calculation.

	CCMP1375	**CCMP1986**	**MIT9313**	**MIT9303**	**MIT9312**	**MIT9515**	**NATL1A**	**NATL2A**	**CC9311**	**CC9605**	**CC9902**	**WH8102**	**MIT9211**	**MIT9215**	**RCC307**	**WH7803**	**MIT9301**	**PCC7002**
P. marinus AS9601	60.08	79.86	58.60	58.62	90.30	79.44	60.20	60.41	57.91	57.74	57.76	57.67	60.10	91.84	56.43	57.70	93.84	50.74
P. marinus CCMP1375	-	59.86	65.54	65.56	60.08	60.03	65.06	65.26	63.18	63.16	62.95	63.16	69.58	60.02	60.18	63.17	60.20	51.51
P. marinus CCMP1986	-	-	58.53	58.48	80.26	87.09	60.34	60.34	57.58	57.60	57.72	57.65	60.10	79.71	56.25	57.53	80.22	50.63
P. marinus MIT9313	-	-	-	95.02	58.62	58.44	64.00	63.92	70.43	70.09	69.68	70.49	66.15	58.41	64.76	70.20	58.55	52.05
P. marinus MIT9303	-	-	-	-	58.34	58.43	63.75	63.72	70.28	69.77	69.51	70.16	66.06	58.48	64.62	69.71	58.54	51.92
P. marinus MIT9312	-	-	-	-	-	79.53	60.28	60.41	57.80	57.68	57.68	57.78	60.05	89.38	56.34	57.72	90.44	50.70
P. marinus MIT9515	-	-	-	-	-	-	60.43	60.36	57.71	57.52	57.52	57.44	60.32	79.66	56.20	57.71	79.38	50.81
P. marinus NATL1A	-	-	-	-	-	-	-	96.53	62.22	62.03	61.99	61.74	65.17	60.20	59.52	61.98	60.22	51.94
P. marinus NATL2A	-	-	-	-	-	-	-	-	62.37	62.10	62.06	61.88	65.22	60.34	59.57	61.97	60.40	51.85
Synechococcus CC9311	-	-	-	-	-	-	-	-	-	70.51	70.34	71.09	63.41	57.68	65.00	74.16	57.73	52.11
Synechococcus CC9605	-	-	-	-	-	-	-	-	-	-	77.62	78.96	63.74	57.78	65.95	71.86	57.86	52.80
Synechococcus CC9902	-	-	-	-	-	-	-	-	-	-	-	77.15	63.68	57.76	65.70	70.68	57.74	52.61
Synechococcus WH8102	-	-	-	-	-	-	-	-	-	-	-	-	63.66	57.82	66.36	72.29	57.83	53.06
P. marinus MIT9211	-	-	-	-	-	-	-	-	-	-	-	-	-	59.97	60.76	63.82	59.98	51.76
P. marinus MIT9215	-	-	-	-	-	-	-	-	-	-	-	-	-	-	56.56	57.82	91.17	50.77
Synechococcus RCC307	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	65.90	56.71	53.73
Synechococcus WH7803	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	57.97	52.68
P. marinus MIT9301	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	50.60
Synechococcus PCC7002	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Supplementary Table 3. Inferred positions of root (see also Figure 3).

Bipartition at which the root is inferred*	Number of supporting gene families
RCC307 | AS9601 CCMP1375 CC9605 CC9902 WH8102 MIT9211 MIT9215 WH7803 MIT9301 CCMP1986 MIT9313 MIT9303 MIT9312 MIT9515 NATL1A NATL2A CC9311	433
CC9605 CC9902 WH8102 RCC307 WH7803 MIT9313 MIT9303 CC9311 | AS9601 CCMP1375 MIT9211 MIT9215 MIT9301 CCMP1986 MIT9312 MIT9515 NATL1A NATL2A	177
AS9601 MIT9215 MIT9301 CCMP1986 MIT9312 MIT9515 | CCMP1375 CC9605 CC9902 WH8102 MIT9211 RCC307 WH7803 MIT9313 MIT9303 NATL1A NATL2A CC9311	107
CC9605 CC9902 WH8102 RCC307 WH7803 CC9311 | AS9601 CCMP1375 MIT9211 MIT9215 MIT9301 CCMP1986 MIT9313 MIT9303 MIT9312 MIT9515 NATL1A NATL2A	37
AS9601 MIT9215 MIT9301 CCMP1986 MIT9312 MIT9515 NATL1A NATL2A | CCMP1375 CC9605 CC9902 WH8102 MIT9211 RCC307 WH7803 MIT9313 MIT9303 CC9311	15
AS9601 CCMP1375 MIT9211 MIT9215 MIT9301 CCMP1986 MIT9312 MIT9515 | CC9605 CC9902 WH8102 RCC307 WH7803 MIT9313 MIT9303 NATL1A NATL2A CC9311	7
CC9605 CC9902 WH8102 WH7803 MIT9313 MIT9303 CC9311 | AS9601 CCMP1375 MIT9211 MIT9215 RCC307 MIT9301 CCMP1986 MIT9312 MIT9515 NATL1A NATL2A	6
CCMP1375 CC9605 CC9902 WH8102 MIT9211 WH7803 MIT9313 MIT9303 NATL1A NATL2A CC9311 | AS9601 MIT9215 RCC307 MIT9301 CCMP1986 MIT9312 MIT9515	5
CC9311 | AS9601 CCMP1375 CC9605 CC9902 WH8102 MIT9211 MIT9215 RCC307 WH7803 MIT9301 CCMP1986 MIT9313 MIT9303 MIT9312 MIT9515 NATL1A NATL2A	4
MIT9313 MIT9303 | AS9601 CCMP1375 CC9605 CC9902 WH8102 MIT9211 MIT9215 RCC307 WH7803 MIT9301 CCMP1986 MIT9312 MIT9515 NATL1A NATL2A CC9311	4
WH7803 | AS9601 CCMP1375 CC9605 CC9902 WH8102 MIT9211 MIT9215 RCC307 MIT9301 CCMP1986 MIT9313 MIT9303 MIT9312 MIT9515 NATL1A NATL2A CC9311	4
CC9605 CC9902 WH8102 WH7803 CC9311 | AS9601 CCMP1375 MIT9211 MIT9215 RCC307 MIT9301 CCMP1986 MIT9313 MIT9303 MIT9312 MIT9515 NATL1A NATL2A	4
CC9605 CC9902 WH8102 RCC307 | AS9601 CCMP1375 MIT9211 MIT9215 WH7803 MIT9301 CCMP1986 MIT9313 MIT9303 MIT9312 MIT9515 NATL1A NATL2A CC9311	3
CC9605 CC9902 WH8102 MIT9211 RCC307 WH7803 MIT9313 MIT9303 CC9311 | AS9601 CCMP1375 MIT9215 MIT9301 CCMP1986 MIT9312 MIT9515 NATL1A NATL2A	3
CC9605 CC9902 WH8102 | AS9601 CCMP1375 MIT9211 MIT9215 RCC307 WH7803 MIT9301 CCMP1986 MIT9313 MIT9303 MIT9312 MIT9515 NATL1A NATL2A CC9311	3
MIT9211 | AS9601 CCMP1375 CC9605 CC9902 WH8102 MIT9215 RCC307 WH7803 MIT9301 CCMP1986 MIT9313 MIT9303 MIT9312 MIT9515 NATL1A NATL2A CC9311	3
AS9601 CCMP1375 MIT9211 MIT9215 WH7803 MIT9301 CCMP1986 MIT9312 MIT9515 NATL1A NATL2A | CC9605 CC9902 WH8102 RCC307 MIT9313 MIT9303 CC9311	2
RCC307 MIT9313 MIT9303 | AS9601 CCMP1375 CC9605 CC9902 WH8102 MIT9211 MIT9215 WH7803 MIT9301 CCMP1986 MIT9312 MIT9515 NATL1A NATL2A CC9311	2
MIT9515 | AS9601 CCMP1375 CC9605 CC9902 WH8102 MIT9211 MIT9215 RCC307 WH7803 MIT9301 CCMP1986 MIT9313 MIT9303 MIT9312 NATL1A NATL2A CC9311	1
CC9605 RCC307 CC9311 | AS9601 CCMP1375 CC9902 WH8102 MIT9211 MIT9215 WH7803 MIT9301 CCMP1986 MIT9313 MIT9303 MIT9312 MIT9515 NATL1A NATL2A	1
CC9605 WH8102 RCC307 WH7803 MIT9313 MIT9303 CC9311 | AS9601 CCMP1375 CC9902 MIT9211 MIT9215 MIT9301 CCMP1986 MIT9312 MIT9515 NATL1A NATL2A	1
CC9605 WH8102 RCC307 | AS9601 CCMP1375 CC9902 MIT9211 MIT9215 WH7803 MIT9301 CCMP1986 MIT9313 MIT9303 MIT9312 MIT9515 NATL1A NATL2A CC9311	1
CCMP1375 CC9605 CC9902 WH8102 RCC307 WH7803 MIT9313 MIT9303 CC9311 | AS9601 MIT9211 MIT9215 MIT9301 CCMP1986 MIT9312 MIT9515 NATL1A NATL2A	1
WH7803 CC9311 | AS9601 CCMP1375 CC9605 CC9902 WH8102 MIT9211 MIT9215 RCC307 MIT9301 CCMP1986 MIT9313 MIT9303 MIT9312 MIT9515 NATL1A NATL2A	1
RCC307 WH7803 CC9311 | AS9601 CCMP1375 CC9605 CC9902 WH8102 MIT9211 MIT9215 MIT9301 CCMP1986 MIT9313 MIT9303 MIT9312 MIT9515 NATL1A NATL2A	1
CC9605 WH8102 RCC307 WH7803 | AS9601 CCMP1375 CC9902 MIT9211 MIT9215 MIT9301 CCMP1986 MIT9313 MIT9303 MIT9312 MIT9515 NATL1A NATL2A CC9311	1
CC9605 CC9902 WH8102 MIT9211 RCC307 WH7803 MIT9313 MIT9303 NATL1A NATL2A CC9311 | AS9601 CCMP1375 MIT9215 MIT9301 CCMP1986 MIT9312 MIT9515	1
CC9605 | AS9601 CCMP1375 CC9902 WH8102 MIT9211 MIT9215 RCC307 WH7803 MIT9301 CCMP1986 MIT9313 MIT9303 MIT9312 MIT9515 NATL1A NATL2A CC9311	1
CC9902 | AS9601 CCMP1375 CC9605 WH8102 MIT9211 MIT9215 RCC307 WH7803 MIT9301 CCMP1986 MIT9313 MIT9303 MIT9312 MIT9515 NATL1A NATL2A CC9311	1

*Genomes are designated by their strain name. The genomes in different parts of bipartitions are separated by “|” symbol

Supplementary Table 4. Gene families that do not support high-light adapted Prochlorococcus as a group. Phylogenies of gene families marked with asterisks are shown in Figure 7 and Supplementary Figures 12 and 13.

Dataset ID	**COG Functional Category**	**Functional Annotation**
68.3.12.0	Not in COGs	hypothetical protein
*143.2.19.0	R	hydrolase of the metallo-beta-lactamase superfamily
188.2.18.0	S	hypothetical protein
*62.2.19.0	J	16S rRNA pseudouridylate synthase
1012.1.17.0	R	Rhomboid family protein
*43.4.19.0	C	Aromatic-ring hydroxylase (flavoprotein monooxygenase)
2.27.8.0	Q	SAM-dependent methyltransferase
in1220.1.8.0	D	hypothetical protein
7.6.17.0	M	NAD dependent epimerase/dehydratase
7.9.12.0	M	Nucleotide-diphosphate-sugar epimerase, membrane-associated protein
34.6.10.0	R	Glycerol-3-phosphate oxidase
27.6.6.0	E	imidazoleglycerol-phosphate synthase, cyclase subunit
11.11.8.0	R	hypothetical protein
17.8.4.0	P	hypothetical protein
9.12.6.0	I	short chain dehydrogenase
11.12.5.0	M	Nucleoside-diphosphate-sugar pyrophosphorylase

Supplementary Table 5. Gene families without in-paralogs that support low-light *Prochlorococcus* as a group (i.e., show at least one quartet at >=80% bootstrap support that conflicts plurality signal). Datasets marked with asterisks are gene families used as examples in Figure 8 and Supplementary Figure 14.

Dataset ID	**Represenative GI number**	**COG Functional Category**	**Functional annotation**
36.5.17.0	33860964	E	O-acetylserine (thiol)-lyase A – cysteine biosynthesis
1353.1.8.0	123968545	Not in COGs	hypothetical protein, probable transmembrane helix$^{\$}$
705.1.19.0	33861828	X	hypothetical protein, probable membrane protein$^{\$}$
52.4.11.0	33861652	H	Member of radical SAM superfamily, contains 4Fe-4S cluster, putative Lipoyl synthase
*60.1.19.0	33861455	C	Ferredoxin petF, 2Fe-2S iron-sulfur electron carrier in photosynthesis
*746.1.19.0	33862008	C	ATP synthase, delta (OSCP) subunit

$^{\$}$predicted with SPLIT 4.0 SERVER, http://split.pmfst.hr/split/4/

Supplementary Table 6. Gene families from Transcription and Translation functional categories (COG categories J and K) that do not support marine *Synechococcus* spp. as a group. In many cases (shown in bold) the conflict is due to two largest *Prochlorococcus* spp. genomes (strains MIT9303 and MIT9313) grouping within marine *Synechococcus* spp. For the datasets marked with an asterisk the support for the latter scenario is poor, but there are highly supported conflicts involving other *Prochlorococcus* genomes. Phylogenies of gene families highlighted in yellow are shown in Supplementary Figures 15 and 16.

Dataset ID	COG Functional Category	Functional Annotation
72.1.19.0*	J	Methionyl-tRNA formyltransferase
165.2.19.0	J	Tyrosyl-tRNA synthetase
820.1.19.0	J	possible 16S rRNA processing protein RimM
24.5.19.0	J	Leucyl-tRNA synthetase
764.1.19.0	J	S1 RNA binding domain:Ribonuclease E and G
167.2.19.0	J	histidyl-tRNA synthetase
20.3.18.0	K	Type II alternative RNA polymerase sigma factor, sigma-70 family
38.3.19.0	J	probable tRNA/rRNA methyltransferase
420.1.19.0	J	Alanyl-tRNA synthetase
32.3.19.0	K	probable ribonuclease II
713.1.19.0	J	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
38.5.19.0	J	tRNA/rRNA methyltransferase (SpoU)
478.1.19.0	J	Glutamyl-tRNA(Gln) amidotransferase subunit C
24.4.19.0	J	Valyl-tRNA synthetase
462.1.19.0	J	Arginyl-tRNA synthetase
147.2.19.0	J	Lysyl-tRNA synthetase
940.1.18.0	J	30S Ribosomal protein S16
525.1.19.0	J	Queuosine biosynthesis protein
833.1.19.0	K	cyanobacteria-specific GntR-like HTH domain containing transcriptional regulator
61.2.19.0*	J	2-methylthioadenine synthetase
20.4.19.0	K	RNA polymerase sigma factor
85.1.19.0*	K	putative acetyltransferase, GNAT family
24.1.19.0	J	Methionyl-tRNA synthetase
569.1.19.0	J	translation initiation inhibitor, yjgF family
735.1.19.0	J	Putative tRNA pseudouridine 55 synthase
779.1.19.0*	J	tRNA pseudouridine synthase A
111.2.19.0*	J	probable GTP-binding protein
81.2.19.0	J	Seryl-tRNA synthetase
2.2.19.0	J	methyltransferase for Ribosomal protein L11
816.1.19.0	J	putative ribonuclease D
187.1.19.0	K	SOS function regulatory protein, LexA repressor
778.1.19.0	J	50S ribosomal protein L13
437.1.19.0	J	Ribosome-binding factor A
787.1.19.0	J	30S ribosomal protein S5
697.1.19.0*	J	ribonuclease P protein component
704.1.19.0	J	Phenylalanyl-tRNA synthetase alpha chain
507.1.19.0*	K	Predicted transcriptional regulator, consists of a Zn-ribbon and ATP-cone domains
625.1.19.0	J	putative rRNA (adenine-N6,N6)-dimethyltransferase

141.2.19.0	J	putative methionine aminopeptidase
191.1.19.0	J	tRNA-dihydrouridine synthase
794.1.19.0*	J	30S Ribosomal protein S17
61.4.17.0	J	Fe-S oxidoreductase
691.1.19.0*	J	Glycyl-tRNA synthetase alpha subunit
221.1.19.0	J	Glutamyl/glutaminyl-tRNA synthetase
467.1.19.0*	J	50S ribosomal protein L10
931.1.18.0	K	GCN5-related N-acetyltransferase
32.2.19.0	J	30S ribosomal protein S1
796.1.19.0*	J	50S ribosomal protein L16
191.2.17.0	J	putative nitrogen regulation protein NifR3 family-like protein
803.1.19.0	J	50S ribosomal protein L3
842.1.19.0	J	50S ribosomal protein L20
800.1.19.0	J	50S ribosomal protein L2
783.1.19.0*	J	30S ribosomal protein S13
854.1.19.0	J	30S ribosomal protein S6
974.1.17.0*	K	GCN5-related N-acetyltransferase
928.1.18.0	J	50S ribosomal protein L32
150.2.19.0*	J	translation initiation factor IF-1
832.1.19.0	J	Translation Initiation factor 3
733.1.19.0	J	50S ribosomal protein L21
798.1.19.0*	J	50S ribosomal protein L22
138.2.19.0	J	protein methyltransferase
547.1.19.0*	J	50S ribosomal protein L19
466.1.19.0	J	50S ribosomal protein L7/L12
768.1.19.0*	J	30S ribosomal protein S10

Supplementary Table 7. 35 gene families with at least one homolog in nine sequenced cyanophages. Phylogenies of gene families highlighted in bold are shown in Supplementary Figures 17 and 18.

Dataset ID	COG Functional Category	Functional annotation
15.2.19.0	L	primosomal protein N'
253.1.19.0	X	hypothetical protein
329.1.19.0	X	hypothetical protein
44.2.19.0	H	Carbon-nitrogen hydrolase:NAD+ synthase
491.1.19.0	F	AICARFT/IMPCHase bienzyme:Methylglyoxal synthase-like domain-containing protein
53.3.19.0	M	Putative NADH-flavin reductase
6.11.18.0	L	putative ATPase, AAA family
6.2.19.0	O	FtsH ATP-dependent protease-like protein
6.5.19.0	L	DNA polymerase, gamma and tau subunits
6.8.19.0	L	Holliday junction DNA helicase RuvB
60.1.19.0	C	ferredoxin, petF-like protein
60.3.14.0	C	ferredoxin
610.1.19.0	F	Phosphoribosylformylglycinamidine synthetase PurS
624.1.19.0	L	DNA primase
682.1.19.0	L	DNA polymerase I
690.1.19.0	C	Ferredoxin
7.11.19.0	M	NAD dependent epimerase/dehydratase
7.2.19.0	M	Putative nucleotide sugar epimerase
7.6.17.0	M	NAD dependent epimerase/dehydratase
7.7.13.0	M	dTDP-D-glucose 4,6-dehydratase
7.9.12.0	M	Nucleotide-diphosphate-sugar epimerase, membrane-associated protein
72.1.19.0	J	putative Methionyl-tRNA formyltransferase
728.1.19.0	X	hypothetical protein
73.2.19.0	L	RecA bacterial DNA recombination protein
809.1.19.0	E	S-adenosylmethionine decarboxylase proenzyme
816.1.19.0	J	ribonuclease D
858.1.19.0	F	thymidylate synthase
90.1.19.0	X	phycocyanobilin:ferredoxin oxidoreductase
90.2.18.0	X	15,16 dihydrobiliverdin:ferredoxin oxidoreductase
90.3.18.0	X	phycoerythrobilin:ferredoxin oxidoreductase
913.1.18.0	L	Uracil-DNA glycosylase
926.1.18.0	R	Predicted pyrophosphatase
in1060.1.8.0	G	possible endolysin
in865.1.13.0	X	CpeT homolog
in868.1.15.0	X	Predicted dioxygenase