

Table 1s. Strain descriptions. Abbreviations are as follows: PUB, phycourobilin chromophore; *gvp*, genes for gas vesicle formation; *nif*, genes for nitrogen fixation.

Organisms		Complete annotation July 2005	Genome size (Mb)	Estimated gene number	Strain properties	
Marine strains	Unicellular	<i>Crocospaera watsonii</i> WH 8501 (<i>Cyanothece. Synechococcus</i>)	No	6.29	6011	Nitrogen fixer Phycocyanin with PUB
		<i>Prochlorococcus marinus</i> MIT9313	Yes	2.41	2326	Isolated from Gulf Stream deep ocean water. Low-light adapted strain. High divinyl-Chlb/a <i>Prochlorococcus</i> clade IV. No phycobilisome. Atypic phycoerythrin
		<i>Prochlorococcus marinus</i> MIT9312	No	1.71	1811	High-light adapted strain from Gulf Stream. Low divinyl-Chlb/a <i>Prochlorococcus</i> clade II. No phycobilisome.
		<i>Prochlorococcus marinus</i> SS120 (CCMP1375)	Yes	1.75	1926	Low-light adapted strain from Sargasso Sea. High divinyl-Chlb/a <i>Prochlorococcus</i> clade II. No phycobilisome. Atypic phycoerythrin
		<i>Prochlorococcus marinus</i> MED4	Yes	1.66	1757	Isolated from surface water of an oligotrophic mediterranean area, High-light adapted strain. Low-divinyl-Chlb/a <i>Prochlorococcus</i> clade I. No phycobilisome.
		<i>Synechococcus</i> sp. CC9902	No	2.23	2373	Coastal (California current)
		<i>Synechococcus</i> sp. CC9605	No	2.51	2747	Isolated from an oligotrophic area (California current)
		<i>Synechococcus</i> sp. WH8102	Yes	2.43	2578	Obligate marine. Marine A <i>Synechococcus</i> clade III from Western Caribbean. Phycoerythrin with PUB, chromophorylated linker polypeptides. Elevated growth requirements for Na ⁺ , but also for Cl ⁻ , Mg ²⁺ , and Ca ²⁺ . Swimming motility
	Filamentous	<i>Trichodesmium erythraeum</i> IMS101	No	8.19	4801	Nitrogen fixer. Upper levels of tropical oceans. Synchronous cell divisions within a trichome. Phycoerythrin with PUB. Gas vesicle (<i>gvp</i>) genes. Gliding motility
Others	Unicellular	<i>Gloeobacter violaceus</i> PCC 7421	Yes	4.66	4478	No thylakoid membrane. Rod-shaped phycobilisomes directly attached to the cytoplasmic membrane. Phycoerythrin with PUB. Isolated from calcareous rock
		<i>Synechocystis</i> sp. PCC 6803	Yes	3.95	3614	Euryhaline. Facultative heterotroph. Twitching (gliding) motility
		<i>Thermosynechococcus elongatus</i> BP-1	Yes	2.59	2521	Thermophile
		<i>Synechococcus elongatus</i> PCC 7942	No	2.69	2663	Strict photoautotroph. Phycobilisome core with only two cylinders of allophycocyanin
	Filamentous	<i>Anabaena</i> sp. PCC 7120	Yes	7.21	6217	Heterocyst differentiation (hormogonium?). Nitrogen fixer. Phycoerythrocyanin. Gas vesicle (<i>gvp</i>) genes
		<i>Anabaena variabilis</i> ATCC 29413 (PCC7937, <i>Anabaena flos-aquae</i> CCAP 1403/13A. UTEX1444)	No	7.11	5809	Heterocyst, hormogonium and akinete differentiation. Nitrogen fixer (3 <i>nif</i> gene sets). Facultative heterotroph. Phycoerythrocyanin. Photosynthetic H ₂ production. gas vesicle (<i>gvp</i>) genes
		<i>Nostoc punctiforme</i> PCC 73102 (ATCC 29133)	Yes	9.06	7465	Heterocyst, hormogonium and akinete differentiation. Nitrogen fixer. Endosymbiont. Gas vesicle (<i>gvp</i>) genes. Facultative photo- and chemoheterotroph. Phycoerythrin. Complementary Chromatic Adaptation type II

Table 2s. Structural domains that may be found in two component regulatory systems cyanobacterial proteins.

Domain acronym	IterProScan identifier	Known properties and/or function
AraC	IPR000005	Related to the arabinose operon regulatory protein AraC, usually positive transcriptional factors, spans roughly 100 residues and comprises two HTH subdomains
Cache	IPR004010	Signaling domain found in a certain class of prokaryotic chemotaxis receptors
CBS	IPR000644	CBS (cystathionine-beta-synthase) are small intracellular modules, mostly found in two or four copies which can act as binding domains for adenosine derivatives and sensors of cellular energy status, may function to form pores in membranes
CHASE	IPR006189	Extracellular domain of 200-230 amino acids found in transmembrane receptors, named from its presence in diverse receptor-like proteins with histidine kinase and nucleotide cyclase domains (Cyclases/Histidine kinases Associated Sensory Extracellular), always N-terminally in extracellular or periplasmic locations, predicted to bind diverse low molecular weight ligands
CHASE2	IPR007890	CHASE2 domains are specifically found in histidine kinases, adenylate cyclases, serine/threonine kinases and predicted diguanylate cyclases/phosphodiesterases.
CHASE3	IPR007891	ibid
CheB	IPR000673	N-terminal domain found in CheB methyl-esterase which removes a methyl group in the methyl-accepting chemotaxis proteins (MCP). CheB is regulated through phosphorylation by CheA.
CheR	IPR000780	CheR are methylating enzymes (MCP methyltransferase)
CheW	IPR002545	CheW-like protein, CheW interacts with the methyl accepting chemotaxis proteins (MCPs) and relays signals to the response regulator CheY which affects flagellar rotation
cNMP	IPR000595	Domain of about 120 residues shared by proteins which bind cyclic nucleotides (cAMP or cGMP)
EAL	IPR001633	Called EAL after its conserved residues, good candidate for a diguanylate phosphodiesterase function, often along with PAS. Also known as DUF2.
FHA	IPR000253	Forkhead-associated (FHA) is a phosphopeptide recognition domain, spans approximately 80-100 amino acid residues, present in a diverse range of proteins, from kinases and phosphatases, to transcription factors and metabolic enzymes involved in many different cellular processes
GAF	IPR003018	Present in phytochromes (Phyt_2) and cGMP-specific phosphodiesterases, also found in NifA, a transcriptional activator required for activation of most <i>nif</i> operons

GGDEF	IPR000160	Ubiquitous in bacteria and often linked to a regulatory domain, such as a phosphorylation receiver or oxygen sensing domain, and in proteins which synthesizes cyclic di-GMP. Also known as DUF1.
GuC	IPR001054	Catalytic domain of class III adenylyl /guanylyl cyclase (AC/GC)
HD	IPR008328	Domain found in a superfamily of enzymes with a predicted or known phosphohydrolase activity (Metal-dependent phosphohydrolase), such as 3',5'-cGMP phosphodiesterase
HATP	IPR003594	HATPase_c domains are found in several ATP-binding proteins such as: histidine kinase, DNA gyrase B, topoisomerases, heat shock protein HSP90, phytochrome-like ATPases and DNA mismatch repair proteins
HisKA	IPR003661	Histidine kinase A (phosphoacceptor) N-terminal domain is a dimerisation and phosphoacceptor domain of histidine kinases
HisKA_2 (HWE)	IPR011495	Dimerisation and phosphoacceptor domain of a subfamily of histidine kinases
HisKA_3	IPR011712	Dimerisation and phosphoacceptor domain of a sub-family of histidine kinases
Hkd	IPR004105	homodimer interface of the signal transducing histidine kinase family
HMP	IPR003660	HAMP region found in front of the HisKA domains in proteins which are usually integral membrane proteins
Hpt	IPR008207	Present, usually at the N terminus, in proteins which undergo autophosphorylation, can mediate phosphotransfer reactions in multistep phosphorelay two-component systems
HTH	IPR010985	DNA-binding domains found in several bacterial repressors which are usually obligate dimers by interactions of their N-Terminal domains
HNOBA	IPR011645	The HNOBA (HNOBAssociated) domain is found associated with the HNOB (<u>H</u> eme- <u>N</u> itric <u>O</u> xide <u>B</u> inding) domain and IPR001054 in soluble cyclases and signalling proteins. The HNOB domain is predicted to function as a heme-dependent sensor for gaseous ligands, and transduce diverse downstream signals in both bacteria and animals. Function guanylate cyclase activity
KaiB	IPR011649	The cyanobacterial clock proteins KaiA and KaiB are proposed as regulators of the circadian rhythm in cyanobacteria. Mutations in both proteins have been reported to alter or abolish circadian rhythmicity. KaiB adopts an alpha-beta meander motif and is found to be a dimer
LuxR	IPR000792	DNA-binding HTH domain of about 65 amino acids, present in transcription regulators of the LuxR/FixJ/NarL family of response regulators
MASE1	IPR007895	Domain of unknown function found in proteins of unknown function
MHYT	IPR005330	Named after its conserved amino acid motif, methionine, histidine, and tyrosine (~190-residue), consists of six predicted transmembrane (TM) segments, connected by short arginine-rich

		cytoplasmic and periplasmic loops rich in charged residues
PAC	IPR000700	40-45 amino acids usually situated C-terminal to PAS (PAC = PAS-associated, C-terminal), contributes to PAS domain fold
PAS	IPR000014	Sensor domain (~60 aa-long) of many signalling proteins, several detect their signal by way of an associated cofactor, such as heme or flavin, often associated with PAC domains
PHYT	IPR001294	Found in plant phytochromes, senses the red and far-red light conditions through a covalently bound linear tetrapyrrole which is responsible for the reversible photochemical conversion
Pkinase	IPR000719	Conserved catalytic core common with both serine/threonine and tyrosine protein kinases
PP2C_SIG	IPR001932	Protein phosphatase 2C-like domain found in protein phosphatase 2C, pyruvate dehydrogenase (lipoamide)-phosphatase and adenylate cyclase. (SpoIIE sometimes used as it is a Protein phosphatase 2C-like)
RR	IPR001789	CheY-, PatA-, and OmpR-like domains
Treg	IPR001867	C-terminal part of transcriptional regulatory protein
Trk	IPR006037	TrkA-C is often found next to the TrkA-N domain, exact function of this domain is unknown, may bind an unidentified ligand
UPF	IPR005360	Domain with a conserved GRY motif, found in a family of proteins about 80 amino acids-long of unknown function. This family appears to be related to ribbon-helix-helix DNA-binding proteins.

Table 3s. Two-component proteins from the 16 complete cyanobacterial genomes available as of May 205, for which no cyanobacterial orthologue could be retrieved. The *Anabaena* 7120 and *Escherichia coli* K12 sequences used in the searches to retrieve the putative proteins containing histidine kinase (HK) and response regulator (RR) domains were as follows: RR-CheY, 7120 alr0442 and *E. coli* NP_416396.1; RR-OmpR, 7120 all3822 (aa 7-124) and *E. coli* NP_417864.1 (aa 5-124); T_reg, 7120 all3822 (aa 151-233) and *E. coli* NP_417864.1 (aa 156-232); HisKA-HATPase, 7120 all0330 (aa 429-656) and *E. coli* NP_417863.1 (aa 234-439); and Hpt, 7120 all1068 (aa 8-111) and *E. coli* NP_415513.1 (aa 815-896). Analyses performed at the IMG site (Integrated Microbial Genomes, <http://img.jgi.doe.gov/cgi-bin/pub/main.cgi>), and at European Bioinformatics Institute (<http://www.ebi.ac.uk/InterProScan/>) and <http://www.cbs.dtu.dk/services/TMHMM/> for the domain structure assignments. Whenever a single gene or a member of a group has been assigned a function in the literature and/or an annotation in databases, the gene name is mentioned. (⊠) indicates proteins that do not have any bacterial orthologues.

A. Two component system proteins (singletons)

Gene Name	Classification	Functional or Temporary gene name
NpunNpF0536	HKA	<i>chk113</i>
NpunNpR3010	HKA	<i>chk114</i>
Cwat_400841320	HKA	<i>chk115</i>
TBP1tr2426	HATPase	<i>chk116</i>
NpunNpF3113, 100% id with NpunNpF2204	HATPase	<i>chk117</i>
NpF2204, 100% id with NpunNpF3113	HATPase	<i>chk118</i>
NpunNpF2205, 100% id with NpunNpF3114	HATPase	<i>chk119</i>
NpunNpF3114, 100% id with NpunNpF2205	HATPase	<i>chk120</i>
Cwat_400841350	HATPase	<i>chk121</i>
Cwat_400874080	HATPase	<i>chk122</i>
NpunpNPBF140	HKI	<i>chk123</i>
7421glr3138	HKI	<i>chk124</i>
7421gll0380	HKI	<i>chk125</i>
6803sll0790	HKI	<i>chk31, hik31</i>
7120all7605	HKI	<i>chk126</i>
7421glr3138	HKI	<i>chk127</i>
NpunpNPBF140	HKI	<i>chk128</i>
SS12Pro1543	HKI (N-ter (TM) ₆)	<i>chk129</i>
Avar_400185180	HKI	<i>chk130</i>
NpunNpF1775	HKI (HisKA_3)	<i>chk131</i>
7421gll1816	HKII + GAF	<i>chk132</i>
NpunNpR2209	HKII + GAF	<i>chk133</i>
7421gll2627	HYII + GAF	<i>chk134</i>
Tery_403225400	HKII +GAF	<i>chk135</i>
Tery_403256460	HKII + GAF	<i>chk136</i>

NpunNpR1759	HKII + 2 GAF	chk137
NpunNpR5313	HKII + 2 GAF	chk138
NpunpNPAF142	HKII + 3 GAF	chk139
NpunNpR5113	HKII + 4 GAF	chk140
Cwat_400862060	HKII + Phytochrome	chk141
Cwat_400870170	HKII + PAC-GAF	chk142
7421glr1239	HKII + PASPAC-GAF	chk143
6803sll1124	HKII + GAF-mPASPAC	chk3, hik3
7120alr2123	HKII + PASPAC	chk144 (α)
NpunNpR3198	HKII + PASPAC	chk145 (α)
NpunNpF4953	HKII + PASPAC	chk146
NpunNpR3581	HKII + PASPAC	chk147
NpunpNPAR133	HKII + PASPAC	chk148
Avar_400177700	HKII + (PASPAC) ₂	chk149
NpunNpF5383	HKII + PASPAC	chk150
6803sll1003	HKII + (PASPAC) ₃	chk13, hik13
NpunNpF2121	HKII + (PASPAC) ₃	chk151
Tery_403260760	HKII + PAS	chk153
NpunNpR6525	HKII + PAS	chk154
Avar_400193580	HKIII + Cache	chk155
7421glr3675	HKIII + HAMP	chk156
NpunNpR6338	HKIII + HAMP	chk157
NpunpNPAR140	HKIII + HAMP	chk158 (α)
7421gll0814	HKIII + HAMP-PAC	chk159
NpunNpR0401	HKIII + HAMP-Cache	chk160
NpunNpF3779	HKIII + Cache	chk177
NpunNpF6040	HKIII + HAMP-GAF-Cache	chk161
7120alr0709	HKIV	chk162
7120all1625	HKIV	chk163
7120all2282	HKIV	chk164
Avar_400222710	HKIV	chk165
NpunNpR2272	HKIV	chk166
NpunNpF0342	HKIV	chk167
NpunNpR6545	HKIV	chk168
Tery_403260730	HKV + CheW	chk169
NpunNpR3054	HKV + CHASE	chk170
NpunNpF3675	HKV + CHASE3	chk171
NpunNpR0610	HKV + CHASE3-(PASPAC) ₂	chk172
7421glr2059	HKV + Trk	chk173
Tery_403252270	HKV + cNMP	chk175
7421gll1661	HKV + cNMP	chk109
Tery_403220860	HKV + CHASE	chk176
TBP1dl12438	RRI CheY	crr103
Cwat_400845910	RRI CheY	crr104
7120all4927	RRI CheY	crr105 (α)
7421glr1534	RRI CheY	crr106
7421glr2272	RRI CheY	crr107
7421glr2257	RRI CheY	crr108
7421glr1532	RRI CheY	crr109 (α)
Tery_403244930	RRI CheY	crr110 (α)
Tery_403262790	RRI CheY	crr111
Tery_403235360	RRI Other (>200aa-long)	crr112
Tery_403235350	RRI CheY	crr113
Avar_400193570	RRI CheY	crr114 (α)
Avar_400193560	RRI CheY	crr115 (α)
NpunNpR1758	RRI CheY	crr116 (α)
NpunNpF3566	RRI CheY	crr117
NpunNpF5044	RRI CheY	crr118
NpunNpF1085	RRI CheY	crr119
NpunpNPBR170	RRI CheY	crr120 (α)
NpunpNPBR131	RRI CheY	crr121 (α)
NpunNpF3676	RRI CheY	crr122
7942_403102070	RRI CheY	crr123
Cwat_400880570	RRI CheY	crr124 (α)

Cwat_400836210	RRI CheY	<i>crr125</i>
NpunNpR1757	RRI CheY	<i>crr175</i>
6803slr1214	RRI PatA	<i>crr15, rre15</i> (α)
6803slr1594	RRI PatA	<i>crr5, rre5</i> (α)
6803slr1693	RRI PatA	<i>crr24, rre24</i> (α)
7120all3348	RRI other (>200aa-long)	<i>crr126</i>
Cwat_400874090	RRI other (>200aa-long)	<i>crr127</i>
Cwat_400841340	RRI other (>200aa-long)	<i>crr128</i>
7120all2281	RRI other (>200aa-long)	<i>crr174</i>
6803sll0797	RRI OmpR	<i>crr33, rre33, nrsR, rppA</i>
6803sll0789	RRI OmpR	<i>crr34, rre34, copR</i>
7120alr1170	RRI OmpR	<i>crr129</i>
7120all7606	RRI OmpR	<i>crr130</i>
7421glr0281	RRI OmpR	<i>crr131</i>
7421glr3139	RRI OmpR	<i>crr132</i>
NpunNpR3241	RRI OmpR	<i>crr133</i>
NpunNpR1903	RRI OmpR	<i>crr134</i>
NpunNpF3064	RRI OmpR	<i>crr135</i>
NpunNpR6204	RRI OmpR	<i>crr136</i>
NpunNpF6321	RRI OmpR	<i>crr137</i>
NpunpNPBF139	RRI OmpR	<i>crr138</i>
Avar_400216090	RRII NarL	<i>crr139</i>
7421gll0379	RRII NarL	<i>crr140</i>
NpunNpF1800	RRII NarL	<i>crr141</i>
NpunNpF1776	RRII NarL	<i>crr142</i>
7120alr9013	RRII NarL	<i>crr143</i>
7120alr8531	RRII NarL	<i>crr144</i>
7421glr0273	RRII NarL	<i>crr145</i>
7421glr0773	RRII NarL	<i>crr146</i>
7421gll2210	RRII NarL	<i>crr147</i>
7421glr0244	RRII NarL	<i>crr148</i>
7421gll0062	RRII NarL	<i>crr149</i>
SS12Pro1542	RRII NarL	<i>crr150</i>
NpunNpR0956	RRII NarL	<i>crr151</i>
6803slr1213	RRII AraC	<i>crr14, rre14</i>
7120alr0429	RRII AraC	<i>crr152</i>
NpunNpF1278	RRII AraC	<i>crr153</i>
NpunNpR3197	RRII AraC	<i>crr154</i>
NpunpNPAF075	RRII AraC	<i>crr155</i>
NpunNpR2091	RRII AraC	<i>crr156</i>
6803slr0687	RRIV + GGDEF	<i>crr4, rre4</i>
Tery_403260660	RRIV +GGDEF	<i>crr157</i>
7421gll4160	RRIV + GGDEF	<i>crr158</i>
Cwat_400861080	RRIV + GGDEF	<i>crr159</i>
7942_403103420	RRIV + PASPAC-GGDEF	<i>crr160</i>
6803slr1305	RRIV + PASPAC-EAL-GGDEF	<i>crr41, rre41</i>
NpunNpR3564	RRIV + PAC-EAL-GGDEF	<i>crr161</i>
NpunNpF3678	RRIV + PASPAC-EAL-GGDEF	<i>crr162</i>
6803slr2100	RRIV + HD	<i>crr20, rre20</i>
Cwat_400840880	RRIV + PACPAS	<i>crr163</i>
Tery_403262700	RRIV + GuC	<i>crr164</i>
7120alr2280	RRIV + PP2C_SIG	<i>crr165</i>
NpunNpR0244	RRIV + CheB	<i>crr166</i>
Tery_403260700	RRIV + CheB	<i>crr167</i>
Tery_403260740	RRIV + CheC	<i>crr168</i>
NpunNpF5527	RRIV + IF2	<i>crr169</i>
Cwat_400887000	RRIV + CheW	<i>crr170</i>
7421glr1768	RRIV + cNMP	<i>crr171</i>
6803slr2099	HYI	<i>chy40, hik40</i>
7120all0637	HYI	<i>chy95</i>
NpunNpF5480	HYI	<i>chy96</i>
7942_403108060	HYI	<i>chy97</i>
Tery_403232010	HYI	<i>chy98</i>
NpunNpF0957	HYI	<i>chy99</i>

Tery_403234400	HYI	<i>chy100</i>
NpunpNPAR130	HYI	<i>chy101</i>
NpunNpF1601	HYI	<i>chy102</i>
NpunNpR2271	HYI	<i>chy103</i>
NpunNpF5043	HYI + PAS	<i>chy104</i>
7421gll1533	HYI + PASPAC	<i>chy105</i>
6803slr0222	HYI + (PAS) ₆ (PAC) ₅	<i>chy25, hik25</i>
NpunNpF1185	HYI + GAF	<i>chy106</i>
NpunNpR2903	HYI + GAF-PASPAC	<i>chy107</i>
NpunNpF3677	HYI (HKA_2)	<i>chy108</i>
NpunNpF1799	HYI (HKA_3)	<i>chy109</i>
NpunNpF5092	HYII (PASPAC) ₃ -GAF	<i>chy110</i>
NpunNpR1760	HYII + GAF	<i>chy111</i>
NpunNpR2268	HYII + (PASPAC) ₂	<i>chy112</i>
Avar_400190700	HYII + PASPAC	<i>chy113</i>
7421gll4411	HYII + (PASPAC) ₂	<i>chy114</i>
NpunpNPAR131	HYII + PASPAC	<i>chy115</i>
7120all4097	HYII + (PASPAC) ₃	<i>chy116</i>
NpunNpF1600	HYII	<i>chy117</i>
NpunNpR2035	HYII + PAS-(PASPAC) ₃	<i>chy118</i>
NpunNpR3591	HYII + PAS-PAC-(PASPAC) ₂	<i>chy119</i>
NpunNpR4211	HYII + PAS-(PASPAC) ₄	<i>chy120</i>
NpunNpF4131	HYII + GAF	<i>chy121</i>
NpunNpR6464	HYII + PAS-GAF	<i>chy122</i>
TBP1tl11282	HYIII + GAF-(PASPAC) ₂ -(GAF) ₂ -HK-(CheY) ₂	<i>chy123</i>
Avar_400176240	HYII + GAF-PASPAC	<i>chy124</i>
7421gll0099	HYII + PASPAC-GAF	<i>chy125</i>
NpunpNPDR038	HYII + GAF ₂ -(PASPAC) ₂ -GAF-PASPAC	<i>chy126</i>
NpunNpF5035	HYII + PAS-PASPAC-PAS-(PASPAC) ₄ -GAF-PASPAC-GAF	<i>chy127</i>
NpunNpF5360	HYII + Cache-HAMP-PAC	<i>chy128</i>
Tery_403234780	HYII + GuC-HAMP	<i>chy129</i>
Tery_403234790	HYII + GuC-HAMP	<i>chy130</i>
Tery_403245090	HYII + GuC-HAMP	<i>chy131</i>
NpunNpR2263	HYIII	<i>chy132</i>
6803slr2098	HYIII + Hpt-MHYT-(PASPAC) ₂	<i>chy21, hik21</i>
7120alr2279	HYIII + PACPAS-GAF-Hpt (N-ter HNOBA)	<i>chy133</i>
6803slr1759	HYIII + PACPAS-GAF-Hpt	<i>chy14, hik14</i>
7421glr1016	HYIII + HAMP-GAF-CHASE3-PACPAS-2RR	<i>chy134</i>
Avar_400173090	HYIV	<i>chy135</i>
NpunNpR2262	HYIV PASPAC	<i>chy136</i>
NpunNpR5611	HYIV PASPAC	<i>chy137</i>
NpunNpR3572	HYIV + GAF-(PAS) ₆ (PAC) ₅	<i>chy138</i>
6803sll1905	HYIV + PAS-Hpt	<i>chy19, hik19</i>
NpunNpF2346	HYV + RHH_2	<i>chy139</i>
6803slr1983	HYVII (HisKA) + PP2C_SIG	<i>chy180, rre22, ppcE</i>
Cwat_400862070	HYVII (HATPase C)	<i>chy140</i>
Cwat_400866330	HYVI + CheW (no Hkd)	<i>chy141 (α)</i>
NpunNpR0245	HYVI + CheW-Hpt	<i>chy142</i>
Cwat_400841330	HYVII (HATPase C)	<i>chy143</i>

Table 4s. Closely linked two-component genes and probable operons that contain two or more two-component genes in the nine complete annotated genomes. This was constructed from the Chromosome map images at Cyanobase (<http://www.kazusa.or.jp/cyano/>) and IMG (<http://img.jgi.doe.gov/cgi-bin/pub/main.cgi>) sites. Genes are identified by their locus tags.

Synechocystis 6803

Assigned gene number	Classification	Functional name	Orthologue number	Additional information
sll0038 <i>rre36</i> sll0039 <i>rre35</i> sll0043 <i>hik18</i>	RRI PatA RRI CheY HYVI+CheW	<i>pixL, taxAY1</i>	<i>crr36</i> <i>crr35</i> <i>chy18</i>	
sll0789 <i>rre34</i> sll0790 <i>hik31</i>	RRII OmpR HKI	<i>copR</i>	<i>crr34</i> <i>chk31</i>	
sll0797 <i>rre33</i> sll0798 <i>hik30</i>	RRII OmpR HKI	<i>rppA, nrsR</i> <i>rppB, nrsS</i>	<i>crr33</i> <i>chk30</i>	
sll1228 <i>hik4</i> sll1229 <i>hik41</i>	HYII+GAF HYI		<i>chy4</i> <i>chy41</i>	
sll1291 <i>rre12</i> sll1292 <i>rre11</i> sll1296 <i>hik39</i>	RRI PatA RRI CheY HYVI+CheW		<i>crr12</i> <i>crr11</i> <i>chy39</i>	sll1294 MCP
sll1590 <i>hik20</i> sll1592 <i>rre19</i>	HKI RRII NarL		<i>chk2</i> <i>crr23</i>	
sll1672 <i>hik12</i> sll1673 <i>rre2</i>	HYII+PAS-MASE1 RRIV+GGDEF		<i>chy12</i> <i>crr2</i>	
sll5059 sll5060	RRIOther (>200aa) HYIII+PACPAS-GAF-Hpt		<i>crr43</i> <i>chy44</i>	Plasmid encoded (pSYSM)
slr0311 <i>hik29</i> slr0312 <i>rre32</i>	HKIII RRII NarL		<i>chk29</i> <i>crr32</i>	
slr0473 <i>hik35</i> slr0474 <i>rre27</i>	HKII+Phytochrome RRI CheY	<i>cph1, aphA</i> <i>rcp1</i>	<i>chk35</i> <i>crr27</i>	
slr1041 <i>rre6</i> slr1042 <i>rre7</i>	RRI PatA RRI CheY	<i>pilH</i>	<i>rre6</i> <i>crr7</i>	
slr1213 <i>rre14</i> slr1214 <i>rre15</i>	RRII AraC RRI PatA		<i>rre14</i> <i>rre15</i>	
slr1759 <i>hik14</i> slr1760 <i>rre8</i>	HYIII+PACPAS-GAF-Hpt RRIV+GGDEF		<i>chy14</i> <i>crr8</i>	
slr1982 <i>rre21</i> slr1983 <i>rre22</i>	RRI CheY HYVI-HisKA C		<i>crr21</i> <i>chy180</i>	
slr2098 <i>hik21</i> slr2099 <i>hik40</i> slr2100 <i>rre20</i> slr2104 <i>hik22</i>	HYIII+Hpt-MHYT HYI RRIV+HD HYII+GAF-PAC-Hpt		<i>chy21</i> <i>chy40</i> <i>crr20</i> <i>chy22</i>	
slr6040 slr6041	RRII OmpR HKI		<i>crr44</i> <i>chk70</i>	Plasmid encoded (pSYSX)

Anabaena 7120

Assigned gene number	Classification	Functional name	Orthologue number	Additional information
----------------------	----------------	-----------------	-------------------	------------------------

all0637 all0638	HYI HYII		<i>chy95</i> <i>chy71</i>	
all0825 all0824 all0823	HKII+GAF HYII+PAS(PAC) ₃ RRI CheY		<i>chk61</i> <i>chy66</i> <i>crr10</i>	
all0930 all0929 all0926	RRI PatA RRI CheY HYVI+CheW	<i>piH</i> <i>cheA</i>	<i>crr6</i> <i>crr7</i> <i>chy43</i>	(all928 CheW, all927 – methyl accepting protein)
all1072 all1071 all1068	RRI PatA RRI CheY HYVI+CheW	<i>pixH, pisH</i> <i>cheA</i>	<i>crr36</i> <i>crr35</i> <i>chy92</i>	
all1178 all1177	HYII+CHASE2 HYI		<i>chy77</i> <i>chy52</i>	
all1281 all1280 all1279	RRI Other (>200aa) HKII+GAF-PAS-HWE HYI+PAS		<i>crr13</i> <i>chk45</i> <i>chy55</i>	
all1389 all1388	HYII+Cache HYI		<i>chy69</i> <i>chy45</i>	
all1704 all1703	RRII NarL RRIV+GAF		<i>crr25</i> <i>crr100</i>	
all2095 all2094	HKVI+CBS-HWE-mPASPAC HYI+PASPAC		<i>chk108</i> <i>chy57</i>	
all2165 all2164 all2161	RRI PatA RRI CheY HYVI+CheW	<i>cheA</i>	<i>crr64</i> <i>crr51</i> <i>chy94</i>	
all2282 all2281	HKV RRI Other (>200aa)		<i>chk164</i> <i>crr174</i>	
all2875 all2874	HYIII+PACPAS-GAF-Hpt RRIV+GGDEF		<i>chy44</i> <i>crr8</i>	
all2899 all2898 all2897	HKII+Phytochrome RRI CheY HYI+PAS(PAC) ₂	<i>aphB</i>	<i>chk73</i> <i>crr46</i> <i>chy56</i>	
all2956 all2955	HKI RRI CheY		<i>chk59</i> <i>crr79</i>	
all3564 all3563	HKII HKIII+PAC		<i>chk16</i> <i>chk74</i>	
all3767 all3766 all3765 all3764	HKII RRI CheY HYI HYI		<i>chk77</i> <i>crr54</i> <i>chy54</i> <i>chy51</i>	
all4097 all4096	HYII+(PAC) ₃ PAS HYI		<i>chy116</i> <i>chy38</i>	
all4503 all4502	RRII OmpR HKI	<i>phoB</i> <i>sphS, phoR</i>	<i>crr29</i> <i>chk7</i>	(all4501 <i>phoU</i>)
all4727 all4726	RRII OmpR HKIV+HAMP		<i>crr3</i> <i>chk10</i>	
all5174 all5173 all5172	RRIII+Hpt-GGDEF HYIII+mPASPAC RRI CheY		<i>crr91</i> <i>chy80</i> <i>crr59</i>	
all5309 all5308	HYII+CHASE2 HYI		<i>chy76</i> <i>chy146</i>	
all7584 all7583	RRII OmpR HKI	<i>copR</i>	<i>crr44</i> <i>chk70</i>	Plasmid encoded (Beta)
all7606 all7605	RRII OmpR HKI		<i>crr130</i> <i>chk126</i>	Plasmid encoded (Beta)

alr0354 alr0356	HKIV S/Tkin-GAF RRI Other (>200aa)		<i>chk97</i> <i>crr70</i>	
alr0428 alr0429	HKII+(PAS) ₃ RRII AraC		<i>chk79</i> <i>crr152</i>	
alr0709 alr0710	HKIV HKV+ S/Tkin-GAF		<i>chk162</i> <i>chk107</i>	
alr1170 alr1171	RRII OmpR HKI		<i>crr129</i> <i>chk55</i>	
alr1229 alr1230 alr1231	HKII+ GAF-mPASPAC RRIV+EAL HYI		<i>chk32</i> <i>crr39</i> <i>chy42</i>	
alr1966 alr1967 alr1968	HKII+GAF-PASPAC RRI CheY HYIV+PACPAS		<i>chk71</i> <i>crr53</i> <i>chy88</i>	
alr2137 alr2138	HKIII+GAF-(PACPAS)-GAF RRII NarL		<i>chk29</i> <i>crr32</i>	
alr2240 alr2241	RRI CheY HYI		<i>crr47</i> <i>chy48</i>	
alr2279 alr2280	HYIII+PACPAS-GAF-Hpt RRIV+ PP2C_SIG		<i>chy133</i> <i>crr165</i>	
alr2306 alr2307	RRIV+EAL-GGDEF HYI		<i>crr98</i> <i>chy81</i>	
alr2428 alr2429	HYIV Polydomain RRI CheY		<i>chy90</i> <i>crr56</i>	
alr3120 alr3121	HYII+CBS-GAF-PAC HYI		<i>chy28</i> <i>chy49</i>	possibly a single protein (alr3120-1)
alr3155 alr3156 alr3157 alr3158 alr3159	HKI RRII NarL HKII+Phytochrome RRI CheY HYI	<i>cph1</i> , <i>aphA</i> <i>rcl1</i>	<i>chk47</i> <i>crr19</i> <i>chk35</i> <i>crr27</i> <i>chy23</i>	
alr4878 alr4879 alr4880 alr4882	HYII+HAMP-Cache HYI HYI HKI		<i>chy145</i> <i>chy47</i> <i>chy50</i> <i>chk51</i>	
alr5188 alr5189	RRII OmpR HKIII+HAMP		<i>crr78</i> <i>chk87</i>	

N. punctiforme

Assigned gene number	Classification	Functional name	Orthologue number	Additional information
NpF0020 NpF0021 NpF0022	HKII+Phytochrome RRI CheY HYI	<i>cph1</i> , <i>aphA</i> <i>rcl1</i>	<i>chk35</i> <i>crr27</i> <i>chy23</i>	
NpF0303 NpF0304	HKI RRII NarL		<i>chk20</i> <i>crr19</i>	
NpF1084 NpF1085	RRI PatA RRI CheY		<i>crr66</i> <i>crr119</i>	
NpF1183 NpF1184 NpF1185	HKII+Phytochrome RRI CheY HYI+GAF	<i>aphB</i>	<i>chk73</i> <i>crr46</i> <i>chy106</i>	
NpF1211 NpF1212	HYIII+HAMP HYI		<i>chy71</i> <i>chy146</i>	NpF1210

NpF1277 NpF1278	HKIII+PAS RRI AraC		<i>chk79</i> <i>crr153</i>	NpF1279
NpF1600 NpF1601	HYII HYI		<i>chy117</i> <i>chy102</i>	
NpF1775 NpF1776	HKI-HKA3 RRII NarL		<i>chk131</i> <i>crr142</i>	
NpF1799 NpF1800	HYI+HKA3 RRII+NarL		<i>chy109</i> <i>crr141</i>	
NpF2161 NpF2162 NpF2165 NpF2168	RRI >200 RRI CheY HYVI + Hpt-CheW HATPase+CheW		<i>crr67</i> <i>crr103</i> <i>chy93</i>	NpF2163 CheW NpF2164 MCP-GAF NpF2166 CheW NpF2167 MCP-GAF
NpF2204 NpF2205	HisKA HATPase		<i>chk118</i> <i>chk119</i>	10bp apart, possible frameshift
NpF2363 NpF2364	HYIII+PACPAS-GAF-Hpt RRIV+GGDEF		<i>chy44</i> <i>crr8</i>	
NpF3113 NpF3114	HisKA HATPase		<i>chk118</i> <i>chk119</i>	10bp apart, possible frameshift
NpF3565 NpF3566	HKV+GAF RRI CheY		<i>chk105</i> <i>crr117</i>	
NpF3675 NpF3676 NpF3677 NpF3678	HKV+CHASE3 RRI CheY HYI HKA_2 RRIV+PACPAS-EAL-GGDEF		<i>chk171</i> <i>crr122</i> <i>chy108</i> <i>crr162</i>	
NpF3778 NpF3779	HKIV S/Tkin-GAF-PASPAC HKIII+Cache		<i>chk103</i> <i>chk177</i>	
NpF3833 NpF3835 NpF3837	HYI HKI HKII		<i>chy50</i> <i>chk51</i> <i>chk78</i>	
NpF4214 NpF4215	RRII OmpR HKIII+PAS	<i>phoB, sphR</i> <i>sphS, phoR</i>	<i>crr29</i> <i>chk7</i>	NpF4216 PhoU
NpF5034 NpF5035	RRIV+Hpt-GGDEF HYII+GAF-PASPAC		<i>crr93</i> <i>chy127</i>	
NpF5043 NpF5044	HYI+PASPAC RRI CheY		<i>chy104</i> <i>crr118</i>	
NpF5360 NpF5361 NpF5362	HYII+HAMP-Cache RRIV+EAL-GGDEF HYI		<i>chy128</i> <i>crr98</i> <i>chy81</i>	
NpF5479 NpF5480	HYII+HAMP-Cache HYI		<i>chy145</i> <i>chy96</i>	
NpF5636 NpF5637 NpF5640	RRI PatA RRI CheY HYVI+CheW		<i>crr64</i> <i>crr51</i> <i>chy94</i>	NpF5638 CheW NpF5639 MCP-GAF
NpF5960 NpF5961 NpF5964	RRI PatA RRI CheY HYVI+CheW	<i>pilH</i> <i>taxAY3</i>	<i>crr6</i> <i>crr7</i> <i>chy43</i>	NpF5962 CheW NpF5963 MCP
NpF6001 NpF6002	HKII+GAF-PASPAC HYI+PASPAC		<i>chk45</i> <i>chy55</i>	
NpF6362 NpF6363 NpF6364	HKII+GAF-PASPAC RRI CheY HYIV+PASPAC	NtrY	<i>chk71</i> <i>crr53</i> <i>chy88</i>	
NpR0244 NpR0245	RRIV+CheB HYVI+CheW		<i>crr166</i> <i>chy143</i>	NpR0246 CheW NpR0247 MCP NpR0248 CheR NpR0249 MCP

				NpR0250 CheW
NpR0564 NpR0565	HisKA HKIII+HAMP		<i>chk34</i> <i>chk85</i>	
NpR1109 NpR1110	RRII NarL HKI		<i>crr89</i> <i>chk174</i>	
NpR1448 NpR1449	HYI HYI		<i>chy51</i> <i>chy54</i>	
NpR1757 NpR1758 NpR1759 NpR1760	RRI CheY RRI CheY HKII HYII+GAF		<i>crr175</i> <i>crr116</i> <i>chk137</i> <i>chy111</i>	
NpR1778 NpR1779	HKI RRII OmpR		<i>chk70</i> <i>crr44</i>	
NpR2262 NpR2263 NpR2268	HYIV HYIII HYII+PASPAC		<i>chy136</i> <i>chy132</i> <i>chy112</i>	NpR2266/7 STAS NpR2264-PP2C NpR2265-ATPase
NpR2271 NpR2272	HYI HKIV		<i>chy103</i> <i>chk166</i>	
NpR2407 NpR2408	RRII NarL HKIV+HAMP-Cache- (PASPAC) ₂ HisKA3		<i>crr32</i> <i>chk29</i>	
NpR2420 NpR2421	RRII NarL HKI-HKA3		<i>crr84</i> <i>chk58</i>	
NpR2901 NpR2902 NpR2903	HYIV+PASPAC RRI CheY HYI+GAF-PASPAC		<i>chy87</i> <i>crr49</i> <i>chy107</i>	
NpR3052 NpR3053 NpR3054	HKIII+HAMP RRI CheY HKV+CHASE		<i>chk82</i> <i>crr50</i> <i>chk170</i>	
NpR3197 NpR3198	RRII AraC HKIII+PASPAC		<i>crr154</i> <i>chk145</i>	
NpR3562 NpR3564	HYII++ (PASPAC) ₂ -(GAF) ₂ RRIV+PACPAS-EAL-GGDEF		<i>chy64</i> <i>crr161</i>	NpR3563 GAF- PASPAC, overlaps with NpR3562
NpR4743 NpR4744 NpR4745	RRI CheY HYII+PASPAC HKII+GAF		<i>crr10</i> <i>chy66</i> <i>chk61</i>	
NpR4768 NpR4769	RRI CheY HYIV Polydomain		<i>crr56</i> <i>chy90</i>	
NpR5315 NpR5316	HKIII+HAMP RRII OmpR		<i>chk92</i> <i>crr77</i>	
NpR5896 NpR5897	RRIV+PACPAS HYII+GAF-PASPAC		<i>crr68</i> <i>chy4</i>	
NpR6010 NpR6014 NpR6015	HYVI+CheW RRI CheY RRI PatA	<i>pixH, pisH</i> <i>pixG, pisG</i>	<i>chy92</i> <i>crr35</i> <i>crr36</i>	NpR6012 MCP-GAF NpR6013 CheW
NpR6148 NpR6149	HKII HKII		<i>chk74</i> <i>chk16</i>	
NpR6203 NpR6204	HKI RRII OmpR		<i>chk55</i> <i>crr136</i>	
NpR6227 NpR6228	HKIII+HAMP RRII OmpR		<i>chk10</i> <i>crr3</i>	
NpR6346 NpR6347	HYI HYII+Cache		<i>chy45</i> <i>chy69</i>	NpR6348 Ser/Thr kinase
pNPAR130	HYI		<i>chy101</i>	Plasmid encoded

pNPAR131 pNPAR133	HYII+PASPAC HKII		<i>chy115</i> <i>chk148</i>	pNPAR132 ModA molybdate binding sensor protein
pNPBF139 pNPBF140	RRI OmpR HKI		<i>crr138</i> <i>chk123</i>	Plasmid encoded
pNPBF145 pNPBF146	HKI-HKA_3 RRII NarL		<i>chk57</i> <i>crr82</i>	Plasmid encoded

T. elongatus

Assigned gene number	Classification	Functional name	Orthologue number	Additional information
tll0572 tll0571 tll0568	RRI PatA RRI CheY HYVI+CheW	<i>cheA</i>	<i>crr66</i> <i>crr51</i> <i>chy94</i>	tll0569 GAF-MCP
tll1025 tll1024 tll1021	RRI PatA RRI CheY HYVI+CheW		<i>crr12</i> <i>crr11</i> <i>chy142</i>	tll1022 MCP
tll1049 tll1048	RRIV+GGDEF-(RR) ₂ RRII NarL		<i>crr173</i> <i>crr9</i>	
tll1910 tll1909	RRII OmpR HKI		<i>crr38</i> <i>chk55</i>	
tlr0345 tlr0346 tlr0349	RRI PatA RRI CheY HYVI+CheW	<i>piH</i> <i>taxAY3</i>	<i>crr6</i> <i>crr7</i> <i>chy43</i>	tlr0347 CheW tlr0348 MCP

G. violaceus

Assigned gene number	Classification	Functional name	Orthologue number
gll0023 gll0022	HKIII+HAMP RRII OmpR		<i>chk87</i> <i>crr78</i>
gll0572 gll0571 gll0570	HKI RRII NarL KdpD		<i>chk20</i> <i>crr19</i> <i>ctc1</i>
gll0635 gll0634	HYIII+GAF-mHAMP HYIV+PACPAS		<i>chy83</i> <i>chy86</i>
gll0814 gll0813	HKIII+HAMP-PAC RRII NarL		<i>chk159</i> <i>crr17</i>
gll1662 gll1661	HKV (cNMP-HK) HKVI (cNMP-HK)		<i>chk59</i> <i>chk109</i>
gll4098 gll4097	HKI RRII NarL		<i>chk174</i> <i>crr89</i>
gll4162 gll4161 gll4160	RRII OmpR HKIII+PAS RRIV+GGDEF	<i>rpaA</i>	<i>crr31</i> <i>chk84</i> <i>crr158</i>
glr0281 glr0282	RRII OmpR HKI		<i>crr131</i> <i>chk55</i>
glr0347 glr0348	HATPase RRII NarL		<i>chk57</i> <i>crr82</i>
glr0860 glr0861	RRII OmpR HKIII+HAMP		<i>crr3</i> <i>chk10</i>

glr1586 glr1587 glr1588	HisKA HKI RRII NarL	<i>ycf29</i>	<i>chk34</i> <i>chk2</i> <i>crr1</i>
glr1767 glr1768	HKV+GAF-PAS-HWE RRIV+ cNMP		<i>chk108</i> <i>crr171</i>
glr2257 glr2258	RRI CheY HKI		<i>crr108</i> <i>chk48</i>
glr2272 glr2274	RRI CheY RRII OmpR		<i>crr107</i> <i>crr37</i>
glr3138 glr3139	HKI RRII OmpR		<i>chk124</i> <i>crr132</i>
glr3432 glr3433 glr3434	HKII+GAF-PASPAC RRI CheY HYIV+PACPAS		<i>chk71</i> <i>crr53</i> <i>chy88</i>
glr3562 glr3563	HKII RRIV-EAL		<i>chk32</i> <i>crr39</i>
glr4211 glr4212 glr4213	RRIII+Hptm-RR HYII+(PASPAC) ₂ RRI CheY		<i>crr93</i> <i>chy90</i> <i>crr56</i>

P. marinus SS120

Assigned gene number	Classification	Functional name	Orthologue number
Pro1542 Pro1543	RRII NarL HATPase		<i>crr150</i> <i>chk129</i>

P. marinus MED4

Assigned gene number	Classification	Functional name	Orthologue number
PMM0705 PMM0706	RRII OmpR HKI	<i>phoB, sphR</i> <i>sphS, phoR</i>	<i>crr29</i> <i>chk7</i>

P. marinus MIT9313

Assigned gene number	Classification	Functional name	Orthologue number
PMT0805 PMT0804	RRII OmpR HKIII+HAMP		<i>crr72</i> <i>chk91</i>
PMT1356 PMT1357	RRI CheY RRII OmpR		<i>crr48</i> <i>crr28</i>

Synechococcus WH8102

Assigned gene number	Classification	Functional name	Orthologue number
SYNW0808 SYNW0807	RRII OmpR HKIII+HAMP		<i>crr72</i> <i>chk91</i>
SYNW0947 SYNW0948	RRII OmpR HKI	<i>phoB, sphR</i> <i>sphS, phoR</i>	<i>crr29</i> <i>chk7</i>