Fig. 1. An alignment of candidate promoters and NtcA binding sites in cyanobacteria. In bold are the promoter (-10)-boxes, in dark green font – alternative (-10)-boxes, in yellow – conserved regions within repressor or activator binding sites with the consensus GTA-8N-TAC, where substitutions are in green, and transcription initiation sites – in red. Deletions are marked with «*». Complementary binding regions with maximal length are underlined. NtcA is a CAP (catabolite activator protein) family member. The gene's properties and organism taxonomy are shown. Five cyanobacterial orders were studied: Nostocales, Oscillatoriales, Prochlorales, Chroococcales, Gloeobacterales, and the unclassified genus *Acaryochloris* of cyanobacteriae. Genes under putative activation and repression are separated. In the notion "Cyanothece sp. ATCC 51142_i" the i-th ortholog comes from the strain Cyanothece sp. ATCC 51142. Different sites within a leader region in one species/strain are separated by «–"–». The species *Anabaena* sp. PCC 7120 is synonymized into *Nostoc* sp. PCC 7120, *Anabaena azollae* – into *Nostoc azollae*. Number at the end of a line is the distance in a special metric between the candidate of 20 b.p. length and the set of experimentally proved sites. These numbers are preceded by locus tags, the gene identifiers within the genome.

Putatively activated genes:

***nrtA*,** nitrate transport 45kD protein

Chroococcales:

TTAGTTACAAACTATACAAATATTTACAAGGAAAAA*ATC**CAGTCT**AGGAATGC Synechocystis sp. PCC 6803; locus_tag="sll1450"; 18.82
TTTGTAACAATCTATACGAATTTTCCTGTTGAAGA**CTG**TAAATT**TAAGTTAA** Microcystis aeruginosa; MAE_14800; 17.76
TTTGTAGCAAATAATACGAAAAATTCAAAATCCCCA*TGT**AATGCT**TTTATCTA Cyanothece sp. PCC 8802; Cyan8802_4459; 15.76
TTTGTAGCAAATAATACGAAAAATTCAAAATCCCCA*TGT**AATGCT**TTTATCTA Cyanothece sp. PCC 8801; PCC8801_4396; 15.76
TTCGTATCTTTTGTTACAAAAATTTATAAATAATG**GGT**TAATTG**CATAAACG** Cyanothece sp. PCC 7822; Cyan7822DRAFT_4837; 17.18
TTCGTATCTTTTGTTACCAAAAGTAACTTGAAATC**GAT**TAATTA**CATATAAA** Cyanothece sp. PCC 7424; PCC7424_3527; 18.35
CTGGTAGCAATCACTACGCTGATGTGTTGCGGATTG*TGG**TGATGT**AAGTGTGT Synechococcus sp. WH 5701; WH5701_04715; 21.06
GATGTAATGAAAGTTACAGCATAATTTGTCCGAAAACAGGA**TAAACA**ACAGGCA Synechococcus sp. PCC 7002; SYNPCC7002_A1748; 18.12
TGAGTATCACCTGATACAACATCCGCGTTCGCTTTCCAAC**TATAAA**TAAAGAAG Synechococcus sp. WH 8102; SYNW2487; 17.06

Oscillatoriales:

ATTGTAGTTGCCGATACATAAACCCCTGGTGATTTT*TCT**TAGATT**GTGGACAC Arthrospira maxima; AmaxDRAFT_3857; 17.76
ATTGTAGTTGCCGATACATAAACCCCTGGTGATTTT*TCT**TAGATT**GTGGACAC Arthrospira platensis; AplaP_010100019381; 17.76
ATTGTAGTTGCCGATACATAAACCCCTGGTGATTTT*TCT**TAGATT**GTGGACAC Arthrospira sp. PCC 8005; APCC8_010100008251; 17.76

Genus *Acaryochloris*, unclassified order:

CAGGTATCAAATACTACAAAAAAATATTTTTTAGTC*CTT**TAATTT**AAGTTTAT Acaryochloris marina; AM1_5168; 16.35

Gloeobacterales:

TGTGTATCAGAAAACACGATTAGGCCGCGGTACAGC*GTG**TTTCCT**GATGTAAA Gloeobacter violaceus; gvip214; 20.47

Prochlorales:

AATGTATTTGTTAATACAACACAAGTGAAATTTAAA*ATA**AATACT**GGAGCTGT Prochlorococcus marinus CCMP1986; PMM0370; 17.06

***narB*,** nitrate reductase

Chroococcales:

TTGGTATGAAAAGTTACAGTTTACTTCAAGGATTCT***CAA**GAGGAT**AAAGAA Cyanothece sp. PCC 7424; PCC7424_3462; 18.94
CCTGTATTACCAAATACAGCCACTTTTCTCATCCCTT**GAC**TCCTTT**ATTGAG** Cyanothece sp. PCC 8801; PCC8801_2463; 18.71
CCTGTATTACCAAATACAGCCACTTTTCTCATCCCTT**GAC**TCCTTT**ATTGAG** Cyanothece sp. PCC 8802; Cyan8802_3646; 18.71
CAAATAAACTTTATTACAAAGCATCCGCCTTAAGTTTA*ATA**TATAAT**AATACT Cyanothece sp. PCC 7822; Cyan7822DRAFT_4001; 20.71

TGAGTAATAACTGTTAGTAACATCCAACCAACCTTCGA*TCG**CCTCTT**TTAAGT Cyanothece sp. PCC 7425; Cyan7425_4566; 18.47
AATGTAATCTTTGGTGCTTTCTTATTTTAAAATAGTCC*ACC**TCTTCT**GTTTAT Cyanothece sp. CCY0110; CY0110_23416; 22.47
CAAGTGTTTGTCTTTACGAATGGTCAGTTTTCTGGCA**CTT**TGTCCC**CCATTC Cyanothece sp. ATCC 51142; GeneID:6168255; 22.12
ACTGTAGGGTGTATTAGCATAGGGCAATACACCAGA***AAA**CATGAT**TTTGGT Crocosphaera watsonii; CwatDRAFT_3495; 22
ATTGTATTAATAACTACAATTTGGCCGAGAAAATCTGCCGAC**TATCTT**AAGTAT Microcystis aeruginosa; MAE_53960; 16.82
TTTGTATCATTAAGTACCTTTAAGAGGGCTCCAAGC***TGA**TCCAAT**AGCGGG Synechococcus sp. RCC307; SynRCC307_2489; 20.12

Genus *Acaryochloris*, unclassified order:

ACCGTAGCCAATCTTACAATTCTCGATTGAAATATT***AAT**GATTAT**AGAGTT Acaryochloris marina; AM1_2987; 18.24

***narK/nrtP*,** nitrate transporter

Chroococcales:

TCTGTAACAACAACTACTTTTTTGCCATCAAGGGCT**TTT**TGTGAT**GCTTGCA Synechococcus sp. WH 7805; WH7805_09724; 18.71
AGAGTATCAGCGGTTACGAATTTAGCGAAGAAAGAA**TGT**GATTCT**TTATCAC Synechococcus sp. PCC 7002; SYNPCC7002_A1313; 18
GATGTAACACAGGATACAAAAATAAATTTATTGGT***AGC**TATTTG**CACCAAA Synechococcus sp. CC9902; Syncc9902_2271p; 16.94;
GATGTAATGCAAGATACAAAAATGAAATTTTAAGT***AGC**CATTCA**AACCAAA Synechococcus sp. BL107; BL107_06904; 17.53
AAAGTTGCAATCGCTACCAAATCGCTGCGGCACAAGA*TCA**CCTGAT**CCGATCA Synechococcus sp. CC9605; Syncc9605_2642; 19.18
TCTGTAACAACGGCTACTCCTTGACCAAGAGGACCC**GCC**TGTGAT**TCTGGCG Synechococcus sp. WH 7803; SynWH7803_2480; 18.71
GCGGTGACAACCGCTACCAAGCAAAAACTGAACGCG**GAT**GCTGAT**CAATCGG Synechococcus sp. WH 8102; SYNW2462; 20.35

***ntcA*,** transcriptional regulator, global nitrogen regulator

Chroococcales:

AAAGTAGCAGTTGCTACAAGCAGCAGCTAGGCTAGGCCG**TACGGT**AACG Synechococcus elongatus PCC 7942; Synpcc7942_0127; 16.59
TGAGTAATAACCGCTACATGACACATCCAGACAGATCCG**TAGCTT**TCAG Synechococcus sp. BL107; BL107_07964; 18.59
TGCGTAATAACCGCTACAGGACACATCCAGGCAGATCCG**TAACTT**TCAA Synechococcus sp. CC9902; Syncc9902_2075; 18.71
ACGGTGTTCATTGCTACAAGCACTGTCAAGCCGGCTCTT**TACCTT**CCTT Synechococcus sp. RCC307; SynRCC307_2344; 20.12
TTAGTGTCGGTGACTACAGGGCTCGCCATGGGCGATCCT**TAGCGT**CCGT Synechococcus sp. WH 5701; WH5701_01395; 21.88
TTCGTAATAACTACTACACGGCTCGTCAGGGCACATTCG**TAATTT**CCCA Synechococcus sp. WH 8102; SYNW0275; 17.88
TACGTAATAACCGCTACATGGCGGCCACCGCCGGGTCCG**TAATTT**CGCT Synechococcus sp. CC9605; Syncc9605_0269; 18.71
TTCGTGTGGATTGCTACAAGGTGGAGCCAGGGCCATCCG**TAGCGT**CCAG Synechococcus sp. RS9917; RS9917_03753; 20.82
ATCGTGTGCGTTGCTACAGGGTGAGAGTCAATCGCTCCT**TAACTT**CCTA Synechococcus sp. WH 7805; WH7805_07576; 22.24
TTCGTGATGGTTGCTACACGGCACTACGGAACTGATCCG**TAACTT**CCAG Synechococcus sp. RS9916; RS9916_39216; 21.65
ACCGTGTGCGTTGCTACAGGGTGGGAATCGATCGCTCCT**TAATTT**CCTT Synechococcus sp. WH 7803; SynWH7803_0319; 22.24
TTCGTGTTCCTTGCTACACGGGCTGGGTGGGCTGATCCA**TAACAT**CTCG Synechococcus sp. CC9311; sync_0316; 22
AAAGTGTGGCATGATACAAAAAAATACCCCTGAATTGGT**TTTACT**GATC Thermosynechococcus elongatus; tll1650; 19.76
TTTGTATCCGATGATACAAAAAAACGATCGGGAGTTCGC**TTAAGT**GTTG Cyanothece sp. PCC 7425; PCC7424_0372; 16.24
TGGTAAAAAGGTTATACCAAGATTGGCTCCGGCTTTGTG**TAACTG**CTGG Cyanothece sp. PCC 7822; Cyan7822DRAFT_4473; 23.53

Nostocales:

TGAGTATAGGAAAGTACAGAAAGGTTGACGGTGCTTTAT**TCATTT**TTCG Nostoc punctiforme; Npun_F5511; 19.41
TGAGTATAGGAAAGTACAGAAAGGTTAACGGTG***CTT**TATTGA**TTTT Anabaena variabilis ATCC 29413; Ava_3283; 19.41
TGAGTATAGGAAAGTACAGAAAGGTTAACGGTG***CTT**TATTGA**TTTT Nostoc sp. PCC 7120 (Anabaena sp.); alr4392; 19.41
TGAGTATAGGAATGTACAGAAAGGTTGACGGTG***CTT**TATTGA**TTTT Nodularia spumigena; N9414_19492; 19.76
TACGTATATTTAAATACAAATAAAATCCAGCTT***TGT**TTTTGA**TCAA Nostoc azollae (Anabaena azollae); AazoDRAFT_5432; 18.47
AATTTATAGGTAGTTACAGTAGCCATTATTGCTTAAGAT**TAAATA**TTAT -"-; AazoDRAFT_5432; 21.18

TGAGTATAGTAAAAGACAGAAAGATTGACGGTGCG*TTC**TTGATT**CTTC Cylindrospermopsis raciborskii; CRC_00858; 21.06

Genus *Acaryochloris*, unclassified order:

TATGTATATTTTGATACATAAAGAAACCCATCAGTCCGC**TTAATT**TATG Acaryochloris marina; AM1_3155; 17.76

Gloeobacterales:

TTAGTAGTCCCTAATACAAAAAAATTGCAGGGCTCGGCT**TAATTG**TTAA Gloeobacter violaceus; gvip454; 17.06

Prochlorales:

ATTGTTACAGTTGATACAAGATATTTCATGGTGTCTTCT**TAAGTT**TTTT Prochlorococcus marinus MIT 9312; PMT9312_0248; 18.59
ACCGTCACCATTGCTACATGGTGGTGGAGGTCCACTCCC**TAACTT**CATA Prochlorococcus marinus MIT 9303; P9303_24561; 21.18
ACCGTCACCATTGCTACATGGTAGTGGAGGTCCACTCCC**TAACTT**CATA Prochlorococcus marinus MIT 9313; PMT1831; 21.18
ACTGTTACTGTTGATACAAGATATTTTCATCTGTCTCCT**TAAATT**TTTA Prochlorococcus marinus MIT 9515; P9515_02791; 19.29
AAAGTCGCCATGGCTACATATCGTTTTTTCTTGACTTAT**TAATGT**TCTT Prochlorococcus marinus NATL1A; NATL1_03241; 20.12
AAAGTAGTTTTAAATACTTTTTGAATTGCGTCGCAATTTCTTGAGCGTA -"-; NATL1_03241; 20.59
AAAGTCGCCATGGCTACATATCGTTTTTTCTTGACTTAT**TAATGT**TCTT Prochlorococcus marinus NATL2A; PMN2A_1612; 20.12
AAAGTAGTTTTAAATACTTTTTGAATTGCGTCGCAATTTCTTGAGCGTA -"-; PMN2A_1612; 20.59
GAAGTATTTATATTTACAACATCAAGATTATTTTTTGAA**TAAATA**TGAG Prochlorococcus marinus AS9601; A9601_02681; 18.35
ACTGTTACTGTTGATACAAGATATTTTAGAGTGTCTTCT**TAAATT**TTTT -"-; A9601_02681; 19.29
GAAGTATTTATATTTACAACATCAAGATTACTTTTGGAA**TATAGA**TGAG Prochlorococcus marinus MIT 9301; P9301_02691; 18.35
ATTGTTACTGTTGATACAAGAAATTTTAGAGTGCCTCCT**TAAATT**TTTT -"-; P9301_02691; 19.29
TGTGTAAATCTTTTTACGTTAATTTTTTCAGCAG**TAT**TATTAACAAC Prochlorococcus marinus MIT 9215; P9215_02691; 21.29
ATTGTTACTGTTGATACAAGATATTTTAAGGCGCCTCCT**TAAATT**TTTT -"-; P9215_02691; 19.29
AAAGTCATTTTTGATACAAGAGGTTTTAGCTTCACTTCT**TATCGT**CCTT Prochlorococcus marinus CCMP1375; Pro0277; 19.88
AAAGTCATTATTGCTACAAGTGGTTTTATAACCCGTCCT**TAGCGT**CGAT Prochlorococcus marinus MIT 9211; P9211_02711; 19.18
TTTTTATTTATGTTTACAACTTCATTATTAGTTGTTAAA**TAACTA**TTTG Prochlorococcus marinus CCMP1986; PMM0246; 20
AATGTTACTGTTGATACAAGTTATTTCTTGCTGTCTTCT**TAAGTT**TTTT -"-; PMM0246; 19.06

***ntcB*,** nitrate assimilation transcriptional activator

Chroococcales:

GCTGTATCACTATCTACATAATTGCTGGGGGCAAATCCGT**TAAGGT**CACAGAA Cyanothece sp. PCC 7425; Cyan7425_1599; 18.35
TGTGTATCAAGATCTACGGAACATTAATCAAAAACC*TGA**TAGTGT**CTTGAGA Cyanothece sp. PCC 7424; PCC7424_0664; 18.47
GTTGTATCAAGATCTACTGAAAGCGAACTCAAAACC*TGT**TACTGT**CATGAGA Cyanothece sp. PCC 7822; Cyan7822DRAFT_4828; 18.71
GGTGTAGAAACTTTTACTTGTTCGTAAGGCATAGT**TAG**TATTAGATATATT -"-; Cyan7822DRAFT_4828; 19.1
TTTGTATCAAATTCTACTAAAATTAAAGTGTGAATC*CCC**TATTGT**GCAATAA Cyanothece sp. PCC 8802; Cyan8802_2380; 16.24
TTTGTATCAAATTCTACTAAAATTAAAGTGTGAATC*CCC**TATTGT**GCAATAA Cyanothece sp. PCC 8801; PCC8801_2329; 16.24
TTAGTATCAATTTCTACGAAATCGGTGACAAAAATA*AGA**TAAGGT**TAAGATA Cyanothece sp. ATCC 51142; cce_0198; 17.18
TTAGTATCAATTTCTACGAAATCGATAGCAAAAATA*AGA**TAAGGT**TAAGATA Cyanothece sp. CCY0110; CY0110_21295; 17.18
TTAGTATAAATTTCTACGAAATAAGGAACAAAGCTCGTGA**TATGAT**AAAGTAT Crocosphaera watsonii; CwatDRAFT_3020; 17.88
ACGGTAGCAATTTCTACCAAAAGTCAGAATAAAATT*CGT**TATCGT**CATAAGG Microcystis aeruginosa; MAE_14790; 17.29
TTTGTAACATCCGGAACAAAAAGCCCCAAAATCAATGCAG**TAAACT**CATTGGG Synechococcus sp. PCC 7002; SYNPCC7002_A1632; 18.82
ACTGTATCAATCAACACACTTTACCATGGGAAAGTT*GCA**TAGATT**GAAGGTA Thermosynechococcus elongatus; tll1359; 20.12

Nostocales:

GCTGTAACAAAATCTACCAAATTGGGGAGCAAAATC*AGC**TAACTT**AATTGAA Nostoc sp. PCC 7120 (Anabaena sp.); all0602; 17.41
ACTGTAACTAACTATACTAAATTGCGGAGAATAAAC*CGT**TAACTT**AGTGAAA Nostoc punctiforme; Npun_F1534; 18.12
AATGTAACCAACTTTACTAAATAGCAGATAAAAAAC*CGT**TAATTT**AGTAAAG Nodularia spumigena; N9414_13370; 18.12

TATGTAGCAGAATCTACCAAATTGGGGAACAAAATC*CGC**TAACTT**AAATTAG Anabaena variabilis ATCC 29413; Ava_4535; 16.94
TTTGTAAACTGTTTCACAGACCAACTCATGCCAAAC*TGA**CATTAG**GGGGTGA Nostoc azollae (Anabaena azollae); AazoDRAFT_2534; 20.94
TTTGTATTATAACCTACTAAAATAATGGAATCAGAA*TTT**TAAGTC**AAAAACT Raphidiopsis brookii; CRD_01496; 17.88
TTTGTATCATAACCTACTAAAATGATGGAATCAGAA*TTT**TAAGTC**AAAAACT Cylindrospermopsis raciborskii; CRC_01749; 17.76

Gloeobacterales:

TACGTAACCCCAGATACATTCGTGGCCGTGCGGGCC*GGA**TATGGT**CTG---- Gloeobacter violaceus; gvip213; 20.24

Genus *Acaryochloris*, unclassified order:

GTTGTATCGTTTCTTACCGAATAGCCGGGCCAAATT*CGT**TACTGT**GCTCTCA Acaryochloris marina; AM1_4387; 19.41

Oscillatoriales:

GTTGTAGCAATTTTTACCAAATTATATCGAGAAATT*GGT**TAAGAT**GAAACCC Arthrospira maxima; AmaxDRAFT_4787; 16.59
GTTGTAGCAATTTTTACCAAATTATATCGAGAAATT*GGT**TAAGAT**GAAACCC Arthrospira platensis; AplaP_010100013008; 16.59
GCGGTATCAAAGTGTACTAAATTGTTTGAGGAAATC*GGT**TAAGGT**CATTCAG Oscillatoria sp. PCC 6506; OSCI_3060008; 19.29
GTTGTAGCAATTTTTACCAAATTATATCGAGAAATT*GGT**TAAGAT**GAAACCC Arthrospira sp. PCC 8005; APCC8_010100003876; 16.59
GCTGTATCAACCTCTACTAAACTCAGAATGTAGATC*AAG**TACCAT**AAATTTG Lyngbya sp. PCC 8106; L8106_02207; 17.65
TTTATATTTTTGGTAATCTATACAGTGCTCCTGCGTAAGT**TATGAT**AATATTT Trichodesmium erythraeum; Tery_4333; 25.53

Two NtcA binding sites in *Nostoc* sp. PCC 7120, *Nostoc punctiforme*, *Nodularia spumigena*, *Anabaena variabilis* ATCC 29413 upstream gene *ntcB*.

Nostocales:

Nostoc sp. PCC 7120 (Anabaena sp. PCC 7120)
AAAGCTGTAACAAAATCTACCAAATTGGGGAGCAAAATC*AGC**TAACTT**AATTGAATGCCGTAAATGTAATTAAGGCTACATACAA
Nostoc punctiforme
ATTACTGTAACTAACTATACTAAATTGCGGAGAATAAAC*CGT**TAACTT**AGT*GAAAGCTATAAATGTAATTCAGGCTACATAAAA
Nodularia spumigena
AAAAATGTAACCAACTTTACTAAATAGCAGATAAAAAAC*CGT**TAATTT**AGT*AAAGGCCATAAATGTAATTCAGGCTACATAAAA
Anabaena variabilis ATCC 29413
AAATATGTAGCAGAATCTACCAAATTGGGGAACAAAATC*CGC**TAACTT**AAATTAGTGCGATAAATGTAATTCAGGCTACATACAA

***nirB*,** nitrite reductase related protein

Chroococcales:

TTAGTAGCAATTGCTACAAGCCTTGACTCTGAAGCC*CGC**TTAGGT**GGAGCCA Synechococcus elongatus; syc0309_c; 16.24

***glnA*,** glutamine synthetase, glutamate-ammonia ligase

Chroococcales:

ATGGTAGCGAAAAATACATTTTCTAACTACTTGA*CTCTT**TACGAT**GGATAGTC Synechocystis sp. PCC 6803; slr1756; 19.29
TTTGTATCGAAAAATACAAAACCGGCCAAAGTGA*CTGTT**TAGGAT**GGATTTTT Microcystis aeruginosa; MAE_19270; 16.24
ACAGTATCAAATTATACAAAATGAAATAAAGCCT*ATCTT**TACTAT**CAATACTT Cyanothece sp. CCY0110; CY0110_09196; 15.65
ACAGTATCCAATTATACAAAATGAAGTAAAGCCT*ATCTT**TACTAT**CAATCATT Cyanothece sp. ATCC 51142; GeneID:6171510; 16.71
TCCGTATCATGGTATACAAAATCTCAGCTATCCT*ATCCC**TAGGAT**CAGTGATT Cyanothece sp. PCC 8801; PCC8801_0808; 18.59
TCCGTATCATGGTATACAAAATCTCAGCTATCCT*ATCCC**TAGGAT**CAGTGATT Cyanothece sp. PCC 8802; Cyan8802_0837; 18.59
ATTGTAACTCATGATACAAAAGCTGTTTCGTCCAAGTCCC**TAGGAT**ATTACATC Cyanothece sp. PCC 7425; Cyan7425_1839; 16.35
TTCGTATCAACAGATACTTTTTTTCAATGTTTACAGATTC**AAAGAG**TGACAATA Cyanothece sp. PCC 7424; PCC7424_2203; 19.53
TCTGTATTGAGTAGCACAAAAGTTAAAGAAACCT*ATCCC**TACGAT**TGATGATT -"-; PCC7424_2203; 18.94

TCTGTATTGAGTAGAACAAAAAATAAAGAAACCT*ATCCC**TACGAT**TGATGATT Cyanothece sp. PCC 7822; Cyan7822DRAFT_1561; 18.94
ACAGTAAAGTAAAATACATTTAGCCGACGATATT*TAAGG**TATGAC**CTTTTATC -"-; Cyan7822DRAFT_1561; 20.59
ACAGTAGCAAATTATACAAAATATGGTAAAGTGA*GTCTT**TACCAT**CAATGACT Crocosphaera watsonii; CwatDRAFT_5271; 15.41
AGTGTGCCTGCTGATACAAAAAGGGCCCTGCGGG*GTGTT**TACGTT**TCTCATGT Synechococcus sp. RS9916; RS9916_29064; 18.35
AAAGTGCGCGTTGATACAAAAGGGTGTCAGATCG*ATGCC**TACGGT**TAGAAAAG Synechococcus sp. CC9311; sync_1569; 20.24
AATGTGCGCGTTGATACAAAACAGGGCATAACGG*CTCCT**TACGGT**CGTTCTTA Synechococcus sp. WH 8102; SYNW1073; 19.65
GATGTGACCCAGACTACAACGAACTGCCATAACT*CTTTC**TAGGAT**CAAGCGAA Synechococcus sp. PCC 7002; SYNPCC7002_A1630; 21.06
AATGTACCCATTGATACAAAACCATCCCACGGCG*ATCCC**TACGGT**CATGCGGT Synechococcus sp. WH 7803; SynWH7803_1347; 16.12
AATGTACTGACGGATACAAAACGGGTGCACATCG*ATCCC**TACGTT**CAAATTGC Synechococcus sp. WH 7805; WH7805_02267; 16.82
AACGTACCCCTTGATACAAAATGGTGTCTTGTGG*CTGTT**TACGGT**CACATCAC Synechococcus sp. CC9605; Syncc9605_1205; 18.12
TTTGTATCAGCTGATACACAGAGGGTCCAAAGTG*CTCCG**TACGTT**CAGCTTTC Synechococcus sp. RCC307; SynRCC307_1296; 16
AGGGTTACAACGGGCACCAAAGGCCAGCGGGACC*TCTCC**AAGCAT**CGATCCCG Synechococcus sp. WH 5701; WH5701_13575; 22.47
GAGGTAAGACTCAATACTGTCGATCGCCGATCTGTCGAAA**TAGCGG**CCCAGCTG Synechococcus sp. CC9902; Syncc9902_1265; 22.35
AATGTCCCTCTTGATACAAAATGGAGCCATCGAG*ATCTT**TACGGT**CTCTCTAA -"-; Syncc9902_1265; 18.71
AATGTCCCCCTTGATACAAGATGGAGCCATCGAG*ATCTT**TACGGT**CTCTCTAA Synechococcus sp. BL107; BL107_13740; 20.12
AATGTACGCATTGATACAAAAAGAGGGTCGATCC*CTCCC**TACGGT**CCGCCACG Synechococcus sp. RS9917; RS9917_11790; 17.06
TATGTATCAGCTGTTACAAAAGTGCCGTTTCGGG*CTACC**TAGGAT**GAAAGCGG Synechococcus elongatus; Synpcc7942_2156; 14.71
TTTGTAGTCCTTGATACAAAAAACTCTTGGGGTCACTTTC**TACCAT**AGATTGTC Thermosynechococcus elongatus; GeneID:1010718; 16.71

Prochlorales:

AAAGTACTTATTGATACAAAATAAGTTTGCTACG*GTAAT**TAATTT**TATTTATA Prochlorococcus marinus CCMP1375; Pro1038; 16.47
AAAGTACTTGTTGATACAAAAACAGATCCAAACG*CTAAT**TAATTT**CATATTTA Prochlorococcus marinus MIT 9211; P9211_10271; 17.06
CAGGTACCTGCTGCTACAAAAAGAAGGCCATGTGGCTGAT**TATGGT**CGCTCTCT Prochlorococcus marinus MIT 9303; P9303_16421; 17.88
AAGGTACCTGTTGCTACAAAAAGGGGGCCATGCGGCTGAT**TACGGT**CGCTCTCT Prochlorococcus marinus MIT 9313; PMT0601; 17.41
AAAGTAACCGTTGATACATTATAAGAATCATCAA*GTTAT**TAATTT**CATAGAAT Prochlorococcus marinus MIT 9215; P9215_09701; 19.18
AAAGTTACCGTTGATACATAACAAGAATCATCAA*GTTAT**TAATTT**CATATAAG Prochlorococcus marinus AS9601; A9601_09401; 20
AAAGTTACCGTTGATACATAACAAGAATCATCAA*GTTAT**TAATTT**CATATAAG Prochlorococcus marinus MIT 9312; PMT9312_0880; 20
AAAGTTATCGTTGATACATAATTAGATTCATCAA*CTTAC**TAATTT**TATTTCAA Prochlorococcus marinus MIT 9515; P9515_10031; 20.12
AAAGTTACCGTTGATACATAACAAGAATCATCAA*GTTAT**TAATTT**CATATAAG Prochlorococcus marinus MIT 9301; P9301_09391; 20
AAAGTGCGTCTTAATACAAAAAGAGATGCGATCG*CTTAT**TACGTT**CATAATTA Prochlorococcus marinus NATL2A; PMN2A_0141; 20.35
AAAGTGCGTCTTAATACAAAAAGAGATGCGATCG*CTCAT**TACGTT**CAAAATTA Prochlorococcus marinus NATL1A; NATL1_07731; 20.35
AAAGTTACCTTTGATACATAATTAGATTCGTCAA*GTTAT**TAATTT**TATTGGAG Prochlorococcus marinus CCMP1986; PMM0920; 20.12

Nostocales:

TCTGTAACAAAGACTACAAAACCATCTAATGTTTAGAATC**TAGGAT**ATTTCAGG Anabaena variabilis ATCC 29413; Ava_0147; 15.88
CCTTTAACGAAAACCACAAACTTTCCTCAAGTCAACTATT**TACGAT**ATTCCAGG Nostoc azollae (Anabaena azollae); AazoDRAFT_1773; 21.18
TCTGTAACAAAGACTACAAAACTGTCTAATGTTTAGAATC**TACGAT**ATTTCAGG Nostoc sp. PCC 7120 (Anabaena sp.); alr2328; 15.88
TCTGTAACATAAGCTACAAAATCCGCTAATGTCTACTATT**TAAGAT**ATTTCTGG Nostoc punctiforme; Npun_R5387; 16
GCTGTAGCAAAAAGTACAAAATCTTCTAATGTCCAGGATT**TATGAT**ATTTCAGG Nodularia spumigena; N9414_03588; 16
ATCGTAACAAACAATACAACCTGGTGGGGA*TCTAGTCTT**TAGGAT**ATTTCAGA Raphidiopsis brookii; CRD_02136; 17.88
ATCGTAACAAACAATACAACCTTGGTAGGACTCTAGTCTT**TAGGAT**ATTTCAGA Cylindrospermopsis raciborskii; CRC_02160; 17.88

Oscillatoriales:

AGTGTAATGCTTGATACAAAACTTAATCAATCTGAGTCAC**TATGAT**CAACCAGA Lyngbya sp. PCC 8106; L8106_18726; 17.06
AGTGTAGCACTCAATACAAAACTTAATCAATTGAAGTCAC**TATGAT**CAACCAGA Arthrospira maxima; AmaxDRAFT_3347; 16.82
TCTGTATCCCCTGCTACCAAAGCATAAGTTGCATAGCGCA**GATAAT**AATCCATG Arthrospira platensis; AplaP_010100011066; 17.65
TCTGTATCGCCTGCTACCAAAGCATAAGTTGCATAGCGCA**GATAAT**AATCCATG Arthrospira sp. PCC 8005; APCC8_010100012680; 17.53

AGTGTAGCACTCAATACAAAACTTAATCAATTGAAGTCAC**TATGAT**CAACCAGA -"-; APCC8_010100012680; 16.82
TTTGTATCAAGAAATACAAACTTGTGACAGTCGAA*TGCC**TATGAT**CGCTCAAA Oscillatoria sp. PCC 6506; OSCI_3060031; 16
AGCGTAACACCCGATACAAACATTCAATTTAAATTCTTAT**TATTAT**TTGAGTGT Trichodesmium erythraeum; Tery_3834; 17.76

Genus *Acaryochloris*, unclassified order:

ACTGTATTTTCTGATACAAAATCTCTGTGATCCAGTCTCT**TAGGAT**ATTAGGCT Acaryochloris marina; AM1_1497; 16

Gloeobacterales:

AGCGTATCTCAGACTACAAAACACCTCCCCGACC*CTCCT**TAGTAT**GGGTTGGC Gloeobacter violaceus; gvip146; 18.59

***nirA*,** ferredoxin-nitrite reductase

Strains *Prochlorococcus marinus* MIT 9215, MIT 9515, MIT 9301, MIT 9312, MIT 9211, CCMP1986, CCMP1375, AS9601, MED4 lack gene *nirA*. In strains *Prochlorococcus marinus* MIT 9313, MIT 9303 the site consists of long degenerate tandem repeats.

Chroococcales:

AGTGTAATTTACGTTACAAATTTTAACGAAACGGGAACCC**TATATT**GATCTCTA Synechocystis sp. PCC 6803; slr0898; 17.29
TCTGTTACAGACAATACAAAATTACCCAAGTTAAAACGAA**TTAAAT**AACAGATA Microcystis aeruginosa; MAE_18410; 18.47
TTGGTAACAAAGTACACAAAGTAAGCCGTGTTTTCT*CCG**TAAGTT**ACGAAGTA Cyanothece sp. PCC 7822; Cyan7822DRAFT_1390; 19.29
TTGGTAACAAAAGACACAAAGTAGTCGGGAAATTCTCCGG**TAAGTT**ACACAAGT Cyanothece sp. PCC 7424; PCC7424_1683; 18.35
TTTGTAACTTAATATACCAATTACCCTCTAGGACT*CCCC**TAAGTT**AGCCAGAT Cyanothece sp. PCC 7425; Cyan7425_4573; 18.35
TTTGTTACATTAGCTACAAAATATCTC*AAATGGTAGAGG**TTAAAT**AGGTACAA Cyanothece sp. ATCC 51142; cce_1223; 18.24
TTTGTTACAACTGATACAAAAATTATCCAAAACAAA*CCA**TTAAAT**GAGAACAA Cyanothece sp. PCC 8802; Cyan8802_3641; 16.59
TTTGTTACAATTAATACAAAAATTATCCAAAACAAA*CCA**TTAAAT**GAGAACAA Cyanothece sp. PCC 8801; PCC8801_2468; 17.06
TTTGTTACATTAGCTACAAAATTTCTCGGATTGTAG*AGT**TTAAAT**AGGTACAA Cyanothece sp. CCY0110; CY0110_23451; 18.24
TTTGTTACATTAGCTACATAAATTCTCCTATTGTGAAGTT**TAAATT**AAATACAA Crocosphaera watsonii; CwatDRAFT_3683; 19.29
AAAGTAGCAAAATTTACAAATGTTCATGATTCATCT*GGC**TAAATT**GGATGTTC Thermosynechococcus elongatus; tlr1349; 15.65
CCCGTAGTTAACACTACAAAATCTCCCTGACTGAAT*GTT**TACCTT**GGTCAAAT Synechococcus sp. PCC 7002; SYNPCC7002_A1827; 18.24
TTCGTAGCAACGGTTACTTCCTCCAGCGGGCTCAAT*CCA**TTAAGT**AGTCAAAA Synechococcus sp. RCC 307; SynRCC307_2482; 19.41
GACGTAGCCATTGTTACAAATTTGCGGGACCGAATCCTCT**GAAAAA**GGGCAACT Synechococcus sp. CC9605; Syncc9605_2656; 16.82
AACGTGGCCACTCGTACAGCTCAGCAGCCTGTGACTGTAT**TGAAAG**CCCGTCAG Synechococcus sp. WH 5701; WH5701_13695; 20.94
ACGGTCTAATTCGCTACCGACTGAAAGGGATACGCA*CGG**TTAATT**GTGTGAGT Synechococcus sp. CC9902; Syncc9902_2284; 22.24
ACGGTCTAATTCGCTACCGACCAAAAGGGCATTGAC*GAG**GTTAAT**GGTATCGA Synechococcus sp. BL107; BL107_06834; 22.24
TTGGTTCCAACCGATACCGAAAGATCACACCACAGATCGC**TATAAA**CCTTTTCT Synechococcus sp. CC9311; sync_2898; 20.94
TTGGTTCCAACCGCTACCGATGGATCCCCAGGCCAACCGC**TAGAAG**ACTTCTGC Synechococcus sp. RS9917; RS9917_06020; 20.94
ACTGTTCAAACTGCTACCGACAGCAACGGTACGAAC*CAG**TTCACT**ACATCTAA Synechococcus sp. WH 7805; WH7805_09654; 19.53
AATGTTCGAACTGCTACCGACAGCAATACAACGAAA*CGG**AAAACT**GTTGTTGG Synechococcus sp. WH 7803; SynWH7803_2492; 19.53
GTTGTAGTTTCTGTTACCAATTGCGAATCGAGAACT*GCC**TAATCT**GCCGAGTA Synechococcus elongatus; syc0310_d; 17.29

Prochlorales:

GATGTATTGAATGATACCAAAATAAATAAAAAATTTGGTA**AAAACT**GATACATC Prochlorococcus marinus NATL1A; NATL1_21711; 17.06
TTGGTAAAAACTGATACATCTAAAGTTTAGATTTTAAAGG**TTAATT**ATAAATGT -"-; NATL1_21711; 18.35
GATGTATTGAATGATACCAAAATAAATAAAAAATTTGGTA**AAAACT**GATACATC Prochlorococcus marinus NETL2A; PMN2A_1298; 17.06
TTGGTAAAAACTGATACATCTAAAGTTTAGATTTTAAAGG**TTAATT**ATAAATGT -"-; PMN2A_1298; 18.35
TTTGTTCAATCTGATACCACTAATACCCTCTTCGTTGAGG**CTTAAT**GGCTCCAT Prochlorococcus marinus MIT 9303; P9303_29861; 20.47
TTTGTTCAATCTGATACCGCCAATGCCCTCTTCGCTGAGG**CTTAAT**GGCTCCAT -"-; P9303_29861; 21.18
TTTGTTCAATCTGATACCACTAATACCCTCTTCGCTGAGG**CTTAAT**GGCCCCAT -"-; P9303_29861; 20.47
TTTGTTCAATCTGATACCGCCAATACCCTCTTCCATCAGG**CTTAAT**GACTCCAT Prochlorococcus marinus MIT 9313; PMT2239; 21.18

TTTGTTCAATCTGATACCGCCAATACCCTCTTCCATCAGG**CTTAAT**GACTCCAT -"-; PMT2239; 21.18
TTTGTTCAATCTGATACCGCCAATACCCTCTTCCATCAGG**CTTAAT**GACTCCAT -"-; PMT2239; 21.18

Nostocales:

ATCGTAACAATTTATACGATTTTAACAGAAATCTCG*TCT**TAAGTT**ATGAGTAT Nostoc punctiforme; Npun_R1528; 18.94
TTTGTAGCTACTTATACTATTTTACCTGAGATCCCG*ACA**TAACCT**TAGAAGTA Nostoc sp. PCC 7120 (Anabaena sp.); alr0607; 17.76
AAGGTAGCAATGTATACTATTTTAACTATAAACTCA*CCT**TAACCT**GATAAGTT Nodularia spumigena; N9414_05289; 19.29
TAAGTAATTATGACTACGCCACTCTACGAGAGCAGG*AAT**CAAACT**GGATTCCT Nostoc azollae (Anabaena azollae); AazoDRAFT_0101; 20.47
GGTTCTTCCAGAATTACTGAATTGCTATCACTAGACCTCG**TAGAAC**CAGCAAGA Anabaena variabilis ATCC 29413; Ava_4539; 25.41
TCTGTAACTGAGTATACGAATCCTTCCCGAAATTTG*CCC**TAGGTT**ATCAATAT Cylindrospermopsisraciborskii; CRC_00047; 18.82
TCCGTAACTGAGTATACGAATCCCTCCCGAAATTTG*CCC**TAAATT**ATCAATAT Raphidiopsis brookii; CRD_02595; 19.88

Oscillatoriales:

TTTGTAACTTTTGCTACGATTAAACTC*AAATTTTCCTTC**TACTGT**AGGTAGTA Lyngbya sp. PCC 8106; L8106_10091; 18.82
AATGTTAAATCTAATACAGACTGAAAG*AAACTAATTTGT**TAACGT**TTGGCCAA -"-; L8106_10091; 18.59
TTTGTTATATAAGCTACGAATTCTTTATTTCCTTTC*CTT**TAAATT**GATAAATC Trichodesmium erythraeum; Tery_1068; 19.41
CCTCGTTAATTAATTACTAAATAATACCCAGACATT*AGA**TAAGTT**TAAACTTT Arthrospira maxima; AmaxDRAFT_3713; 21.82
TATGTATCTTCAACTACAAATTATCACTGATTCTGA*ATT**TAATCT**ACAAATCA Arthrospira platensis; AplaP_010100015448; 16.35
TTGGTATTCCAAAATACTAAAACCTGGCAAG*********TATGCG**TTTTTTAC -"-; AplaP_010100015448; 19.65
TCTGTAACATTAGCTACGAACAACTCTCGATCCACG*CCC**TACAGT**GTTCCTTA Oscillatoria sp. PCC 6506; OSCI_2880001; 17.53
AAGGTAGTAATTAATACTCTTTGCTGACTAAATATTTCTC**TGTAAT**CTTGACCA -"-; OSCI_2880001; 18.94
TGAGTAATTAAATATACGGTTTTAGTTAAG**********TATAAC**CACTAAGG -"-; OSCI_2880001; 20.47
TATGTATCTTAAACTACAAATTATCACTGATTCTGA*ATT**TAATCT**ACAAATCA Arthrospira sp. PCC 8005; APCC8_010100020361; 16.47

Genus *Acaryochloris*, unclassified order:

TTTGTAACGACTGATACGAACGATCCCAATGGGAGT*CGC**TAAGTT**AGTATCCC Acaryochloris marina; AM1_2984; 16.82

Gloeobacterales:

AATGTATCTGGGGTTACGTAAAGTTAAGTTCCTTCAAAGC**ATTCGT**GGGACGCT Gloeobacter violaceus; gvip212; 19.65

***glnN*,** glutamine synthetase type III

Chroococcales: -35 box

TTTGTATCTATA**TTGTCT**ATTTTAAAAAATCATCT*TGCG**TATGAT**TGGGGGG Synechocystis sp. PCC 6803; slr0288; 22.82
TCTGTATAGATCTTATCTGTTTTTAGGATTAGTCT*TGGG**TAGTAT**GGTGCGG Synechococcus sp. PCC 7002; SYNPCC7002_A0246; 24.94
TCTGTAGCAACGGCTTCATCTCAGGGGTTGCCTGC*CGAC**TTGGAT**GTCTTCC Synechococcus sp. WH 5701; WH5701_08944; 19.29
TCCGTAACAACACCCACATTCAGCTGTCACCGCGCTCCTG**TTCTGT**GTGCTCA Synechococcus sp. RS9917; RS9917_10876; 21.18
TCTGTATCTTTTCTAGCGATCGAGCTGGTCACCAT*TGAG**TACGAT**CAATTGA Synechococcus elongatus; syc1338_c; 23.29
CTCGTACATGGTGCTGGCTTACACCGATGAGGAGA*ACGT**GAGCCT**CAATGTG Synechococcus sp. WH 7805; WH7805_04111; 27.06
CTCCTACATGGTGTTGGCTTACACGGATGCGGAGA*GTGT**GAGCCT**CAACGTG Synechococcus sp. WH 7803; SynWH7803_1458; 28.94
TTGGTGAGCAAAAATACGCACCTGAGGCCCATAAACCATG**CCCAAT**CCATCAC Synechococcus sp. CC9311; sync_1253; 22.71
TTTGTAATGTTTGATACAAATCTCTCGATCATCTG*TTCT**TAAGAT**GAAGGCT Microcystis aeruginosa; MAE_09050; 16.71
TGAGTAGCTCGTGATACAAAAAGCGGAATTTCTC*CCCC**TACTGT**ATGAAAT Cyanothece sp. PCC 7425; Cyan7425_4041; 17.41

Gloeobacterales:

CGAGTACCGGCCCATCCCCTCGATCTGGGGCCACCGCGCC**GGCAAT**CCCTATC Gloeobacter violaceus; gll2499; 25.06

Genus *Acaryochloris*, unclassified order:

TCGGTATCCGTTCTATCCATTTTAAGGAATAAAT*TTGCG**TACCAT**CAGGAAG Acaryochloris marina; AM1_0336; 24.47

***glnB*/*glnK*,** nitrogen regulatory protein P-II

Cyanothece sp. PCC 7424 and *Crocosphaera watsonii* are closely related. *Crocosphaera watsonii* might have acquired substitutions (or indels) within the *glnB* gene regulatory region, which lead to the loss of conservativity:

```
Cyanothece sp. PCC 7424 ATT*GTACTAACTTTTAC**AAAAATTTTTGGGTTAATGG*TGATAGAAAGTGACAG PCC7424_0456; 15.41
Crocosphaera watsonii   ATCAATACTCATCAATACTTAAAAATCTTAGGGTTTATGG**GCAACTGGGTGGGAA CwatDRAFT_5924; 22.47
```

Substitutions in the *glnB* gene regulatory region in *Nostoc azollae* (*Anabaena azollae*) led to the loss of conserved binding sites:

```
Nostoc azollae          AAC*AAAAAAAGAGGTGATATAAAATACATAACACTCATGTTAAATAAGTAAAGCTA AazoDRAFT_6171; 28.35
```

Chroococcales:

```
ACGGTACTGATTTTTACAAAAAAAC*TTTTGGAGAACATGTTAAAAGTGTCTGG Synechocystis sp. PCC 6803; ssl0707; 17.76
AATGTAGCTATTGCTACACGCTATACTCTTTTGCCAA*TGCTAAATAAAGATTA Thermosynechococcus elongatus; tll0591; 17.06
TCGGTATTCCTCTGTACCTAGAGATGATTAAAATCATTGGCTGTGGAGTCTAAA Microcystis aeruginosa; MAE_59130; 22.24
ATTGTACTAACTTTTACAAAAAATTTTTGGGTTAATGG*TGATAGAAAGTGACAG Cyanothece sp. PCC 7424; PCC7424_0456; 15.41
ACAAAACTAAATCTTACCAATGCTTAAAAATCTTAGG*GTTTATAGTTAAGTCT Cyanothece sp. ATCC 51142; cce_1775; 21.53
CGGGTAATCAGCTTTATGTTAACCCTGGTATTGGTAT*GAGTATCGCACCAATC Cyanothece sp. CCY0110; CY0110_30376; 25.53
CCCAAACTGGGTTTTACAAAAACTTTCGGATCTTAGA*TGATCTAAGCTTAAGA Cyanothece sp. PCC 8801; PCC8801_1848; 23.65
CCCAAACTGGGTTTTACAAAAACTTTCGGATCTTAGA*TGATCTAAGCTTAAGA Cyanothece sp. PCC 8802; Cyan8802_1874; 23.65
CAAGGCAAAACCATTACGCTAAAATAATTTTCAGAAAACATTATGATTAAATTC Cyanothece sp. PCC 7822; Cyan7822DRAFT_4871; 24.94
GCTGTAGCAGTAACTACAACTGTGGTCTAGTCAGCGG*TGTTACCAAAGAGTCG Synechococcus elongatus; syc1192_d; 17.29
ACCGTAACGACCGTTACGCCTTCCCCAGGACTCGCT**CGATGGAATGAGCCGG Synechococcus sp. RCC307; SynRCC307_1895; 20.71
TTGGTATCAACACGTACAACTCATAGACCGGAATTCAGGCCTAATAGGGGTGTG Synechococcus sp. CC9311; sync_2331; 18.24
CCCGTAACAACAGCTACGCCGAAGCCCTGCGCACTCAGGCCTAATGGCCGCAGT Synechococcus sp. WH 5701; WH5701_15331; 20.24
TCGGTATCAAAACGAACAACGATTGGACCAAAGACCAGTCGTAATGGGTGTAGA Synechococcus sp. WH 7805; WH7805_12458; 20.35
TCGGTAGCGAATCGAACAACGACTGGGCCTGGTACCAGCGGTAATGGGGACACA Synechococcus sp. WH 7803; SynWH7803_2035; 20.47
TCTGTAACAGCAATCACACCCCAAGGTGTGTTCCGAGCCATTCTGGATTCAGTT Synechococcus sp. BL107; BL107_16790; 19.65
TCTGTAACAGCGATGACACCCCTGGGTGTGTTCCG***AGCCATCCTGGATTCA Synechococcus sp. CC9902; Syncc9902_0470; 20.12
TCTGTAACAGCCGCCACGCACGGCTCCGAGAACCG***GGGCATCCTTGATACC Synechococcus sp. CC9605; Syncc9605_2206; 20.47
```

Nostocales:

```
AGAGCAGATACGGTTAAAAAAAAGTTGCAATTCTCATAAGTGTCTTTAAAATAAG Anabaena variabilis ATCC 29413; Ava_0137; 21.29
AGAGCAGATACGGTTAAAAAAAAGTTGCAATTCTCATAAGTGTCTTTAAAATAAG Nostoc sp. PCC 7120 (Anabaena sp.); all2319; 21.29
TCATACAGAGCAGATACGGTTAAAAAAAGTTG******CAATTCTCATAAGTGT -"-; all2319; 26.47
AAAGAATCTACACTTTAAAAAAGTGGTAATTA*TTACATATCTTTTAGGAAAAG Nostoc punctiforme; Npun_F4466; 22.71
CTGGTAGTTACAGCTTCAGACAAAGGATCTGG******TTAGATATATAAACGA -"-; Npun_F4466; 20.24
TTGGTAGGGCGCATACAGTTTAAAAAAAGTGG******CAATTCTTACATACTT Nodularia spumigena; N9414_06139; 29.53
AAGGTGATATAAAATACATAAGATATATTTCT****CATCAGTGTTAAATGAAT Raphidiopsis brookii; CRD_00313; 20.12
AAGGTGACATAAAATACATAAGATGGATTTCT****CACCAGTGTTAAATGAAT Cylindrospermopsis raciborskii; CRC_02206; 20
```

Oscillatoriales:

```
AGTGTATCATATGTTACTAATTTTTTTGATGACGCGG*TGATATCAAGCTTATG Lyngbya sp. PCC 8106; L8106_23605; 16.47
ATTGTATCACGATTTACCAATTGCGGGGCTTTGATAG*TGATATCAAATGGTTT Arthrospira maxima; AmaxDRAFT_1137; 18.35
ATTGTATCACGATTTACCAATTGCGGGGCTTTGATAG*TGATATCAAATGGTTT Arthrospira platensis; AplaP_010100021771; 18.35
ATTGTATCACGATTTACCAATTGCGGGGCTTTGATAG*TGATATCAAATGGTTT Arthrospira sp. PCC 8005; APCC8_010100026158; 18.35
GATGTAGCAAAACATACACCTACTTCACCAGT*AGCACGGGCATATCCATCAGC Trichodesmium erythraeum; Tery_2842; 17.76
```

Prochlorales:

TTTGTATCATTGGGTACAAAACTTATCTTCGCTATCCCCCC**TAATGG**ATATAAC Prochlorococcus marinus NATL2A; PMN2A_0993; 17.06
TTTGTATCATTGGGTACAAAACTTATCTTCGCTATCCCCCC**TAATGG**ATATAAC Prochlorococcus marinus NATL1A; NATL1_18621; 17.06
GCTGTACTCATTTTTAGTAGAACTTATTTACTAAACCTATA**TAAAAA**GATCAAA Prochlorococcus marinus MIT 9312; PMT9312_1556; 21.53
TTAGTAGAACTTATTTACTAAACCTATA***TAAAACTATA**TAAAAA**GACCAAA Prochlorococcus marinus MIT 9215; P9215_17311; 23.06
GCATTATCAAATAGTACGGTTCTTATTTCTGAAGAAAATTC**CACATT**GTTAATT -"-; P9215_17311; 21.41
TATGTATCATCAATTACAAACATTATTCTCGCTAGAGGCGC**TAATAG**GTGCAAC Prochlorococcus marinus CCMP1375; Pro1616; 15.53
TTTGTATCAACAGTTACAGTTACAAGCTAATTACAAAGGCT**TATATA**TATCTAG Prochlorococcus marinus MIT 9515; P9515_16421; 19.06
TTTTTATTCAAAAATACCTTTACTAAAACTATATAAAAAGC**TAATAA**TTCGCAA Prochlorococcus marinus CCMP1986; PMM1463; 21.88
TTTGTATCAATGATAACAATCTTAGTATCGTTAAAAGGGTT**TATATG**TATTTAG -"-; PMM1463; 16.82

Gloeobacterales:

ATGGTAGTCTCACCAACACCACCGGCTGGCATCGGG**TGC**TAGTTAGTGGTCT Gloeobacter violaceus; gvip021; 22.35

Genus *Acaryochloris*, unclassified order:

TAAGTAGCCTGGCCCACAAA**CTTCCAATTTGGAA**AAT**GATGAT**ATTTCCA Acaryochloris marina; AM1_3158; 20.94

***icd*,** isocitrate dehydrogenase

Chroococcales:

TTCGTAACAGCCAATGCAATCAGAGCCTCCAGAAAG*GAT**TATGAT**CTGCTCCG Synechocystis sp. PCC 6803; slr1289; 20.59
GACGTAACTGG*ACTCGAACCAGTGACCCCATCGAT*GTC**AACGAT**GTACTCTA Cyanothece sp. CCY0110; CY0110_04106; 27.06
TCTGTAGCGGTATTTACCGAACAAAAACTGCTAT***GAA**TAAATA**GCGACTGT Cyanothece sp. PCC 7424; PCC7424_0313; 18.94
GGTGTTTTAGGGATTACAAAAAGTTGACCTTCATGC*CAG**TATATA**AAGCAGTC Cyanothece sp. PCC 8801; PCC8801_3969; 19.65
ACTCTAAAATGTTGTACGATAACATTAAATGGTACTAATA**TATTAT**ACATTTAT Cyanothece sp. PCC 8802; Cyan8802_4014; 22.24
CAAGTAAATGCCGATACTTTAATTTCTCTTGGCAA**CGA**TGAGTTAGCTACAC Crocosphaera watsonii; CwatDRAFT_3583; 22.24
AATGTATATGGTTTTACCTGTGGCTGTAACATTTA**TTA**TATGTTTACTATGG -"-; CwatDRAFT_3583; 20.71

Oscillatoriales:

CTAGTAGTCTATCTTACTAAAATTCGCCATAATTACAAAC**TTAACT**AGAGAAAT Trichodesmium erythraeum; Tery_0071; 19.41
ACCGAATCTTGTTATACAAAATTCACTAATTTTA***AGA**TAATTT**TCAATAAA Lyngbya sp. PCC 8106; L8106_03474; 20.24
GATCTACCTCCACTTACAAATACTACACCGATCG***CCC**TAATTC**TGAAGTGA Arthrospira maxima; AmaxDRAFT_4514; 20
ACGGTCGAAAGCGGTACTACATTTCCGGCTCAATT**GGG**TAGAAT**GGGGTTTA Oscillatoria sp. PCC 6506; OSCI_4110009; 22.47

Nostocales:

TCAGTAAAAATTCCTGCAACACCTCTAATCCCT****CAT**TACCCA**AGTCAGCT Nostoc azollae (Anabaena azollae); AazoDRAFT_0802; 19.88
GACATAGCAAAATATACAAGCCTGCCACATAAGGT**TTT**TTGAAT**AGCAGCGT Nostoc punctiforme; Npun_R5474; 19.76
ACTATTTTTGATCGTACAAAATTGCGGCGGTTTCACTGAA**AATCCT**ATATCAGA Anabaena variabilis ATCC 29413; Ava_4831; 21.18
GCCGAAGGCAATCGTACAAAATTGCAATCGCTTCACTGAA**TTCAAT**AGCTTAAC Nodularia spumigena; N9414_14022; 21.06
ACTATTTTTGATCGTACAAAATTGCGGCGGTTTCACTGAA**AATCCT**ATACCGAA Nostoc sp. PCC 7120 (Anabaena sp.); alr1827; 21.18

Prochlorales:

TCCCCAATAGTAGCTACTCTCCATCTTTCAGCAGG**AAT**TATCAT**TCCGCGCA Prochlorococcus marinus CCMP1986; PMM1596; 24.59
AACTCATCACAAATTACAATACGAAGAACTTTAATG*AGT**TATCTT**TGTTCTTA Prochlorococcus marinus MIT 9312; PMT9312_1688; 22.12
AACTCATCGCAAATTACAATACGAGGAACTTCAATG*AGT**TATATT**TTATTAAT Prochlorococcus marinus MIT 9215; P9215_18691; 23.06
AACTCATCGCAAATTACAATACGAGGAACTTTAATG*AGT**TATCTT**TTATTAAT Prochlorococcus marinus AS9601; A9601_18051; 23.06
AACTCATCGCAAATTACAATACGAGGAACTTCAATG*AGT**TATCTT**TTATTAAT Prochlorococcus marinus MIT 9301; P9301_17881; 23.06
TTCAGACCAATTACTACCAAAGTAGAACGACAAATCTTGC**TCTAAT**TTTTTCTT Prochlorococcus marinus NATL1A; NATL1_20451; 21.65
TTCAGACCAATTACTACCAAAGTAGAACGACAAATCTTGC**TCTAAT**TTTTTCTT Prochlorococcus marinus NATL2A; PMN2A_1170; 21.65

***amt1*/*amtB*,** ammonium/methylammonium permease

Chroococcales:

```
AAAGTAGTAAATCATACAGAAAACAATCATGTAAAAA*****TTGAATACTCTAA Synechocystis sp. PCC 6803; sll0108; 16.24
AATGTATCAACGAATACAGAAATCTCAGATCAAAAAA***ACAAAGATCGAAAAA Microcystis aeruginosa; MAE_40010; 16.12
GATGTATAACGTGATACAGAACTGCTTAAGTCAGACC**TGTTGGAATCTAAAAA Cyanothece sp. PCC 7425; Cyan7425_0782; 17.88
AAAGTATCTTATGTTACAGAAAACAAACTT*AAAAAG**CATAAGAATTTTTAAG Cyanothece sp. PCC 7424; PCC7424_0499; 16.94
AATGTATTTTTTGATACAGAAATTTTCCCTAAAAAAG**AGTAATAAAAAAACA Cyanothece sp. PCC 8801; PCC8801_1229; 16.82
AATGTATTTTTTGATACAGAAATTTTCCCTAAAAAAG**AGTAATAAAAAAACA Cyanothece sp. PCC 8802; Cyan8802_1259; 16.82
ACAGTATGATTTGTTACAGAAAAAAGCTTTAAAAAC***CATAAAAATCTTGGCG Cyanothece sp. PCC 7822; Cyan7822DRAFT_4819; 17.65
ACTTTTATGCAAGCTACAGAATTATCTCCTAATGAAG**CTCAAGATTATTTAAA Cyanothece sp. CCY0110; CY0110_08851; 22
AAGGGTTTTTGTTTTACCGAAACGAGTCTGTTGTTTA*ACGGCAAAGTTAAAGTC Cyanothece sp. PCC 7822; Cyan7822DRAFT_2270; 23.88
CTCTTTTAGAGTTTATACACTAAGTAAGGAAATGAC****TGCTAAAAAAACCTTT Cyanothece sp. ATCC 51142; cce_3261; 24.94
AAAGTAGCAAAAGTTACGTATATCACCAGTCTGCCT***AGCCAGAGTTGTGAGA Synechococcus elongatus; syc1821_c; 17.53
AGCGTATCATTTTATACAGAATGTTAGTGGTAAACA***TTTTAAAAACCAGGGT Synechococcus sp. PCC 7002; SYNPCC7002_A2208; 17.88
GGTGTACCTGCAAGTACATCTTAGGTCAGTGAGACGG*TCATCACGACAGTCCGA Synechococcus sp. WH 7805; WH7805_07696; 20
AATGTATAACGTGATACAGATTTTCCAATCTGAGAAC**TGGTGCAATCAAAAAC Thermosynechococcus elongatus; tll1985; 17.41
CTGGTATTATAAACTACAGGCTGCTGAATCAATTTT***TCCTATGGCCACTAAT Crocosphaera watsonii; CwatDRAFT_6540; 19.76
```

Oscillatoriales:

```
AAAGTAGCATTTGATACAGAAAGTTAAGACAATTAA***CAGTGCGATATTTTTT Trichodesmium erythraeum; Tery_4477; 16.35
AGTGTAGTCTGAAATACAGAAAATTACTCTAATCTGT**TTTCAAGATCATCAAC Lyngbya sp. PCC 8106; L8106_22019; 18.47
TCTGTAGTTTGTGATACAAACTTTCGGTCGTTCTAC***CGCTTTAATCGAGCCA -"-; L8106_22019; 16.59
TCTGTATTGCAGTATACAGAAATCTGATTAA**********GATAATGTTAACA Arthrospira maxima; AmaxDRAFT_4502; 18.82
TCTGTATCATGGAATACAGAATTTAGGTTAA**********TATAGTGTTAAGA Arthrospira platensis; AplaP_010100009550; 17.88
TCTGTATTGCAGTATACAGAAATCTGATTAA**********GATAATGTTAACA Arthrospira sp. PCC 8005; APCC8_010100023060; 18.82
AATGTATATTTGGATACAGAAGGTAAGGAACAAATAT**TGATGGAATCTTTGAA Oscillatoria sp. PCC 6506; OSCI_2480010; 18.47
```

Nostocales:

```
-TAGTATCTGATCATACAGAA***TTACAGCTTGAGTAAATAGAAAATCAATCAA Nodularia spumigena; N9414_16726; 17.88
AATGTATTAACCAATACAGAA***TTAATGTTTAGGTAAATAGAAAATCAATCAA Anabaena variabilis ATCC 29413; Ava_2954; 16.24
AATGTATTAACTAATACAGAA***TTAATGTTTAGGTAAATAGACAATCAATCAA Nostoc sp. PCC 7120 (Anabaena sp.); alr0990; 15.18
ATAGTATCAAACTCTACGTAA****CTTCAATTAAGTA*ATGTAAATTAACTTGT Nostoc azollae (Anabaena azollae); AazoDRAFT_0734; 19.06
AATGTAGTTAAGGATACAAATCGTTGGTGCTTAGGAAG*AAAAAGAAGTGTTGAA Nostoc punctiforme; Npun_R3288; 16
ATTGTGATTGAACATACAGAACTGAACTATAGATAACAGTTTGATAATTTACAGT Cylindrospermopsis raciborskii; CRC_00109; 20.12
ATTGTGATTGAACATACAGAACTGAACCATAGATAACAGTTTGATAATTTACAGT Raphidiopsis brookii; CRD_01628; 20.12
```

Gloeobacterales:

```
CATGTAGTCATTGATACAAATGAGTCGCGACAGATTCG*CTTAACTCTCCTTATT Gloeobacter violaceus; glr3061; 16.35
```

Genus *Acaryochloris*, unclassified order:

```
AATGTATCTAAAAATACAGATTTACAGTACCTAGCA***CTGTTGAATCGATTAG Acaryochloris marina; AM1_3533; 16.47
```

Prochlorales:

```
TATGTATTAATGATTACCAGATAGCAACTCTATTTT*ACCCATATTTGTATTTGT Prochlorococcus marinus CCMP1375; Pro0295; 18
TTCGTATAAACGATTACCAGAAGGATTACAAGATTT*CTCCATATTTAGAATTGT Prochlorococcus marinus MIT 9211; P9211_02901; 19.53
TTAGTACTTTGCATTACCAGAAAAGACGCTTGTTTC**TTTGCCCCTGGAGATGA Prochlorococcus marinus MIT 9215; P9215_02871; 21.18
TTGGTAGCAACGACTACTAGACACTGATAAGTTTTTT**GAGCAGAATTGTATAT -"-; P9215_02871; 18.71
TTGGTATCAACAACTACTAGACACTGATAAGTTTTTT**GAATAAATTTAAATTT Prochlorococcus marinus MIT 9301; P9301_02861; 18.47
```

TTGGTAACAACGACTACTAGACACTAATAAGTTTTTT**GAG**CAAAAT**TAAATAT Prochlorococcus marinus AS9601; A9601_02851; 19.06
TTGGTAGCAACGACTACTAGACACTAATAAGTTTTTT**GAG**CAAAAT**TGAATAT Prochlorococcus marinus MIT 9312; PMT9312_0265; 18.71
TTGGTATAAACGACTACTAGACATACATAAGTTTTCT**GAG**CAAAAT**TATATAT Prochlorococcus marinus CCMP1986; PMM0263; 19.65
TTGGTATAAATGAATACTCGACATAGACA*GTTTTCA**GAG**CAAAAT**TATGAAT Prochlorococcus marinus MIT 9515; P9515_02961; 20.94
TTTGTATCAACACTTACATCTTGTAAGTATATGTTGC**ATA**CATATT**TAATATT Prochlorococcus marinus NATL2A; PMN2A_1629; 17.65
TTTGTATCAACACTGACATCTTGTAAGTATAAGTTGC**GTA**CATATT**TAATATT Prochlorococcus marinus NATL1A; NATL1_03411; 19.65
GAAGGCCTTAGGCCTACCAAGACGACACAGAGTTGAC*ACAG**CAAAAA**GAGCTCC Prochlorococcus marinus MIT 9313; PMT1853; 24.35
GAGGGCCTTAGGCCTACCAAGACGACACAGAGTTGAC*ACAG**CAAAAA**GAGCTCC Prochlorococcus marinus MIT 9303; P9303_24801; 24.35

apcF, phycobilisome core component/allophycocyanin
Chroococcales:
TTTGTATACCATGATACGGATTTTCTCGATCAAAC**CTG**TAAAGA**TCACTAAA Cyanothece sp. PCC 8801; PCC8801_0807; 19.29
TTTGTATACCATGATACGGATTTTCTCGATCAAAC**CTG**TAAAGA**TCACTAAA Cyanothece sp. PCC 8802; Cyan8802_0836; 19.29
TTTGTATCATGAGTTACAATTTTATTTTTATTCCCCGGTTT**TATATT**ATCTTTT Cyanothece sp. PCC 7425; Cyan7425_1838; 17.65
GCGATAAAAGGTCATACCTTAAATATCGTCGGCTAA*ATG**TATTTT**ACTTTACT Cyanothece sp. PCC 7822; Cyan7822DRAFT_1560; 24.35
AGTGTGATAAGAAATACACTACAGGAAAATCGGTAAAATC**TATAAT**TTTGATTA Cyanothece sp. CCY0110; CY0110_24111; 20.47
CGTGTGATAAGAAATACACTACAGGAAAATTGCTAAAATC**TATAAT**TTTGATTA Cyanothece sp. ATCC 51142; cce_1944; 21.29
AAGGAAAAACCATGTACTTCAGTCATGGGAGTATGT*CAA**TATAAA**ACGAAGCT Synechococcus sp. PCC 7002; SYNPCC7002_A1631; 23.76
CGTGTAGTGTCAAGTGCTGCCAAATCCCCGGCATGT*CGA**TAAATT**TCAAATCG Synechococcus elongatus; syc1936_c; 22.94
TTTGTATCCGTCAGTACATTTTTTCTCTTGT******CAC**TATTCA**TGAGGGCT Synechococcus sp. WH 7805; WH7805_02272; 20.35
TTCGTATAAGATGTAACTATTAGCGCATCTA******GTT**TAAATT**ATTGGGAG Crocosphaera watsonii; CwatDRAFT_4305; 21.18
Nostocales:
TTTGTAGTCTTTGTTACAGAACGTCTGGATTAC****AGG**TAATATTAA**CCTTT Nostoc sp. PCC 7120 (Anabaena sp.); all2327; 17.06
TTTGTAGTCTTTGTTACAGAACGTCTGGATTAC****AGG**TAATATTAA**CCTTT Anabaena variabilis ATCC 29413; Ava_0146; 17.06
GGGGTTGTCATTGTTACTCCTTGTGTACCCAATT***TCC**TAAAGT**AAAACCAC Nostoc azollae (Anabaena azollae); AazoDRAFT_1772; 22.35
AAGGTTGTCTGGATTACAGGTTATATTAACCTTTCTTTCT**TAATCT**GCACAAAA -"-; AazoDRAFT_1772; 22.59
TTTGTAGCTTATGTTACAGATATCTGGATTATCA***GGT**TATATT**AACCTTTC Nostoc punctiforme; Npun_F5388; 16.35
TTTGTACTTTTTGCTACAGCTTATCTGGATTACA***GGT**TATATT**AACCTTTT Nodularia spumigena; N9414_03583; 18.47
GTTGTATTGTTTGTTACGATTGTTCAAGTTTTTG***GGT**TATGTT**AAGTTTTT Raphidiopsis brookii; CRD_02137; 19.41
GTTGTATTGTTTGTTACGATTGTTCAAGTTTCTG***GGT**TATGTT**AAGTTTTT Cylindrospermopsis raciborskii; CRC_02161; 19.41
Oscillatoriales:
TTTGTATCAAGCATTACACTATAACACGACATGGCT*CAG**AATAAT**TGAAAGAA Lyngbya sp. PCC 8106; L8106_18731; 18.47
TTTGTATTGAGTGCTACACTATGACGCAACCTGGAT*CAA**AATAAT**TTAAAGAT Arthrospira maxima; AmaxDRAFT_3348; 18.47
TTTGTATTGAGTGCTACACTATGACGCAACCTGGAT*CAA**AATAAT**TGAAAGAT Arthrospira platensis; AplaP_010100011061; 18.47
TTTGTATTGAGTGCTACACTATGACGCAACCTGGAT*CAA**AATAAT**TTAAAGAT Arthrospira sp. PCC 8005; APCC8_010100012675; 18.47

psaI, photosystem I reaction center subunit VIII
Prochlorales:
CCAGTAATAAAAGTTACACTATTTGTCCTTTAAA***TAC**TAAAGT**GTTCAAAG Prochlorococcus marinus CCMP1986; PMM1520; 18.82
TTCGTATTGGATTGAACTTTTATGTCACCAATGGGGTCTT**TACGTT**TCTATTAC Prochlorococcus marinus CCMP1375; Pro1678; 23.65
CAAGTAATAAAATTAACACTATTTCTTAGAAAAA***TAC**TAGAGT**ATTTATAG Prochlorococcus marinus MIT 9215; P9215_17881; 21.06
CCAGTAATAAAATTAACACTATTTCGCAGATAAA***TAC**TAAAGT**GTTTAAAG Prochlorococcus marinus AS9601; A9601_17241; 21.29
CCAGTAATAAAATTAACACTATTTCGCAGATAAA***TAC**TAAAGT**GTTTAAAG Prochlorococcus marinus MIT 9301; P9301_17121; 21.29

CCAGTAATAAAATTAACACTATTTATCAGATAAA***TAC**TAAAGT**GTTTAAAG Prochlorococcus marinus MIT 9312; PMT9312_1612; 21.29
CTTGTAAGAAACACTATTCTTTTTAGATATATTTT**TAT**TATTAT**TATTTATA -"-; PMT9312_1612; 22.94

petF*/*fdx, ferredoxin, 2Fe-2S type

Chroococcales:

AATGTATATACAAATACTCCCTCGACCCGAGACATCGAAT**TATGGA**CTTACTTT Synechococcus sp. WH 8102; SYNW1274; 19.41
AAGATATAAACAAATACTGCCTGGCTTGTAGATCT**AAA**TATTGT**GTCGTTTT Synechococcus sp. BL107; BL107_12815; 21.53
CATGCAGCAGGCGGTACGGGTGTCGGACTACACCGGCTTC**TTCAAT**GCCGAAGC Synechococcus sp. WH 5701; WH5701_13215; 22.82
TCAGTAACCAAGTGGACTATGAGCAATTGCAGACCCTTTCCT**TAAAAT**TAGAAG Synechococcus sp. PCC 7002; SYNPCC7002_A0624; 22.71
TTGGTGCTAAAAAATACTAGCTAATCAGAAC******CGT**AATAAT**CCCTAAAC Cyanothece sp. ATCC 51142; cce_1665; 21.18
TTTGTAAAATTTTGTAAAGTGATATTGACA*******AAA**TATTTT**TAATGTCA -"-; cce_1665; 21.18
TAAATAGTGTTATTTACACAAAATAAAGTTTGT****AAT**GATAAT**TTTGCTGT Cyanothece sp. PCC 7424; PCC7424_0512; 20.59
AAAATATTAATATTTACAGATGATAACGAATTATTA*CAT**TATCTT**ATAAGGAT Cyanothece sp. PCC 7822; Cyan7822DRAFT_5105; 18.94
TATGTAAGTATAAATATACTGTAATTTGTTTGTTGGTTGA**AATAAT**ACAAAAAC -"-; Cyan7822DRAFT_5105; 21.18
TTTGTAAAGATGAAGACTATTTGAGCAATTATTGA**TGA**TATAAT**CAAGGTGG Microcystis aeruginosa; MAE_41230; 22.12

Prochlorales:

CCTGTCAAAAAGATTACCATCTACTTTAAA*******AGA**TATTTT**AAATACAA Prochlorococcus marinus MIT 9301; P9301_09611; 21.65
TTGGTTTCAAAATATACTTTGAAAAATTTTGA*****TTT**TATTTT**GTGTAATA Prochlorococcus marinus CCMP1986; PMM1352; 22.12
--CGTCCCAATCCCTACCTACGAATCTTTTTA*****AGT**TATAAA**ATGTATTT Prochlorococcus marinus MIT 9515_1; P9515_09801; 21.65
ATTGTTTCATTTTTTACGCCAAAAACTCACAAAT***GGT**TATGAT**CAATGTTA Prochlorococcus marinus MIT 9515_2; P9515_15131; 21.06

Oscillatoriales:

AATTTAACCGCAATTACACTATAATAGCAGAAACA**ACT**TAAATT**TAGCAACC Trichodesmium erythraeum; Tery_0754; 20.71
CGAGTATAAGTATATAAGCTACCTTTTAG********TCA**TATCAT**CAATAACT -"-; Tery_0754; 24.12

Nostocales:

TAACTAAATTTCATTACTAATCGTGGTGGATTTCAAGTTA**TATACT**TGGATTTT Nostoc sp. PCC 7120 (Anabaena sp.); all4148; 21.65
CCCATAGGGCGGTCTACCCTTCGGGAACTCCGCAG**GAG**TACAAT**TAGTGGTG Anabaena variabilis ATCC 29413; Ava_0981; 26.59
TTGGTACGAAATACTAGTAATAAACCTAA********AAC**TATACT**GGGAATAA Nostoc azollae (Anabaena azollae); AazoDRAFT_4699; 20.35

som, porin

Chroococcales:

TATGTATCAGGCGAAACATTTATTGTGG*********TTT**TATGTT**CAGTTAGC Synechococcus sp. CC9902_1; Syncc9902_0323; 20.94
TCAGTAGTGGGTTATACCGTTCGGGTCTTC*******CAA**TATTCA**CAGGTGGC Synechococcus sp. CC9902_2; Syncc9902_0324; 20.82
ACGGTAGTCAGGTGTCCTCATGGCACGG*********TCC**TATAAT**ACAGTCCA Thermosynechococcocus elongatus; tlr1246; 23.41
TAGATATATTGCCATACTTGTGATTTGTATAGAAC**CAA**TATTAG**ACGATTTT Cyanothece sp. PCC 7425; Cyan7425_2214; 25.29
CATGTATCATTGAATACCATCGACAGTA*********AAG**TATTGC**GTCTTTAG -"-; Cyan7425_2214; 19.65

Genus *Acaryochloris*, unclassified order:

TAGTTAGAGAAAGTTACAAGTTTAATTATTTCTGAATTTG**TATTGA**ATTTTACA Acaryochloris marina; AM1_0866; 20.35
TTTGTATTGAATTTTACAAAATAAGCACCCACCA**TAACGT**AGT**TATATC**TGGGG -"-; AM1_0866; 16

Prochlorales:

TATGTAGCAGTTAATACTTGATTTCGTATACCCAAATACAC**TATCTT**GAGGTGT Prochlorococcus marinus AS9601; A9601_12271; 18.35
TTCGTATACCCAAATACACTATCTTGAGGTGTA****CTT**TAAATT**TCGTGTGA -"-; A9601_12271; 21.06
TTAGTAATTTGTACTACATGATAGAAAAATG******AGG**TATTTT**TTTTATCT Prochlorococcus marinus CCMP1986_1; PMM1121; 19.76

```
AAAGTATCGTTTGATACTTGTGATTATACTC******AGTTATACAGTATCTTA -"-; PMM1121; 20.12
TGAGCAGTTTTTTGTACCCAATGCCTAT*********CCTTATGATTGATCTTA Prochlorococcus marinus MIT 9313_1; PMT0802; 21.88
CTAGTAGTCAGTGCTCCTCAGTCTCAA**********TACTATCTTGCCGAATT -"-; PMT0802; 21.76
AATGTAGTCGTTGATACTTGATTTAGTATATCGAAGTACATTATCTTAAAGTGT Prochlorococcus marinus MIT 9215_1; P9215_12571; 19.41
TTAGTATATCGAAGTACATTATCTTAAAGTGTC****CTTTAAATTTCGTGTGA -"-; P9215_12571; 21.65
TTACTAGGCATGTTTACCAAATTATGACTTGCGAAA*ATCTATCCTCTACTTAG Prochlorococcus marinus MIT 9211; P9211_08291; 21.76
TATTTAATTATATGTACTGTCCTGTTATTAAT*****AAATATTTTTAGTAATTC Prochlorococcus marinus MIT 9215_2; P9215_12551; 22.12
AATGTAGTCTTTGATACTTGATTTTGTATGCCCAAATACACTATCTTGAGGTGT Prochlorococcus marinus MIT 9312_1; PMT9312_1131; 19.53
TTTGTATGCCCAAATACACTATCTTGAGGTGTA****CTCTAAAATTTCGTGTG -"-; PMT9312_1131; 20.24
AAAGTATCAATGGTTACATCTCCTTGAAACTT*****TTATATCATTTCAAGAT Prochlorococcus marinus MIT 9312_2; PMT9312_1264; 18.12
AATGTAGCGGTTGCTACTTGATTTCGTATGGGTTAATACACTATCTTAAGGTGT Prochlorococcus marinus MIT 9301_1; P9301_12271; 19.18
TTCGTATGGGTTAATACACTATCTTAAGGTGTA****CTTTAAATTTCGTGTGA -"-; P9301_12271; 20.59
CTTGTATAAAAAGTTACAACGCCTTGAA*********ACTTATATTTCTTTCCT Prochlorococcus marinus MIT 9301_2; P9301_13641; 17.53
GGTGTAACATAAGCTACTTTTAGCCTCCTCTTGACATTGTTAAATTTGATACAA Prochlorococcus marinus NATL1A_1; NATL1_04791; 20
TTTGTAATTTTTTATACAAAAGATATAGATAA*****AAATATAAACTTTTTCT Prochlorococcus marinus NATL1A_2; NATL1_19301; 16.94
GGTGTAACATAAGATACTTTTACTCTCCTCAACGAAACGTTAAATTCGATACAA Prochlorococcus marinus NATL2A; PMN2A_1757; 20
CTTGCAACCATTGATACATCTGTCACACCAAGTG***GTGTTTAATCAGAAGTT Prochlorococcus marinus MIT 9303_1; P9303_26321; 20.82
TGAGCAGTTTTTTGTACCCAATTCCTAT*********CCCTATGATTGATCAAT Prochlorococcus marinus MIT 9303_2; P9303_14071; 21.88
ACTGTATACGGGACTACTGCTTTATTCTCTTAAGTA*ATCGATAATTGAATTGT Prochlorococcus marinus CCMP1375; Pro0737; 21.65
AATGTAACTGATTATACTTTTTTATGTAGC*******TTTTATGGTACCATTTG Prochlorococcus marinus MIT 9515_1; P9515_12121; 19.65
TCCGTATCATTTGCTACATAGAAGTGAATCTATTAA*TCCTACTTAAAAATGTT Prochlorococcus marinus MIT 9515_2; P9515_13341; 17.65
```

Nostocales:

```
TTAGTAGCAAAAAGTACTTTTTACAGTTAATA*****ATGTATATTAACTTATT Anabaena variabilis ATCC 29413_1; Ava_4059; 19.53
CTCCTATCTTGCCATACAGTCGAAATG**********CTCTATAATAATCAGCT Anabaena variabilis ATCC 29413_2; Ava_2510; 24.24
CTTGTAATGTGAGATAAAATTCTTAAACTCCTGT****TAATATTTGGAAAAGAC Nodularia spumigena; N9414_09004; 22.82
ATTGTACATTGTTATATTATAATAATTCTCATTC***TCATATACTGAAGACAA Nostoc sp. PCC 7120 (Anabaena sp.); alr0834; 22.71
ATACTAAAAAACATTACAAGGAAATATTATAAT****AATTATCATTATCATGC Raphidiopsis brookii; CRD_01401; 20.12
AAAGTGTTTGGATTTACCCATTAGTGGGTAAGGACGCTAATTATAAAGAGGTGT -"-; CRD_01401; 22.12
TTGGTACATTAACCTGCGATCGCTCACAA********CCATATATTAAAGAAAT Cylindrospermopsis raciborskii; CRC_01524; 24
```

Oscillatoriales:

```
TATGTACAAGCTATTACTTTTGTAGTTCTCAGCTTGAA**TATACCTAAAATAT Trichodesmium erythraeum; Tery_1234; 19.41
```

urtA, putative urea ABC transporter

Chroococcales:

```
TCGGTTCCGGTTGATACCAAAGCGGTGGGGGGCCCTTTTT******TACCTTCC Synechococcus sp. WH 8102; SYNW2442; 21.06
TTAGTTCGTTTTGATACGAAAGAAGGGTGCTGTTGATCCG******TACCTTCC Synechococcus sp. CC9902; Syncc9902_2250; 21.53
TTAGTTCGTTTTGATACGAAAGAAGGATGCTGTTGATCCG******TACCTTCC Synechococcus sp. BL107; BL107_06989; 21.53
ACGGTAGCTGTCATTACAAGAAGGGCATTTGTGATCCG********TATTGTCG Synechococcus sp. RCC307; SynRCC307_2454; 19.06
TTGGTAGTCAGCGCTACTAGTGAAGGTTTGCCTCGCTGAGAA****TGTCATGA Synechococcus sp. WH 7805; WH7805_09844; 20.59
AAGGTGCGATTGAATACATTTCTGCAGTTCTCGCCCTCCA******TACCTTTT Synechococcus sp. CC9605; Syncc9605_2620; 23.18
ACGGTATCCTATGCTACATAATTTTTTCTTAACCTACTCG******TAGTATGG Synechocystis sp. 6803; slr0447; 18.47
AAAGTAGCGAAAGTTACAAAAAATCTGCTAACCTAGCCG*******TAATATCA Microcystis aeruginosa; MAE_06220; 16.12
TTTGTATCTAAAGTTACAATTCCCAGTCGGGTCTAGCCT**TAACATACAAACT Cyanothece sp. PCC 7424; PCC7424_2862; 16.94
AGTGTATAATAGATCACATCTTTAAGGAGGGGGGAG*******TTTTATCATCT Cyanothece sp. PCC 7425; Cyan7425_2260; 21.18
```

ACTGTATCCAATGTTACAATTTTTCGGTTAACCTAGTTTTA**CCA**TACAAT**CT Cyanothece sp. PCC 7822; Cyan7822_1357; 16.71
CTTGTAGCCTAGCATACAGTTTGCCCCAATAACCTCCTCTTA*AGC**TATGGG**GA Thermosynechococcus elongatus; tlr1120; 20.59

Oscillatoriales:

TGGGTAGTAATTAATACCCTTGTCTGTTGTTTATCCAT*****TTT**TAACCT**AA Lyngbya sp. PCC 8106; L8106_19671; 19.29
TTTGTATCTAAAAATACAACTAAACACTTAACCTAGTCC****TAA**TATCAC**AC Oscillatoria sp. PCC 6506; OSCI_2210004; 16.94
ACTGTATTAATAGCTACATTTTTTGCGGTTGACTTA*******GCC**TAGGAT**AC Arthrospira sp. PCC 8005; APCC8_010100028378; 17.76
ACTGTATTAATAGCTACATTTTTTGGGGTTGACTTA*******GCC**TAGGAT**AC Arthrospira platensis; AplaP_010100001300; 17.76

Nostocales:

ATTGTAAATTTGTACACTTTTTTCTCGTAAAATCCCATAC***TAA**TATCTG**TT Cylindrospermopsis raciborskii; CRC_00451; 24
TGGGTATTACTCATTACAAATATTAATATTGTAA*********AGT**TATTTT**TT Nostoc punctiforme; Npun_F4615; 18.12
CTAGTAAACTATCGCACTATCATCACACTTTAAAGTT******TGA**TATTGT**CA -"-; Npun_F4615; 24.12
TTGGTATCCAAGATAACTTTTTACTAGTAAACTATCG******CAC**TATCAT**CA -"-; Npun_F4615; 23.18
TTAGTATCAAAAATAACAATTCAATGGTTAAATATCAAAC***TAA**TATCAC**AA Nostoc sp. 7120 (Anabaena sp.); all1951; 18.94
TTAGTATCAAAAATAACAATTCAGTGGTTAAATATCAAAC***TAA**TATCAC**AA Anabaena variabilis; Ava_4361; 18.94

Prochlorales:

ATGGTATAACCGGATACGGAAACTTATCTTCGTT*********TAC**TGAAAT**AG Prochlorococcus marinus NATL1A; NATL1_19191; 20.24
TTTGTATCATTCACTACAGAATCATGCCTGGCTGAT*******CTT**TACGTT**CC Prochlorococcus marinus MIT9313; PMT2229; 17.41
TTTGTATCATTCACTACAGAATCATGCCTGGCTGAT*******CTT**TACGTT**CC Prochlorococcus marinus MIT9303; P9303_29741; 17.41
AATGTTACCTATGCTACAAAACTAAATCCCACTC*********GTT**TATAAC**TT Prochlorococcus marinus CCMP1986; PMM0970; 18.24
CTTGTAGATTTATCTATTTACGTGATTGATGTATCGGCTGG**GGA**TAAAAT**TC -"-; PMM0970; 22.82
AAAGTTACGACCGATACAAAACGAATCCTCACTCGTTAATAACTTT**TACTTT**AT Prochlorococcus marinus AS9601; A9601_08881; 18.94
AAAGTTACGACAGATACAAAACTAATCCTCACTCGTTAATAACTTT**TACTTT**AT Prochlorococcus marinus MIT9215; P9215_09181; 18.47
TCATTAATAACTTTTACTTTATCCCCCTAATGAGGGTCAATT*ACC**TAATTT**AT Prochlorococcus marinus MIT9312; PMT9312_0829; 21.53
AAAGTTACGACTGATACAAAACTAATCCTCACTCGTTAATAACTTT**TATTTT**AT Prochlorococcus marinus MIT9301; P9301_08861; 17.88

cynA, cyanate ABC transporter substrate-binding protein

Chroococcales:

TGAGTATCACCTGATACAACATCCGCGTTCGCTTTCCAAC**TATAAA**TAAAGAAG Synechococcus sp. WH 8102; SYNW2487; 17.06
GTTGTAACGACGGCTACATTTTGACCCTGGGGTTAC*TAC**TACCAT**TCGCCCTT Synechococcus elongatus; syc1986_d; 19.29

speB, agmatinase

Chroococcales:

TCCGTAGCTGATCACACCGCGCGTGCCACCGGTGCC*GCA**GACAGT**GGAAGGGC Synechococcus sp. WH 8102; SYNW1412; 22.24
TTTGTAGCGGTTCACACTCCGCAGAGCATCAAAGCC*GCA**GACAGT**TAAATTGC Synechococcus sp. CC9605; Syncc9605_1082; 21.88
CGGGTATCGCCTGTTACCGCATTGGGGTCTATCGCAGCCG**GAGATT**GGATAGGC Synechococcus sp. RCC307; SynRCC307_2478; 21.06
CTGGTAGCAACCGTCACAACGCCACCGCTCTGATGG*GGT**GAAGGT**TCAGGGGC Synechococcus sp. WH 5701; WH5701_03684; 20
TATGTAGTCTGTGATACTGTTTTTGGGATTGAGATC*GTC**TAAGGT**CAAGGAAA Microcystis aeruginosa; MAE_18840; 19.76
AAAGAAATGGCAGCTACGGGCATATCCAGCACC****CGA**CATAAT**AGCCTAGT Synechocystis sp. PCC 6803; sll1077; 22.24
AAAGTAGCACAGAACCACAGTTAACTTAATCATCCCT*AGC**TAGGGT**AGGGGATA -"-; sll1077; 20.71
TTAGTAATTTTTGATACTGTTTGTAGAGAGCCAAAT*TGA**TAAACT**TTGCAACG Cyanothece sp. PCC 7822; Cyan7822_2040; 20.12

Oscillatoriales:

ACGGTAACTTTTAATACGGTTTGTCTTCGCAAGAAT*TGC**TATAAC**TGTAATAT Arthrospira sp. PCC 8005; APCC8_010100011683; 20.71
ATAGTATAGGGTTCTACATCTTTGGTAACTACAGCATGGGCGGC**TATTAT**CGCG Arthrospira platensis; AplaP_010100008946; 20.59

ATCGTATAGGGTTCTACATTCTTAGTTACTACAGAATGAGCCGC**TATTAT**TGCC Arthrospira maxima; AmaxDRAFT_4146; 21.29

Nostocales:

TTTGTAGTTTTATATACAGTTTATTACCATTAACA**TGT**TGT**TATGAT**CGCAACTA Cylindrospermopsis raciborskii; CRC_02094; 19.06

TTTGTAGTTTTATATACAGTTTATTACCATTAACA**TGT**TGT**TATTAT**CCCAGTTA Raphidiopsis brookii; CRD_02353; 19.06

Genus *Acaryochloris*, unclassified order:

TTTGTATCCTAAAATACAATTATTCATATCTAGACT*TGA**TATTTC**TTGTATAG Acaryochloris marina; AM1_2833; 18.24

***mutS*,** recombination and DNA strand exchange inhibitor protein

Chroococcales:

TCTGTTGATGGAATTACGATTCCGCTCGTATTGCCGT*CTG**TAACTC**TTGTCGT Synechococcus sp. WH 8102; SYNW2476; 22.12

GATGTAAAACTCTGTTCAGCATGCGGGCAGCTCCCGA*GCT**GATCAT**TTGCCAA –"–; SYNW2476; 22.47

TTGGTTGCTTTGGCTACATCTTGTGGCTGGCGTGGAG*ACG**TAGAAG**GCAGCTT Synechococcus sp. CC9902; Syncc9902_2283; 22.12

AATGTGACTTCAGATACAGCTTCGAAGCAGATGCAG**CCG**TAGAAGGCACCAT Synechococcus sp. WH 7803; SynWH7803_2490; 20

AGTGTTGCTTCGGCTACACCTTGGAAGCATTGGCAG**CCG**TAGAAGGCTTTCC Synechococcus sp. WH 7805; WH7805_09664; 20.71

Prochlorales:

GATGTATCAGTTTTTACCAAATTTTTTATTTATTT***TGG**TATCAT**TCAATAC Prochlorococcus marinus NATL1A; NATL1_21691; 17.18

TTGGTATCATTCAATACATCCTATCTATCCATTTA***ATT**GATTAA**TTGGTTT –"–; NATL1_21691; 19.76

GATGTATCAGTTTTTACCAAATTTTTTATTTATCT***TGG**TATCAT**TCAATAC Prochlorococcus marinus NATL2A; PMN2A_1297; 17.18

TTGGTATCATTCAATACATCCTATCTATCCATTTA***ATT**GATTAA**TTGGTTT –"–; PMN2A_1297; 19.76

***rnc*,** ribonuclease III

Chroococcales:

GATGTAGCGGCGACTACCGATATCGGCGCTC***CTGACGGGGC**TGGCGG**GACT Synechococcus sp. WH 8102; SYNW0152; 19.06

GATGTAGCGTTGGCTACCGGCAGAAAAGCTC***GTGACGGGGC**TGGCGG**GACT Synechococcus sp. CC9902; Syncc9902_0179; 20.47

GATGTAGCGCTCACTACTTCTTCAGTAATTC****TCACGGGGC**TGGCGG**GACT Synechococcus sp. CC9605; Syncc9605_0149; 21.06

GATGTAGCGTTGGCTACCGGCAGAAAAGCTC***GTGACGGGGC**TGGCGG**GACT Synechococcus sp. BL107; BL107_05764; 20.47

CCGGTAGCGTTCCCTACCACCTGGA*********TCGACGGGGC**TGGCGG**GACT Synechococcus sp. WH 5701; WH5701_03639; 22

AATGTAGTAGTGACTACCGGCGCCGCTGTTCAATCTCACGGGGC**TGGCGG**GACT Synechococcus sp. WH 7803; SynWH7803_0204; 19.18

AATGTAGTAGTGACTACCGGTGTCAGTCTT**TGTCAT**CACGGGGC**TGGCGG**GACT Synechococcus sp. WH 7805; WH7805_08251; 19.18

Prochlorales:

GATGTAGTACCGGCTACCGCGAACAGTTGGCAAAA*AACGGGGC**TGGCGG**GACT Prochlorococcus marinus MIT 9303; P9303_25991; 19.41

GATGTAGTACCGGCTACCGCGAACAGTTGGCAAAA*AACGGGGC**TGGCGG**GACT Prochlorococcus marinus MIT 9313; PMT1948; 19.41

ACTATATATAGGGCTGAAGATTGTGATAATTACCTT*ACGGGGC**TGGCGG**GACT Prochlorococcus marinus MIT 9211; P9211_17271; 24.47

TAGGTAGTGAACAGAACCCGGCTTATGACCTGCTTT***CACCC**TATTAA**TTAA –"–; P9211_17271; 23.06

CATGTAGCTTATGCAACTATTTTTAGCATTT******ACGGGGC**TGGCGG**GATT Prochlorococccus marinus CCMP1375; Pro1762; 20.71

GCAGTATCGAGAGGTACTGGCTTGGTAAACCCC****GGTTGGG**TGCAAG**GCAA –"–; Pro1762; 21.29

Oscillatoriales:

AAGGTCGGAGGAACTACGGTTGGTCTTTTACCAG****TTCCGC**TTGAAT**GCAC Oscillatoria sp. PCC 6506; OSCI_3820007; 23.29

ATCCTATTGTGAGATACTATGAATTAGAGAGAGT***CAGCGAA**GGCAAT**TGCG Lyngbya sp. PCC 8106; L8106_28616; 23.06

GTTGTTGGTGGAGATACTTCTGAGGTCAGAAAGC***GGGAAAT**TAGAAG**TCAG Trichodesmium erythraeum; Tery_1105; 23.53

GATGTTGTGAGCAATACGCATGGTGATCT********CAATAAC**TACAGT**TAAC Arthrospira maxima; AmaxDRAFT_5300; 21.76

GATGTTGTGAGCAATACGCATGGTGATCT********CAATAAC**TACAGT**TAAC Arthrospira sp. PCC 8005; APCC8_010100014955; 21.76

Nostocales:

CGTGTAAAGCAATTTACACAGAGAACAGAA*******CCCGGCT**TACCAC**CAAC Nostoc sp. PCC 7120 (Anabaena sp.); all4107; 20.12
CGTGTAAAGCAATTTACACAGAGAACAGAA*******CCCGGCT**TACTAC**CAAC Anabaena variabilis ATCC 29413; Ava_0797; 20.12
TCGGTAAACCCCGGTAGGGAGCAAGGCCAG*******AGGAAAC**TACGGT**TGGT Nodularia spumigena; N9414_22623; 23.88
GGTTTAGCTTCCGGTACAAATATACCAC*********CTATGTC**TATAAT**AAGC Raphidiopsis brookii; CRD_02283; 19.76
GGTTTAGCTTCTGGTACAAATATACCAC*********CTGTGTC**TATCAT**AAAC Cylindrospermopsis raciborskii; CRC_02811; 18.71

hupS, hydrogenase small subunit

Chroococcales:

TCTGTAAAATTTAATACATTTCATAAATTCTCT*TGT**CATGAT**TTCATGGTAAT Cyanothece sp. ATCC 51142; cce_1063; 18.71
TTTGTAGCTAATGCTACTTTTAGGAGGAATAAGGTTT**TATAAC**GTGAAATTTAT Cyanothece sp. PCC 7822; Cyan7822_1668; 18.47
TCTGTAAATTCTAATACACTTTTATAATCTTTG*TGT**CATGAT**TTCATGGTGAA Cyanothece sp. CCY0110; CY0110_26288; 18.94
TCTGTAACTGATGATACGTTTAATAATAATTAATTTT**TATAAG**GTAAACTTTAA Cyanothece sp. PCC 7424; PCC7424_1817; 19.53
ACAGTAAAAAATAAGACATTATTGCTAAATTTA*TGT**TATTAT**TTCATGACAAA Crocosphaera watsonii; CwatDRAFT_0515; 20.35

Nostocales:

TCTGTCTAAAAAACTACTATTAAAACAAA**********TATAAT**TTTGATTTAAC Anabaena variabilis; Ava_4596; 20.35
TCTGTCTAAAAAACTACTATTAAAACAAA**********TATAAT**TTTGATTTAAC Nostoc sp. PCC 7120 (Anabaena sp.); all0688; 20.35

Oscillatoriales:

ACTGTATTATCTGATACCGTTGCTTAGTTGAATTGCA**TCTCTG**TCAAGATAAGC Lyngbya sp. PCC 8106; L8106_01867; 18.35
AAAGTATCAAAAATAACAAATTTTTAGACAAA**AAT**AACAAT**TAATTAAAAAT Trichodesmium erythraeum; Tery_3369; 17.53

Genes with unknown function:

Chroococcales:

TCGGTAGTCGCCGCTACATCTGGTGGGGTGGGCAGACCGTCCTCCACCGC**TAGGGT**TT Synechococcus sp. WH 8102; SYNW0153; 20.12
CCCGTAGTCGACACTACATTTGGTGGGGTGGTCAGATCGTGCACCCCCGC**TAGCGT**TG Synechococcus sp. RS9917; RS9917_05010; 21.41
CCGGTAGTCACTACTACATTTGGTGGGGTGGTCATACGTCACGCTCCCGC**TAGCGT**TG Synechococcus sp. WH 7805; WH7805_08246; 19.65
CCGGTAGCCAACGCTACATCTGGTGGGGTGGCCAGTTCGCTCACCACCGC**TAGTGT**TA Synechococcus sp. CC9902; Syncc9902_0180; 20.35
CCGGTAGCCAACGCTACATCTGGTGGGGTGGCTAGATCGTTCACCACCGC**TAGTGT**TA Synechococcus sp. BL107; BL107_05754; 20.35
CCGGTAGTCACTACTACATTTAGTGGGGTGGCCAGACGGCACGCTCCCGC**TAGCGT**TG Synechococcus sp. WH 7803; SynWH7803_0205; 19.65
GTGGTAGGGAACGCTACCGGTGGTGGCCTTCCCAAACAGGGCTCCCTGGC**TAGCGT**GG Synechococcus sp. WH 5701; WH5701_03644; 21.76
ACTGATGGATTGTCTTGGATTGATGGGGTAGCCAGATGCCAGACCCCCGC**TATTTT**TG Synechococcus sp. CC9311; sync_0203; 28.24

Prochlorales:

GCGGTAGCCGGTACTACATCTGGTGGGCTAGGCAAATGCCACGCTCTCGC**TAGTGT**TG Prochlorococcus marinus MIT 9313; PMT1948; 20.35
GCGGTAGCCGGTACTACATCTGGTGGGCTAGGCAAATGCCACGCTCTCGC**TAGTGT**TA Prochlorococcus marinus MIT 9303; P9303_25971; 20.35
CCAGTACCTCTCGATACTGCTGGTGCGCTCTTACCGCACCTTTGCA*CCC**TTGCCT**GT Prochlorococcus marinus CCMP1375; Pro1761; 22.24
AAAATATGTAGTCCTATATGTGGTGGGTGTGCAAAATCATTACTCTCTCGC**TAGCGT**T Prochlorococcus marinus NATL2A; PMN2A_1178; 24.12

HNH endonuclease family protein

Chroococcales:

TTGGTGCTCAGCGTTACCAAGGGGGGCTTGGTGGAATCGA**TAGCAT**G Synechococcus sp. WH 8102; SYNW2097; 22.12
ACTGTGCTCGGGGCTACCGAGACAAACGTGGTGTGCCCAA**TAGCAT**G Synechococcus sp. CC9902; Syncc9902_1985; 22.47

ACTGTGCTCGGGGCTACCGAGACAAACGTGGTGTGCCCAA**TAGCAT**G Synechococcus sp. BL107; BL107_08459; 22.47
GCAGTGCTCGGCGCTACCGAGACAGCTTCACATCGC*CGA**TAGCAT**G Synechococcus sp. CC9605; Syncc9605_0346; 23.53
AGGGTAGTGAAGATCACAGTGATGTTCCTGTGATTCTCGA**TAGCAT**G Synechococcus sp. CC9311; sync_0407; 21.53
TTGGTAGTGGGGGGCACAAGGCGTCATTTGTGACCC*CGT**TAGCAT**G Synechococcus sp. WH 7805; WH7805_07131; 22.35
GAGGTAGTGAGGGTCACAAAGCGTTGGTTGTGACCC*CGA**TAGCAT**G Synechococcus sp. WH 7803; SynWH7803_0404; 20.35

Prochlorales:

AATGTAATAAGGGAAACTGCAACAAGGAGTTAATAC*CAC**TAGCAT**G Prochlorococcus marinus NATL1A; NATL1_19741; 21.18
AATGTAATAAGGGAAACTGCAACAAGGAGTCAATAC*CAC**TAGCAT**G Prochlorococcus marinus NATL2A; PMN2A_1099; 21.18
CCAGTATCAACTTTTAGATGCAACGAGAATTTTAAATTAT**TACTCT**T Prochlorococcus marinus CCMP1986; PMM1528; 20.35
TTAGTAATCACGTGAACAGCCGCTTCCGTTTCATGC*CAA**TAGCAT**G Prochlorococcus marinus MIT 9211; P9211_16481; 22

Nostocales:

AAAGTAAGAAGTTTTACTTAAAAACTGCTAACTCTTCACT**GATAAC**T Nostoc punctiforme; Npun_R5632; 19.29
GCAGTAATAAGAGCTAAACTTCTGTTTTTGACTGAGGTAA**TAGCCA**A Nostoc sp. PCC 7120 (Anabaena sp.); all2457; 21.29

Chroococcales:

CAAGTCGCTGGCGCTACCCAAATCACAGGCTGTCAGCG*CAC**TGATGC**CTTC Synechococcus sp. WH 8102; SYNW2456; 22.71
AAAGTCGCCAGGGCTACCTGAATCACAGGCTGCCAGTG*GCT**TGATGC**CTCC Synechococcus sp. WH 7805; WH7805_09759; 22.82
TCGGTAACAAACAACACCTAGTTCAACAGACAGACGGATCTT**TGATGA**GTTG Synechococcus sp. CC9902; Syncc9902_2264; 21.18
TCAGTAACAAACAACACTTAGCTCAGCAGACACACAGTTCTT**TGATGT**ATTG Synechococcus sp. BL107; BL107_06944; 20.71
GGGGTATAAACAGAGACTAAAAAACCCACCACTCGAGCCGGT**TGATGA**CCTG -"-; BL107_06944; 20.59
TTGGTAACCGATGCGACCTGCCCTGGCAGGACAGAAGC*CTT**GGATGG**GCCC Synechococcus sp. RS9917; RS9917_06045; 22.71
TCGGTAGCGATAACCACCAGAGTTGTCAATGCCGCTGT*CCA**TGATCA**AGCC Synechococcus sp. WH 5701; WH5701_03589; 21.53
GCTGTATGACTTGTTACAGCTTCGGTCGCTCAATGCT**TGA**TACATCTAGC Synechococcus sp. PCC 7002; SYNPCC7002_A0469; 18.71
GCCGTATCCCGAACTACAGAAGTGGACTCTGAGCGAT**TTC**TATAGTCCTT Synechococcus elongatus; Synpcc7942_1032; 20.24
TTGGTAGCGACAATGACCTTACCACAAACCTCATGGG**GAA**TATCATGAAT Cyanothece sp. PCC 8801; PCC8801_1913; 22.24

Prochlorales:

AGCGTATTTAAATATACACATCTTAATTCTTTAAAGC**TTT**TAGAGTGCTT Prochlorococcus marinus NATL2A; PMN2A_0733; 18.59
TTTGTATTTTTAGCTACCTAGTTAATTTAAAACTTGCT*CCA**TATTTT**ATAA -"-; PMN2A_0733; 19.18
AGTGTATTGATAAATACATTTATTGTTGTCATATAGTT*TCT**TGCCTT**CTTT Prochlorococcus marinus NATL1A; NATL1_15711; 18.94
TTTGTATTTTTAGATACCTGAGTTATGAATATCTTGAT*TCA**TATTTT**ATAA -"-; NATL1_15711; 20.24
TTGGTAGCCATAAATACGAATAAAGATCTAAAGCTA***AAG**TAAAAG**TTAC Prochlorococcus marinus MIT 9215; P9215_11721; 18.71
TTGGTAGCTATAAATACGAATAAAGATATAAAGCTA***AAG**TAAAAG**TTAC Prochlorococcus marinus MIT 9301; P9301_11431; 18.35
TTGGTAGCCATAAATACGAATAAAGATATAAAGCTA***AAG**TAAAAG**TTAC Prochlorococcus marinus AS9601; A9601_11421; 18.71
TTTGTAGTTTTAAATACTAAATATCGTCCCAAGCTA***AAG**TAAAAA**GTTA Prochlorococcus marinus CCMP1986; PMM1038; 18
TTTGTAGCTATAAATACGAATTAAGATTCGAAGCTA***AAG**TAAAAG**TTAT Prochlorococcus marinus MIT 9312; PMT9312_1049; 17.29

Oscillatoriales:

TCAGTAGTTTATGCTACAAAAACCTGCAAAACCCCAA**AAG**TGGGAT**CGCT Oscillatoria sp. PCC 6506; OSCI_3880026; 16.47
TTCGTATCATTAGTTACAGCCTCTAGACTCAGACTCC**CCT**TATCGT**GGGA Lyngbya sp. PCC 8106; L8106_30150; 18.71
TTGGTATCATTAACTACAGTTCCCGGACGGACTCAGCC*CCT**TATCAT**CAGA Arthrospira platensis; AplaP_010100003884; 19.18
CTAGTACCGCGTTGGACTACACTTCCGGCGGAACAGT**CCC**TAGAAT**TGGC -"-; AplaP_010100003884; 24.12
TTGGTATCATTAACTACAGTTCCCGGACGGACTCAGCC*CCT**TATCAT**CAGA Arthrospira maxima; AmaxDRAFT_2064; 19.18

Genus *Acaryochloris*, unclassified order:

CAGGTATCAAAACATACAGTTTAGAGATATCAAGTTAC*CAG**TAAGAT**ACCC Acaryochloris marina; AM1_1414; 19.41

Putatively repressed genes:

***gifA*,** glutamine synthetase inactivating factor IF7

Chroococcales:

```
AAATTTTTGTTTCTGTATAAAATGTTACAGAATTTG***TGCTATAAATATTAAA Synechocystis sp. PCC 6803; ssl1911; 8.67
GATTAGGGCATTTTGTAAAAATAGATACAGAATTGTA**TGTTATATATATTTAG Microcystis aeruginosa; MAE_49490; 12
TATCTAAAAAAAATGTAACAATAGATACAGAAAATA***AATCGTTCATCCTTCT Cyanothece sp. PCC 7424; PCC7424_2046; 9.33
TAAAGTCCTATTTTGTAGCAAATGTTACAGAATTGTG**TGTTATACATAGTAGT Cyanothece sp. ATCC 51142; cce_0259; 12
CAAAACATCATTTTGTAGCAAATGTTACAGAATTTTA**TGTTATAAATAATAAT Cyanothece sp. CCY0110; CY0110_12762; 9.33
TTAATCAAAAAACTGTTAAATTAGATACAGAATAAAT**CGTTCATCTCTCCAAT Cyanothece sp. PCC 7425; Cyan7425_2387; 13.33
CATTTCAAAAAAATGTAAAAATAGATACAGAATAAAA**TTTTAAGTTTGTCTAG Cyanothece sp. PCC 7822; Cyan7822DRAFT_4909; 12.67
CTTTTCAAGATTCTGTATAAATTGCTACAGTTTAATGTGGTATATATAGTAATAG Cyanothece sp. PCC 8801; PCC8801_2371; 12
CTTTTCAAGATTCTGTATAAATTGCTACAGTTTAATGTGGTATATATAGTAATAG Cyanothece sp. PCC 8802; Cyan8802_2421; 12
AAAGATGTATAACGGTAGCAATTTATACAGAAAAATA**TTGCTAGATTTAAATC Synechococcus elongatus; Synpcc7942_0900; 8.67
AAAATATGTTTTCTGTATCAAATTTTACAGAAAATA***AAATATAATGAGGTAG Synechococcus sp. PCC 7002; SYNPCC7002_A0582; 8.67
ATAATACCAAAAGTGTTAAATTAGATACAGAAATAA***ATCGTTCATCTTCGCT Thermosynechococcus elongatus; tll1670; 15.33
```

Genus *Acaryochloris*, unclassified order:

```
TTCAATTTAAAAATGTTATATTTGATACAGAAATTAA**ATCGTTCATCCCTTAA Acaryochloris marina; AM1_2559; 15.33
```

Nostocales:

```
AAAATTGCAATTCTGTAACTTTAGCTACAGAAATCT***CGCTAATATAAGAACT Nodularia spumigena; N9414_03593; 13.33
GAAATTGCTATTTCGTAGCTTACGATACAGAATTTCAGCTACTATCATAAATGTG Nostoc azollae (Anabaena azollae); AazoDRAFT_1774; 12.67
AAATCTTCAATTCCGTAGCATAAGATACAGAATTCT***TGCTATATTAAATGTG Nostoc sp. PCC 7120 (Anabaena sp); asl2329; 8
AATTTATGTAAAAAGTAGCTTATGTTACAGAATTTT***TGTTATGTTAGTCTTG Nostoc punctiforme; Npun_F5386; 13.33
AGATGCTCAATTTCGTAGCATAAGATACAGAATTTT***CGCTATATTAATTATG Anabaena variabilis ATCC 29413; Ava_0148; 10
CAATTTGGGATTTGGTAGTTTAAGTTACGGAATTCCC**TGTTACTATGAGGAAG Cylindrospermopsis raciborskii; CRC_02159; 16
CAAATTAGAATTTGGTAGTTTAAGTTACGGAATTCCC**TGCTACTATAAAAAAG Raphidiopsis brookii; CRD_02135; 16
```

Oscillatoriales:

```
TTAAAACTCATTTTGTAAAATACGCTACAGAATTTTC**GTGTATGATGAAAGAG Lyngbya sp. PCC 8106; L8106_21554; 14
CTCAGAAGACGGAAGTAAGGAAACATCCTGAAGGAATTTACCTCAAATACTCTGC Trichodesmium erythraeum; Tery_3029; 21.33
ACTTCTGGTATTCTGTAAAATAGACTACAGAATTT**TATAGTATAGTAGTAAGC Arthrospira maxima; AmaxDRAFT_1837; 13.33
ACTTCTGGTATTCTGTAAAATAGACTACAGAATTT**TATAGTATAGTAGTAAGC Arthrospira sp. PCC 8005; APCC8_010100011341; 13.33
AAAAAAGCGATTTTGTAATTTGAATACAGAATTGTC**TGTTATAGTCATAAAA Oscillatoria sp. PCC 6506; OSCI_3000009; 16.67
```

***gifB*,** glutamine synthetase inactivating factor IF17

Chroococcales:

```
ATTTTGTTGACCATTCCTTGACATGATCTTGAAAAACCGTAAAAATGGATACAG Synechocystis sp. PCC 6803; sll1515; 6
ATATGAAGTCTCATTACTTGACATTCTCAAAAAAATCTGTAACAATAGATACAG Cyanothece sp. ATCC 51142; cce_2638; 8
AATTTTTATAAAGCCCTGACAAATTCACCTGCAAAAATGTAACAATAGATACAG Cyanothece sp. PCC 8801; PCC8801_3716; 10
AATTTTTATAAAGCCCTGACAAATTCACCTGCAAAAATGTAACAATAGATACAG Cyanothece sp. PCC 8802; Cyan8802_3768; 10
TTTTAAATGATTTTTTTTTCTTATATAAGAATAAAATGTAAAATTTGATACAG Cyanothece sp. PCC 7424; PCC7424_1285; 11
GCAAAGGTTTACCAATGCACGATTGCCCAGATCAAACGGTATCATATAATACAG Cyanothece sp. PCC 7425_1; Cyan7425_1393; 12
ACAAATAAATCAGAGGTTGCTGACAAACTAGGGAAACCGTATAATTAGATACAG Cyanothece sp. PCC 7425_2; Cyan7425_4644; 7
```

AAACAAAATATTAAGTTTTATAGACAAAACATCATTTTGTAGCAAATGTTACAG Cyanothece sp. CCY0110; CY0110_12762; 13
TAATTTTTGTTAACTTTGGCCATTCATTTCAAAAAAATGTAAAAATAGATACAG Cyanothece sp. PCC 7822; Cyan7822DRAFT_4909; 10
AAATTTTTATAAAGTTTTTTCAATTTCCCTAACAAAATGTAAAAATAGATACAG Microcystis aeruginosa; MAE_26330; 10
TAAATTTTTATGAAGTCCTGACATTTTCTCAAAAAAGTGTAACAATAGATACAG Crocosphaera watsonii; CwatDRAFT_1910; 10
ACAACAACTGGCGTTTACCGTATGAATATTGGAAAATCGTATAATAGAATACAG Thermosynechococcus elongatus; tlr0313; 11

Oscillatoriales:

AAAAAAGTGAGAGATACTTGACATTTTTCTGAAAAAGCGTAACAATAGATACAG Lyngbya sp. PCC 8106; L8106_24935; 8
AAAGTTTTGTAAAATATTGGGGATCACCCGGAAAAAATGTAACAATGAATACAG Arthrospira maxima; AmaxDRAFT_0891; 12
AGAAATCCCATAGACCCTTGACATTTATTTGGGAAACGGTAACAATAGATACAG Arthrospira platensis; AplaP_010100005119; 8
TTAAATTTTGTCAACGAACCCTGTTATTTGGCAAAACAGTAAAATACAATACAG Oscillatoria sp. PCC 6506; OSCI_3750001; 11

Nostocales:

CACTGAAACATAAAGTTTTATTACAAAATTGCAATTCTGTAACTTTAGCTACAG Nodularia spumigena; N9414_03593; 12
TACAATAATATTAAAAATTGTAACAGATGCTCAATTTCGTAGCATAAGATACAG Anabaena variabilis ATCC 29413; Ava_0148; 8
TACATAAGTATTACAATTTATAACAAATCTTCAATTCCGTAGCATAAGATACAG Nostoc sp. PCC 7120 (Anabaena sp.); asl2329; 6
CACTTAAGTGTGAAGTTTTGTGATGAAATTGCTATTTCGTAGCTTACGATACAG Nostoc azollae (Anabaena azollae); AazoDRAFT_1774; 11
ATATTGTCAGCAAAGCAAAACATAAATTTATGTAAAAAGTAGCTTATGTTACAG Nostoc punctiforme; Npun_F5386; 11

Genus *Acaryochloris*, unclassified order:

TTATCGGGTTTTGTAGGCCGATCTTTAATTTGAAAAGGGTATTATTAGATACAG Acaryochloris marina; AM1_3215; 12

Fig. 2. An alignment of candidate NtcB factor binding sites in cyanobacteria. In yellow are conserved regions with the consensus TGCA-5N-TGCA, with varied sites shown in green. Complementary binding regions with maximal length are underlined. Different binding regions within a leader region in one species/strain are separated by «-"-».

***nrtA*,** nitrate transport 45kD protein

Chroococcales:

```
ACCAATGCAGTAATCGCATGAAAATTAATTATTAGTTACAAACTATACAAATATTTACAAGGAAAAAATC*CAGTCT Synechocystis sp. PCC 6803; sll1450; 10
TTTTATGCTTTTATGACATGGCTGACAATATTTTGTAACAATCTATACGAATTTTCCTGTTGAAGACTGTAAATTTA Microcystis aeruginosa; MAE_14800; 19.5
AGCTATGCAAAAACCACATAAAAGCAATGCTTTTATTGCAAAAATTGTAATATTTTGTAGCAAATAATACGAAAAAT Cyanothece sp. PCC 8801; PCC8801_4396; 14
AGCAATGCTTTTATTGCAAAAATTGTAATATTTTGTAGCAAATAATACGAAAAATTCAAAATCCCCATGTAATGCTT -"-; PCC8801_4396; 16.5
AGCTATGCAAAAACCACATAAAAGCAATGCTTTTATTGCAAAAATTGTAATATTTTGTAGCAAATAATACGAAAAAT Cyanothece sp. PCC 8802; Cyan8802_4459; 14
AGCAATGCTTTTATTGCAAAAATTGTAATATTTTGTAGCAAATAATACGAAAAATTCAAAATCCCCATGTAATGCTT -"-; Cyan8802_4459; 16.5
AAAAATGCTTAAATTGCATCAAGGCAAGCTTCTATAGCCATCTTCTCTAAACTTTTGTTACAAATTATACAAAAAAG Cyanothece sp. PCC 7424; PCC7424_3527; 12.5
TATTATGCAAAAAACGCATCAAATAGTTTTTTTTGTTACAAAAACTACAAATTTCTGAATTAAAGACTGTAAATTTT -"-; PCC7424_3527; 12
GAACATGCAAAAATTGCGCTCCTTAGTGTGACGGAAGGAGAAGATAAACCCCTCAAATATCCCGTTATGTTCCAAGA Synechococcus sp. PCC 7002; SYNPCC7002_A1748; 17.5
```

Oscillatoriales:

```
TCAGATGCAAGTTACGCATGACCGAAATGCGTAAAATAAAAAACAATCTAAAAATTGTAGTTGCCGATACATAAACC Arthrospira maxima; AmaxDRAFT_3857; 16.5
TCAGATGCAAGTTACGCATGACCGAAATGCGTAAAATAAAAAACAATCTAAAAATTGTAGTTGCCGATACATAAACC Arthrospira platensis; AplaP_010100019381; 16.5
TCAGATGCAAGTTACGCATGACCGAAATGCGTAAAATAAAAAACAATCTAAAAATTGTAGTTGCCGATACATAAACC Arthrospira sp. PCC 8005; APCC8_010100008251; 16.5
CAAGTTGCAGACATTGGTTTACTCGGTACAAAAGTCCGCAAAAACGGCAAAACATTAGAAGGTGTTGACCTTTATAT Oscillatoria sp. PCC 6506; OSCI_2880002; 18.5
```

***nirA*,** ferredoxin-nitrite reductase

Chroococcales:

```
CTTCTTGCAGAACATGCATGATTTACAAAAAGTTGTAGTTTCTGTTACCAATTGCGAATCGAGAACTGCC*TAATCT Synechococcus elongatus; syc0310_d; 11
CTAAATGCGTAAACTGCATATGCCTTGGCTGAGTGTAATTTACGTTACAAATTTTAACGAAACGGGAACCCTATATT Synechocystis sp. PCC 6803; slr0898; 11
TGTAATACAAAAACTGTATCGAGACAAGAAAAAAGTAGCAAAATTTACAAATGTTCATGATTCATCTGGCTAAATTG Thermosynechococcus elongatus; tlr1349; 16
CTCAATGCACTTACTGCATGATATAAATCCTTAAACACGATGGTTAACGGATTTCTGTTACAGACAATACAAAATTA Microcystis aeruginosa; MAE_18410; 13.5
ATTTATGCAAAATACGCATTATATAAATGCTTATTTATCAAGCCGACCCAAATTTGGTAACAAAGTACACAAAGTAC Cyanothece sp. PCC 7822; Cyan7822DRAFT_1390; 13.5
TATCATGCAATAATCGCATAATTAAAATGCTTTTTTTTCAAGAATAGAAAAGTTTGGTAACAAAAGACACAAAGTAG Cyanothece sp. PCC 7424; PCC7424_1683; 12
GATTATGCAGGTTTGGCATCGATCCCATGACAAAAAGTCACAATAGGTAAATTTTTGTAACTTAATATACCAATTAC Cyanothece sp. PCC 7425; Cyan7425_4573; 18
TAAAATGCAAAAAATCCATAATTAAAATGCAAAAAACGGATTTTTAATACAATTTTGTTACATTAGCTACAAAATAT Cyanothece sp. ATCC 51142; cce_1223; 12
TATTATGCAAAATAATCATCTTTTAAATCCGTTTTTTTCATAAAATCCTTATTTTTGTTACAACTGATACAAAAATT Cyanothece sp.PCC 8802; Cyan8802_3641;16.5
TGTTATGCAAAATAATCATCTTTAAAATCCGCTTTTTACATAAAATCCTTATTTTTGTTACAATTAATACAAAAATT Cyanothece sp. PCC 8801; cce_1223; 16.5
AAAAATGCAAAAAATACATGAAAAAAATGCAAAAGACGAATTTTGCCCAACTTTTTGTTACATTAGCTACAAAATTT Cyanothece sp. CCY0110; CY0110_23451; 12
AGATATGCAAGAAAAACATAATAAAAATGCAAAAAACGAATGTAATTTAACATTTTGTTACATTAGCTACATAAATT Crocosphaera watsonii; CwatDRAFT_3683; 16
TTGAATGCGGCGATCGCATCATCAATATTCTGATTTCGCAAAAGCAAGCCAGGCCCGTAGTTAACACTACAAAATCT Synechococcus sp. PCC 7002; SYNPCC7002_A1827; 15
GGCAATGAAGCAAGTTGAAGGCGAAAGTCAAACGGTCTAATTCGCTACCGACTGAAAGGGATACGCACGGTTAATTG Synechococcus sp. CC9902; Syncc9902_2284; 20
GGCGATGGATGCTCTGCAGTCAGCACTTCCAATTCAACTGGGGCCGTTCGGCGCTCAGCCACAGCATCCACAATCTG Synechococcus sp. CC9311; sync_2898; 22.5
```

Nostocales:

```
TGTAATGCAGAAAACGCATATTCTCTATTAAACTTACGCATTAATACGAGAATTTTGTAGCTACTTATACTATTTTA Nostoc sp. PCC 7120 (Anabaena sp.PCC7120); alr0607;8
CTTGATGCAAAAAATGCATATTATGCATCCATTTTCAGCAATTTTACTAAAAAATCGTAACAATTTATACGATTTTA Nostoc punctiforme; Npun_R1528; 11
AAAAATGCATATTATGCATCCATTTTCAGCAATTTTACTAAAAAATCGTAACAATTTATACGATTTTAACAGAAATC -"-; Npun_R1528; 15
```

TCTCATGCAGAAATTGCATTCATTTGATGCGCTTTGAACAAAAAATTCAAAAAAAGGTAGCAATGTATACTATTTTA Nodularia spumigena; N9414_05289; 11.5
ATCAATGCAAAATTTGCACAACGGCCGAGAATTTCTGAGAATTCTGTAACTGAGTATACGAATCCTTCCCGAAATTT Cylindrospermopsis raciborskii; CRC_00047; 14
ATAAATGCAAAATCTGCATAGCATCCGAGAATTTCTGAGAATTCCGTAACTGAGTATACGAATCCCTCCCGAAATTT Raphidiopsis brookii; CRD_02595; 13

Oscillatoriales:

GTTTATGTAAAAAACGCATTTTTTATATTCTAAAACTTCATATAACTTAAAATTTTGTAACTTTTGCTACGATTAAA Lyngbya sp. PCC 8106; L8106_10091; 14
TTTTATGTAATAAATGCATTTTAACAATGCATAAATGTCATAAAAGACTCTATTTTGTTATATAAGCTACGAATTCT Trichodesmium erythraeum; Tery_1068; 15
AACAATGCATAAATGTCATAAAAGACTCTATTTTGTTATATAAGCTACGAATTCTTTATTTCCTTTCCTTTAAATTG -"-; Tery_1068; 13
ATATATGCAATAATTGCAAAATCAGAGGTTTTATGTATCTTCAACTACAAATTATCACTGATTCTGAATTTAATCTA Arthrospira platensis; AplaP_010100015448; 15
ATATATGCAATAACTGCAAACTCTTGGATTTTATGTATCTTAAACTACAAATTATCACTGATTCTGAATTTAATCTA Arthrospira sp. PCC 8005; APCC8_010100020361; 16
CGCGATGTAGAAACTGCATCGTGCTTATCCGAAAACTTCATAAAACTCCATATTCTGTAACATTAGCTACGAACAAC Oscillatoria sp. PCC 6506; OSCI_2880001; 14.5

Genus *Acaryochloris*, unclassified order:

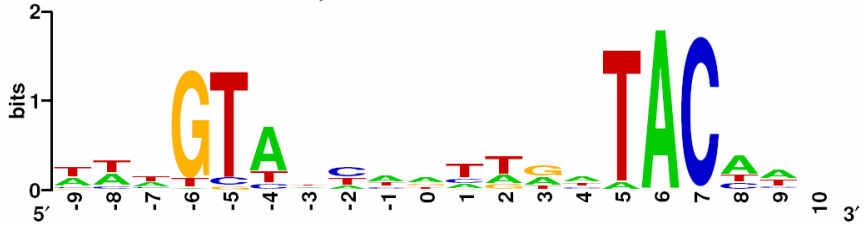
AGCAATGCAGCTAATGCATCTACAGACTGGGGCCCTATTTGGGCTAACATTTGGGCCGCAATTTCAGCAACATCATC Acaryochloris marina; AM1_2984; 13
TTCTATGCAGCTTTAGCATGAATCAATTGCAAGAACTACATTTAGCCTATCACTTTGTAACGACTGATACGAACGAT -"-; AM1_2984; 16.5
TCAATTGCAAGAACTACATTTAGCCTATCACTTTGTAACGACTGATACGAACGATCCCAATGGGAGTCGCTAAGTTA -"-; AM1_2984; 16

Gloeobacterales:

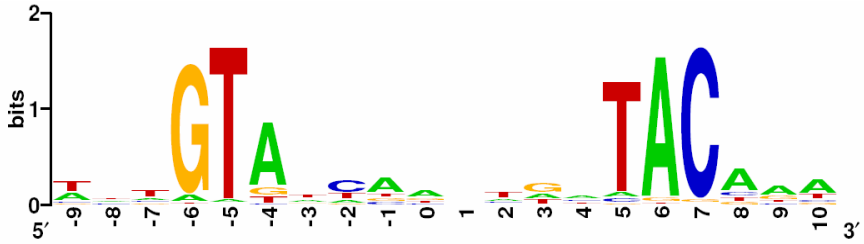
CGTTAGTCACTCCGTGCATCAGACCATATCCGGCCCGCACGGCCACGAATGTATCTGGGGTTACGTAAAGTTAAGTT Gloeobacter violaceus; gvip212; 21

Fig. 3. Nucleotide frequency distributions in the predicted NtcA binding sites in five orders and five species of cyanobacteria. Other genera are sparsely represented.

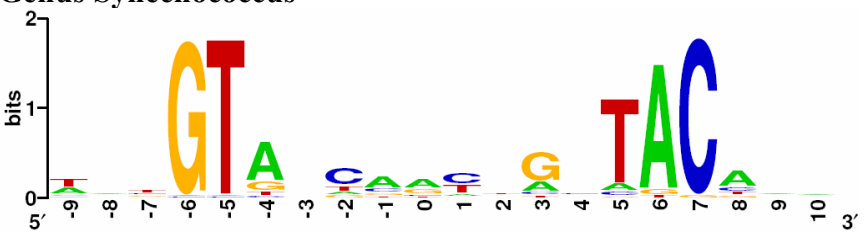
Genus Prochlorococcus, order Prochlorales



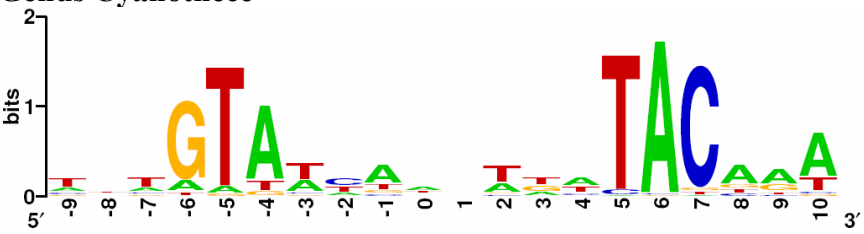
Order Chroococcales



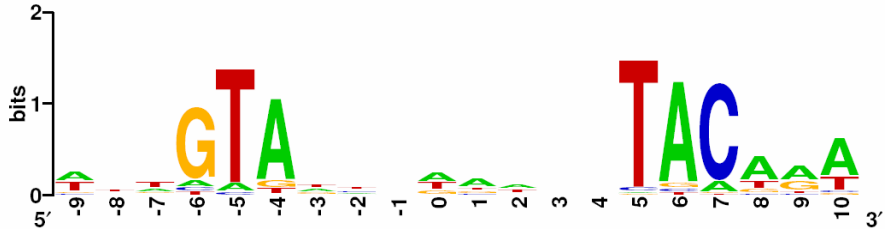
Genus Synechococcus



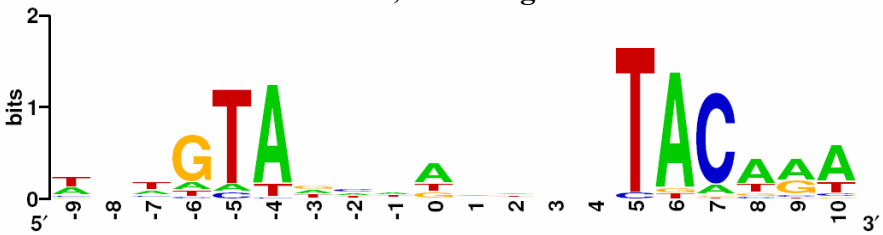
Genus Cyanothece



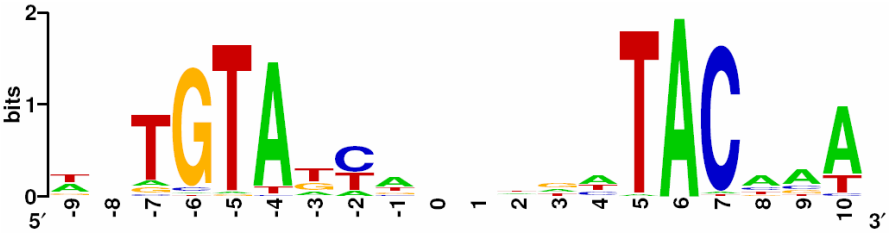
Order Nostocales



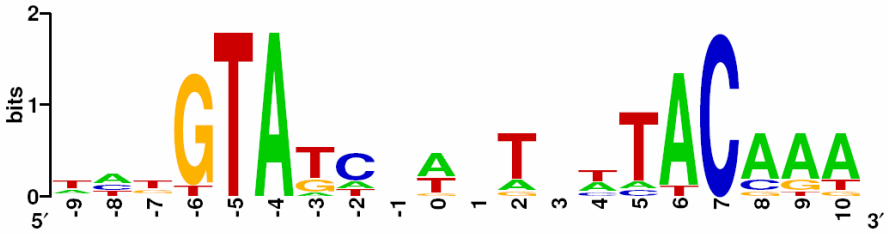
Genus Nostoc and Anabaena, taken together



Order Oscillatoriales



Genus *Acaryochloris*, unclassified order



Order Gloeobacterales

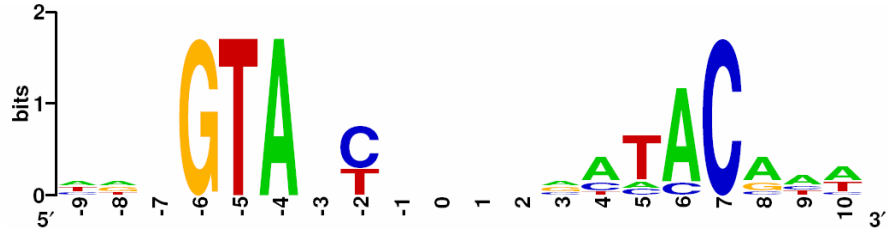


Fig. 4a. Nucleotide frequency distribution in NtcA binding sites in cyanobacteria. The distribution is built for all predicted regulatory sites. Most conserved positions included in the consensus are underlined.

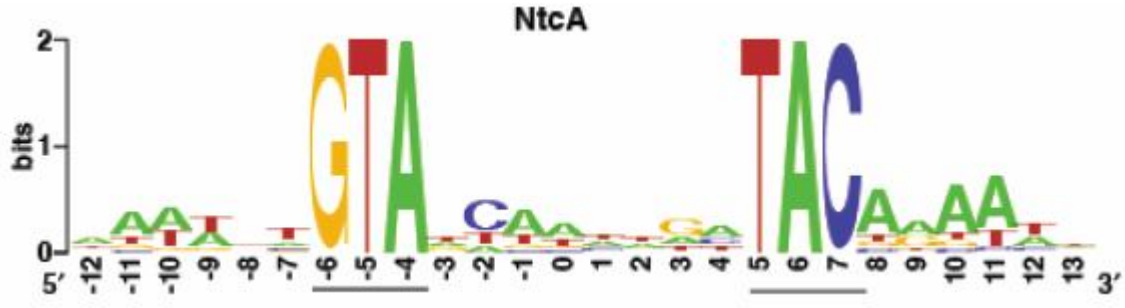


Fig. 4b. Nucleotide frequency distribution in NtcB binding sites in cyanobacteria. The distribution is built for all predicted regulatory sites. Most conserved positions included in the consensus are underlined.

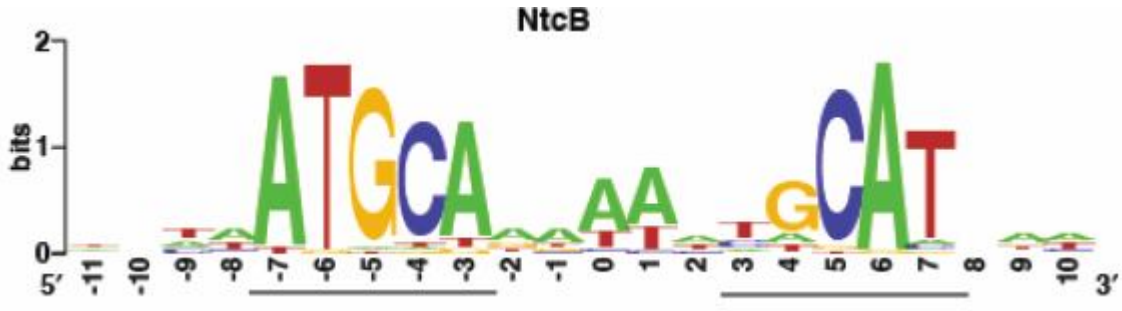


Fig. 5. The cyanobacterial NtcA factor tree. Taxonomy is colored by orders. The factor phylogeny and species taxonomy are in good agreement.

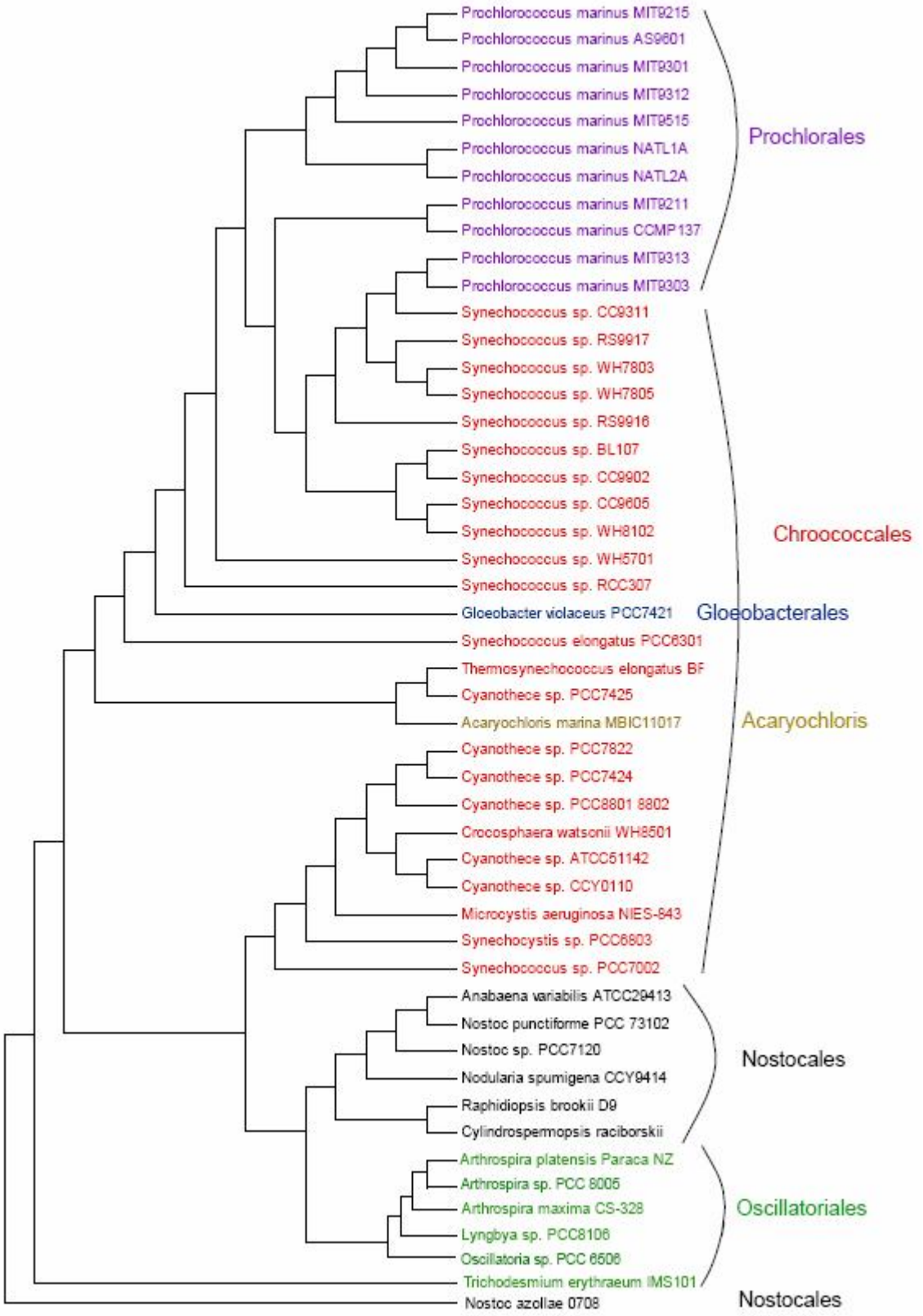


Fig. 6. Candidate promoters upstream gene *rps20*. In bold are (–35)- and (–10)-boxes and 5'-extensions of the (–10)-box. Colored are conserved putative binding sites of the NtcA (Ycf28) factor. In the last column are distances to the start codon. The Ycf28 protein possesses the PF00325 domain exclusively in rhodophyte plastomes shown in rows 2-4.

Cyanidioschyzon merolae	ACTC**TTGCTT**TTTGCCATCTGCT=ATTT**TATCTT**TATGTAGACT	-33
Cyanidium caldarium	AAAT**TTGTTT**ATTTTACTTTAAT=A**TG**A**TACAGT**AATTTATAAC	-32
Porphyra purpurea	GCTA**TTGCCT**ATTCTTTTTTTTTAA**TGT****TATAAT**ACGGCGCATA	-78
Porphyra yezoensis	ACTA**TTGCCT**ATTGTTTTTCTTTAA**TGT****TATAAT**ACGCCGCATA	-78
Gracilaria tenuistipitata	GTTC**TTGTCT**ATTTTAATGTATTAA**TG**A**TATAAT**CCAATTAGAT	-63
Guillardia theta	TTAA**TTTATT**CCATTATTTCTTATA**TGT****TATAAT**CTTTTATTAC	-59

Table 1. **Predicted NtcA-mediated regulation in cyanobacteria.** Designations: «+» – predicted factor binding site adjoins the gene, «-» – both gene and site absent, «ns» – gene present but the site or promoter box not found (the gene may be regulated within an operon or not regulated), «*» – site found but contains deviations from the consensus GTA-8N-TAC. NtcA-activated genes: *glnA, glnB, glnN, nirA, icd, amt1/B, narB, narK*, *ntcA, ntcB, nrtA, psaI, petF/fdx, apcF, som, urtA, cynA, speB, mutS, rnc, hupS*; NtcA-repressed genes: *gifA, gifB*.

Order, Species	**Activator**									
Chroococcales:	glnA	nirA	glnN	glnB	icd	amt1/B	narB	ntcA	ntcB	nrtA
Synechocystis sp. PCC 6803	+	+	*	+	*	+	ns	ns	ns	*
Microcystis aeruginosa	+	*	+	+	ns	+	+	ns	+	+
Cyanothece sp. CCY0110	+	*	-	*	*	*	*	ns	+	ns
Cyanothece sp. ATCC 51142	+	*	-	*	ns	*	*	ns	+	-
Cyanothece sp. PCC 8802	+	*	-	*	*	+	+	ns	+	+
Cyanothece sp. PCC 8801	+	*	-	*	*	+	+	ns	+	+
Cyanothece sp. PCC 7425	+	+	+	ns	ns	+	*	+	+	ns
Cyanothece sp. PCC 7424	+	*	-	+	+	+	+	ns	+	+
Cyanothece sp. PCC 7822	+	*	-	*	-	+	*	+	+	+
Crocosphaera watsonii	+	*	-	ns	+	+	*	ns	+	-
Synechococcus sp. WH 5701	*	*	*	+	ns	ns	ns	*	-	+
Synechococcus sp. RS9916	*	ns	-	ns	ns	ns	ns	*	-	-
Synechococcus sp. CC9311	*	*	*	+	ns	ns	ns	*	-	-
Synechococcus sp. WH 8102	*	ns	-	ns	ns	ns	ns	+	-	+
Synechococcus sp. PCC 7002	*	+	*	ns	ns	+	ns	ns	*	+
Synechococcus sp. WH 7803	+	*	*	*	ns	ns	ns	*	-	-
Synechococcus sp. WH 7805	+	*	*	*	ns	+	ns	*	-	-
Synechococcus sp. CC9605	+	+	-	*	ns	ns	ns	+	-	-
Synechococcus sp. RCC307	+	+	-	+	ns	ns	+	*	-	-
Synechococcus sp. CC9902	+	*	-	*	ns	ns	ns	+	-	-
Synechococcus sp. BL107	*	*	-	*	ns	ns	ns	+	-	-
Synechococcus sp. RS9917	+	*	*	ns	ns	ns	ns	*	-	-
Synechococcus elongatus	+	+	*	+	ns	+	ns	+	ns	ns
Thermosynechococcus elongatus	+	+	-	+	ns	+	ns	*	*	ns
Prochlorales:	glnA	nirA	glnN	glnB	icd	amt1/B	narB	ntcA	ntcB	nrtA
Prochlorococ. marinus CCMP1375	+	-	-	+	ns	+	-	*	-	-
Prochlorococ. marinus CCMP1986	*	-	-	*	*	+	-	*	-	+
Prochlorococ. marinus MIT 9211	+	-	-	+	ns	+	-	*	-	-
Prochlorococ. marinus MIT 9303	+	*	-	+	ns	*	-	*	-	-
Prochlorococ. marinus MIT 9313	+	*	-	+	ns	*	-	*	-	-
Prochlorococ. marinus MIT 9215	+	-	-	*	*	+	-	*	-	-
Prochlorococcus marinus AS9601	*	-	-	+	*	+	-	+	-	-
Prochlorococ. marinus MIT 9312	*	-	-	*	*	+	-	*	-	-
Prochlorococ. marinus MIT 9515	*	-	-	+	ns	+	-	+	-	-
Prochlorococ. marinus MIT 9301	*	-	-	+	*	+	-	+	-	-
Prochlorococcus marinus NATL2A	*	+	-	+	*	+	-	*	-	-
Prochlorococcus marinus NATL1A	*	+	-	+	*	*	-	*	-	-
Nostocales:	glnA	nirA	glnN	glnB	icd	amt1/B	narB	ntcA	ntcB	nrtA
Anabaena variabilis ATCC 29413	+	*	-	*	*	+	ns	+	+	ns
Nostoc azollae (A. azollae)	*	+	-	*	*	+	-	+	*	-
Nostoc sp. PCC 7120 (Anabaena)	+	+	-	*	*	+	ns	+	+	ns
Nostoc punctiforme	+	+	-	*	*	+	ns	+	+	ns
Nodularia spumigena	+	+	-	*	*	+	ns	+	+	-
Raphidiopsis brookii	+	+	-	*	ns	*	ns	ns	+	ns
Cylindrospermopsis raciborskii	+	+	-	*	ns	*	ns	*	+	ns
Oscillatoriales:	glnA	nirA	glnN	glnB	icd	amt1/B	narB	ntcA	ntcB	nrtA
Lyngbya sp. PCC 8106	+	+	-	+	*	+	ns	ns	+	-
Arthrospira maxima	+	*	-	+	*	+	ns	ns	+	+
Arthrospira platensis	+	+	-	+	ns	+	ns	ns	+	+
Arthrospira sp. PCC 8005	+	+	-	+	ns	+	ns	ns	+	+
Oscillatoria sp. PCC 6506	+	+	-	ns	*	+	ns	ns	+	ns
Trichodesmium erythraeum	+	*	-	+	+	+	ns	ns	*	-
Genus Acaryochloris:	glnA	nirA	glnN	glnB	icd	amt1/B	narB	ntcA	ntcB	nrtA
Acaryochloris marina	+	+	*	*	ns	+	+	+	+	+

Gloeobacterales:	glnA	nirA	glnN	glnB	icd	amt1/B	narB	ntcA	ntcB	nrtA
Gloeobacter violaceus	+	+	*	*	ns	+	ns	+	+	*

Table 1 continued:

Order, Species	**Activator**					**Repressor**	
Chroococcales:	psaI	petF/fdx	narK	apcF	som	gifA	gifB
Synechocystis sp. PCC 6803	ns	ns	ns	ns	-	+	+
Microcystis aeruginosa	ns	*	ns	ns	ns	+	+
Cyanothece sp. CCY0110	ns	ns	ns	*	ns	+	+
Cyanothece sp. ATCC 51142	ns	*	ns	*	ns	+	+
Cyanothece sp. PCC 8802	ns	ns	ns	+	ns	+	+
Cyanothece sp. PCC 8801	ns	ns	ns	+	ns	+	+
Cyanothece sp. PCC 7425	ns	ns	ns	+	+	*	+
Cyanothece sp. PCC 7424	ns	*	ns	ns	ns	+	+
Cyanothece sp. PCC 7822	ns	*	ns	*	ns	+	+
Crocosphaera watsonii	-	ns	ns	*	ns	-	+
Synechococcus sp. WH 5701	ns	*	ns	ns	ns	-	-
Synechococcus sp. RS9916	ns	ns	ns	ns	ns	-	-
Synechococcus sp. CC9311	ns	ns	ns	ns	ns	-	-
Synechococcus sp. WH 8102	ns	+	*	ns	ns	-	-
Synechococcus sp. PCC 7002	ns	*	+	*	ns	+	-
Synechococcus sp. WH 7803	ns	ns	+	ns	ns	-	-
Synechococcus sp. WH 7805	ns	ns	+	+	ns	-	-
Synechococcus sp. CC9605	ns	ns	*	ns	ns	-	-
Synechococcus sp. RCC307	ns	ns	ns	ns	ns	-	-
Synechococcus sp. CC9902	ns	ns	+	ns	+	-	-
Synechococcus sp. BL107	ns	*	+	ns	ns	-	-
Synechococcus sp. RS9917	ns	ns	ns	ns	ns	+	-
Synechococcus elongatus	ns	ns	ns	*	ns	+	-
Thermosynechococcus elongatus	ns	ns	ns	ns	*	*	+
Prochlorales:	psaI	petF/fdx	narK	apcF	som	gifA	gifB
Prochlorococ. marinus CCMP1375	*	ns	-	-	+	-	-
Prochlorococ. marinus CCMP1986	+	*	-	-	+	-	-
Prochlorococ. marinus MIT 9211	ns	ns	-	-	*	-	-
Prochlorococ. marinus MIT 9303	ns	ns	-	-	*	-	-
Prochlorococ. marinus MIT 9313	ns	ns	-	-	*	-	-
Prochlorococ. marinus MIT 9215	*	ns	-	-	+	-	-
Prochlorococcus marinus AS9601	*	ns	-	-	+	-	-
Prochlorococ. marinus MIT 9312	*	ns	-	-	+	-	-
Prochlorococ. marinus MIT 9515	ns	*	-	-	+	-	-
Prochlorococ. marinus MIT 9301	*	*	-	-	+	-	-
Prochlorococcus marinus NATL2A	ns	ns	-	-	+	-	-
Prochlorococcus marinus NATL1A	ns	ns	-	-	+	-	-
Nostocales:	psaI	petF/fdx	narK	apcF	som	gifA	gifB
Anabaena variabilis ATCC 29413	ns	*	ns	+	+	+	+
Nostoc azollae (A.azollae)	ns	*	-	*	ns	+	+
Nostoc sp. PCC 7120 (Anabaena)	ns	*	ns	+	*	+	+
Nostoc punctiforme	ns	ns	ns	+	ns	+	+
Nodularia spumigena	-	ns	ns	+	*	+	+
Raphidiopsis brookii	-	ns	ns	+	*	+	-
Cylindrospermopsis raciborskii	-	ns	ns	+	*	+	-
Oscillatoriales:	psaI	petF/fdx	narK	apcF	som	gifA	gifB
Lyngbya sp. PCC 8106	ns	ns	ns	+	ns	+	+
Arthrospira maxima	ns	ns	ns	+	ns	+	+
Arthrospira platensis	ns	ns	ns	+	ns	ns	+
Arthrospira sp. PCC 8005	ns	ns	ns	+	ns	+	ns
Oscillatoria sp. PCC 6506	ns	ns	-	ns	ns	+	+
Trichodesmium erythraeum	ns	*	ns	ns	+	*	-
Genus Acaryochloris:	psaI	petF/fdx	narK	apcF	som	gifA	gifB
Acaryochloris marina	-	ns	ns	ns	+	*	+
Gloeobacterales:	psaI	petF/fdx	narK	apcF	som	gifA	gifB
Gloeobacter violaceus	-	ns	ns	ns	ns	-	-

Table 1 continued:

Order, Species	Activator						
Chroococcales:	urtA	speB	mutS	rnc	hupS	nirB	cynA
Synechocystis sp. PCC 6803	+	*	ns	ns	–	–	–
Microcystis aeruginosa	+	+	ns	ns	–	–	–
Cyanothece sp. CCY0110	–	–	ns	ns	+	–	–
Cyanothece sp. ATCC 51142	–	–	ns	ns	+	–	–
Cyanothece sp. PCC 8802	–	–	ns	ns	ns	–	–
Cyanothece sp. PCC 8801	–	–	ns	ns	ns	–	–
Cyanothece sp. PCC 7425	*	–	ns	ns	–	–	–
Cyanothece sp. PCC 7424	+	ns	ns	ns	+	–	–
Cyanothece sp. PCC 7822	+	+	ns	ns	+	–	–
Crocosphaera watsonii	ns	–	ns	ns	*	–	–
Synechococcus sp. WH 5701	ns	*	ns	+	–	–	–
Synechococcus sp. RS9916	ns	ns	ns	ns	–	–	–
Synechococcus sp. CC9311	ns	ns	ns	ns	–	–	–
Synechococcus sp. WH 8102	*	*	*	+	–	ns	+
Synechococcus sp. PCC 7002	ns	ns	ns	ns	–	–	ns
Synechococcus sp. WH 7803	ns	ns	*	+	–	–	–
Synechococcus sp. WH 7805	+	ns	*	+	–	–	–
Synechococcus sp. CC9605	*	*	ns	+	–	–	–
Synechococcus sp. RCC307	+	+	ns	ns	–	–	–
Synechococcus sp. CC9902	*	ns	*	+	–	–	–
Synechococcus sp. BL107	*	ns	ns	+	–	–	–
Synechococcus sp. RS9917	ns	ns	ns	ns	–	–	–
Synechococcus elongatus	ns	ns	ns	ns	–	+	+
Thermosynechococcus elongatus	+	–	ns	–	–	–	–
Prochlorales:	urtA	speB	mutS	rnc	hupS	nirB	cynA
Prochlorococc. marinus CCMP1375	–	ns	ns	+	–	–	–
Prochlorococc. marinus CCMP1986	*	ns	ns	ns	–	–	–
Prochlorococc. marinus MIT 9211	–	ns	ns	*	–	–	–
Prochlorococc. marinus MIT 9303	+	ns	ns	+	–	–	–
Prochlorococc. marinus MIT 9313	+	ns	ns	+	–	–	–
Prochlorococc. marinus MIT 9215	*	ns	ns	ns	–	–	–
Prochlorococcus marinus AS9601	*	ns	ns	ns	–	–	–
Prochlorococc. marinus MIT 9312	*	ns	ns	ns	–	–	–
Prochlorococc. marinus MIT 9515	–	ns	ns	ns	–	–	–
Prochlorococc. marinus MIT 9301	*	ns	ns	ns	–	–	–
Prochlorococcus marinus NATL2A	ns	ns	+	ns	–	–	–
Prochlorococcus marinus NATL1A	+	ns	+	ns	–	–	–
Nostocales:	urtA	speB	mutS	rnc	hupS	nirB	cynA
Anabaena variabilis ATCC 29413	*	ns	ns	+	*	–	–
Nostoc azollae (A.azollae)	ns	ns	ns	ns	ns	–	–
Nostoc sp. PCC 7120 (Anabaena)	*	ns	ns	+	*	–	–
Nostoc punctiforme	+	ns	ns	ns	ns	–	–
Nodularia spumigena	ns	ns	ns	*	ns	–	–
Raphidiopsis brookii	ns	+	ns	*	–	–	–
Cylindrospermopsis raciborskii	*	+	ns	*	ns	–	–
Oscillatoriales:	urtA	speB	mutS	rnc	hupS	nirB	cynA
Lyngbya sp. PCC 8106	+	ns	ns	*	+	ns	–
Arthrospira maxima	ns	+	ns	*	–	–	–
Arthrospira platensis	+	+	ns	ns	–	–	–
Arthrospira sp. PCC 8005	+	+	ns	*	–	–	–
Oscillatoria sp. PCC 6506	+	ns	ns	*	–	ns	–
Trichodesmium erythraeum	ns	ns	ns	*	*	–	–
Genus Acaryochloris:	urtA	speB	mutS	rnc	hupS	nirB	cynA
Acaryochloris marina	ns	+	ns	ns	–	ns	–
Gloeobacterales:	urtA	speB	mutS	rnc	hupS	nirB	cynA
Gloeobacter violaceus	–	–	ns	ns	–	–	–

Table 2. Typical Crp family domains PF00325 at C-termini of the Ycf28 protein in plastomes of Rhodophyta. In rows: the Crp family consensus is followed by all rhodophyte species with known plastomes. Pairwise alignments of the domains found and the Crp consensus (according to the Pfam data). Numbers are the domain start and end relative to the protein N-terminus. E-values are alignment qualities (e-values >1 not shown).

Consensus of protein family crp			LpmsLRqeIAdylGlTrETVsRlLtrLrekgLI	
Cyanidioschyzon merolae	E=0.9	114	WRLS=QASLARILGTSRAAIGQVLGDWKKQAWL	145
Cyanidium caldarium	E=0.00019	157	IYIS=QHDIASILSTTRSTITRLINQLRKDNII	188
Porphyra purpurea	E=0.00052	184	LTIT=HKVLAQIIGSNRVSITRIISKLIHTKFI	215
Porphyra yezoensis	E=0.0021	184	FTIT=HKILAQIIGSNRVSVTRILANLLKTKLI	215
Gracilaria tenuistipitata	E>1			

Table 3. **Co-existence of regulation and gene**. In columns: (1) gene is mostly present, regulation is seldom, (2) gene is widely present, its regulation is mostly absent, (3) where found, gene is always under regulation.

(1)	(2)	(3)
glnB/glnK (nitrogen regulatory protein P-II, solanesyl diphosphate synthase)	***rbcL*** (ribulose bisophosphate carboxylase; type III RuBisCO)	***glnA*** (glutamine synthetase type I, glutamate-ammonia ligase)
icd (isocitrate dehydrogenase)	***gor*** (glutathione reductase)	***glnN*** (glutamine synthetase type III)
amtB/amt1 (ammonium/methylammonium permease)	***ndhB*** (NAD(P)H-quinone oxidoreductase subunit 2; Proton-translocating NADH-quinone oxidoreductase, chain N)	***gifA*** (glutamine synthetase inactivating factor IF7)
ntcA (transcriptional regulator, global nitrogen regulator, Crp/FNR family transcriptional regulator)	***cmpA*** (bicarbontate transport system substrate-binding protein, ABC-type nitrate/sulfonate/bicarbonate transport systems)	***gifB*** (glutamine synthetase inactivating factor IF17). The gene is often not annotated
narK/nrtP (nitrate transporter)	***petH*** (ferredoxin--NADP(+) reductase, oxidoreductase FAD/NAD(P)-binding)	
nrtA (nitrate transport 45kD protein)	***pcbD*** (chlorophyll a/b binding light harvesting protein)	
ntcB (nitrate assimilation transcriptional activator, transcriptional regulator, LysR family)	***xisA*** (nifD element site-specific recombinase, phage integrase family protein)	
narB (nitrate reductase, Molybdopterin oxidoreductase)	***cpcB*** (phycocyanin beta subunit)	
nirA (ferredoxin-nitrite reductase). *Synechococcus sp*. RS9916 and *Synechococcus sp*. WH8102 may lack regulation	***devB*** (heterocyst specific ABC-transporter, membrane fusion protein, 6-phosphogluconolactonase, ABC exporter membrane fusion protein)	
nirB (nitrite reductase related protein). The gene is annotated only in *Synechococcus elongatus, Acaryochloris marina, Oscillatoria* sp. PCC 6506	***ccmK*** (carbon dioxide concentrating mechanism protein)	
apcF (phycobilisome core component/allophycocyanin)	***hetC*** (heterocyst differentiation protein)	
petF/fdx (ferredoxin, 2Fe-2S type)	***urtB*** (urea ABC transporter, permease protein)	
som (porin)	***rpe*** (ribulose-phosphate 3-epimerase)	
psaI (photosystem I reaction center subunit VIII)	***nifH*** (nitrogenase reductase)	
speB (arginase)	***isiB*** (flavodoxin FldA)	
urtA (putative urea ABC transporter, substrate binding protein)	***isiA*** (iron-stress chlorophyll-binding protein)	
rnc (ribonuclease III)	***ureE*** (urease accessory protein E)	
hupS (hydrogenase small subunit)	***ureG*** (urease accessory protein G)	

Table 3 continued:

mutS (recombination and DNA strand exchange inhibitor protein)	***pcbA*** (chlorophyll a/b binding light harvesting protein; light-harvesting complex protein)	
cynA (cyanate ABC transporter substrate-binding protein)	***apcE*** (phycobilisome core-membrane linker protein)	
	apcA (allophycocyanin a chain)	
	psbB (photosystem II core light harvesting protein)	
	psbO (photosystem II manganese-stabilizing polypeptide)	
	psbZ (photosystem II reaction center protein Z)	
	psbW (photosystem II reaction center protein)	
	psbA3 (photosystem II D1 protein)	
	psbE (cytochrome b559 subunit alpha)	
	psb27 (photosystem II reaction center protein Psb27)	
	psaB (photosystem I P700 chlorophyll a apoprotein A2)	
	psaL (photosystem I reaction center protein subunit XI)	
	psaF (photosystem I subunit III precursor)	
	trxA (thioredoxin A)	
	trxM (thioredoxin M)	
	aarF (protein kinase ABC1 family)	
	cobA (uroporphyrin-III c-methyltransferase)	
	cobB (cobyrinic acid a,c-diamide synthase)	
	hetR (heterocyst differentiation control protein)	
	cynB (cyanate ABC transporter permease protein)	
	cynD (cyanate ABC transporter ATP-binding protein)	
	cynS (cyanate hydratase)	
	futC (ABC transporter ATP-binding protein)	
	gltS (sodium/glutamate symporter)	
	hisH (imidazole glycerol phosphate synthase subunit)	
	nblA (phycobilisome degradation protein)	
	hypA2 (hydrogenase expression/ formation protein, hydrogenase nickel incorporation protein)	
	hypB (hydrogenase isoenzyme formation protein, hydrogenase nickel incorporation protein)	

Table 3 continued:

	metG (methionyl-tRNA synthetase)	
	moaA (molybdenum cofactor biosynthesis protein A)	
	moaC (bifunctional molybdenum cofactor biosynthesis protein C/molybdopterin-binding protein)	
	moeA (molybdopterin biosynthesis)	
	nrtC (nitrate transport protein)	
	rnpB (RNA subunit of RNase P)	
	rpoD/sigE/sigB (RNA polymerase sigma factor)	
	tauA (ABC-type nitrate/sulfonate/ bicarbonate transport system)	
	tauB (ABC-type nitrate/sulfonate/ bicarbonate transport system)	
	tauC (ABC-type nitrate/sulfonate/ bicarbonate transport system)	
	thrC (threonine synthase)	
	urtC (membrane protein of urea ABC transport system)	
	urtD (ATP-binding subunit of urea ABC transport system)	
	urtE (ATP-binding subunit of urea ABC transport system)	

Table 4. Predicted NtcB-mediated regulation in cyanobacteria. Designations: «+» – predicted factor binding site adjoins the gene, «-» – both gene and site absent, «ns» – gene present but the site not found (the gene may be regulated within an operon or not regulated), «*» – site contains deviations from the consensus GTA-8N-TAC. NtcB-regulated genes: *nirA, nrtA*.

Order, species	**NtcB**	
Chroococcales:	nirA	nrtA
Synechocystis sp. PCC 6803	*	*
Microcystis aeruginosa	+	*
Cyanothece sp. CCY0110	*	ns
Cyanothece sp. ATCC 51142	*	-
Cyanothece sp. PCC 8802	*	*
Cyanothece sp. PCC 8801	*	*
Cyanothece sp. PCC 7425	*	ns
Cyanothece sp. PCC 7424	*	*
Cyanothece sp. PCC 7822	*	ns
Crocosphaera watsonii	*	-
Synechococcus sp. WH 5701	ns	ns
Synechococcus sp. RS9916	ns	-
Synechococcus sp. CC9311	*	-
Synechococcus sp. WH 8102	ns	ns
Synechococcus sp. PCC 7002	*	*
Synechococcus sp. WH 7803	ns	-
Synechococcus sp. WH 7805	ns	-
Synechococcus sp. CC9605	ns	-
Synechococcus sp. RCC307	ns	-
Synechococcus sp. CC9902	*	-
Synechococcus sp. BL107	ns	-
Synechococcus sp. RS9917	ns	-
Synechococcus elongatus	+	ns
Thermosynechococcus elongatus	*	ns
Prochlorales:	nirA	nrtA
Prochlorococcus marinus CCMP1375	-	-
Prochlorococcus marinus CCMP1986	-	ns
Prochlorococcus marinus MIT 9211	-	-
Prochlorococcus marinus MIT 9303	ns	-
Prochlorococcus marinus MIT 9313	ns	-
Prochlorococcus marinus MIT 9215	-	-
Prochlorococcus marinus AS9601	-	-
Prochlorococcus marinus MIT 9312	-	-
Prochlorococcus marinus MIT 9515	-	-
Prochlorococcus marinus MIT 9301	-	-
Prochlorococcus marinus NATL2A	ns	-
Prochlorococcus marinus NATL1A	ns	-
Nostocales:	nirA	nrtA
Anabaena variabilis ATCC 29413	ns	ns
Nostoc azollae (A.azollae)	ns	-
Nostoc sp. PCC 7120 (Anabaena)	*	ns
Nostoc punctiforme	+	ns
Nodularia spumigena	+	-
Raphidiopsis brookii	+	ns
Cylindrospermopsis raciborskii	+	ns
Oscillatoriales:	nirA	nrtA
Lyngbya sp. PCC 8106	*	-
Arthrospira maxima	ns	*
Arthrospira platensis	+	*
Arthrospira sp. PCC 8005	+	+
Oscillatoria sp. PCC 6506	*	*
Trichodesmium erythraeum	*	-
Genus Acaryochloris:	nirA	nrtA
Acaryochloris marina	+	ns
Gloeobacterales:	nirA	nrtA
Gloeobacter violaceus	*	ns