Table s1, Zhengchang Su, et al

Table s1. Training data sets A1 and B1 obtained by phylogenetic footprinting

Genome	Recovered rank[1]	Transcription unit	Name	NtcA binding site[2]	Downstream of NtcA binding site and -10 like box[3]	NtcA site position[4]
PCC7421	9	*gvip212*	*nirA*	GTAtctggggtTAC	GTAAAGTTAAGTTCCTTCAAAGCAT**TCGTGG**	-131
PCC7120	2	*alr0607*	*nirA*	**GTAgctacttaTAC**	TATTTTACCTGAGATCCCGACA**TAACCT**TAG	-498
MIT9313	5	*pmt2239*	*nirA*	GTTcaatctgaTAC	CGCCAATACCCTCTTCCATCAGGCT**TAATGA**	-470
PCC6301	7	*syc0310_d*	*nirA*	GTAgcaattgcTAC	TAAAAACTGCGATCGCTGC**TGAAAT**GAGCTG	-207
WH8102	9	*synw2477*	*nirA*	GTAattccatcAAC	AGAACAACTTTTGAGTACGAAC**TAGAAA**AGG	-349
PCC6803	4	*slr0898*	*nirA*	GTAatttacgtTAC	AAATTTTAACGAAACGGGAACCC**TATATT**GA	-61
Thermosynecho	2	*tlr1349*	*nirA*	GTAgcaaaattTAC	AAATGTTCATGATTCATCTGGC**TAAATT**GGA	-69
PCC7421	10	*gvip021*	*glnB*	GTGtgcaacttTAC	CCTCGTCGGGGAGCAATCGC**TAGCGG**TGGCT	-142
PCC7120	22	*all2319*	*glnB*	GTTacacagacTAC	ATCGAAAGTTTTTGCGGAAAA**TAGCAT**TTTG	-199
CCMP1375	3	*pro1615 pro1616*	*- glnB*	GTAtcatcaatTAC	AAACATTATTCTCGCTAGAGGCGC**TAATAG**G	-52
MED4	27	*pmm1462 pmm1463*	*- glnB*	GTAtcaatgatAAC	AATCTTAGTATCGTTAAAAGGGTT**TATATG**T	-52
MIT9313	10	*pmt1480 pmt1481*	*- glnB*	GTAacaacagcCAC	AACCCAGGCCCGTTGAAAAAGCCC**TAATGG**G	-32
PCC6301	9	*syc1192_d*	*glnB*	GTAgcagtaacTAC	AACTGTGGTCTAGTCAGCGGTGT**TACCAA**AG	-90
WH8102	16	*synw0462*	*glnB*	GTTacaggggcTAC	CCACACCGCCACCATTCACG**TCATGC**TTAAT	-51
PCC6803	8	*ssl0707*	*glnB*	**GTActgattttTAC**	AAAAAAACTTTTGGAGAACATGT**TAAAAG**TG	-169
Thermosynecho	8	*tll0590 tll0591*	*aroE glnB*	GTAgctattgcTAC	ACGCTATACTCTTTTGCCAATGC**TAAATA**AA	-65
PCC7421	1	*gvip146*	*glnA*	GTAtctcagacTAC	AAAACACCTCCCCGACCCTCCT**TAGTAT**GGG	-103
PCC7120	1	*alr2328*	*glnA*	**GTAacaaagacTAC**	AAAACTGTCTAATGTTTAGAATC**TACGAT**AT	-129
CCMP1375	2	*pro1038*	*glnA*	GTActtattgaTAC	AAAATAAGTTTGCTACGGTAAT**TAATTT**TAT	-56
MED4	2	*pmm0920*	*glnA*	GTTacctttgaTAC	ATAATTAGATTCGTCAAGTTAT**TAATTT**TAT	-52
MIT9313	2	*pmt0601*	*glnA*	GTAcctgttgcTAC	AAAAAGGGGGCCATGCGGCTGAT**TACGGT**CG	-66
WH8102	2	*synw1073*	*glnA*	GTGcgcgttgaTAC	AAAACAGGGCATAACGGCTCCT**TACGGT**CGT	-60
PCC6803	2	*slr1756*	*glnA*	**GTAgcgaaaaaTAC**	ATTTTCTAACTACTTGACTCTT**TACGAT**GGA	-85
PCC7421	8	*gvip454*	*ntcA*	GTAgtccctaaTAC	AAAAAAATTGCAGGGCTCGGCT**TAATTG**TTA	-62
PCC7120	23	*alr4392*	*ntcA*	GTAtaggaaagTAC	AGAAAGGTTAACGGTGCTT**TATTGA**TTTTTC	-98
CCMP1375	6	*pro0277*	*ntcA*	GTCattttgaTAC	AAGAGGTTTTAGCTTCACTTCT**TATCGT**CCT	-50
MED4	3	*pmm0246*	*ntcA*	GTTactgttgaTAC	AAGTTATTTCTTGCTGTCTTCT**TAAGTT**TTT	-50
MIT9313	3	*pmt1831*	*ntcA*	GTCaccattgcTAC	ATGGTAGTGGAGGTCCACTCCC**TAACTT**CAT	-123
PCC6301	10	*syc1377_c*	*ntcA*	GTAgcagttgcTAC	AAGCAGCAGCTAGGCTAGGCCG**TACGGT**AAC	-142
CCMP1375	4	*pro0295*	*amt1*	GTAttaatgatTAC	CAGATAGCAACTCTATTT**TACCCA**TATTTGT	-89
MED4	11	*pmm0263*	*amt1*	GTAtaaacgacTAC	TAGACATACATAAGTTTTC**TGAGCA**AAATTA	-85
MIT9313	28	*pmt1853*	*amt1*	GTAacaaaatgAAC	TTTCACCCCCGCAAATAGACC**TAAAAG**CGTC	-147
PCC6301	35	*syc1076_c*	*amt1*	GTTacatcgatTAC	AAAACAACCTTGAGTCTCGC**TGAATG**CTTAC	-112
WH8102	10	*synw0253*	*amt1*	GTTcagtcggaTAC	ACCATCCGGCGTGACCAGCAGCTC**TGCACT**C	-37
PCC6803	9	*sll0108*	*amt1*	**GTAgtaaatcaTAC**	AGAAAACAATCATGTAAAAATTGAA**TACTCT**	-174
Thermosynecho	9	*tll1985*	*amt1*	GTAtaacgtgaTAC	AGATTTTCCAATCTGAGAACTGG**TGCAAT**CA	-175
PCC7120	15	*all1951*	*urt1*	**GTAtcaaaaatAAC**	AATTCAATGGTTAAATATCAAAC**TAATAT**CA	-101
MED4	4	*pmm0970*	*urtA*	GTTacctatgcTAC	AAAACTAAATCCCACTCGTTTA**TAACTT**TTA	-48
MIT9313	1	*pmt2229*	*urtA*	GTAtcattcacTAC	AGAATCATGCCTGGCTGATCTT**TACGTT**CCC	-80
WH8102	1	*synw2442*	*urtA1*	GTTccggttgaTAC	CAAAGCGGTGGGGGGCCCTTTTT**TACCTT**CC	-52
PCC6803	3	*slr0447*	*urtA*	GTAtcctatgcTAC	ATAATTTTTTCTTAACCTACTCG**TAGTAT**GG	-198
Thermosynecho	3	*tlr1120*	*urt1*	GTAgcctagcaTAC	AGTTTGCCCCAATAACCTCCTCT**TAAGCT**AT	-68
PCC7120	1888	*alr2817*	*hetC*	**GTAacatgagaTAC**	ACAATAGCATTTATATTTGCTT**TAGTAT**CTC	-607
PCC7120	84	*alr3710 alr3711*	*devB devC*	**GTAcagtctgtTAC**	CTTTACCTGAAACAGATGAATG**TAGAAT**TTA	-738
PCC7421	3	*gvip213 glr1559*	*ntcB -*	GTAaccccagaTAC	ATTCGTGGCCGTGCGGGCCGGA**TATGGT**CTG	-33
PCC7120	6	*all0602*	*ntcB*	**GTAacaaaatcTAC**	CAAATTGGGGAGCAAAATCAGCT**AACTT**AAT	-65
PCC6301	15	*syc0308_c syc0309_c*	*ntcB nirB*	GTAacagaaacTAC	AACTTTTTGTAAATCATGCA**TGTTCT**GCAAG	-207
Thermosynecho	14	*tll1357 tll1358 tll1359*	*- - ntcB*	GTAtcaatcaaCAC	ACTTTACCATGGGAAAGTTGCA**TAGATT**GAA	-107
PCC7120	739	*alr1827*	*icd*	GTAgaggcgaaAAC	AACAAAAAATCCTAATCCCCAG**TACCCA**ATC	-84
PCC6803	12	*sll1689*	*rpoD*	**GTAtcacgaatTAC**	ACTGCCGTGAAAATTTAACGA**TATTTT**GGAC	-296
PCC7120	5	*alr4249*	*sigF*	GTAattcaaaaTAC	TTTTATTAGTTATGATACTTGA**TAAGTT**TCC	-685

[1] The rank recovered by the scanning process ; [2] bold, experimentally verified binding sites; [3] bold, -10 like boxes; [4]start from the first codon.

Table s2. Training data sets A2 and B2 collected from literatures

Genome	Transcription unit	NtcA binding site[1]	-10 like box
WH7803	*ntcA*	GTGtgcgttgcTAC	TAATTT
WH8103	*nrtP*	GTAgcaattccAAC	TTAACT
WH8103	*nirA*	GTAattccatcAAC	TAGAAA
PCC7942	*nir_operon*	GTAgtttctgtTAC	TAATCT
PCC7942	*nirB-ntcB*	GTAgcaattgcTAC	TTAGGT
PCC7942	*ntcA*	GTAgcagttgaTAC	TACGGT
PCC7942	*glnB*	GTAgaagtaacTAC	TACCAA
PCC7942	*glnA*	GTAtcagctgtTAC	TAGGAT
PCC7942	*amit1*	GTTacatcgatTAC	AATGCT
PCC7601	*glnA*	GTAacaaagacTAC	TAGGAT
PCC6903	*glnN*	GTGcgcgtttaTAC	TATCTT
PCC7002	*nrtP*	GTAtcagcggtTAC	TTCTTT
PCC6803	*icd* -	**GTAacagccaaTGC**	*TATGAT*

[1] Bold: this site together with those in Table s1 is used to compute prediction sensitivity.

Table s3. Predicted NtcA promoters in CCMP1375

Rank	Transcription unit	Name	NtcA site	Downstream of NtcA binding site and -10 like box[1]	NtcA site position	Score
1	*pro0731*	-	GTTgtttttttaTAC	AGTATTTTCAGAGATTATGGAATT**TAATGT**T	-50	15.61912
2	*pro1038*	*glnA*	GTActtattgaTAC	AAAATAAGTTTGCTACGGTAAT**TAATTT**TAT	-56	14.43201
3	*pro1615 pro1616*	*- glnB*	GTAtcatcaatTAC	AAACATTATTCTCGCTAGAGGCGC**TAATAG**G	-52	13.95702
4	*pro0295*	*amt1*	GTAttaatgatTAC	CAGATAGCAACTCTATTT**TACCCA**TATTTGT	-89	13.39198
5	*pro0658*	-	GTTtgtgcaagTAC	ACTTAATAAATAGCTAAC**TATGAT**TAAATCA	-51	13.30912
6	*pro0277*	*ntcA*	GTCatttttgaTAC	AAGAGGTTTTAGCTTCACTTCT**TATCGT**CCT	-50	13.16134
7	*pro1761*	-	GTAcctctcgaTAC	TGCTGGTGCGCTCTTACCGCACCT**TTGCAC**C	-407	13.15936
8	*pro0892*	*pcbG*	GTCaccttttttTAC	AGCAGATAGTAGCGATCGA**TAACAG**TTTCAA	-108	12.97992
9	*pro1527*	*kch*	GTAtaaataccTAC	ATATTTCACAATGAAGAC**TTGTTT**GACAGAT	-198	12.91811
10	*pro1762*	*rnc*	GTAtcgagaggTAC	TGGCTTGGTAAACCCCGG**TTGGGT**GCAAGGC	-218	12.90905
11	*pro0398 pro0399*	*lipB fAA1*	GTAttacaccaTAC	CCTCCATTTTTTCTTCTATAAA**TAACCT**GCA	-587	12.85354
12	*pro0839*	*rpe*	GTGgaaggaatTAC	CTTTCTAGTTGCACCTTCAATTAA**TAGAAT**C	-53	12.81605
13	*pro0737*	-	GTAacagacatTAC	TTAAACTGTATACGGGAC**TACTGC**TTTATTC	-79	12.74994
14	*pro1317*	-	GTAatctcattTAC	ATCTAATAAATAATAATAATTT**TACTAT**TAT	-181	12.69293
15	*pro0811 pro0812*	- -	GTCattagtaaTAC	CAATACACCTCAATGCAC**TAACAC**CATTAGC	-229	12.66551
16	*pro0020 pro0021 pro0022*	*- murB murC*	GTCaatgatccTAC	AGAAAAGTGGCGCAATAAC**TAAATT**TTTTGA	-139	12.56876
17	*pro1446*	-	GTAattttttttTAC	AGCACAAACCATTGATTATTAA**TACATT**ATT	-91	12.55323
18	*pro1614*	-	GTAattgatgaTAC	ATATTTGTGGTTTTAGTTT**TAAATT**TCCTAA	-129	12.50052
19	*pro1462 pro1463*	*gloB -*	GTAaaatcaaaTTC	ATCTTGCTCCCTTTCAGCTCTTC**TAAAGT**AA	-46	12.37224
20	*pro1488 pro1489*	- -	GTGactgagcaAAC	TAATTAATTTCTGTTGAGA**TAATCC**CTACAT	-136	12.37189
21	*pro0197 pro0198*	*ndhH fcbC*	GTTaatagctaTAC	GCAATGCTTGAAAAGTTC**TAGTTG**CAGGATG	-701	12.35928
22	*pro0334*	*pcnB*	GTTgaggaaatTAC	TTTTTAAAATGATCGAAGAACAACC**TATCTT**	-40	12.32315
23	*pro1453*	-	GTAtaaaatatTAC	TAAATATAGTTGCAATCAAT**TGCCCT**GATTA	-542	12.28661
24	*pro0783*	*pcbA*	GTTtatttttaTAC	TTTCAGGTTCTAATTGATG**TAGAAA**ATTATC	-561	12.18944
25	*pro1037*	-	GTAtcaataagTAC	TTTTTGTGACTATTGGCCTCCATT**TCCCAT**T	-170	12.18029

[1] Bold, -10 like boxes

Table s4. Predicted NtcA promoters in PCC7421

Rank	Trascription unit	Name	NtcA site	Downstream of NtcA binding site and -10 like box[1]	NtcA site position	Score
1	*gvip146*	*glnA*	GTAtctcagacTAC	AAAACACCTCCCCGACCCTCCT**TAGTAT**GGG	-103	14.53822
2	*gll2499 gll2500 gll2501*	- - -	GTAgtcccggcTAC	AAAACTGCGCGAATCGTAATCGTA**TACCAG**T	-49	14.20753
3	*gvip213 glr1559*	*ntcB* -	GTAaccccagaTAC	ATTCGTGGCCGTGCGGGCCGGA**TATGGT**CTG	-33	14.10536
4	*glr3061*	*amt1*	GTAgtcattgaTAC	AAATGAGTCGCGACAGATTCGCT**TAACTC**TC	-88	13.75101
5	*glr2690 glr2691*	- -	GGAgcggcggaTAC	GGCGGTTCTTGCGGACTGT**TAACCT**GAGGGC	-57	13.64837
6	*glr2602*	-	GTTtattacgcTAC	ATTTTAATTGGGGAGCGATT**TATAAT**GGGTC	-57	13.63074
7	*glr0369*	-	GTAgttacggcTAC	TCCTGATTGAGCAGCAGTT**TATCTT**ATACCC	-323	13.56199
8	*gvip454*	*ntcA*	GTAgtccctaaTAC	AAAAAAATTGCAGGGCTCGGCT**TAATTG**TTA	-62	13.30728
9	*gvip212*	*nirA*	GTAtctggggtTAC	GTAAAGTTAAGTTCCTTCAAAGCAT**TCGTGG**	-131	13.22382
10	*gvip021*	*glnB*	GTGtgcaacttTAC	CCTCGTCGGGGAGCAATCGC**TAGCGG**TGGCT	-142	12.93903
11	*gll1330*	-	GTAgtgaggatGAC	GTCGGTGACGGCCACACGCCCG**TAGAAC**AAC	-419	12.8756
12	*gll1734 gll1735*	- -	GTActtttctgTAC	TTTTAGTCATTGTCTGAAAACC**TATGCT**AGG	-268	12.70844
13	*gll2867*	-	GTGgagccgacTAC	TATCGTTGCGCAGGGAAG**TATTTT**CGCTGTA	-585	12.69388
14	*gll1254 gll1255 gsl1256*	- - -	GTAgcagttttTAC	CATTTTTAACAATGATTTGCT**TCCCGG**TACA	-36	12.47082
15	*glr1973 glr1974 glr1975*	- - -	GTAatcctgaaCAC	TCAACCATTGCTTTTCAAAACACTT**TCACAA**	-373	12.44562
16	*glr3270*	-	GTAaccgtatcTAC	GCCGATTTTTTGGGGAGACTGAT**TATTTC**TT	-216	12.44111
17	*gll4013*	-	GTTccagacatTAC	AATGGCGGCAATTCCCAGA**TGGGGA**GCGTAC	-32	12.40925
18	*glr0200*	-	GTGacgaagttTAC	ATCCAAGCTTCCAGATGCGT**TACAGT**GTGAA	-294	12.38995
19	*gll1907*	-	GTGtcatcggcTAC	CAAAGTTAGTTTTTCAGCTGCCCG**TAAAAC**G	-105	12.33962
20	*gll0937*	-	GTGagctaaaaTAC	ACGGCCAAAGTTTGAGACCG**TATATT**CTCAT	-132	12.33636
21	*gll1029*	-	GTAgcaaagacTAC	GCGCGTGTTCACCAATGTAGG**TAACGA**CCCG	-610	12.325
22	*gvip214 gvip215 gvip216 gvip217 gvip218 glr1572*	*nrtA nrtB nrtC nrtD narB* -	GTAtcagaaaaCAC	GATTAGGCCGCGGTACAGCGTG**TTTCCT**GAT	-64	12.30207
23	*glr1552*	-	GTGtcgcacgaTAC	GACTTCCGCATCTTTCGCGAGCAT**TAATTC**A	-236	12.23584
24	*glr2112*	-	GTAgcgtaagcTAC	GGATTCTTGTTTTGTTGTG**TATCTT**TAGATA	-107	12.19255
25	*gll0158*	-	GTTgtatgagtTAT	AGCGGTGGCCCGGTGAAAAAT**TAACCG**CAT	-31	12.1132
26	*gll2294 gvip312*	- *petH*	GTGgcgatggtTAC	TTTCTTTGCTATCTATAG**TAAGGT**AAAAATA	-218	12.04155
27	*gll2225 gll2226*	- -	GTAgcgttgccTAC	GATGAACACAGTTTTGTGATGG**TAGCTG**GCG	-224	12.03374
28	*glr0398*	-	GTCcggcggcgGAC	GGCGACTGGCCTGCGCTTTCGG**TCAGGG**GGA	-86	12.03244
29	*gll3955*	-	GACggatcgccTAC	CGGCACCATTGGTCTCGTCATGC**TTGGAT**GA	-146	12.00524
30	*gll0878*	-	GTTagtttcggTAC	AGCTTGGGAAAACCCACTGCGT**TGAGTT**GTC	-105	11.99292
31	*gll0368*	-	GTAagtatcaaTAC	GAAGGCAAACACCAGGATGAACGCA**TTGAAA**	-176	11.95922
32	*gll1655*	-	GTAagcgccccTAC	CTGATTGCTCTGATTGTACAG**TAAGCC**CAGG	-168	11.90786
33	*glr1664 glr1665*	- -	GTGtccccagaTAC	GTTCGATCGATTCGCGGAAGC**TACAATAT**GG	-58	11.89084
34	*glr0719*	-	GTAcgcattttGAC	TGCGATCCATTTGAGAAC**TAACCT**GAGGGGT	-64	11.87019
35	*gvip453*	*purA*	GTAgttctcccTAC	CGCCCTGGAGCGAGCGACGAG**TAAGCG**ACAA	-469	11.85558
36	*glr2944 glr2945 gsr2946*	- - -	GTGttggtggcTAC	GTCCAGCCACACTGCCGG**TATTTT**TAGA	-29	11.84654
37	*gll1477*	-	GTCggcaggagGAC	GCTGTGCTCCAGGCGGCCTATGGGT**TAATCT**	-37	11.73635
38	*glr2787 glr2788 glr2789*	- - -	GTTtggccgcgGAC	TGGCGTGTCTTCATCTTAGGCA**TGGTGT**TAC	-524	11.73536
39	*glr0098*	-	GTTtggacggcTAC	CCGGAATGATGTTGCAGATCCGT**TACACT**GT	-58	11.71579
40	*glr1771*	-	GTAgaggcaggGAC	TTTTCTTCCATCCTCTGAAGGT**TCTGAT**ACT	-245	11.71145
41	*gll1545 gll1546*	- -	GTGcccgcccgCAC	TGGGAGGGTCGGGGTTGAGCCGT**TACACT**GA	-69	11.69152
42	*glr2813*	-	GTTtgcagtcgTTC	CGAGAGCATGGCACCGGCGC**TGCATT**GCTAC	-58	11.68814
43	*gll3195*	-	GTAgcaaaagcTAC	AGTTGTCTGGATAGTGTA**TCAAAA**TGTAAAG	-38	11.67703
44	*gvip395*	*thrS*	GTAcccagcaaTAC	GCAAACGCATTGCATCGTAGTGGAG**TGACAT**	-59	11.66554
45	*gll0203*	*rpoD*	GTAccgaccacCAC	CCCCGCGCTCGAGTAGACCAC**CAACAC**CGCC	-235	11.6371
46	*glr0422*	-	GTAtcctacgtTAC	GATTATTCCTTTTCAAGGCAGAT**TCAACC**CC	-119	11.62389

[1] Bold, -10 like boxes.

Table s5, Zhengchang Su, et al

Table s5. Predicted NtcA promoters in MED4

Rank	Transcription unit	Name	NtcA site	Downstream of NtcA binding site and -10 like box[1]	NtcA site position	Score
1	*pmm0958*	-	GTTgcttttaTAC	ATTTTAAATCATTTTATT**TAGTGT**TTAATTA	-53	15.61409
2	*pmm0920*	*glnA*	GTTacctttgaTAC	ATAATTAGATTCGTCAAGTTAT**TAATTT**TAT	-52	14.60082
3	*pmm0246*	*ntcA*	GTTactgttgaTAC	AAGTTATTTCTTGCTGTCTTCT**TAAGTT**TTT	-50	14.18353
4	*pmm0970*	*urtA*	GTTacctatgcTAC	AAAACTAAATCCCACTCGTTTA**TAACTT**TTA	-48	13.85116
5	*pmm1171*	*isiB*	GTAtaaaaataTAC	ATGAAAAAGTTTTTATAAAAAT**TAATTT**ATC	-72	13.72387
6	*pmm0028 pmm0029*	*pdxA -*	GTAtaattaaaTAC	ATAAAATATATTCAATCCAAGAAA**TATAGT**C	-782	13.62519
7	*pmm0370 pmm0371 pmm0372 pmm0373*	*cynA cynB cynD cynS*	GTAtttgttaaTAC	AACACAAGTGAAATTTAAAA**TAAATA**CTGGA	-55	13.15913
8	*pmm0803*	*futC*	GTAtaaaattaTAC	AAGTGATATTTATGGTGAA**TATTTT**CTGTTA	-64	13.14794
9	*pmm0579*	-	GTTagacctctTAC	TCTTTGTTTTACTAAAGTT**TATCTT**ATCAGT	-39	13.08737
10	*pmm1519 pmm1520*	*psaL psaI*	GTAataaaagtTAC	ACTATTTGTCCTTTAAATAC**TAAAGT**GTTCA	-48	13.02118
11	*pmm0263*	*amt1*	GTAtaaacgacTAC	TAGACATACATAAGTTTTC**TGAGCA**AAATTA	-85	12.92402
12	*pmm0339*	*T8L23.23*	GTTcttagtatCAC	AACAACAACAATCACTTTTGTT**TAAGCT**TTA	-55	12.92013
13	*pmm0619*	*acs*	GTAgttaaaaaTAC	ATTATGATTTTTTGAAAATTATAAT**TAACTT**	-63	12.8663
14	*pmm0141*	-	GTTctcgccaaTAC	AATTATTTGTCTGGAAGCAT**TAAAGT**TTGAT	-501	12.85815
15	*pmm0020 pmm0021 pmm0022*	*- murB murC*	GTTgctgatcaTAC	AGAATTTTGTGCAATAACGTAAG**TATTTT**TT	-153	12.82735
16	*pmm0781*	*cbbA,cfxA,fbaA,fda*	GTAacaattatTAC	TCGATTTTTTAGTGATTTGTATATC**TAATGA**	-112	12.74246
17	*pmm1164*	*futA*	GTAttaattgaTAC	AAGGTTATGTGTAGGAAATTTCTTT**TAGGAT**	-81	12.71619
18	*pmm0709*	*som*	GTAaatttcttAAC	CTTACAAAGTTATAAAAAGT**TTGGAT**TAACA	-58	12.69323
19	*pmm0015*	-	GTGtttgtctgTAC	AATCCATCATTGGAGCAACACT**TATTTT**ATG	-96	12.69013
20	*pmm0278 pmm0279*	- -	GTTtaaattttTAC	TTATATAAATTAATATAGA**TAATTA**TTGGAA	-52	12.59038
21	*pmm0919*	*spt*	GTAtcaaaggtAAC	TTTTTTCAACACTTGCTGGG**TTCTTT**GAGTA	-150	12.58175
22	*pmm0600 pmm0601 pmm0602 pmm0603 pmm0604*	- - - - -	GTCtattaactTAC	CACCTAGTTATGCAAGAAGTCT**TAAATT**GAG	-130	12.57007
23	*pmm1114 pmm1115 pmm1116 pmm1117*	*- crtH gidA psbY*	GTAaaattgagTAC	ATATGAAATTATCCCTTCTAAAGG**TAAAAA**A	-156	12.5626
24	*pmm1059 pmm1060 pmm1061*	*- - trxA*	GTCtgttttgtTAC	TTATATGTTTTTATCTGTAA**TGAAAT**CATCA	-77	12.55726
25	*pmm0650 pmm0651*	- -	GTCtaaacataTTC	TGAATAGAATCTATAGACAATT**TAAATA**GAT	-62	12.55306
26	*pmm0357 pmm0358*	*- thiD*	GTAtcgtcccaTTC	AAATCTATCAGACATTATTGC**TAAAAT**TTAA	-138	12.51801
27	*pmm1462 pmm1463*	*- glnB*	GTAtcaatgatAAC	AATCTTAGTATCGTTAAAAGGGTT**TATATG**T	-52	12.49891

[1] Bold, -10 like boxes.

Table s6, Zhengchang Su, et al

Table s6. Predicted NtcA promoters in MIT9313

Rank	Transcription unit	Name	NtcA site	Downstream of NtcA binding site and -10 like box[1]	NtcA site position
1	*pmt2229*	*urtA*	GTAtcattcacTAC	AGAATCATGCCTGGCTGATCTT**TACGTT**CCC	-80
2	*pmt0601*	*glnA*	GTAcctgttgcTAC	AAAAAGGGGGCCATGCGGCTGAT**TACGGT**CG	-66
3	*pmt1831*	*ntcA*	GTCaccattgcTAC	ATGGTAGTGGAGGTCCACTCCC**TAACTT**CAT	-123
4	*pmt1947 pmt1948*	- -	GTAgccggtacTAC	ATCTGGTGGGCTAGGCAAA**TGCCAC**GCTCTC	-60
5	*pmt2239*	*nirA*	GTTcaatctgaTAC	CGCCAATACCCTCTTCCATCAGGCT**TAATGA**	-470
6	*pmt1127*	*trxA*	GTCacagttccTAC	CTGTGGAGGCTCTTAAGCATGAA**TAAAAT**AC	-73
7	*pmt1023 pmt1024*	- -	GTCttgccggcTAC	TTGCTCAGATCATGCTTGTCGGAG**TGCGTT**C	-464
8	*pmt0496*	*pcbB*	GTAgtaaatatTAC	CTGATAATGTAGCGCCGAATTC**TCGCTT**ATG	-234
9	*pmt0830 pmt0831*	- -	GTTcaactgccTAC	CACACCTAACTACCACAG**TAAAAC**CCCCTCC	-155
10	*pmt1480 pmt1481*	- *glnB*	GTAacaacagcCAC	AACCCAGGCCCGTTGAAAAAGCCC**TAATGG**G	-32
11	*pmt0454 pmt0455*	- -	GTAaggaagagTAC	GGAATTAACTTTTTATGCGTTAT**TAAACA**GT	-170
12	*pmt2246 pmt2247*	*sig* -	GTActctttgcTAC	CTAAAAAGGAATAATCAGCA**TGCAAA**ACAGT	-161
13	*pmt0284*	*som*	GTAtcaattgaTAC	TTAGGGCTGTTAATTGTG**TCCGTT**GTGAACA	-92
14	*pmt0091 pmt0092 pmt0093 pmt0094* *pmt0095 pmt0096 pmt0097 pmt0098* *pmt0099 pmt0100 pmt0101*	- - - - - - - - *hisH hisF* -	GTCcttgccgaTAC	TTAATCATTCCCCTTGCTTGG**TATAAT**TACA	-51
15	*pmt2161 pmt2162 pmt2163*	- - -	GTTtaaaaaacTAC	CAAGCTGACCCTGATATTCAATGCT**TAAGGG**	-440
16	*pmt0600*	*spt*	GTAgcaacaggTAC	CTTCTACTGAAGCCTTGTT**TCCTTT**CTGGGT	-472
17	*pmt0456*	*sigB*	GTActcttcctTAC	GCAGTCAGGATGGAAACACT**TTGGCT**CTAGG	-56
18	*pmt2237 pmt2238*	*cobA* -	GTGtctttgtgTAC	ATTCTGAACCCGGCTGATGT**TGGCAA**CTAGA	-591
19	*pmt0629*	-	GTCtagtcttgAAC	TCGCTCATGTTCATTGCAGA**TAGCTC**TCAAT	-135
20	*pmt0842*	-	GTGcgctgatgAAC	TCCAACACCGTGGAACAGT**TGAGAC**CCTTTT	-632
21	*pmt0034*	-	GTAttactcaaAAC	TGCGTTCCAGCCTGTCTATCCTG**TAACGG**TC	-486
22	*pmt0868 pmt0869*	- -	GTCcatatggtTAC	CTAACCAACCTAACAAGACCCAA**TAATAT**AG	-466
23	*pmt1564*	-	GTGaccaaatcAAC	TATCCAAACAGAAATCTAGC**TATCGA**TAAAA	-175
24	*pmt0743 pmt0744*	*metG* -	GTCtgagagatGAC	GGCAGACGATCAATAAGC**TGAAAT**GAGTTGT	-79
25	*pmt0391*	-	GTTaaaccattGAC	AACTTGCATCCAATCGAAAA**TAACTG**AGCTG	-79
26	*pmt1631*	-	GCAcacctctgTAC	ACTTGTAAGTAAGACTTACTTTTT**TAGGGT**	-31
27	*pmt1839*	*ams1*	GTAattaacgcTAC	CCCTGGCATAAAGGGTGTTGAAGT**TGCTGT**C	-634
28	*pmt1853*	*amt1*	GTAacaaaatgAAC	TTTCACCCCCGCAAATAGACC**TAAAAG**CGTC	-147

[1] Bold, -10 like boxes.

Table s7. Predicted NtcA promoters in PCC7120 (to be continued)

Rank	Transcription unit[1]	Name	Putative ntcA site[2]	Downstream of NtcA binding site and -10 like box[3]	ntcA site position	Score
1	*alr2328*	*glnA*	**GTAacaaagacTAC**	AAAACTGTCTAATGTTTAGAATC**TACGAT**AT	*-129*	*14.619*
2	*alr0607*	*nirA*	**GTAgctacttaTAC**	TATTTTACCTGAGATCCCGACA**TAACCT**TAG	*-498*	*14.515*
3	***all4312***	-	GTAacaaagacTAC	AAAACCTTGGGCATGGGCTTGT**TACTTT**GAA	-60	14.478
4	*all1075*	-	GTAcacacaaaTAC	TATATACAAGGGTTTCAGGCGTTA**TAGCCC**C	-298	14.1783
5	*alr4249*	*sigF*	GTAattcaaaaTAC	TTTTATTAGTTATGATACTTGA**TAAGTT**TCC	*-685*	*14.164*
6	*all0602*	*ntcB*	**GTAacaaaatcTAC**	CAAATTGGGGAGCAAAATCAGC**TAACTT**AAT	*-65*	*14.102*
7	*asr0064*	-	GTAaccttggcTAC	GCAACAGTTTATACCTTTAT**TTTCAT**CATAA	-435	14.0627
8	***alr3982***	-	GTAgtgatagcTAC	AAAATCCGTCGTAATTATA**TAAGTT**CTAGTT	-237	13.9351
9	*all2509*	-	GTAaatgattaTAC	TGATTTTTTCCCAATAAAAGAT**TACCCT**CAT	-713	13.8553
10	*asr4313*	-	GTAgtctttgtTAC	TCTAGCCACAATTACCTC**TACGTT**TTTTTAT	-465	13.8443
11	*all1616*	-	GTAataaatccTAC	AGCCAAAAACACGCTAAACACGC**TAGGCT**A	-31	13.7858
12	*alr4308*	-	GTAgtcggagtTAC	AACATATTGGGATCAGCCAGTGT**TAACTT**TG	-51	13.7468
13	*alr0992*	*amt1*	GTAactaagtaTAC	AAGGTTACATTTCATCAGGGAG**TGGGAA**CAT	-37	13.7253
14	*alr5117*	-	GTAtgtttaatTAC	ACATGATCTCTATATTTTT**TACAGT**AATCAT	-153	13.5645
15	*all1951*	*urtA*	**GTAtcaaaaatAAC**	AATTCAATGGTTAAATATCAAAC**TAATAT**CA	*-101*	*13.446*
16	*all1127*	-	GTAgcttaaatTAC	TCAATGTTTGTCTTTTTGG**TAACTT**TAAGCA	-131	13.438
17	***all5069***	-	GTGttcatgtaTAC	TTTGTTTAAGAATAGCTATC**TATACCT**TAGT	-214	13.4139
18	*alr0543*	-	GTAagcgatcaTAC	TGAAACAAAAAGTTTAAAGTTTTTA**TAACTT**	-221	13.3997
19	***all0706 all0707***	- -	GTGgaaagactTAC	ACCCCATACCCTACCCTCAACAC**TATTGC**CT	-111	13.3546
20	*all3084*	-	GTAataagtgaTAC	TCCCATTACCAATTACCATGAA**TCTGCT**TGG	-190	13.3401
21	*alr5370*	-	GTTtaaaataaGAC	AACAGAGACTACACCCCAATCAA**TAACTT**CG	-67	13.3164
22	*all2319*	*glnB*	GTTacacagacTAC	ATCGAAAGTTTTTGCGGAAAA**TAGCAT**TTTG	-199	13.305
23	*alr4392*	*ntcA*	GTAtaggaaagTAC	AGAAAGGTTAACGGTGCTT**TATTGA**TTTTTC	-98	12.9867
24	*all1248*	-	GTAttatatatTAC	CCATTTTTAAGATAAACAAGTA**TATCGT**CAG	-134	12.9852
25	*all1284*	-	GTAtattagctCAC	AATTCTTTTTTGTCGATAAATCA**TAACTT**AT	-119	12.954
26	*all2327*	*apcF*	GTAgtctttgtTAC	AGAACGTCTGGATTACAGG**TAATAT**TAACCT	-218	12.8922
27	*alr4029 alr4030*	- -	GTAcgcaacacCAC	GTTTAATTATGGGACAAA**TATTAC**TAGAGAA	-105	12.863
28	*alr2825*	-	GTAacgcacccTAC	TTGTATTGCAAAATTCCAAA**TAATGT**TCAAT	-290	12.7889
29	*alr4156*	*ndhF*	GTCttataaaaTAC	ATAGATTTTTGTCTATGATTAGTAT**TAACTT**	-247	12.7842
30	*alr0276*	-	GTTaattatttTAC	CTTCTCCTATTATTCTCT**TAATCT**GACTTCA	-191	12.7679
31	***alr1231***	-	GTAatcatccaTAC	AATCTTATGCAATTATTAAAAACT**TATTTT**G	-42	12.7297
32	*alr4965*	-	GTGctagcctcTAC	GGACAATAGCTATGAATCCT**TATGTG**AGGAT	-511	12.7219
33	*alr3810*	*sigD*	GTAacaaaagcTAC	AATTTGCTGAATTTATAAAGTTTT**TAGTCT**C	-285	12.7079
34	***alr3920***	-	GTAttgtaataTAC	AATATTACCTGTGAATTGAAAT**TATGTG**ATG	-420	12.7052
35	***all3660***	-	GTAttacgaatTAC	AACTGAGAAGTCATCATCTA**TACGTT**TCTCG	-737	12.7033
36	*alr0487*	-	GTAgtagatgcTAC	AAATAGCGATAAAATTAA**TAGGCT**ATTTGCC	-109	12.6945
37	*all2170 all2171*	- -	GTAactttgcaAAC	ATCCTCTTATAAATAATTGATT**TAAAAA**ATA	-106	12.6945
38	*alr2311*	*rbpF*	GTCattgcctaGAC	ACCAACTTTTAATTCCCATCATC**TATCCT**GA	-372	12.6889
39	*all3184*	*lrtA*	GTCtcttttcaTAC	AAAAACACCATCACATCAAA**TATTGT**GCTAT	-48	12.6317
40	*all4950*	-	GTTacttcttcTGC	TCAGAGTTACTCATCAGCCTGTTT**TAAGAC**A	-51	12.6308
41	*all4813*	*pknC*	GAGatatttatTAC	CATCGGCAGAATCTAGGTAA**TAACTC**TGGGG	-38	12.6264
42	*all4006*	-	GTCaaacagctTAC	GATATAAAGCTTTAACCTG**TCAAAT**TTATCT	-369	12.5808
43	*alr2514 alr2515*	*coxB coxA*	GTAccaaaaaaTAC	CGAGAATCACCAAAAACTTCACGG**TAAACA**A	-410	12.5528
44	*alr3280*	*sigE*	GTTggtattatTAC	TAATAATAGAGTCTAATAGAGC**TAGTTT**CTG	-636	12.5423
45	*alr3979*	*pex*	GTAgcgctataTAC	TTACACATTTGCAAGTAGCGCCC**TATACT**GA	-86	12.5394
46	*all0909 all0910 all0911*	- - -	GTTatcctagtTAC	GGATAGGCATTATTTCTCATAAA**TAAGCT**CT	-349	12.5238
47	*asr1275 alr1276*	- -	GTAgctatactTAC	ATAAAACCCACTTTTGTTAT**TAACAA**CACAT	-681	12.5183
48	*all1020 all1021*	- -	GTActtttataTAC	AAAGTTATGTGGTTTACA**TAAGAT**ACGCATA	-90	12.5175
49	*alr2877*	*cmpA*	GTAagtataaaAAC	TATATCTATAGTTTTAGATGA**TAATTT**GCAT	-480	12.499
50	*all2364*	*accD*	GTAacgttcatTAC	CCCATTGATACTTGAGATTTTGA**TAATAT**CA	-589	12.4973
51	*alr0165*	-	GTTtttgctatTAC	CATCTTGAATATTTAAATAAA**TAAAAT**CATC	-240	12.4797
52	*all5185*	-	GTTttcagcgtTAC	TTTAAGTAGTTAAGGATAATGTCA**TACAAC**A	-157	12.459
53	*all3680*	-	GTTaacgcctttTAC	AGTAAACCAGTATTGGATTTTGAA**TAGACT**A	-643	12.4574

Table s7. Predicted NtcA promoters in PCC7120 (continued)

Rank	Transcription unit^1	Name	Putative ntcA site^2	Downstream of NtcA binding site and -10 like box^3	ntcA site position	Score
54	*alr3376*	-	GTTtttcccatTAC	CTATGACATTAGATCCAAAAGT**TAATTT**ATT	-177	12.44031
55	*alr0021*	*apcA*	GTCacttctttTAC	TGCTGTGTCTACAAGGCGAGAGT**TAATTC**TC	-104	12.40309
56	***all0330***	-	GTTgtaatttgTAC	AATGCAGCTGTGACGTGGATC**TAGTTC**AGAT	-52	12.38822
57	*alr0440*	-	GTGttttcagtTAC	AGTTATGCCAGATGCAAT**TAAGCC**ACAATGT	-83	12.37565
58	*all1327*	-	GTAgtgtatgtTAC	AGAAAAAATTAATTTAATATAGA**TATTTT**TA	-37	12.36931
59	*all3375*	-	GTTagtactcaTAC	ATCTAGACAATAGATGTAATG**TAACTT**TGGT	-84	12.34861
60	*alr0080*	-	GTTtttaaaatTAC	CTACAATCAATTGCCAAA**TCCACA**GCTATTA	-310	12.34684
61	*all3040*	-	GTTgtataatcTAC	TTTTTGAATTTTGAATTTTGAATTT**TAAATT**	-118	12.3115
62	*all4107*	-	GTAtcgagaggTAC	TGGCTCGGTAAACCCCGGT**TGGGAG**CAAGGC	-461	12.30874
63	*all0041 all0042*	- -	GTAtttatctaTAC	AAAAAAACATTACAAATGA**TGAGGT**AGTATC	-735	12.29755
64	*alr3882 alr3883*	- -	GTActtttagaTAC	TATTGTGAAAGCTCATCTG**TACTTT**TACCCT	-98	12.29166
65	*alr2785*	-	GTGaaatcctaTAC	CCTTTTAGGGATGGGATGAGCT**TAAAGT**TAT	-78	12.28042
66	*alr4239*	-	GTTggagggtaTAC	ATTTTGATCTGAATAATTTCAGAT**TAGTTT**A	-215	12.27095
67	*alr2471*	-	GTAcctgtggcTAC	ACATTGTTACTGATTTTAA**TCAAAA**TTGCTG	-164	12.25365
68	*all0866 alr0867*	*ccmL ccmK*	GTAgtttcttTAC	ACTGCGCTGTGCCAAAAGCCGAG**TAAAAA**AA	-105	12.23615
69	*all4756*	-	GTAgaccgtcgCAC	CCTAGGTCATTAAAGACCGCTTTG**TAGCAT**G	-188	12.2266
70	*all1601*	-	GTAgtcagtcgTAC	TAACCACTGAAGACAAGAAC**TAGTCA**AGAAG	-248	12.21125
71	*all1235*	-	GTAaaaacaacTAC	TCCAAGTACTTCTATTGTG**TAAACT**TTGAGA	-165	12.20566
72	*alr4641*	-	GTCcccaatgcTAC	GACTTTTCCTGAGACTTC**TAAACT**AGCTTCT	-386	12.20396
73	*alr0786*	-	GTAttgatcaaTAC	TGGGGGTTAACCCATCTATCCGCAA**TAATTT**	-60	12.1998
74	*all0187*	-	GTGcggtcttaCAC	TGGCAAGATTGGTAGCTTGCTT**TAACAT**AAA	-62	12.19843
75	*alr0038 alr0039*	*rfbB* -	GTAatcaacagTAC	TACCGATAGCAGAAAGGCGCTGGT**TAGCAT**C	-630	12.19774
76	*alr3718*	-	GTGattgttcaCAC	AACAACTTCCAAGCATGGAGGG**TATCTG**CTG	-755	12.19582
77	*all3538*	-	GTAatcgatttTAC	TCGGTTCAATAATTATCTCCAGT**TAAAAA**AT	-231	12.17417
78	*all3116*	-	GTAggggatgtTAC	GGCAAAATTAGCCAAAGCGATCGC**TCAATT**G	-196	12.09504
79	*all3582*	-	GTTtagtcttgTAC	ACTAGAACAAAGACCTCT**TAACTG**AAGTCTT	-79	12.09369
80	*alr2921 alr2922*	- -	GTGtgtcttgaTAC	AATACTTTTAAGATGACTGTTA**TATGCT**TTC	-300	12.08668
81	*alr3705*	-	GTTttaggtgaTAC	TAAAGGCGATCGCAATATTT**TAATAT**GAGCG	-42	12.06532
82	*alr3738*	-	GTTagccctttAAC	AGTTTAGCCTCATGAAAGC**TAACTT**AGTATT	-54	12.05389
83	*all3408 asl3409 all3410*	- - -	GTAaatagaaaCAC	ACTTAGTAAGTGATTTTGTTTC**TAAATA**TGT	-464	12.05173
84	*alr3710 alr3711*	*devB devC*	**GTAcagtctgtTAC**	CTTTACCTGAAACAGATGAATG**TAGAAT**TTA	-738	12.04803
85	*all4236*	-	GTTctagccaaTAC	TTCGCAATAGAGATGCACAA**TAATGG**CTCTC	-418	12.04208
86	*alr0892 alr0893*	- -	GTTaggagtttTTC	ATGCCCAGTAAAACTGATCAA**TAATTT**TTAA	-124	12.03502
87	*all3903*	-	GTCtactatacTAC	TATATAGTAATATAATATT**TATGTT**TTATAA	-81	12.02717
88	***alr1712***	-	GTGataaaaaaTAC	CTTAATTTCGATAATAAACATTGA**TAAACA**T	-511	12.0257
89	*alr1537*	-	GTTacaaattaTAC	CAATCAAAAGCAGTATTACAAAC**TATTTG**CT	-107	12.02084
90	*alr3790*	-	GACaatattcaTAC	CACAGCCAGAGAAAATAAACGGA**TAAACT**TA	-75	12.01766
91	*alr1785*	-	GTAtgtattaaTAC	TGAAGTTTTAGTATAAAGTTTTTGC**TAGGTT**	-59	12.01261
92	*all1291*	*cynS*	GTAgaagagggAAC	TACAATGACCTACGATCACC**TACGGT**GGAGC	-107	11.99881
93	***all1070 all1071***	- -	GTTaaacgtggTAC	ATAATTATTTATAAGAAATAAG**TAAATT**TCA	-187	11.99609
94	*alr3874*	-	GTTcggttaatTAC	CCAGCAAATACGCATTTA**TAGAGT**TGTTAAA	-32	11.99544
95	*alr1240 alr1241*	*zam* -	GTAttttccccTAC	ATAGCGCTTTTGGTCAGCG**TAGAGA**GTAAAT	-385	11.9935
96	*all4248*	-	GTAttttgaatTAC	AATTTTATTCTAAAGAAAGAAAAAT**TAAAAT**	-103	11.97364
97	*all5097*	-	GTCaaaagtcaAAC	CTCTTGACCTTTGACTATTGAATTT**TGAAAA**	-309	11.95972
98	*alr0522*	-	GTTcagcaattTAC	AGTTAAGTCAACACAGTTAGTCTT**TATTTT**T	-722	11.94098
99	*all1731*	-	GTTcagataagTAC	TAATATCTGTTGGTATCTACCAGC**TAAAAT**C	-138	11.92767
100	***alr0295 alr0296***	- -	GTAataactatTAC	TCAATAGATTAGGCTAAATA**TAGCGT**TTTAA	-71	11.92562
101	*alr3955*	-	GTCaaaagcggTAC	TGAAATATAGCTGCTATAAAACAT**TAACTA**G	-95	11.92449
102	*alr1250*	-	GTAatatataaTAC	ATGAATAAATCTAAAATACAGG**TATGGT**GAG	-309	11.92065
103	*alr4354*	-	GTCctcgatttTAC	TGTGCGAGGATTTTAGTAAAGACT**TAAATG**A	-160	11.91353
104	*all0443*	-	GTGacaaaagtCAC	TAATATTTGCTTATTTATATATGAC**TATAGT**	-480	11.91314
105	*alr1602*	*devB*	GTAcgactgacTAC	TGACAAGCAGCTAGAGAAAGTGA**TAAGGA**TA	-151	11.90962
106	*all1837 all1838 asl1839 asl1840 all1841 all1842 all1843*	- - - - - - -	GTAccggatacCAC	CTATTCTTGATATGAAATCCGTG**TAAACT**CA	-250	11.90219

[1] Bold, putative sensor kinase or response regulator; [2]bold, experimentally verified binding sites; [3]bold, putative -10 like boxes .

Table s8. Predicted NtcA promoters in PCC6301

Rank	Transcription unit	Name	NtcA site	Downstream of NtcA binding site and -10 like box[1]	NtcA site position	Score
1	*syc1985_d syc1986_d syc1987_d* *syc1988_d syc1989_d*	- *cynA cynB* *cynD cynS*	GTAacgacggcTAC	ATTTTGACCCTGGGGTTACTAC**TACCAT**TCG	-54	15.0095
2	*syc1638_c*	-	GTAaaggcgaaTAC	AGAAGCCACAATGGACAGCTTGC**TAGGTT**AA	-146	14.7153
3	*syc1639_d*	-	GTAttcgccttTAC	GAGCGCGACTTCGAGCACCCGT**TGTCTT**CCT	-174	14.5871
4	*syc0921_d syc0922_d*	*glgC* -	GTAgccttggtTAC	CTCCGATCGATCCCGGAGCTGCG**TAGGGT**AA	-121	14.5638
5	*syc1821_c*	*amt1*	GTAgcaaaagtTAC	GTATATCACCAGTCTGCC**TAGCCA**GAGTTGT	-72	14.5596
6	*syc2377_c syc2378_c*	*metX* -	GTAgcgatcgcTAC	AGCAGCAACAAAGATTGACACGA**TTGGTT**AG	-116	13.9901
7	*syc0310_d*	*nirA*	GTAgcaattgcTAC	TAAAAACTGCGATCGCTGC**TGAAAT**GAGCTG	-207	13.9832
8	*syc0700_d syc0701_d syc0702_d* *syc0703_d syc0704_d*	- *merR* - - -	GTAgtcaccgtTAC	AGTCATTCCTTAGAACTTGTTT**TAGTTA**ATT	-80	13.8317
9	*syc1192_d*	*glnB*	GTAgcagtaacTAC	AACTGTGGTCTAGTCAGCGGTGT**TACCAAA**G	-90	13.7626
10	*syc1377_c*	*ntcA*	GTAgcagttgcTAC	AAGCAGCAGCTAGGCTAGGCCG**TACGGT**AAC	-142	13.5778
11	*syc0699_c*	-	GTAacggtgacTAC	TAAATGATTGCCTGTTTAACCTGAC**TTAATT**	-105	13.5477
12	*syc0006_c*	-	GTAacaaaagcTAC	ACACCCGCAGCAATCCTTAACAAA**TAATTA**A	-67	13.429
13	*syc0513_c syc0514_c syc0515_c*	- - -	GTAtcccgaacTAC	AGAAGTGGACTCTGAGCGATTTC**TATAGTC**C	-60	13.3282
14	*syc2346_d*	-	GTAgcaaataaTAC	AAAATGGCACCGCGATCGC**TCAGAC**GGCTGC	-98	13.3005
15	*syc0308_c syc0309_c*	*ntcB nirB*	GTAacagaaacTAC	AACTTTTTGTAAATCATGCA**TGTTCT**GCAAG	-207	13.1924
16	*syc2474_d*	*cmpA*	GTAaatgtttcTAC	ATGCAAAACGCAAAGTCTATTGACT**TAGCTC**	-206	13.0616
17	*syc0649_c*	-	GTAactggataCAC	AGCTTGCCCCTAGAGCGGGTGT**TAGCGT**GAT	-67	13.0216
18	*syc0953_d syc0954_d syc0955_d*	*rpoD4 codA* -	GTAtcccaaacTAC	TAAATGCAACAAAGACTGTCATC**TACTTA**AC	-146	12.9658
19	*syc0508_d*	-	GTAgccttgctTAC	ACTTTGCCGATCTGGCCTCTTC**TAGGAT**GCC	-93	12.7814
20	*syc0545_d*	-	GTAgcctctgtTAC	TGTTCATTAGACGTTCAA**TGAAAT**CCCTAAA	-242	12.7237
21	*syc2196_d*	-	GTTctgatattTAC	CAAAGCGATCGGCGTAGG**TGGTGT**CGCAGGG	-554	12.5616
22	*syc1440_d syc1441_d*	*rfbF rfbG*	GTGattccaaaTAC	TGTCTTAGGTTGCAAAGT**TAATGT**TATTTAA	-125	12.424
23	*syc2140_d syc2141_d*	*pyrG* -	GTAactcacgcTAC	CGAGCAGAACAGTCCGGAACAGGGC**TCGGCG**	-275	12.4066
24	*syc0416_d syc0417_d*	- -	GTAcacatttaAAC	TATTTTATTTGCGATTATATTTCTT**TAGTTT**	-413	12.399
25	*syc0002_c*	*isiA*	GTCggctagaaTAC	CTAGAGTTCGTCGTAAAACT**TAATAT**CAACT	-77	12.2816
26	*syc1620_d*	-	GTGtagctagtTAC	GGATCAGGTTGACTAAAG**TCCGGC**AGAATAG	-66	12.2728
27	*syc2410_d*	*cysA*	GTAtcatgccaGAC	ATAAGGTGCAACTAAATTCTCC**TAAATT**TCC	-71	12.242
28	*syc1403_c*	*ndbA*	GTGacttaggcTAC	GGTCGTGCCTTACAATTCTTGACC**TACTGC**G	-104	12.1682
29	*syc0172_d*	*ppa*	GTTgctccgaaTAC	CAGCGCACTAAGCTCAATCCCAAAG**TACTGC**	-185	12.1667
30	*syc1049_d syc1050_d syc1051_d*	*pgm secG* -	GTAagcctcaaTAC	AATTTGGGGGATCGAGCAAT**TACAGC**TTGGG	-373	12.1515
31	*syc0822_c*	-	GTCgttttctaTAC	TGTTGTTACAAGCTTCTGACAATC**TAGGTT**A	-37	12.146
32	*syc1389_d*	-	GTAgcgatcgcCAC	AAATATTTACAGAATATTGGC**TAAGCT**GTGG	-58	12.1316
33	*syc0512_d*	-	GTAgctgtagcTAC	TGTTTGCGCTGGCAGCGC**TAGGGT**TGTTGCA	-123	12.1236
34	*syc1918_c*	*natD*	GTAggtaccgaGAC	TGCCTTGAACTTGGGGGTGAA**TGGCTT**TAAG	-503	12.1178
35	*syc1076_c*	*amt1*	GTTacatcgatTAC	AAAACAACCTTGAGTCTCGC**TGAATG**CTTAC	-112	12.092
36	*syc0509_c syc0510_c syc0511_c*	- - -	GTAgctacagcTAC	GAATTGAAAGCTGGTGCCAGCC**TAGCTT**GGA	-58	12.0575
37	*syc0934_c*	-	GTAatcgattcTAC	CGAGCGGACGCCAATGCTT**TAAGCA**TCGTTA	-96	12.0356
38	*syc2171_d*	*sigG*	GTTgcccagccCAC	ATCCCAGTGGCACCGCGATCGGC**TACTGT**TG	-221	11.9328
39	*syc2184_d syc2185_d syc2186_d*	*zam ubiX* -	GTTcacctgtcTAC	TCCGGCTGGCTATTCTCGTTGCCT**TAGATT**G	-171	11.9236
40	*syc0204_d*	-	GTTggcgaactGAC	AGACCTCGCGATGCAGTTGCGCG**TAGGTC**AG	-380	11.9233

[1] Bold, -10 like boxes.

Table s9. Predicted NtcA promoters in PCC6803

Rank	Transcription unit	Name	NtcA site[1]	Downstream of NtcA binding site and -10 like box[2]	NtcA site position	Score
1	*sll1330*	-	GTAactgttgtTAC	AAAGCCTTGACATTGACTTTGT**TAGATT**AAC	-60	14.7137
2	*slr1756*	*glnA*	***GTAgcgaaaaaTAC***	*ATTTTCTAACTACTTGACTCTT* ***TACGAT*** *GGA*	*-85*	*14.5902*
3	*slr0447*	*urtA*	GTAtcctatgcTAC	ATAATTTTTTCTTAACCTACTCG**TAGTAT**GG	-198	14.5627
4	*slr0898*	*nirA*	GTAatttacgtTAC	AAATTTTAACGAAACGGGAACCC**TATATT**GA	-61	14.523
5	*ssr1562*	-	GTAatcgataaTAC	AGCAAAGCCCCCCTTTGTTTG**TATTTG**GGGT	-65	14.1297
6	*slr0801*	-	GTAgcgaaacaTAC	AAGTGTTAGAAATTTCGAGGCT**TACAGT**TGA	-73	14.0459
7	*sll1294*	*pilJ*	GTCttaatcctTAC	TTTTTCACCTTAACAGAGTCT**TAATTT**TCAT	-92	13.9365
8	*ssl0707*	*glnB*	***GTActgattttTAC***	*AAAAAAACTTTTGGAGAACATGT* ***TAAAAGT*** *G*	*-169*	*13.9072*
9	*sll0108*	*amt1*	***GTAgtaaatcaTAC***	*AGAAAACAATCATGTAAAAATTGAA* ***TACTCT***	*-174*	*13.8698*
10	*slr1597*	-	GTAtccattttTAC	GGTTTTTCAAGATCATGTCAAGGAA**TGGTCA**	-226	13.8343
11	*sll0783*	-	GTAtgtttcgcTAC	AAATTGATGATCTCTGATTCTAT**TAATAG**GA	-379	13.7161
12	*sll1689*	*rpoD2-V/sigE*	***GTAtcacgaatTAC***	*ACTGCCGTGAAAATTTAACGA* ***TATTTT*** *GGAC*	*-296*	*13.5867*
13	*sll1291*	-	GTAattagggaTAC	TTAACTAAAAAGAGTTGAACC**TATATC**AAAA	-342	13.4693
14	*sll1077*	*speB*	GTAgcacagaaCAC	AGTTAACTTAATCATCCCTAGC**TAGGGT**AGG	-226	13.2741
15	*slr1147*	-	GTAatgatagtTAC	AAAAGCCGTGCCGATTATA**TCCGGT**CTAATT	-201	13.2353
16	*slr0851*	*ndh*	GTAacaaccatTAC	CGGTAATCCCCAAGGAAATGAT**TACCAT**TAA	-59	13.153
17	*slr0789 slr0790*	- *umuC*	GTAaagtaagcTAC	TTTTTATTTGTCTTTTTGGCA**AAACGT**TATG	-186	13.1432
18	*slr1686*	-	GTAcaatgggaAAC	GATTAATCAAATGCACGCA**TAACTT**GAA	-29	13.0763
19	*slr0623*	*trxA*	GTCccaacctgTAC	AATGAAGAGGGCGGAGACG**TAAGTT**TCCGTT	-232	13.019
20	*slr0949 slr0950*	*urtB* -	GTAgggtagctTAC	GGCGACATCGCCCTTGGTCTCACGT**TAGGAA**	-36	12.9839
21	*slr0054*	*dgkA*	GTAatccaggtTAC	ATAGGAGAATTGTTCTGTCGAC**TACACT**GCT	-47	12.8624
22	*sll1699 sll1698*	- -	GTCaaataggcTAC	ATAAAGCCTACCGATCCGGCCT**TAACCT**GAA	-248	12.8554
23	*sll0837*	-	GTGttgggcccTAC	CCTATTTAACCTAATGAT**TAGGTT**TTGCTTT	-84	12.699
24	*sll1624 sll1623*	- -	GTAcctaaaccCAC	CAAGGCCGGACCATTTTCCTGCCC**TAATCC**A	-114	12.6251
25	*sll0947*	*lrtA*	GTCttattcaaTAC	ATAGTGCTAATCTGAAGA**TAGTCT**TAGGAGT	-331	12.6022
26	*slr0257*	*ctpB*	GTTtttcttatAAC	TAAAGTTTTAATTAATCAGTCAACT**TAATCT**	-270	12.5681
27	*slr1622*	*ppa*	GTTcttctaatTAC	TCGCAATTGGGGATAGGTGTCGCT**TAAGTT**A	-238	12.5541
28	*sll0807*	*cfxE*	GTTtgggggccTAC	TACAATTGACCGTAATTGTT**TACACT**GAATT	-72	12.5109
29	*slr1849*	*merA*	GTTttttcattTAC	ATTCACGTATCCATCGTTATC**TAAGAA**AGCG	-286	12.4384
30	*r1923 slr1924 slr192*	- - *cbiB*	GTTgttgaacgTAC	CTTTGCTTGGTTTGGTCGC**TACCGT**CGTTTA	-279	12.4196
31	*slr1906 slr1907*	- -	GTActtctcatTAC	TTTCTTGAAACTCTCTAAAT**TCCCGG**CAAGG	-252	12.3395
32	*ssr2047*	-	GTAaagtctttTCC	TCAATAATCCTAAAACAACAC**TAAAAC**TGAA	-325	12.3203
33	*slr0670*	-	GTActagcaacAAC	CATGCCAGCGGAGGGCAGA**TTTTAT**TA	-28	12.3202
34	*sll0638*	-	GTAgtgattccCAC	CGTAGTCAAAGAAAATCT**TAATCA**GGAATCG	-53	12.2868
35	*sll1566*	*otsA*	GTAattagtgtTAC	GGCATTTTACTGACGGGTTAAG**TAATCT**TTA	-269	12.2712
36	*slr1379 slr1380*	*cydA cydB*	GTAgctgacaaTAC	AGAGACTTTACCTGGGATTTCTGT**TAGTTT**T	-115	12.2343
37	*sll1247*	-	GTGatctagatCAC	CTCCCATGGGTGCTTGCATCACT**TACATT**GG	-152	12.2249
38	*slr0935 slr0936*	- *nadC*	GTAcaagatacTAC	GGGATTATCCCTACCTCACTCCGCA**TCCCAG**	-674	12.2247
39	*slr1428*	-	GTAtatctagaGAC	AAATCCTGGTCGGGATAGGG**TAGTCA**AGTAG	-140	12.189
40	*sll0261*	-	GTGacattttTTC	CTTCAGATTCCCCTTTTTT**TAGGAC**GT	-28	12.1721
41	*slr1919*	-	GTCgcggttttTAC	ATTTCTTTAAGTTTTAGGTAGT**TAAAAA**TTC	-66	12.1434
42	*slr1697*	*pknA*	GTTtcctttcTAC	TGATTTGTATTAATGCTTGCT**TAGTTT**GCGA	-269	12.1168
43	*sll0900 sll0899*	*hisG glmU*	GTAttatcgatTAC	AATTCTTTGCACTTCCGTA**TAAATA**CTTAAT	-165	12.1068

[1] Bold, experimentally verified binding sites; [2] bold, -10 like boxes.

Table s10. Predicted NtcA promoters in thermosynechococcus

Rank	Transcription unit	Name	NtcA site	Downstream of NtcA binding site and -10 like box[1]	NtcA site position	Score
1	*tlr1330*	-	GTAttattcgtTAC	GAAATGATAGGTATTAGATTTGC**TAGATT**AG	-125	14.75439
2	*tlr1349*	*nirA*	GTAgcaaaattTAC	AAATGTTCATGATTCATCTGGC**TAAATT**GGA	-69	14.74845
3	*tlr1120*	*urtA*	GTAgcctagcaTAC	AGTTTGCCCCAATAACCTCCTCT**TAAGCT**AT	-68	14.69383
4	*tsr2448*	-	GTAttggatatTAC	AATTTGCCAATCCTGAATCC**TAGGGT**CTCTT	-50	14.60239
5	*tll1021 tll1022 tll1023 tll1024 tll1025*	- - - - -	GTTaagcgggaTAC	ATTTTTTGGGAAGGCATGAGCTT**TAAAAT**GT	-352	13.94199
6	*tlr0592 tlr0593*	- -	GTAgcaatagcTAC	ATTATCTCTGATCGGCGAGCGA**TCGCCC**AAG	-79	13.8677
7	*tlr1468*	-	GTAgttgagttTAC	AACTGCATCGAATTGACAATGT**TAGGCT**AAC	-72	13.82521
8	*tll0590 tll0591*	*aroE glnB*	GTAgctattgcTAC	ACGCTATACTCTTTTGCCAATGC**TAAATA**AA	-65	13.76371
9	*tll1985*	*amt*	GTAtaacgtgaTAC	AGATTTTCCAATCTGAGAACTGG**TGCAAT**CA	-175	13.61188
10	*tlr1254*	-	GTGtccaaaaaTAC	AACGACTCTCGCTCAAATCGGT**TATGCT**AGT	-118	13.51412
11	*tlr0066*	-	GTTtgcttctcTAC	AATATCAGGGTTACGCTC**TTTGCT**GCTGCGA	-77	13.50423
12	*tll2090 tll2091*	- -	GTGaagaaacaTAC	AAATTGAGAGCCCAACCAAGGTC**TAAATT**AA	-89	13.37053
13	*tll2446 tll2447*	- -	GTAatatccaaTAC	TGAACTATTTTAATTTTTGCCC**TAGGGA**GCT	-175	13.17517
14	*tll1357 tll1358 tll1359*	- - *ntcB*	GTAtcaatcaaCAC	ACTTTACCATGGGAAAGTTGCA**TAGATT**GAA	-107	13.16372
15	*tll0306 tll0307*	- *clpC*	GTAaattggcaTAC	AAAAACACGCCAAAGTGGACGTGT**TAAGAT**G	-69	13.1492
16	*tll1975*	-	GTAgagcagagGAC	TGAAAATCCTCGTGTCGCCAGT**TCAATT**CTG	-52	13.02452
17	*tll1659*	-	GTAtagaaaacTAC	ACATAAGCCCTGCTAATTAA**TAAAAT**TCTTA	-173	12.97041
18	*tlr1592 tlr1593 tlr1594 tlr1595*	- - *ycf53* -	GTTattatcgcTAC	ATTTTGTCAATTTAAGACGTGA**TAAAAT**TTA	-175	12.78389
19	*tll0830 tll0831*	- *sigB*	GTAgattaggaTAC	AGGCTAACTCCCTCTCCCCTGC**TAGGCT**AAC	-195	12.71235
20	*tll2333 tll2334 tll2335 tll2336*	- *pilin* - -	GTAagtttgttTAC	ACTGAAGGCAAGTGTGGA**TCAACT**GGATAAG	-309	12.61023
21	*tll1334*	-	GTAgtcgtggaTAC	ATTGTTGACTTTAATGATGAGCCTG**TATCCA**	-267	12.59979
22	*tlr1660*	-	GTAgttttctaTAC	TATATTTAAGCAAACCCAACAAGA**TCAGCT**C	-162	12.43139
23	*tll0396 tll0397 tll0398*	- - -	GAAgtccccatTAC	TGGTAACGTGGTATGGAAGCCA**TATGAT**GGC	-65	12.40654
24	*tlr1050*	*isiA*	GTTtctggaaaTAC	AACGCCCCTGGCCGTCAGT**TACATA**TCTGTT	-182	12.36656
25	*tll0312*	-	GTAttctattaTAC	GATTTTCCAATATTCATACGG**TAAACG**CCAG	-70	12.34701
26	*tll1486 tll1487 tll1488*	- - -	GTAagtagcacAAC	TATACTCGGCAGTCATAGAC**TGCCGC**TCTTA	-231	12.34164
27	*tlr0833*	*lrtA*	GTAatccccgaTAC	GAAGATGTTAGGGTGGAAG**TAGATC**AGCCTT	-67	12.33556
28	*tlr1530 tsr1531*	*psbB psbT*	GTAgctgaagtTAC	GGGTAAATTGGCCGTTAAACTCT**TAGTAG**GG	-115	12.29831
29	*tll1048*	-	GTAcaactgcgCAC	CAGCACCGTTGACCGTCC**TAAATG**	-25	12.2856
30	*tll0832*	-	GTAtcggggatTAC	AGGGCATCCTGAGGCTCTTCACCTT**TATTCG**	-80	12.27722
31	*tlr2380 tlr2381*	*desC1* -	GTAgctcgttaTAC	TAAAGTCTGATAGGGGCTGT**TCAAGC**TCAGC	-70	12.17003
32	*tll0616 tll0617*	- *sigA*	GTTagaaaggcTAC	AATCAATGAGGTTTTAGT**TAAAAT**TACTCAA	-198	12.14681
33	*tlr0782*	-	GTAatttaatgTAC	ATTTCACAAATCTGTATTAGCCT**TAAGTT**AA	-66	12.11459
34	*tll0559*	-	GTGtaaattgtTAC	AAACCATTACGAGTGCGCAC**TAGACT**GGGAA	-58	12.1118

[1] Bold, -10 like boxes.

Table s11. Predicted NtcA promoters in WH8102

Rank	Transcription unit	Name	NtcA site	Downstream of NtcA binding site and -10 like box[1]	NtcA site position	Score
1	*synw2442*	*urtA*	GTTccggttgaTAC	CAAAGCGGTGGGGGGCCCTTTTT**TACCTT**CC	-52	14.1879
2	*synw1073*	*glnA*	GTGcgcgttgaTAC	AAAACAGGGCATAACGGCTCCT**TACGGT**CGT	-60	13.7574
3	*synw2485 synw2486 synw2487*	*cynA cynB cynC*	GTAtcacctgaTAC	AACATCCGCGTTCGCTTTCCAACTA**TAAATA**	-51	13.692
4	*synw0165*	-	GTAgtttaggaAAC	ATATGGGTTAAAGATTTTTCGT**TATCAA**GAG	-39	13.1547
5	*synw1105*	-	GTGttagttaaTAC	ACAAGCATGTACTAGACTGCGC**TAGTTT**AAT	-64	13.1128
6	*synw0153 synw0154*	- -	GTAgtcgccgcTAC	ATCTGGTGGGGTGGGCAGACCG**TCCTCC**ACC	-62	13.1126
7	*synw0347*	-	GTAgctaatttTAC	TCTAGCCTTGTTTTCTATATAGC**TAGCAC**TA	-173	13.0991
8	*synw1434 synw1435*	- -	GTAacaacaccTAC	AGCCTGAACCAGTCCACTCGG**TAACAC**TATG	-143	12.7911
9	*synw2477*	*nirA*	GTAattccatcAAC	AGAACAACTTTTGAGTACGAAC**TAGAAA**AGG	-349	12.7111
10	*synw0253*	*amt1*	GTTcagtcggaTAC	ACCATCCGGCGTGACCAGCAGCTC**TGCACT**C	-37	12.7063
11	*synw1412 synw1413 synw1414* *synw1415 synw1416 synw1417* *synw1418*	- *hypA2 hypB* - - - -	GTAgctgatcaCAC	CGCGCGTGCCACCGGTGCCGCA**GACAGT**GGA	-69	12.6782
12	*synw2475 synw2476*	*cobA* -	GTTgatggaatTAC	GATTCCGCTCGTATTGCCGTCTG**TAACTC**TT	-199	12.5851
13	*synw1507 synw1508*	- -	GTAataaagacTGC	GGAATTAATATTTCGGCAATACTTA**TACCTT**	-111	12.5376
14	*synw0152*	*rnc*	GTAgcggcgacTAC	CGATATCGGCGCTCCTGACGGGGC**TGGCGG**G	-527	12.4768
15	*synw1422 synw1423 synw1424* *synw1425 synw1426*	- - - - -	GTCgtatttcgTAC	ATTTTTTGTGGGCCGACCGGAGC**CAGTCT**TT	-52	12.4149
16	*synw0462*	*glnB*	GTTacaggggcTAC	CCACACCGCCACCATTCACG**TCATGC**TTAAT	-51	12.2099
17	*synw2171*	-	GTCatggatacTAC	CCTTGCCCACCTCTGTACACTT**TCGGGT**AGC	-56	12.1677

[1] Bold, -10 like boxes.

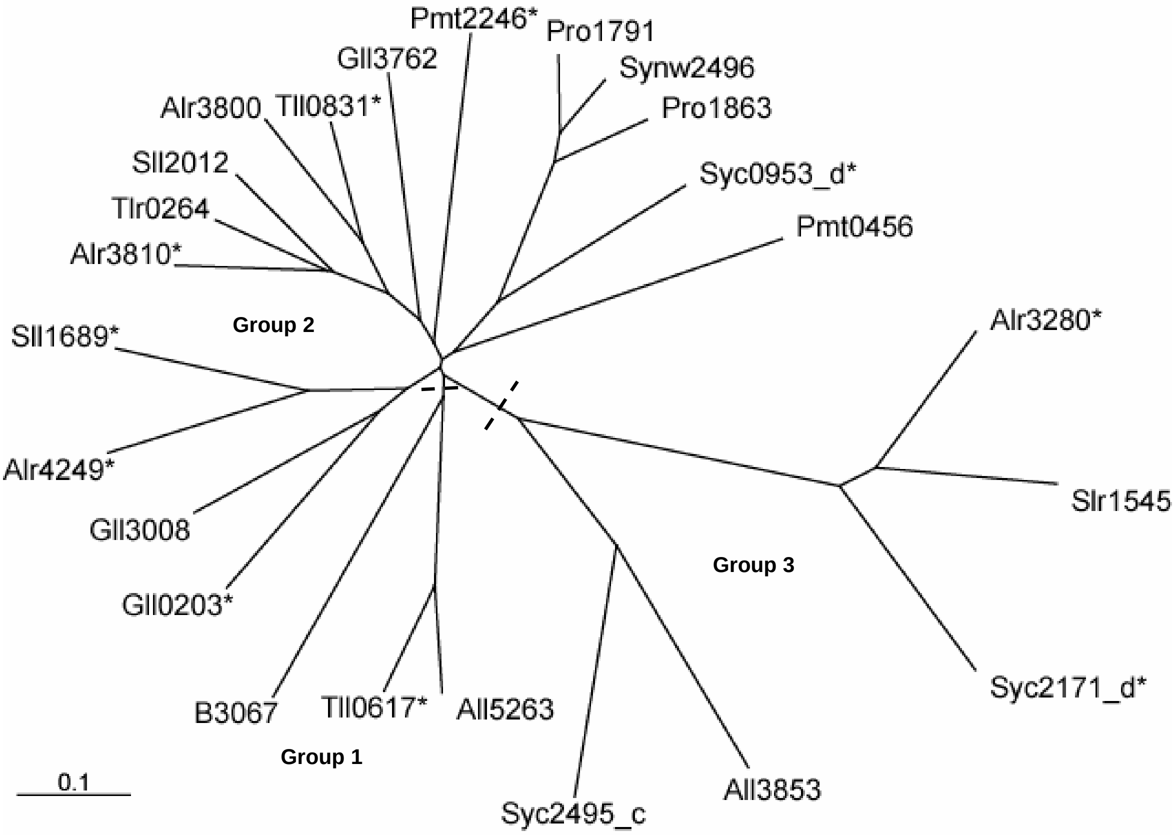

Figure s1. Phylogenetic relationships of the σ-factors bearing a putative NtcA promoter in their regulatory regions of the 9 genomes in our analyses. The unrooted tree was generated by the neighbor-joining method based on the multiple amino acid sequence alignments by the ClustalX program using the default settings. The σ^{70} of *E. coli* is included in this tree (B3067). The proteins are labeled by their synonyms. The dotted lines show the division of the groups. A star (*) indicates that the putative promoter is identified at p-value <0.01, otherwise at p-value <0.05. The gnomes that these proteins are from are as follows. CCMP1375: Pro1791 and Pro1863; PCC7421: Gll0203, Gll3762 and Gll3008; MIT9313: Pmt2246 and Pmt0456; PCC7120, All4249, Alr3810, Alr3280, All3853, All3800 and All5263; PCC6301, Syc2171_d, Syc2495_c and Syc0953_d; PCC6803, Sll1689, Sll2012 and Slr1545; thermosynechococcus: Tll0831, Tll0617 and Tlr0264; WH8102: Synw2496. Scale bar, substitutions per position.