

Supplementary Material

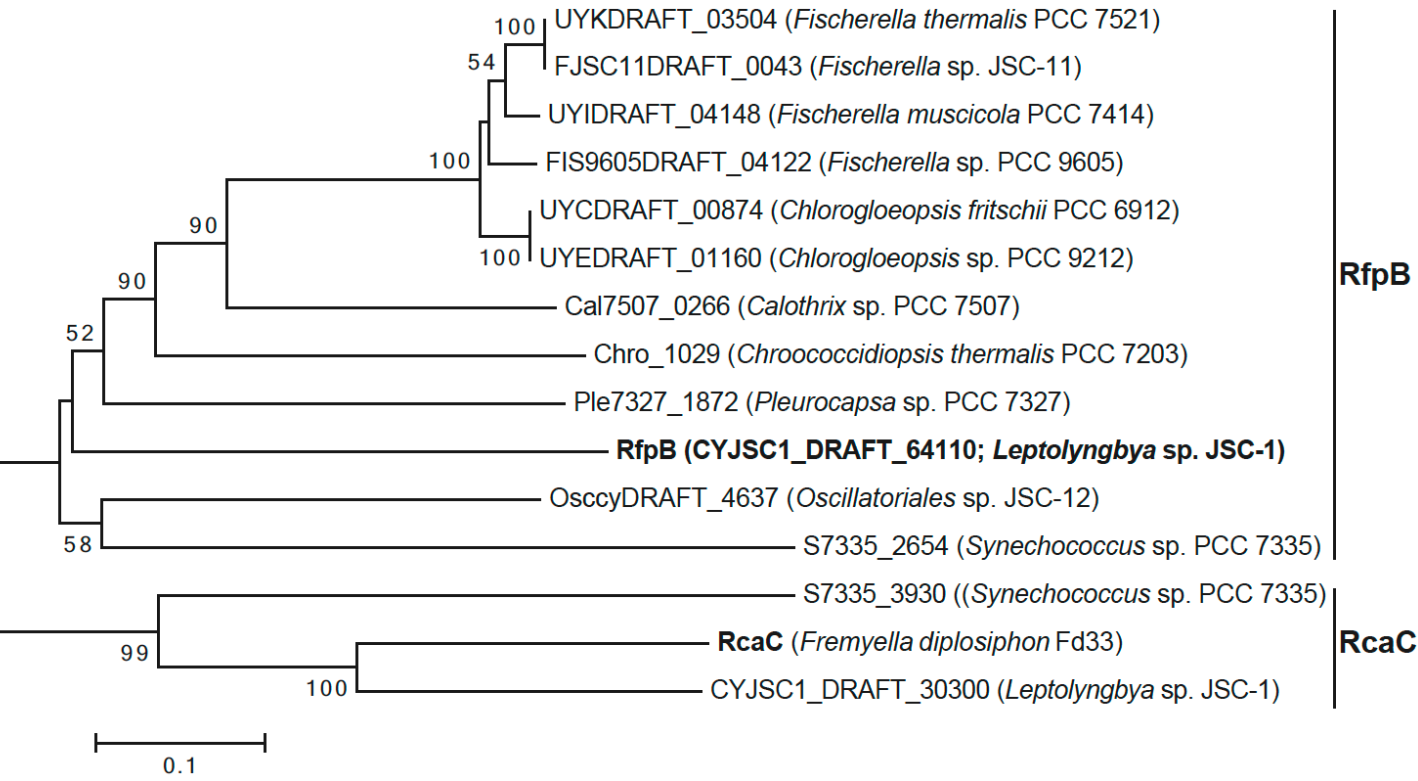


Figure S1. Phylogenetic tree for RfpB sequences from strains with photosynthetic gene clusters similar to that in *Leptolyngbya* JSC-1. Sequences were aligned using the the MUSCLE module implemented in MEGA6, and a neighbor-joining tree was built from the resulting sequence alignment. Numbers at the nodes indicate the bootstrap support based upon 100 resamplings. A few RcaC sequences, the signal receiver/response regulator protein for CCA, were included as outgroup.

	10	20	30	40	50	60	70	80	90	100
57335_2654	MRILLVEDDES	VAKILHKVL	TDEHYVDV	AYTGHIGW	QLVSSFN	YDLIILDV	LLPELDGI	KFCQQLR	AQYETL	PVLFV
Chro_1029	MRILLIEDDEC	VAKAKKNV	LQQQNYVD	TATDGTG	WDLIQTFT	YDILLD	VILPKLDG	IKLCQRL	RDRKYQ	TPVLL
Ple7327_1872	MRILVVEDDK	SIAETL	ERVLGEQ	HYVDAAT	DGEMGW	ELVEFAY	DLIVL	DKMLPK	NGIEFC	QRLSR
CYJSC1_DRAFT_64110	MKVLLVEDDE	CTAKTLE	TALARE	NYAIDV	AGDGLG	WRLEAF	AYDLIL	LDVILP	KLDGIS	LCRQIR
OscCyDRAFT_4637	MRILLVEDDE	SISRAL	EAILTE	HHYVVD	AAADQIG	WECVEA	FTYDIL	VLDVLL	PRLDGI	QFCQQL
Ca17507_0266	MRILLVEDDE	CIAKTLE	RVLG	NQNYVD	IAADGL	LWELVE	AFSYDL	ILLDIIL	PKLDGI	KLCQRL
UYEDRAFT_01160	MRLLLVEDDE	CIADTLE	KILGNH	HYVVD	VANDG	ELGWEL	VEAFNY	DLILLD	VMLPK	LDGIQL
UYCDRAFT_00874	MRLLLVEDDE	CIADTLE	KILGNH	HYVVD	VANDG	ELGWEL	VEAFNY	DLILLD	VMLPK	LDGIQL
FIS9605DRAFT_04122	MRLLLVEDDE	CIANTLE	ILGNH	HYVVD	VANDG	ELGWEL	VEAFNY	DLILLD	VMLPK	LDGIQL
FJSC11DRAFT_0043	MRLLLVEDDE	CIANTLE	ILGNH	HYVVD	VANDG	ELGWEL	VEAFNY	DLILLD	VMLPK	LDGIQL
UYKDRAFT_04148	MRLLLVEDDE	CIANTLE	ILGNH	HYVVD	VANDG	ELGWEL	VEAFNY	DLILLD	VMLPK	LDGIQL
UYKDRAFT_03504	MRLLLVEDDE	CIANTLE	ILGNH	HYVVD	VANDG	ELGWEL	VEAFNY	DLILLD	VMLPK	LDGIQL
57335_2654	KPFEEELLAR	VRALLRR	GKTSVVL	TLEWEG	RLLEVNS	RELTyr	NNKLSL	TPKEHAM	LELFLR	NPAQAF
Chro_1029	KPFDMPELLA	RIRVLLRR	-QSKPIL	QALWGD	RLRQC	GTCEVT	YGDRV	LHLTP	NEYRLE	LELFLR
Ple7327_1872	KPFDPPELLA	RIRVLLRR	-KNSLIP	PVLEW	GDRLR	DPNTE	CEASYS	QPLRL	TPKEYR	LELFLR
CYJSC1_DRAFT_64110	KPFELPELLA	RIRVLLRR	-GSSFVL	PVLEW	HDRLD	LAACQV	TYGDL	LQLTP	KEYRML	ELFLR
OscCyDRAFT_4637	KPFEPQELLA	RIRVLLRR	-KGSAR	PAVLE	WMALR	DLPSL	CEVTY	DGHVSL	TPKEYR	LELFLR
Ca17507_0266	KPFELSELLA	RIRVLLRR	-RNSPIL	SVMKWE	KLRINS	GSCEVT	YGCGL	NLTP	KEYRLE	LELFLR
UYEDRAFT_01160	KPFDMSELLA	RIRVLLRR	-RNSPIQ	TVLEW	ENLRD	PGKCEV	LYNNH	LNLTP	KEYRLE	LELFLR
UYCDRAFT_00874	KPFDMSELLA	RIRVLLRR	-RNSPIQ	TVLEW	ENLRD	PGKCEV	LYNNH	LNLTP	KEYRLE	LELFLR
FIS9605DRAFT_04122	KPFDMSELLA	RIRVLLRR	-RHSPIQ	TVLEW	ENLRD	PGKCEV	TYNSH	LNLTP	KEYRLE	LELFLR
FJSC11DRAFT_0043	KPFDMSELLA	RIRVLLRR	-RNSTLQ	TVLEW	ENLRD	PGKCEV	LYNNH	LNLTP	KEYRLE	LELFLR
UYKDRAFT_04148	KPFDMSELLA	RIRVLLRR	-RNSTLQ	TVLEW	ENLRD	PGKCEV	LYNNH	LNLTP	KEYRLE	LELFLR
UYKDRAFT_03504	KPFDMSELLA	RIRVLLRR	-RNSTLQ	TVLEW	ENLRD	PGKCEV	LYNNH	LNLTP	KEYRLE	LELFLR
57335_2654	KLRTVGAPAD	LIIETVYG	GYRLKPI	ETEVKPI	ETEEELC	SVDSRSA	-----	-----	-----	-----
Chro_1029	KLKQVGAPAD	LIIETVYG	GYRLKQ	PAPTTK	TKLRGAS	LPSQSDS	-----	-----	-----	-----
Ple7327_1872	KLKRAGSP	PDIIATVYG	IGYRLK	SSAPSA	SEASCAP	SATDKRAIS	-----	-----	-----	-----
CYJSC1_DRAFT_64110	KLKQAGAP	FNIIETVYG	MGYRLR	QSSK	-----	EPSSQV	SPGNAQ	AAAGQK	QAPIIL	NGNGS
OscCyDRAFT_4637	KLKRQAGAP	PDIIQTVYG	GQGYRL	KQPS	SEFTAS	PPALPG	WVRDGL	VQ	-----	-----
Ca17507_0266	KLQAGAPT	DFIETVYG	LYRLKAF	ADEEK	PSVASR	QTKKSQ	SKR	-----	-----	-----
UYEDRAFT_01160	KLKQAGAEN	DFIETVYG	LYRLKV	FTSSQ	KMSQKT	KENHRL	NGAK	-----	-----	-----
UYCDRAFT_00874	KLKQAGAEN	DFIETVYG	LYRLKV	FTSSQ	KMSQKT	KENHRL	NGAK	-----	-----	-----
FIS9605DRAFT_04122	KLKQAGAEN	DFIETVYG	LYRLKV	FTSSQ	KMLPQ	QTKENH	RLNGVK	-----	-----	-----
FJSC11DRAFT_0043	KLKQAGAEN	DFIETVYG	LYRLKV	FTSSQ	KMLPQ	QTKENH	RLNGVK	-----	-----	-----
UYKDRAFT_04148	KLKQAGAEN	DFIETVYG	LYRLKV	FTSSQ	KMSQKT	KENHRL	NGAK	-----	-----	-----
UYKDRAFT_03504	KLKQAGAEN	DFIETVYG	LYRLKV	FTSSQ	KMLPQ	QTKENH	RLNGVK	-----	-----	-----
57335_2654	VKLRH-----	-----	QHTQAAL	SALWKS	VKSQQT	NRLAIL	QQA	FEQLQ	RGTLS	KEMRES
Chro_1029	-----	-----	IQQTKAAL	AEVWEK	IKYQS	SDRVAV	LEQAT	ISL	LDNKL	TEELRE
Ple7327_1872	ISYSD-----	-----	ATRSRIA	QQT	KSGLKV	WEKFG	QGT	KARV	TILEQ	ATAAL
CYJSC1_DRAFT_64110	QALANPT	QIQDLE	TGRQK	QIQV	GLW	SKYA	AEHNF	NRLRL	IEQ	AEALQ
OscCyDRAFT_4637	-----	-----	QKTRAG	LADV	WAKYK	HLNCD	RCLV	LEQ	AGASML	VGTL
Ca17507_0266	-----	-----	KPQIRE	ALAV	VWERL	KVHSS	DRLE	IIQ	QASLE	LQANNL
UYEDRAFT_01160	QQAEDLF	YEVQDL	HNH	-----	IQQTKA	ALT	KLWEK	IKDKS	CDRI	II
UYCDRAFT_00874	QQAEDLF	YEVQDL	HNH	-----	IQQTKA	ALT	KLWEK	IKDKS	CDRI	II
FIS9605DRAFT_04122	QQTRDLC	YETKDL	QNG	-----	IQQTKA	ALT	KLWEK	IKDKS	CDRI	II
FJSC11DRAFT_0043	QQIKDLF	YEAQDL	HDN	-----	IQQTKA	ALT	KLWEK	IKDKS	CDRI	II
UYKDRAFT_04148	QQIRDLF	YEAQDL	HDN	-----	IQQTKA	ALT	KLWEK	IKDKS	CDRI	II
UYKDRAFT_03504	QQIKDLF	YEAQDL	HDN	-----	IQQTKA	ALT	KLWEK	IKDKS	CDRI	II
57335_2654	LADDQARLD	ALIRAL	QATIDE	-AIEHR	GRTAVE	IPKVP	LLVL	VDPQL	RMPAL	VGMRL
Chro_1029	DRSEALH	LYEMV	MALKRE	IQR-PS	LSQLDR	PLRQ	QSIV	MLAID	GDAGL	AEQIV
Ple7327_1872	-LTRSSHL	SKRV	VALRQL	LEE-SS	ASQIAE	PIARH	RSVLL	VISD	DAKA	ARRLV
CYJSC1_DRAFT_64110	PREKHAAL	IAIIAEL	RQVLQ	AGAPL	PEVGT	-----	TSLLL	VIDH	PTLL	LEQVR
OscCyDRAFT_4637	TPNKAVQ	LLDLLN	LKQKAI	QS-PP	-SHHNR	-----	VPLL	VIED	DEELAD	QVFL
Ca17507_0266	DQSEALR	LRHLV	ELLSG	ELQ	-PAFRE	LDK	-----	-----	-----	-----
UYEDRAFT_01160	NQNLILH	LCLNL	VIALK	QEIQQ	-PAFAE	LEK	-----	-----	-----	-----
UYCDRAFT_00874	NQNLILH	LCLNL	VIALK	QEIQQ	-PAFAE	LEK	-----	-----	-----	-----
FIS9605DRAFT_04122	NQSVLHL	LYNLV	IALKQ	EIQ	-PAFAE	LEK	-----	-----	-----	-----
FJSC11DRAFT_0043	NQSVLHL	LYNLV	IALKQ	EIQ	-PAFAE	LEK	-----	-----	-----	-----
UYKDRAFT_04148	NHNLVHL	LYNLV	IALKQ	EIQ	-PAFAE	LEK	-----	-----	-----	-----
UYKDRAFT_03504	NQSVLHL	LYNLV	IALKQ	EIQ	-PAFAE	LEK	-----	-----	-----	-----

Figure S2. Cont.

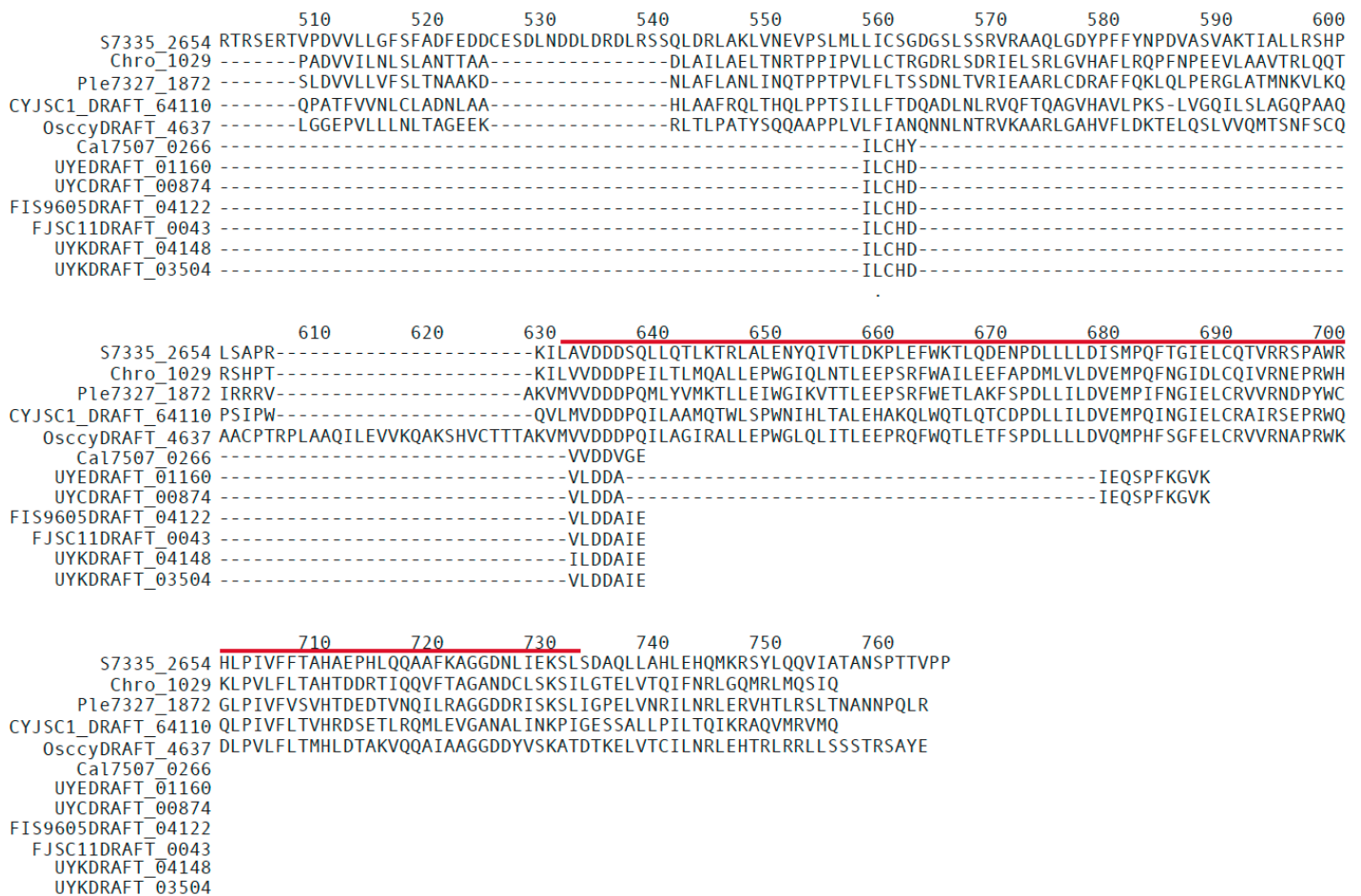
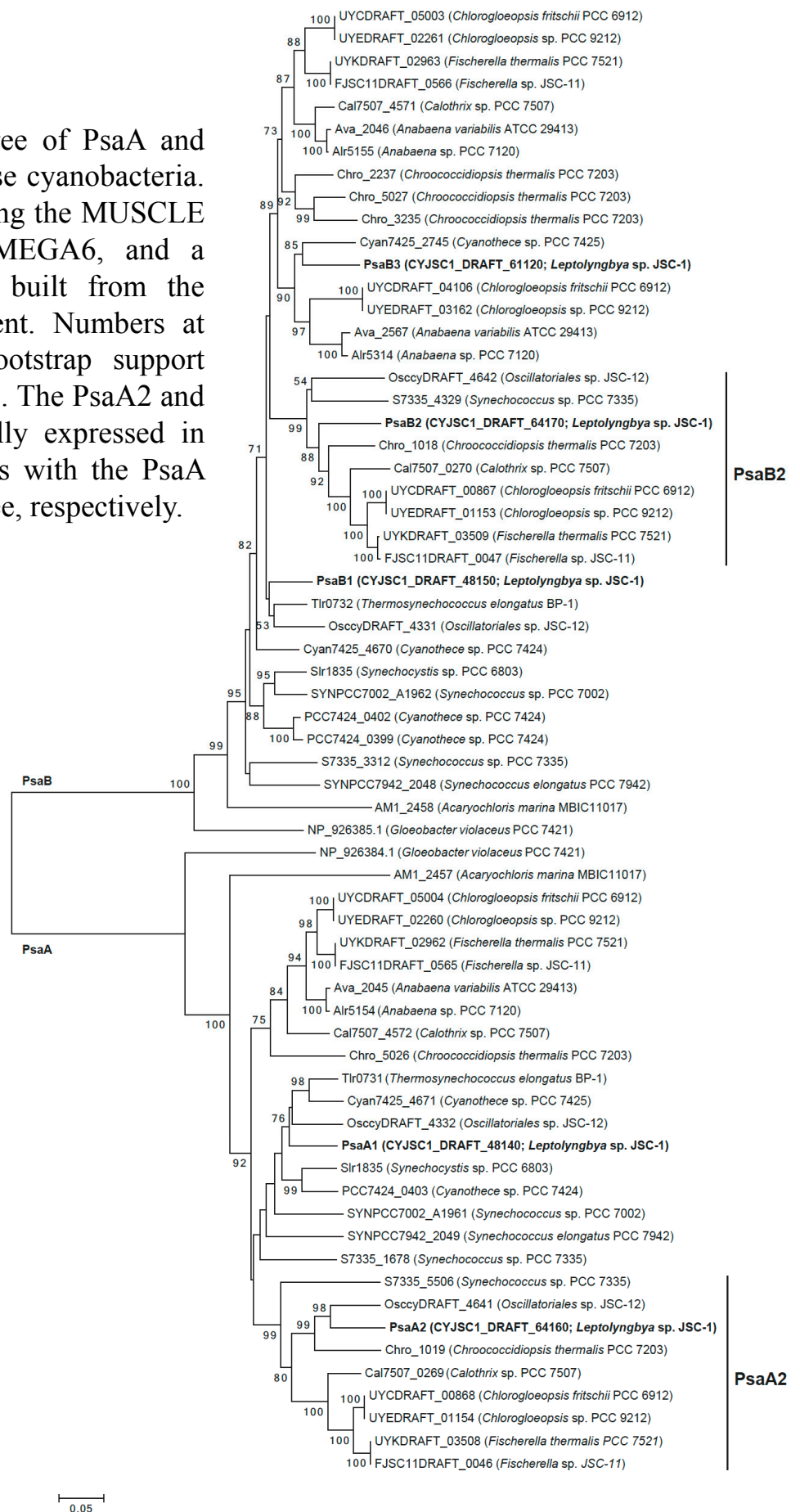


Figure S2. Sequence alignment for RfpB sequences from various strains containing photosynthetic gene clusters similar to that in *Leptolyngbya* JSC-1. Sequences were aligned using the MUSCLE module implemented in MEGA6. The red lines indicate the CheY-like signal-receiver domains at the N-terminus (residues 1 to 120) and C-terminus (residues 630 to 735). The green line indicates the DNA-binding, winged-helix domain (residues 128 to 233). The blue line indicates the histidine-containing phosphotransferase (HPT) domain (residues ~300 to 430).

Figure S3. Phylogenetic tree of PsaA and PsaB sequences from diverse cyanobacteria. Sequences were aligned using the MUSCLE module implemented in MEGA6, and a neighbor-joining tree was built from the resulting sequence alignment. Numbers at the nodes indicate the bootstrap support based upon 100 resamplings. The PsaA2 and PsaB2 sequences specifically expressed in FRL form distinctive clades with the PsaA and PsaB branches of the tree, respectively.



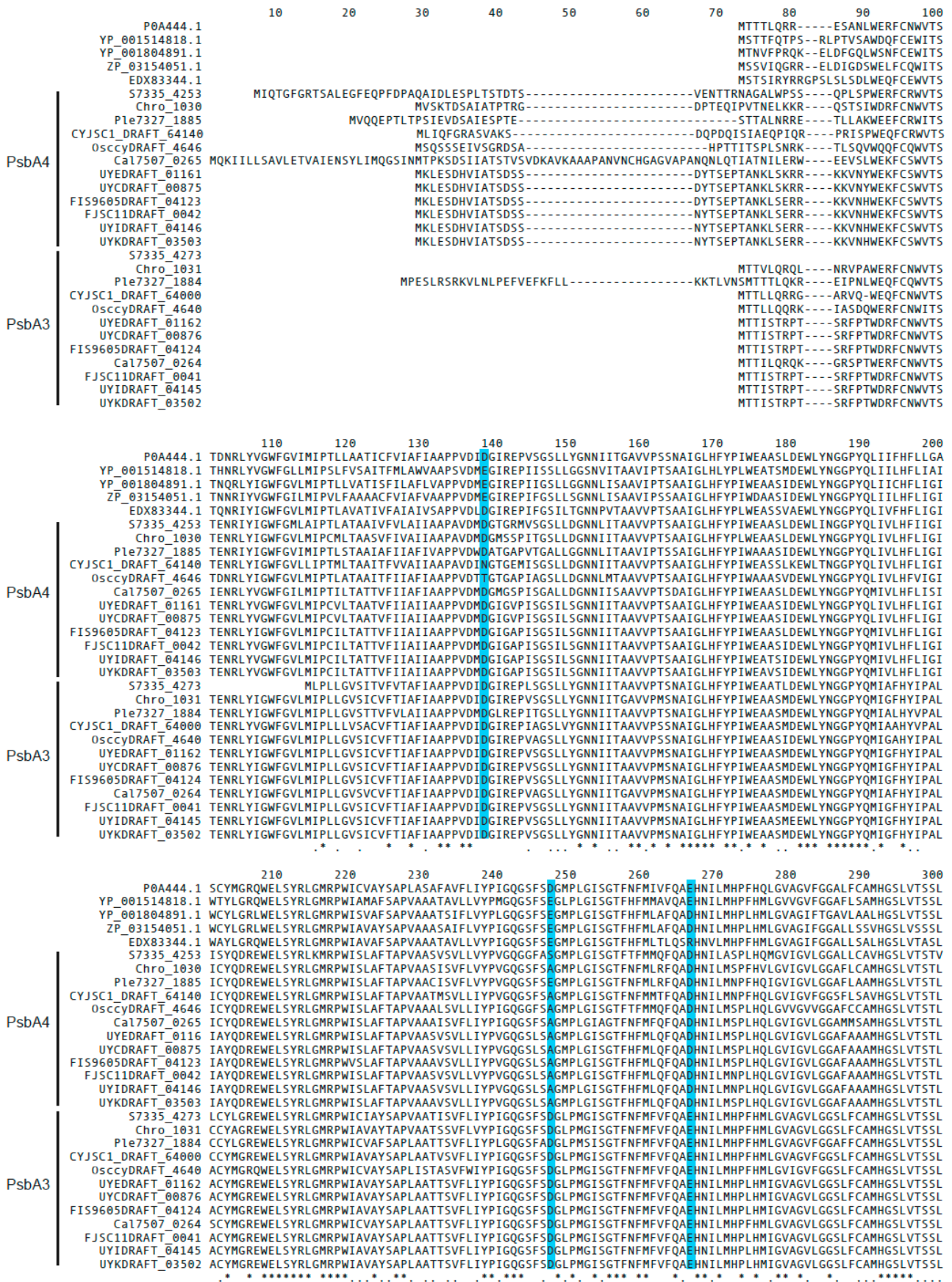


Figure S4. Cont.

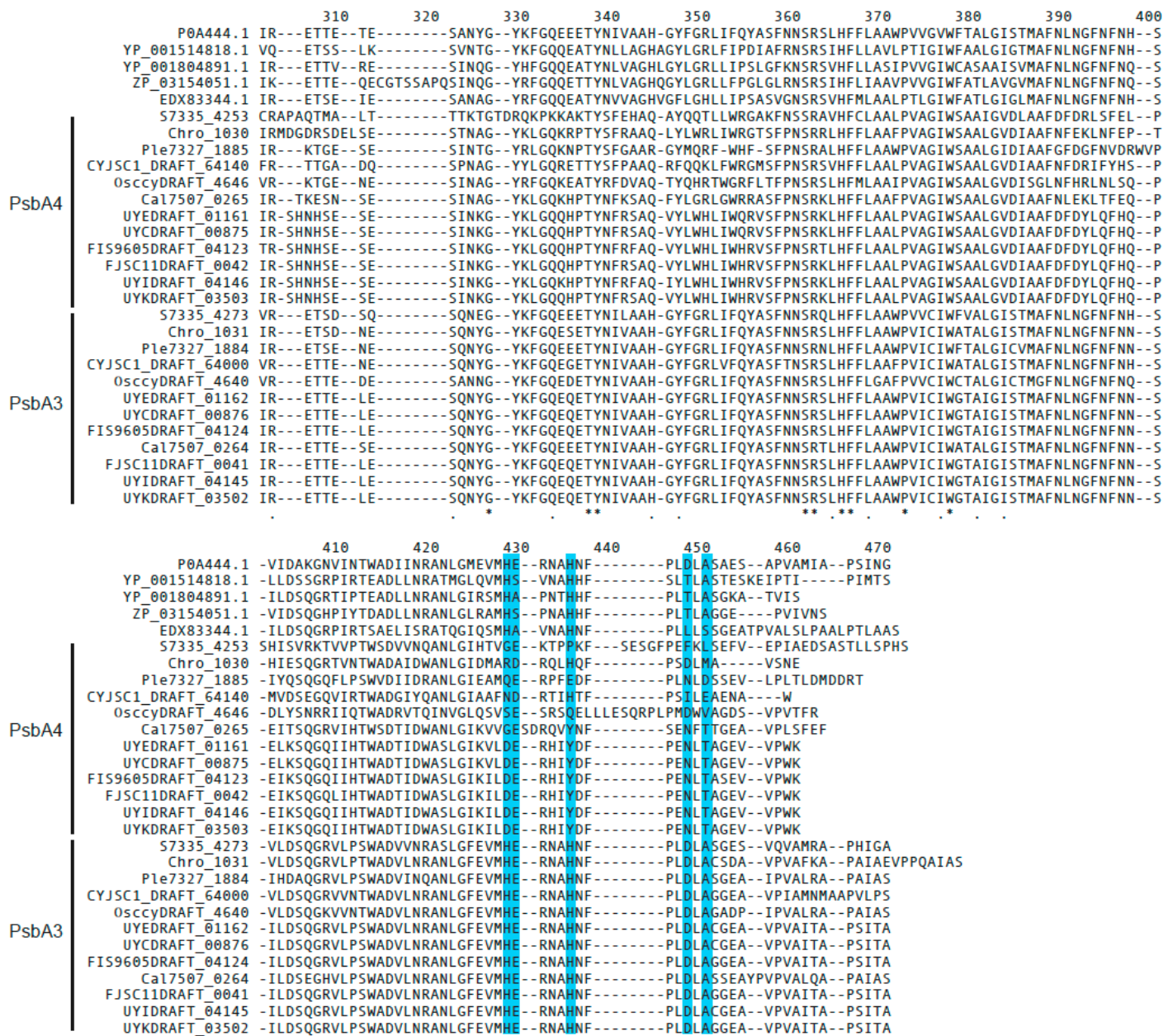


Figure S4. Sequence alignment for PsbA3 and PsbA4 sequences from various strains containing photosynthetic gene clusters similar to that in *Leptolyngbya* JSC-1. A few sequences from other strains are also included for comparison (also see Figure 5). Sequences were aligned using the MUSCLE module implemented in MEGA6. The residues shaded in blue are functionally important residues that provide ligands to important prosthetic groups for electron transfer (for details, see [41]). The D1 protein encoded by *psbA3* includes most of the highly conserved, functionally important residues found in PsbA proteins. The D1 protein encoded by *psbA4* is a “super-rogue” D1 protein that lacks most of the highly conserved, functionally important residues for binding prosthetic groups, in particular the Mn₄ cluster of the oxygen-evolving complex.

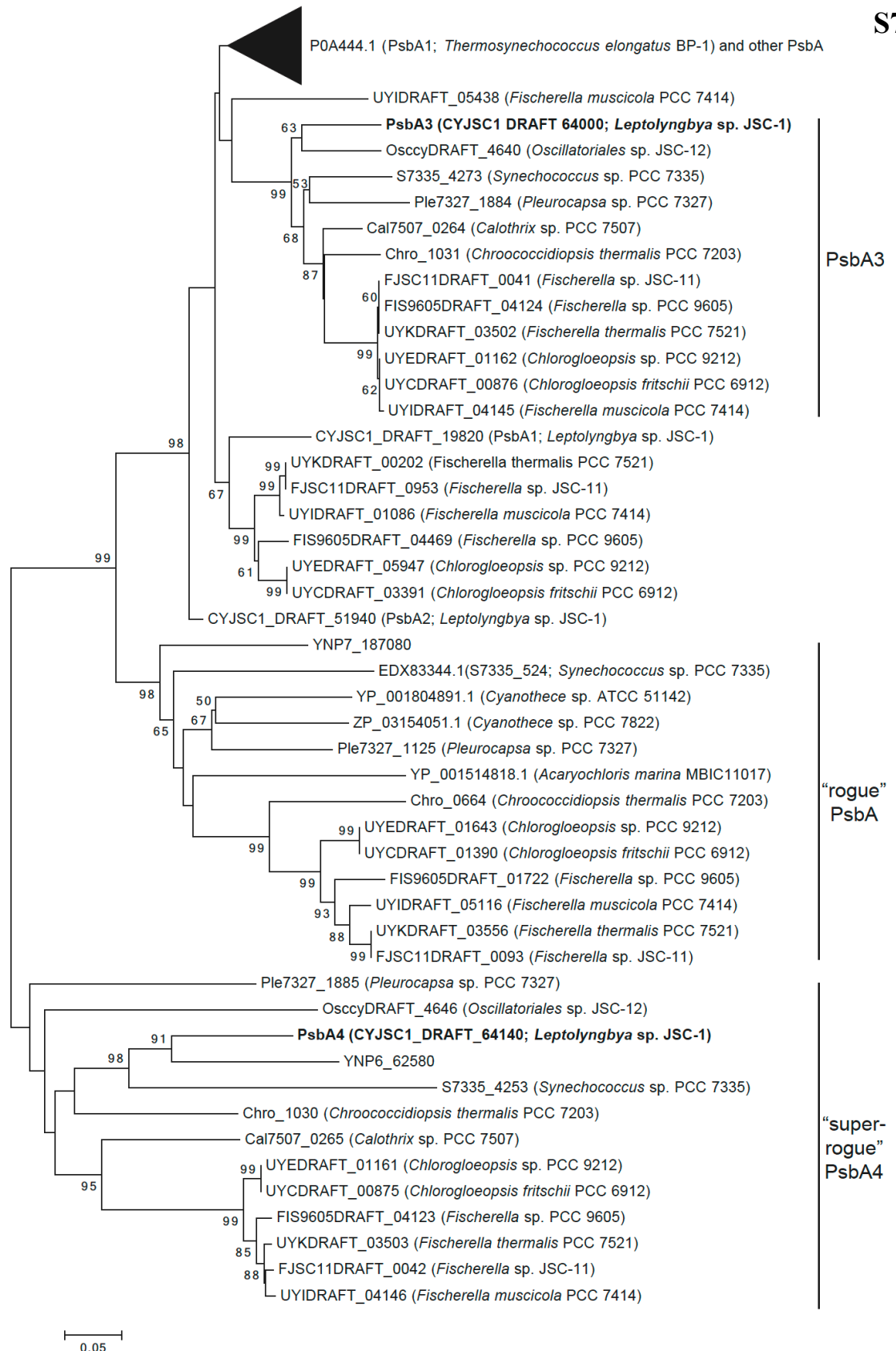


Figure S5. Phylogenetic tree for PsbA1, PsbA3, "rogue" and "super-rogue" PsbA4 sequences. The included PsbA3 and PsbA4 sequences are found in the photosynthetic gene clusters that are similar to the 21-gene cluster found in *Leptolyngbya* JSC-1 [28]. Sequences were aligned using the MUSCLE module implemented in MEGA6. The designations of "rogue" PsbA and "Super-rogue" PsbA sequences are taken from the analysis of Murray [41]. For additional details, see main text.

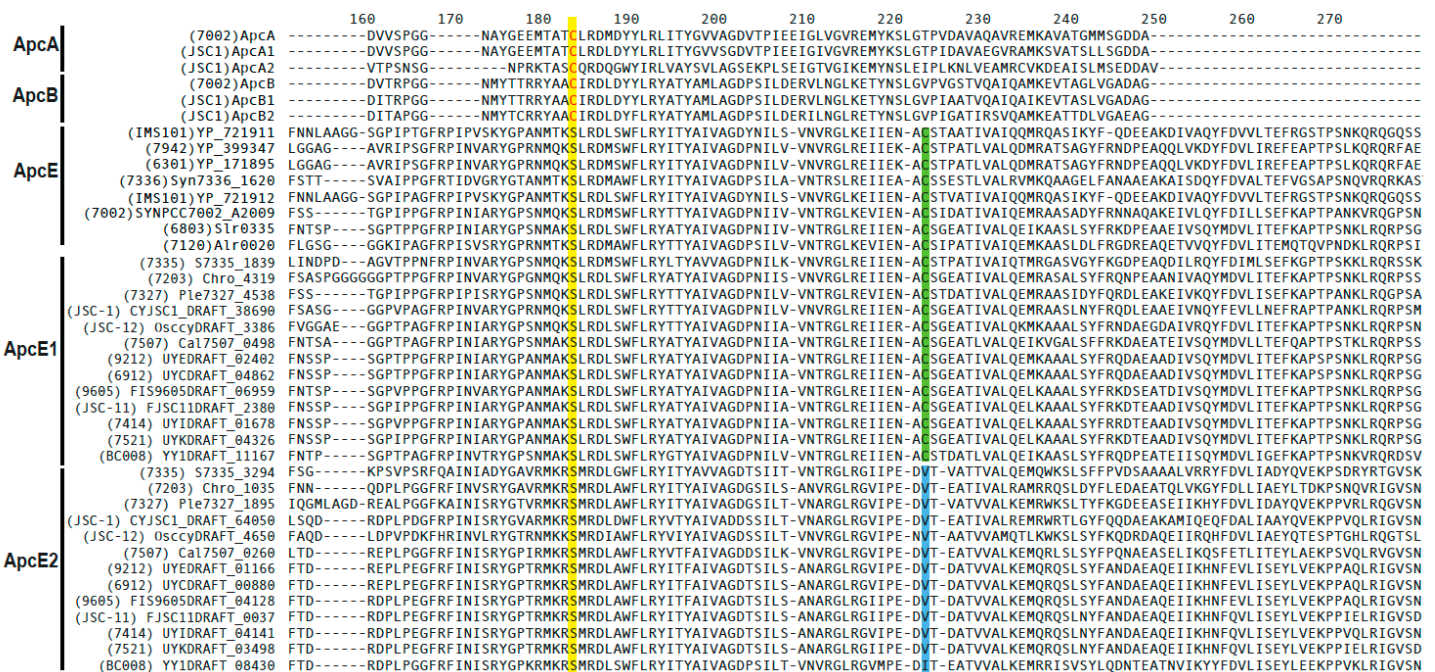


Figure S6. Sequence alignment of a portion of the N-terminal, phycobiliproteins-like domains of ApcE1 and ApcE2 proteins (compared to the corresponding regions of selected ApcA and ApcB sequences). Sequences were aligned using the MUSCLE module implemented in MEGA6. The phycocyanobilin-binding cysteine residue in ApcA and ApcB is highlighted in yellow and red font. The corresponding residue is serine in all ApcE1 and ApcE2 sequences. The chromophore binding cysteine residue for ApcE1 sequences is shaded in green. In all ApcE2 sequences, cysteine does not occur at either position as shown by the yellow and blue shading. Because the chromophore binding pocket is still present (residues 180 to 230), it is probable that the phycocyanobilin associated with ApcE2 is non-covalently bound.

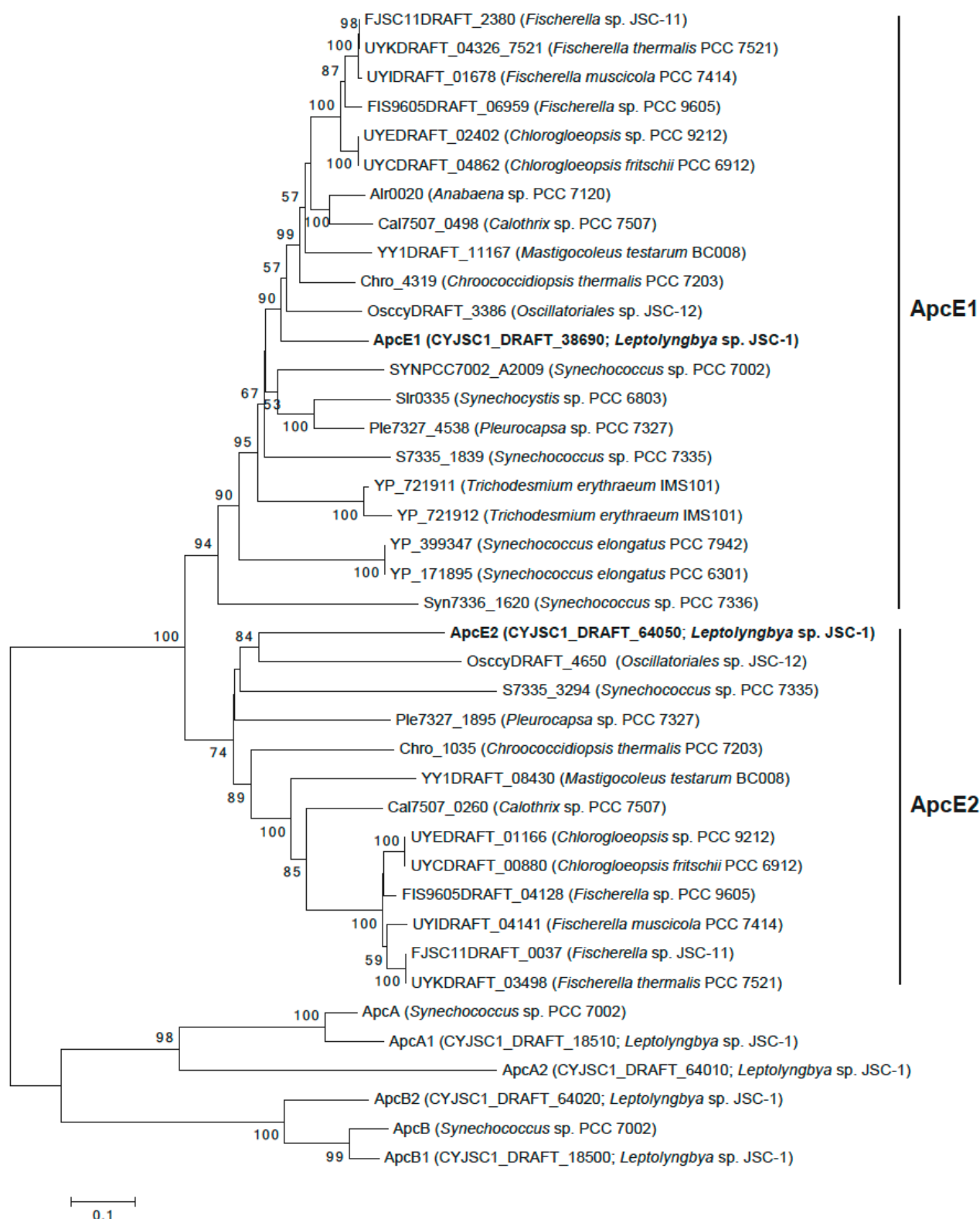
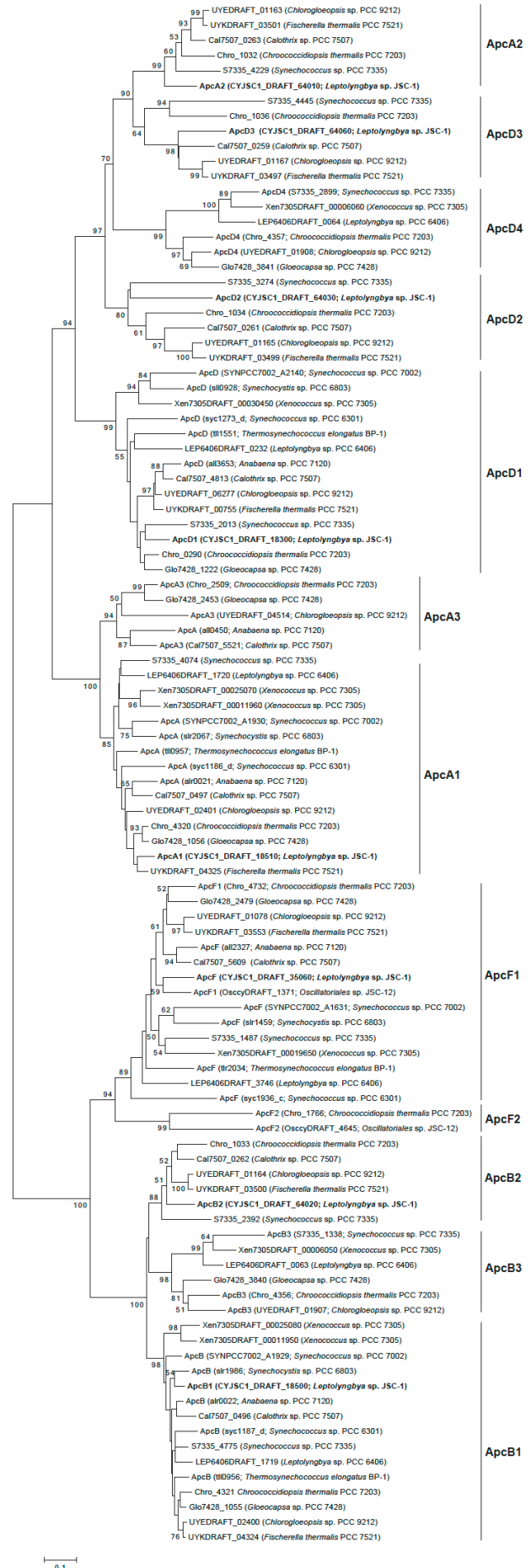


Figure S7. Phylogenetic tree for ApcE1 and ApcE2 sequences from strains with photosynthetic gene clusters similar to that in *Leptolyngbya* JSC-1 and from a few other selected strains with ApcE/ApcE1-like sequences. Sequences were aligned using the MUSCLE module implemented in MEGA6, and a neighbor-joining tree was built from the resulting sequence alignment. Numbers at the nodes indicate the bootstrap support based upon 100 resamplings. Only the N-terminal phycobiliprotein-like domain and REP domains 1 and 2 were included in the analysis. For additional details, see main text.

Figure 8. Phylogenetic tree for ApcA, ApcB, ApcD, and ApcF sequences from strains with photosynthetic gene clusters similar to that in *Leptolyngbya* JSC-1 and from a few other selected strains. Sequences were aligned using the MUSCLE module implemented in MEGA6, and a neighbor-joining tree was built from the resulting sequence alignment. Numbers at the nodes indicate the bootstrap support based upon 100 resamplings. Seven clades of allophycocyanin alpha subunits (ApcA1, ApcA2, ApcA3, ApcD1, ApcD2, ApcD3 and ApcD4) were identified. ApcA1 and ApcD1 are replaced by ApcA2, ApcD2, and ApcD3 when cells are grown in FRL. Four clades of beta subunits (ApcB1, ApcB2, ApcB3 and ApcF) were identified. ApcB2 replaces ApcB1 when FaRLiP strains are grown in FRL. In most strains, the *apcF* gene is present as a single-copy gene and presumably the product of that gene is used with both ApcE1 and ApcE2. However, two FaRLiP organisms, *Chr. thermalis* PCC 7203 and *Oscillatoria* sp. JSC-12, have two ApcF genes, and in the latter organism, the *apcF2* gene is localized in the FaRLiP gene cluster (see [28]). The functional roles of ApcA3, ApcB3, and ApcD4 are unknown. Additional details can be found in the main text.



		10	20	30	40	50	60	70	80	90	100
ApcA1	(7002)ApcA	MSIVTKSVINADAEARYLSPGELDRKAFVTSGESRLRIAE	TLTGSRRRIKISAGDALFQKRPDVSPGGNAYGEEMTAT	CLRDMDYLLRLITYGVVAGD							
	S7335_4074	MSIVTKSVINADAEARYLSPGELDRKIGFVTSGERRRVIA	AQVLTESRERIVKTAGDQLFQKRPDVSPGGNAYGEEMTAT	CLRDMDYLLRLITYGVVAGD							
	Chro_4320	MSIVTKSVINADAEARYLSPGELDRKISFVTSGERRRRIA	AQALTDNRERIVKQAGDQLFQKRPDVSPGGNAYGEEMTAT	CLRDLDYLLRLITYGVVAGD							
	CYJSC1_DRAFT_18510	MSIVTKSVINADAEARYLSPGELDRKIGFVTSGERRRRIA	AQVLTDNRERIVKQAGDQLFQKRPDVSPGGNAYGEEMTAT	CLRDLDYLLRLITYGVVAGD							
	Cal7507_0497	MSIVTKSVINADAEARYLSPGELDRKISFVTSGERRRRIA	AQVLTDNRERIVKQAGDQLFQKRPDVSPGGNAYGEEMTAT	CLRDLDYLLRLITYGVVAGD							
ApcA2	UYEDRAFT_02401	MSIVTKSVINADAEARYLSPGELDRKISFVTSGERRRRIA	AQVLTDNRERIVKQAGDQLFQKRPDVSPGGNAYGEEMTAT	CLRDLDYLLRLITYGVVAGD							
	UYKDRAFT_04325	MSIVTKSVINADAEARYLSPGELDRKISFVTSGERRRRIA	AQVLTDNRERIVKQAGDQLFQKRPDVSPGGNAYGEEMTAT	CLRDLDYLLRLITYGVVAGD							
	S7335_4229	MSIVTKSVINADAEARYLSPGELDRKISFVTSGERRRRIA	AQVLTDNRERIVKQAGDQLFQKRPDVSPGGNAYGEEMTAT	CLRDLDYLLRLITYGVVAGD							
	Chro_1032	MSIVTKSVINADAEARYLSPGELDRKISFVTSGERRRRIA	AQVLTDNRERIVKQAGDQLFQKRPDVSPGGNAYGEEMTAT	CLRDLDYLLRLITYGVVAGD							
	CYJSC1_DRAFT_64010	MSIVTKSVINADAEARYLSPGELDRKISFVTSGERRRRIA	AQVLTDNRERIVKQAGDQLFQKRPDVSPGGNAYGEEMTAT	CLRDLDYLLRLITYGVVAGD							
ApcA3	Cal7507_0263	MSIVTKSVINADAEARYLSPGELDRKISFVTSGERRRRIA	AQVLTDNRERIVKQAGDQLFQKRPDVSPGGNAYGEEMTAT	CLRDLDYLLRLITYGVVAGD							
	UYEDRAFT_01163	MSIVTKSVINADAEARYLSPGELDRKISFVTSGERRRRIA	AQVLTDNRERIVKQAGDQLFQKRPDVSPGGNAYGEEMTAT	CLRDLDYLLRLITYGVVAGD							
	UYKDRAFT_03501	MSIVTKSVINADAEARYLSPGELDRKISFVTSGERRRRIA	AQVLTDNRERIVKQAGDQLFQKRPDVSPGGNAYGEEMTAT	CLRDLDYLLRLITYGVVAGD							
	Chro_2509	MSIVTKSVINADAEARYLSPGELDRKISFVTSGERRRRIA	AQVLTDNRERIVKQAGDQLFQKRPDVSPGGNAYGEEMTAT	CLRDLDYLLRLITYGVVAGD							
	CYJSC1_DRAFT_04521	MSIVTKSVINADAEARYLSPGELDRKISFVTSGERRRRIA	AQVLTDNRERIVKQAGDQLFQKRPDVSPGGNAYGEEMTAT	CLRDLDYLLRLITYGVVAGD							
ApcD1	(7002)ApcD	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
	S7335_2033	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
	Chro_0290	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
	CYJSC1_DRAFT_18300	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
	Cal7507_4813	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
ApcD2	UYEDRAFT_01165	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
	UYKDRAFT_03499	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
	S7335_4445	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
	Chro_1036	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
	CYJSC1_DRAFT_64060	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
ApcD3	Cal7507_0259	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
	UYEDRAFT_01167	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
	UYKDRAFT_03497	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
	S7335_2899	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
	Chro_4357	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
ApcD4	UYEDRAFT_01908	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
	Xen7305DRAFT_0006060	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
	LEP6406DRAFT_0064	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
	G107428_3841	MSVSVQVILRADDELRYPSGELKSIQNFVLTGQRRRIA	ATLAEANEKIVQEQSKLFWERCPNTPSNSGN---	ERKTASCLRDQGWYRLVAYSILAGS							
ApcA1	(7002)ApcA	VTPIIEIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	S7335_4074	VTPIIEIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	Chro_4320	VTPIIEIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	CYJSC1_DRAFT_18510	VTPIIEIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	Cal7507_0497	VTPIIEIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
ApcA2	UYEDRAFT_02401	VTPIIEIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	UYKDRAFT_04325	VTPIIEIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	S7335_4229	VTPIIEIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	Chro_1032	VTPIIEIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	CYJSC1_DRAFT_64010	VTPIIEIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
ApcA3	Cal7507_0263	VTPIIEIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	UYEDRAFT_01163	VTPIIEIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	UYKDRAFT_03501	VTPIIEIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	Chro_2509	VTPIIEIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	CYJSC1_DRAFT_04521	VTPIIEIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
ApcD1	(7002)ApcD	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	S7335_2033	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	Chro_0290	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	CYJSC1_DRAFT_18300	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	Cal7507_4813	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
ApcD2	UYEDRAFT_01165	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	UYKDRAFT_03499	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	S7335_4445	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	Chro_1036	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	CYJSC1_DRAFT_64060	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
ApcD3	Cal7507_0259	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	UYEDRAFT_01167	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	UYKDRAFT_03497	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	S7335_2899	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	Chro_4357	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
ApcD4	UYEDRAFT_01908	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	Xen7305DRAFT_0006060	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	LEP6406DRAFT_0064	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								
	G107428_3841	KDPIERIGLVGVMYKSLGTPVDVAQAVREMKAVATGMM	SGDDAAEAGAYFDYVIGAME								

Figure S9. Sequence alignment of ApcA and ApcD allophycocyanin subunits. ApcA1 is the major allophycocyanin alpha subunit produced in WL, GL, and RL, and ApcA2 (ApcD5) is specifically produced when cells are grown in FRL. ApcD1 is the allophycocyanin-B alpha subunit produced in WL, GL, and RL, and ApcD2 and ApcD3 are subunits produced in cells grown in FRL. Sequences were aligned using the MUSCLE module implemented in MEGA6. Yellow shading indicates conserved cysteine residues. The conserved cysteine residue at position 81 typically forms a covalent thioether bond to phycocyanobilin. Additional details can be found in the main text.

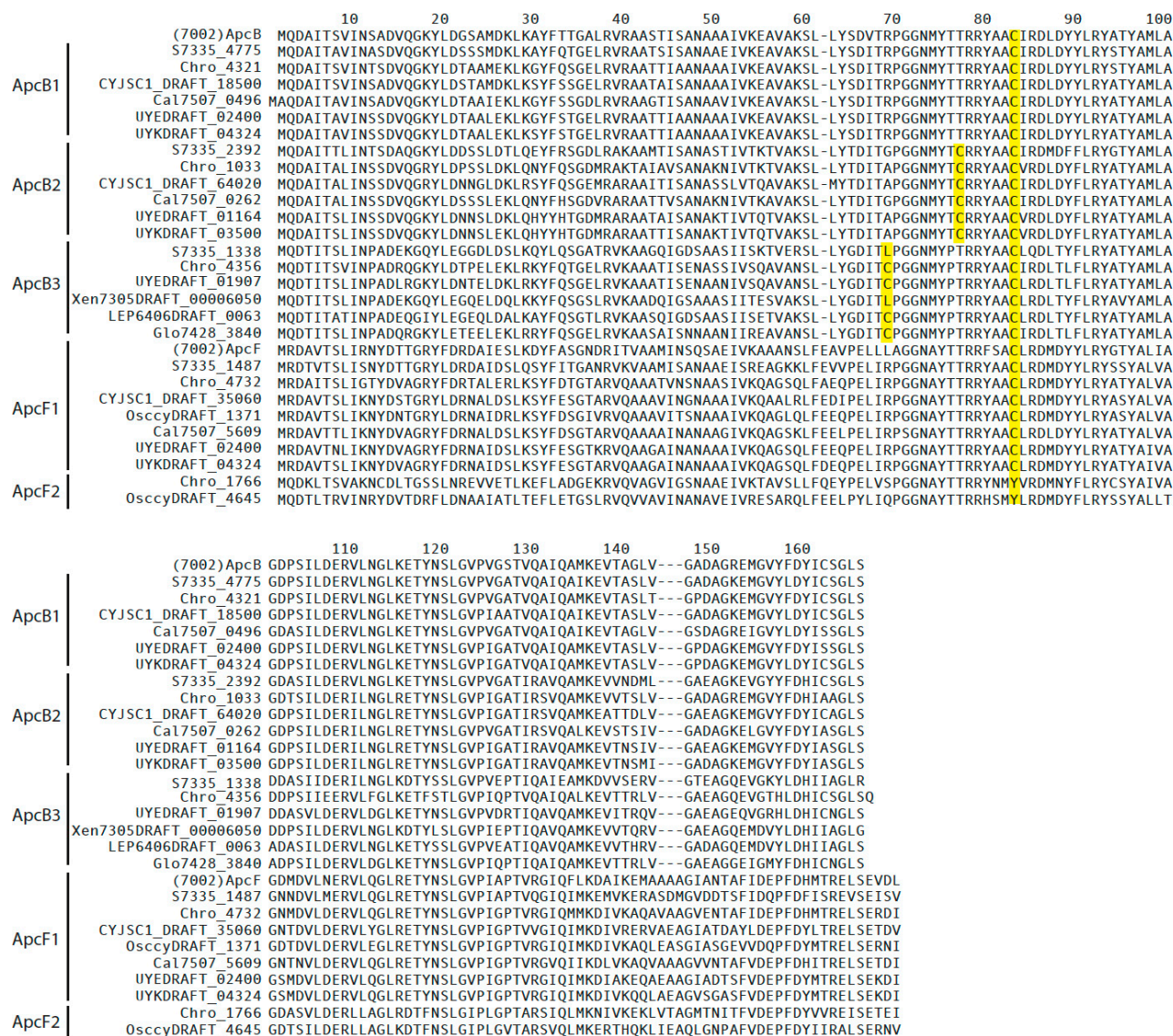


Figure S10. Sequence alignment of ApcB and ApcF subunits of allophycocyanin. ApcB1 is the major allophycocyanin beta subunit produced in WL, GL, or RL, and ApcB2 is the beta subunit of allophycocyanin produced in FRL. In most cyanobacteria, ApcF is a single-copy gene, the product of which apparently functions with both ApcE1 and ApcE2 subunits. Sequences were aligned using the MUSCLE module implemented in MEGA6. Yellow shading indicates conserved cysteine residues. The conserved cysteine residue at position 83 typically forms a covalent thioether bond to phycocyanobilin. Additional details can be found in the main text.