

Supplementary Material

Sequential splicing of a group II twintron in the marine cyanobacterium *Trichodesmium*

Ulrike Pfreundt and Wolfgang R. Hess

University of Freiburg, Faculty of Biology, Schänzlestr. 1, D-79104 Freiburg, Germany

*Corresponding author:

Prof. Wolfgang R. Hess, University Freiburg, Faculty of Biology, Inst. Biology III,
Schänzlestr. 1, D-79104 Freiburg, Germany

Tel. +49-(0)761-203-2796

FAX: +49-(0)761-203-2745

Email: wolfgang.hess@biologie.uni-freiburg.de

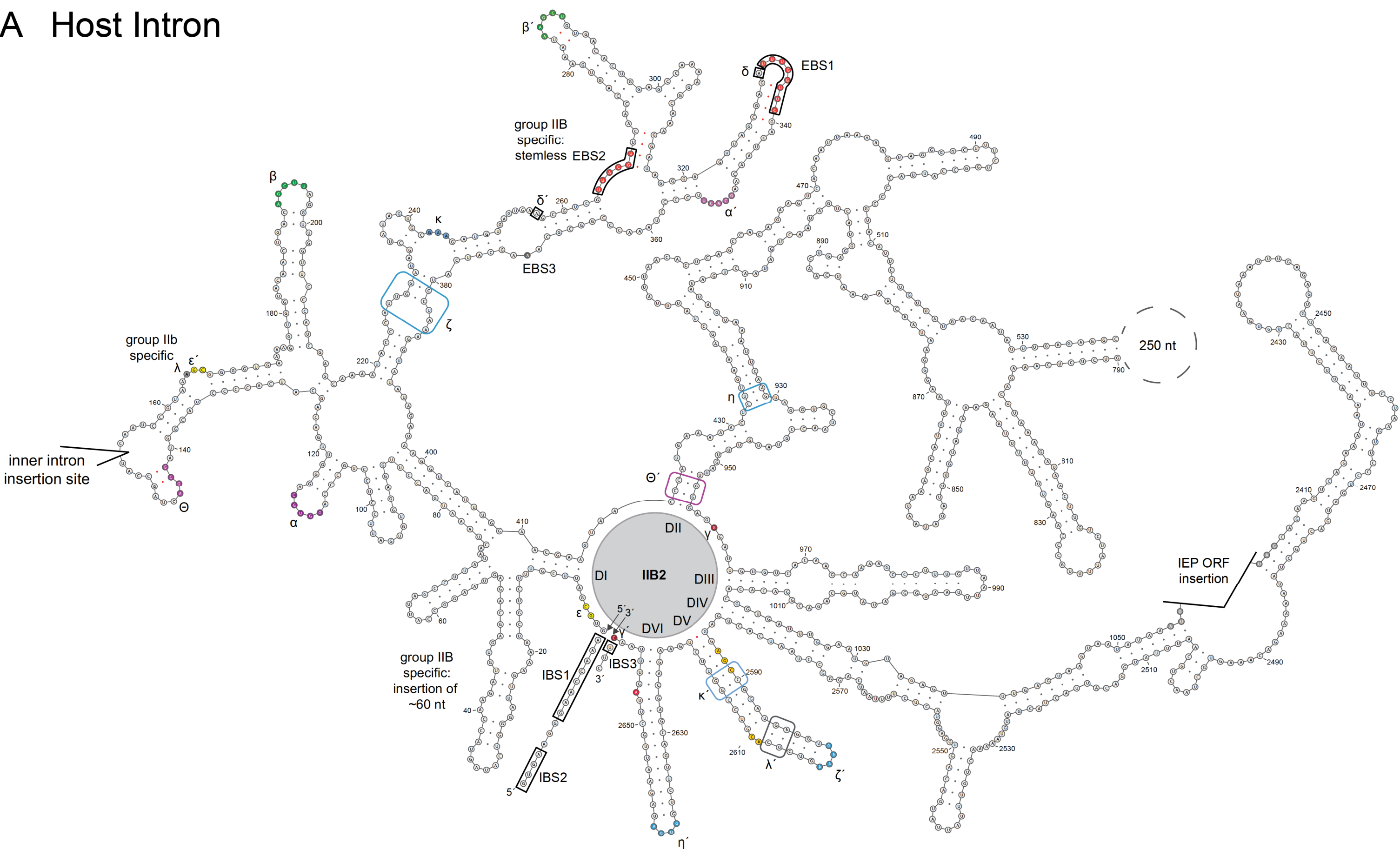
Supplementary Figure 1

RNA secondary structure of the group II host intron and the inner intron comprising the Tery_4732 twintron.

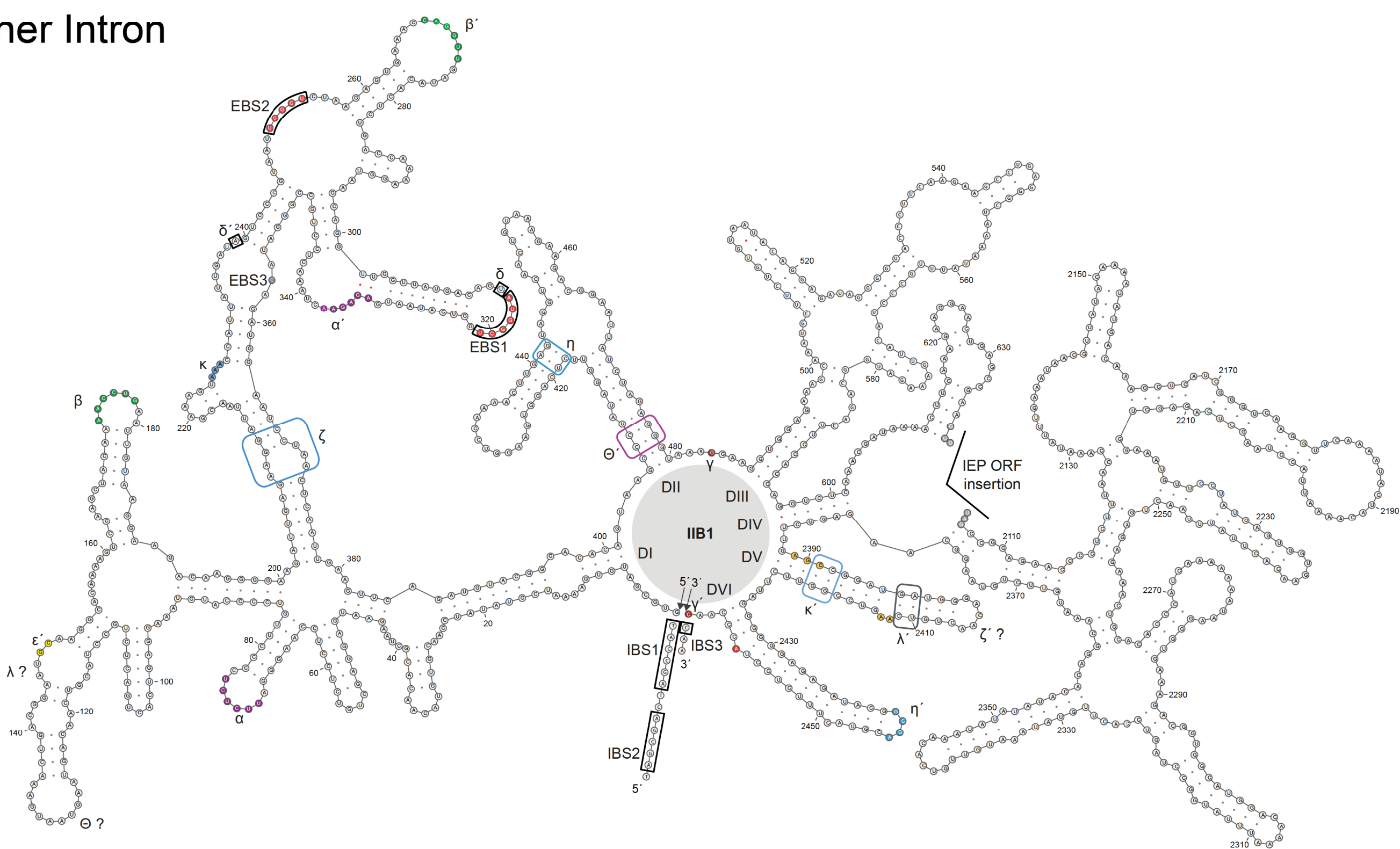
A) Mature host intron. The insertion site of the inner intron, within domain I (DI), and the IEP ORF situated in domain IV, are marked. A large extension of 250 nt within DII is indicated by a dashed loop. Structure prediction led to the classification of this intron as group II B2. Greek letters mark sequence regions that pair to form tertiary structures (for example β and β'). The bulging A in domain VI that usually initiates the first nucleophilic attack on the intron-exon boundary is colored red. EBS = exon binding site, IBS = intron binding site. Base pairs are indicated by gray dots; reds dots indicate paired bases that fit the consensus better if not paired.

B) Inner intron. Annotation is the same as for the host intron. This intron was classified as group II B1. Notably, DV does not fit the consensus structure; the 5-nt loop is atypical, so are λ and λ' including the adjacent bulge.

A Host Intron



B Inner Intron



Supplementary Alignment 1

Alignment of all full-length cyanobacterial hybrid RNase HI proteins together with all separately encoded RNase HI and *Tery_4732+* homologs that were concatenated for this alignment. To distinguish the latter, the term “concat” was added to the species name. Note that all proteins start on page 1 and the alignment is continued on the following pages. Species are sorted alphabetically. The amino acid similarity is color-coded, the darker the more similar.

The first part of the protein is reasonably well conserved (page 1-3), followed by a putative linker region of varying length and sequence. The C-terminal part (page 6), corresponding to *Tery_4732+*, the host gene of the *T. erythraeum* twintron, is highly conserved in all species except the picocyanobacteria and *Gloeobacter*.

					10	20	30	40	50	60	70																																																					
Acaryochloris_marina_MBIC11017/1-257	1	M	P	-	-	-	A	F	-	-	S	-	-	-	-	-	-	-	D	I	D	S	I	Y	T	D	G	A	C	S	G	N	P	G	P	G	G	W	S	V	I	Y	F	K	D	G	S	C	H	-	-	E	I	G	G	F	V	-	P	Q	T	46		
Acaryochloris_sp._CCMEE_5410/1-257	1	M	P	-	-	-	A	F	-	-	S	-	-	-	-	-	-	-	-	D	I	D	S	I	Y	T	D	G	A	C	S	G	N	P	G	P	G	G	W	S	V	I	Y	F	K	D	G	S	C	H	-	-	E	I	G	G	F	V	-	P	Q	T	46	
Anabaena_cylindrica/1-302	1	M	S	-	-	-	T	Q	-	-	P	-	-	-	-	-	-	-	-	I	Q	S	I	Y	T	D	G	A	C	T	G	N	P	G	P	G	G	W	V	V	V	F	Q	D	G	S	I	H	-	-	E	M	G	D	A	-	S	H	T	46				
Anabaena_sp._90/1-302	1	M	S	-	-	-	T	Q	-	-	S	-	-	-	-	-	-	-	-	L	I	Q	S	I	Y	T	D	G	A	C	T	G	N	P	G	P	G	G	W	T	V	V	Y	F	H	D	G	S	V	H	-	-	E	M	G	D	S	-	Q	H	T	46		
Anabaena_sp._PCC_7108/1-302	1	M	S	-	-	-	T	Q	-	-	R	-	-	-	-	-	-	-	-	I	Q	S	I	Y	T	D	G	A	C	T	G	N	P	G	P	G	G	W	A	V	V	V	F	N	D	G	S	I	H	-	-	E	M	G	D	A	-	R	H	T	46			
Anabaena_variabilis/1-302	1	M	S	-	-	-	T	Q	-	-	P	-	-	-	-	-	-	-	-	I	I	E	S	I	Y	T	D	G	A	C	T	G	N	P	G	P	G	G	W	V	V	V	F	S	D	G	S	V	H	-	-	E	M	G	D	A	-	K	H	T	46			
Aphanizomenon_flos-aquae/1-305	1	M	S	-	-	-	T	Q	-	-	R	-	-	-	-	-	-	-	-	L	I	Q	S	I	Y	T	D	G	A	C	T	G	N	P	G	P	G	G	W	T	V	V	Y	F	N	D	G	S	V	H	-	-	E	M	G	D	S	-	Q	H	T	46		
Arthrospira_sp._PCC_8005_concat/1-259	1	M	N	-	-	-	-	M	G	-	-	I	-	-	-	-	-	-	-	T	K	V	T	I	Y	T	D	G	A	C	S	G	N	P	G	K	G	Y	G	A	V	L	M	C	G	S	H	R	-	K	-	E	I	S	G	G	F	-	R	L	T	45		
Calothrix_parietina/1-301	1	M	S	-	-	-	-	A	R	-	-	-	-	-	-	-	-	-	-	T	I	E	S	I	Y	T	D	G	A	C	T	G	N	P	G	P	G	G	W	V	V	V	L	N	D	G	S	V	Y	-	-	E	I	G	D	G	-	R	Q	T	45			
Calothrix_sp._336/3/1-303	1	M	P	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	I	I	E	S	I	Y	T	D	G	A	C	T	G	N	P	G	P	G	G	W	T	V	V	F	Q	D	G	S	I	H	-	-	E	M	G	D	S	S	H	G	K	T	44		
Calothrix_sp._PCC_7103/1-301	1	M	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	I	E	S	I	Y	T	D	G	A	C	T	G	N	P	G	P	G	G	W	V	V	V	F	D	D	G	S	V	H	-	-	E	L	G	D	G	-	R	Q	T	45			
Calothrix_sp._PCC_7507/1-303	1	M	S	-	-	-	-	T	Q	-	-	H	-	-	-	-	-	-	-	K	I	Q	S	I	Y	T	D	G	A	C	T	G	N	P	G	P	G	G	W	T	V	V	F	S	D	S	I	H	-	-	E	M	G	D	A	-	S	H	T	46				
Candidatus_Atelocyanobacterium_thalassa_SIO64986_concat/1-249	1	M	K	-	-	-	-	K	-	-	-	-	-	-	-	-	-	-	-	L	K	I	Y	T	D	G	A	C	S	G	N	P	G	K	G	Y	G	T	I	L	V	F	N	K	Y	R	-	K	-	E	L	S	G	G	Y	-	R	L	T	41				
Candidatus_Atelocyanobacterium_thalassa_concat/1-249	1	M	K	-	-	-	-	K	-	-	-	-	-	-	-	-	-	-	-	L	K	I	Y	T	D	G	A	C	S	G	N	P	G	K	G	Y	G	V	I	L	V	F	N	K	Y	R	-	K	-	E	L	S	G	G	Y	-	R	L	T	41				
Candidatus_Synechococcus_spongiarum/1-242	1	M	T	-	-	-	-	R	P	-	-	-	-	-	-	-	-	-	-	H	L	I	T	A	A	C	D	G	A	C	S	G	N	P	G	P	G	G	W	A	Y	L	L	R	F	D	D	G	R	V	E	-	-	E	E	G	G	F	S	-	E	N	T	45
Chamaesiphon_minutus/1-286	1	M	S	-	-	-	-	R	-	-	-	-	-	-	-	-	-	-	-	I	T	N	L	Y	T	D	G	A	C	S	G	N	P	G	P	G	G	W	A	A	V	A	F	D	D	G	S	C	H	-	-	E	L	G	G	R	D	-	S	A	T	43		
Chlorogloeopsis_fritschii/1-304	1	M	S	-	-	-	-	S	V	-	-	P	-	-	-	-	-	-	-	T	I	E	S	I	Y	T	D	G	A	C	T	G	N	P	G	P	G	G	W	I	V	V	F	F	D	G	S	I	H	-	-	E	I	G	D	A	-	K	Q	T	46			
Chroococciopsis_thermalis/1-306	1	M	P	-	-	-	-	I	S	-	-	A	-	-	-	-	-	-	-	N	I	Q	S	I	Y	T	D	G	A	C	T	G	N	P	G	P	G	G	W	A	A	V	V	F	T	D	G	S	L	H	-	-	E	I	G	G	S	A	-	Q	T	46		
Coleofasciculus_chthonoplastes/1-302	1	M	P	-	-	-	-	V	T	-	-	R	-	-	-	-	-	-	-	Q	I	K	S	V	Y	T	D	G	A	C	S	G	N	P	G	P	G	G	W	T	V	V	Y	F	T	D	G	S	V	H	-	-	E	I	G	G	A	T	-	A	K	T	46	
Crinalium_epipsammum/1-305	1	M	F	-	-	-	-	-	S	D	-	-	R	-	-	-	-	-	-	T	I	K	S	I	Y	T	D	G	A	C	S	G	N	P	G	P	G	G	W	V	V	V	Y	S	D	N	A	V	A	-	-	E	M	G	D	G	E	-	R	A	T	46		
Crocospaera_watsonii_concat/1-248	1	M	N	-	-	-	-	K	-	-	-	-	-	-	-	-	-	-	-	V	Q	I	Y	T	D	G	A	C	S	G	N	P	G	K	G	Y	G	I	I	L	A	Y	N	E	H	R	-	K	-	E	L	S	G	G	Y	-	R	L	T	41				
Cyanobacterium_stanieri_PCC_7202_concat/1-256	1	M	N	-	-	-	-	K	E	-	-	S	-	-	-	-	-	-	-	L	Q	E	V	E	Y	T	D	G	S	C	L	H	N	P	G	A	G	Y	G	V	V	L	L	Y	G	K	H	R	-	K	-	E	F	S	G	G	F	-	S	L	T	46		
Cyanobium_gracile/1-269	1	M	G	-	-	-	-	-	D	R	-	-	P	-	-	-	-	-	-	V	R	V	V	A	A	A	C	D	G	A	C	S	G	N	P	G	P	G	G	W	A	L	L	R	F	E	D	G	S	R	R	-	-	E	L	G	G	A	D	-	A	A	T	47
Cyanothece_sp._CCY0110_concat/1-246	1	M	K	-	-	-	-	K	-	-	-	-	-	-	-	-	-	-	-	V	Q	I	Y	T	D	G	A	C	S	G	N	P	G	K	G	Y	G	I	I	L	V	Y	N	E	H	R	-	K	-	E	L	S	G	G	Y	-	R	L	T	41				
Cyanothece_sp._PCC_7425_concat/1-254	1	M	N	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	E	I	S	K	I	Y	T	D	G	A	C	S	G	N	P	G	P	G	G	W	V	V	V	F	F	D	G	S	V	H	-	-	E	M	G	G	A	-	A	A	T	44			
Cylindrospermopsis_raciborskii/1-306	1	M	S	-	-	-	-	N	Q	-	-	R	-	-	-	-	-	-	-	I	I	K	S	I	Y	T	D	G	A	C	K	G	N	P	G	P	G	G	W	V	V	V	F	H	D	G	S	I	H	-	-	E	M	G	E	A	-	D	Y	T	46			
Cylindrospermum_stagnale/1-301	1	M	S	-	-	-	-	T	Q	-	-	R	-	-	-	-	-	-	-	T	I	Q	S	I	Y	T	D	G	A	C	T	G	N	P	G	P	G	G	W	A	V	V	V	F	N	D	G	S	V	H	-	-	E	M	G	D	R	A	-	S	H	T	46	
Dactylococcopsis_salina/1-282	1	M	K	-	-	-	-	P	-	-	-	-	-	-	-	-	-	-	-	V	I	D	C	I	Y	T	D	G	A	C	L	G	N	P	G	P	G	G	W	A	V	L	Y	F	Q	D	G	S	V	Y	-	-	E	I	G	G	A	E	-	R	E	T	44	
Dolichospermum_circinale1/1-305	1	M	S	-	-	-	-	N	Q	-	-	R	-	-	-	-	-	-	-	I	I	E	S	I	Y	T	D	G	A	C	T	G	N	P	G	P	G	G	W	T	V	I	Y	F	C	D	G	S	V	H	-	-	E	M	G	D	S	-	Q	H	T	46		
Dolichospermum_circinale2/1-303	1	M	S	-	-	-	-	N	Q	-	-	R	-	-	-	-	-	-	-	I	I	E	S	I	Y	T	D	G	A	C	T	G	N	P	G	P	G	G	W	T	V	I	Y	F	C	D	G	S	V	H	-	-	E	M	G	D	S	-	Q	H	T	46		
Fischerella_muscicola1/1-304	1	M	S	-	-	-	-	F	Q	-	-	H	-	-	-	-	-	-	-	T	I	A	S	I	Y	T	D	G	A	C	I	G	N	P	G	P	G	G	W	V	V	V	F	D	D	G	S	V	H	-	-	E	M	G	D	A	-	A	K	T	46			
Fischerella_muscicola2/1-304	1	M	S	-	-	-	-	S	Q	-	-	R	-	-	-	-	-	-	-	T	I	Q	S	I	Y	T	D	G	A	C	I	G	N	P	G	P	G	G	W	V	V	V	F	S	D	G	S	V	H	-	-	E	I	G	D	T	A	-	A	K	T	46		
Fischerella_sp._PCC_9339/1-304	1	M	S	-	-	-	-	S	Q	-	-	R	-	-	-	-	-	-	-	T	I	Q	S	I	Y	T	D	G	A	C	I	G	N	P	G	P	G	G	W	V	V	V	F	S	D	G	S	V	H	-	-	E	I	G	D	A	-	A	K	T	46			
Fischerella_sp._PCC_9431/1-304	1	M	S	-	-	-	-	S	Q	-	-	R	-	-	-	-	-	-	-	T	I	Q	S	I	Y	T	D	G	A	C	I	G	N	P	G	P	G	G	W	V	V	V	F	S	D	G	S	V	H	-	-	E	I	G	D	A	-	A	K	T	46			
Fischerella_sp._PCC_9605/1-304	1	M	S	-	-	-	-	Y	Q	-	-	R	-	-	-	-	-	-	-	T	I	Q	S	I	Y	T	D	G	A	C	T	G	N	P	G	P	G	G	W	V	V	V	F	S	D	G	S	I	H	-	-	E	I	G	D	A	-	P	K	T	46			
Geitlerinema_sp._PCC_7407/1-282	1	M	G	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	L	I	T	T	I	Y	T	D	G	A	C	S	G	N	P	G	P	G	G	W	V	V	N	F	A	D	G	T	V	Q	-	-	E	L	G	G	H	E	-	A	Q	T	44		
Geminocystis_herdmanii_concat/1-253	1	M	S	-	-	-	-	D	K	-																																																						

	80	90	100	110	120	130	140	150
Acaryochloris_marina_MBIC11017/1-257	47	TNNRME	LQAA	I A I L	ELLQDQP	L P K T I P L Y	T D S E Y V K N G I T K	- W L A G W K R K G W K T A K G K A V L N Q D L W Q Q L D Q L - N D 119
Acaryochloris_sp._CCMEE_5410/1-257	47	TNNRME	LQAA	I A I L	ELLQDQP	- L P K T I P L Y	T D S E Y V K N G I T K	- W L A G W K R K G W K T A K G K A V L N Q D L W Q Q L D Q L - N D 119
Anabaena_cylindrica/1-302	47	TNN	KMEMQA	A A I A A L	E F L E T S G - Q T Q P I T L Y	T D S E Y L I N S V T K	- W M P G W K K K G W K K S D G N P V Q N Q D L L E T L D Q L - N S 119	
Anabaena_sp._90/1-302	47	TNN	KMEMQA	A A I A A L	E F F K S S R - Q T E P I T L Y	T D S K Y V I D C V T K	- W V K G W K K K G W K K A D G N P V Q N Q D L L E I L D Q L - N S 119	
Anabaena_sp._PCC_7108/1-302	47	TNN	KMEMQA	A A I A A L	Q Y L E A S G - Q N E P I T L Y	T D S E Y L I N C V T K	- W V K G W K K K G W K K S D G N P V Q N Q D L L E I L D E L - N S 119	
Anabaena_variabilis/1-302	47	TNN	KMEMQA	A A I A A L	K F L H D S G - Q T E P I T L Y	T D S E Y L I N C V T K	- W V K G W K K K G W K K S D G N P V Q N Q D L L E T L D E L - N S 119	
Aphanizomenon_flos-aquae/1-305	47	TNN	KMEMQA	A A I A A L	E F F Q S S G - Q T Q P I T L Y	T D S K Y V I D C V T K	- W V K S W K K K G W K K S D G N P V Q N Q D L L E I L D Q L - N S 119	
Arthrospira_sp._PCC_8005_concat/1-259	46	TNNRME	MMAA	I A A L	R A L K F P - - - C S V T L Y	S D S K Y L V D A M T L G W A K R W Q K N G W R N Q K E W A K N P D L W A Q L L G L C E E 117		
Calothrix_parietina/1-301	46	TNNRME	MMQA	A A I A A L	Q Y F Q Q S G - Q S E P I T L Y	T D S E Y L I N C V T K	- W V K S W K K K D W K K A D G K P V N Q D L L E V L D E L - N N 118	
Calothrix_sp._336/3/1-303	45	TNN	KMEMQA	A A I A A L	E Y F H A S G - Q T E P I T L Y	T D S E Y L I N C V T K	- W V K G W K R K G W K K A D G K P V L N Q D L L E K L D D L - N N 117	
Calothrix_sp._PCC_7103/1-301	46	TNNRME	MMQA	A A I A A L	Q Y L E A S G - Q S Y P I T L Y	T D S E Y L I N C V T K	- W V K S W K K K G W K K A D G K P V N Q D L L E I L D G L - N S 118	
Calothrix_sp._PCC_7507/1-303	47	TNN	KME I QAA	I A A L	Q Y L Q A S G - Q T E P I T L Y	T D S E Y V I K S I T Q	- W M K G W K K R G W K K A D G N P V Q N Q D L L Q I L D D L - N S 119	
Candidatus_Atelocyano bacterium_thalassa_SIO64986_concat/1-249	42	TNNRME	I I S I I	I G L E S L I E S - - - C N V S I Y	S D S R Y V V D S M N R G W A S K W Q S N D W K R N K Q E I A K N S D L W K R L L D L C T K 113			
Candidatus_Atelocyano bacterium_thalassa_concat/1-249	42	TNNRME	I I S I I	I G L E A L T K L - - - C N V T I Y	S D S R Y V V D T M N K G W A S K W Q S N N W K R N K Q E I A K N S D L W K R L L D L C T K 113			
Candidatus_Synechococcus_spongjarum/1-242	46	TNNRME	MMQA	A A L L K R I A Q E P - R Q P G F A I R T D	S K Y L I Q G F S Q	- W L K G W K Q K G W K K A D G Q C V K N R D L W E E L D Q A - R L 118		
Chamaesiphon_minutus/1-286	44	TNNRME	MMQA	A I S A L E F F V T N D - R A E P C V V Y T D	S E Y L I K G I T Q	- W V K G W K K K G W K T A S G K P V L N Q D L W E T L D R L - D R 116		
Chlorogloeopsis_fritschii/1-304	47	TNNRME	MMQA	A A I A A L	E F L K A S G - Q S Q P I T L Y	T D S E Y L I N C V T K	- W V T K W R K G W K K A D G K P V L N Q D L L E I L D Q L - N T 119	
Chroococcidiopsis_thermalis/1-306	47	TNNRME	MMQA	A A I A A L	Q L I A E S S - H L S A V T I H T D	S E Y L I N C V T K	- W V K G W K K K G W K T A Q G K P V L N Q D L W E T L D R L - N H 119	
Coleofasciculus_chthonoplastes/1-302	47	TNNRME	L L A A I	E A L I L V Q G S G - Q T Q P I T L Y	T D S E Y V K N G I T K	- W I K G W K K K G W K T S Q G K A V L N Q D L W E K L D A V - N S 119		
Crinalium_epipsammum/1-305	47	TNNRME	M L A A I	A A L K F I A K S Q - Q T E P I T L Y	T D S E Y V L K G I T Q	- W V K G W K K K G W K T S Q G K A V L N Q D I W E I L D Q L - N S 119		
Crocospaera_watsonii_concat/1-248	42	TNNRME	MMAA	I I A L E A L N K P - - - C D V L Y	T D S R Y V V D A I T K G W A K K W Q A N D W R N K K E Q A K N P D L W Q R L L D L C E Q 113			
Cyanobacterium_stanieri_PCC_7202_concat/1-256	47	TNNRME	I F A T I	V G L R S L K K P - - - C K V R L Y	T D S R Y V V D A I N K G W A K Q K W E K N G W K R N S K E A K N P D L W A D L L A L C R S 118			
Cyanobium_gracile/1-269	48	TNNRME	L T A A L	A V L E A L R E L P - R H P D L V L R	T D S R Y L I D G F E K	- W M A G W K R K G W R T A S G G A V L N R D L W E R L D A A - R L 120		
Cyanothece_sp._CCY0110_concat/1-246	42	TNNRME	MMAA	I I G L E A L K T P - - - C E V T L Y	T D S R Y L V D A I T K G W A K K W Q A N G W K R N N K E A A K N P D L W Q K L L D L C K K 113			
Cyanothece_sp._PCC_7425_concat/1-254	45	TNNRME	MMQA	A A L E F C Q A A G - - H P P V I L Y	T D S E Y V K K G I T E	- W L P G W K R K G W K T A Q G K P V L N Q D L W Q V L D Q L - N T 116		
Cylindrospermopsis_raciborskii/1-306	47	TNNQME	I QAA I	G A L E L L A Q S G - Q V E P I T L Y	S D S E Y L I N S V T K	- W I H G W K K K G W K K S D G T P V Q N Q E L L K I I D K L - N T 119		
Cylindrospermum_stagnale/1-301	47	TNN	KMEMQA	A A I A A L	Q F L E T S G - Q T E P I T L Y	T D S E Y L I N C V T K	- W V T G W K K K G W K K S D G N P V Q N Q D L L E T L D E L - N S 119	
Dactylococcopsis_salina/1-282	45	TNNRME	L QAA I	A S L Q L F R D S G - Q T K P I Q L Y	S D S Q Y L I K G L T Q	- W L K G W K K K G W K T S T G K A V L N Q D L W Q L L D Q L - N N 117		
Dolichospermum_circinale1/1-305	47	TNN	KMEMQA	A A I A A L	E F F Q S S G - Q T Q P I T L F	T D S E Y L I N S I T K	- W M P G W K R K G W K K S D G N P V Q N Q D L L Q T L D Q L - N S 119	
Dolichospermum_circinale2/1-303	47	TNN	KMEMQA	A A I A A L	E F F K S S G - Q T Q P I T L F	T D S E Y L I N S I T K	- W M P G W K R K G W K K S D G N P V Q N Q D L L Q T L D Q L - N S 119	
Fischerella_muscicola1/1-304	47	TNN	KMEMQA	A A I A A L	S F L K A S G - Q T Q P I T L Y	S D S E Y L I N C V T K	- W I K T W K R K G W K K A D G S P V L N R D L L E Q L E Q L - N S 119	
Fischerella_muscicola2/1-304	47	TNN	KMEMQA	A A I A A L	S F L K A S G - H T E P I T L Y	T D S E Y L I N C V T K	- W I K T W K R K G W K K A D G N P V L N Q D L L E Q L E Q L - N S 119	
Fischerella_sp._PCC_9339/1-304	47	TNN	KMEMQA	A A I A A L	S F L K A S G - H T E P I T L Y	T D S E Y L I N C V T K	- W I K T W K R K G W K K A D G N P V L N Q D L L E Q L E Q L - N S 119	
Fischerella_sp._PCC_9431/1-304	47	TNN	KMEMQA	A A I A A L	S F L K A S G - Y T E P I T L Y	T D S E Y L I N C V T K	- W I K T W K R K G W K K A D G N P V L N Q D L L E Q L E Q L - N S 119	
Fischerella_sp._PCC_9605/1-304	47	TNN	KMEMQA	A A I A A L	Q F L K A S G - Q T E P I T L Y	T D S E Y L I N C V T K	- W V T K W R K G W K K A D G N P V L N Q D L L E E L D Q L - N S 119	
Geitlerinema_sp._PCC_7407/1-282	45	TNNRME	MMQA	A A L E F L R D R P - Q S E P I A L H T D	S E Y V K K G V T Q	- W L A G W K R K G W K T S T G S A V L N Q D L W E T L D R L - N S 117		
Geminocystis_herdmanii_concat/1-253	46	TNNRME	I L A V I	K G L Q A L K K P - - - C R V N L Y	T D S Q Y V V N A I S K G W A K K W Q K N N W Q R N A K E K A K N P D L W Q I L L E L C E Q 117			
Gloeobacter_kilauensis/1-220	40	TNNRME	MQAV I	E G L R R V - - - G - A T V P V T V I	S D S Q Y V I K G M S E	- W L P G W K R R Q W R T A A G K P V E N R D L W E A L E K V - A A 109		
Gloeobacter_violaceus/1-213	46	TNNRME	MR A V I	E G L T R V - - - P - A D A K V K V L	T D S Q Y V I K G M S E	- W L P G W K R R G W I T S T G K P V E N R D L W E A L E R A - V G 109		
Gloeocapsa_sp._PCC_7428/1-297	46	TNNRME	MMQA	A A L E L L S T V P - T N E P I T L Y	T D S E Y L I K G I T Q	- W V K G W K K K G W K T A Q G K P V L N Q D L W E L D E R - N 117		
Halothece_sp._PCC_7418/1-283	47	TNNRME	MMQA	A S L E L F Q A S E - Q K T P I V L Y	S D S Q Y L I K G L T Q	- W L K G W K K K G W K T S T G K A V L N Q D L W T T L D Q L - N H 119		
Leptolyngbya_boryana/1-263	46	TNNRME	MMQA	A A L K F F K D S G - Q S K V I A L Y	T D S E Y V K N G I T K	- W L P G W K K K G W K T S A G K P V L N Q D L W E Q L D E L - D S 118		
Leptolyngbya_sp._PCC_7375/1-288	47	TNNRME	MMQA	A A L Q F L Q H N D - Y D S P I E L N T D	S K Y L I Q G I T Q	- W M S G W K R K G W K T S A G K A V L N Q D L W M T L D K L - K Q 119		
Leptolyngbya_sp._PCC_7376/1-293	42	TNNRME	MMQA	A I A L K Q F R D S D - Q R E P I Q L L	T D S K Y V I D G I T K	- W I K G W K R K G W K K S S G A P V L N K D L W Q E M D K L - N H 114		
Lyngbya_aestuarii_putative_concat/1-271	60	TNNRME	MMAA I	V G L E V L E K K - - - S K V T L Y	S D S K Y V V D A I E K G W A E R W Q T N G W K R N K K E L A M N P D L W E Q L L K L C S Q 131			
Lyngbya_sp._PCC_8106_concat/1-283	61	TNNRME	MMAA I	V G L E V L D K K - - - S K V K L Y	S D S K Y V V D A I E K G W A E R W Q A N G W K R N K K E L A M N P D L W E Q L L K L C S Q 132			
Mastigocladopsis_repens/1-301	47	TNN	KMEMQA	A A I A A L	E F L K D S G - Q T E P I P L Y	T D S E Y L I N C V T K	- W V R G W K K R G W K K A D G S P V L N Q D L L E T L D D L - N S 119	
Mastigocoleus_testarum/1-302	47	TNNRME	MMQA	A A L Q F L K T T E - Q R D S I T L Y	T D S E Y L I N C V T K	- W V G W K K K G W K K S N G K P V L N Q D L L V T L D L - N T 119		
Microchaete_sp._PCC_7126/1-302	47	TNN	KMEMQA	A A I A A L	K F F K T S G - Q T Q S I T L Y	T D S E Y L I N C V T K	- W I K G W K K K G W K K A D G N P V Q N Q D L L E I L D D L - N S 119	
Microcoleus_sp._PCC_7113/1-303	47	TNNRME	MMQA	A I Q A L K V L K A S A - Q T E S V I L Y	T D S E Y V K N G V T K	- W V K G W K K K G W K T A Q G K A V L N Q D L W E T L D E L - N S 119		
Microcoleus_vaginatus/1-305	46	TNNRME	MMQA	A V A A L E F F A N S G - Q T E P V T V Y	T D S E Y V K N G I T K	- W I Q G W K N K G W K T S T G K D V V N Q D L W H L L D K L - N S 118		
Neosynechococcus_sphagnicola_sy1_concat/1-227	43	TNNRME	MMAA I	Q A L Q S L T R P - - - C Q V R L H	T D S Q L L V N T M T R N W K -			

		160	170	180	190	200	210	220		
Acaryochloris_marina_MBIC11017/1-257	120	-----A-----	TIDWRYIRAH	GGHEGNERCDQ	IA-----	RAFSQYQT-P-DLH	-----Q-----	I-T---	158	
Acaryochloris_sp._CCMEE_5410/1-257	120	-----A-----	TIDWRYIRAH	AGHEGNERCDQ	IA-----	RAFSQHQT-P-DLH	-----Q-----	I-T---	158	
Anabaena_cylindrica/1-302	120	-----R-----	L VKWQHVRGH	SGNIGNERCDA	IA-----	RSFASGKI-P-SLQ	-----Q-----	L-S---	158	
Anabaena_sp._90/1-302	120	-----R-----	LQVWEHVRGH	SGNIGNERCDA	IA-----	RSFANHKI-P-HLK	-----Q-----	L-L---	158	
Anabaena_sp._PCC_7108/1-302	120	-----R-----	L VKWQHVRGH	SGNIGNERCDA	IA-----	RSFASSKI-P-SLQ	-----Q-----	L-S---	158	
Anabaena_variabilis/1-302	120	-----R-----	KVNWHHVRGH	SGNIGNERCDA	IA-----	RCFATGRM-P-SLQ	-----Q-----	L-S---	158	
Aphanizomenon_flos-aquae/1-305	120	-----Q-----	RVQWQHVRGH	SGNVGNERCDL	IA-----	RSFANHKI-P-HLK	-----Q-----	S-L---	158	
Arthrospira_sp._PCC_8005_concat/1-259	118	-----H-----	QVRFVWVKGH	AGDRENEICDR	LA-----	VEATHRDS-L-PPD	-----A-----	G-Y---	156	
Calothrix_parietina/1-301	119	-----R-----	KVKWQHVRGH	AGNIGNERCDA	IA-----	RAFATGKK-P-QLK	-----Q-----	E-S---	157	
Calothrix_sp._336/3/1-303	118	-----R-----	KI KWEYVRGH	AGNIGNERCDA	IA-----	RAFASGNV-P-SLQ	-----S-----	I-S---	156	
Calothrix_sp._PCC_7103/1-301	119	-----R-----	KVRWEHVRGH	AGNIGNERCDA	IA-----	RAYAAGKV-P-VLK	-----Q-----	S-Q---	157	
Calothrix_sp._PCC_7507/1-303	120	-----D-----	LAVNWQHVRGH	SGNIGNERCDA	IA-----	RAFANGKI-P-SLQ	-----Q-----	S-S---	159	
Candidatus_Atelocyanobacterium_thalassa_SIO64986_concat/1-249	114	-----H-----	NVQFIWVKGH	AGHPENERCDQ	LA-----	VLASKKME-L-FID	-----K-----	V-F---	152	
Candidatus_Atelocyanobacterium_thalassa_concat/1-249	114	-----H-----	NVQFIWVKGH	SGHPENERCDQ	LA-----	VLASKRME-L-FID	-----K-----	V-F---	152	
Candidatus_Synechococcus_spongiarum/1-242	119	-----P-----	GVALTHVKGH	SGDPDNDRVD	IA-----	VAFRDGKQ-P-ALR	-----S-----	G-S---	157	
Chamaesiphon_minutus/1-286	117	-----Q-----	QVQSQT KAPV	SWQHVRGH	SGNVGNERCDT	IA-----	RGFSSGQP-P-KLR	-----Q-----	R-S---	162
Chlorogloeopsis_fritschii/1-304	120	-----R-----	QVKWQHVRGH	AGNEGNERCDA	IA-----	RAFACGKT-P-LLK	-----Q-----	I-T---	158	
Chroococcidiopsis_thermalis/1-306	120	-----P-----	SI EWRHVRGH	AGIIGNERCDA	IA-----	RAFASGKP-P-HLQ	-----Q-----	Q-S---	158	
Coleofasciculus_chthonoplastes/1-302	120	-----P-----	LVTWQYVRGH	SGNQGNERCDQ	IA-----	NAFANGKS-P-KLT	-----Q-----	V-L---	158	
Crinalium_epipsammum/1-305	120	-----D-----	DIQ-QQGTSLN	WQYVRGH	AGNVGNERCDA	IA-----	RGFAIGKP-P-LLK	-----Q-----	L-T---	164
Crocospaera_watsonii_concat/1-248	114	-----H-----	QVEFVWVKGH	AGHPENEQCDR	LA-----	VAAQCEVE-L-SID	-----A-----	V-Y---	152	
Cyanobacterium_stanieri_PCC_7202_concat/1-256	119	-----H-----	EVEFIWVRGH	AGNKENERCDR	LA-----	FDAAHGNN-L-PPD	-----E-----	N-Y---	157	
Cyanobium_gracile/1-269	121	-----P-----	GLALRHVRGH	SGDPDNDRCDA	IA-----	VAFSRGQS-P-ALA	-----T-----	E-A---	159	
Cyanothece_sp._CCY0110_concat/1-246	114	-----H-----	EVKFVWVKGH	AGHPENEQCDR	LA-----	VTATQQLT-L-AID	-----E-----	V-Y---	152	
Cyanothece_sp._PCC_7425_concat/1-254	117	-----K-----	SVNWQYVRGH	SGNAGNERCDA	IA-----	RAFAAGRP-P-ELR	-----Q-----	S-R---	155	
Cylindrospermopsis_raciborskii/1-306	120	-----K-----	LVRWQHVRGH	AGNIGNERCDA	IA-----	RSFASGEI-P-PLQ	-----Q-----	L-S---	158	
Cylindrospermum_stagnale/1-301	120	-----K-----	RVKWQHVRGH	SGNIGNERCDA	IA-----	RSYASGKI-P-SLQ	-----Q-----	L-P---	158	
Dactylococcopsis_salina/1-282	118	-----S-----	LVTWKHVP	AHQGIAGNERCDA	IA-----	RHFAQQGQ-P-NLK	-----K-----	T-L---	156	
Dolichospermum_circinale1/1-305	120	-----Q-----	L I KWQHVRGH	SGNIGNERCDA	IA-----	RCFANGKT-P-NLQ	-----Q-----	L-F---	158	
Dolichospermum_circinale2/1-303	120	-----Q-----	L I KWQHVRGH	SGNIGNERCDA	IA-----	RCFANGKT-P-NLQ	-----Q-----	L-F---	158	
Fischerella_muscolola1/1-304	120	-----R-----	QVKWQHVRGH	AGNEGNERCDA	IA-----	RAFASGKT-P-VLR	-----K-----	I-S---	158	
Fischerella_muscolola2/1-304	120	-----R-----	QVKWQHVRGH	TGNEGNERCDA	IA-----	RAFAGGKN-P-SLQ	-----Q-----	I-S---	158	
Fischerella_sp._PCC_9339/1-304	120	-----R-----	QVKWQHVRGH	TGNEGNERCDA	IA-----	RAFASGKN-P-SLQ	-----Q-----	I-S---	158	
Fischerella_sp._PCC_9431/1-304	120	-----R-----	QVRWQHVRGH	TGNEGNERCDA	IA-----	RAFAGGKN-P-SLQ	-----Q-----	I-S---	158	
Fischerella_sp._PCC_9605/1-304	120	-----S-----	Q I KWQHVRGH	AGNEGNERCDA	IA-----	RAFADGKT-P-SLR	-----Q-----	V-S---	158	
Geitlerinema_sp._PCC_7407/1-282	118	-----S-----	KVSWHYVRGH	AGNKDNERCDA	IA-----	RVFAQGKA-P-NLR	-----Q-----	S-S---	156	
Geminocystis_herdmanii_concat/1-253	118	-----H-----	QVNF I WVRGH	SGNRENERCDR	LA-----	FNAAQSSD-L-EED	-----S-----	N-Y---	156	
Gloeobacter_kilaueensis/1-220	110	-----G-----	RVQWQHVVYGH	RHGAENERAND	LA-----	QAAAARRP-VPP	AVEPAAFAT-----	S-----	154	
Gloeobacter_violaceus/1-213	110	-----G-----	RVLWEHVRGH	TGHPENERANT	LA-----	QTAARGGA-P-AVG	-----A-----	R-----	147	
Gloeocapsa_sp._PCC_7428/1-297	118	-----K-----	NVQWQYVRGH	AGNVGNERCDK	IA-----	RAFATGKT-P-TLQ	-----Q-----	P-K---	156	
Halothece_sp._PCC_7418/1-283	120	-----S-----	L I EWKHVRGH	EGIAGNERCDA	IA-----	RTFAAGKR-P-HLK	-----K-----	N-L---	158	
Leptolyngbya_boryana/1-263	119	-----P-----	KVEWRYVRGH	SGDPGNERCDV	IA-----	RAFSGQKM-P-QLK	-----M-----	P-----	156	
Leptolyngbya_sp._PCC_7375/1-288	120	-----E-----	ELATSQGI	DIDWRYVKGH	AGNVGNERCDA	IA-----	RAYSHQKS-I-RLK	-----Q-----	L-T---	165
Leptolyngbya_sp._PCC_7376/1-293	115	-----P-----	AVQWTHV	KGHSGEPGNERAD	IA-----	NSFARGKA-P-I LK	-----T-----	P-R---	153	
Lyngbya_aestuarii_putative_concat/1-271	132	-----H-----	QVKFVWVKGH	AGNRENECCDR	LA-----	VQACQQQN-L-LQD	-----L-----	G-Y---	170	
Lyngbya_sp._PCC_8106_concat/1-283	133	-----H-----	QVKFVWVKGH	AGNRENECCDR	LA-----	VQGCQQQN-L-LQD	-----V-----	G-YENP	174	
Mastigocladopsis_repens/1-301	120	-----P-----	RVKWEHVRGH	AGNVGNERCDA	IA-----	RAFANGKT-L-SLQ	-----Q-----	Q-D---	158	
Mastigocoleus_testarum/1-302	120	-----K-----	Q I KWQHVRGH	SGNTGNERCDQ	IA-----	TTFASGKY-P-QLK	-----Q-----	S-L---	158	
Microchaete_sp._PCC_7126/1-302	120	-----N-----	L VNWQHVRGH	SGNIGNERCDA	IA-----	RAFASGKT-P-SLQ	-----Q-----	S-F---	158	
Microcoleus_sp._PCC_7113/1-303	120	-----P-----	LVEWQYVRGH	SGNEGNERCDA	IA-----	RAFAVGKP-L-VLD	-----Q-----	A-F---	158	
Microcoleus_vaginatus/1-305	119	-----R-----	Q I KWEHVRGH	SGDFYNDRCDA	IA-----	RSFSHNLT-P-ELQ	-----Q-----	Q-K---	157	
Neosynechococcus_sphagnicola_sy1_concat/1-227	101	-----H-----	QVEWTWVKGH	GSDPWN I	RADELA-----	VKAAQAI I-Q-R	-----	-----	134	
Nodosilinea_nodulosa_concat/1-258	117	EVNQ	TLDR-----	PLQWVYVRGH	SGDVGNERCDT	IA-----	RAFAANRA-I-ALA	-----T-----	A-S---	162
Nodularia_spumigena/1-312	120	-----R-----	QVNWQHVRGH	SGNIGNERCDA	IA-----	RSFASGQI-P-SLQ	-----E-----	L-S---	158	
Nostoc_punctiforme/1-302	120	-----Q-----	QVKWQHVRGH	SGNIGNERCDA	IA-----	RTYASGKI-P-SLQ	-----Q-----	F-T---	158	
Nostoc_sp._PCC_7107/1-302	120	-----R-----	Q I SWQHVRGH	AGNVGNERCDA	IA-----	RSFAAGKI-P-SLK	-----E-----	L-S---	158	
Nostoc_sp._PCC_7120/1-302	120	-----R-----	KVNWHHVRGH	SGNIGNERCDA	IA-----	RCFATGRM-P-SLQ	-----Q-----	L-S---	158	
Nostoc_sp._PCC_7524/1-302	120	-----K-----	QVNWHHVRGH	SGNVGNERCDV	IA-----	RSFATGKM-P-SLQ	-----Q-----	I-S---	158	
Oscillatoria_acuminata_putative_concat/1-241	102	-----FN-----	QFESYAPPEH	VRRADNALADAE	A-----	NRAID-----	-----	-----	131	
Oscillatoria_nigro-viridis/1-301	119	-----R-----	Q I KWEHVRGH	SGDFYNDRCDA	IA-----	RSFSRNLTP-ELQ	-----Q-----	Q-K---	157	
Oscillatoria_sp._PCC_10802/1-284	120	-----R-----	Q I QWQHVR	AHAGNFYNERCDA	IA-----	RAFSLGQT-P-SLK	-----K-----	P-P---	158	
Oscillatoriales_cyanobacterium_JSC-12_concat/1-254	120	-----H-----	DVTFKWVRGH	AGNLENERCDV	LA-----	VAAAKQSD-L-LVD	-----E-----	G-Y---	158	
Planktothrix_agardhii_NIVA-CYA_126/8_concat/1-205	118	-----H-----	QVQFLWVKGH	SGDEENECCDR	LA-----	VKASHQPD-L-LPD	-----L-----	G-----	155	
Planktothrix_rubescens_concat/1-250	118	-----H-----	QVQFLWVKGH	SGDEENECCDR	LA-----	FKASHQPD-L-LPD	-----L-----	G-Y---	156	
Pleurocapsa_minor_concat/1-255	118	-----H-----	EVKFIWVKGH	AGDAENEYCDR	LA-----	VSAALQKD-L-PPD	-----T-----	A-Y---	156	
Pleurocapsa_sp._PCC_7319_concat/1-243	111	-----H-----	DVEFVWVKGH	AGI PENERCDR	LA-----	VSAAQGN-L-ATD	-----Q-----	G-Y---	149	
Prochlorococcus_marinus_str._MIT_9303/1-257	121	-----D-----	DVPLAYVKGH	SGDPDNERVDQ	IA-----	VSFSKGGGAA-ALM	-----S-----	G-Q---	160	
Prochlorococcus_marinus_str._MIT_9313/1-257	121	-----D-----	HVPLAYVKGH	SGDPDNERVDQ	IA-----	VSFSKGGGTP-ALM	-----S-----	G-Q---	160	
Prochlorococcus_marinus_str._MIT_9515/1-239	121	-----Q-----	GLEMEFVKGH	SGDKQNERVDL	IA-----	TSYSKGI P-L-DSE	-----V-----	K---	158	
Prochlorococcus_sp._MIT_0701/1-257	121	-----D-----	DVPLAYVKGH	SGDPDNERVDQ	IA-----					

	230	240	250	260	270	280	290	300	
Acaryochloris_marina_MBIC11017/1-257	159	-	-	-	-	-	-	-	161
Acaryochloris_sp._CCMEE_5410/1-257	159	-	-	-	-	-	-	-	161
Anabaena_cylindrica/1-302	159	-	-	-	-	-	-	-	171
Anabaena_sp._90/1-302	159	-	-	-	-	-	-	-	171
Anabaena_sp._PCC_7108/1-302	159	-	-	-	-	-	-	-	171
Anabaena_variabilis/1-302	159	-	-	-	-	-	-	-	171
Aphanizomenon_flos-aquae/1-305	159	-	-	-	-	-	-	-	171
Arthrospira_sp._PCC_8005_concat/1-259	157	-	-	-	-	-	-	-	166
Calothrix_parietina/1-301	158	-	-	-	-	-	-	-	170
Calothrix_sp._336/3/1-303	157	-	-	-	-	-	-	-	170
Calothrix_sp._PCC_7103/1-301	158	-	-	-	-	-	-	-	169
Calothrix_sp._PCC_7507/1-303	160	-	-	-	-	-	-	-	172
Candidatus_Atelocyanobacterium_thalassa_SIO64986_concat/1-249	153	-	-	-	-	-	-	-	154
Candidatus_Atelocyanobacterium_thalassa_concat/1-249	153	-	-	-	-	-	-	-	154
Candidatus_Synechococcus_spongiarum/1-242	158	-	-	-	-	-	-	-	158
Chamaesiphon_minutus/1-286	163	-	-	-	-	-	-	-	168
Chlorogloeopsis_fritschii/1-304	159	-	-	-	-	-	-	-	173
Chroococcidiopsis_thermalis/1-306	159	-	-	-	-	-	-	-	171
Coleofasciculus_chthonoplastes/1-302	159	-	-	-	-	-	-	-	172
Crinalium_epipsammum/1-305	165	-	-	-	-	-	-	-	170
Crocospaera_watsonii_concat/1-248	153	-	-	-	-	-	-	-	153
Cyanobacterium_stanieri_PCC_7202_concat/1-256	-	-	-	-	-	-	-	-	-
Cyanobium_gracile/1-269	160	-	-	-	-	-	-	-	175
Cyanothece_sp._CCY0110_concat/1-246	-	-	-	-	-	-	-	-	-
Cyanothece_sp._PCC_7425_concat/1-254	156	-	-	-	-	-	-	-	157
Cylindrospermopsis_raciborskii/1-306	159	-	-	-	-	-	-	-	175
Cylindrospermum_stagnale/1-301	159	-	-	-	-	-	-	-	171
Dactylococcopsis_salina/1-282	157	-	-	-	-	-	-	-	162
Dolichospermum_circinale1/1-305	159	-	-	-	-	-	-	-	173
Dolichospermum_circinale2/1-303	159	-	-	-	-	-	-	-	171
Fischerella_muscicola1/1-304	159	-	-	-	-	-	-	-	172
Fischerella_muscicola2/1-304	159	-	-	-	-	-	-	-	172
Fischerella_sp._PCC_9339/1-304	159	-	-	-	-	-	-	-	172
Fischerella_sp._PCC_9431/1-304	159	-	-	-	-	-	-	-	172
Fischerella_sp._PCC_9605/1-304	159	-	-	-	-	-	-	-	172
Geitlerinema_sp._PCC_7407/1-282	157	-	-	-	-	-	-	-	162
Geminocystis_herdmanii_concat/1-253	-	-	-	-	-	-	-	-	-
Gloeobacter_kilaueensis/1-220	-	-	-	-	-	-	-	-	-
Gloeobacter_violaceus/1-213	-	-	-	-	-	-	-	-	-
Gloeocapsa_sp._PCC_7428/1-297	157	-	-	-	-	-	-	-	159
Halothece_sp._PCC_7418/1-283	159	-	-	-	-	-	-	-	163
Leptolyngbya_boryana/1-263	157	-	-	-	-	-	-	-	157
Leptolyngbya_sp._PCC_7375/1-288	166	-	-	-	-	-	-	-	169
Leptolyngbya_sp._PCC_7376/1-293	154	-	-	-	-	-	-	-	168
Lyngbya_aestuarii_putative_concat/1-271	171	-	-	-	-	-	-	-	178
Lyngbya_sp._PCC_8106_concat/1-283	175	-	-	-	-	-	-	-	190
Mastigocladopsis_repens/1-301	159	-	-	-	-	-	-	-	171
Mastigocoleus_testarum/1-302	159	-	-	-	-	-	-	-	172
Microchaete_sp._PCC_7126/1-302	159	-	-	-	-	-	-	-	171
Microcoleus_sp._PCC_7113/1-303	159	-	-	-	-	-	-	-	171
Microcoleus_vaginatus/1-305	158	-	-	-	-	-	-	-	183
Neosynechococcus_sphagnicola_sy1_concat/1-227	-	-	-	-	-	-	-	-	-
Nodosilinea_nodulosa_concat/1-258	163	-	-	-	-	-	-	-	164
Nodularia_spumigena/1-312	159	-	-	-	-	-	-	-	171
Nostoc_punctiforme/1-302	159	-	-	-	-	-	-	-	171
Nostoc_sp._PCC_7107/1-302	159	-	-	-	-	-	-	-	171
Nostoc_sp._PCC_7120/1-302	159	-	-	-	-	-	-	-	171
Nostoc_sp._PCC_7524/1-302	159	-	-	-	-	-	-	-	171
Oscillatoria_acuminata_putative_concat/1-241	132	-	-	-	-	-	-	-	132
Oscillatoria_nigro-viridis/1-301	158	-	-	-	-	-	-	-	179
Oscillatoria_sp._PCC_10802/1-284	159	-	-	-	-	-	-	-	164
Oscillatoriales_cyanobacterium_JSC-12_concat/1-254	-	-	-	-	-	-	-	-	-
Planktothrix_agardhii_NIVA-CYA_126/8_concat/1-205	-	-	-	-	-	-	-	-	-
Planktothrix_rubescens_concat/1-250	157	-	-	-	-	-	-	-	157
Pleurocapsa_minor_concat/1-255	157	-	-	-	-	-	-	-	160
Pleurocapsa_sp._PCC_7319_concat/1-243	-	-	-	-	-	-	-	-	-
Prochlorococcus_marinus_str._MIT_9303/1-257	161	-	-	-	-	-	-	-	164
Prochlorococcus_marinus_str._MIT_9313/1-257	161	-	-	-	-	-	-	-	164
Prochlorococcus_marinus_str._MIT_9515/1-239	-	-	-	-	-	-	-	-	-
Prochlorococcus_sp._MIT_0701/1-257	161	-	-	-	-	-	-	-	164
Prochlorococcus_sp._W11/1-241	-	-	-	-	-	-	-	-	-
Prochlorococcus_sp._W2/1-241	-	-	-	-	-	-	-	-	-
Prochlorococcus_sp._W8/1-241	-	-	-	-	-	-	-	-	-
Prochlorococcus_subsp._pastoris_str._CCMP1986/1-240	-	-	-	-	-	-	-	-	-
Prochlorothrix_hollandica_concat/1-247	-	-	-	-	-	-	-	-	-
Raphidiopsis_brookii/1-309	159	-	-	-	-	-	-	-	175
Richelia_intracellularis_HH01/1-299	159	-	-	-	-	-	-	-	167
Rivularia_sp._PCC_7116/1-306	159	-	-	-	-	-	-	-	172
Rubidibacter_lacunae/1-266	155	-	-	-	-	-	-	-	156
Scytonema_hofmanni/1-303	159	-	-	-	-	-	-	-	171
Scytonema_hofmanni/1-302	159	-	-	-	-	-	-	-	167
Spirulina_subsalsa_concat/1-221	-	-	-	-	-	-	-	-	-
Stanieria_cyanosphaera_PCC_7437_concat/1-249	-	-	-	-	-	-	-	-	-
Synechococcus_elongatus_PCC_6301/1-261	159	-	-	-	-	-	-	-	160
Synechococcus_sp._BL107/1-247	160	-	-	-	-	-	-	-	160
Synechococcus_sp._CB0101/1-248	153	-	-	-	-	-	-	-	154
Synechococcus_sp._CC9311/1-250	-	-	-	-	-	-	-	-	-
Synechococcus_sp._CC9605/1-252	160	-	-	-	-	-	-	-	161
Synechococcus_sp._CC9616/1-243	-	-	-	-	-	-	-	-	-
Synechococcus_sp._KORDI-100/1-247	-	-	-	-	-	-	-	-	-
Synechococcus_sp._KORDI-49/1-240	-	-	-	-	-	-	-	-	-
Synechococcus_sp._NKBG15041c_concat/1-243	-	-	-	-	-	-	-	-	-
Synechococcus_sp._PCC_7002/1-293	159	-	-	-	-	-	-	-	168
Synechococcus_sp._PCC_7335/1-312	164	-	-	-	-	-	-	-	189
Synechococcus_sp._RCC307/1-247	167	-	-	-	-	-	-	-	167
Synechococcus_sp._RS9916/1-275	160	-	-	-	-	-	-	-	166
Synechococcus_sp._WH_5701/1-258	160	-	-	-	-	-	-	-	164
Synechococcus_sp._WH_7803/1-240	-	-	-	-	-	-	-	-	-
Synechococcus_sp._WH_7805/1-241	162	-	-	-	-	-	-	-	162
Synechococcus_sp._WH_8016/1-256	-	-	-	-	-	-	-	-	-
Synechococcus_sp._WH_8102/1-248	160	-	-	-	-	-	-	-	161
Synechocystis_sp._PCC_6714_concat/1-255	161	-	-	-	-	-	-	-	161
Synechocystis_sp._PCC_6803_concat/1-258	157	-	-	-	-	-	-	-	164
Synechocystis_sp._PCC_7509/1-298	159	-	-	-	-	-	-	-	160
Thermosynechococcus_elongatus_BP-1_concat/1-269	156	-	-	-	-	-	-	-	168
Tolypothrix_boutillei_VB521301/1-303	159	-	-	-	-	-	-	-	169
Trichodesmium_erythraeum_IMS101_concat/1-247	-	-	-	-	-	-	-	-	-
Trichormus_azollae/1-302	159	-	-	-	-	-	-	-	171
Xenococcus_sp._PCC_7305_concat/1-260	-	-	-	-	-	-	-	-	-
cyanobacterium_PCC_7702/1-302	158	-	-	-	-	-	-	-	171
cyanobacterium_endosymbiont_of_Epithemia_turgida/1-252	-	-	-	-	-	-	-	-	-
filamentous_cyanobacterium_ESFC-1/1-291	156	-	-	-	-	-	-	-	168

	310	320	330	340	350	360	370		
Acaryochloris_marina_MBIC11017/1-257	162	- - - - -	- - - - -	MP - - -	DS - - -	PE - VNQ - H - - -	P - PDHAEH - -	PASESR I SH L QD L VET 194	
Acaryochloris_sp._CCMEE_5410/1-257	162	- - - - -	- - - - -	MP - - -	DS - - -	PE - VNQ - H - - -	P - PDHAEH - -	PASESR I SH L QD L VET 194	
Anabaena_cylindrica/1-302	172	LNVI KGSDSAINSTK -	INKN ILESSTFPSDT -	TT - - -	- MEPSTA - S - - -	A - ATTTDE - -	KPSEMRVLQLRNLVET 233		
Anabaena_sp._90/1-302	172	ISPAKVSNDPINSTI -	TNKNTLESSPSASDT -	ST - - -	- MELSTT - P - - -	T - ATTTNE - -	QPTEIRVLQLRNLVET 233		
Anabaena_sp._PCC_7108/1-302	172	VNVVKESDSTRNSTI -	IPENIPESITFPLEI -	TT - - -	- MEPSTT - S - - -	A - TTTNDE - -	KLPEMRVLQLRNLVET 233		
Anabaena_variabilis/1-302	172	INVVKVPDHSVTTSTI -	VHKTTQEISTSASNI -	NT - - -	- MEPSTA - H - - -	V - AATIEE - -	NPPEKRVEQLRNLVET 233		
Aphanizomenon_flos-aquae/1-305	172	ISVVKVPNDPINSTI -	TNKNTLESSLSASDT -	ST - - -	- MELSTT - P - - -	T - ATTTNE - -	QPTEIRVLQLRNLVET 233		
Arthrospira_sp._PCC_8005_concat/1-259	167	- - - - -	- - - - -	SMS - - -	XM - - -	TEPI - - N - - - -	ESNES - -	LPREFRVNQLRNL I ET 197	
Calothrix_parietina/1-301	171	ITVAKEPEVTANLTT -	NNPKSTAVNTVEHDM -	TM - - -	- MEPSLA - P - - -	A - THTHEN - -	VPHETRVVLQRNLVET 232		
Calothrix_sp._336/3/1-303	171	VAVTRVSDSLTESTI -	INKKQQQISTLASDI -	TN - - -	- MELPTV - S - - -	A - TSTNEE - -	LPRETRVTQLRNLVET 232		
Calothrix_sp._PCC_7103/1-301	170	VYVANASEIMANSTI -	INEVTSDVD - IAPDN -	TM - - -	- IDASISYA - - -	S - NNTNEE - -	I PREMRVTQLRNLVDT 231		
Calothrix_sp._PCC_7507/1-303	173	INVAKVPDSHANSTI -	INNKEQEISTLASDI -	TI - - -	- MESHIT - P - - -	T - AAAIDE - -	KPELRVTQLRNLVET 234		
Candidatus_Atelocyanobacterium_thalassa_SIO64986_concat/1-24955	155	- - - - -	- - - - -	- IAN - - -	XM - - -	IDAT - - V - - -	N - TIEEFD - -	LSREERV I EIQNLI ET 187	
Candidatus_Atelocyanobacterium_thalassa_concat/1-249	155	- - - - -	- - - - -	- TVN - - -	XM - - -	IDAA - - V - - -	N - TIEEFD - -	LSREERV I EIQNLI ET 187	
Candidatus_Synechococcus_spongiorum/1-242	159	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	PPDAG - -	APMDLAPPPLQRLLTR 179	
Chamaesiphon_minutus/1-286	169	- - - - -	- PAPVTRTPS - EPLPIPS -	LPDLHTA - -	PM - - -	LE - EQI - I - - -	H - PDSTEH - -	I PREVRVAHLHHLIDT 222	
Chlorogloeopsis_fritschii/1-304	174	VAVASLSDSSDELM I -	MNQNTQDISSLAPDT -	NM - - -	- IETSHT - S - - -	A - ALNNEE - -	LPREMRVIQLRNLVET 235		
Chroococcidiopsis_thermalis/1-306	172	APVSEVADFSSNDLI -	IDKSNTN - SHLASNT -	SM - - -	- IDSST - P - - -	S - SISNEE - -	LPREVRVAQLRNLLET 231		
Coleofasciculus_chthonoplastes/1-302	173	TPVTGVSDEKTTFDV -	VYKDVPS - SLPTSET -	PM - - -	- IE - IPD - S - - -	P - NNSNDE - -	LPREIRVAQLGNLI ET 232		
Crinalium_epipsammum/1-305	171	SQ - - -	ETDGSSNKTI - QS -	TCLSSDPPTLDF -	VM - - -	- VDSTTS - L - - -	S - SDTLDN -	LPSEARVAHLHSLI ET 228	
Crocospaera_watsonii_concat/1-248	154	- - - - -	- - - - -	- EQK - - -	XM - - -	IDSS - - V - - -	N - TIEEFD - -	LSREERV I EIQNLI ET 186	
Cyanobacterium_stanieri_PCC_7202_concat/1-256	158	- - - - -	- - - - -	- LN - - -	XM - - -	- NESY - - N - - -	N - AIEEQS - -	LNRDERVKQLRNMIDT 189	
Cyanobium_gracile/1-269	176	- - - - -	- - - - -	- - - - -	- EA - EVP - A - - -	A - ESPALA - -	DPLDPAPPSLQQLLSR 204		
Cyanothece_sp._CCY0110_concat/1-246	153	- - - - -	- - - - -	- EL - - -	XM - - -	IDSS - - V - - -	N - TIEEFD - -	LSREERV I EIQNLI ET 184	
Cyanothece_sp._PCC_7425_concat/1-254	158	- - - - -	- - - - -	- - - - -	- XMSELQELIAT -	V - - -	T - TDSEDH - -	LSRTLVRGQLKNLI ET 192	
Cylindrospermopsis_raciborskii/1-306	176	RDVTKVSDYPTKSAK -	I AQNMLEGRCTLLEK -	RT - - -	- METSTG - I - - -	S - VVTQEE - -	KPPEMRILQIRNLMET 237		
Cylindrospermum_stagnale/1-301	172	INVAKVAGSDLHSSI -	THKQTQESSTFAPDI -	TF - - -	- MEPSTA - P - - -	S - ATTIEA - -	QPSEIRVSQRLRNLVET 233		
Dactylococcopsis_salina/1-282	163	NS - - -	DTVAEVKQTI - QS -	ETPEPIASKTLA -	NPM - - -	- EELE - - E - - -	R - SLEDNR - -	LIQGDRTGLNHLIES 219	
Dolichospermum_circinale1/1-305	174	MSESLVSDYPINSRI I	INKNTLESSSSVSDI -	TI - - -	- MEPSTA - P - - -	T - ATTTNE - -	QPIEMRVSQRLRNLVET 236		
Dolichospermum_circinale2/1-303	172	MSESLVSDYPINSTI I	INQNTLESSSSVSDI -	TI - - -	- MEPSTA - P - - -	T - ATTTNE - -	QPIEIRVSQRLRNLVET 234		
Fischerella_muscicola1/1-304	173	VVASVSDSDPESMI -	INKNTREISTLASEI -	TM - - -	- TDTSNH - S - - -	T - ALSHEE - -	LPREMRVVQLRNLVET 234		
Fischerella_muscicola2/1-304	173	VAVASVSDSDPESMI -	INKNIPEISNFASEI -	NM - - -	- TEISNH - S - - -	T - SLSHEE - -	LPREMRVVQLRNLVET 234		
Fischerella_sp._PCC_9339/1-304	173	ITVASISDSDPESMI -	INKNIREISTLASKI -	NM - - -	- TDTSNH - S - - -	T - PFSHEE - -	LPREMRVVQLRNLVET 234		
Fischerella_sp._PCC_9431/1-304	173	VAVASVSDSDPESMI -	INKNIREISNFASEI -	NM - - -	- TEISNH - S - - -	T - SLSHEE - -	LPREMRVVQLRNLVET 234		
Fischerella_sp._PCC_9605/1-304	173	AVVASVSDCNHESTI -	INKNTQDISALASEI -	TM - - -	- MDTSNH - S - - -	T - TLSHEE - -	LPREMRVVQLHNLVET 234		
Geitlerinema_sp._PCC_7407/1-282	163	SV - - -	ATDRKNDNSV - ES -	ISNLAAPPI I ET -	PM - - -	- VESTAA - T - - -	N - QDSLDS -	VPREVRVTQLRNLVET 220	
Geminocystis_herdmanii_concat/1-253	157	- - - - -	- - - - -	- I - - -	XM - - -	- NESL - - -	T - - N -	SIEEQS -	LNRDERVKQLKNTIDT 187
Gloeobacter_kilaueensis/1-220	155	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	APR - -	SDGITY 163
Gloeobacter_violaceus/1-213	148	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	TVP - -	TNGTTY 156
Gloeocapsa_sp._PCC_7428/1-297	160	VLVTEVSDYVPQSTI -	SYEKTTS SHLVSDM -	TM - - -	- MESTAP - N - - -	S - VVTSEE - -	I PRETRVALLRSLVET 221		
Halothece_sp._PCC_7418/1-283	164	AN - - -	SDTVMSVNKT - IQ -	SETPAQTP I SDG -	TNM - - -	- QELE - - E - - -	R - TIEDSR - -	LVQGDROMEGLNHLIES 220	
Leptolyngbya_boryana/1-263	158	- - - - -	- - - - -	- LQ - VSLTLIPNAAEV -	VMSETP - S - - -	S - IDAEIV - -	VPREVRVEQLRNLVDM 201		
Leptolyngbya_sp._PCC_7375/1-288	170	- - - - -	- - - - -	- ASKQLS -	STRTADN - GPKSEEA -	DC - - -	DR - MTA - D - - -	A - PDRQNQMNT EAGDTLQRLSHILNR 222	
Leptolyngbya_sp._PCC_7376/1-293	169	IL - - -	SPSVTELSAD - SN -	ENSIVIQNESRQ -	NPM - - -	- TDSA - - T - - -	N - STEELV - -	LSRQERVHQLSALIES 225	
Lyngbya_aestuarii_putative_concat/1-271	179	- - - - -	- - - - -	- SLF - - -	XM - - -	- NDIT - - N - - - -	- DAQEV - -	LPREVRVNQLRNL I ET 209	
Lyngbya_sp._PCC_8106_concat/1-283	191	- - - - -	- - - - -	- SDF - - -	AM - - -	- NDIT - - N - - - -	- DAQEV - -	LPREVRVNQLRNL I ET 221	
Mastigocladopsis_repens/1-301	172	VNVSTGADFATNSTI -	INTKKQEISTLGTDI -	TT - - -	- MEPQTH - S - - -	A - SSVSEE - -	VPRETRVTQLQNLVET 233		
Mastigocoleus_testarum/1-302	173	I EAVKVSDSGSKSTI -	INKDTHQTSTFASDT -	TN - - -	- MDSSNA - S - - -	A - VAVSED - -	KSGEIRIAHLHNLVET 234		
Microchaete_sp._PCC_7126/1-302	172	VNVAKVSDFQVDSTI -	INTHEQEFSTLLSDL -	TI - - -	- MESHIT - P - - -	T - TATIDE - -	KPPELRITQLRNLVET 233		
Microcoleus_sp._PCC_7113/1-303	172	SPVAGVSDSNVSKL -	AYETIQM - SLPSSSEV -	AM - - -	- MDSPVS - S - - -	P - AGAIED - -	LPREVRVAQLRNLVET 232		
Microcoleus_vaginatius/1-305	184	SE - - -	GTDTTGNKTI - ES -	YVITSDLPI SNS -	TM - - -	- MD - - - -	S - - T -	ADTL EN -	LPREVRVQDLRNL I ET 237
Neosynechococcus_sphagnicola_sy1_concat/1-227	135	- - - - -	- - - - -	- - - - -	XM - - -	- IDPL - - T - -	TL - SSDLEN -	LPRDVRVTQLRHLVET 165	
Nodosilinea_nodulosa_concat/1-258	165	- - - - -	- - - - -	- - - - -	XM - - -	- ADSGTS - P - - -	I - SENFDS - -	LPREVRVDSLRLNVLET 196	
Nodularia_spumigena/1-312	172	ANVEKVADYESKSTI -	LHKSTQDNSTLISDI -	NA - - -	- MEPSSA - Q - - -	A - AAANEE - -	KLSTTRVKQLKNLLET 233		
Nostoc_punctiforme/1-302	172	LNVAKVADYEKKSR I -	INQSNQEISTSALEI -	TV - - -	- MEPSTG - P - - -	T - AATIDE - -	KPAEMRVSQLRSLVET 233		
Nostoc_sp._PCC_7107/1-302	172	LNVAKVSQDKAHSTI -	ITKEKSATITSASDI -	TV - - -	- MEPSTA - Q - - -	T - AAAIDT -	TTPGIRVEQLRHLVET 233		
Nostoc_sp._PCC_7120/1-302	172	INVVKVPDHSVTTSTI -	VHKTTQEISTSASNI -	NT - - -	- MEPSTA - H - - -	V - AATIEE - -	NPPEKRVEQLRNLVET 233		
Nostoc_sp._PCC_7524/1-302	172	NSVTKVPNYVTTSTI -	VNKKTQEVGTSAIN T -	TT - - -	- MEPSTA - H - - -	A - AATIEE - -	QPPEHRVEQLRNL I ET 233		
Oscillatoria_acuminata_putative_concat/1-241	133	- - - - -	- - - - -	- YQP - - -	XM - - -	- DTTA - - H -	ST - TESLDN -	LPREVRVTQLQYLMQT 166	
Oscillatoria_nigro-viridis/1-301	180	WE - - -	ETDTTGNKTI - ES -	YVITSDLPI SNS -	TM - - -	- MD - - - -	S - - T -	ADTL EN -	LPREVRVQDLRNL I ET 233
Oscillatoria_sp._PCC_10802/1-284	165	LA - - -	LTDGAVNKT I - QS -	NFTLTDFSPSDV -	AM - - -	- TDSNAP - L - - -	T - AESPES - -	LPREVRVTQLRDLVET 222	
Oscillatoriales_cyanobacterium_JSC-12_concat/1-254	159	- - - - -	- - - - -	- KGG - - -	XM - - -	- TDLH - - S - - -	SAPTDFES - -	LPREIRVAQVRCMVEA 192	
Planktothrix_agardhii_NIVA-CYA_126/8_concat/1-205	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
Planktothrix_rubescens_concat/1-250	158	- - - - -	- - - - -	- QEN - - -	XM - - -	- TDL - - - - -	- - - - -	INDSLEN - -	LPREVRVNQLRNL I ET 188
Pleurocapsa_minor_concat/1-255	161	- - - - -	- - - - -	- KSK - - -	XM - - -	- IESS - - - V - - -	N - AIDEPQ - -	LPREERVIQLKNV I ET 193	
Pleurocapsa_sp._PCC_7319_concat/1-243	150	- - - - -	- - - - -	- IS - - -	XM - - -	- TESS - - N - - -	S - TMEEIR - -	LSSEERVTQLRDLIDT 181	
Prochlorococcus_marinus_str._MIT_9303/1-257	165	- - - - -	- - - - -	- - - - -	- IA - LGL - V - - -	E - PIKEAG - -	DGEDDAPSDLQRLLTR 193		
Prochlorococcus_marinus_str._MIT_9313/1-257	165	- - - - -	- - - - -	- - - - -	- IA - LGL - V - - -	E - PIKEAG - -	DGEDFAPSDLQRLLTR 193		
Prochlorococcus_marinus_str._MIT_9515/1-239	159	- - - - -	- - - - -	- - - - -	- - - - -	- KL SSEI - -	DSKETGSSKIQNLF LK 180		
Prochlorococcus_sp._MIT_0701/1-257	165	- - - - -	- - - - -	- - - - -	- IA - LDL - V - - -	E - PIKETG -	DGEDYAPAH LQRLLTR 193		
Prochlorococcus_sp._W11/1-241	160	- - - - -	- - - - -	- - - - -	- - - - -	- ENID - -	QLDLAAPQEIINLYSR 179		
Prochlorococcus_sp._W2/1-241	160	- - - - -	- - - - -	- - - - -	- - - - -	- ENID - -	QLEIAAPQEIINLYSR 179		
Prochlorococcus_sp._W8/1-241	161	- - - - -	- - - - -	- - - - -	- - - - -	- TFNL - -	DKKEIAPEN LKNLFYR 180		
Prochlorococcus_subsp._pastoris_str._CCMP1986/1-240	159	- - - - -	- - - - -	- - - - -	- - - - -	- K - SFEAID -	FSKKKTENKIQKLF SK 181		
Prochlorothrix_hollandica_concat/1-247	157	- - - - -	- - - - -	- QP - - -	XM - - -	- TELD - - - - -	- - - - -	NGSHG -	LGREVRVEQIRNLVDT 185
Raphidiopsis_brookii/1-309	176	RDVTKVSDYPTKSPK -	I AENMLEGRCTLLEE -	RT - - -	- MEPSTG - ITGIS -	VVTQEE - -	KPPEMRILQIRNLMET 240		
Richelia_intracellularis_HH01/1-299	168	YDLAAASSGNSNNKP -	GNENTQNFEFVLQK -	NM - - -	- IDSSNA - T - - -	A - SLNSET - -	ISSELRMVHIRNLLET 229		
Rivularia_sp._PCC_7116/1-306	173	VEVSSTSNSHITPPV -							

		390	400	410	420	430	440	450		
Acaryochloris_marina_MBIC11017/1-257	195	LEIADHIAKEGYLISSEELADLMAVHANAVTSRGDEWVWRNWLISRVVRE	-----GN-QI--	LWQLERLN	-	256				
Acaryochloris_sp._CCMEE_5410/1-257	195	LEIADHIAKEGYLISSEELADLMAVHANAVTSRGDEWVWRNWLISRVVRE	-----GN-QI--	LWQLERLN	-	256				
Anabaena_cylindrica/1-302	234	LRIADEISEKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGDQ	296					
Anabaena_sp._90/1-302	234	LRIADEIAEKGYLITSSSELADLMGVHSSAVTSRGDEWRWRNWI	-----GN-QI--	LWELERGDQ	296					
Anabaena_sp._PCC_7108/1-302	234	LHIADEISEKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGDQ	296					
Anabaena_variabilis/1-302	234	LHIADEIAAKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGDQ	296					
Aphanizomenon_flos-aquae/1-305	234	LRIADEIAEKGYLITSSSELADLMGVHSSAVTSRGDEWRWRNWI	-----GN-QI--	LWELERVD	T	296				
Arthrospira_sp._PCC_8005_concat/1-259	198	LHIADEIASKGYLISSEELADLMDINASAVTSRGDNWAWRNWE	-----GN-QI--	LWQLERV	-	258				
Calothrix_parietina/1-301	233	LRIADEVAEKGFLLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGDQ	295					
Calothrix_sp._336/3/1-303	233	LHIADEIAEKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWV	-----GN-QI--	LWELERGD	L	295				
Calothrix_sp._PCC_7103/1-301	232	LRIADEIADKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWV	-----GN-QI--	LWELERGD	K	294				
Calothrix_sp._PCC_7507/1-303	235	LRIADEIANKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGDQ	297					
Candidatus_Atelocyanobacterium_thalassa_SIO64986_concat/1-248	188	LRIADEIAEKGYLITSSSELADLMDINASAVTSRGDHSWRNWV	-----GN-QI--	LWQLEKI	-	248				
Candidatus_Atelocyanobacterium_thalassa_concat/1-249	188	LRIADEIAEKGYLITSSSELADLMDINASAVTSRGDHSWRNWV	-----GN-QI--	LWQLEKI	-	248				
Candidatus_Synechococcus_spongiorum/1-242	180	LEMADRLAAGGHVLAAMELAQLVDMPHGELRQRRQPWI	WRDWRVLP	PCSDG	-----	GWRLQRKSP	238			
Chamaesiphon_minutus/1-286	223	LRIADEIAKSSYLIASSSELADLMDINASAVTSRGDRWAWRNWI	-----GN-QI--	LWQLERV	D	284				
Chlorogloeopsis_fritschii/1-304	236	LRIADEIAEKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGD	R	298				
Chroococcidiopsis_thermalis/1-306	232	LRSADEIATKGYLITSSSELADLMDVHASAVTSRGEWRWRNWI	-----GN-QI--	LWELERDA	A	294				
Coleofasciculus_chthonoplastes/1-302	233	LRIADEIASEGYLISSEELADLMDINASAVTSRGDHWPWRNWI	-----GN-QI--	LWQLERV	DQ	295				
Crinalium_epipsammum/1-305	229	LRIADEVADQGYLITSSSELADLMDVNASAVTSRGDEWRWRNWI	-----GN-QI--	LWQLERWDQ	291					
Crocospaera_watsonii_concat/1-248	187	LRVAEEIASKGYLITSSSELADLMDINASAVTSRGDHSWRNWV	-----GN-QI--	LWQLEKI	-	247				
Cyanobacterium_stanieri_PCC_7202_concat/1-256	190	LHIAEDISSNGYLITSSSELADLMDINASAVTSRGEAWRNWI	-----GN-QI--	LWQLERVEK	252					
Cyanobium_gracile/1-269	205	LELADRLAAGGYSRAVELARLLEVPTAGLEGKVEPWLRD	DWLRPLGNG	-----	CWRLERGAG	263				
Cyanothece_sp._CCY0110_concat/1-246	185	LRIADEIADKGYLITSSSELADLMDINASAVTSRGDHSWRNWV	-----GN-QI--	LWQLEKI	-	245				
Cyanothece_sp._PCC_7425_concat/1-254	193	LRIAEESI	SKQRYLISSEELADLMDVNASAVTSRGDEWVWRNWV	-----GN-QI--	LWQLERV	-	253			
Cylindrospermopsis_raciborskii/1-306	238	LKIADEIAQKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGDQ	300					
Cylindrospermum_stagnale/1-301	234	LRIADEIADKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGDQ	296					
Dactylococcopsis_salina/1-282	220	LRIADEIAQKGYLISSEELAE	LMNVNASAVTSRGSHWSWRNWV	-----GN-QI--	LWQLERA	-	280			
Dolichospermum_circinale1/1-305	237	LRIADEIAEKGYLITSSSELADLMSVHSSAVTSRGDEWRWRNWI	-----GN-QI--	LWELERGD	K	299				
Dolichospermum_circinale2/1-303	235	LRIADEIAEKGYLITSSSELADLMSVHSSAVTSRGDEWRWRNWI	-----GN-QI--	LWELERGD	K	297				
Fischerella_muscolicola1/1-304	235	LRIADEIAEKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGDH	297					
Fischerella_muscolicola2/1-304	235	LRIADEIAEKGYLVTSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGD	R	297				
Fischerella_sp._PCC_9339/1-304	235	LRIADEIAEKGYLVTSSSELADLMDVHPSAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGD	R	297				
Fischerella_sp._PCC_9431/1-304	235	LRIADEIAEKGYLVTSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGD	R	297				
Fischerella_sp._PCC_9605/1-304	235	LRIADEIAEKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGD	R	297				
Geitlerinema_sp._PCC_7407/1-282	221	LHMADEIAKQGYLISSEELADLMDVNASAVTSRGDNWVWRNWV	-----GN-QI--	LWQLERV	-	281				
Geminocystis_herdmanii_concat/1-253	188	LHIAEEISSNGYLITSSSELADLMDINASAVTSRGEWSWRNWI	-----GN-QI--	LWQLEKVEK	250					
Gloeobacter_kilaueensis/1-220	164	LSLID	-----GQ----	IERFSSWPACQAR	QGVAKARYKKCRSPEER	TVLAEWGMASYPP	215			
Gloeobacter_violaceus/1-213	157	LSLVD	-----GH----	LSRSTGWSACQAL	QGVSGARYKKCRNRAEEL	DTISAWGLPPESL	208			
Gloeocapsa_sp._PCC_7428/1-297	222	LRIADEIAAKGYLITSSSELADLMDVHASAVTSRGDEWRWRNWI	-----GN-QI--	LWELERADQ	284					
Halothece_sp._PCC_7418/1-283	221	LRIADEIAEKGYLISSEELADLMNVNASAVTSRGNHWSWRNWV	-----GN-QI--	LWQLERA	-	281				
Leptolyngbya_boryana/1-263	202	LHMADEIASKGYLISSEELADLMDINANAVTSRGDHWVWRNWV	-----GN-QI--	LWQLERV	-	262				
Leptolyngbya_sp._PCC_7375/1-288	223	LQMADTLAEHGYLIT	TQELADLVNETVSDIAKRGAWTWRNWS	VAQVGGN	-----GSDKD	FWQLERLDS	286			
Leptolyngbya_sp._PCC_7376/1-293	226	LRIADEIAEKGYFISSEELADLMDVNASAVTSRGHEWTWRNWI	-----GN-QI--	LWQLDRIM	G	288				
Lyngbya_aestuarii_putative_concat/1-271	210	LHIADEIASKGYLISSEELADLMDVNASAVTSRGDNWAWRNWE	-----GN-QI--	LWQLERV	-	270				
Lyngbya_sp._PCC_8106_concat/1-283	222	LHIADEIASKGYLISSEELADLMDVNASAVTSRGDNWAWRNWE	-----GN-QI--	LWQLERV	-	282				
Mastigocladopsis_repens/1-301	234	LRIADEIAEKGYLITSSSELADLMDVHASAVTSRGDEWRWRNWI	-----GN-QI--	LWELERGD	R	296				
Mastigocoleus_testarum/1-302	235	LRIADEIAEKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWV	-----GN-QI--	LWELERAD	L	297				
Microchaete_sp._PCC_7126/1-302	234	LRIADEIATKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGDQ	296					
Microcoleus_sp._PCC_7113/1-303	233	LHIADEIARQGYLISSEELADLMDVNASAVTSRGDHWPWRNWI	-----GN-QI--	LWQLERV	DQ	295				
Microcoleus_vaginatus/1-305	238	LHIADEVATKGYLITSSSELADLMDVNASAVTSRGDNWVWRNWV	-----GN-QI--	LWQLERID	H	300				
Neosynechococcus_sphagnicola_sy1_concat/1-227	166	LRMADEIAVKGYLITSSSELADLMDVNASAVTSRGDNWVWRNWV	-----GN-QI--	LWQLERV	-	226				
Nodosilinea_nodulosa_concat/1-258	197	LQIADEIAKQGYLITSSSELADLMDVNASAVTSRGEFWARNWS	-----GN-QI--	LWQIERI	-	257				
Nodularia_spumigena/1-312	234	LRIAEESI	SEKGYLITSSSELADLMDVHASAVTSRGDEWRWRNWI	-----GN-QI--	LWELERGD	C	296			
Nostoc_punctiforme/1-302	234	LRIADEISEKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGDQ	296					
Nostoc_sp._PCC_7107/1-302	234	LRIADEIADKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGD	R	296				
Nostoc_sp._PCC_7120/1-302	234	LHIADEIAAKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGD	R	296				
Nostoc_sp._PCC_7524/1-302	234	LHIADEIATKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGD	R	296				
Oscillatoria_acuminata_putative_concat/1-241	167	LEIADRIATKGYLITSSSELADLMDVNASAVTSRGDNWPWRNWI	-----GN-QI--	LWQLERID	L	229				
Oscillatoria_nigro-viridis/1-301	234	LHIADEVATKGYLITSSSELADLMDVNASAVTSRGDNWVWRNWV	-----GN-QI--	LWQLERID	H	296				
Oscillatoria_sp._PCC_10802/1-284	223	LRIADEIAAKRYLITSSSELADLMDVNASAVTSRGDNWVWRNWI	-----GN-QI--	LWQLERV	-	283				
Oscillatoriales_cyanobacterium_JSC-12_concat/1-254	193	LRIADEIAEKGYLITSAELADLMDINASAVTSRGDHWVWRNWV	-----GN-QI--	LWQLERV	-	253				
Planktothrix_agardhii_NIVA-CYA_126/8_concat/1-205	156	-----YSQENXLADLMDINASAVTSRGDNWAWRNWE	-----GN-QI--	LWQLERV	-	204				
Planktothrix_rubescens_concat/1-250	189	LHIADEIATKGYLISSEELADLMDINASAVTSRGDNWAWRNWE	-----GN-QI--	LWQLERV	-	249				
Pleurocapsa_minor_concat/1-255	194	LRIADEIATQGYLITSSSELADLMDINASAVTSRGDNWSWRNWV	-----GN-QI--	LWQLERV	-	254				
Pleurocapsa_sp._PCC_7319_concat/1-243	182	LRIAEEI	ASKGYLISSEELADLMDINASAVTSRGDHSWRNWV	-----GN-QI--	LWQLERV	-	242			
Prochlorococcus_marinus_str_MIT_9303/1-257	194	LELADRF	AVRGYGLSLVELAQLVEQPIEQLASKTAPWR	WRDWLLEPMGEN	-----	RWRLRRC	AA	252		
Prochlorococcus_marinus_str_MIT_9313/1-257	194	LELADRF	AVRGYGLSLIEVAQLVEQPIEQLANKTAPWR	WRDWLLEPVGEN	-----	RWRLHR	CAA	252		
Prochlorococcus_marinus_str_MIT_9515/1-239	181	NELLEKFAEKKYFLN	SSSELNELLGEESYPLIKKYELFEWRNWRLIPK	-----N-QK--	YWI	IENK	-	237		
Prochlorococcus_sp._MIT_0701/1-257	194	LELADRLA	ERGYSLSLVELAQLVEQPIEQLAGKTA	PHWRDWLLEPVGEN	-----	RWRLR	CAA	252		
Prochlorococcus_sp._W11/1-241	180	IELVSKFAQKGFLLT	SRELNLLSIEENNYIREMKRFNWRNWI	IKPTSKK	-----	LWI	IQNN	237		
Prochlorococcus_sp._W2/1-241	180	IELVSKFAQKGFLLT	TRELNLLSIEENDYIKEMKQFNWRNWI	IKPTSKK	-----	LWI	IENN	237		
Prochlorococcus_sp._W8/1-241	181	FDLIDN	FAKKKYLLNSNELFDLLDLKKN	SNIENYIEFKWRNWI	FKPINNQ	-----	HWQI	IKN	237	
Prochlorococcus_subsp._pastoris_str_CCMP1986/1-240	182	NELLEQFAQKKYFLN	SSSELNELLGEESYILVEKYELFEWRNWRLIPK	-----N-QK--	YWI	IENI	-	238		
Prochlorothrix_hollandica_concat/1-247	186	LHMADEIASKGYLISSEELADLMDVNASAVTSRGDHWVWRNWV	-----GN-QI--	LWQLERV	-	246				
Raphidiopsis_brookii/1-309	241	LKIADEIAQKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGDQ	303					
Richelia_intracellularis_HH01/1-299	230	LHVADEIAAKGYLITSSSELADLMGVHASAVTSRGDEWRWRNWV	-----GN-QI--	LWELERGD	L	292				
Rivularia_sp._PCC_7116/1-306	237	LRIADEIAEKGYLITSSSELADLMNVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERAD	K	299				
Rubidibacter_lacunae/1-266	204	LRIADEIAREGYLISSEELADLMDVNASAVTSRGDHWAWRNWV	-----GN-QI--	LWQLERT	D	265				
Scytonema_hofmanni/1-303	230	LHIADEIATKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGD	R	292				
Scytonema_hofmanni/1-302	230	LRIADEIAEKGYLITSSSELADLMDVHASAVTSRGDEWRWRNWI	-----GN-QI--	LWELERGD	K	292				
Spirulina_subsalsa_concat/1-221	160	LRIAEEI	ATKGYLISSEELADLMDVNASAVTSRGDHSWRNWV	ISRVVRE	-----GN-QI--	LWQLERV	-	220		
Stanieria_cyanosphaera_PCC_7437_concat/1-249	188	LRIADEIASKGYLISSEELADLMDINASAVTSRGDHSWRNWV	-----GN-QI--	LWQLERV	-	248				
Synechococcus_elongatus_PCC_6301/1-261	200	LQIAAAIAEQGFLLITSSSELADLMDVNASAVTSRGEEWI	WRNWVRKDRRE	-----GN-QI--	LWQLERI	-	260			
Synechococcus_sp._BL107/1-247	184	LELADRLA	EAGFALTVELAQLVEQPIKQLDAKDG	PWIRDWLVS	PA-----D-TG-	RWRLQR	RAG	242		
Synechococcus_sp._CB0101/1-248	179	FELADRLA	EGGYGLSLLLEAQLVEQPLKKLEN	RSEPWRWDWI	VQPSDDG	-----	RWRLAR	DAS	237	
Synechococcus_sp._CC9311/1-250	187	MELADRLA	SGGYALTAVELAQLVEQPLNRLSER	QGSWRWDWMVEPVEND	-----	RWRLRR	DP	A	245	
Synechococcus_sp._CC9605/1-252	189	LELADRLA	QGGYSLSLLLEAQLVEKPLKQLET	KAESWIRDWTIEPQ	-----A-EG-	RWTLQR	REA	247		
Synechococcus_sp._CC9616/1-243	180	LELAERLA	AGAFTLTAVELAQLVEQPLQQLQER	SSSWRWDWSVDPVQAG	-----	RWRLRR	REG	238		
Synechococcus_sp._KORDI-100/1-247	184	LELAERLA	AGAFTLTAAELAQLVEQPLQNLQER	SSSWRWDWSVDPVQAG	-----	RWRLRR	REG	242		
Synechococcus_sp._KORDI-49/1-240	177	LELADRLA	AAGFTLQVTELAQLVEQPMRQLEER	KEPWIWRDWLVEPLQGG	-----	RWRLQR	REA	235		
Synechococcus_sp._NKBG15041c_concat/1-243	175	LRIADEIA	ENGFYISSEELADLMDVNASAVTSRGHEWTWRNWI	ISRVKRE	-----GN-QI--	LWQLDR	I	G	237	
Synechococcus_sp._PCC_7002/1-293	226	LRIADEIA	ENGFYISSEELADLMDVNASAVTSRGHEWTWRNWI	ISRVKRE	-----GN-QI--	LWQLDR	I	G	288	
Synechococcus_sp._PCC_7335/1-312	250	LQVTD	ELAKQGYLITSAELADLMDVNASAVTSRGDYAWRNWS	ISRIIRE	-----GN-QI--	LWQLER	ID	311		
Synechococcus_sp._RCC307/1-247	185	LEVADRLA	SGGYGLTSVELAQLVEQPLRQLES	RDGPWQWRD	FQVQPLGDG	-----	RWQLQR	SKP	243	
Synechococcus_sp._RS9916/1-275	205	LELADRLA	SGGYALTAEI	LAQLVEQPLSKLGQHGGAWR	WRDWEVQPIEEG	-----S-	RWRLHR	DTA	264	
Synechococcus_sp._WH_5701/1-258	194	LEVADRLA	SGGYALTAAELARLVEQPLARLQER	QAPWIRDWRV	SPTESG	-----	RWRLAR	AE	252	
Synechococcus_sp._WH_7803/1-240	179	LELADRLA	SGGFTLTAVELAQLVEQPLTN	YLERQQPWRWRDWMVEPIEAD	-----	RWRLRR	AE	236		
Synechococcus_sp._WH_7805/1-241	180	LELADRLA	SGGFTLTAVELAQLVEQPLTNLLER	QQPWRWRDWMVEPIKGD	-----	RWRLRR	AE	237		
Synechococcus_sp._WH_8016/1-256	193	MELADRLA	SGGYALTAVELAQLVEQPLNRLSER	QGSWRWDWMVEPVEND	-----	RWRLRR	DP	A	251	
Synechococcus_sp._WH_8102/1-248	185	LELADRLA	DGGFSLSLVELAQLVELPLKQLE	QRRESWSWRDWRVHPD	-----K-EG-	RWTLQR	REA	243		
Synechocystis_sp._PCC_6714_concat/1-255	194	LRAAEI	IAKQGYLITSSSELADLMDINPSAVTSRGDHSWRNWV	ISRVVRE	-----GN-QI--	LWQLERV	-	254		
Synechocystis_sp._PCC_6803_concat/1-258	197	LRAAEI	IAKQGYLITSSSELADLMDINPSAVTSRGDHSWRNWV	ISRVVRE	-----GN-QI--	LWQLERV	-	257		
Synechocystis_sp._PCC_7509/1-298	222	LRIAEI	ATKGYLITSSSELADLMDVHASAVTSRGDEWRWRNWI	ISRVVRE	-----GN-QI--	LWELERGD	D	284		
Thermosynechococcus_elongatus_BP-1_concat/1-269	196	LRIADELA	AGKYLITSGELAE	LMDVHASAVTSRGEWVWRNWQ	ISRVVRE	-----SN-QI--	LWQLER	ID	V	258
Tolypothrix_bouteillei_VB521301/1-303	233	LRIADEIADKGYLITSSSELADLMDVHASAVTSRGDEWRWRNWI	-----GN-QI--	LWELERGD	K	295				
Trichodesmium_erythraeum_IMS101_concat/1-247	186	LHIADEVAN	KGYLITSSSELADLMDINASAVTSRGDHWVWRNWV	ISRVVRE	-----GN-QI--	LWQLERV	-	246		
Trichormus_azollae/1-302	234	LRVADEIAEKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERGDQ	296					
Xenococcus_sp._PCC_7305_concat/1-260	196	LRIADEIAQKGYLISSEELADLMDINASAVTSRGDHSWRNWV	-----GN-QI--	LWQLEN	I	EA	258			
cyanobacterium_PCC_7702/1-302	234	LRIADEIAEKGYLITSSSELADLMDVHASAVTSRGDQWRWRNWI	-----GN-QI--	LWELERSD	R	296				
cyanobacterium_endosymbiont_of_Epithemia_turgida/1-252	189	LRIADEISNRGYLITSSSELADLMDVNASAVTSRGDHSWRNWV	-----GN-QI--	LWQLERT	D	250				
filamentous_cyanobacterium_ESFC-1/1-291	230	LRIADEIAGKGYLISSEELADLMDVNASAVTSRGDHWPWRNWV	-----GN-QI--	LWQLERV	-	290				

	460	470	480		
Acaryochloris_marina_MBIC11017/1-257	-	-	-	-	S
Acaryochloris_sp._CCMEE_5410/1-257	-	-	-	-	S
Anabaena_cylindrica/1-302	297	-	-	-	EAD
Anabaena_sp._90/1-302	297	-	-	-	EID
Anabaena_sp._PCC_7108/1-302	297	-	-	-	DTD
Anabaena_variabilis/1-302	297	-	-	-	EDE
Aphanizomenon_flos-aquae/1-305	297	-	-	-	QIL
Arthrospira_sp._PCC_8005_concat/1-259	-	-	-	-	D
Calothrix_parietina/1-301	296	-	-	-	EEE
Calothrix_sp._336/3/1-303	296	-	-	-	EES
Calothrix_sp._PCC_7103/1-301	295	-	-	-	SNE
Calothrix_sp._PCC_7507/1-303	298	-	-	-	ESE
Candidatus_Atelocyanobacterium_thalassa_SIO64986_concat/1-249	-	-	-	-	D
Candidatus_Atelocyanobacterium_thalassa_concat/1-249	-	-	-	-	D
Candidatus_Synechococcus_spongiarum/1-242	239	-	-	-	PHH
Chamaesiphon_minutus/1-286	285	-	-	-	GE
Chlorogloeopsis_fritschii/1-304	299	-	-	-	GDE
Chroococcidiopsis_thermalis/1-306	295	-	-	-	ATE
Coleofasciculus_chthonoplastes/1-302	296	-	-	-	DEE
Crinalium_epipsammum/1-305	292	-	-	-	SEHNSDLATE
Crocospaera_watsonii_concat/1-248	-	-	-	-	D
Cyanobacterium_stanieri_PCC_7202_concat/1-256	253	-	-	-	MDD
Cyanobium_gracile/1-269	264	-	-	-	RPQ
Cyanothece_sp._CCY0110_concat/1-246	-	-	-	-	D
Cyanothece_sp._PCC_7425_concat/1-254	-	-	-	-	D
Cylindrospermopsis_raciborskii/1-306	301	-	-	-	DDN
Cylindrospermum_stagnale/1-301	297	-	-	-	SEE
Dactylococcopsis_salina/1-282	281	-	-	-	EG
Dolichospermum_circinale1/1-305	300	-	-	-	EVE
Dolichospermum_circinale2/1-303	298	-	-	-	EIE
Fischerella_muscicola1/1-304	298	-	-	-	NDS
Fischerella_muscicola2/1-304	298	-	-	-	NDD
Fischerella_sp._PCC_9339/1-304	298	-	-	-	NDD
Fischerella_sp._PCC_9431/1-304	298	-	-	-	NDD
Fischerella_sp._PCC_9605/1-304	298	-	-	-	NDE
Geitlerinema_sp._PCC_7407/1-282	-	-	-	-	D
Geminocystis_herdmanii_concat/1-253	251	-	-	-	AE
Gloeobacter_kilaueensis/1-220	216	-	-	-	LKS
Gloeobacter_violaceus/1-213	209	-	-	-	LEG
Gloeocapsa_sp._PCC_7428/1-297	285	-	-	-	ATE
Halothece_sp._PCC_7418/1-283	282	-	-	-	EG
Leptolyngbya_boryana/1-263	-	-	-	-	D
Leptolyngbya_sp._PCC_7375/1-288	287	-	-	-	LY
Leptolyngbya_sp._PCC_7376/1-293	289	-	-	-	EAS
Lyngbya_aestuarii_putative_concat/1-271	-	-	-	-	D
Lyngbya_sp._PCC_8106_concat/1-283	-	-	-	-	D
Mastigocladopsis_repens/1-301	297	-	-	-	NEE
Mastigocoleus_testarum/1-302	298	-	-	-	DED
Microchaete_sp._PCC_7126/1-302	297	-	-	-	EAE
Microcoleus_sp._PCC_7113/1-303	296	-	-	-	DEG
Microcoleus_vaginatus/1-305	301	-	-	-	TTD
Neosynechococcus_sphagnicola_sy1_concat/1-227	-	-	-	-	D
Nodosilinea_nodulosa_concat/1-258	-	-	-	-	D
Nodularia_spumigena/1-312	297	G	S	R	SAGGDRLETE
Nostoc_punctiforme/1-302	297	-	-	-	EEE
Nostoc_sp._PCC_7107/1-302	297	-	-	-	EEE
Nostoc_sp._PCC_7120/1-302	297	-	-	-	EDE
Nostoc_sp._PCC_7524/1-302	297	-	-	-	EED
Oscillatoria_acuminata_putative_concat/1-241	230	-	-	-	AGE
Oscillatoria_nigro-viridis/1-301	297	-	-	-	TTD
Oscillatoria_sp._PCC_10802/1-284	-	-	-	-	D
Oscillatoriales_cyanobacterium_JSC-12_concat/1-254	-	-	-	-	D
Planktothrix_agardhii_NIVA-CYA_126/8_concat/1-205	-	-	-	-	N
Planktothrix_rubescens_concat/1-250	-	-	-	-	N
Pleurocapsa_minor_concat/1-255	-	-	-	-	D
Pleurocapsa_sp._PCC_7319_concat/1-243	-	-	-	-	D
Prochlorococcus_marinus_str._MIT_9303/1-257	253	-	-	-	DQI
Prochlorococcus_marinus_str._MIT_9313/1-257	253	-	-	-	DQI
Prochlorococcus_marinus_str._MIT_9515/1-239	238	-	-	-	HA
Prochlorococcus_sp._MIT_0701/1-257	253	-	-	-	DQI
Prochlorococcus_sp._W11/1-241	238	-	-	-	NPQ
Prochlorococcus_sp._W2/1-241	238	-	-	-	NPQ
Prochlorococcus_sp._W8/1-241	238	-	-	-	DIE
Prochlorococcus_subsp._pastoris_str._CCMP1986/1-240	239	-	-	-	NT
Prochlorothrix_hollandica_concat/1-247	-	-	-	-	D
Raphidiopsis_brookii/1-309	304	-	-	-	DDN
Richelia_intracellularis_HH01/1-299	293	-	-	-	DEN
Rivularia_sp._PCC_7116/1-306	300	-	-	-	EGE
Rubidibacter_lacunae/1-266	-	-	-	-	C
Scytonema_hofmannii/1-303	293	-	-	-	DAAIGG
Scytonema_hofmannii/1-302	293	-	-	-	EGSRE
Spirulina_subsalsa_concat/1-221	-	-	-	-	D
Stanieria_cyanosphaera_PCC_7437_concat/1-249	-	-	-	-	D
Synechococcus_elongatus_PCC_6301/1-261	-	-	-	-	H
Synechococcus_sp._BL107/1-247	243	-	-	-	GES
Synechococcus_sp._CB0101/1-248	238	-	-	-	RDG
Synechococcus_sp._CC9311/1-250	246	-	-	-	TET
Synechococcus_sp._CC9605/1-252	248	-	-	-	EQS
Synechococcus_sp._CC9616/1-243	239	-	-	-	EQS
Synechococcus_sp._KORDI-100/1-247	243	-	-	-	EQS
Synechococcus_sp._KORDI-49/1-240	236	-	-	-	GEP
Synechococcus_sp._NKBG15041c_concat/1-243	238	-	-	-	TTA
Synechococcus_sp._PCC_7002/1-293	289	-	-	-	TTT
Synechococcus_sp._PCC_7335/1-312	-	-	-	-	E
Synechococcus_sp._RCC307/1-247	244	-	-	-	ATP
Synechococcus_sp._RS9916/1-275	265	-	-	-	GHG
Synechococcus_sp._WH_5701/1-258	253	-	-	-	EFP
Synechococcus_sp._WH_7803/1-240	237	-	-	-	GSR
Synechococcus_sp._WH_7805/1-241	238	-	-	-	GSR
Synechococcus_sp._WH_8016/1-256	252	-	-	-	KQT
Synechococcus_sp._WH_8102/1-248	244	-	-	-	EQS
Synechocystis_sp._PCC_6714_concat/1-255	-	-	-	-	D
Synechocystis_sp._PCC_6803_concat/1-258	-	-	-	-	D
Synechocystis_sp._PCC_7509/1-298	285	-	-	-	ATE
Thermosynechococcus_elongatus_BP-1_concat/1-269	259	-	-	-	PPT
Tolypothrix_bouteillei_VB521301/1-303	296	-	-	-	EGD
Trichodesmium_erythraeum_IMS101_concat/1-247	-	-	-	-	D
Trichormus_azollae/1-302	297	-	-	-	EAE
Xenococcus_sp._PCC_7305_concat/1-260	259	-	-	-	VE
cyanobacterium_PCC_7702/1-302	297	-	-	-	EDE
cyanobacterium_endosymbiont_of_Epithemia_turgida/1-252	251	-	-	-	LE
filamentous_cyanobacterium_ESFC-1/1-291	-	-	-	-	D

Supplementary Alignment 2

Aligned homologs of *Trichodesmium erythraeum* RNase HI.
These alignments were done with ProbCons (see Methods) and used to build the Bayesian tree shown in Figure 2A.

```
>Synechococcus_sp_NKBG15041c
AT-GGC-ACA-AC-A-ACAATTGATTG--ACTGTAT-TTACACTGATGGTGGCTGCAT-C
CATAACCCTGGCCCCGGCGGGTGGGGTGTGGTTGTCCA--ATTTACC-GATGGTT-----
-----CT--GTGCTTGAAATGGGCGATCACCAACCAGAGACCACCAACAATCGCATGGA
AATGC-AAGCTGCGATCCAAGCCCTTGAGTATGTTCTGTGA----ATGCGATCAACAG---
-----GAGC-CCATGCGTCTCTTAACGGACAGTAAATATGTGATCGATGGGATCACGAA
---G---TGGATCAAAGGCTGGCAAAAAAAGGCTGGAAAACCGCCGCCGGCAAACCAGT
CCTCAACCAAGACCTCTGGCAACGACTCCATCAGCT-CA-----A-T-CATCC-----
-----CCAAATCTCCTGGTGCCATGTCAAGGGCCACAGCGGTGAAC-CAG
GCAACGA--GCGGGCCGATGCGATCG---CCAA-TGGTTT---TGC-CCGGGG-----
-----CC---AA---AAG---CC---C-C--ACCTC-----C-
----GGAAC-----A-----
-----ACCAA--AGC-----GCCTCCA
TGCCGAAAAAA-TC-TACC--AAAAT---C-A-A--GAGGTAA
>Synechococcus_sp_PCC_7002
AT-GGC-ACA-AC-A-ACGCACAATTG--ACCGCAT-TTATACCGACGGTGGCTGTAT-T
CGCAATCCCGGCCCGCGGTGGTTGGGGAGTGGTGATCCA--GTTCAACC-GATGGCA-----
-----CC--GTTTTGGAAATGGGCGATCGCCAACCCGAAACCACCAATAACCGCATGGA
AATGC-AAGCCGCAATCCAGGCCCTCGAATTTGTCCGCGA----CTGTCCCCAGCCA---
-----GAAC-CAGTGCGCCTTTTTTACCGATAGTAAATATGTCATCGACGGCATCACTAA
---A---TGGATTAAAGGTTGGCAAAAAAAGGGCTGGAAAACCGCCACTGGCAACCCCGT
TTTAAATCAAGACCTCTGGCAGCAGCTCAATCAACT-GA-----A-T-CAACC-----
-----GACCATCTCTTGGTGTACGTCAAAGGCCATAGCGGTGAGG-CGG
GGAATGA--ACGGGCCGATGCGATCG---CCAA-TGGCTT---TGC-CCGGGG-----
-----TC---AA---GCT---CC---T-A--ACCTA-----C-
-----GC--AG--C-----
-----CCCAA--TGC-----A----
-ACC--AAAGAT-C-TACC--AAAAT---C-T-A--GAGGTAC
>Acaryochloris_marina_MBIC11017
AT-GCC-GGC-TT-T-TTCTGATATTG--ACAGTAT-CTATACTGATGGCGCTTGTTT-A
GGGAATCCTGGTCCAGGCGGATGGGGCAGTGTCATCTA--CTTCAAA-GATGGAT-----
-----CC--TGCCATGAAATTGGCGGGTTTGTGCCTCAAACGACGAATAACCGGATGGA
ATTGC-AGGCTGCGATCGCAATCTTGGAATTGTTGCAAGA----CCAACCGTTACCT---
-----AAGA-CCATTCCCCTCTATACTGACAGCGAATACGTCAAAAATGGCATTACTAA
---G---TGGTTGGCAGGATGGAAACGTAAAGGCTGGAAAACAGCTAAAGGTAAAGCGGT
TCTGAATCAAGATTTGTGGCAACAGCTCGATCAACT-TA-----A-T-GATGC-----
-----GACCATAGACTGGCGATATATCCGCGCTCATGGAGGCCATG-AAG
GCAATGA--GCGATGTGATCAGATCG---CTCG-TGCTTT---TTC-CCAGTA-----
-----TC---AA---ACC---CC---T-G--ACCTT-----C-
-----
-----ACCAAATTAC-----C----
-CCC--AG-TACCA-TGCC-----AGACTC
>Acaryochloris_sp_CCME_5410
AT-GCC-GGC-TT-T-TTCTGATATTG--ACAGTAT-CTATACTGATGGCGCTTGCTC-G
GGGAATCCCGGTCCAGGCGGATGGGGCAGTGTCATCTA--CTTCAAG-GATGGAT-----
-----CC--TGCCATGAAATTGGCGGGTTTGTGCCTCAAACGACGAATAACCGGATGGA
```


ATTGC-AGGCTGCGATCGCAATCTTGGAATTGTTGCAAGA----CCAACCGTTACCC---
-----AAGA-CCATTCCCCTCTATACTGACAGCGAGTACGTCAAAAATGGCATTACTAA
---G---TGGTTGGCAGGATGGAAACGCAAAGGCTGGAAAACAGCTAAAGGTAAAGCGGT
TCTGAATCAAGATTTGTGGCAACAGCTCGATCAACT-TA-----A-T-GATGC-----
-----GACCATAGACTGGCGATATATCCGCGCTCATGCAGGCCATG-AAG
GCAATGA--GCGATGTGATCAGATCG---CCCG-TGCTTT---TTC-CCAGCA-----
-----TC---AA---ACC----CC---T-G--ACCTT-----C-

-----ACCAAATTAC-----C-----
-CCC--AG-TGCCA-TGCC-----AGACTC
>Thermosynechococcus_elongatus_BP-1
AT-GAC-CCA-TA-----TTC--GTGCCAT-CTATACCGATGGTGCCTGTGA-A
GGCAATCCCGGACCGGGGGGGTGGGGCGTGGTTATTTA--CTTTACC-GATGGCT-----
-----CT--GTGCACGAGCTCGGGGGACACCACCCGGCCACGACGAATAATCGCATGGA
ATTAC-AGGCTGCGATTGAAGCTCTCAAGGCATGGCGCCA----AC---TGGCCCCT---
----GGGAGCGCGATCGCCCTTTACACCGACAGCGAGTATGTCCTCAGGGGAATTACTGA
---G---TGGATTCACCACTGGAAACGGCGGGGCTGGAAAACAGCCGCCAAAAAGCCGGT
TCTCAACCAAGACCTCTGGCAGGAGTTAGATGCCCT-CA-----A-T-GATCC-----
-----CTTAGTCCAATGGCACCATGTCCGCGGCCATAGGGGAGATG-TCG
GCAATGA--ACGCTGCGATCTTATTG---CCCG-CCG-TT---TGAGTCGTGG-----
-----GC---A----GCC-----AA-
---TT-GCC--CT--A-----
-----AGAACGCTGC-----C-----
-ACC-----ACC--GC-----T--GGGGTAA
>Synechococcus_elongatus_PCC_6301
AT-GTC-TGA-TG-C-----CATCC--TCCGCCT-TTACACCGATGGCGCTTGTAC-G
GGCAATCCAGGGCCGGGCGGTGGGGGGTGCAAATTCA--ATTTGCC-GATGGCA-----
-----CG--GTGCAAGAGCTGGGTGGGCGTGACGCTGCCACAATAAACC GGATGGA
GTTGC-AGGCTGCGATCGCGGCCTTAGAGTTCTGGCGCGA----TCACTCTGACC-A---
----G-CAAC-CCGTCACCTCTACACCGACAGCGAATACGTGATCAAAGGGATCACCCA
---A---TGGCTGAAGGGGTGGCAGCGCAAGGGTTGGAAAACGGCGGGCGGGTAAGCCGGT
GCTCAACCAGGACTTATGGATGCACTTGGATGACCT-CA-----A-T-TCACC-----
-----CGCAGCGCAATGGCAATATGTGCGCGGCCATAGCGGCGATC-CGG
GCAATGA--GCGTTGCGATCAAATTG---CCCG-TAGTTT---TGC-TCAGCA-----
-----AC---G----TCCGTTACC-----ACTCTGCCAACGGTCTGATG-
---CA-CCA--CT--A-----
-----CAATCGGGGC-----
-----AACCGCA-GCA--GGAGC---CCG-T--TCGGATG
>Trichodesmium_erythraeum_IMS101
AT-GAC-AGA-AA-A-AAGAACCGAAA--TTACTAT-CTATACTGATGGTGCTTGTAG-T
GGAAATCCCGGCCCTGGTGGTTATGGTATAATTATTTT--ATCTGAG-AAA--AA-----
-----AC--G-CCAAGAATTATCTGGAGGTTATAAATTAACCACTAATAATCGTATGGA
ATTAA-TGGCAGTAATAGTAGGATTAGAACAAC TAGAAAT----T-----CCAT---
-----CAA-TTGTCAATTTATATACTGACTCTAAATATATAGTTGATGCTGTTACAAA
---AGGTTGGGCAAAACGTTGGCGAGCTAACAGTTGGAAACGGAATAAAAAAGATAAAGC
AATGAACCCCGATCTATGGGGAAAATTATTAGATTT-ATGC-A-G-T-AAACA-----
-----TCAAGTTGAATTTTCTTGGGTACGGGGTCATTCTGGAAATA-TAG
AAAATGA--ACGCTGTGATAAATTAG--CGGT-CAAAGC---TTC-TCA-----
-----A---AAA---CT---T-G--ATTTA-----C-
-----C--CT--C-----
-----AG----ACC-----T-----
-----T-----GG---T-T-A--TCAATAA
>Planktothrix_NIVA-CYA_98
AT-GAA-CGA-CA-A-TTTAAAAAAG--TTACAAT-TTATACCGATGGAGCTTGTAG-T
GGGAATCCAGGGCCAGGAGGATATGGCGTGGTGTGAT--GTTTGGC-CCT--CA-----

-----CC--G-CAAGGAATTATCAGGGGGATATCAACTTACTACTAATAATCGCATGGA
ATTAA-TCGCGGCTATTATAGGATTAGAAGCCTTAAAGTT----T-----CCTT---
-----GTG-CTGTTAATTTATATACGGATTCTCGCTATTTAGTTGATGCCATGAATAA
---AGGATGGGCGAAGAAATGGCAGAAAAATAATTGGCAACGTAGTCAAAAAGAATCGGC
TAAAAATCCTGATTTGTGGGAAAGATTACTCACTAT-ATCC-G-A-A-CAACA-----
-----TCAAGTTCAATTTCTTTGGGTAAAGGTCATTCGGGAGATG-AAG
AAAATGA--ATGTTGCGATCGCTTGG---CTGT-CAAGGC---TTC-ACA-----
-----T---CAA---CC---G-G--ATTTA-----C-
-----T--CC--C-----
-----AG---ATTTGG-----GT-----TATTC-----
-----C-----CA---A-G-A--GAATTAA

>Cyanobacterium_sp_UCYN-A2

TT-GAA-AAA-----AT--TAAAAAT-TTACACTGATGGTGCCTGTTC-T
GGTAACCCTGGCAAAGGGGGTTATGGAACAATATTAGT--TTTTAAT-AAG--TA-----
-----TC--G-GAAAGAATTGTCAGGAGGTTATAGACTAACAATAATCGTATGGA
GATCA-TATCTATTATCATTGGCTTAGAATCTTTAATAGA----A-----TCTT---
-----GCA-ATGTAAGTATTTATTCTGATTCTCGGTATGTAGTAGACAGTATGAATAG
---AGGATGGGCAAGTAAGTGGCAGTCTAACGACTGGAAACGTAACAAACAAGAAATAGC
TAAAAATTCTGATCTATGGAAAAGATTATTAGATTT-ATGC-A-C-T-AAGCA-----
-----TAATGTCCAGTTTATATGGGTAAAGGACATGCTGGACATC-CTG
AAAACGA--ACGTTGCGATCAATTAG---CTGT-TCTTGC---CTC-TAA-----
-----A---AAA---AT---G-G--AATTA-----T-
-----T--TA--T-----
-----AG---ACAAAG-----TT-----TTT-G-----
-----AA--AGTAT---A-G-C--TAATTGA

>Cyanobacterium_UCYN-A

TT-GAA-AAA-----AT--TAAAAAT-TTATACTGATGGTGCCTTGCTC-T
GGAAATCCAGGTAAAGGAGGTTATGGAGTAATATTAGT--GTTTAAC-AAA--TA-----
-----TC--G-AAAAGAATTATCAGGAGGCTATAGACTAACAATAATCGTATGGA
AATAA-TATCGATTATTATTGGGTTAGAGGCTTTAACAAA----A-----TTGT---
-----GTA-ATGTAACTATTTATTCTGACTCTCGTTACGTAGTAGACACTATGAATAA
---AGGATGGGCAAGTAAATGGCAATCTAATAATTGGAAACGCAATAAACAAGAAATAGC
CAAGAATTCTGATCTATGGAAAAGGTTGTTAGATCT-ATGC-A-C-T-AAACA-----
-----TAATGTTCAATTTATTTGGGTCAAAGGGCATTCTGGACATC-CTG
AGAACGA--ACGATGCGACCAATTAG---CTGT-CCTTGC---TTC-TAA-----
-----A---AGA---AT---G-G--AATTA-----T-
-----T--CA--T-----
-----AG---ACAAGG-----TT-----TTT-G-----
-----AA--AGTAC---A-G-T--TAACTAA

>Crocospaera_watsonii_WH_8501

AT-GAA-CAA-----AG--TACAAAT-TTATACCGACGGGGCTTGTTTC-T
GGCAACCCAGGGAAAGGCGGTTACGGGATTATTTTAGC--CTACAAT-GAA--CA-----
-----CC--G-AAAAGAACTCTCCGGCGGTTATCGTTTAAACCACTAATAACCGTATGGA
AATGA-TGGCCGCTATTATTGCCCTCGAAGCCCTAAATAA----A-----CCCT---
-----GTG-ATGTAATTTTATATACTGATTCCCGTTATGTAGTCGATGCTATTACCAA
---AGGATGGGCTAAAAAATGGCAAGCCAATGATTGGCAAAGAAACAAAAAAGAACAAGC
CAAGAATCCTGATTTATGGCAAAGATTATTAGACCT-ATGT-G-A-A-CAACA-----
-----TCAAGTAGAATTTGTCTGGGTAAAGGTCACGCCGGACACC-CCG
AAAATGA--ACAATGCGATCGCTTAG---CAGT-GGCCGC---TTG-TCA-----
-----A---GAA---GT---A-G--AATTA-----T-
-----C--CA--T-----
-----TG---ATGCAG-----TT-----TAT-G-----
-----AA--G-----A-GCA--AAAGTAA

>Cyanotheca_sp_CCY_0110

AT-GAA-AAA-----AG--TACAAAT-TTACACAGATGGGGCTTGTTTC-T

GGTAATCCCGGAAAAGGGGGTTATGGCATCATTTTAGT--CTATAAC-GAA--CA-----
-----CC--G-AAAAGAACTATCAGGAGGGTATCGCTTAACCACTAATAATCGCATGGA
AATGA-TGGCCGCTATCATTTGGTTTAGAAGCTTTAAAAAC----C-----CCCT---
-----GTG-AGGTGACATTATACACAGACTCTCGTTACTTAGTTGATGCTATCACTAA
---AGGATGGGCAAAAAAATGGCAAGCCAACGGTTGGAAAAGAAATAACAAAGAAGCTGC
AAAAAATCCTGATTTTATGGCAAAAACTGTTAGATTT-GTGT-A-A-A-AAACA-----
-----TGAAGTAAAATTTGTCTGGGTAAAAGGTCACGCAGGCCACC-CAG
AAAATGA--ACAATGCGATCGCTTAG---CTGT-CACAGC---AAC-CCA-----
-----A---CAA---TT---G-A--CACTT-----G-
-----C--CA--T-----
-----TG---ATGAAG-----TG-----TAT-G-----
-----A--ACTTTGA
>Arthrospira_sp_PCC_8005
AT-GT-----
-----GTGGT-AGT--CA-----
-----CC--G-GAAGGAAATATCGGGAGGCTTCCGACTAACTACTAATAACCGCATGGA
AATGA-TGGCAGCGATCGCCGCCTTACGAGCCCTCAAATT----T-----CCCT---
-----GTT-CTGTGACCCTCTATAGTGATTCTAAATATCTCGTAGATGCTATGACCCT
---GGGATGGGCTAAACGTTGGCAAAAGAATGGATGGCGACGTAATCAGAAGGAATGGGC
GAAAAATCCCGATTTGTGGGCGCAATTATTAGGTTT-GTGC-G-A-G-GAACA-----
-----TCAAGTCCGGTTTGTCTGGGTAAAGGGTCATGCAGGCGATC-GCG
AAAATGA--AATTTGCGATCGTCTTG---CTGT-AGAAGC---TAC-TCA-----
-----T---AGA---GA---T-A--GTTTA-----C-
-----C--GC--C-----
-----AG---ATGCAG-----GA-----TAC-GAAAACCCACCCC
AACC---TCAAGA-CATC--GAT-T---C-TAT--GAGTTAA
>Lyngbya_aestuarii_BL_J
AA-AAT-TCATCAGA-ACTTCAAGAAG--TAATGTT-ATATACCGATGGGGCTTGTC-A-G
GGAAATCCGGGGCCAGGGGGTTATGGAATTGTATTAAT--TCGTGGA-GAT--CA-----
-----TC--G-AGAAGAATTATCAGGAGGTTTTTCAGTTCACCACCAATAATCGCATGGA
AATGA-TGGCGGCAATTGTGGGGTTGGAAGTATTAGAGAA----A-----AAAA---
-----GCA-AAGTAACGCTTTATAGTGATTCTAAATATGTGGTTGATGCGATTGAAAA
---AGGGTGGGCTGAACGTTGGCAAACCAATGGCTGGAAACGAAATAAAAAAGAATTAGC
CATGAACCCCGATCTTTGGGAACAGTTATTAAACT-GTGT-T-C-A-CAGCA-----
-----TCAAGTAAAATTTGTTTGGGTCAAAGGTCATGCAGGCAATC-GTG
AAAATGA--ATGTTGCGATCGCCTAG---CAGT-TCAAGC---CTG-TCA-----
-----A---CAG---CA---A-A--ATCTA-----C-
-----T--TC--A-----
-----AG---ATTTAG-----GC-----TAT-G----AAAA----
-TCC--AGAGATGC-AGCA--AATTT---C-GTT--GTTTTAA
>Lyngbya_sp_CCY_8106
AA-AAT-TCATCAAA-ACTTCAAGAAG--TAATATT-GTATACTGATGGGGCTTGTC-A-G
GGAAATCCGGGGCCAGGTGGTTATGGGATTGTGTTAAT--TCGTGGA-GAT--CA-----
-----TC--G-AGAAGAATTATCGGGGGGTTTTTCAGTTCACAACCAATAATCGCATGGA
AATGA-TGGCGGCGATTGTTGGACTGGAAGTTTTAGATAA----A-----AAAA---
-----GTA-AAGTTAAGCTTTATAGTGATTCTAAATATGTGGTTGATGCAATTGAAAA
---AGGGTGGGCAGAACGTTGGCAAGCTAATGGCTGGAAACGAAATAAAAAAGAATTAGC
CATGAATCCGGATCTTTGGGAACAGTTACTAAACT-TTGT-T-C-A-CAGCA-----
-----TCAAGTAAAATTTGTTTGGGTCAAAGGTCATGCAGGAAATC-GCG
AAAATGA--ATGTTGTGATCGCCTAG---CAGT-TCAAGG---CTG-TCA-----
-----A---CAG---CA---A-A--ATCTA-----C-
-----T--GC--A-----
-----AG---ATGTAG-----GC-----TAT-G----AAAA----
-TCC--AGAGATGC-AGCA--AATTT---C-GTT--GTTTTAA
>Roseobacter_denitrificans

AT-GCC-TGA-GT-T-----GTTTCGC-CTATACGGATGGTGCCTGTTC-G
GGCAACCCGGGCCCCGGCGGCTGGGGCGTTTTGCTGCAGGCCAAAGAGGG---CGACAGG
C-TGGTCAAGGAACGCGCGCTCAAGGGCGGTGAGGCACATACCACCAACAACCGGATGGA
GCTTT-TGGCCGCCATCAACGCGCTGGAATCCCTCAGCCG----C-----GCCA---
-----GCA-CGATCACCGTCGTGACCGACAGCAACTACGTCAAAAATGGCATCACCGG
---C---TGGATTCATGGCTGGAAGCGAAACGGCTGGAAAAATGCCGCCAAGAAACCCGT
CGCGAACGCAGAACTCTGGCAAAGGCTGGATGAGGC-CAA--C-GCA-CGTCA-----
-----CGACGTCACGTGGAATGGGTCAAAGGGGCACGCCGGGCATG-CGG
AAAACGA--ACGCGCGGACGAATTGG--CGCG-CG---C---G---G--GC-----
-----AT--GGCC---CC-----G-----T-
-----T-----

-----C-----A-A--ACCATGA
>Roseobacter_sp_MED193
AT-GGC-TGA-TC-T-----CTATGC-CTATACCGATGGCGCCTGCTC-T
GGCAATCCCGGCCCTGGCGGCTGGGGCGCCCTGTTGCAGGCCAAAGATGG---TGGTA-G
CGTCATCAAGGAAAAAGAACTCAAAGGCGGCGAGGCCAATACCACCAACAACCGGATGGA
GCTAC-TAGCTGCCATCAATGCACTGGAAAGTCTAGACCG----C-----CCCT---
-----CGG-CGCTGACCGTGGTGACCGACAGCAATTACGTGAAAAACGGCATCACCGG
---C---TGGATCTTTGGCTGGAAGAAAAACGGCTGGAAAAACGCCGCCAAAAAACCGGT
TAAAAATGCCGAGCTGTGGCAGCGTCTGGATGCGGCGCAA--A-G-C-CGCCA-----
-----TCAAGTGACCTGGGAATGGGTCAAAGGTCACGCAGGCCATC-CCG
AAAACGA--ACGCGCCGATGAGCTGG--CCCG-CG---C---A---G--GT-----
-----AT--GGCC---CC-----G-----T-
-----TC-----
-----A---AGAAAA-----GC-----A-----
-----A-ATCC-----A-A--AGCATGA
>Escherichia_coli_K12-_W3110
AT-GCT-TAA-AC-A-----GG--TAGAAAT-TTTCACCGATGGTTCGTGTCT-G
GGCAATCCAGGACCTGGGGGTTACGGCGCTATTTTACG--CTATCGCGGA---CG-----
-----CG---AGAAAACCTTTAGCGCTGGCTACACCCGCACCACCAACAACCGTATGGA
GTTGA-TGGCCGCTATTGTGCGCGCTGGAGGCGTTAAAAGA----A-----CATT---
-----GCG-AAGTCATTTTGAGTACCGACAGCCAGTATGTCCGCCAGGGTATCACCCA
---G---TGGATCCATAACTGGAAAAACGTGGCTGGAAAACCGCAGACAAAAAACAGT
AAAAAATGTCGATCTCTGGCAACGTCTTGATGCTGCATTG--G-G-G-CAGCA-----
-----TCAAATCAAATGGGAATGGGTAAAGGCCATGCCGGACACC-CGG
AAAACGA--ACGCTGTGATGAACTGG--CTCG-TG---C---CGC-G--GCG-----
-----AT---GA--AT-C---CC-----A-----C-
-----ACTGG--A-----
-----AG---ATACAG-----GC-----T-----
-----ACC--AAG-T---T-G-A--AGTTTAA
>Alteromonas_macleodii_English_Channel_615
GT-GGC-GCA-AA-A-GA---C---CA--TTCATAT-TTATACCGACGGCTCCTGTCT-A
GGCAACCCAGGCCCTGGTGGGTATGGTGCTGTATTAAT--TTATAAACAG---CA-----
-----TA---AGAAAGAGCTAAGCGACGGATTTGCGCATAACCACAAATAACCGCATGGA
GCTGC-TGGCCCCTATCGAAGCGCTAAACAGCCTGACTGA---G-----CCTT---
-----GTG-CGGTAGAACTCACTACCGACAGCCAGTATGTAAAAAATGGGATTAACCA
---A---TGGATCCATAACTGGCGCAAAAACGGTTGGCGTACGACTGATAAAAAGCCAGT
GAAAAATGCCGATTTATGGCAACGGTTAGATGAAGCAGTG--A-A-A-AAGCA-----
-----TCAAGTGAATTGGCACTGGGTCAAAGGCCACTCGGGGCACC-CCG
AAAATGA--ACGCTGCGATGAGCTTG--CTCG-CGGGGC---CGC-C--GAA-----
-----GC---GA--AA-C---CA-----A-----C-
-----TCAA--T-----
-----CG---ATGAAG-----GT-----T-----
-----T-----TG-T---T-G-G--CAACTAA

>Candidatus_Pelagibacter_ubique_HTCC1062
AT-GAT-T-----AAAAT-TTATACAGATGGATCATGTTT-A
AATAACCCCGGTAATGGTGGTTGGGCTGCAATTATAAA--CCAAAAT-GG---TGA----
-----A--GTCAAAGAAATTAGTGGGAGCGTAAAAGACACTACTAATAATAAAATGGA
ATTGA-TGGCTCCTATTATGGCTCTAAAACAAATTAAACA----A-----AATG---
-----ATC-AAATTGAAATCTATACCGACTCACAATATGTAAGATTAGGAATTACTGA
---A---TGGGTGCATAAATGGATAAAAAATAATTGGCAAACCTTCAAAGAAAGAACCTGT
TAAGAATAAAGAACTTTGGATACAACCTATACGAGCT-AACA-A-A-T-CAATT-----
-----TGAAATAAAATGGATTTGGGTAAAGCACATGCAGGTAATG-CCC
TAAATGA--AGAAGTAGATTTACTCG---CAA-AAA--A---G-----CTG-----

-----CTG-----A-----
-----AT-----T--AAACTAA
>Alpha_proteobacterium_IMCC1322
AT-GAG-CAC-TA-G-----TG--TGACGAT-CCATACCGACGGGTCATGCCT-G
AGCAATCCTGGCCCCGGCGGCTGGGCAGCGATTCTGAA--TTGGCGG-GG---CAC----
-----G--GAAAAAGAATTATCAGGTAGTGTGGCCGATACGACCAACAATCGCATGGA
ATTGC-AAGCCGCCATTGAAGGACTGAATGCGCTAAGCCG----T-----CCAA---
-----TGC-AGGTGCAACTTCATACTGACAGCAAATATGTCATGAATGGCGTCACTGA
---C---TGGATGGCACGCTGGAAAGCAAATGGCTGGCGTACGGCGGCCAAGAAACCTGT
CGCCAATCAGGATTTATGGGAACAGCTTGATGCGGC-CACC-C-A-A-CGTCA-----
-----TCAGATCACCTGGCATTGGGTAAAGGTCATGCAGGTAACG-AAC
TCAATGA--ACGTTGTGACGTGCTGG--CACG-AAG--C---G-----CCG-----

-----CCG-----A-----A-----
-----AGCC-----T--GAAGTGA
>Marinobacter_sp_ELB17
AT-GAG-CGG-TA-A-----CG--TGATTAT-GTACACTGACGGCGCCTGCAA-G
GGCAACCCCGGGCGCGGCGGTTGGGGAGTGGTGCTGCG--CTGGGGC-GA---GGT----
-----TT--G-CAAAACCTTGCACGGTGGCGAGCAGCACACCACCAATAACCGCATGGA
GCTGA-TGGCCGCGATTGAAGGGCTGAAAGCGTTGAAGCG----T-----GATT---
-----GTG-ACGTAGAGCTATATACCGACTCCCAGTACGTGCGCAAAGGCATCACCGA
---G---TGGCTGGCCGGCTGGAAACGCAACGGTTGGAAAACAGCGGCTAAAAAGCCGGT
TAAAAATGACGATCTTTGGAAAGCCCTGGATGAACA-AAGC-G-A-G-CGCCA-----
-----CCGGGTAACTGGCACTGGGTGAAAGGGCACGCCGGCGTGC-CGG
ATAACGA--ACTGGCCGACCAGCTGG--CCAA-CCA--A---G-----GCG-----

-----TTGAAG-----AGC-----T-----
-----GACC-----GGTTAA
>Marine_bacterium_HP15
AT-GAC-CGG-CA-A-----GG--TAACCCT-CTATACGGATGGCGCCTGCAA-A
GGCAACCCCGGGCCCGGCGGCTGGGGCGTGGTTTTGCG--CTATGGC-GA---CAA----
-----TC--G-GAAGACCCTGCACGGCGGGGAGACGAATACCACGAACAACCGCATGGA
ATTAA-TGGCAGCCATCCGTGGTCTGGAGGCCCTGAAGCG----C-----CCGT---
-----GCG-AGGTGGAGCTGTTTACGGATTTCGCAGTACGTGCGCAAAGGCATTACGGA
---A---TGGATGGCAGGCTGGAAGCGGAACGGCTGGAAGACCTCCGCGAAAAAGCCCGT
GAAGAACGAAGATCTTTGGCGCGAGCTGGATGC-CG-AAGTGG-C-C-CGGCA-----
-----CCAGATTAACCTGGCACTGGGTAAAAGGGCACTCTGGTGTTC-CTG
ACAATGA--GCTGGCGGATGAACTGG--CCAA-TCG--C---G-----GCG-----

-----TCGACG-----AGC-----T-----

-----TTCC-----G-G--CGCATAG
>SAR86_cluster_bacterium_SAR86A
AT-GAA-AAC-----AG--TAAATAT-ATATACAGACGGTGCATGCAG-A
GGAAATCCGGGAGATGGTGGATGGGGTGCTTTAGTAGA--GTATCAA-AA---TTT----
-----TA--G-TAAAGAGTATTTTGGTGGCCAGAAAAATACTACTAATAATCAAATGGA
ATTAA-AAGCTGCAATTGAGGGTCTTAAAGCATTGAAAGA----A-----ACAT---
-----GCA-ATGTTAATTTAACCACTGATTCGAAATATGTTATGGATGGAATTACTTC
---T---TGGATAGATAATTGGAAAAAAATAATTGGAAAAGTTCATCAAAAAGGATGT
TAAAAATAAAGATCTATGGATAGAGTTGGATAAATA-CAGG-T-C-T-ATGCA-----
-----TAATGTAAAATGGATTTGGGTAAAGGCCACTCAGGGCATG-AGA
AAAATGA--AATCGCTGATATGTTGG---CAAA-CAA--G---G-----GTA-----

-----TAG-----A-----T-----
-----A-G--CATTTAA
>Synechocystis_sp_PCC_6803
AT-GGC-CTC-TA-C-CCCCAATTCCG--TAACTCT-CTACACCGATGGTGCCTGCTC-C
ATGAATCCTGGCCCTGGTGGCTATGGCGCAGTGATTCT--CTACGGT-GATGGTC-----
-----GT--CGGGAGGAATTGTCCGCCGTTATAAAATGACCACCAATAATCGCATGGA
AATTA-TGGGGGCGATCGCCGCCTTGAGTCATTTACAA-G---A--A-----CCGT-C-
C-----C-AAGTGTTGCTCTATACCGATTCCCGCTATATGGTCGATGCCATGAGCAA
GGGT---TGGGCCAAGAAATGGAAAGCCAACGGATGGCAAAGGAATGCCAAGGAAAAGGC
CAAAAACCCCGACCTGTGGGAAACGATGTTAACGCT-GTGC-G-A-A-AAACA-----
-----TCAAGTTACATTCCAGTGGGTCAAAGCCCACGCTGGCAATA-AGG
AAAATGA--GCGGTGCGATCGCCTAG---CGG-----T---GGC-CGCCTA-----
-----TC---AA---AAT---A-----
-----AT-----C--CCA---A-
----CTTAGTAG----ATGAAG-----GC-----TT-----
-----C-----GG---C-A-A--ATTCTAG
>Synechocystis_sp_PCC_6714
AT-GGC-TCC-CA-C-TGCAAACCTCTG--TAACCCT-CTACACCGATGGTGCCTGCTC-G
AGAAACCCTGGCCCTGGAGGCTATGGGGCAGTGATTCT--TTACGGT-GATGGTC-----
-----GC--CGAGAGGAGTTGTCTGCCGTTATAAAATGACCACCAATAACCGCATGGA
AATGA-TGGGGGCGATCGCCGCTTTAAGTTATTTGGAA-G---A--A-----CCCT-C-
C-----C-AGGTGTTGCTCTACAGCGATTCCCGCTATGTGGTGGATGCAATGACTAA
GGGC---TGGGCCAAGAAATGGCAAGCGCAAGGATGGCAAAGAAACGCCAAGGAAAAGGC
TAAAAATCCGGATCTCTGGGAAATAATGTTAGCCCT-ATGC-A-A-A-AAACA-----
-----CCAGGTCACCTTTTCAGTGGGTCAAAGCCCATGCTGGCAATA-AAG
AAAATGA--ACGGTGCGATCGCCTAG---CGG-----T---GGC-GGCCTA-----
-----CC---AA---AAC---A-----
-----AT-----C--CCA---G-
----CTTAGTGG----ATAAGG-----GC-----TT-----
-----T-----GG---C-A-A--GTTCTGA
>Chloroflexus_sp_Y-400-fl
TT-CTTCCTC-CA-C-AAACGGGATTG--TGATGAT-GTACACCGACGGTAG-TGCTCTC
GGCAATCCTGGGCCGGGCGGGTACGGCGTTGTGCTGCG--CTACAAT-GACCA-----
-----T--TACCAGGAACCTACCGGTGGTTTTTCGGCTCACGACCAACAACCGCATGGA
GTAA-TGGCCTGTATTGTCTGGGCTGCGTGCCTCAAG-C---G--A-----CCGG-T-
G-----T-CGGTACGCCTGCACAGCGACTCGCGCTACGTCGTCGATGCTTTTCGCAA
AGGC---TGGATACACCGCTGGCAGGCACAAAACCTGGATGCGCAGCGAGACCGAACCGGT
AAAAAATGCCGATCTCTGGGCTGAATTGCGCTTGCT-GTGT-G-A-A-GGTCA-----
-----TCAGGTTGAGTTTCGTCTGGGTGCCCGGTCATCAGGGCGTAC-CTG
ACAATGA--ACGTTGTCACGAACCTGG---CCA-----T---CGC-TGCCGC-----
-----AG---CT---CGC---C-----
-----CA-----G--ACC---TG

CCGCC-----GG----ATGTTG-----CCTTTGAGATG-----
-----CAGTGGAGT---C---G--CACATAA
>Synechococcus_sp_JA-3-3Ab
CACGGC-TGGCCG-ATTGCCCTCCTAG---TGGCAT-TTTTACCGACGGGGGCTGTTC-C
CCCAATCCAGGGCCAGGGGGATGGGCGGTGGTGGTGGT--TGAGGAT-GACCAGA-----
-----TT--GTGGATGAGTTTTACGGAGGGGATCCCGCCAGCACCAACAACCGTATGGA
ATTGACCGGCCTC-ATTCG-GGCTTTGCAATATTGTCGCGA----C-A-----CCCAGAG
CAA-CAGC-A-TAGACACCATCTACAGCGATTCCCACCTTTGCGTACAGACCTACAACGA
---T---TGGATGGAGCGCTGGGCAGCGCAGGGGTGGCGGCGG--CGC-ACGGGGCCGGT
GGAAAACCTTGATCTGGTGCAGGAGCTGTATCAACTCAAGC-AGC-AGTGCCC-----
-----GCAGGTGCGGGTGGCCTGGATCCCGGCCCATCGGGGTCTGCGCTG
G-AACGAGTATGTG--GATGAGCTGGTTCG-----

-----CC-----
-----GC-GCCCGA---C---C--CAAGTAG
>Synechococcus_sp_JA-2-3Ba
GCCGTC-TGGCTG-CCTGCCCCCGCAG---CGGCAT-CTTCACGGACGGGGGATGTTC-T
CCCAATCCGGGGGCCAGGGGGATGGGCCGTGGTGGTGGT--TCAAGAT-GATCGGG-----
-----TT--GTGGAAGAGTTTTGGGGAGGGGATCCCGATAGCACCAGCAACCGCATGGA
GTTGACCGGCCTC-ATCCG-GGCTTTGCGGTATTGCCGCCA----C-----CCCAGGG
CAA-CGGC-G-TCGATACCATCTACAGCGACTCCCACCTTTGCGTGCAGACTTACAACAA
---T---TGGATGGAGCGCTGGGCAGCCCAGGGATGGCGGCGG--CGC-ACAGGGCCAGT
GGAAAACCTTGATCTGGTTGAGGAGTTATATCAACTCAAGC-AGG-AGTGCCC-----
-----TCAGGTGCGGGTGGTCTGGATCCCAGCCCACCAAGGTCTGCGCTG
G-AACGAGTACGTG--GATGGGTGGTTGCC-----A---GGC-----

-----TCGAGC----CAGAGT-----GACCGTTCCTC-----
-----CAG-AAGTT---A---G--GGGTTGA
>Cyanothece_sp_PCC_7425
AT-GAA-C-G-----CAGAAATTA--GCAAAAT-TTATACTGATGGAGCTTGTTTC-C
GGCAATCCTGGTCCGGGGGGGATGGGGGGTGTGGTTTA--CTTTGGG-GATGGCT-----
-----CT--GTGCATGAGATGGGGGGGGCAGCAGCGGCCACCACAAACAATCGGATGGA
AATGC-AGGCTGCGATCGCCGCGCTGGAATTTTGTCAAGC----T--G-----CCGG-C-
CATC---CCC-CAGTTATCCTCTATAACCGACAGTGAGTATGTCAAAAAAGGGATTACAGA
---G---TGGCTACCAGGCTGGAAACGCAAGGGATGGAAAACCGCTCAGGGTAAACCCGT
GCTCAATCAAGACCTCTGGCAGGTGCTGGATCAACT-G----A-A-C-ACCAA-----
-----AAGTGTGAACTGGCAGTACGTGCGGGGCCACTCTGGTAATG-CTG
GCAATGA--GCGCTGTGATGCGATCG---CCCG-AGCCTT---TGC-CGCCGG-----
-----AA---GG---CAA---CC---A-G-----

-----A---G-----CTGC-----
-----G-TC--AGAGC---C-G-AGACCTGTAA
>Halothece_sp_PCC_7418
AT-GTC-TCA-AC-T-CCCCGAAATTG--ATTGTAT-TTATACTGACGGCGCTTGTTAC-G
GGAAACCCAGGCCCTGGCGGGTGGGGCGCTGTCTTTTA--CTTTAAA-AATGGTT-----
-----CA--GTATATGAAATCGGTGGCGCGGAGCCGGAAACGACCAATAATCGGATGGA
AATGC-AAGCAGCGATCGCGTCATTAGAGTTATTTCAAGC----C--A-----GTGA-A-
CAAAAAACAC-CGATTGTTCTCTACAGTGACAGTCAGTATCTGATTAAAGGTCTCACCCA
---A---TGGCTCAAAGGCTGGAAGAAAAAAGGCTGGAAAACGTCAACGGGCAAAGCTGT
TTTAAATCAAGACCTGTGGACAACATTAGACCAACT-G----A-ATC-ACTCT-----
-----CTTATTGAGTGGAAGCACGTGCGCGGTACGAAGGAATTG-CAG
GGAATGA--ACGGTGCGACGCGATCG---CGCG-GACTTT---TGC-TGCTGG-----
-----AA---AA---CGT---CC---T-C--ATTT-----

-----AA--A-GAAAAATTT-----AGATT----T-
-----TCCTC----TTCCCG-C---T-----AAC-----
-----T-CC--GATAC---C-G-TAAT-GAGTG
>Microcoleus_sp_PCC_7113
AT-GCC-TTC-TA-C-TCGCCAAATTA--AATGTTT-TTACACCGATGGCGCTTGCGC-T
GGCAACCCTGGTCCCGGCGGCTGGGGTACTGTAGTTTA--TTTTACC-GATGGCT-----
-----CC--GTTTATGAAATGGGTGGTGGGACTGCTCAAACCACAAACAACCGAATGGA
AATGC-AAGCAGCCATTCAAGGCGTTGAAGGTTCTCAAGGC----A--T-----CGGC-T-
CAAACCGAAT-CCGTCATCCTCTACACCGATAGTGAATATGTCAAGAATGGCGTCACCAA
---A---TGGGTGAAAGGCTGGAAAAAGAAAGGCTGGAAAGACAGCACAGGGAAAAGCTGT
TTTAAATCAAGATTTGTGGGAACTCTTGATGAGCT-G---A-A-T-TCTCC-----
-----CTTGGTGGAGTGGCAGTATGTTCCGGGGGCACAGTGGCAATG-AGG
GCAATGA--GCGTTGTGATACAATCG--CTCG-TGCTTT---TGC-TGTTGG-----
-----CA---AG---C-----CC---T-TGGTACTC-----G-
-----A--TC--A-GGCGTTCGA-----CCTGA----G-
-----TAC--TC---CACCTCC--A-----ATC-----
-----A-CTCA--AAC-----TCCAACCTCA
>Synechocystis_sp_PCC_7509
AT-GTC-TTC-TT-G-CGGCACAATTC--AAAGCGT-GTATACAGATGGAGCTTGACAC-T
GGCAATCCTGGCCCTGGGGGTGGGGCGTAGTTGTTTA--CTACACA-GATAAAT-----
-----CA--GTGCGCGAATTTGGCGGCAATGAAAAGCAAACGACCAATAATCGTATGGA
GATGC-AGGCAGCGATCGCAGCATTAACCTTTTTTAACCGA----A--A-----CACC-G-
CCCACAGAAC-CAATAACGTTGTATACCGATAGCGAATATTTAATTAAAGGTGTAACCTCA
---G---TGGGTAAAAAATTGGAAAAAAAAGGCTGGAAAACCGCCACTGGTAAGCCCGT
ACTAAATCAAGACCTGTGGGAAAACCTCGACCAATT-G---A-A-T-AGTAA-----
-----ATTGGTATTGTGGCAACACGTCCGGGGTCATGCGGGAAATG-TGG
GTAATGA--AAGGTGCGATGCGATCG--CACG-GGCGTT---TGC-TAATGG-----
-----AA---AA---A-----CC---T-TATCGCTA-----C-
-----A--AC--A---AGGCAAA-----CC-----
-----AAAC---CTCCA-C---A-----AAC-----
-----AACA--GAAT-----TAGCTGAAAA
>Cylindrospermopsis_raciborskii_CS-505
AT-GTC-AAA-TC-A-ACGCATAATTA--AAAGTAT-ATATACCGATGGAGCTTGCAA-G
GGAAACCCAGGTCCTGGTGGTTGGGGTGTGGTTGTTTA--CTTTCAT-GATGGTT-----
-----CA--ATTCATGAAATGGGGGAAGCATCTGACTACACTACAAACAATCAAATGGA
AATCC-AAGCTGCGATCGGGGCACTTGAATTATTAGCTCA----A--T-----CCGG-A-
CAAGTTGAAC-CTATTACTCTATATAGCGATAGCGAATATCTAATTAATAGTGTGACTAA
---A---TGGATACATGGATGGAAAAAAAAGGGTTGGAAAAAATCAGATGGCACTCCGGT
GCAAAACCAAGAGCTACTAAAAATTATAGACAAGCT-G---A-A-C-ACCAA-----
-----GCTAGTCAGGTGGCAACATGTACGTGGACATGCGGGTAATA-TAG
GTAATGA--AAGATGCGATGTTATTG---CCCG-TTCTTT---TGC-GAGTGG-----
-----CG---AA---ATT---CC---C-C--CTTTG-----C-
-----A--AC--A---A--TTAT-----C--CCC---C-
----C---AA-----CAGTTG-C---TAAC--TAA-----TT-----
-----G-TCAA--GAAAC---C-T-T--AACCCAA
>Raphidiopsis_brookii_D9
AT-GTC-AAA-TC-A-ACGCATAATTA--AAAGTAT-ATATACTGATGGAGCTTGCAA-G
GGAAACCCAGGTCCTGGTGGTTGGGGTGTGGTTGCTTA--CTTTCAT-GATGGTT-----
-----CA--ATTCATGAAATGGGGGAAGCATCTGATCACACTACAAACAATCAAATGGA
AATCC-AAGCTGCGATCGCAGCCCTTCAATTATTCGCTCA----A--T-----CCGG-A-
CAAGTTGAAC-CTATTACTCTGTATAGTGTAGCGAATATCTAATTAACAGTGTGACTAA
---A---TGGATACATGGATGGAAAAAAAAGGGGTGGAAAAAATCAGATGGCACTTCGGT
GCAAAACCAAGATCTACTAAAAGTTATAGACAAGCT-A---A-A-C-ACCAA-----
-----TCTGGTCAAGTGGCAGCATGTACGTGGACATACGGGTAATA-TAG
GTAATGA--AAGATGCGATGTTATTG---CCCG-TTCTTT---TGC-GAGTGG-----

-----CA---AA---ATT----CC---C-C--CTTTG-----C-
-----A--AC--A---G--TTAT-----C--CCC----C-
----C---AA-----CAGTTG-C---TAAC--TAA-----TT-----
-----GTCAA--GAAAC---C-T-T--AACCCAA
>Nostoc_sp_PCC_7120
AT-GTC-TAC-CC-A-ACCGATAATCG--AAAGTAT-ATACACTGATGGTGCTTGCAC-T
GGCAATCCTGGCCCTGGCGGCTGGGGTGTAGTTGTCTA--TTTTAGT-GATGGTT-----
-----CA--GTTACGAAATGGGCGATGCTGCCAAGCATACCACCAATAATAAGATGGA
AATGC-AGGCGGCGATCGCCGCTCTCAAATTTCTCCACGA----C--T-----CCGG-A-
CAAGCTGAAC-CCATTACCCTCTATACCGATAGTGAGTATCTAATTAAGTGCATCACC
---G---TGGGTAAAAGGTTGGAAAAAAGGGTTGGAAAAAATCTGATGGGAATCCAGT
TCAGAATCAAGACCTTCTAGAACTTTAGATGAATT-A----A-A-T-AGCCG-----
-----TAAGGTAAATTGGCATCATGTTTCGAGGTCATTCTGGCAACA-TCG
GTAACGA--GCGTTGTGATGTGATTG---CTCG-CTGTTT---TGC-AACTGG-----
-----TA---GA---ATG----CC---C-T--CTCTA-----C-
-----A--AC--A---G--TTAT-----C--TCC----C-
----C---GT-----CA--TG-CTCA---TAAATC-----TT-----
-----T-ACCT--GACAT---C-A-G--TAAAATA
>Rivularia_sp_PCC_7116
AT-GTC-TTC-TC-A-GCCTACAATTTC--AAAATAT-ATATACTGATGGCGCTTGTAC-C
GGCAATCCAGGTCCCGGTGGTTGGGGTGTGTTGTTTGA--CTTCGAT-GATGGTT-----
-----CG--ATTCACGAAATCGGTGATGCATCAAAAAAACCACCAATAACAGGATGGA
AATGC-AAGCCGCGATCGCAGCTTTGGAATTTTTAGAAATC----T--T-----CCGG-G-
CAAAACAAAC-CCGTTACTCTTTACACCGACAGCGAATATCTAATTAAGTGCATACGAA
---G---TGGGTGAAAGGATGGAAGCGTAAAGGTTGGAAGAAATCCGACGGAAACCCGT
GCTTAACCAAGATTTATTGGAATTTCTAGATAATCT-T----A-A-T-AATCA-----
-----AAAGATTAAATGGCAACACGTTTCGCGGTACGCTGGTAATA-TAG
GTAACGA--ACGTTGCGACACAATCG---CCCG-TGGTTT---TGC-TTCAGG-----
-----GA---AA---ATA---CC---A-T--CGCTA-----A-
-----A--GC--A---A--ATGC-----T-G-----G-
--ATA-----TAAATT---TCAGTGATGC-----TA-----
-----A-TT-----TCGCACAA-A-A--AGCCGAA
>Nodularia_spumigena_CCY9414
AT-GTC-CAA-CC-A-ACGCAAAATTTC--AAAGCAT-CTACACCGATGGCGCTTGTAC-T
GGGAACCCTGGCCCTGGGGGTGGGGTGTGGTTCGCTTA--TTTCAGT-GATGGCA-----
-----AA--ATCCACGAAATGGGTGATGCTGCCAGCTATACGACTAATAATAAAATGGA
AATGC-AAGCGGCGATCGCAGCTTTAAATTTCTTCCACGC----A--T-----CAGG-A-
CAAGCTGAAC-CCATCACCTCTATACCGATAGCGAATACTTGATCAACTGCGTGAATAA
---G---TGGTTACCAGGATGGAAAAAAGAGGCTGGAAAAAAGCAGACGGTAAACCTGT
CCAGAACCAAGAACTTTTAGAACTTCTCGATCAACT-C----A-A-T-ACTCG-----
-----CCAAGTAAATTGGCAACACGTCCGGGGACATTCTGGTAACA-TAG
GTAACGA--GCGCTGTGATGTGATTG---CTCG-CTCCTT---TGC-TAGCGG-----
-----TC---AA---ATC---CC---A-T--CTCTA-----C-
-----A--AG--A---A--TTAT-----C--C-----G-
----C---TA-----CTGAGT-C---CTACAAATC-----TT-----
-----T-AC-----ATTCTTTA-G-G--TGAAGCA
>Calothrix_sp_PCC_7103
AT-GGC-TTC-AC-T----ACAATTG--AGAGTAT-ATACACAGACGGCGCCTGTAC-A
GGGAACCCAGGACCAGGAGGTGGGGTGTGGTGGTTTA--TTTTGAT-GATGGTT-----
-----CG--GTTTCATGAACCTTGGAGATGGTTGTGCTCAAACAACGAATAACAGGATGGA
AATGC-AAGCTGCCATCGCAGCACTTCAATATTTGGAAGC----G--T-----CAGG-G-
CAAAGTTACC-CTATTACTTTGTACACAGATAGTGAGTATTTGATCAATTGCGTTACCAA
---G---TGGGTAAAAGCTGGAAAAAGAAGGGTTGGAAAGAAAGCTGATGGTAAGCCAGT
TCAGAATCAGGATTTACTCGAAATTCTTGACGGACT-T----A-A-C-AGTCG-----
-----TAAGGTAGATGGGAACATGTTTCGAGGTCATGCTGGTAATA-TAG

GTAACGA--AAGATGCGATACCATTG---CGCG-GGCTTA---TGC-GGCTGG-----
-----AA---AA---GTG---CC---A-G--TATTA-----A-
-----A--GC--AAAGC--CAAG-----ACG-----C-
--AGA-----AA---CTATTG-C---TTTC--G-A-----TG-----
-----T-AAAA--ACAAA---A-G-T--TTATGTA
>Richelia_intracellularis_HH01
AT-GTC-TCC-AG-T-TCATAAAATCC--AAAGTAT-ATACACCGATGGTGCCTGCAC-T
GGGAACCCTGGTCCTGGTGGTTGGGGTGTGGTTATCTA--TTTTAAT-GATGGTT-----
-----CG--ATTCATGAAATAGGCGAGCCATGTGATTGCACTACTAATAACCGGATGGA
GATAG-AAGCTGCAATTGCAGCTTTGCGATATCTGCAAAC----T--A-----GTGA-A-
AAAAATGAAC-AAATTACCCTTTACACAGATAGTGAATACTTAATTAAGTGTGTAACAAA
---A---TGGGTAACAGGCTGGAAAGCAAAGGGGTGGAAGAAAGCCGATGGCAAACCAAGT
TGTTAATCAAGACTTGCTAGTAACTTTAGATGAACT-T----A-A-C-AGTCA-----
-----TCACATAAAATGGCAACATGTCAGGGGTCATTCAGGAAATG-TGG
GAAATGA--ACGTTGTGATATGATCG---CTCG-TGCATT---TGC-CAATGG-----
-----TA---CC---GAA---CC---A-T--TCCTA-----A-
-----C--AC--A---A--ATTT-----CT-A-----G-
----C---TGCA----ATTCAC-A---ACC-AC-A-----AA-----
-----ATGAA--TATGA---T-T-T--GGCTGCA
>Leptolyngbya_sp_PCC_7104
AT-GGC-AAT-TA-T-----TG--AGCGAAT-CTATACAGACGGAGCCTGCTC-C
GGCAATCCTGGCCCCGGCGGCTGGGGAACCGTGCTGTA--CCTGGCC-GAGGGCG-----
-----GC--GTGCATGAGCTGGGCGGCGGCGAGGCCCAAACCAACAACCGGATGGA
AATGC-AGGCTGCGATCGCAGGTCTGACCCTGCTGCGCGATAGCG--G-----CCA----
--GACCCAGC-CTGTCACCCTGCATACCGATAGCGAGTACGTGCTCAAGGGCATCACCCA
---G---TGGATTGCCGGCTGGAAGCGGCGGGGCTGGGTGAACTCGGCCAAAAAGCCGGT
CCTCAACCGCGACCTGTGGGAGGTTTTGGATGCCTT-G----A---C-CACCGAGGTCAA
TCAAACCCTCGATCGCCCCCTCCAGTGGGTCTACGTGCGCGGCCATTCGGGGCGA-T-GTG
GGC-----AACGAGCGCTGCGA-CA---C-----CAT---TGC-CCGGGC-----
-----CT---TT---GCG---GC-----C-----
-----AACCGCGCGATC--GCC----C-
-----TGGCTA-----C-----
-----GG--CATCG---T---T--GCCCTAG
>Leptolyngbya_sp_PCC_7376
AT-GAT-CA-----AACGAAT-TTATACCGACGGCGGCTGCAT-C
AATAATCCCGGTTCCGGTGGCTGGGGCGTTGTGATTCA--ATATCAA-AATGGCT-----
-----CT--GTTTCCGAACCTTGGCGATCATGAAGCCCAAACCACAAATAATCGCATGGA
GATGC-AAGCTGCCATCGAAGCCCTCAAACAATTTTCGAGATAGCG--A-----TCA----
--GCGCGAAC-CCATCCAGCTCCTTACCGATAGTAAATACGTAATTGATGGCATTACGAA
---G---TGGATTAAAGGCTGGAAACGGAAAGGCTGGAAAAATCATCCGGTGCACCAGT
CCTCAATAAAGATCTCTGGCAGGAAATGGATAAACT-G----A-A-T-CACCC-----
-----TGCTGTGCAATGGACTCACGTTAAAGGCCATAGCGGTGAGC-CTG
GGAATGA--ACGTGCCGATGCGATCG---CCAA-CAGTTT---TGC-CCGTGG-----
-----CA---AA---GCA---CC-----C-----A-
-TTTT-AA--AA--C-----ACCTCGTTTACAAAACGGCGAAAAA-AT-ACA----A-
-----AATCTT-----C-----
-----CC--AAATC---C---A--GAGATAT
>Gloeobacter_kilaueensis_JS1
AT-GAT-C-----CGGAT-CGTTACCGATGGCGCGTGCCT-G
GGCAATCCCGGCCCGGCGGTTGGGCGGCGTTGAT-----CTTGCAG-GACGATC-----
-----AA--ATCGAAGAATACAGCGGCTTCGAGGCCAGACCACCAACAACCGCATGGA
GATGC-AGGCCGTGATCGAGGGGCTGCGCCGCGTGGGCG---CT--A-----CTG----
-----TCC-CGGTGACGGTAATCTCCGATAGCCAGTACGTCATCAAAGGCATGAGCGA
---A---TGGCTGCCCCGGCTGGAAACGCCGCCAGTGGCGCACCGCCGCGGTAAGCCCGT
CGAGAACCGCGACCTCTGGGAAGCACTCGAAAAAGT-G----G-C-G--GCGG-----

-----GCCGGGTGCAGTGGCAACACGTCTACGGCCATCGCGGCCATG-CTG
AGAACGA--GCGGGCAAACGATCTGG---CCCA-GGCTGCCGCTGC-CCGCAGACCGGTG
CCGCCCCGCTGTCTG---AA---CCT----GC---TGC--TTTTG-----C-
--AACCAGC--GC--T-----CCCCGCTCAGACGGTATTACCTATC-----TC----A-
-----GCTTGA-----T-----
-----CG--ATGGC---C---A--GATCGAG
>Gloeobacter_violaceus_PCC_7421
AT-GTT-T-----CGGAT-AGCGACCGACGGGGCCTGCAG-C
GGCAATCCCGGTCCGGGGGGGATGGGCGGCCCTGGT-----CGTGGGG-GAGGGCA-----
-----CC--TACGAAGAAGTGTTTCGGTTTCGAGCCGCACACCACCAATAACCGCATGGA
GATGC-GCGCCGTCATCGAAGGTTTGACGCGCGTACCCG----CG--G-----ACG----
-----CCA-AGGTGAAAGTGCTCACCGACAGCCAGTACGTGATCAAAGGGATGAGCGA
---G---TGGCTGCCCGGCTGGAAGCGGCGCGGCTGGATCACTTCCACCGGCAAACCCGT
CGAGAACCGGGATCTGTGGGAAGCGCTGGAGCGGGC-G----G-T-A--GGGG-----
-----GGCGCGTGCTCTGGGAGCATGTGCGCGGCCACACCGGCCACC-CCG
AGAATGA--GCGGGCCAACACGCTCG---CCCA-GACGGC---AGC-CCGCGG-----
-----CG---GT---GCT----CC---GGC--CGTTG-----G-
C-GCC-CGG--AC--T-----GTGCCCACCAATGGAACGACCTATCTG--AGT----C-
-----TGGTGG-----A-----
-----CG--GGCAC---C---T--CAG-CCG
>Prochlorococcus_marinus_str_NATL2A
AT-GTC-TAA-AG-A-AGAAAAATTAGCGATTGCTGCAGCAACCGATGGTGCATGCAG-T
GGTAATCCTGGTCCGGGTGGCTGGGGAGCATTAATCAG--ATTCCAG-GATGGAA-----
-----GC--GAGATTGAATTCGGTGGGAACTCCCCCGAGACAACGAATAATCGCATGGA
ATTAC-AGGCTGCATTATTTATTCTAGAAAAATTAAAA-AA--CA--T-----AAA--A-
TTTGCC-CCTTCACTAACTATTAAACTGATAGCAAATATTTAATTGATGGAATGGACAA
---A---TGGATGCCAAATTGGAAAAAGAAAGGCTGGAAACTGCCTCTGGCAAACCAGT
TCTCAATCAAGATTTGTGGAAAGCTCTTGACCATCC-T----G-A-A-CTACC-----
-----AAAAATCAAAC TTCAATATGTCAAAGGTCATAGCGGAGAAA-AGG
ATAATGA--CAGAGTCGATGCAATTG--CGGT-AGCTTT---TTC-TAAAGG-----
-----TA---G-GAAAAT---AC---A-A--CT-----
-----TA-A----A----G-
-----ATTTTG-----C-----
-----C-A-A--CAACTAA
>Prochlorococcus_marinus_str_MIT_9313
AT-GGC-TGA-TG-G-TCGTGGACGGGTTGTGGCTGCTGCCACTGATGGGGCTTGTAG-T
GGAAATCCTGGCCCTGGGGGCTGGGGGGGCACTGCTGCG--CTTTGAG-GATGGGA-----
-----GC--GTGGAGGAGTTTCGGAGGCTATGCGCCGGCCACGACCAACAACCGCATGGA
ACTTC-AGGCGGCCCTTCATGTGCTGGAGCAGCTCAAAGAG--CT--T-----CCT--T-
GTCACC-CCG-ATTTAAAGATCCGTACCGACAGCAAATATTTAATTGATGGTTTGAGCAA
---A---TGGATGGCGGGATGGAAACGTAAGGGTTGGCGTACGGCGGCTGGCAAGCCTGT
GCTGAATCAAGATCTCTGGCGAGCTCTAGACCGTGC-T---C-G-C-TTGGA-----
-----TCATGTGCCCCTGGCTTATGTGAAAGGTCATAGCGGTGACC-CAG
ACAATGA--AAGGGTTGACCAAATCG---CAGTGAGCTTT---T-C-CAAAGG-----
-----AG---GTGGTACT----CC---A-G--CGTTG-----A-
-TGTC-TGG--AC--A---G--GGGTCAGC-----C----T-
-----CTATTG-----C-----
-----ACTC--GGTCT---T-G-T--TGAGCCG
>Prochlorococcus_sp_MIT9303
AT-GGC-TGA-TG-G-TCGTGGACGGGTTGTTGCTGCTGCCACTGATGGGGCTTGTAG-T
GGAAATCCTGGCCCTGGGGGCTGGGGGGGCACTGCTGCG--CTTTGAG-GATGGGA-----
-----GC--GTGGAGGAGTTTCGGAGGCCATGCGCCGGCCACGACCAACAACCGCATGGA
ACTTC-AGGCGGCCCTTCATGTGCTGGAGCAGCTCAAAGAG--CT--T-----CCT--T-
GTCACC-CCG-ATTTAAAGATCCGTACCGACAGCAAATATTTAATTGATGGTTTAAGCAA
---A---TGGATGGCGGGATGGAAACGTAAGGGTTGGCGTACGGCGGCTGGCAAGCCTGT

GCTGAATCAGGATCTCTGGCGAGCTCTAGACCGTGC-G----C-G-C-TTGGA-----
-----TGATGTGCCCCCTGGCTTATGTGAAAGGTCACAGCGGTGACC-CAG
ACAATGA--AAGGGTTGACCAAATCG---CAGTGAGCTTT---T-C-CAAAGG-----
-----AG---GTGGTGCT----GC---A-G--CGTTG-----A-
-TGTC-TGG--AC--A---G--GGGTCAGC-----C----T-
-----CTATTG-----C-----
-----ACTC--GGTCT---T-G-T--TGAGCCG

>Synechococcus_sp_WH_8102

AT-GGC-TGA-AG-A-ACGTGGTCGGGTTGTGGCTGCAGCCACCGATGGAGCCTGCAG-T
GGCAACCCGGGTCCAGGCGGTTGGGGCGCTTTGCTGCG--ATTTGAA-GATGGCA-----
-----GT--GTTGAGGAATTCGGCGGCCATGACCCGGCAACGACCAACAACCGCATGGA
ACTGC-AGGCGGCGCTGGAGTTGCTGCAGCGCCTGAAACAA--CT--G-----CCA--C-
GCCATC-CCG-ATCTGACGATTCGCACCGACAGCAAGTACCTGATCGATGGGCTCGGCTC
---C---TGGATGAAGGGCTGGAAGCGCAAGGGTTGGAAGACAGCGGCTGGCAAGCCTGT
GCTCAACCAGGATCTCTGGAAAGCCCTGGATGCAGC-T----C-G-G-CTGGA-----
-----TGATGTGCCGCTGGCCTACGTGAAAGGTCACAGCGGTGATC-CGG
ACAACGA--CCGTGTGACCGGATCG---CAGT-TGCTTT---TTC-CCAGGG-----
-----GCGTTCGT-T--GG----AT---C-T--CGGCC-----AG
---CCAGCC--GT--T---G--AGAG-----C----GA-----C----G-
-----TTGCGG-----C-----
-----AGTC--GATGT---C-G-A--TGCTGAT

>Synechococcus_sp_CC9605

AT-GGC-GGA-AG-G-ACGGGGACGGGTCGTGGCTGCAGCAACGGATGGTGCCTGCAG-C
GGCAATCCAGGTCCGGGGGGTTGGGGCGCGTTGTTGCG--TTTCGAA-GATGGCA-----
-----GC--GTTGAGGAATTCGGTGGCCATGCGCCCGACACTACCAACAACCGCATGGA
ACTGC-AGGCGGCATTGGAGGTGCTGCAACGGCTGGAGCAG--CT--G-----CCC--C-
GTCATC-CGG-ATCTCACCCCTGCGCACTGACAGCAAATATCTGATCGATGGCCTGGGCTC
---C---TGGATCAAAGGCTGGAAGCGCAAAGGTTGGAAAACGGCTGCTGGCAAGCCCGT
GCTCAACCAGGACCTCTGGAAGGCCCTGGATGCTGC-C----C-G-T-CTGGA-----
-----TGACGTCCCCCTGAGCTACGTCAAAGGCCACAGCGGTGATC-CCG
ACAACGA--ACGGGTGGATCGCATCG---CCGT-GGCGTA---CTC-CCATAA-----
-----CG---CTC---AG----CC---G-C--CCTTG-----G-
C--TCTGAA--GC--A---G--GGATC-----G--T----C----G-
-----GAGCTG-----C-----CCT-----
-----C-CTCC--AAGGC---C-G-C--ACCAGGC

>Prochlorococcus_sp_W11

AT-GAA-TAG-TA-G-TCAAAATATTGCAATTGAAGCTGCTACTGATGGAGCCTGCAG-T
GGTAACCCTGGTAACGGAGGATGGGGAGGTTTAATTTT--ATTTGAG-GATGGAA-----
-----GT--GAATTAGAAATTGGAGGTTTTGAAAAAATACTACCAATAATAGGATGGA
ACTTA-TGGCTGCCATCAAACTCTAGAAAAATTAAACAA---T--T-----CAA--A-
CTAAAAAAA-ATTTTAAATTAAGAACAGATAGTAAATACTTAATTGATGGCTATAGTAA
---T---TGGATATATAATTGGAAGAAAAACGGTTGGAAAACAAGTACTGGTAAATTAGT
TCAGAATCTAGATTTATGGCAAAGATTGATGGATT-A----A-G-A-ATTAA-----
-----TGAAGTAAGCATGGAATTTGTAAAGGGCATAGTGGAGATA-AAT
ATAATGA--AAGAGTAGATCTAATAG---CAACTAATTAT---A-G-TAAAG-----
-----TA---T-----T----TATAAA---G---T-----AGAAAAA
AAGCG-A-----CAAG---A--AAAT-----ATT--G-----ATCAAC-
-----TTGATT-----T-----
-----GGCA--G-----C-A-C--CCCAAGA

>Prochlorococcus_sp_W8

AT-GGA-TAT-CA-A-AAATAACATTGCAATCGAAGCTGCTACTGATGGAGCTTGTAG-T
GGTAATCCTGGGAATGGTGGATGGGGTGGCTTAATAAT--TTTTTCT-GATGGCA-----
-----GT--GAAATTGAAATTGGCGGTGCTGCAGAAAACACTACCAATAATAGAATGGA
ACTTA-TTGCAGCAATAAAAACAATTGAAAACTTAAAAAT---T--T-----CAG--A-
TTAAAGGAAA-ATTTTAAATTAAGAACTGATAGTAAATATCTTATAGAAGGTTATACAAA

---T---TGGATTAAAACTGGAAAAAAAAATGGTTGGAAAACAAGTAATGGAAAATCTGT
ACAAAATTTAGATTTATGGCAACAAATTGACTCTTT-A---A-G-A-ATTCC-----
-----AGGTTTAACTATGGAATTTGTAAAAGGTCATAGTGGAGATC-CTT
CTAATGA--AAGAGTTGATTTTATTG---CAACAAACTAC---A-G-CAAAGG-----
-----AA---TAAAT-AT---CG---A---C---C-----AG
AATTC-A-----AAAA---A--AAAGA-----CAT-----TTAATT-
-----TAGATA-----A-----
-----AAAA--GAAAT---T-G-C--GCCAGAA
>Prochlorococcus_sp_MIT9515
AT-GAA-TAG-TA-A-ACAAATATTGCTATTGAAGCTGCAACAGATGGTGCTTGCAG-T
GGAAACCCGGGACCTGGTGGATGGGGAGGTTTAATAAT--CTTCATT-GACAATA-----
-----GC--GAAATAGAAATAGGTGGTTCAGAGCAGAATACTACTAATAATCGAATGGA
ACTCA-CTGCAGCAATAAAAACTCTCGAAAAGTTAAAAAAT--TA--T-----CAA--T-
-TAAAAGAAA-ACTTTAAATTAAGAACAGATAGTAAATATGTTATTGAAGGTTATACAAA
---A---TGGATCAGTAATTGGAAGAAAAATGGATGGAAAACGAGTGCAGGAAAATCTGT
ACAGAACTTAGATTTATGGCAAAAAATAGATGAATT-A---A-G-A-ATACA-----
-----AGGCTTAGAGATGGAATTCGTAAAAGGTCATAGCGGGGATA-AAC
AGAATGA--GAGAGTTGATTTAATTG---CCACAAGCTAT---A-G-TAAGGG-----
-----AA---T-----T---CC---A---C---T-----AGATAGT
GAAGT-T-----AAAA---A--ATTA-----T--CTT-----CAGAAA-
-----TTGATT-----C-----
-----AAAA--GAAAC---C-G-G--ATCTTCG

Supplementary Alignment 3

Aligned homologs of the *Trichodesmium erythraeum* twintron host gene *Tery_4732*. These alignments were done with ProbCons (see Methods) and used to build the Bayesian tree shown in Figure 2B.

```
>Trichodesmium_erythraeum_IMS101
ATG--A--A-TGA-----A-----ACAACTAAC-----
GA---A-----TTA-----GTC-----GAA-AATA--TGCCTAG
GGAGCAGAGGGTT--GGGCAACTGCGTAACTTAATTGAAACCCTACATATAGCCGATGAA
GTGGCCAATAAAGGCTATCTAATCACTAGCTCTGAACTGGCAGATTTAATGGATATCAAT
GCCAGTGCAGTTACTAGTCGTGGAGATCATTGGGTATGGCGGAACTGGGTTGTATCTAGA
GTCCGTCGTGAAGGGAACCAAAGT-----GCGATTCTGTTT-----
-----C-----
-----T-----
-----GCTAA
>Cyanothece_sp_PCC_7425
ATG--T--C-TGA-----GTTGCAAGAACTGATCGCTACCGTC-----
AC---G-----ACC-----GATTCC---GAGGAT-CACC--TCAGTCG
CACTCTCCGGGTC--GGACAGTTGAAGAATTTAATTGAAACCCTGCGGATTGCGGAGGAA
ATTTCTAAACAGCGCTATTTAATTAGTAGTTCCGAACTGGCAGATCTGATGGATGTCAAT
GCCAGCGCCGTAACCAGTCGGGGGGATGAATGGGTCTGGCGGAACTGGGTTGTTTCCAGG
GTGCGGCGGGAAGGGAACCAGATTCTCTGGCAATTGGAACGGGT-----
-----G-----
-----G-----
-----ATTAG
>Cyanobacterium_sp_UCYN-A2
ATG--A--T-TGA-----C-----GCAACTGTC-----
AA---T-----ACA-----ATTGAA---GAATTT-GATC--TTTCTCG
TGAAGAGAGAGTC--ATTGAAATTCAAAACCTCATTGAAACCTTGCGTATTGCTGATGAA
ATAGCTGAGAAAGGCTATCTAATTACCAGTTCTGAATTAGCTGATTTAATGGATATTAAT
GCCAGTGCTGTTACAAGTCGTGGCGATCACTGGTCATGGCGAAACTGGGTAGTTTCTAGA
GTTCTGTAGAGAAGGAAACCAAATTCTCTGGCAACTAGAAAAGAT-----
-----T-----
-----G-----
-----ATTAA
>Synechocystis_sp_PCC_6714
ATG--G--A-AAA-----T-----ATTATCAAT-----
AA---T-----GTG-----GAA-----GAGTAT-CTGT--TATCGAG
GGAAGAACGGATC--CTCCAGGTGCGTAAATTTAGTGGAACCCTACGGGCCGCTGAAGAA
ATTGCCAAGCAGGGCTACTTAATTACCAGTTCAGAGTTGGCAGATTTAATGGATATTAAC
CCCAGTGCTGTCACTAGCCGGGGAGATCATTGGTCTTGGCGCAATTGGGTGGTGTCTCGG
GTGAGAAGGGAGGGGAACCAAATTCTCTGGCAGTTGGAGCGGGT-----
-----G-----
-----G-----
-----ATTAG
>Lyngbya_aestuarii_BL_J
-----
-----
-----TTGCATATTGCTGATGAA
ATTGCAAGTAAGGGTTACTTGATCAGTAGTTCGGAATTGGCTGATCTGATGGATGTCAAT
GCTAGTGCTGTCACCAGTCGCGGTGATAACTGGGCTTGGCGGAAATTGGGAAGTTTCCAGA
```


GTTTCGCCGTGAGGGAAATCAAATCCTTTGGCAATTGGAACGGGT-----
-----A-----
-----G-----
-----ACTAA
>Synechococcus_sp_NKBG15041c
ATG--A--C-TGA-----A-----ACCCAGAGT-----
TT---T-----GTTT-----CT---GACGAT---GCCACCA-CCG--ATTCTCA
CCATGAGCGGGTG--CAACGGTTAACCGATGTCATCGAAAGCCTCCGCATTGCCGACGAA
ATTGCAGAAAATGGCTATTTTCATCAGTAGTTCTGAGCTGGCAGATTTGATGGATGTTAAC
GCCAGTGCTGTCACCAGTCGCGGCCATGAGTGGACATGGCGCAATTGGATTGTTTCCCGT
GTGAAACGTGAAGGCAATCAAATTCTCTGGCAACTAGACCGCAT-----
-----CATTG-----GTGGTGA-----
-----TG---AGACGACGG-----
-----CCTAA
>Cyanobacterium_UCYN-A
ATG--A--T-TGA-----C-----GCAGCTGTC-----
AA---T-----ACA-----ATTGAA---GAATTT-GATC--TTTCTCG
TGAAGAGAGAGTT--ATTGAAATTCAAATCTCATTGAAACCTTGCGTATTGCTGATGAA
ATAGCTGAAAAAGGCTATTTAATTACTAGTTCTGAATTAGCCGATTTAATGGATATTAAT
GCTAGTGCTGTTACGAGTCGTGGTGATCATTGGTCATGGCGAAACTGGGTAGTTTCTAGA
GTTTCGTAGAGAAGGAAACCAAATTCTTTGGCAACTAGAGAAGAT-----
-----T-----
-----G-----
-----ATTAA
>Synechocystis_sp_PCC_6803
ATG--G--A-AAA-----T-----ATTGTCAAT-----
AA---T-----GTG-----GAA-----GAGTAT-CTGT--TATCGAG
GGAAGAACGGATT--CTCCAGGTGCGTAACTTGGTAGAAACCTCCGGGCCGCGGAAGAA
ATTGCCAAGCAGGGGTATTTAATTACCAGTTCGGAGTTGGCTGACTTGATGGATATTAAC
CCCAGTGCCGTCACCAGTCGGGGGGATCATTGGTCATGGCGTAACTGGGTCTGTTTCCAGG
GTGAGAAGGGAAGGGAACCAAATCCTCTGGCAGTTGGAGCGGGT-----
-----A-----
-----G-----
-----ACTAA
>Thermosynechococcus_elongatus_BP-1
ATG--G--T-GCA-----A-----CCATCCTTGGGGCTTGATGGTGAC
GG---TGAATGGTTGG-----GGG-----GAA---AATC--ACCCAG
TGGTCTGAGCGCCGTGCCCGCCTTGAAGAACTCATTGTGACACTGCGAATTGCCGATGAA
TTGGCGGCTGGTAAATACTTGATTACCAGTGGCGAATTGGCAGAATTGATGGATGTCCAC
GCCAGTGCCGTCACTAGTCGCGGTGAAGAATGGGTATGGCGCAACTGGCAGATCTCGCGG
GTACGCCGTGAGAGTAACCAAATTCTCTGGCAATTGGAACGCAT-----
-----T-----
-----GATGTGGC---CAAAGAGCAGTCCTCAGAATTA
CCCCCGACGTAG
>Leptolyngbya_sp_PCC_7104
ATG--G--C-AGA-----C-----TCTGGAAC-----
TCTCCC-----ATT-----AGCGAAAACTTTGAT-TCTT--TACCCCG
AGAGGTCAGGGTT--GATAGCCTCCGCAACGTGCTCGAAACCTGCAAATTGCCGACGAA
ATTGCCAAACAGGGCTATTTAATCACCAGTTCAGAACTGGCCGATTTAATGGATGTGAAC
GCCAGTGCTGTCACCAGCCGAGGAGAATTTTGGGCCTGGCGAAACTGGTCTGGTGTACCGG
GTGCGGCGAGAAGGGAACCAAATTCTTTGGCAAATTGAGCGCAT-----
-----C-----
-----G-----
-----ATTAG
>Lyngbya_sp_CCY_8106

ATG--A--A-CGA-----C-----ATAACGAAT-----
GA---T-----GCA-----CAA-----GAA---GTGT--TACCCCG
TGAGGTTAGGGTT--AACCAACTTCGTAATTTGATAGAAACCTTGCATATTGCTGATGAA
ATTGCCAGCAAGGGTACTTGATCAGTAGTTCTGAATTGGCTGATCTGATGGATGTCAAT
GCTAGTGCTGTCACCAGTCGCGGTGATAACTGGGCTTGGCGGAATTGGGAAGTCTCTAGA
GTTTCGCCGTGAGGGAAATCAAATCCTTTGGCAATTGGAACGGGT-----
-----A-----
-----G-----
-----ACTAA
>Cyanothece_sp_CCY_0110
ATG--A--T-TGA-----C-----TCATCTGTC-----
AA---T-----ACA-----ATTGAA---GAATTT-GATC--TCTCCCG
TGAAGAGAGAGTC--ATCGAAATTCAAACCTCATTGAAACCCTACGCATTGCTGATGAA
ATCGCTGATAAGGGTACTTAATCACTAGCTCAGAATTAGCGGATTTAATGGATATAAAT
GCCAGTGCTGTCACCAGTCGGGGTGATCATTGGTCTTGGCGTAACTGGGTTGTTTCACGG
GTTTCGCCGAGAAGGAAATCAAATTCCTTTGGCAACTCGAAAAAAT-----
-----T-----
-----G-----
-----ATTAA
>Arthrospira_sp_PCC_8005
ATG--A--C-TGA-----G-----CCAATCAAC-----
GA---G-----TCC-----AAT-----GAG-AGTT--TACCCCG
GGAATTTTCGAGTC--AATCAACTTAGGAATCTCATAGAAACCCTGCATATTGCCGACGAA
ATTGCTAGTAAAGGCTATTTGATCAGTAGTTCAGAACTTGCTGATTTAATGGATATTAAC
GCTAGTGCGGTCACCAGTCGCGGCGATAACTGGGCGTGGCGAAATTGGGAAGTTTCCCGG
GTTAGACGGGAAGGCAATCAAATCCTCTGGCAATTGGAGAGGGT-----
-----T-----
-----G-----
-----ATTAG
>Planktothrix_NIVA-CYA_98
ATG--A--C-TGA-----C-----TTAATTAAT-----
GA---T-----TCA-----CTA-----GAA-AACC--TACCGCG
CGAGGTTAGGGTG--AACCAACTTCGCAATTTGATCGAGACCCTGCATATTGCTGATGAA
ATTGCTACCAAGGGCTACTTAATCAGTAGTTCTGAACTGGCCGATCTCATGGATATTAAT
GCCAGTGCTGTCACCAGTCGAGGGGATAACTGGGCATGGCGCAATTGGGAAGTATCCAGA
GTGCGGCGAGAAGGTAATCAAATTCCTCTGGCAATTAGAACGAGT-----
-----T-----
-----A-----
-----ACTAG
>Oscillatoria_acuminata_PCC_6304
ATG--G--A-CAC-----A-----ACTGCCCATTCCTACTA-----CT
GA---A-----TCC-----CTT-----GAC-AACT--TACCCCG
TGAGGTGAGGGTC--ACTCAACTCCAGTACCTGATGCAAACCCTGGAAATCGCCGATCGC
ATCGCCACCAAAGGCTATCTGATCACGAGTTCCGAAGTACCGACCTGATGGATGTCAAT
GCCAGCGCCGTCACCAGTCGCGGCGACAACCTGGCCTTGGCGTAACTGGATCGTCTCCCGG
GTGCGACGAGAAGGAAACCAAATTCTCTGGCAACTAGAACGAAT-----
-----T-----
-----GATTTATTTAACGGATCCGGTGAGGATGATGGG
GCCGGAGAATAA
>Crocospaera_watsonii_WH_8501
ATG--A--T-TGA-----C-----TCATCTGTC-----
AA---T-----ACA-----ATTGAAGAATTC----GATC--TCTCCCG
TGAAGAGAGAGTC--ATTGAAATTCAAACCTCATTGAAACCCTACGGGTTGCTGAGGAG
ATCGCTAGTAAAGGTTACTTAATTACTAGCTCAGAATTAGCTGATTTAATGGACATCAAT
GCCAGTGCACTCACCAGTCGTGGTGATCATTGGTCTTGGCGTAACTGGGTTGTTTCGAGG

GTTCGCCGAGAAGGAAACCAAATTCTTTGGCAACTGGAAAAAAT-----
-----C-----
-----G-----
-----ATTAA
>Acaryochloris_marina_MBIC11017
CTC--C--C-GAAGTGAATCA-----ACATCCT---C-----CC
GA---T-----CAT-----GCA-----GAA---CATC--CTGCCAG
TGAATCGAGGATC--TCTCATCTGCAGGATTTAGTCGAAACGTTGGAAATTGCCGATCAC
ATTGCGAAAGAGGGTTATCTAATCAGCAGTTCTGAATTAGCAGATCTAATGGCGGTCCAC
GCCAACGCGGTCACTAGCAGAGGGGATGAATGGGTGTGGCGCAATTGGCTAATCTCCCGT
GTCCGTCGCGGAAGGCAACCAGATTCTGTGGCAGTTAGAGCGTCT-----
-----T-----
-----AAC-----
-----TCCTAG
>Synechococcus_sp_PCC_7002
ATG--A--C-AGA-----T-----CTCACCAGC-----
CT---T-----G-----CC---CCTGAA---GACATTA-CCC--TGCCCTA
CCATGAACGAGTT--AAGCGCTTAACAGACCTCATCGAAAGTCTCCGTATTGCCGATGAA
ATTGCGGAAAATGGTTATTTTATTAGTAGTTCTGAGTTGGCAGATCTAATGGATGTGAAT
GCCAGTGCCGTCACCAGTCGGGGCCATGAATGGACTTGGCGCAATTGGATTGTCTCCCGG
GTCAAACGAGAAGGTAATCAGATCCTTTGGCAACTCGACCGCAT-----
-----TATC-----GGTGA-----
-----TG---AGACGACGA-----
-----CCTAA
>Cylindrospermopsis_raciborskii_CS-505
ACT--A--T-GGA-----A-----ACATCTACA-----
GG---A-----ATAT-----CT---GTGGTGACACAGGAA-GAAA--AACCCCC
GGAAATGAGAATA--TTACAAATCCGTAATCTCATGGAAACCCTAAAGATTGCGGACGAA
ATTGCCCAGAAAGGTTACTTAATCACCAGTTCCGAAGTAGCAGATTTAATGGATGTTTCAT
GCCAGTGCAAGTCAAGTCGCGGTGATCAGTGGAGATGGCGTAAGTGGATAGTGTACGGA
GTCAGACGAGAAGGGAATCAGATTTTGTGGGAGCTGGAAAGGGG-----
-----CGATC-----AAGTGGA-----
-----AA---AGGATGATA-----
-----ATTAA
>Raphidiopsis_brookii_D9
ACT--A--T-GGA-----A-----CCATCTACA-----
GG---A-----ATAACAGGAATATCT---GTGGTGACACAGGAA-GAAA--AACCCCC
GGAAATGAGAATC--TTACAAATCCGTAATCTCATGGAAACCCTAAAGATTGCAGACGAA
ATTGCCCAGAAAGGTTACTTAATCACCAGTTCTGAACTAGCAGATTTAATGGATGTTTCAT
GCCAGTGCCGTGACAAGTCGCGGTGATCAATGGCGATGGCGAAACTGGATAGTGTACGG
GTTAGACGTGAGGGGAATCAGATTTTGTGGGAGCTGGAAAGGGG-----
-----CGATC-----AAGTGGA-----
-----AA---AGGATGATA-----
-----ATTAA
>Halotheca_sp_PCC_7418
ATATGC--AAGAACTAGAGGA-----ACGGACA-----ATC-----
GA---G-----GAT-----TCT-----CGCCTGGTGCAAG
--GCGATCGGATG--GAAGGACTGAATCATTTAATTGAATCCCTCAGAATTGCCGATGAA
ATTGCCGAAAAAGGCTATCTCATTAGTAGTTCTGAACTTGGCGATTTAATGAATGTTAAC
GCCAGCGCCGTTACCAGTAGAGGCAATCACTGGTCTTGGCGCAATTGGGTCGTTTCACGA
GTTTCGACGGGAGGGAAACCAAATTCTCTGGCAACTAGAACGAGC-----
-----A-----
-----GAA-----
-----GGCTAA
>Leptolyngbya_sp_PCC_7376

ATG--A--C-TGA-----C-----TCCGCTACA-----
AA---C-----T-----CT---ACAGAAGAACTTG----TTT--TATCCCG
TCAAGAGCGCGTT--CACCAGCTATCTGCCTTAATTGAGAGCCTCCGTATCGCCGATGAA
ATCGCAGAGAAAGGCTACTTTATCAGTAGCTCCGAACCTTGCTGATCTCATGGATGTAAAT
GCCAGTGCTGTAACTAGCCGTGGCCACGAATGGACATGGCGTAATTGGATCGTTTCCCGT
GTGAAGCGTGAAGGTAACCAAATTTTGTGGCAACTAGACCGCAT-----
-----TATG-----GGCGA-----
-----TG---CCGAGGCTT-----
-----CCTAA
>Calothrix_sp_PCC_7103
ATG--A--T-AGA-----T-----GCTTCTATCTCCTATGCGAGCAAC
AA---T-----ACT-----AAT-----GAA-GAGA--TACCACG
CGAAATGCGGGTA--ACGCAACTTCGCAACCTCGTTGATACCCTCCGTATCGCCGATGAA
ATTGCTGATAAAGGTTATTTAATAACCAGTTCTGAACTAGCTGATTTAATGGATGTTTCAT
GCCAGCGCTGTTACCAGTCGTGGAGATCAGTGGCGCTGGCGAAATTGGGTTGTCTCTCGC
GTTCGTAGAGAAGGTAATCAAATACTTTGGGAACTCGAACGTGGAGATAAAATTCAGGAA
GAGAGCAAT-----
-----G-----
-----AGTAA
>Microcoleus_sp_PCC_7113
ATG--A--T-GGA-----T-----TCACCTGTT-----
TC---T-----TCCC-----CTGCTGGTGCA---ATTGAA-GACT--TACCTCG
CGAGGTGAGAGTT--GCTCAACTCCGTAACCTGGTGGAACTCTCCACATCGCCGATGAG
ATAGCCCGCCAAGGCTACCTGATTAGCAGTTCTGAGCTGGCTGACTTAATGGATGTTAAT
GCCAGTGCCGTGACTAGTCGCGGCGACCACTGGCCTTGGCGAAATTGGATTGTGTCACGA
GTCCGACGTGAAGGCAATCAAATTCTTTGGCAACTTGAACGAGT-----
-----GGATC-----AAGTGAATACAG-----
-----CAG---ATGATGAAG-----
-----GTTAA
>Rivularia_sp_PCC_7116
ACT--A--T-GGA-----A-----TCATCTAGC-----
GC---T-----TCTG-----CTACCTTGCAAAACGAA----GATT--TACCACG
CGAAACAAGAGTC--TCGCAACTCCGCAGTTTAGTTGAACTCTTCGCATTGCTGACGAA
ATAGCCGAAAAAGGTTACTTAATCACAAGTTCCGAATTAGCGGACTTAATGAACGTCCAT
GCTAGTGCTGTTACTAGCCGTGGAGACCAATGGCGTTGGAGAACTGGATAGTTTCCCGA
GTACGTCGCGAGGGTAATCAAATACTTTGGGAACTCGAACGCGC-----
-----CGATA-----AACTTGAC---G-----
-----ATG---ATGAAGGAG-----
-----AGTAA
>Acaryochloris_sp_CCME_5410
CTC--C--C-GAAGTGAATCA-----ACATCCT---C-----CC
GA---T-----CAT-----GCA-----GAA-CATC--CTGCCAG
TGAATCAAGGATC--TCTCATCTGCAGGATTTAGTCGAAACGTTGGAAATTGCCGATCAC
ATTGCGAAAGAGGGTTATCTAATCAGCAGTTCTGAATTAGCAGATCTAATGGCGGTCCAC
GCCAACGCGGTCACTAGCAGAGGGGATGAATGGGTGTGGCGCAATTGGCTAATCTCCCGT
GTCCGTCGCGGAAGGCAACCAGATTTTGTGGCAGTTAGAGCGTCT-----
-----T-----
-----AAC-----
-----TCCTAA
>Synechocystis_sp_PCC_7509
AT---A--T-GGA-----T-----TCACCAGCC-----
CC---A-----ATAA-----ATACTCATGC-----AGAC-GAAT--TACCCCG
CGAGGTCAAAGTG--GGACAGTTGCGCCACTTGATTGATACGCTGCGTATAGCGGAAGAA
ATTGCGACCAAAGGCTACCTAATTACAAGTTCCGAACCTTGCCGATTTGATGGATGTCCAC
GCCAGCGCCGTCACCAGTCGCGGAGACGAGTGGCGCTGGCGCAACTGGATTGTGTCGAGA

GTAAGACGAGAAGGCAATCAAATTCTGTGGGAATTAGAACGAGG-----
-----AGACG-----ATGTTAGT---TCGGCAACT
AGCTATAACTCCGACC---CCGCCACTG-----
-----AGTAG
>Richelia_intracellularis_HH01
ATG--A--T-TGA-----C-----TCATCCAAT-----
GC---C-----ACTG-----CTTCACTAAAT---AGTGAG-ACGA--TTTCTAG
TGAATTAAGGATG--GTGCACATCCGTAAACCTATTGGAAACCCTTCATGTCGCTGACGAA
ATTGCAGCTAAAGGTTATTTGATAACCAGTTCCGAACTGGCAGACTTAATGGGTGTTTCAT
GCCAGTGCTGTTACTAGTAGGGGAGATGAATGGCGCTGGCGTAACCTGGGTAGTTTCACGT
GTTAGAAGAGAAGGTAATCAAATTCTTTGGGAAGTAGAGCGGGG-----
-----TGATC-----TATTTGAT---G-----
-----AGG---AAGATGAAA-----
-----ATTAA
>Nodularia_spumigena_CCY9414
GCT--A--T-GGA-----A-----CCATCCAGC-----
GC---A-----CAAG-----CT---GCGGCCGCAAATGAA-GAAA--AACTGTC
TACAACCAGGGTA--AAGCAACTCAAAAACCTTGCTGGGAGACTCTGCGTATAGCTGAGGAA
ATTTTCGGAAAAAGGCTACTTAATCACCAGTTCTGAACTAGCAGACTTGATGGATGTCCAC
GCCAGTGCTGTAACTAGTCGGGGAGACGAGTGGCGCTGGCGGAACCTGGATAGTGTACGCG
GTAAGACGCGAAGGAAATCAAATTCTCTGGGAACCTGGAACGAGG-----
-----CGATTGCGGTTTCGCGCAGTGCCGGAGGCGATCGCCTAGA-----
-----AACCGAAGACGAAA-----
-----TATGA
>Nostoc_sp_PCC_7120
ACT--A--T-GGA-----A-----CCATCTACC-----
GC---A-----CACG-----TT---GCGGCTACAATCGAG-GAAA--ACCCGCC
GGAAAAGCGGGTG--GAACAACTCCGCAACTTAGTAGAACTCTGCATATTGCAGACGAA
ATTGCGGCGAAAGGTTATTTAATCACTAGTTCTGAGTTGGCAGACTTAATGGATGTCCAC
GCCAGTGCTGTTACTAGTAGGGGAGATCAGTGGCGCTGGCGCAACTGGATTGTTTCACGG
GTACGCCGTGAGGGTAATCAAATTCTCTGGGAATTAGAACGAGG-----
-----CGATC-----GCTTGGG-----
-----GG---GTGAAGATG-----
-----AATAA
>Synechococcus_elongatus_PCC_6301
GTT--CGGA-TGC-----C-----TGAAGTTGT-----
GA---C-----CCA-----AGAGTG---G-----AGCAG--TAACCCT
CGCGATCGCAGT---CAACGCCTGCGCGATGTGCTGGACTGCTTGCAAATTGCTGCAGCA
ATTGCTGAGCAGCAATTTTTGATCACGAGTTTCAGAGCTGGCTGACCTGATGGATGTCAAC
GCCAGTGCGGTCACCAGTCGGGGCGAAGAGTGGATTTGGCGCAACTGGCGGGTGAAGCGC
GATCGCCGTGAAGGCAATCAAATCCTCTGGCAGCTCGAGCGCAT-----
-----T-----
-----C-----
-----ATTGA

Supplementary Alignment 4 Alignment of the *T. erythraeum* IMS101 twintron and its host gene exons with the respective sequences from metagenomic datasets. TrichoMGDRRAFT_c100672 originates from Hawaii (Pacific Ocean), the TCCM contigs originate from Bermuda (Atlantic Ocean). Identical residues in all three sequences are represented by a dot. The black box and the sequence below it show the consensus of the alignment. The transcription start site (TSS) of the pre-transcript in *T. erythraeum* IMS101 is indicated by an arrow, the mRNA exons (green) and intron IEP ORFs (red and blue for the inner and host intron, respectively) are boxed. The starts of the host intron (blue) and the inner intron (red) are indicated by open boxes. Note that interestingly, the intron sequence is generally better conserved than the two IEP ORF sequences.

