

Comparative genomic analysis of the marine cyanobacterium *Acaryochloris marina* MBIC10699 reveals the impact of phycobiliprotein reacquisition and the diversity of *Acaryochloris* plasmids.

Haruki Yamamoto^{1*}, Kazuma Uesaka², Yuki Tsuzuki¹, Hisanori Yamakawa¹, Shigeru Itoh³, and Yuichi Fujita¹

Supplementary materials

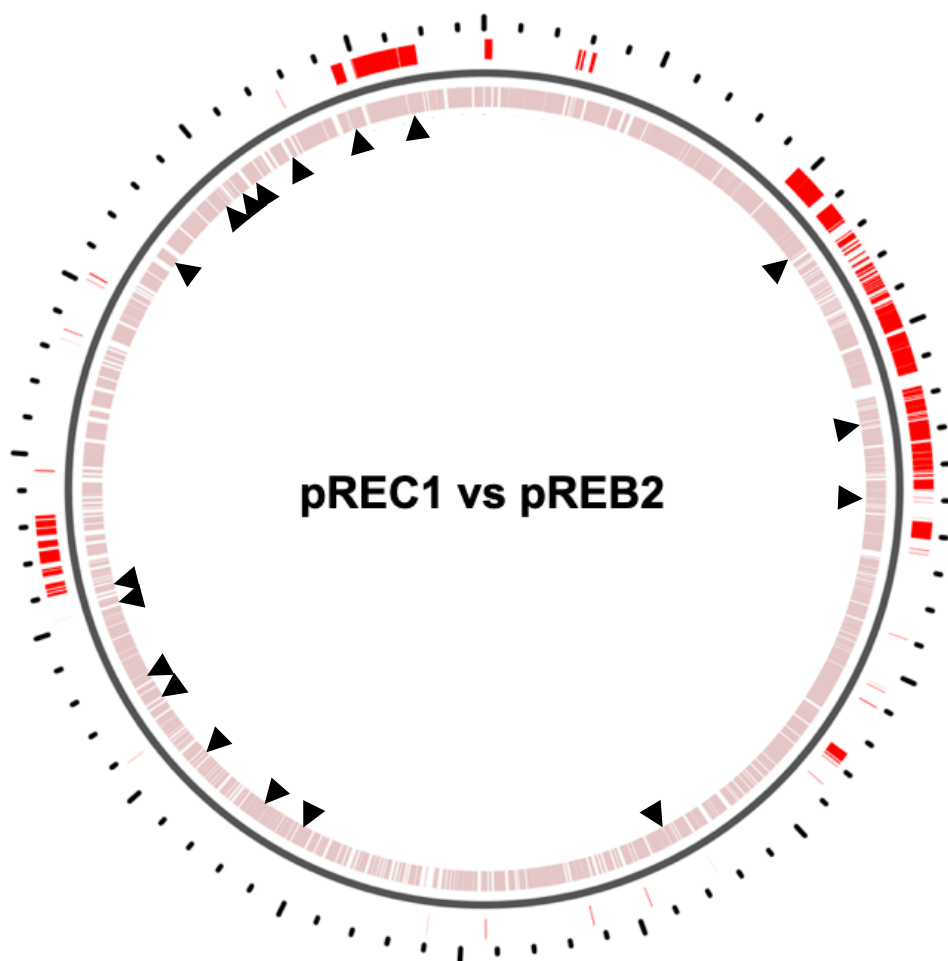


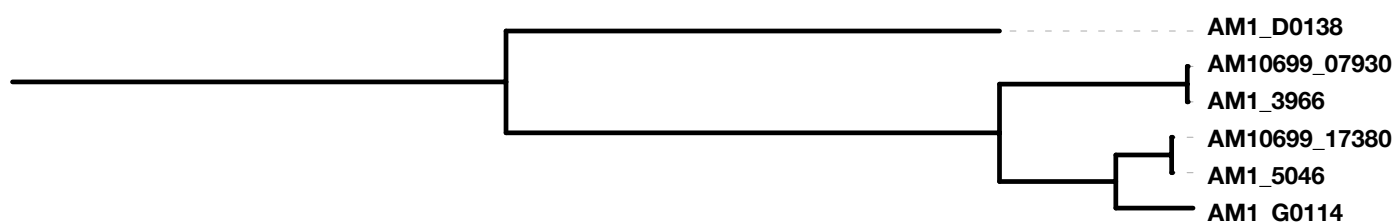
Figure S1 Loci of pREC1-specific genes in the comparison with pREC1 and pREB2.

Two plasmids with significant homology, pREC1 and pREB2, were compared and genes only present in pREC1 were mapped in red on the pREC1 sequence (light purple). Specific genes were extracted by blastn

(e-value cut off $1e^{-10}$, identity cut off 80%). Black triangles show the loci of transposase genes.

A

Tree scale: 0.1



B

AM1_D0138	MKHFRKQLLVGLILTCMGMTFSFQVPVYAGEILEIQPINSVQVSDPQSI--SKESCIPI	59
AM10699_07930	MHSFVRRFGLIGLALATCLSF--WSQPVNALNLQS--GPTIVLADTYRNPIDEKLGTD	57
AM1_3966	MHSFVRRFGLIGLALATCLSF--WSQPVNALNLQS--GPTIVLADTYRNPIDEKLGTD	57
AM10699_17380	MRLFMRRFAVLTLVLASCLGLTG--WLQPANALDIN-----LATTYRNPVDAKLET	51
AM1_5046	MRLFMRRFAVLTLVLASCLGLTG--WLQPANALDIN-----LATTYRNPVDAKLET	51
AM1_G0114	MRLFIRRFALLTLVVASCLSLMG--WLQPVSAIDIQQSNSPGLLAAYRNPVDAKLET	59
	*: * : : : *.: :*: : : : : * . * : : : : * . :	
AM1_D0138	GEKIDLNANALAFKDCPGYYPTLAKKIINNGPFETVDEVLNIRGLNLKQQLLEDKLD	119
AM10699_07930	GQKIDVNNTNVLAFMKYKGLYPTVAGKVVQNAPYSSVEDVLQIDGLTDQKQKVLQDNM	117
AM1_3966	GQKIDVNNTNVLAFMKYKGLYPTVAGKVVQNAPYSSVEDVLQIDGLTDQKQKVLQDNM	117
AM10699_17380	GQKIDLNNTNVIASFQYKGLYPTIAGKVVQNAPYSSVEEVVLNIDGLTDVQKKMLQQLH	111
AM1_5046	GQKIDLNNTNVIASFQYKGLYPTIAGKVVQNAPYSSVEEVVLNIDGLTDVQKKMLQQLH	111
AM1_G0114	GQKIDLNNTNVIASFQYKGLYPTIAGKVVQNAPYSSVEEVVLNIDGLTEPQKKLLQQHMG	119
	*:***:*.:.** . * ***:* *: :*.:.*: :*: ** . ** :*: : : .	
AM1_D0138	FTVSEPKTDLATRMP---P---RPMMR-	140
AM10699_07930	FTISDPDPSLVGGQDRYNPGSYKPFMK-	144
AM1_3966	FTISDPDPSLVGGQDRYNPGSYKPFMK-	144
AM10699_17380	FTITDPDPALVGGQDRYNPGAYYPVRRS	139
AM1_5046	FTITDPDPALVGGQDRYNPGAYYPVRRS	139
AM1_G0114	FTVTDPDPALVGGQDRYNTGAYYPVRRS	147
	*: : : * . * . : : * . :	

Figure S2 Phylogenic analysis of PsbU paralogs conserved in two *Acaryochloris* strains.

A Phylogenetic tree (A) and an alignment (B) based on amino acid sequences of the two PsbUs conserved in *A. marina* MBIC10699 (AM10699_07930, 17380) and the four PsbUs conserved in *A. marina* MBIC11017 (AM1_3966, 5046, D0138, G0114). Plasmid-encoded PsbUs are AM1_D0138 (pREB4) and AM1_G0114 (pREB7).

Table S1 List of genes for the concatenated protein sequences

A core gene search was conducted on 26 *Acaryochloris* and *Cyanothece* sp. PCC7425 and *A. thomasi* RCC1774 species, listing 897 genes that are conserved in more than 90% of their strains.

Table S2 Chromosome specific gene in *A. marina* MBIC10699

Genes encoded in the chromosome of *A. marina* MBIC10699 that are not present in the chromosome of *A. marina* MBIC11017 were extracted. Genes in red indicate specific genes that are not present in *A. marina* MBIC11017 at all, while genes in black indicate that they are present in the plasmids of *A. marina* MBIC11017.

Table S3 Chromosome specific gene in *A. marina* MBIC11017

Genes encoded in the chromosome of *A. marina* MBIC11017 that are not present in the chromosome of *A. marina* MBIC10699 were extracted. Genes in blue indicate specific genes that are not present in *A. marina* MBIC10699 at all, while genes in black indicate that they are present in the plasmids of *A. marina* MBIC10699. * indicates that the same KEGG ID was annotated in multiple genes.

Table S4 Plasmids specific gene in *A. marina* MBIC10699

Genes encoded in the plasmids of *A. marina* MBIC10699 that are not present in the plasmid of *A. marina* MBIC11017 were extracted. Genes in red indicate specific genes that are not present in *A. marina* MBIC11017 at all, while genes in black indicate that they are present in the chromosome of *A. marina* MBIC11017. * indicates that the same KEGG ID was annotated in multiple genes.

Table S5 Plasmids specific gene in *A. marina* MBIC11017

Genes encoded in the plasmids of *A. marina* MBIC11017 that are not present in the plasmids of *A. marina* MBIC10699 were extracted. Genes in blue indicate specific genes that are not present in *A. marina* MBIC10699 at all, while genes in black indicate that they are present in the chromosome of *A. marina* MBIC10699. * indicates that the same KEGG ID was annotated in multiple genes.