Characterization of the complete plastid genome of *Porphyridium purpureum* strain CCMP1328

Guiqi Bi*ab

*Key Laboratory of Marine Genetics and Breeding (OUC), Ministry of Education, Qingdao, China; †College of Marine Life Sciences, Ocean University of China, Qingdao, China

**ABSTRACT**

In this study, the complete plastid genome of *Porphyridium purpureum* strain CCMP1328 was recovered through Illumina sequencing data. This complete plastid genome of *P. purpureum* was 220,483 bp in length and contained a pair of IR regions (4604 bp). The pt genome of *P. purpureum* encoded 234 genes including 199 protein-coding genes, 29 tRNA genes, one tmRNA, and six ribosomal RNA genes in IR regions. The overall GC content of *P. purpureum* cp genome is 30.4%. By phylogenetic analysis using 18S DNA fragments through NJ method, *P. purpureum* strain CCMP1328 was grouped in the *P. purpureum* cluster without further distinction. This complete plastid genomes can be subsequently used for evolution studies of red algae and provide valuable insight into dynamic evolution of group II introns.

---

*Porphyridium cruentum* is a species of red alga in the family Porphyridiophyceae.

The genome sequence of *Porphyridium cruentum* strain CCMP1328 was published in 2013 (Bhattacharya et al. 2013). The organelle genome information of *P. cruentum* strain CCMP1328 is still limited, even though another two plastid genomes from *Porphyridium* species have been available. In this study, the complete plastid genome of *P. purpureum* strain CCMP1328 was recovered through Illumina sequencing data. This complete plastid genomes can be subsequently used for evolution studies of red algae and provide valuable insight into dynamic evolution of group II introns.

The sample of *P. purpureum* strain CCMP1328, collected from Eel Pond, Woods Hole, MA, USA (41.5264°N, 70.67°W), can be obtained from the National Center for Marine Algae and Microbiota, East Boothbay, ME. The raw reads sequenced by Illumina Hiseq 2000 platform were retrieved from NCBI Sequence Read Archive database under Accession number: SRR747671 and SRR747672. After reads quality filtration, the clean reads were assembled by SPAdes 3.6.1 (Bankevich et al. 2012) based on default settings. We used another plastid genome of *Porphyridium cruentum* (AP012987) (Tajima et al. 2014) as a reference sequence to align the contigs and identify gaps. To fill the gap, Price (Ruby et al. 2013) and MitoBim v1.8 (Hahn et al. 2013) were applied and Bandage (Wick et al. 2015) was used to identify the borders of the IR regions. The mean sequencing coverage of this pt genome is 253X and exhibits an obvious positive relationship with GC content. The complete sequence was primarily annotated by Plann (Huang and Cronk 2015) combined with manual correction. All tRNAs were confirmed using the tRNAscan-SE search server (Lowe et al. 1997). Other protein coding genes were verified by BLAST search on the NCBI website (http://blast.ncbi.nlm.nih.gov/), and manual correction for start and stop codons was conducted. The circular plastid genome map was drawn using OrganellarGenomeDRAW (Lohse et al. 2007). This complete plastid genome sequence together with gene annotations were submitted to GenBank under the accession numbers of MF401423.

The plastid genome of *Porphyridium cruentum* strain CCMP1328, with a length of 220,483 bp, has an unusual quadripartite structure compared with other red algae. The whole pt genome contains a two inverted repeat (IR) regions of 4604 bp, in which encodes two rRNA operons. The pt genome possesses 234 genes, including 199 protein coding genes, eight ribosomal RNA genes (four rRNA species), and 29 tRNA genes and one tmRNA. The overall GC content of the pt genome is 30.4%. The genome structure, gene order, GC content are similar to another pt genome of *Porphyridium cruentum* (AP012987), but the number of introns and its locations were different. Significantly, both the number of introns and genome synteny were diverse when comparisons between the two pt genomes of *Porphyridium cruentum* and *Porphyridium sordidum* (NC_031175) (Lee et al. 2016) were made.

For phylogenetic analysis assessing the relationship of this species, we selected 18S fragments from 12 *Porphyridium* species as DNA maker due to the lack of enough *Porphyridium* pt genomes. The sequence alignment was done by MUSCLE aligner (Edgar 2004) in MEGA 6 (Tamura et al. 2013) and the
The tree was constructed by NJ method under Maximum Composite Likelihood method with 1000 bootstrap calculations. As expected, the NJ tree exhibited three unique clusters, but *P. purpureum* strain CCMP1328 was grouped in the *P. purpureum* cluster without further distinction, caused by the shortness of 18S fragment (Figure 1).

**Disclosure statement**

The authors have declared that no competing interests exist.

**References**

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pyribelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19:455–477.

Bhattacharya D, Price DC, Chan CX, Oiu H, Rose N, Ball S, Weber APM, Arias MC, Henrispat B, Coutinho PM, et al. 2013. Genome of the red alga *Porphyridium purpureum*. Nat Commun. 4:1941.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–1797.

Hahn C, Bachmann L, Chevreux B. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. Nucleic Acids Res. 41:e129.

Huang DI, Cronk QCB. 2015. Plann: a command-line application for annotating plastome sequences. Appl Plant Sci. 3:1500026.

Lee JMO, Cho CH, Park SI, Choi JW, Song HS, West JA, Bhattacharya D, Yoon HS. 2016. Parallel evolution of highly conserved plastid genome architecture in red seaweeds and seed plants. BMC Biol. 14:75.

Lohse M, Drechsel O, Bock R. 2007. OrganelarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. Curr Genet. 52:267–274.

Lowe Todd M, Sean R, Eddy 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25:955–964.

Ruby JG, Priya B, DeRisi JL. 2013. PRICE: software for the targeted assembly of components of (Meta) genomic sequence data. G3: Genes Genomes Genet. 3:865–880.

Tajima N, Sato S, Maruyama F, Kurokawa K, Ohta H, Tabata S, Sekine K, Moriyama T, Sato N. 2014. Analysis of the complete plastid genome of the unicellular red alga *Porphyridium purpureum*. J Plant Res. 127:389.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 30:2725–2729.

Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics. 31:3350–3352.