

*Genetic Resources*

## Complete Nucleotide Sequence of the *Porphyra purpurea* Chloroplast Genome

*Michael Reith and Janet Munholland*

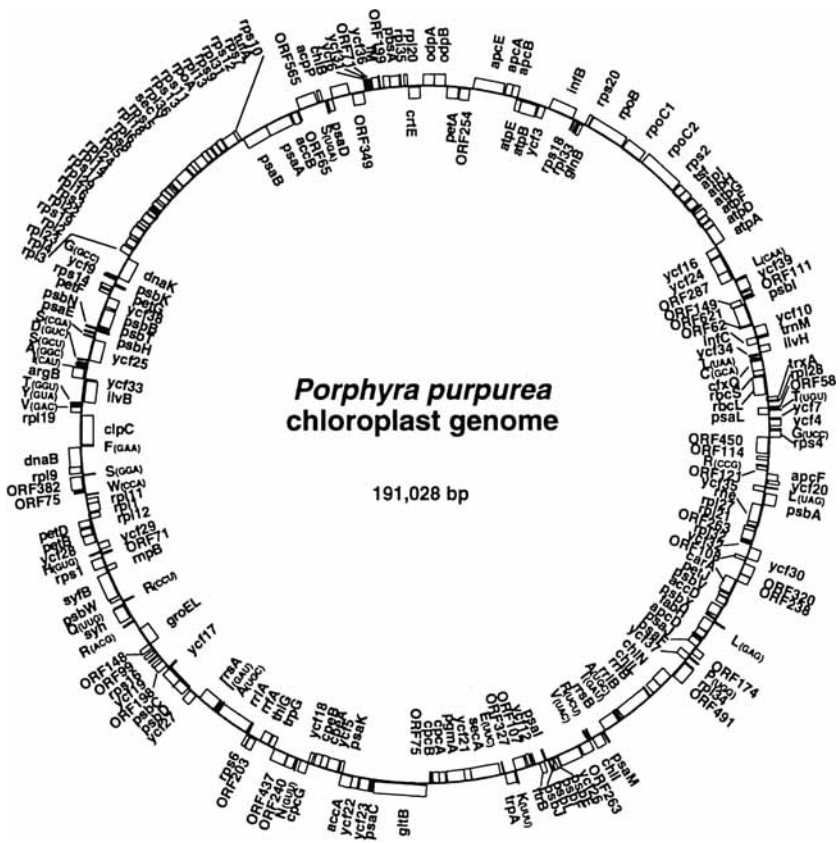
Email: reithm@imb.lan.nrc.ca  
NRC Institute for Marine Biosciences, 1411 Oxford St., Halifax,  
Nova Scotia, B3H 3Z1 Canada

**Key Words:** chloroplast, nucleotide sequence, *Porphyra purpurea*, red alga.

**Abstract:** The complete nucleotide sequence of the chloroplast genome of the red alga *Porphyra purpurea* has been determined (accession number = U38804). The circular genome is 191,028 bp in length and encodes approximately 250 genes.

The complete nucleotide sequence of the chloroplast genome of the red alga *Porphyra purpurea* (Rhodophyceae, Bangiophycidae), Avonport strain, has been determined (accession number = U38804). The circular genome is 191,028 bp in length, making it, to date, the largest chloroplast genome to be completely sequenced (see Fig. 1). Similar to other chloroplast genomes, the *P. purpurea* chromosome has a low G+C content (33 percent) and contains two copies of the ribosomal RNA operon. Unlike other chloroplast genomes, the two rRNA operons are organized as direct, rather than inverted repeats (Reith and Munholland, 1993a,b) and the two copies are not identical, with 41 of 4,820 (0.85 percent) positions differing in the two versions. Only the genes of the ribosomal RNA operon are present in these repeats.

Approximately 251 genes and open reading frames have been identified on the *P. purpurea* chloroplast genome, more than double the number of genes present on chloroplast genomes of land plants. As is the case for land-plant chloroplast genomes, many of these genes encode photosynthetic proteins and proteins involved in gene expression. However, many genes of these two categories that have been transferred to the



**Fig. 1. The *Porphyra purpurea* chloroplast genome.** Genes on the outside of the circle are transcribed in a clockwise direction; those inside the circle are transcribed counterclockwise. Gene names were designated according to recommendations (Hallick, 1989; Hallick and Bairoch, 1994; Hallick and Bottomley, 1983). For a list of genes and gene products, see the table on p. 324.

nucleus in land plants have been retained in the *P. purpurea* chloroplast genome. In addition, there are many more genes encoding proteins involved in biosynthetic functions. Among the more interesting features of the *P. purpurea* chloroplast genome are:

- 53 genes encoding photosynthetic proteins: 11 photosystem I, 16 photosystem II, 8 ATPase, 10 phycobilisome, 4 cytochrome  $b_6/f$  complex, 2 RuBisCo and 2 mobile electron carriers (ferredoxin and cytochrome  $c_{553}$ )

- 47 ribosomal protein genes: 28 large subunit, 19 small subunit
- 35 tRNAs
- 9 additional genes encoding proteins involved in gene expression: 2 initiation factors, 2 tRNA synthetases, 2 elongation factor subunits, a replication helicase subunit, the RNA component of RNase P and an RNase E homolog
- 5 ORFs with probable functions as transcriptional regulatory proteins
- 24 genes for biosynthetic functions, including synthesis of amino acids, fatty acids, pigments and thiamine
- 65 ORF or *ycf* genes, only 19 of which are unique to *P. purpurea*
- no introns and no apparent RNA editing
- no genes (or pseudogenes) for subunits of the NADH dehydrogenase complex

With the recent completion of the chloroplast genome sequences of *P. purpurea*, the glaucocystophyte *Cyanophora paradoxa* (Stirewalt et al., 1995), and the diatom *Odontella sinensis* (Kowallik et al., 1995), algal chloroplast genomes have been found to encode many more genes than those of land plants. These sequences should provide interesting insights into the evolution of chloroplasts.

**Acknowledgments.** We wish to thank Don Bryant and Klaus Kowallik for sharing information prior to publication and Susan Douglas for critically reading the manuscript. This is NRCC publication number 38112.

## References

- Hallick, R.B. 1989. Proposals for the naming of chloroplast genes. II. Update to the nomenclature of genes for thylakoid membrane polypeptides. *Plant Mol. Biol. Repr.* 7:266-275.
- Hallick, R.B. and A. Bairoch. 1994. Proposals for the naming of chloroplast genes. III. Nomenclature for open reading frames encoded in chloroplast genomes. *Plant Mol. Biol. Repr.* 12:529-30.
- Hallick, R.B. and W. Bottomley. 1983. Proposals for the naming of chloroplast genes. *Plant Mol. Biol. Repr.* 1:38-43.
- Kowallik, K.V., B. Stoebe, I. Schaffran, U. Freier. 1995. The chloroplast genome of a chlorophyll *a+c*-containing alga, *Odontella sinensis*. *Plant Mol. Biol. Repr.* 13:336-342.
- Reith, M. and J. Munholland. 1993a. A high-resolution gene map of the chloroplast genome of the red alga *Porphyra purpurea*. *Plant Cell* 5:465-475.
- Reith, M. and J. Munholland. 1993b. The ribosomal RNA repeats are non-identical and directly oriented in the chloroplast genome of the red alga *Porphyra purpurea*. *Curr. Genet.* 24:443-450.
- Stirewalt, V.L., C.B. Michalowski, W. Löffelhardt, H.J. Bohnert and D.A. Bryant. 1995. Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*. *Plant Mol. Biol. Repr.* 13:327-332.