
Chloroplast Genomes of Bryophytes: A Review

Asheesh Shanker

Department of Bioscience and Biotechnology, Banasthali University, Rajasthan, India

Abstract: Since the availability of first complete DNA sequence of chloroplast genome of *Marchantia polymorpha* several studies to determine the structure and organization of chloroplast genomes of various plant groups have been made. However the lineage sampling of bryophytes are poorly represented in complete chloroplast genome sequences (only three liverworts, two mosses and one hornwort). This review presents features of sequenced chloroplast genomes of bryophytes.

Introduction

Chloroplasts are intracellular organelles present in green plants which contain their own autonomously replicating DNA genome to encode a number of components for the process of photosynthesis (Palmer 1985; Sugiura 1989). Since the availability of first complete DNA sequence of chloroplast genome of *Marchantia polymorpha* (Ohyama et al. 1986; Ohyama et al. 1988) several studies to determine the structure and organization of chloroplast genomes have been conducted (Nishiyama et al. 2004; Turmel et al. 2006; Jansen et al. 2007). Consequently the nucleotide sequences of many chloroplast genomes which belong to algae, bryophytes, pteridophytes, gymnosperms and angiosperms have been worked out (NCBI Organelle Genome Resources; 248 chloroplast genomes of Viridiplantae as on 08.09.12). Despite of the fact that the first complete chloroplast genome sequenced was of a bryophyte only a small proportion (6 out of 248) of available chloroplast genomes sequenced belong to bryophytes (<http://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?opt=plastid&taxid=33090>).

Chloroplast genomes are typically represented as circular with an inverted repeat region (containing rRNA and some other genes) dividing the chloroplast genome into large single copy (LSC) and small single copy (SSC) regions. Table 1 presents information of sequenced chloroplast genomes of bryophytes. The present review deals with important features of these chloroplast genomes.

Chloroplast Genomes of Liverworts

Marchantia polymorpha

The first complete chloroplast genome sequence determined belongs to *Marchantia polymorpha* (Ohyama et al. 1986; Ohyama et al. 1988). The chloroplast DNA of this liverwort contains 121024 base pairs (bp) out of which 20116 bp belongs to large inverted repeats (IRA and IRB, each of 10058 bp) separated by a SSC region of 19813 bp and a LSC region of 81095 bp. Considering the

universal genetic code, the computational analysis of sequence identified 103 gene products which are related to protein molecules or RNA. Stable RNA genes for 4 types of ribosomal RNA and 32 types of tRNA were located. Several open reading frames (ORFs) show structural similarities to *Escherichia coli* sequences. Seven ORFs were found to be similar with human mitochondrial NADH dehydrogenase genes. A total of 20 genes (8 tRNA genes and 12 ORFs) were found to contain introns. Throughout the *Marchantia* chloroplast genome only 28.8% G+C content was calculated (Ohyama et al. 1988).

Aneura mirabilis

The complete plastid genome of parasitic liverwort *Aneura mirabilis*, the only known heterotrophic lineage of bryophytes, was sequenced to study the effect of its nonphotosynthetic life history on chloroplast genome organization. The chloroplast genome of *A. mirabilis* contains 108007 bp. The two inverted repeats (each of 8240 bp) separates 77553 bp LSC region and 13974 bp SSC region. Chloroplast genome comparison of *A. mirabilis* with *M. polymorpha* shows absence of all *ndh* genes. Additionally sequences similar to *ndh* genes, *cysA*, *cysT*, *ccsA*, and *ycf3* were found in the form of pseudogenes. Moreover 5 of 15 *psb* genes, 2 of 6 *psa* genes, 2 of 6 *pet* genes were also found as pseudogenes. The remaining open reading frames of *A. mirabilis* were found complete (Wickett et al. 2008).

Ptilidium pulcherrimum

Among bryophytes the chloroplast genome sequence of *Ptilidium pulcherrimum*, a leafy liverwort, is the first plastid genome to be sequenced using next generation sequencing (NGS) technology. The plastid genome of *P. pulcherrimum* was found to be 119007 bp long which contains a total of 122 genes excluding the genes present on second inverted repeat region. Out of 122 genes, 88 genes codes for proteins, 4 encodes rRNAs and 30 codes for tRNAs. Two protein-coding genes (*cysA* and *cysT*) used for sulphate import were found in the form of pseudogenes (Forrest et al. 2011).

Table 1. Information of sequenced chloroplast genomes of bryophytes.

S. No.	Organism Name	*Accession No.	Genome Size (bp)	Reference
	Liverworts			
1.	<i>Aneura mirabilis</i>	NC_010359	108007	Wickett et al. 2008
2.	<i>Marchantia polymorpha</i>	NC_001319	121024	Ohyama et al. 1986
3.	<i>Ptilidium pulcherrimum</i>	NC_015402	119007	Forrest et al. 2011
	Mosses			
4.	<i>Physcomitrella patens</i>	NC_005087	122890	Sugiura et al. 2003
5.	<i>Syntrichia ruralis</i>	NC_012052	122630	Oliver et al. 2010
	Hornworts			
6.	<i>Anthoceros formosae</i>	NC_004543	161162	Kugita et al. 2003

*Complete genome sequence at NCBI will be accessed using this number.

Chloroplast Genomes of Mosses

Physcomitrella patens

The complete nucleotide sequence of moss *Physcomitrella patens* chloroplast genome contains 122890 bp with 83 protein, 31 tRNA and 4 rRNA genes. A sequence similar to tRNA was classified as a pseudogene. The IRA and IRB consists of 9589 bp each. A LSC region of 85212 bp and a SSC region of 18501 bp separate these inverted repeats. The overall GC content determined was 28.5%. Four protein coding genes (rpoA, cysA, cysT and ccsA) present in *Marchantia polymorpha* found to be absent in *P. patens* chloroplast genome (Sugiura et al. 2003).

Syntrichia ruralis

Among land plants *Syntrichia ruralis*, also known as *Tortula ruralis*, is one of the most desiccation tolerant plant and act as a model organism to study plants survival strategies under dehydration stress (Alpert and Oliver 2002; Oliver et al. 1993). Since the role of chloroplast is important to recover the vegetative plant cells from desiccation (Proctor et al. 2007) therefore the chloroplast genome of *S. ruralis* was determined (Oliver et al. 2010). The complete nucleotide sequence of *S. ruralis* contains ~122630 bp. However the exact length of the chloroplast genome remains unknown due to a gap of ~750 bp. Though several attempt was made to sequence this region. In comparison to *Physcomitrella* chloroplast genome, *S. ruralis* chloroplast genome lacks ~71 kb inversion in the LSC region. Moreover it lacks petN gene which is found in all known land plants plastid genomes (Oliver et al. 2010).

Chloroplast Genome of Hornwort

Anthoceros formosae

The complete chloroplast genome sequence of *Anthoceros formosae* consists of 161162 bp. The LSC (107503 bp) and SSC (22171 bp) regions were found to be divided by a pair of IR regions of 15744 bp each. A total of 112 genes were found which includes 76 protein coding, 4 rRNA and 32 tRNA genes. Two genes, matK and rps15, which are very common in the chloroplast genomes of land plants were identified as pseudogenes. The overall GC content calculated was 32.9% (Kugita et al. 2003).

It is evident from this review that earliest land plants are poorly represented in organellar genome database at NCBI with only three liverworts, two mosses and one hornwort have completely sequenced, annotated chloroplast genomes. Given the importance of bryophytes more sequences will be valuable to unravel the mysteries of early land plants evolution.

References

- Alpert P, Oliver MJ. 2002. Drying without Dying. In: Black M, Pritchard HW (Eds.) Desiccation and survival in plants: Drying without Dying. CABI Publishing, Wallingford, 3-43.
- Forrest LL, Wickett NJ, Cox CJ, Goffinet B. 2011. Deep sequencing of *Ptilidium* (Ptilidiaceae) suggests evolutionary stasis in liverwort plastid genome structure. *Plant Ecol Evol* 144: 29-43.

- Jansen RK, Cai Z, Raubeson LA, Daniell H, dePamphilis CW, Leebens-Mack J, Muller KF, Guisinger-Bellian M, Haberle RC, Hansen AK, Chumley TW, Lee SB, Peery R, McNeal JR, Kuehl JV, Boore JL. 2007. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proc Natl Acad Sci USA* 104: 19369-19374.
- Kugita M, Kaneko A, Yamamoto Y, Takeya Y, Matsumoto T, Yoshinaga K. 2003 The complete nucleotide sequence of the hornwort (*Anthoceros formosae*) chloroplast genome: insight into the earliest land plants. *Nucleic Acids Res* 31: 716-721.
- Nishiyama T, Wolf PG, Kugita M, Sinclair RB, Sugita M, Sugiura C, Wakasugi T, Yamada K, Yoshinaga K, Yamaguchi K, Ueda K, Hasebe M. 2004. Chloroplast phylogeny indicates that bryophytes are monophyletic. *Mol Biol Evol* 21: 1813-1819.
- Ohyama K, Fukuzawa H, Kohchi T, Sano T, Sano S, Shirai H, Umesono K, Shiki Y, Takeuchi M, Chang Z, Inokuchi H, Ozeki H. 1988. Structure and organization of *Marchantia polymorpha* chloroplast genome. I. Cloning and gene identification. *J Mol Biol* 203: 281-298.
- Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Sano S, Umesono K, Shiki Y, Takeuchi M, Chang Z, Aota S, Inokuchi H, Ozeki H. 1986. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* 322: 572- 574.
- Oliver MJ, Mishler BD, Quisenberry JE. 1993. Comparative measures of desiccation-tolerance in the *Tortula ruralis* complex. I. Variation in damage control and repair. *Am J Bot* 80: 127-136.
- Oliver MJ, Murdock AG, Mishler BD, Kuehl JV, Boore JL, Mandoli DF, Everett KD, Wolf PG, Duffy AM, Karol KG. 2010. Chloroplast genome sequence of the moss *Tortula ruralis*: gene content, polymorphism, and structural arrangement relative to other green plant chloroplast genomes. *BMC Genomics* 11: 143.
- Palmer JD. 1985. Comparative organization of chloroplast genomes. *Ann Rev Genet* 19: 325-354.
- Proctor MCF, Oliver MJ, Wood AJ, Alpert P, Stark LR, Cleavitt N, Mishler BD. 2007. Desiccation tolerance in bryophytes. *The Bryologist* 110: 595-621.
- Sugiura C, Kobayashi Y, Aoki S, Sugita C, Sugita M. 2003. Complete chloroplast DNA sequence of the moss *Physcomitrella patens*: evidence for the loss and relocation of rpoA from the chloroplast to the nucleus. *Nucleic Acids Res* 31: 5324-5331.
- Sugiura M. 1989. The chloroplast chromosomes in land plants. *Annu Rev Cell Biol* 5: 51-70.
- Turmel M, Otis C, Lemieux C. 2006. The chloroplast genome sequence of *Chara vulgaris* sheds new light into the closest green algal relatives of land plants. *Mol Biol Evol* 23: 1324-1338.
- Wickett NJ, Zhang Y, Hansen SK, Roper JM, Kuehl JV, Plock SA, Wolf PG, DePamphilis CW, Boore JL, Goffinet B. 2008. Functional gene losses occur with minimal size reduction in

the plastid genome of the parasitic liverwort *Aneura mirabilis*. *Mol Biol Evol* 25: 393-401.

Online Sept. 9, 2012