Bangladesh J. Bot. **53**(1): 109-113, 2024 (March) DOI: https://doi.org/10.3329/bjb.v53i1.72249

COMPLETE CHLOROPLAST GENOMES OF *PALLAVICINIA LONGISPINA* STEPHANI AND *PLEUROZIA SUBINFLATA* (AUSTIN) AUSTIN

ZHE-MING SONG, BING-XIN LIU, MEI-YING FAN AND YING YU*

College of Life and Environmental Sciences, Huangshan University, Huangshan 245041, China

Keywords: Pallavicinia longispina, Pleurozia subinflata, Chloroplast genomes, RNA editing sites, SSR

Abstract

Recent progress in sequencing technology and the increasing number of chloroplast genomes available in liverworts have greatly improved our understanding of not only genome structure and evolution, but also its phylogeny and evolution. However, the plastome resource of this group is still ineffective, as the plastomes have been sequenced from a small number of species. In the present study, two newly sequenced chloroplast genomes from liverworts, one from *Pallavicinia longispina* Stephani and the other from *Pleurozia subinflata* (Austin) Austin were reported. The former genome has a length of 122,410 bp, with GC content 36.90%, and the latter has a length of 118,441 bp, with GC content 32.22%. A total of 132 and 127 genes were determined in the plastomes of *P. longispina* and *P. subinflata*, respectively. Furthermore, 41 short sequence repeats (SSRs), 419 editing sites and 15 non-adjacent repeats in the plastome of *P. longispina*, and 65 SSRs, 288 editing sites, and nine non-adjacent repeats in the plastome of *P. subinflata* were detected. Valuable genomic resources for further research on liverwort phylogeny, evolution, and population genetics are provided from this study.

Introduction

Liverworts, having ca. 6,000 species with a dominant tropical/subtropical distribution, represent an important branch of early land plants (Simpson 2010). This group, together with mosses and hornworts, comprises bryophytes, which are key to the study of early land plant evolution (Puttick *et al*. 2018, de Sousa *et al*. 2019, Harris *et al*. 2022). Recent progress of sequencing technology and increasing number of chloroplast genomes available in this group have not only greatly improved our understating of genome structure and evolution (Yu *et al*. 2019, Dong *et al*. 2021), but also enhanced the research of liverwort taxonomy, phylogeny and macroevolution (Yu *et al*. 2020, Dong *et al*. 2021, Xiang *et al*. 2022). However, compared to its counterpart-vascular plants, the plastome recourse of liverworts is still ineffective, as the complete chloroplast genomes have been sequenced from a small number of species so far. At the time of writing this paper (30/June/2023), over 20,000 chloroplast genomes from a variety of plant groups are available in several digital repositories, e.g., GeneBank (https://www.ncbi.nlm.nih.gov/genbank/), FigShare (https://figshare.com/) and Chloroplast Genome Information Resource (https://ngdc.cncb.ac.cn/cgir/). Of these, only 0.5% (ca. 100) are from liverworts. Insufficient genomic data hinders the research on phylogenomics and evolution in this group. Therefore, more plastomes sequencing is urgently needed for the studies that aim to answer some fundamental and complex questions concerning liverwort phylogeny and evolution.

Pallavicinia (Pallaviciniaceae, Pelliidae) and *Pleurozia* (Pleuroziaceae, Metzgeriidae) represent two independent and basal-diverging lineages of Jungermanniopsida (Yu *et al*. 2020, Dong *et al*. 2021). Based on morphological and molecular evidence (Thiers 1993, Crandall-Stotler *et al*. 2009, Mamontov *et al*. 2015), there are 11 and 13 species currently accepted in *Pallavicinia*

*Author for correspondence: <yuying@hsu.edu.cn>.

and *Pleurozia*, respectively (http://www.worldfloraonline.org). In these two genera, most species are distributed in tropical and subtropical forests (Gradstein *et al*. 2001). *Pallavicinia* species are well characterized by prostrate or erect thalli, the presence of midrib with one or several vascular strands, unistratose margins, and anacrogynous gynoecia on the dorsal surface of the thallus (Mamontov *et al*. 2015), while *Pleurozia* species are well defined by two-sided apical cell, bilobed leaves, absence of underleaves, and presence of fertile gynoecia (Thiers 1993). The relationships among and within these two genera have been largely clarified in recent phylogenetic studies using DNA fragments or loci (Forrest *et al*. 2006, He *et al*. 2006, Mamontov *et al*. 2015). However, some deep nodes in these genera remain weakly supported, such as the relationships within the *P. subciliata* complex in *Pallavicinia* (Mamontov *et al*. 2015), and those between *Pleurozia conchifolia*, *P. purpurea*, and *P. gigantea* in *Pleurozia* (He *et al*. 2006). *Pallavicinia* and *Pleurozia* had also been sampled in recent phylogenomic studies (Yu *et al*. 2020, Dong *et al*. 2021), but the samplings were all rather small. Currently, only three plastomes are available in these two genera, including two from *Pallavicinia lyellii* (Hook.) Gray (*Pallavicinia*, MN433344 and MN433345) and one from *Pleurozia purpurea* Lindb. (*Pleurozia*, MK645838). Here, two newly finished chloroplast genomes: one from *P. longispina* Stephani in *Pallavicinia* (GenBank accession number: ON494491) and the other one from *P. subinflata* (Austin) Austin in *Pleurozia* (GenBank accession number: ON357688) were reported. Furthermore, some important genomic parameters of finished genomes using the advanced method, including short sequence repeats (SSRs), RNA editing sites, and non-adjacent repeats were explored.

Materials and Methods

The plants of *Pallavicinia longispina* and *Pleurozia subinflata* were both collected from the Gao-Li-Gong-Shan area (Du-Long Xiang, Yunnan province, China). Two voucher specimens were deposited in the herbarium of Kunming Institute of Botany (KUN, Chinese Academy of Sciences, D. G. Long 36198 of *P. longispina*, and D. G. Long 36229 of *P. subinflata*). The total DNA from 10-50 milligrams material using a modified CTAB method was extracted. The quantity and quality of DNA using the method that has been described in Huang *et al*. (2019) were evaluated. Using the Covaris S220 sonicator (Covaris, Woburn, MA, USA.), approximately 1μg of high-quality genomic DNA was cut ultrasonically. Following the manufacturer's instruction manual (Illumina, San Diego, CA, USA.), multiple 500bp DNA fragments were obtained by cutting the complete genome. The genome library was constructed using the Nextera XT DNA Library Preparation Kit. Subsequently, the genome was sequenced using IlluminaHiSeq2000 to generate 90bp double-ended sequences, in BGI-Shenzhen, accumulating about 2Gb of original sequences.

Complete chloroplast genomes were assembled using GetOrganelle (Jin *et al*. 2020). The genomes were annotated using Geneious Prime (https://www.geneious.com/). Short sequence repeats (SSR) in two genomes, including mononucleotide (Mono-, minimum number of repeating units ≥ 10), dinucleotide (Di-, minimum number of repeating units ≥ 5) and trinucleotide (Tri-, minimum number of repeating units≥4), using GMATo v1.2 (Wang *et al*. 2013) were detected, and the RNA editing sites using PREPACT 3.0 (Lenz *et al*. 2018) were estimated. In addition, nonadjacent repeats (with a length longer than 50 bp) in two genomes using REputer (Kurtz *et al*. 2001) were detected. The maps of two plastomes were generated using OGDRAW (Greiner *et al*. 2019). Two chloroplast genomes were deposited in GenBank (GenBank accession number: ON357688 and ON494491).

Results and Discussion

The plastome of *P. longispina* has a length of 122,410 bp (Fig. 1), including a large single copy (LSC) region of 85,087 bp, a small single copy (SSC) region of 19,791 bp, and a pair of

inverted repeats (IR) regions of 8,996 bp for each. A total of 132 genes were determined in this genome, including 83 protein-coding genes, 38 tRNA and 9 rRNA genes. The GC content of this genome is 36.90%. This genome shared a great similarity to those of *P. lyellii* (MN433344 and MN433345) in gene content and structure (Dong *et al.* 2021). The plastome of *P. subinflata* has a length of 118,441 bp (Fig. 2), including an LSC of 81481 bp, an SSC of 19768 bp, and two IRs of 8,596 bp for each. A total of 127 genes were determined, including 87 protein-coding, 31 tRNA, and eight rRNA genes. The GC content of this genome was 32.22%. The newly sequenced genome was nearly identical with the one of *P. purpurea* (MK645838), with the exception of a minor difference of 275bp in the length of noncoding regions (the latter genome has a length of 118166 bp).

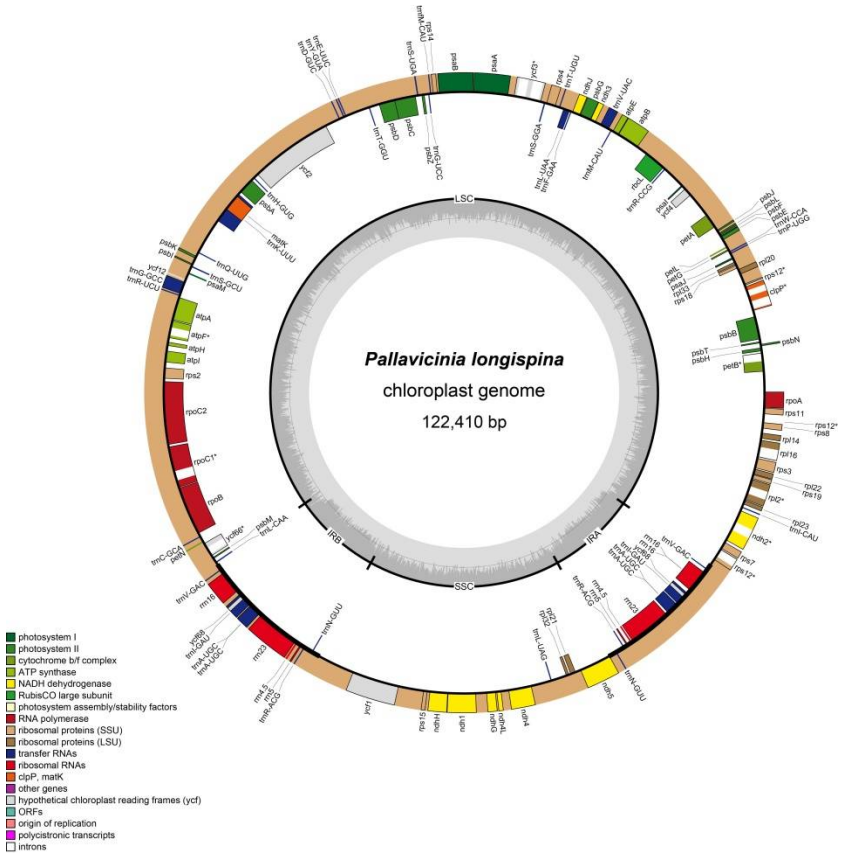

Fig. 1. Circular visualization chloroplast genome of *Pallavicinia longispina*. Genes (exons are shown as closed boxes) shown on the outside of the circle are transcribed clockwise, whereas on the inside are transcribed counter-clockwise.

Short sequence repeats (SSRs) or microsatellite DNAs, a class of molecular markers of short DNA repeat tandem sequences (1-6bp), is considered as one of the important genomic components that are involved in many important biological processes (Vieira *et al.* 2016). Using GMATo, 41 SSRs in the genome of *P. longispina* (Table S1), including 12 mono-, 27 di-, one tri-, and one quad-nucleotide tandem repeats, and 65 SSRs in the genome of *P. subinflata* (Table S2), including 45 mono-, 14 di-, and six tri-nucleotide tandem repeats were detected.

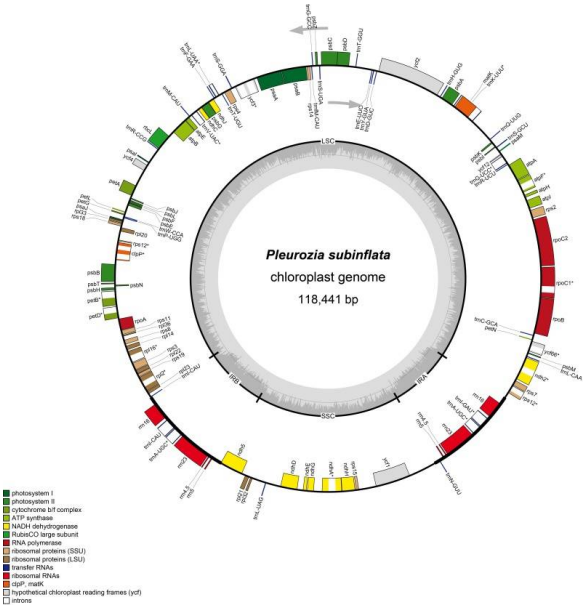

Fig. 2. Circular visualization chloroplast genome of *Pleurozia subinflata*. Genes (exons are shown as closed boxes) shown on the outside of the circle are transcribed clockwise, whereas on the inside are transcribed counter-clockwise.

The RNA editing is a post-transcriptional modification mechanism that changes the identity of nucleotides in an RNA sequence, playing an important role in the stability of the structure of proteins (Nishikura 2010). Taking the chloroplast genome of *Marchantia paleacea* Bertol. as reference (GenBank accession number: NC_001319.1), 419 (Table S3) and 288 editing sites (Table S4) in the plastomes of *P. longispina* and *P. subinflata*, respectively, were estimated. Nevertheless, the deduced editing sites need to be checked using the transcriptomic data. Furthermore, 15 non-adjacent repeats, including three palindromic, six forward, and six reverse, in the plastome of *P. longispina*, and nine, including four palindromic, two forward, and three reverse, in the plastome of *P. subinflata* using REputer, were detected. In general, the two newly finished chloroplast genomes from *Pallavicinia* and *Pleurozia* as well as estimation of some important genomic parameters provide valuable resources for further research of liverwort phylogeny, evolution, and population genetics.

Acknowledgments

This study was finally supported by National Natural Science Foundation of China (No. 31970227), Anhui *Provincial Department of Education* (No. 2022AH051962); and Huangshan University (No. 2022xzx007 and No. 2021xkjq004). The authors are grateful to the Innovative Center of Plant Physiology and Biochemistry at Huangshan University for assistance in field works and specimen examination.

References

Crandall-Stotler B, Stotler R, Long D 2009. Phylogeny and classification of the Marchantiophyta. Edinb. J. Bot. **66**(1): 155-198.

de Sousa F, Foster PG, Donoghue PCJ, Schneider H, Cox CJ 2019. Nuclear protein phylogenies support the monophyly of the three bryophyte groups (Bryophyta Schimp.). New Phytol. **222**(1): 565-575.

Dong S, Zhang S, Zhang L, Wu H, Goffinet B, Liu Y 2021. Plastid genomes and phylogenomics of liverworts (Marchantiophyta): Conserved genome structure but highest relative plastid substitution rate in land plants. Mol. Phylogenet. Evol. **161**: 107171.

Forrest LL, Davis EC, Long DG, Crandall-Stotler BJ, Clark A, Hollingsworth ML 2006. Unraveling the evolutionary history of the liverworts (Marchantiophyta): multiple taxa, genomes and analyses. Bryologist **109**(3): 303-334.

Gradstein S, Churchill S, Salazar Allen N 2001. Guide to the bryophytes of tropical America. Mem. N. Y. Bot. Gard. **86**: 1-577.

Greiner S, Lehwark P, Bock R 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. **47**(W1): W59-W64.

Harris BJ, Clark JW, Schrempf D, Szöllősi GJ, Donoghue PCJ, Hetherington AM, Williams TA 2022. Divergent evolutionary trajectories of bryophytes and tracheophytes from a complex common ancestor of land plants. Nat. Ecol. Evol. **6**(11): 1634-1643.

He X, Juslén A, Ahonen I, Glenny D, Piippo S. 2006. Illuminating the evolutionary history of liverworts (Marchantiophyta) - Towards a natural classification. Cladistics **22**(1): 1-31.

Huang W-Z, Ma W-Z, Schneider H, Yu Y, Wu Y-H 2019. Mitochondrial genome from *Andreaea wangiana* reveals structural conservatism and a trend of size reduction in mosses. Bryologist **122**(4): 597-606.

Jin JJ, Yu WB, Yang JB, Song Y, de Pamphilis CW, Yi TS, Li DZ 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. **21**(1): 241.

Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. Nucleic Acids Res. **29**(22): 4633-4642.

Lenz H, Hein A, Knoop V 2018. Plant organelle RNA editing and its specificity factors: enhancements of analyses and new database features in PREPACT 3.0. BMC Bioinform. **19**(1): 255.

Mamontov Y, Konstantinova N, Vilnet A, Bakalin V 2015. On the phylogeny and taxonomy of Pallaviciniales (Marchantiophyta), with overview of Russian species. Arctoa **24**(1): 98-123.

Nishikura K 2010. Functions and regulation of RNA editing by ADAR deaminases. Ann. Rev. Biochem. **79**: 321-49.

Puttick MN, Morris JL, Williams TA, Cox CJ, Edwards D, Kenrick P, Pressel S, Wellman CH, Schneider H, Pisani D, Donoghue PCJ 2018. The interrelationships of land plants and the nature of the ancestral embryophyte. Curr. Biol. **28**(5): 733-745.e2.

Simpson MG 2010. 3 - Evolution and Diversity of Green and Land Plants. *In*: Plant Systematics (Second Edition), Simpson MG (Ed), pp. 55-74. Academic Press, San Diego.

Thiers BM 1993. A monograph of *Pleurozia* (Hepaticae; Pleuroziaceae). The Bryologist **96**(4): 517-554.

Vieira ML, Santini L, Diniz AL, Munhoz Cde F. 2016. Microsatellite markers: what they mean and why they are so useful. Genet. Mol. Biol. **39**(3): 312-28.

Wang X, Lu P, Luo Z 2013. GMATo: A novel tool for the identification and analysis of microsatellites in large genomes. Bioinformation **9**(10): 541-544.

Xiang Y-L, Jin X-J, Shen C, Cheng X-F, Shu L, Zhu R-L 2022. New insights into the phylogeny of the complex thalloid liverworts (Marchantiopsida) based on chloroplast genomes. Cladistics **38**(6): 649-662.

Yu Y, Liu H-M, Yang J-B, Ma W-Z, Pressel S, Wu Y-H, Schneider H 2019. Exploring the plastid genome disparity of liverworts. J. Syst. Evol. **57**(4): 382-394.

Yu Y, Yang J-B, Ma W-Z, Pressel S, Liu H-M, Wu Y-H, Schneider H 2020. Chloroplast phylogenomics of liverworts: a reappraisal of the backbone phylogeny of liverworts with emphasis on Ptilidiales. Cladistics **36**(2): 184-193.

(Manuscript received on 22 May 2023; revised on 18 March, 2024)