

Characterization of the complete chloroplast genome of *Platycarya strobilacea* (Juglandaceae)

Jing Yan¹ · Kai Han¹ · Shuyun Zeng¹ · Peng Zhao¹ · Keith Woeste² · Jianfang Li¹ · Zhan-Lin Liu¹ 

Received: 1 September 2016 / Accepted: 4 October 2016 / Published online: 6 October 2016
© Springer Science+Business Media Dordrecht 2016

Abstract. The whole chloroplast genome (cp genome) sequence of *Platycarya strobilacea* was characterized from Illumina pair-end sequencing data. The complete cp genome was 160,994 bp in length and contained a large single copy region (LSC) of 90,225 bp and a small single copy region (SSC) of 18,371 bp, which were separated by a pair of inverted repeat regions (IRs, 26,199 bp). The genome contained 130 genes, including 85 protein-coding genes (80 PCG species), 36 tRNA genes (29 tRNA species) and 8 ribosomal RNA genes (4 rRNA species). Most genes occur as a single copy, but 15 genes are duplicated.

The overall AT content of the cp genome is 64%, and the corresponding values of the LSC, SSC and IR regions are 66.4, 70.1 and 57.5%, respectively. Phylogenetic analysis confirmed the placement of *P. strobilacea* near to *Juglans*.

Keywords *Platycarya strobilacea* · Hickory · China · Conservation · Chloroplast genome

Platycarya strobilacea, the only species in the monotypic genus *Platycarya* (Juglandaceae), mainly grows as scattered individuals in evergreen forests of South China (Chen et al. 2012). As one of main components of the local vegetation, this species plays a key role in the forest ecosystems. Extracts of its fruit contain compounds used for anti-aging cosmetics (Kim et al. 2010). However, due to anthropogenic over-exploitation and fragmentation of its natural habitat, wild resources of *P. strobilacea* have dramatically declined and need urgent protection and restoration. An improved understanding of its genetics would contribute to the formulation of conservation strategies. In this study, we assembled and characterized the complete chloroplast genome of *P. strobilacea* from Illumina sequencing data.

Fresh leaves of an individual were collected from Zhengpin (Shaanxi, China; 31°53'N, 109°31'E), and were used for total genomic DNA isolation. A shotgun library was prepared and sequenced on the Illumina Hiseq 2500 Sequencing System (Illumina, CA, USA). In total, 8.2G raw reads were obtained, quality-trimmed and assembled for the chloroplast genome using the program MITObim v1.8 (Hahn et al. 2013) with *Juglans regia* (GenBank: NC_028617.1) (Potter et al. 2002) as the initial reference. The genome was annotated by aligning with the reference chloroplast genome. A map of the genome was generated using OGDRAW (<http://ogdraw.mpimp-golm.mpg.de/>) (Lohse et

✉ Zhan-Lin Liu
liuzl@nwu.edu.cn

Jing Yan
734037784@qq.com

Kai Han
hank.nwu@outlook.com

Shuyun Zeng
zengsy.nwu@outlook.com

Peng Zhao
pengzhao@nwu.edu.cn

Keith Woeste
kwoeste@fs.fed.us

Jianfang Li
jianfangli2010@qq.com

¹ Key Laboratory of Resource Biology and Biotechnology in Western China (Ministry of Education), College of Life Science, Northwest University, Xi'an 710069, Shaanxi Province, China

² USDA Forest Service Hardwood Tree Improvement and Regeneration Center (HTIRC), Department of Forestry and Natural Resources, Purdue University, 715 West State Street, West Lafayette, IN 47907, USA

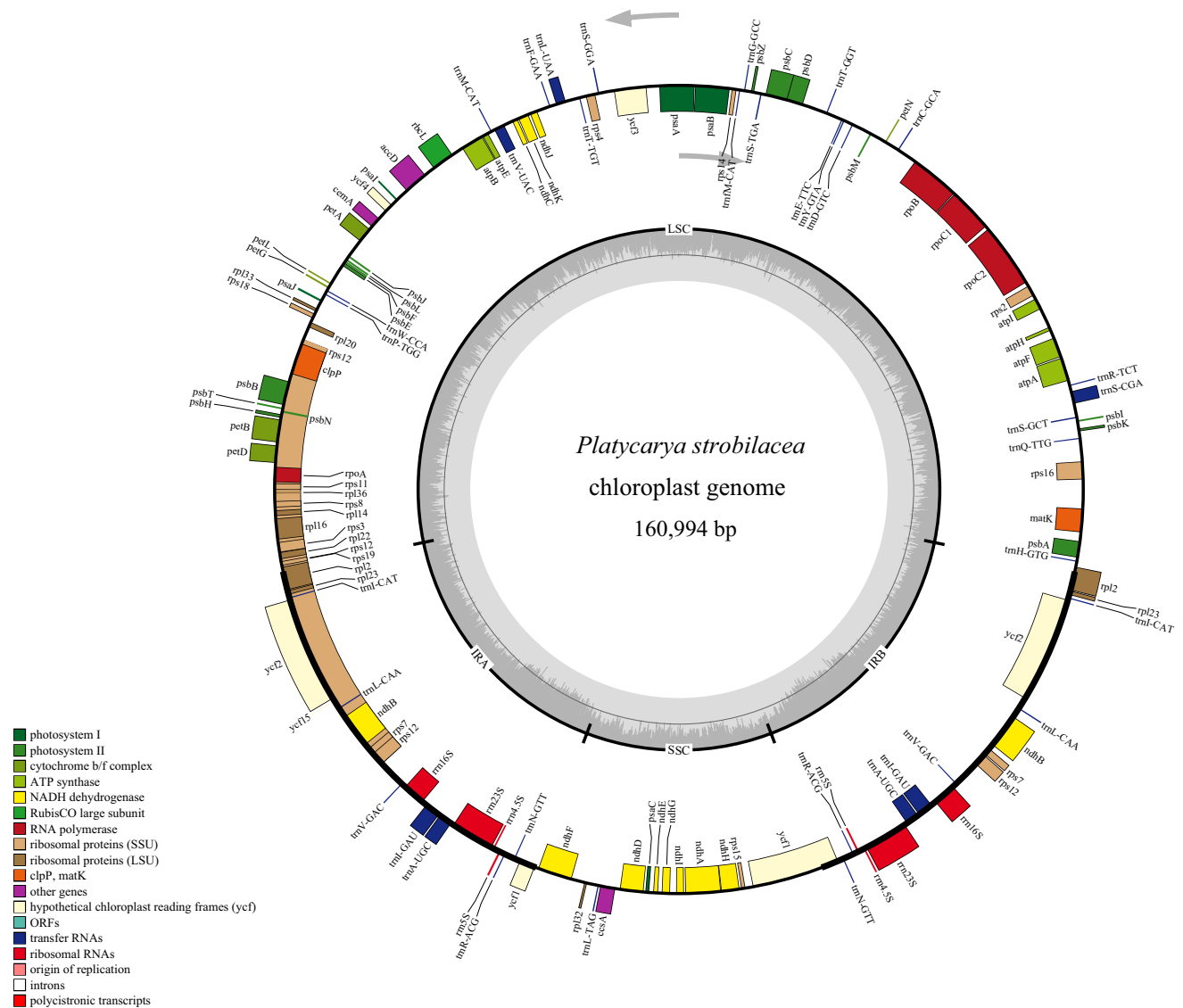


Fig. 1 Gene map of the whole chloroplast genome of *Platycarya strobilacea*

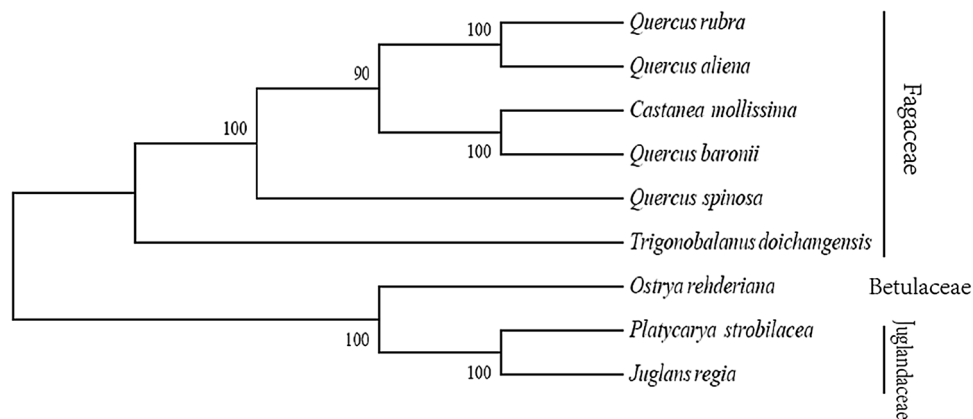


Fig. 2 Phylogeny of nine species in Fagales based on the neighbor-joining (NJ) analysis of the concatenated coding sequences of 80 chloroplast PCGs. The bootstrap values are indicated next to the branches. *Quercus rubra* JX970937, *Quercus aliena* NC_026790.1, *Castanea*

mollissima HQ336406.1, *Quercus baronii* NC_029490.1, *Quercus spinosa* KM841421.1, *Trigonobalanus doichangensis* NC_023959.1, *Ostrya rehderiana* NC_028349.1, *Platycarya strobilacea* KX868670, *Juglans regia* NC_028617.1

al. 2013). The annotated cp genome has been deposited into GenBank with the accession number KX868670.

The cp genome of *P. strobilacea* is 160,994 bp in length, and is comprised of a pair of inverted repeat (IR) regions (26,199 bp each), a large single-copy (LSC) region of 90,225 bp, and a small single-copy (SSC) region of 18,371 bp (Fig. 1). The chloroplast genome harbors 130 genes, including 85 protein-coding genes (80 PCG species), 36 transfer RNA genes (29 tRNA species) and 8 ribosomal RNA genes (4 rRNA species). The majority of the gene species are single copy, but 15 gene species located within the IR regions are duplicated, including 5 PCG species (*rpl2*, *rpl23*, *rps7*, *rps12*, *ycf2*), 6 tRNA species (*trnA-UGC*, *trnI-CAT*, *trnL-CAA*, *trnN-GTT*, *trnR-ACG*, *trnV-GAC*), and all rRNA species (4.5S, 5S, 16S, 23S rRNA). The base composition is asymmetric (31.7% A, 18.3% C, 17.7% G, 32.3% T) with an overall A+T content of 64%. The A+T contents of the LSC, SSC and IR regions are 66.4, 70.1 and 57.5%, respectively.

To identify the phylogenetic position of *Platycarya strobilacea*, a neighbor-joining (NJ) phylogenetic tree was reconstructed using the concatenated coding sequences of 80 chloroplast PCGs from nine taxa of Fagales (see Fig. 2 for details) with MEGA6 (Tamura et al. 2013). The topology of the phylogenetic tree was consistent with the traditional tribe-level taxonomy of Fagales. The two Juglandaceae species (*Platycarya strobilacea*, *Juglans regia*) grouped into one clade and formed a sister to *Ostrya rehderiana*.

Acknowledgments This study was financially supported by the National Natural Science Foundation of China (31670219, 31370353, J1210063) and Natural Science Basic Research Plan in Shaanxi Province of China (2015JM3106). Mention of a trademark, proprietary product, or vendor does not constitute a guarantee or warranty of the product by the U.S. Department of Agriculture and does not imply its approval to the exclusion of other products or vendors that also may be suitable.

References

- Chen SC, Zhang L, Zeng J, Shi F, Yang H, Mao YR, Fu CX (2012) Geographic variation of chloroplast DNA in *Platycarya strobilacea* (Juglandaceae). *J Syst Evol* 50:374–385
- 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
- Kim YH, Kim HK, Han CS, Yang HC, Park SH, Jang HI, Kim JW, Choi YS, Lee NH (2010) Anti-wrinkle activity of *Platycarya strobilacea* extract and its application as a cosmeceutical ingredient. *J Cosmet Sci* 61:211–223
- Lohse M, Drechsel O, Kahlau S, Bock R (2013) OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res* 41:W575–W581
- Potter D, Gao FY, Baggett S, McKenna RJ, McGranahan HG (2002) Defining the sources of Paradox: DNA sequence markers for North American walnut (*Juglans* L.) species and hybrids. *Sci Hortic* 94:157–170
- 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