

The complete chloroplast genome sequence of the biofuel plant Sacha Inchi, *Plukenetia volubilis*

Xiao-Di Hu^{a,b}, Bang-Zhen Pan^a, Qiantang Fu^a, Mao-Sheng Chen^a and Zeng-Fu Xu^a 

^aKey Laboratory of Tropical Plant Resources and Sustainable Use, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Menglun, Yunnan, China; ^bCollege of Life Sciences, University of Chinese Academy of Sciences, Beijing, China

ABSTRACT

Sacha Inchi (*Plukenetia volubilis*) is a potential woody oil seed plant for producing healthy vegetable oil due to high content of α -linolenic acid in its seeds. In this study, we report the structure of the complete chloroplast genome of *P. volubilis* using high-throughput next-generation sequencing technology. The circular chloroplast genome is 161,733 bp in size, containing a pair of inverted repeat regions (IR) of 27,382 bp each, which were separated by a large single copy region (LSC) of 88,843 bp and a small single copy region (SSC) of 18,126 bp. The chloroplast genome harbors 135 genes, including 92 protein-coding genes, 35 tRNA genes and 8 rRNA genes. Based on the phylogenetic relationships between the chloroplast genome of *P. volubilis* and those of the other species, *P. volubilis* is most closely related to castor bean (*Ricinus communis*).

ARTICLE HISTORY

Received 17 January 2018
Accepted 1 February 2018





KEYWORDS

Plukenetia volubilis;
chloroplast genome;
phylogenetic relationship

Sacha Inchi (*Plukenetia volubilis*), also known as Inca peanut, belongs to the genus *Plukenetia* of tribe Plukenetieae in the family Euphorbiaceae (Gillespie 2007). Sacha Inchi seeds have high content of α -linolenic acid, which is an essential fatty acid in our diet (Burdge 2006; Gutierrez et al. 2011; Chirinos et al. 2013). So far little genetic information on Sacha Inchi is reported. The chloroplast genome information was extensively applied in understanding plant genetic diversity and evolution (Ye et al. 2014). In this study, we assembled and annotated the chloroplast genome of *P. volubilis* using high-throughput next-generation sequencing technology. *P. volubilis* was introduced and cultivated to xishuangbanna of China in 2006 (Cai 2011), and young leaves were collected at the Xishuangbanna Tropical Botanical Garden, the Chinese Academy of Sciences, Menglun, Yunnan, China (21°54'N, 101°46'E, 580 m above sea level). Total genomic DNA was extracted with the CTAB method (Doyle 1986). According to the manufacturer's manual (Illumina, San Diego, CA), the pair-end and long mate-pair libraries were constructed and then sequenced using Hiseq2000 platform. *De novo* assembly of chloroplast genome was performed using the SPAdes with the default parameters (Bankevich et al. 2012). The assembled contigs were aligned and ordered with the chloroplast genome of castor bean (*Ricinus communis*, GenBank accession number: NC_016736), whose genome is closest to that of Sacha Inchi. The plastome was annotated by Dual Organellar GenoMe Annotator (DOGMA) (Wyman et al. 2004), and the annotated genomic sequence was deposited in GenBank with the accession number MF062253.

The circular genome of *P. volubilis* is 161,733 bp in size, and comprises a pair of inverted repeat (IR) regions of 27,382 bp each, a large single-copy region (LSC) of 88,843 bp and a small single-copy region (SSC) of 18,126 bp. The chloroplast genome harbours 135 genes, including 92 protein-coding genes, 35 tRNA genes and 8 rRNA genes. The overall GC content of the chloroplast genome is 36.2% that is similar to other Euphorbiaceae chloroplast genomes (Table 1), and those of the LSC and SSC regions are 33.7% and 30.5%, respectively. The size of Euphorbiaceae chloroplast genomes is in the range of 161–164 kb (Table 1).

To determine the phylogenetic position of *P. volubilis*, a phylogenetic analysis was carried out among eight complete chloroplast genomes that are derived from *P. volubilis*, *Ricinus communis*, *Hevea brasiliensis*, *Manihot esculenta*, *Jatropha curcas*, *Euphorbia esula*, *Arabidopsis thaliana* and *Oryza sativa*. The chloroplast genome sequences of the other seven species except *P. volubilis* were downloaded from National Center for Biotechnology Information (NCBI, <https://www.ncbi.nlm.nih.gov/>). Alignment of sequences was performed using MAFFT (Katoh and Standley 2013). The maximum-likelihood analyses were accomplished using the program MEGA 7 (Kumar et al. 2016). The results showed that all six species of the family Euphorbiaceae formed a single cluster, and *P. volubilis* is closest to *R. communis* (Figure 1). The assembled chloroplast genome of *P. volubilis* reported in this paper will provide valuable information on the genetic diversity research of *P. volubilis*, and enrich the resources of chloroplast genomes in plants.

CONTACT Mao-Sheng Chen  chenms@xtbg.org.cn  Key Laboratory of Tropical Plant Resources and Sustainable Use, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Menglun, Yunnan, China; Zeng-Fu Xu  zfxu@xtbg.ac.cn  Key Laboratory of Tropical Plant Resources and Sustainable Use, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Menglun, Yunnan, China

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Table 1. Comparison of Euphorbiaceae chloroplast genomes reported in this study.

	<i>P. volubilis</i>	<i>R. communis</i>	<i>H. brasiliensis</i>	<i>M. esculenta</i>	<i>J. curcas</i>	<i>E. esula</i>
Size (bp)	161,733	163,161	161,191	161,453	163,856	160,512
Number of genes	135	131	129	128	129	132
GC content (%)	36	36	36	36	35	36

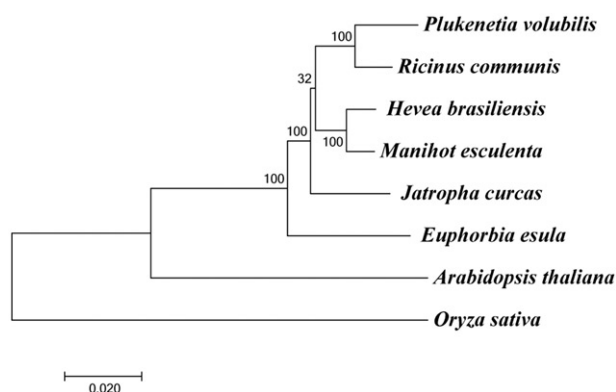


Figure 1. Maximum-likelihood phylogenetic tree of eight species. These species include *Plukenetia volubilis* (MF062253), *Ricinus communis* (NC_016736.1), *Hevea brasiliensis* (NC_015308.1), *Manihot esculenta* (NC_010433.1), *Jatropha curcas* (NC_012224.1), *Euphorbia esula* (NC_033910.1), *Arabidopsis thaliana* (NC_000932.1), and *Oryza sativa* (NC_027678.1). The numbers at the nodes are bootstrap values with 1000 replicates.

Acknowledgements

The authors gratefully acknowledge Cheng-Wen Gao and Kui Li for their help during data analysis. The authors also acknowledge the Central Laboratory of the Xishuangbanna Tropical Botanical Garden for providing high-performance computing and other research facilities.


Disclosure statement

The authors declare no competing financial interests.

Funding

This work was supported by funding from the '135' Program of Chinese Academy of Sciences (2017XTBG-T02) and the National Natural Science Foundation of China (31670612).

ORCID

Zeng-Fu Xu  <http://orcid.org/0000-0001-6045-5865>

References

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19:455–477.
- Burdge GC. 2006. Metabolism of alpha-linolenic acid in humans. *Prostaglandins Leukot. Essent. Fatty Acids.* 75:161–168.
- Cai Z. 2011. Advance in research on a special woody oilseed crop of *Plukenetia volubilis* L. *China Oils Fats.* 36:1–6.
- Chirinos R, Zuloeta G, Pedreschi R, Mignolet E, Larondelle Y, Campos D. 2013. Sacha Inchi (*Plukenetia volubilis*): a seed source of polyunsaturated fatty acids, tocopherols, phytosterols, phenolic compounds and antioxidant capacity. *Food Chem.* 141:1732–1739.
- Doyle JJ. 1986. A rapid DNA isolation procedure for small amounts of fresh leaf tissue. *Phytochem Bull.* 19:11–15.
- Gillespie LJ. 2007. A revision of paleotropical *Plukenetia* (Euphorbiaceae) including two new species from Madagascar. *Syst Bot.* 32:780–802.
- Gutierrez LF, Rosada LM, Jimenez A. 2011. Chemical composition of Sacha Inchi (*Plukenetia volubilis* L.) seeds and characteristics of their lipid fraction. *Grasas Aceites.* 62:76–83.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30:772–780.
- Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 33:1870–1874.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics.* 20:3252–3255.
- Ye C-Y, Lin Z, Li G, Wang Y-Y, Qiu J, Fu F, Zhang H, Chen L, Ye S, Song W, et al. 2014. Echinochloa chloroplast genomes: insights into the evolution and taxonomic identification of two weedy species. *PLoS One.* 9:e113657.