

## The complete chloroplast genome of *Strobilanthes crispus* (Acanthaceae)

Guoqing Wan<sup>a,b\*</sup>, Yanhui Liu<sup>b\*</sup>, Yanjun Zheng<sup>a</sup>, Xuefeng Gu<sup>a,b</sup> and Changlian Lu<sup>a,b</sup>

<sup>a</sup>Shanghai Key Laboratory of Molecular Imaging, Zhoupu Hospital, Shanghai University of Medicine and Health Sciences, Shanghai, PR China;

<sup>b</sup>School of Pharmacy, Shanghai University of Medicine and Health Sciences, Shanghai, PR China

### ABSTRACT

We reported and characterized the complete chloroplast genome sequence of *Strobilanthes crispus* Blume 1826. *Strobilanthes crispus* belongs to the Acanthaceae family and has a number of local names including Batuzin, Bayam Karang, Kotz Bellin, and Pekka Batu, which is native to Malaysian with diverse beneficial uses. Green leaves were determined using next-generation sequencing. We found that the entire chloroplast genome of *S. crispus* was 144,987 bp in length, included four segments, named a large single-copy (LSC) region (92,556 bp), a small single-copy (SSC) region (17,783 bp), and a pair inverted repeat regions (IRs) (17,324 bp in each), respectively. The chloroplast genome of *S. crispus* contained a total of 129 functional genes, including 84 protein-coding genes, 37 transfer RNAs (tRNAs), and eight ribosomal RNA (rRNA) genes. The phylogenetic tree reconstructed by nine chloroplast genomes reveals that *S. crispus* is most closely related to *Strobilanthes bantonensis* and *Strobilanthes cusia*.

### ARTICLE HISTORY

Received 19 September 2021  
Accepted 7 August 2022

### KEYWORDS

*Strobilanthes crispus*;  
complete chloroplast  
genome; phylogenetic  
analysis

*Strobilanthes crispus*, known as the black faced general, is native to Malaysia. In folk medicine, *S. crispus* has been used to treat some diseases including diuretic, anticancer, antidiabetic, antilytic, and laxative for decades (Sunarto 1977; Goh 2004). It belongs to a family of *Acanthopanax senticosus* and has many local names, including Batuzin, Bayam Karang, Kotz Bellin, and Pekka Batu. In addition, in previous studies, its phylogenetic position is unclear because of the lack of genomic information. Here, we characterized the complete chloroplast genome sequence of *S. crispus* based on the genome skimming sequencing data. These genomic data can provide intragenic information to clarify the taxonomic characteristics and valuable information about the evolution of *Strobilanthes*.

The samples of *S. crispus* were collected from Zhao'an County, Zhangzhou, Fujian Province, China (23°71N, 117°18E). The voucher specimen was deposited in Shanghai University of Medicine and Health Sciences under the voucher number SUMHS020190707 (Guoqing Wan, [wangq@sumhs.edu.cn](mailto:wangq@sumhs.edu.cn)). The total genomic DNA was extracted from the fresh leaves of *S. crispus* using the DNeasy Plant Mini Kit (Qiagen, Valencia, CA). The DNA was stored at 80 °C in our lab. The whole-genome sequencing was conducted on the BGISEQ-500 sequencing platform from Hefei Biodata Biotechnologies Inc. (Hefei, China). The complete chloroplast genome was assembled using the program SPAdes assembler 3.10.0 using default settings (Bankevich et al. 2012). Software GeSeq (Tillich et al. 2017) was used to annotate the

chloroplast genes. All transfer RNA (tRNA) genes were further verified by using tRNAscan-SE 2.0 program (Chan et al. 2021).

The chloroplast genome sequence of *S. crispus* was submitted to NCBI (GenBank accession no. MW411449), consisting of two inverted repeat regions (IRs) of 17,324 bp, a large single-copy (LSC) with 92,556 bp, and a small single-copy (SSC) region with 17,783 bp. The content of overall GC in the complete chloroplast genome of the *S. crispus* is 38.2%, and the corresponding values in LSC, SSC, and IR regions are 36.5%, 32.6%, and 45.6%, respectively. The chloroplast genome of *S. crispus* is predicted to contain 129 genes with different functional classifications, including 84 protein-coding genes, 37 tRNA genes, and eight ribosomal RNA (rRNA) genes. Among these genes, five protein-coding, six tRNA, and four rRNA genes were duplicated in the IR regions.

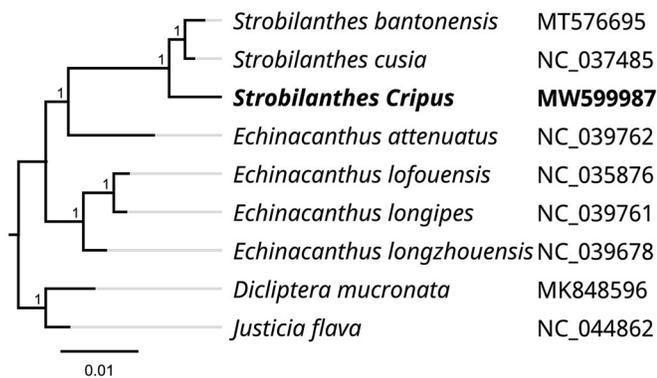
To investigate its taxonomic status, alignment was performed on the nine complete chloroplast genome sequences using MAFFT v7.307 (Katoh and Standley 2013), and a maximum-likelihood (ML) tree was constructed with 1000 bootstrap replicates in FastTree v2.1.10 (Price et al. 2010). *Dicliptera mucronata* (MK848596) and *Justicia flava* (NC\_044862) were used as outgroups to construct the phylogenetic tree. The ML phylogenetic tree shows that *S. crispus* is most closely related to *Strobilanthes bantonensis* (MT576695) and *Strobilanthes cusia* (NC\_037485) in Acanthaceae (Figure 1).

**CONTACT** Changlian Lu  [lvcl@sumhs.edu.cn](mailto:lvcl@sumhs.edu.cn)  Shanghai Key Laboratory of Molecular Imaging, Zhoupu Hospital, Shanghai University of Medicine and Health Sciences, 279 Zhouzhu Road, Shanghai 201318, PR China

\*Both authors contributed equally to this work.

© 2022 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.



**Figure 1.** Phylogenetic tree inferred by maximum-likelihood (ML) method based on the complete chloroplast genome of nine representative species. A total of 1000 bootstrap replicates were computed and the bootstrap support values were shown at the branches. Accession numbers are shown in the figure.

## Acknowledgements

This study did not involve endangered or protected species, and the plant was collected under special permission from Shanghai University of Medicine and Health Sciences. We wish to thank Ningqiu Zhang, Huiqing Pan, and Wei Xie for advice on critical revisions.

## Author contributions

Guoqing Wan and Changlian Lu drafted or revised the manuscript. Yanjun Zheng, Xuefeng Gu, and Changlian Lu were involved in the conception and design. Guoqing Wan and Yanhui Liu participated in the data collection analysis. Guoqing Wan, Yanhui Liu, and Changlian Lu revised the paper critically for intellectual content. All authors have reviewed the final version of the manuscript and approved it for publication.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This work was supported by the funding from the Shanghai Municipal Education Commission-Young Teacher Training Projection Program [ZZJKYX19007]; “Top-100 Talent Cultivation Plan” of SUMHS; Scientific Research Foundation of SUMHS; Construction Project of Shanghai Key Laboratory of Molecular Imaging [18DZ2260400].

## Data availability statement

The genome sequence data that support the findings of this study are openly available in GenBank of NCBI (<https://www.ncbi.nlm.nih.gov/nucleotide/MW599987/>) under the accession no. MW599987. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA701834, SRR13720062, and SAMN17914799, respectively.

## References

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19(5):455–477.
- Chan PP, Lin BY, Mak AJ, Lowe TM. 2021. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res.* 49:9077–9096.
- Goh KL. 2004. Malaysian herbs. Vol. 2. Klang: Goh Kong Ling; p. 249.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
- Price MN, Dehal PS, Arkin AP. 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLOS One.* 5(3):e9490.
- Sunarto PA. 1977. *Materia Medica Indonesia*. 1st ed. Jakarta, Indonesia: Penerbitan Direktorat Jenderal Pengawasan Obat dan Makanan; p. 95–99.
- Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. GeSeq – versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* 45(W1):W6–W11.