

## The complete chloroplast genome assembly of *Amorphophallus krausei* Engler, Pflanzenr 1911 (Araceae) from southwestern China

Si Yin  and Yong Gao 

College of Biological Resource and Food Engineering, Qujing Normal University, Qujing, Yunnan, China

### ABSTRACT

Plants in the genus *Amorphophallus*, many of which possess high konjac glucomannan content, are considered important cash crops in many Asian countries. Wild relatives of cultivated *Amorphophallus* species are valuable resources for the genetic improvement of these crops. To aid in future genetic research of wild germplasm resources of *Amorphophallus*, a single individual of *Amorphophallus krausei* Engler, Pflanzenr 1911 was collected from southwestern China, and its chloroplast genome was sequenced using next-generation sequencing technologies. The assembled chloroplast genome was 172,418 bp in length with a GC content of 35.23% (GenBank accession no. OR416863). A typical quadripartite structure was found in the genome, which was comprised of one large single-copy (LSC), one small single-copy (SSC), and two inverted repeats (IRs), with lengths of 91,983 bp, 15,591 bp, 32,422 bp, and 32,422 bp, respectively. A total of 132 genes were annotated in the genome, including 86 protein-coding genes, 38 tRNAs, and 8 rRNAs. A maximum likelihood (ML) tree of *A. krausei* and 17 other species in the family Araceae suggested that all *Amorphophallus* species formed a single monophyletic clade. A close relationship among *A. konjac*, *A. albus*, and *A. krausei* was also revealed by the phylogenetic tree. The newly sequenced chloroplast genome of *A. krausei* will support future genetic studies, particularly the assessment of genetic diversity, resource conservation, and phylogeographic research.

### ARTICLE HISTORY

Received 5 September 2023  
Accepted 24 November 2023

### KEYWORDS



*Amorphophallus krausei*;  
Araceae; chloroplast  
genome; phylogenetic  
analysis


### Introduction

The genus *Amorphophallus* is a group of about 200 perennial plants that are mainly distributed throughout West Africa, subtropical regions of Asia, and islands in the Pacific Ocean (Li et al. 2010). Several *Amorphophallus* species, possessing high konjac glucomannan content in their corms, have been cultivated as food and cash crops in Asian countries such as China, Japan, and India (Gao et al. 2017). *Amorphophallus krausei* Engler, Pflanzenr 1911 typically lives in shaded forest margins in Southeast and East Asia, including southwestern China, Myanmar, Laos, and northern Vietnam (Li et al. 2010). Less adaptable than other *Amorphophallus* species, *A. krausei* has not been widely planted throughout China, except by some ethnic minorities who utilize *A. krausei* as a food resource. However, the wild populations of *A. krausei* are declining due to disturbance from human activities in China. Assessing the population genetic diversity and demographic dynamics of wild germplasm resources is important for the future breeding of *Amorphophallus* (Srzednicki and Borompichaichartkul 2020). Here, we report the first complete chloroplast genome assembly of *A. krausei*, which will support further phylogeographic and population genetic studies of this species.



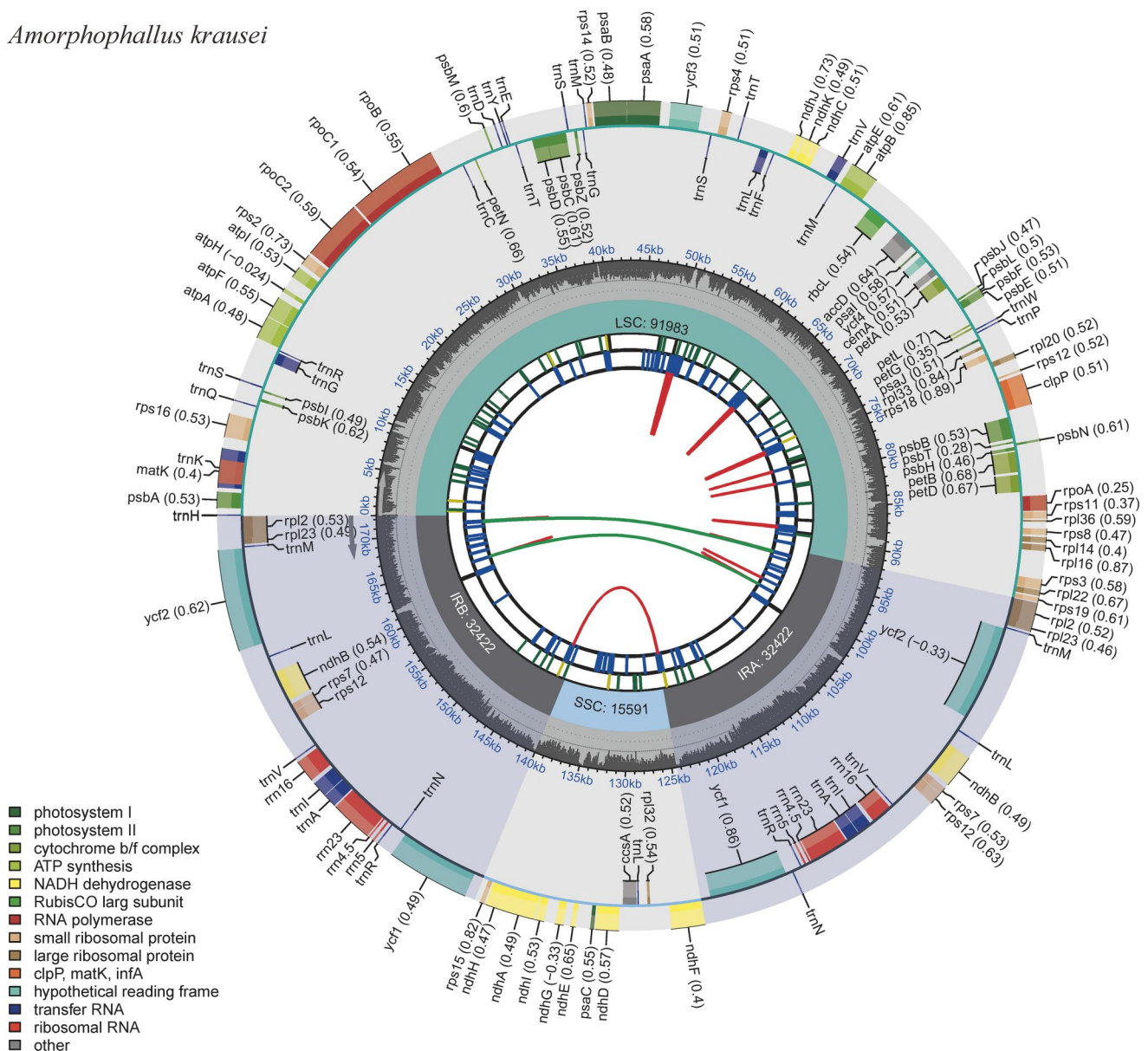
**Figure 1.** Morphological characteristics of the flower of *Amorphophallus krausei*. The photo was taken by the author Yong Gao at the greenhouse of Qujing Normal University. The Spadix of *Amorphophallus krausei* is nearly as long as the spathe. The appendix is fusiform or fusiform-conic, sometimes slightly laterally compressed.

**CONTACT** Yong Gao  562698574@qq.com  College of Biological Resource and Food Engineering, Qujing Normal University, Qujing, Yunnan 655011, China.

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/23802359.2023.2288889>.

© 2023 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-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

*Amorphophallus krausei*

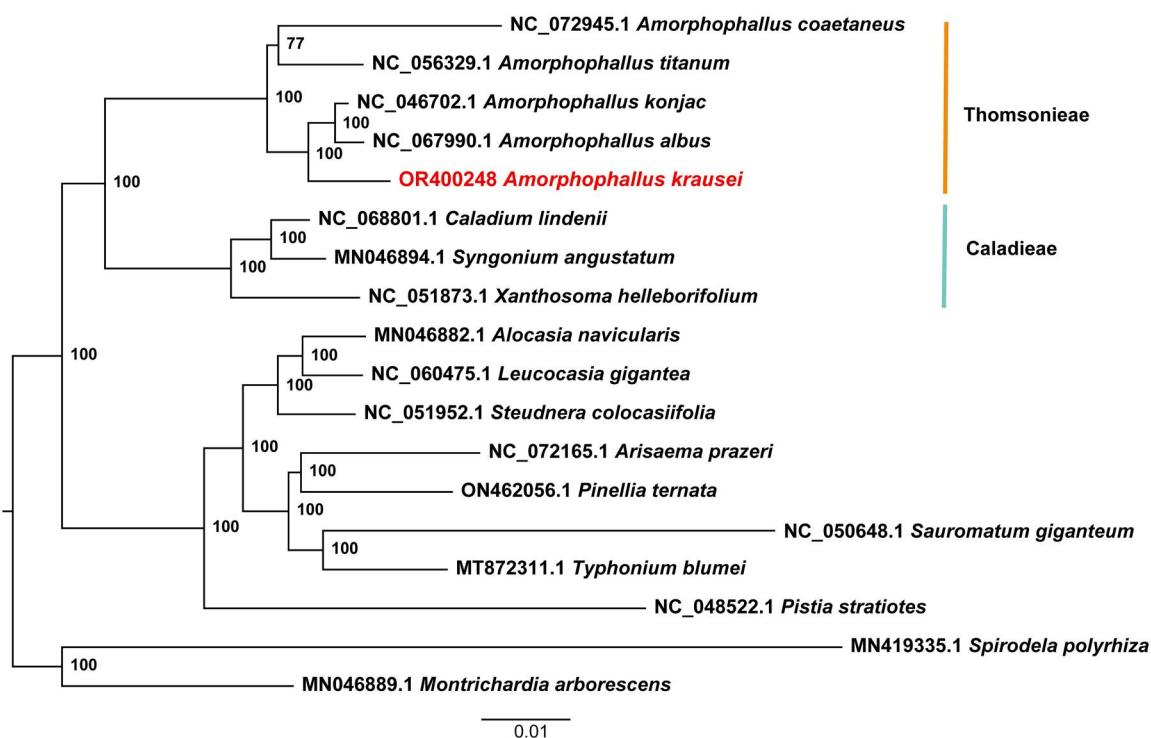
**Figure 2.** The circular map of the chloroplast genome of *Amorphophallus krausei* produced by the software CPGVIEW. Genes belonging to different functional groups are plotted in the outer circle. The functional classification of the genes is shown at the left bottom. The dark gray in the inner circle indicates the GC content of the chloroplast genome. The quadripartite structure, which consists of the LSC, the SSC, and two IR regions, is shown. The inner track shows the forward and reverse repeats connected with red and green arcs.

## Materials and methods

A single individual of *A. krausei* was collected from Xishuangbanna, Yunnan province, China (N 22°01'45", E 101°15'12") (Figure 1). Approximately 5 g of leaf material was stored in a plastic bag with silica gels until DNA extraction. Genomic DNA was isolated using a commercial plant DNA isolation kit (Tiangen, Beijing, China). The sample (yinsi\_XLSD20190726; Si Yin, July 2019) was deposited in the herbarium of Qujing Normal University (Yong Gao, [gaoyong@mail.qjnu.edu.cn](mailto:gaoyong@mail.qjnu.edu.cn)). Whole genome sequencing of *A. krausei* was conducted by Novogene (Beijing, China). A genomic sequencing library of about 350 bp was constructed using established protocols and then sequenced on the NovaSeq 6000 platform (Illumina, California, US), producing 150 bp paired-end (PE) reads.

The chloroplast genome of *A. krausei* was assembled using GetOrganelle v1.7.8.1 with default parameters, except

k-mers were set to 75, 95, 115, and 127 (Jin et al. 2020). The features of the chloroplast genome, such as protein-coding genes, tRNAs, and rRNAs, were annotated with the online software CPGAVAS2 (<http://47.96.249.172:16019/analyzer/home>) and GeSeq (<https://chlorobox.mpimp-golm.mpg.de/geseq.html>) (Tillich et al. 2017; Shi et al. 2019). Annotations were manually checked and adjusted when necessary. The map of the chloroplast genome was drawn using the online tool CPGview (<http://www.1kmpg.cn/cpgview/>) (Liu et al. 2023). The coverage depth was generated using Burrow-Wheeler Aligner (BWA) by aligning sequencing data onto the chloroplast genome (Li and Durbin 2009). The cis-splicing genes and trans-splicing genes were processed using CPGview (Liu et al. 2023). To analyze the phylogenetic position of *A. krausei*, the chloroplast genomes of 17 other species in the Araceae family were downloaded from the NCBI GeneBank database. The



**Figure 3.** Maximum-likelihood phylogenetic tree of *Amorphophallus krausei* and 17 related taxa from the family Araceae. *Spirodela polyrhiza* and *Montrichardia arborescens* are used as outgroups. The scale bar represents the number of substitutions at each locus. Accession numbers: *Amorphophallus coetaneus*, NC\_072945 (Gao et al. 2023); *Amorphophallus titanum*, NC\_056329 (Abdullah et al. 2021); *Amorphophallus konjac*, MK611803 (Hu et al. 2019); *Amorphophallus albus*, NC\_067990 (Shan et al. 2023); *Amorphophallus krausei*, OR400248 (this study); *Caladium lindenii*, NC068801 (reference not available); *Syngonium angustatum*, MN046894 (Henriquez et al. 2020); *Xanthosoma helleborifolium*, NC\_051873 (reference not available); *Alocasia navicularis*, MN046882 (Henriquez et al. 2020); *Leucocasia gigantea*, NC\_060475 (reference not available); *Stuednera colocasifolia*, NC\_051952 (reference not available); *Arisaema prazeri*, NC\_072165 (reference not available); *Pinellia ternata*, ON462056 (Cai et al. 2020); *Sauromatum giganteum*, NC\_050648 (Kim et al. 2020); *Typhonium blumei*, MT872311 (Low et al. 2021); *Pistia stratiotes*, NC\_048522 (Quan and Chen 2020); *Spirodela polyrhiza*, MN419335 (reference not available); *Montrichardia arborescens*, MN046889 (Henriquez et al. 2020).

genome sequences of these 18 species were aligned using mafft v7.475 (Katoh and Standley 2013). The best nucleotide substitution model was identified using ModelFinder (Kalyaanamoorthy et al. 2017). A maximum likelihood (ML) tree was constructed using IQTREE v1.6.12 with 1000 bootstraps, with *Spirodela polyrhiza* and *Montrichardia arborescens* treated as outgroups (Nguyen et al. 2015).

## Results

A total of 8.42 Gb of raw data was produced by next-generation sequencing (NGS), and 8.38 Gb of clean data was retained after filtering reads with low quality. The reads were then used for *de novo* assembly of the chloroplast genome and provided an average coverage depth of  $945 \times$  (Supplementary material, Figure S1). The length of the assembled genome of *A. krausei* was 172,418 bp with an average GC content of 35.23%. This genome sequence has been deposited into the NCBI GenBank database with the accession number OR416863. The *A. krausei* chloroplast genome showed a typical quadripartite structure comprised of one large single-copy (LSC), one small single-copy (SSC), and two inverted repeats (IRs), with lengths of 91,983 bp, 15,591 bp, 32,422 bp, and 32,422 bp, respectively (Figure 2). A total of 132 genes were annotated in the genome, including 86 protein-coding genes, 38 tRNAs, and 8 rRNAs. Among these genes, 15 cis-splicing genes, including *rps16*, *atpF*, *rpoC1*, *ycf3*, *clpP*, *petB*, *petD*, *rpl16*, and of *rpl2* (two copies),

*rpl23* (two copies), *ndhB* (two copies), and *ndhA*, were discovered (Supplementary material, Figure S2). The trans-splicing gene *rps12* had three unique exons (Supplementary material, Figure S3). The ML phylogenetic tree of 18 Araceae species grouped all *Amorphophallus* species into a single monophyletic clade and suggested a close relationship among *A. konjac*, *A. albus*, and *A. krausei*. This phylogeny also indicated that the tribe Thomsonieae is sister to Caladieae (Figure 3).

## Discussion and conclusion

Exploration of excellent genetic resources from wild relatives of cultivated plants is an effective approach for the genetic improvement of cultivars. Species in the genus *Amorphophallus* have newly developed as cash crops and have great economic potential (Gao et al. 2017). A total of 132 genes were found in the chloroplast genome of *A. krausei*, which was similar to previous reports in *A. titanum* (NC\_056329), and *A. coetaneus* (NC\_072945) (Abdullah et al. 2021; Gao et al. 2023). However, the study on chloroplast genomes of four *Amorphophallus* species (*A. albus*, *A. bulbifer*, *A. konjac*, and *A. muelleri*) has reported much fewer gene numbers that ranged from 111 to 113 (Liu et al. 2019). Although the structure of chloroplast genomes is considered conserved, the expansion and contraction of IRs also occur commonly in chloroplast genomes and lead to variation in the number of genes among different species (Abdullah et al. 2020a, 2020b; Ahmed et al. 2012). We observed larger IR



regions of *A. krausei* compared to four *Amorphophallus* species in the previous study. Therefore, IR expansion might contribute to the large number of genes found in this study.

In conclusion, we sequenced and annotated the chloroplast genome of *A. krausei* for the first time. The chloroplast genomic data reported here will support further studies of genetic diversity, resource conservation, evolution, and phylogeographic research of *Amorphophallus*.

## Ethical approval

This study includes no human, animal, or endangered plant samples, and the sampling site was not in the natural reserve. No permissions are needed during the collection of samples.

## Disclosure statement

The authors declare no potential conflict of interest.

## Author contributions

Si Yin performed the molecular experiment and analyzed the data. Yong Gao and Si Yin wrote the manuscript. The authors have revised and approved the final version of this manuscript.

## Funding

This work was supported by the Special Basic Cooperative Research Programs of Yunnan Provincial Undergraduate Universities [202101BA070001-011].

## ORCID

Si Yin  <http://orcid.org/0000-0001-7788-0788>

Yong Gao  <http://orcid.org/0000-0002-0993-231X>

## Data availability statement

The chloroplast genome sequence of *Amorphophallus krausei* was deposited into the NCBI GenBank database under the accession number OR416863. The associated BioProject, Bio-Sample, and SRA numbers are PRJNA1006503, SAMN37041997, and SRR25679975, respectively.

## References

- Abdullah, Henriquez CL, Mehmood F, Carlsen MM, Islam M, Waheed MT, Poczaï P, Croat TB, Ahmed I. 2020a. Complete chloroplast genomes of *Anthurium huixtlense* and *Pothos scandens* (Pothoideae, Araceae): unique inverted repeat expansion and contraction affect rate of evolution. *J Mol Evol.* 88(7):562–574. doi:10.1007/s00239-020-09958-w.
- Abdullah, Henriquez CL, Mehmood F, Hayat A, Sammad A, Waseem S, Waheed MT, Matthews PJ, Croat TB, Poczaï P, Ahmed I. 2021. Chloroplast genome evolution in the *Dracunculus* clade (Aroideae, Araceae). *Genomics.* 113(1 Pt 1):183–192. doi:10.1016/j.ygeno.2020.12.016.
- Abdullah, Henriquez CL, Mehmood F, Shahzadi I, Ali Z, Waheed MT, Croat TB, Poczaï P, Ahmed I. 2020b. Comparison of chloroplast genomes among species of unisexual and bisexual clades of the monocot family Araceae. *Plants.* 9(6):737. doi:10.3390/plants9060737.
- Ahmed I, Biggs PJ, Matthews PJ, Collins LJ, Hendy MD, Lockhart PJ. 2012. Mutational dynamics of aroid chloroplast genomes. *Genome Biol Evol.* 4(12):1316–1323. doi:10.1093/gbe/evs110.
- Cai Z, Wang H, Wang G. 2020. Complete chloroplast genome sequence of *Pinellia ternata* (Thunb.) Breit, a medicinal plant to China. *Mitochondrial DNA B Resour.* 5(3):2107–2108. doi:10.1080/23802359.2020.1765207.
- Gao Y, Dong K, Xiao P, Wu W, Yin S. 2023. Complete assembly of the chloroplast genome of *Amorphophallus coaetaneus* S. Y. Liu & S. J. Wei 1986 (Araceae) from southwestern China. *Mitochondrial DNA B Resour.* 8(7):766–770. doi:10.1080/23802359.2023.2238939.
- Gao Y, Yin S, Wu L, Dai D, Wang H, Liu C, Tang L. 2017. Genetic diversity and structure of wild and cultivated *Amorphophallus paeoniifolius* populations in southwestern China as revealed by RAD-seq. *Sci Rep.* 7(1):14183. doi:10.1038/s41598-017-14738-6.
- Henriquez CL, Abdullah, Ahmed I, Carlsen MM, Zuluaga A, Croat TB, McKain MR. 2020. Evolutionary dynamics of chloroplast genomes in subfamily Aroideae (Araceae). *Genomics.* 112(3):2349–2360. doi:10.1016/j.ygeno.2020.01.006.
- Hu H, Liu J, Wang B, An J, Wang Q. 2019. Characterization of the complete chloroplast genome of *Amorphophallus konjac* (Araceae) and its phylogenetic analysis. *Mitochondrial DNA Part B.* 4(1):1658–1659. doi:10.1080/23802359.2019.1606683.
- Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis 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. doi:10.1186/s13059-020-02154-5.
- Kalyaanamoorthy S, Minh BQ, Wong T, Haeseler AV, Jermin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods.* 14(6):587–589. doi:10.1038/nmeth.4285.
- Katoh K, Standley D. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780. doi:10.1093/molbev/mst010.
- Kim J, Lee J, Choi S, Um S, Kim H, Chun HS, Nah G. 2020. The complete chloroplast genome of *Typhonium giganteum* (Araceae). *Mitochondrial DNA B Resour.* 5(3):2994–2995. doi:10.1080/23802359.2020.1797571.
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 25(14):1754–1760. doi:10.1093/bioinformatics/btp324.
- Li H, Zhu G, Boyce PC, Jin M, Hettterscheid WLA, Bogner J, Jacobsen N. 2010. Araceae. Flora of China. Beijing (China): Science Press.
- Liu E, Yang C, Liu J, Jin S, Harijati N, Hu Z, Diao Y, Zhao L. 2019. Comparative analysis of complete chloroplast genome sequences of four major *Amorphophallus* species. *Sci Rep.* 9(1):809. doi:10.1038/s41598-018-37456-z.
- Liu S, Ni Y, Li J, Zhang X, Yang H, Chen H, Liu C. 2023. CPGView: a package for visualizing detailed chloroplast genome structures. *Mol Ecol Resour.* 23(3):694–704. doi:10.1111/1755-0998.13729.
- Low SL, Yu CC, Ooi IH, Eiadthong W, Galloway A, Zhou ZK, Xing YW. 2021. Extensive Miocene speciation in and out of Indochina: the biogeographic history of *Typhonium sensu stricto* (Araceae) and its implication for the assembly of Indochina flora. *J of Systematics Evolution.* 59(3):419–428. doi:10.1111/jse.12689.
- Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 32(1):268–274. doi:10.1093/molbev/msu300.
- Quan G, Chen L. 2020. Characterization of the complete chloroplast genome sequence of *Pistia stratiotes* (Araceae) and its phylogenetic implications. *Mitochondrial DNA B Resour.* 5(3):2168–2169. doi:10.1080/23802359.2020.1768935.
- Shan Y, Li J, Zhang X, Yu J. 2023. The complete mitochondrial genome of *Amorphophallus albus* and development of molecular markers for five *Amorphophallus* species based on mitochondrial DNA. *Front Plant Sci.* 14:1180417. doi:10.3389/fpls.2023.1180417.
- Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, Liu C. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* 47(W1):W65–W73. doi:10.1093/nar/gkz345.
- Szrednicki G, Borompichaichartkul C. 2020. Konjac glucomannan-production, processing, and functional applications. Boca Raton (FL): CRC Press.
- 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. doi:10.1093/nar/gkx391.