MITOCHONDRIAL DNA PART B: RESOURCES
2024, VOL. 9, NO. 5, 592–596
https://doi.org/10.1080/23802359.2024.2349771

Taylor & Francis
Taylor & Francis Group

PLASTOME REPORT

OPEN ACCESS Check for updates

The complete chloroplast genome assembly of *Amorphophallus tonkinensis* Engler and Gehrmann 1911 from southwestern China

Si Yin*, Huanhuan Chen*, Weijia Wu and Yong Gao

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

ABSTRACT

Species in the *Amorphophallus* genus are important cash crops in many tropical and subtropical Asian countries. Although several molecular markers have been employed to determine relationships and assess genetic variation in the *Amorphophallus* genus, some conflicts remain in infrageneric classification and evolution. To aid in the phylogenetic research of the *Amorphophallus* genus, we collected one sample of *Amorphophallus tonkinensis* Engler and Gehrmann 1911 from southwestern China. We assembled the first chloroplast genome of this species using high-throughput sequencing. The assembled genome was 169,341 bp long with a typical quadripartite structure (GenBank accession number: PP234804). The lengths of the large single-copy region, small single-copy region, and two inverted repeats were 90,705 bp, 15,640 bp, 31,498 bp, and 31,498 bp, respectively. We annotated 129 genes across the chloroplast genome of *A. tonkinensis*. The phylogenetic trees suggested that the *Amorphophallus* species distributed in continental Asia split into two main clades. The chloroplast genome reported in our study provided valuable genomic resources for the future phylogenetic research of the *Amorphophallus* genus.

ARTICLE HISTORY
Received 22 February 2024
Accepted 25 April 2024

KEYWORDS
Chloroplast genome; *Amorphophallus tonkinensis*; Araceae; phylogenetic analysis

Introduction

The plant genus *Amorphophallus*, belonging to the family Araceae, is distributed from tropical Africa throughout subtropical Asia into the tropical western Pacific and northeastern Australia (Li et al. 2010). *Amorphophallus tonkinensis* Engler and Gehrmann 1911 is naturally distributed in the dense tropical forests and moist shaded places with an altitude of 800–900 m in southwestern China, northern Vietnam, and Laos (Srzednicki and Borompichaichartkul 2020). Among all *Amorphophallus* species, *A. tonkinensis* is the only species that could inhabit ever-wet forests (Srzednicki and Borompichaichartkul 2020). The remarkable morphologic feature of this species is the solitary leaf that rises from the tuber, consisting of a vertical petiole and a horizontal leaf blade (Figure 1) (Henriquez et al. 2020a).

Various species of the genus *Amorphophallus* have been used over the centuries in tropical and subtropical Asia as a food source and traditional medicine (Liu 2004). There are estimated to be more than 200 *Amorphophallus* species worldwide (Srzednicki and Borompichaichartkul 2020). Several molecular markers have been employed to determine relationships and assess genetic variation in this genus (Grob et al. 2004; Claudel et al. 2017). Nevertheless, the present understanding of genetic relationships and evolution among *Amorphophallus* species provides only baseline information. There are still some conflicts in infrageneric classification and evolution based on complex morphological features (Srzednicki and Borompichaichartkul 2020).

The chloroplast is an important organelle for photosynthesis in green plants. More and more chloroplast genomes have been revealed with the aid of newly developed sequencing technologies, which enhance our understanding of the chloroplast DNA variation, intracellular gene transfer, and genomic basis of adaptation in plant species (Mehmood et al. 2020). For the low recombination rate and high transferability, chloroplast genomes have been widely utilized in phylogenetic research (Henriquez et al. 2020b). To aid in the phylogeny research of the *Amorphophallus* genus, we reported the first chloroplast genome of *A. tonkinensis* in this study.

Materials and methods

One individual of *A. tonkinensis* was sampled from Hekou County, Yunnan province, China (E 103°51′42.6″, N 22°37′30.6″) during the field investigation of 2022. The specimen has been deposited into the herbarium of the College of Biological Resource and Food Engineering, Qujing Normal University (BSNC_18_Yinsi20220809, Yong Gao, 562698574@

CONTACT Yong Gao 562698574@qq.com College of Biological Resource and Food Engineering, Qujing Normal University, Qujing, Yunnan 655011, China.
*Authors contributed equally to this work.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/23802359.2024.2349771.

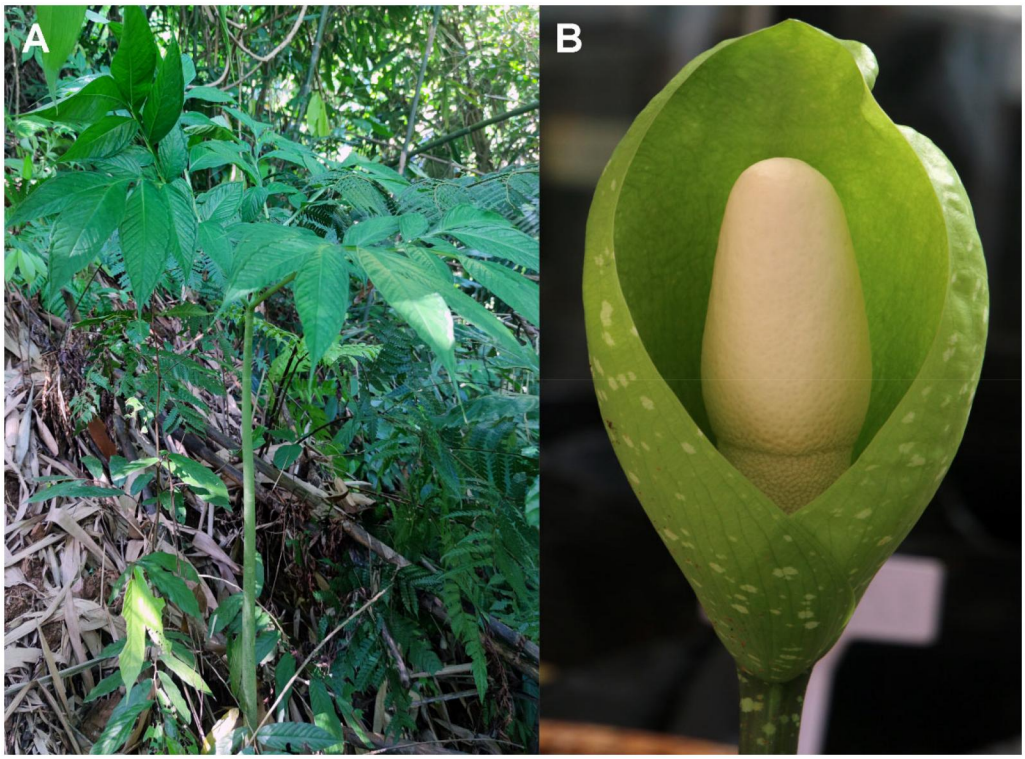

Figure 1. Morphological characteristics of the leaf (A) and flower (B) of *A. tonkinensis*. The photos were taken by the author Yong Gao. The spathe of *A. tonkinensis* is up to 20 cm long, with a shortly convolute base, strongly concave, erect, or arching over spadix apically. The appendix is oval-elliptic, obtuse, white, base with sterile stamens, gradually disappearing upwards. The spadix is sessile, shorter than or nearly as long as spathe.

qq.com, gaoyong@mail.qjnu.edu.cn). Genomic DNA was isolated from 0.3 g of fresh leaves using a modified CTAB protocol (Doyle and Doyle 1987). For the construction of the genome sequencing library, the DNA sample was randomly fragmented by sonication, and fragments with a size of approximately 350 bp were selected and amplified by PCR. The DNA library was sequenced (paired-end 150 bp, PE 150) on the Novaseq 6000 platform (Illumina, CA).

The quality control of raw sequencing reads of *A. tonkinensis* was conducted using the software fastp with default parameters (Chen et al. 2018). We employed GetOrganelle v1.7.8 to assemble the chloroplast genome (Jin et al. 2020). Default parameters were applied except the k-mers (k) were set as 75, 95, 115, and 127, and the maximum extension rounds (R) were set to 40. Two software, Cpgavas 2 (http://47.96.249.172:16019/analyzer/home) and Geseq (https://chlorobox.mpimp-golm.mpg.de/geseq.html) were adopted for the annotation of the chloroplast genome (Michael et al. 2017; Shi et al. 2019). We manually adjusted the annotation results when necessary. The assembly quality was assessed by mapping the sequencing reads against the genome using BWA v0.7.17 (Li and Durbin 2009). The circular map of the chloroplast genome and the detailed transcript structure were drawn using Cpgview (http://www.1kmpg.cn/cpgview/) (Liu et al. 2023). To assess the phylogenetic position of *A. tonkinensis*, we downloaded chloroplast genomes of the *Amorphophallus* species and 11 Acraea species from the GenBank database. The genome sequences were aligned using MAFFT v7.475 (Katoh and Standley 2013). We tested the best nucleotide substitution model of these sequences using ModelFinder (Kalyaanamoorthy et al. 2017). Finally, a maximum likelihood (ML) phylogeny was constructed by IQ-TREE v1.6.12 with 1000 ultra-fast bootstraps (Nguyen et al. 2015).

Results

A total of 10.08 Gb of raw sequencing data was produced for *A. tonkinensis* by high-throughput sequencing. After filtering low-quality reads, 10.04 Gb of clean data was retained for genome assembly. A circular chloroplast DNA molecule with a typical quadripartite structure was assembled using GetOrganelle. The genome sequence has been deposited into the NCBI GenBank database (accession number: PP234804). The assembled genome was 169,341 bp in length with an average GC content of 36%. The lengths of the large single-copy (LSC) region, small single-copy (SSC) region, and two inverted repeats (IRs) were 90,705, 15,640, 31,498, and 31,498 bp, respectively (Figure 2). The average coverage depth of sequencing reads across the genome was 981 (supplemental Figure S1). We found 129 genes (110 unique genes) across the chloroplast genome, including 84 protein-coding genes (77 are unique), eight rRNA genes (four are unique), and 37 tRNA genes (29 are unique). Cpgview detected 13 splicing genes in the chloroplast genomes of *A. tonkinensis* (supplemental Figures S1 and S2). Twenty genes (*rps16*, *atpF*, *rpoC1*, *petB*, *petD*, *rpl16*, *rpl2* (two copies), *ndhA*, *ndhB* (two copies), *trnH-GUG*, *trnG-UCC*, *trnL-UAA*, *trnV-UAC*, *trnI-GAU* (two copies), and *trnA-UGC* (two copies)) have one intron, and four genes (*rps12* (two copies), *ycf3*, and *clpP*) have two introns. According to the Bayesian information criterion, the TVM + F+R3 model was determined as the best-fit

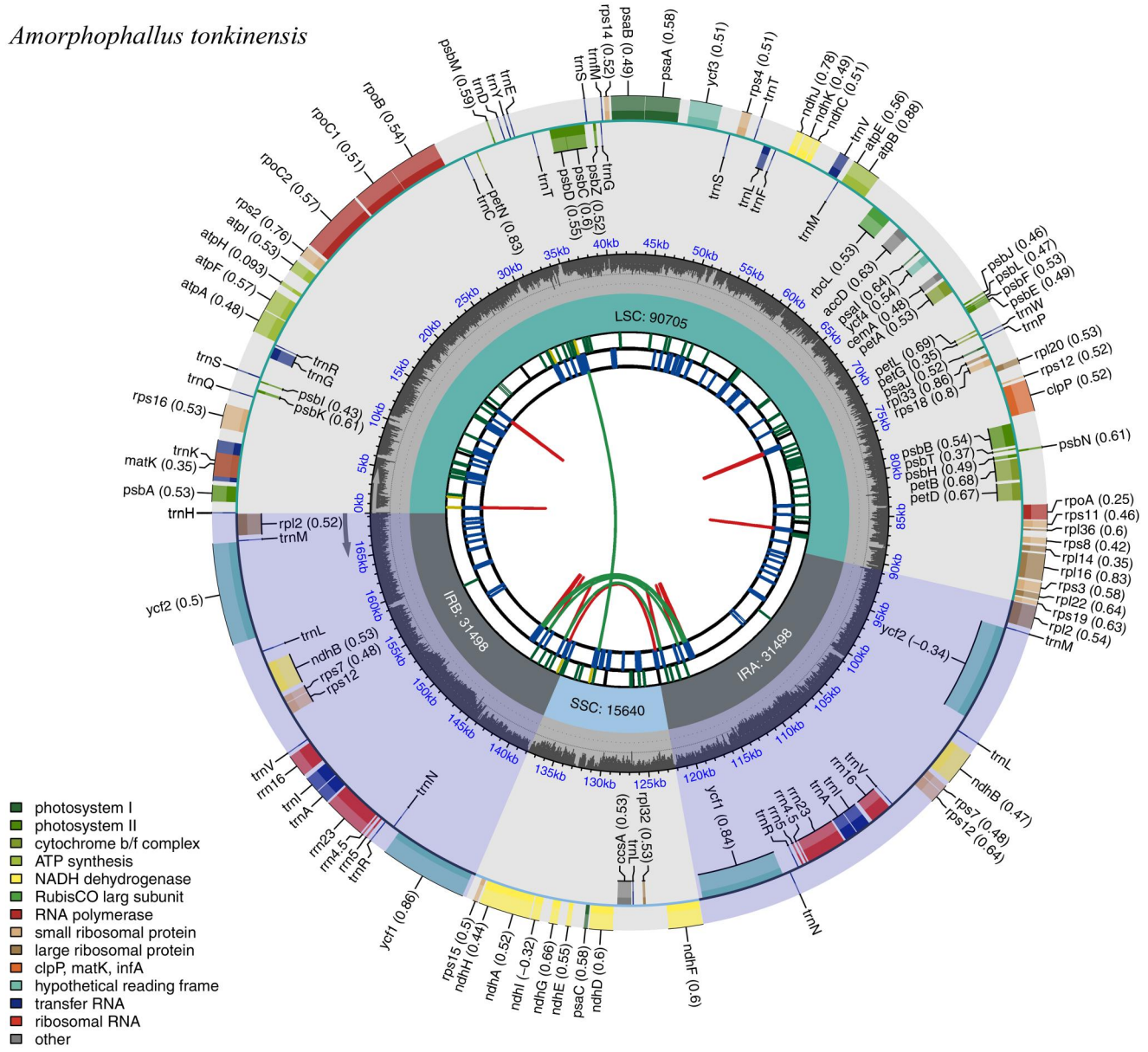

Figure 2. The circular map of the chloroplast genome of *A. tonkinensis*. Genes belonging to different functional groups are plotted in the outer circle. The quadripartite structure, which consists of the LSC, the SSC, and two IR regions, is shown. The dark gray in the inner circle indicates the GC content of the chloroplast genome.

nucleotide substation model. The ML tree supported all nine *Amorphophallus* species forming into one clade with a bootstrap of 100 (Figure 3).

Discussion and conclusion

At present, chloroplast genomes of several *Amorphophallus* species have been sequenced, which gives insight into the chloroplast genomic characteristics of these species (Liu et al. 2019; Li et al. 2024). A typical quadripartite structure and 129 genes are detected across the chloroplast genome of *A. tonkinensis*, which is consistent with the findings in chloroplast genomes of other *Amorphophallus* species (Liu et al. 2019; Yin and Gao 2023a, 2023b; Li et al. 2024). However, we find large differences in the chloroplast genome lengths of these *Amorphophallus* species, ranging from 161,647 to 176,835 bp. The contraction/expansion of IRs usually causes the length variation of the chloroplast genome (Li et al. 2024). With limited genomes available, mechanisms underlying the structural variation in the chloroplast genomes of the *Amorphophallus* genus still need further investigation. The chloroplast genome produced in our study provides additional genomic resources for the upcoming evolution research of the chloroplast genomes in this genus.

Several phylogenetic studies of *Amorphophallus* species have been conducted using chloroplast and nuclear DNA markers (Grob et al. 2004). Claudel et al. (2017) constructed the phylogeny of 157 *Amorphophallus* species using nuclear (*ITS*1) and plastid (*rbcL* and *matK*) regions. The phylogenetic trees suggested that species distributed in continental Asia split into two main clades (subgenus *Metandrium* and *Scutandrium*). Eight *Amorphophallus* species in our

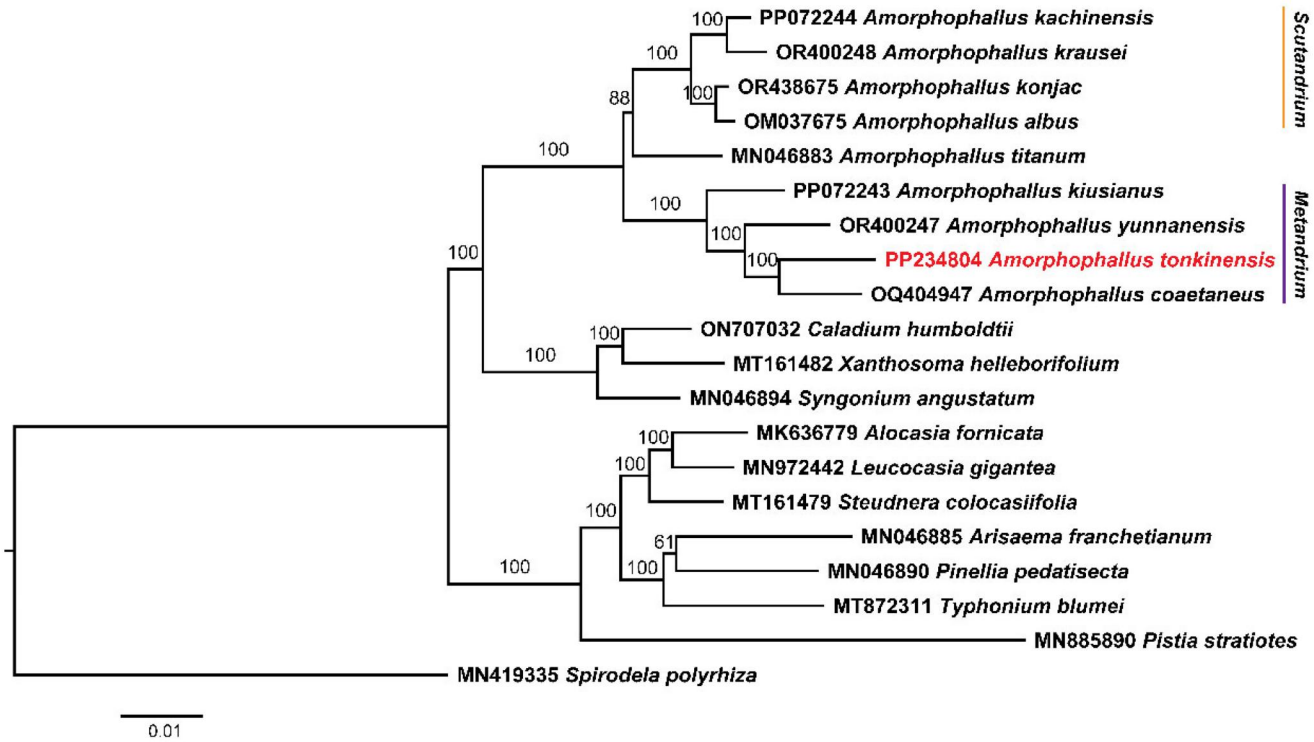

Figure 3. Maximum-likelihood phylogeny of *A. tonkinensis* and 19 related taxa in the family Araceae. *Spirodela polyrhiza* is used as an outgroup. The scale bar represents the number of substitutions at each locus. Accession numbers: *Amorphophallus tonkinensis*, PP234804 (this study); *Amorphophallus kiusianus*, PP072243 (reference not available); *Amorphophallus krausei*, OR400248 (Yin and Gao 2023b); *Amorphophallus konjac*, OR438675 (Li et al. 2024); *Amorphophallus albus*, OM037675 (Shan et al. 2023); *Amorphophallus titanium*, MN046883 (Henriquez et al. 2021); *Amorphophallus kiusianus*, PP072243 (reference not available); *Amorphophallus yunnanensis*, OR400247 (Yin and Gao 2023a); *Amorphophallus coaetaneus*, OQ404947 (Gao et al. 2023); *Caladium humboldtii*, ON707032 (reference not available); *Xanthosoma helleborifolium*, MT161482 (reference not available); *Syngonium angustatum*, MN046894 (Henriquez et al. 2020a); *Alocasia fornicata*, MN636779 (reference not available); *Leucocasia gigantea*, MN972442 (reference not available); *Steudnera colocasiifolia*, MT161479 (reference not available); *Arisaema franchetianum*, MN046885 (Henriquez et al. 2020a); *Pinellia pedatisecta*, MN046890 (Henriquez et al. 2020a); *Typhonium blumei*, MT872311 (Low et al. 2021); *Spirodela polyrhiza*, MN419335 (reference not available).

phylogenetic study were distributed in continental Asia except for *A. titanium* belonging to the Southeast Asia clade. Our study supported the divergence of the continental Asia clade with a high support value (Figure 3). However, we cannot resolve the relationship between the continental Asia and Southeast Asia groups with only one species in the Southeast Asia clade. A clearer phylogeny of *Amorphophallus* will be revealed with more chloroplast genomes of *Amorphophallus* species available in the future.

In conclusion, we sequenced and assembled the first chloroplast genome of *A. tonkinensis* using next-generation sequencing in this study. The chloroplast genome provides valuable genomic resources for the future phylogenetic research of the *Amorphophallus* genus.

Author contributions

Yong Gao conceived the study. Yong Gao and Weijia Wu collected the sample. Huanhuan Chen, Si Yin, and Weijia Wu conducted the molecular experiment and analyzed the data. Si Yin, Huanhuan Chen, and Yong Gao wrote the paper. All authors have approved the final version of this manuscript.

Ethical approval

This study includes no human, animal, or endangered plant species, and the sampling site was not in the natural reserve. No permissions are needed during the collection of plant material. The ethical approval is granted by the ethics committee of Qujing Normal University.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Scientific Research Found Project of the Yunnan Provincial Education Commission [2024J0937], and the Special Basic Cooperative Research Programs of Yunnan Provincial Undergraduate Universities [202101BA070001-011].

Data availability statement

The chloroplast genome assembly of *A. tonkinensis* was deposited into the NCBI GenBank database with accession number PP234804. The associated BioProject, Bio-Sample, and SRA numbers are PRJNA1070842, SAMN39664789, and SRR27775148, respectively.

References

Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 34(17):i884–i890. doi:10.1093/bioinformatics/bty560.

Claudel C, Buerki S, Chatrou LW, Antonelli A, Alvarez N, Hetterscheid W. 2017. Large-scale phylogenetic analysis of *Amorphophallus* (Araceae) derived from nuclear and plastid sequences reveals new subgeneric

delineation. Bot J Linn Soc. 184(1):32–45. doi:10.1093/botlinnean/box013.

Doyle J, Doyle J. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 19(1):11–15.

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.

Grob GBJ, Gravendeel B, Eurlings MCM. 2004. Potential phylogenetic utility of the nuclear *FLORICAULA/LEAFY* second intron: comparison with three chloroplast DNA regions in *Amorphophallus* (Araceae). Mol Phylogenet Evol. 30(1):13–23. doi:10.1016/s1055-7903(03)00183-0.

Henriquez CL, Ahmed I, Carlsen MM, Zuluaga A, Croat TB, McKain MR, Abdullah. 2020a. Evolutionary dynamics of chloroplast genomes in subfamily Aroideae (Araceae). Genomics. 112(3): 2349–2360. doi:10.1016/j.ygeno.2020.01.006.

Henriquez CL, Ahmed I, Carlsen MM, Zuluaga A, Croat TB, McKain MR, Abdullah. 2020b. Molecular evolution of chloroplast genomes in Monsteroideae (Araceae). Planta. 251(3): 72. doi:10.1007/s00425-020-03365-7.

Henriquez CL, Mehmood F, Hayat A, Sammad A, Waseem S, Waheed MT, Matthews PJ, Croat TB, Poczai P, Ahmed I, Abdullah. 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.

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, Jermiin 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.

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 L, Qi Y, Gao P, Yang S, Zhao Y, Guo J, Liu J, Huang F, Yu L. 2024. The complete chloroplast genome sequence of *Amorphophallus konjac* (Araceae) from Yunnan, China and its phylogenetic analysis in the family Araceae. Mitochondrial DNA B Resour. 9(1):41–45. doi:10.1080/23802359.2023.2300471.

Liu PY. 2004. Konjac biology. Beijing: China Agriculture Press.

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.

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.

Li H, Zhu G, Boyce PC, Jin M, Hetterscheid WLA, Bogner J, Jacobsen N. 2010. Araceae. Flora of China. Beijing: Science Press.

Low SL, Yu C-C, Ooi IH, Eiadthong W, Galloway A, Zhou Z-K, Xing Y-W. 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 Sytematics Evolution. 59(3):419–428. doi:10.1111/jse.12689.

Mehmood F, Shahzadi I, Ahmed I, Waheed MT, Mirza B, Abdullah. 2020. Characterization of *Withania somnifera* chloroplast genome and its comparison with other selected species of Solanaceae. Genomics. 112(2): 1522–1530. doi:10.1016/j.ygeno.2019.08.024.

Michael T, Pascal L, Tommaso P, Ulbricht-Jones ES, Axel F, Ralph B, Stephan G. 2017. GeSeq – versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 45(W1):W6–W11. doi:10.1093/nar/gkx391.

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.

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(1):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.

Srzednicki G, Borompichaichartkul C. 2020. Konjac glucomannan-production, processing, and functional applications. Boca Raton: CRC Press.

Yin S, Gao Y. 2023a. Characterization of the complete chloroplast genome assembly of *Amorphophallus yunnanensis* Engler, Pflanzenr (Araceae) from southwestern China. Mitochondrial DNA B Resour. 8(12):1445–1449. doi:10.1080/23802359.2023.2294896.

Yin S, Gao Y. 2023b. The complete chloroplast genome assembly of *Amorphophallus krausei* Engler, Pflanzenr 1911 (Araceae) from southwestern China. Mitochondrial DNA B Resour. 8(12):1339–1342. doi:10.1080/23802359.2023.2288889.