



## Report

## Molecular analyses display the increasing diversity of Podostemaceae in China

Zhi-Zhong Li<sup>a, b, \*, 1</sup>, Zhun Xu<sup>b, c, 1</sup>, Shuang Wu<sup>d</sup>, Lang-Xing Yuan<sup>e</sup>, Chun-Yu Zou<sup>f</sup>, Yan Liu<sup>f</sup>, Jian-Yong Lin<sup>g</sup>, Shi-Chu Liang<sup>a, \*\*, 1</sup><sup>a</sup> Key Laboratory of Ecology of Rare and Endangered Species and Environmental Protection (Guangxi Normal University), Ministry of Education, Guilin 541006, China<sup>b</sup> Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China<sup>c</sup> Sino-Africa Joint Research Center, Chinese Academy of Sciences, Wuhan 430074, China<sup>d</sup> Guangxi Association for Science and Technology, Nanning 530023, China<sup>e</sup> Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou 571101, China<sup>f</sup> Guangxi Institute of Botany, Chinese Academy of Sciences, Guilin 541006, China<sup>g</sup> Guangxi Forestry Research Institute, Nanning 530028, China

## ARTICLE INFO

## Article history:

Received 5 September 2023

Received in revised form

29 January 2024

Accepted 2 February 2024

Available online 7 February 2024

Podostemaceae (riverweed family), one of the largest aquatic plant families in the world, comprises about 350 species in 51 genera placed in three subfamilies, Podostemoideae, Weddellinoideae, and Tristichioideae (Koi et al., 2022). Most Podostemaceae genera (~38) are monotypic or oligotypic, with a high level of endemism (Koi et al., 2018, 2022). Members of Podostemaceae adhere to rocks in fast-flowing rivers, streams, or waterfalls in the subtropics and tropics (Koi et al., 2018). As an adaptation to fast-running currents, the morphological structure of river-weeds tends to be extremely reduced, including a loss or reduction of primary shoots or roots (Koi and Kato, 2007). The vegetative mass of these plants resembles lichens or bryophytes, and are thus known commonly as “dream bryophytes” (Philbrick and Retana, 1998). Podostemaceae plants flower when the water level drops during the dry season. In recent years, anthropogenic factors and climate change have severely affected river weed habitats, decreasing populations of these plants. Most species within the

family are under threat, with 25 species (e.g., *Lebbiea grandiflora* Cheek, *Winklerella dichotoma* Engl.) designated as Critically Endangered on the IUCN Red List (IUCN, 2022).

Recent discoveries indicate that Podostemaceae species diversity in China may be seriously underestimated. About 80 species of Podostemaceae in 19 genera from two subfamilies (Podostemoideae and Tristichioideae) have been recorded in Asia (Kato, 2018). In China, the major reference on plants lists only four species of Podostemaceae from *Cladopus* H. Möller, *Hydrobryum* Endl., and *Terniopsis* H.C. Chao (Qiu and Philbrick, 2003). However, the list of known Podostemaceae species in China has recently increased. Two new species (*C. yinggelingsensis* and *Terniopsis daoyinensis*; Lin et al., 2016) and one newly recorded genus (*Paracladopus* M. Kato; Wu et al., 2022) were reported from Yinggeling National Nature Reserve and Wuzhishan Tropical Rainforest Scenic Area in Hainan Province, respectively. More recently, two novel species (*Terniopsis yongtaiensis* and *Polypleurum chinense* B.Hua Chen & Miao Zhang) were discovered in Fujian Province (Chen et al., 2022; Zhang et al., 2022). To date, 11 species in five genera have been recorded from Yunnan, Guangdong, Fujian, Hainan, Guizhou, and Hong Kong (Chen et al., 2022). The conservation status of most of these newly recorded species has been assessed as Grade II on the List of National Key Protected Wild Plants (Lu et al., 2021). The underestimation of Podostemaceae species diversity can be attributed to the ease of overlooking tiny plant species. To improve our understanding of river-weed diversity in China, additional fieldwork should be conducted in potential distribution regions of the family, such as Guangxi Province.

Previous studies have examined morphology (Moline et al., 2007; Khanduri et al., 2015), developmental biology (Koi and Kato, 2007), phylogeny and biogeography (Koi et al., 2018), and adaptive evolution (Katayama et al., 2022) of Podostemaceae. However, the extremely reduced and modified morphological traits of these species often confound species identification and challenge

\* Corresponding author. Key Laboratory of Ecology of Rare and Endangered Species and Environmental Protection (Guangxi Normal University), Ministry of Education, Guilin 541006, China.

\*\* Corresponding author.

E-mail addresses: [lizhizhong@wbgcas.cn](mailto:lizhizhong@wbgcas.cn) (Z.-Z. Li), [gxlsc@sina.com](mailto:gxlsc@sina.com) (S.-C. Liang).

Peer review under responsibility of Editorial Office of Plant Diversity.

<sup>1</sup> Authors contributed equally to this work.

taxonomic research. One approach that has been widely used to delimit species and analyze phylogenetic relationships is DNA barcoding (Koi et al., 2022). Previous research has indicated that the plastid *matK* sequence is the optimum region for DNA barcoding in Podostemaceae (Kelly et al., 2010). In addition to traditional barcoding, plastomes have been used as “super barcodes” to assist species identification in Podostemaceae (Bedoya et al., 2019; Mwanzia et al., 2020; Chen et al., 2022; Wu et al., 2022; Zhang et al., 2022).

During field investigations from 2020 to 2022, several Podostemaceae populations, including *Hydrobryum*, *Terniopsis* and *Cladopus*, were discovered by a group of aquatic plant enthusiasts and local researchers in Guangxi and Hainan Province, China (Fig. S1). However, owing to the extremely reduced morphological traits of the species in this family, it has not been easy to identify them accurately by morphological evidence alone. In this study, we used molecular data (i.e., *matK* sequence and plastome sequences) as barcodes to classify and determine phylogenetic relationships of 15 newly discovered populations (Fig. S1) and two reported species, *T. daoyinensis* and *C. yinggelingensis*, were also collected from Hainan Province for subsequent analyses.

All 15 newly discovered populations were initially identified by morphology according to Qiu and Philbrick (2003), Kato (2006), (Koi et al. 2018, 2022), Lin et al. (2016), Wu et al. (2022), and Zhang et al. (2022). However, several populations were difficult to identify accurately. Thus, 1–2 silica gel-dried individuals were randomly selected from each population for subsequent molecular phylogenetic analysis.

Approximately 3 Gb of genome skimming data for each sample were generated after trimming and filtering by Fastp v.0.23.2 (Chen et al., 2018) with default settings. Then, whole plastomes were assembled *de novo* using GetOrganelle v.1.7.7 (Jin et al., 2020) with the word size for extension (-w) from 0.6 to 0.9. We annotated the assembled complete plastomes and manually adjusted the start/stop codons in Geneious v.5.6.4 using *Terniopsis yongtaiensis* (OM717943; Zhang et al., 2022) as a reference.

In this study, we generated two data sets for phylogenetic analysis. The first data set consisted of *matK* sequences from the 17 newly sequenced Podostemaceae plastomes, three available plastomes of Asian Podostemaceae [*Polypleurum chiangmaiensis* (NC\_061663), *P. chinense* (NC\_070353) and *T. yongtaiensis* (NC\_066797)], and 81 plastome sequences retrieved from GenBank, including those of *Cladopus*, *Terniopsis*, *Paracladopus*, *Hydrobryum*, *Weddellina* Tul., and *Polypleurum* (Table S1). The outgroups were *Clusia cruiva* Cambess. (AB450037) and *Hypericum calycinum* L. (AB698446). The second data set was generated to determine whether phylogenetic analysis based on *matK* was consistent with that based on plastome data. For this data set, we sequenced common protein-coding genes from all available plastomes in Podostemaceae. The outgroups were *Cratogeomys pruniflorum* (Kurz) Kurz (NC\_062805) and *Hypericum ascyron* Mill. (MZ424306). Each data set was aligned using MAFFT v.7.310 (Katoh et al., 2002) with the “l–INS–i” Strategy. Maximum Likelihood (ML) analysis was carried out in IQTREE v.2.2.0 (Minh et al., 2020) with 1000 bootstrap replicates (BS) under the auto-selected best nucleotide substitution model (TVM + F + G4). Bayesian Inference (BI) was carried out in MrBayes v.3.2.7 (Ronquist et al., 2012) with three million generations, sampling trees every 1000 generations. The first 25% of trees were discarded as “burn-in”, and the remaining trees were utilized to estimate the posterior probabilities (PP).

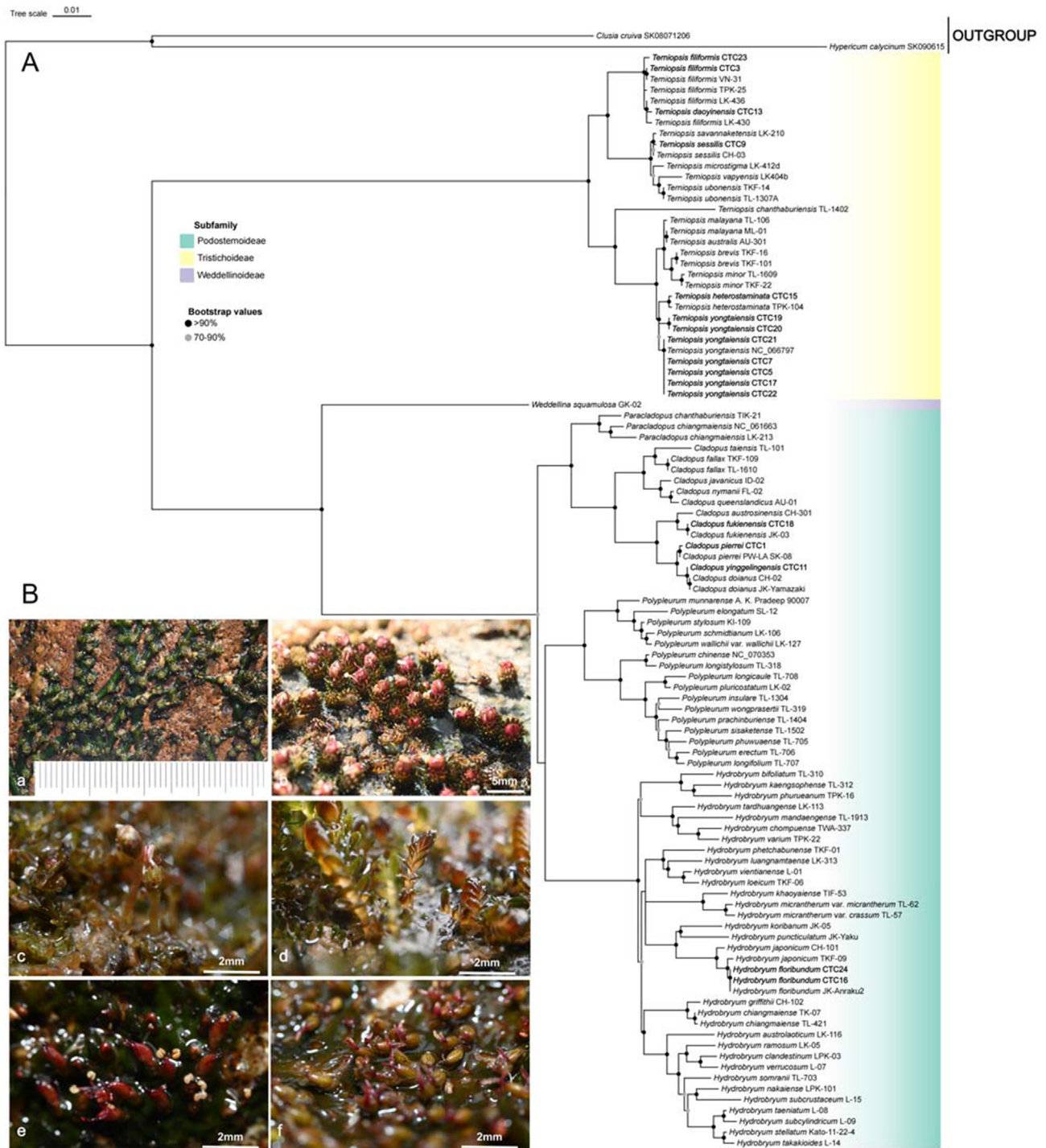
For comparative analysis of Podostemaceae plastomes, we used plastomes of 15 species representing eight genera, and *Byrsonima crassifolia* (L.) Kunth (NC\_037192) and *Garcinia mangostana* L. (NC\_036341) were used as references. Structural variations of these

plastomes were identified in Mauve (Darling et al., 2010) using default parameters. We used CPJSDraw v.1.1.0 (Li et al., 2023) to analyze the genes at the IR boundaries of junction locations (LSC/IR and SSC/IR) in each plastome. To identify hotspots that might serve as novel DNA barcodes of Podostemaceae, we aligned protein-coding sequences from all 23 available Podostemaceae plastomes in MAFFT v.7.310 with default settings. Nucleotide variability ( $P_i$ ) was calculated for each protein-coding sequence using the software DnaSP v.6.12.03 (Rozas et al., 2017).

We assembled and annotated 17 individual plastomes from six species of Podostemaceae (Table S2). All plastomes had a typical quadripartite structure, with conserved size (~130 kb), GC content (~36%), and number of genes (~107), which is consistent with results from previous studies (Table S2; Bedoya et al., 2019; Chen et al., 2022; Wu et al., 2022). Comparative plastome analyses indicated that the *ycf1* and *ycf2* genes were lost or pseudogenized, consistent with all Podostemaceae species, probably resulting in reduced IR regions (Weng et al., 2014). All the newly sequenced plastomes have a large inversion (~50 kb) with about 50 genes from *rbcl* to *trnK* in the LSC regions (Fig. S2), which might be attributed to frequent homologous recombination between repeats (Park et al., 2018). Except for *P. chinense*, the IR boundaries of all Podostemaceae plastomes were identical at JSB (LSC/IRb) and JLA (IRa/LSC), which fall within the genes *rps19* and *trnH*, respectively (Fig. S3). In the *P. chinense* plastome, JLA is located in the *psbA* gene. The JSB (IRb/SSC) and JSA (SSC/IRa) of Podostemaceae plastomes differ in each genus. Specifically, the JSB of all *Terniopsis* plastomes is located in the *ndhF* gene, but in all plastomes of Podostemoideae is located between *rps15* and *ndhF* (Fig. S3). In addition, we found high polymorphism ( $P_i > 0.09$ ) in three protein-coding sequences (*matK*, *ccsA*, and *ndhA*; Fig. S4). This finding is consistent with those of previous studies that indicated *matK* ( $P_i = 0.116$ ) may serve as a barcode for taxa classification in Podostemaceae (Kelly et al., 2010). Alternative barcodes that may be used in studies of population genetics and cryptic diversity in Podostemaceae include *ccsA* and *ndhA*.

The tree topologies based on ML and BI methods were consistent, generating two robust clades corresponding to two subfamilies, which is also in line with the *matK* results (Fig. S5). In both ML and BI analyses based on the *matK* data set, all branches showed high support values and clustered into three distinct clades, corresponding to three subfamilies with robust support (Fig. 1). In agreement with previous study (Koi et al., 2022), Tristchoideae was resolved as a sister to Weddellinoideae and Podostemoideae, with the highest support (BS = 100, PP = 1.00; Fig. 1 and S5). Most newly collected taxa can be identified reasonably based on molecular phylogenetic evidence and supported by morphological evidence. For example, our molecular data indicate that *Cladopus fukienensis* CTC18 and JK-03 form a clade sister to *C. austrosinensis* (Fig. 1). This is also supported by morphological evidence (i.e., bracts 12–20 in CTC18 vs. bracts 10 or fewer in *C. austrosinensis*; Fig. 1). There is one exception: *T. daoyinensis* CTC13 groups with *Terniopsis filiformis* LK-430 and other *T. filiformis* accessions forming a paraphyly. *T. daoyinensis* CTC3 was resolved as a sister to *T. filiformis* VN-31; CTC23 sister to all *T. filiformis* accessions and *T. daoyinensis* CTC13. However, in phylogenomic analysis based on plastid protein-coding sequences (Fig. S3), CTC3 and CTC23 form a clade sister to *T. daoyinensis* CTC13. Therefore, we here update the key to species of Podostemaceae in China (Supplementary data) and recommend that further investigations aim to resolve taxonomic relationships between these two species. In total, our analysis reveals four newly recorded species: *Terniopsis heterostaminata*, *T. filiformis*, *Hydrobryum floribundum* and *Cladopus pierreii*.

The newly recorded species here have been previously reported as endemic to Mainland Southeast Asia (i.e., Thailand, Laos, and



**Fig. 1.** Maximum likelihood phylogenetic inference of Podostemaceae using *matK* region (A) and newly recorded species of Podostemaceae were collected by Shuang Wu during field investigations in Guangxi from 2020 to 2022 (B). Only bootstrap support values (BS) > 70% are shown for nodes. a and b, *Cladopus pierrei*: a. Rhizomes; b. Flowers. c and d, *Terniopsis filiformis*: c. Flowers; d. Ramuli. e and f, *Hydrobryum floribundum*: flowers on thallus.

Vietnam). *T. yongtaiensis* was recorded for the first time in Guangxi Province, China, which is far from the only locality in Fujian Province previously reported by Zhang et al. (2022). Furthermore, this study indicates that the flora of Guangxi and Hainan provinces in China are closely related to those of Mainland Southeast Asia, which both contribute to the Indo-Burma hotspot in Southeast Asia (Tan et al., 2020). Importantly, our new findings reveal that the

species diversity of Podostemaceae has been underestimated in southwestern China.

Our field investigations revealed that most populations of Podostemaceae are distributed in karst streams on the edges of villages. In these areas, water quality and seasonal changes in water levels may be impacted by anthropogenic activities such as agriculture and dam construction (Personal communication with local

people; Zhang et al., 2022). Environmental changes to these habitats increase algae levels (e.g., *Spirogyra*) on the surfaces of stones, which seriously affects the natural growth of Podostemaceae. All species of two genera – *Cladopus* and *Terniopsis* – and a single species of *Hydrobryum* (*H. griffithii*) have been placed under national Grade II protection on the Updated List of National Key Protected Wild Plants (Decree No. 15) in China. However, most of the newly recorded species are not protected. Furthermore, potential populations remain to be discovered, which poses a great challenge for conservation work on Podostemaceae. Future studies aimed at characterizing Podostemaceae diversity should focus on southwestern China and include new approaches such as environmental DNA (eDNA; Tsukamoto et al., 2021). Moreover, the conservation status should be determined for these newly recorded species, and the List of National Key Protected Wild Plants should be promptly updated. The unique habitats (i.e., flowing water) and life cycle (i.e., dependent on seasonal water level changes) of Podostemaceae require *in situ* conservation strategies.

### Declaration of competing interest

The authors declare that they have no conflict of interest.

### CRediT authorship contribution statement

**Zhi-Zhong Li:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Zhun Xu:** Writing – original draft, Formal analysis. **Shuang Wu:** Investigation, Data curation. **Lang-Xing Yuan:** Resources, Investigation, Formal analysis. **Chun-Yu Zou:** Investigation, Formal analysis. **Yan Liu:** Investigation, Formal analysis. **Jian-Yong Lin:** Investigation, Formal analysis. **Shi-Chu Liang:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Conceptualization.

### Acknowledgements

The authors thank Mr. Bin Zhang and Zai-Ji Xiang for their field assistance. This study was supported by grants from the National Natural Science Foundation of China (Nos. 32100186 and 32300182) and the Key Laboratory of Ecology of Rare and Endangered Species and Environmental Protection (Guangxi Normal University), Ministry of Education, China (No. ERESEP2022K03).

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.pld.2024.02.002>.

### References

- Bedoya, A.M., Ruhfel, B.R., Philbrick, C.T., et al., 2019. Plastid genomes of five species of riverweeds (Podostemaceae): structural organization and comparative analysis in Malpighiales. *Front. Plant Sci.* 10, 1035.
- Chen, B.H., Zhang, M., Zhao, K., et al., 2022. *Polypleurum chinense* (Podostemaceae), a new species from Fujian, China, based on morphological and genomic evidence. *PhytoKeys* 199, 167–186.
- Chen, S., Zhou, Y., Chen, Y., et al., 2018. fastp: an ultra-fast all-in-one FASTQ pre-processor. *Bioinformatics* 34, i884–i890.
- Darling, A.E., Mau, B., Perna, N.T., 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5, e11147.
- IUCN, 2022. The IUCN Red List of Threatened Species. Version 2022-2. <https://www.iucnredlist.org>. (Accessed 31 July 2023).
- Jin, J.J., Yu, W.B., Yang, J.B., et al., 2020. GetOrganelle: a fast and versatile toolkit for accurate *de novo* assembly of organelle genomes. *Genome Biol.* 21, 1–31.
- Katayama, N., Koi, S., Sassa, A., et al., 2022. Elevated mutation rates underlie the evolution of the aquatic plant family Podostemaceae. *Commun. Biol.* 5, 75.
- Kato, M., 2006. Distribution and biogeography of Podostemaceae in Asia. *Bull. Natl. Sci. Mus.* 32, 19–27.
- Kato, M., 2018. Podostemaceae. In: Chayamarit, K., Balslev, H. (Eds.), *Flora of Thailand*, 14. Forest Herbarium, Department of National Parks, Wildlife and Plant Conservation, Bangkok, pp. 68–114.
- Katoh, K., Misawa, K., Kuma, K.I., et al., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066.
- Kelly, L.J., Ameka, G.K., Chase, M.W., 2010. DNA barcoding of African Podostemaceae (river-weeds): a test of proposed barcode regions. *Taxon* 59, 251–260.
- Khanduri, P., Tandon, R., Uniyal, P.L., et al., 2015. Comparative morphology and molecular systematics of Indian Podostemaceae. *Plant Syst. Evol.* 301, 861–882.
- Koi, S., Kato, M., 2007. Developmental morphology of the shoot in *Weddellina squamulosa* and implications for shoot evolution in the Podostemaceae. *Ann. Bot.* 99, 1121–1130.
- Koi, S., Uniyal, P.L., Kato, M., 2022. A classification of the aquatic Podostemaceae subfamily Tristichoideae, with a new genus based on *ITS* and *matK* phylogeny and morphological characters. *Taxon* 71, 307–320.
- Koi, S., Won, H., Trần, H., et al., 2018. Molecular and morphological variation in *Terniopsis* (Podostemaceae) show contrasting patterns. *Nord. J. Bot.* 36, e01872.
- Li, H., Guo, Q., Xu, L., et al., 2023. CPJSDraw: analysis and visualization of junction sites of chloroplast genomes. *PeerJ* 11, e15326.
- Lin, Q.W., Lu, G., Li, Z.Y., 2016. Two new species of Podostemaceae from the Yinggeling National Nature Reserve, Hainan, China. *Phytotaxa* 270, 49–55.
- Lu, Z., Qin, H., Jin, X., et al., 2021. On the necessity, principle, and process of updating the list of national key protected wild plants. *Biodivers. Sci.* 29, 1577.
- Minh, B.Q., Schmidt, H.A., Chernomor, O., et al., 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534.
- Moline, P., Thiv, M., Ameka, G.K., et al., 2007. Comparative morphology and molecular systematics of African podostemaceae-podostemoideae, with emphasis on *Dicraeanthus* and *Ledermannia* from Cameroon. *Int. J. Plant Sci.* 168, 159–180.
- Mwanzia, V.M., He, D.X., Gichira, A.W., et al., 2020. The complete plastome sequences of five *Aponogeton* species (Aponogetonaceae): insights into the structural organization and mutational hotspots. *Plant Divers.* 42, 334–342.
- Park, S., An, B., Park, S., 2018. Reconfiguration of the plastid genome in *Lampirocapnos spectabilis*: IR boundary shifting, inversion, and intraspecific variation. *Sci. Rep.* 8, 13568.
- Philbrick, C.T., Retana, A.N., 1998. Flowering phenology, pollen flow, and seed production in *Marathrum rubrum* (Podostemaceae). *Aquat. Bot.* 62, 199–206.
- Qiu, H., Philbrick, T.C., 2003. Podostemaceae. In: Wu, C.Y., Raven, P.H., Hong, D.Y. (Eds.), *Flora of China*, 5. Science Press, Beijing and Missouri Botanical Garden Press, St. Louis, pp. 190–191.
- Ronquist, F., Teslenko, M., Van Der Mark, P., et al., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542.
- Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J.C., et al., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34, 3299–3302.
- Tan, K., Malabrigo Pastor, L., Ren, M.X., 2020. Origin and evolution of biodiversity hotspots in southeast Asia. *Acta Ecol. Sin.* 40, 3866–3877.
- Tsukamoto, Y., Yonezawa, S., Katayama, N., et al., 2021. Detection of endangered aquatic plants in rapid streams using environmental DNA. *Front. Ecol. Evol.* 8, 622291.
- Weng, M.L., Blazier, J.C., Govindu, M., et al., 2014. Reconstruction of the ancestral plastid genome in Geraniaceae reveals a correlation between genome rearrangements, repeats, and nucleotide substitution rates. *Mol. Biol. Evol.* 31, 645–659.
- Wu, M., Zhang, K., Yang, X., et al., 2022. *Paracladopus chiangmaiensis* (Podostemaceae), a new generic record for China and its complete plastid genome. *PhytoKeys* 195, 1–13.
- Zhang, M., Zhang, X.H., Ge, C.L., et al., 2022. *Terniopsis yongtaiensis* (Podostemaceae), a new species from South East China based on morphological and genomic data. *PhytoKeys* 194, 105–122.