

The complete chloroplast genome sequence of *Hypecoum erectum* L. (Papaveraceae)

Haruto Mano^a, Eugeny V. Boltenkov^b, Elena A. Marchuk^b, Koh Nakamura^c and Watanabe Yoichi^a

^aGraduate School of Horticulture, Chiba University, Matsudo, Chiba, Japan; ^bBotanical Garden-Institute, Far Eastern Branch, Russian Academy of Sciences, Vladivostok, Russia; ^cBotanic Garden, Field Science Center for Northern Biosphere, Hokkaido University, Sapporo, Hokkaido, Japan

ABSTRACT

Hypecoum erectum L., a widespread species in northern Eurasia, is a valuable medicinal plant, but its chloroplast genome has not previously been reported. We determined its complete chloroplast genome using a high-throughput sequencing technique. Its total length was 169,241 bp, consisting of a large single-copy region of 93,301 bp and a small single-copy region of 17,316 bp, separated by a pair of inverted repeat regions of 29,312 bp. A total of 140 genes were annotated, including 91 protein coding genes, 41 tRNA genes, and eight rRNA genes. The phylogenetic analysis shows that *H. erectum* and *H. zhukanum* of the subfamily Hypecoideae are monophyletic with the highest support.

ARTICLE HISTORY

Received 28 November 2023
Accepted 25 July 2024

KEYWORDS

Hypecoideae; medicinal plant; phylogenetic relationships; Russia

Introduction

Hypecoum erectum L. (Linnaeus 1753) is a winter annual plant found in Russia, Mongolia, and China (Zhang and Lidén 2008). Many species of the family Papaveraceae are used medicinally due to their high alkaloid content; indeed, *H. erectum* has been used extensively in folk medicine to treat inflammation, fever, and pain (Yuan et al. 2021).

The family Papaveraceae includes four subfamilies: Papaveroideae, Fumarioideae, Hypecoideae, and Pteridophylloideae. *Hypecoum* is the sole genus in Hypecoideae. Recently, the phylogenetic relationships between subfamilies and genera have started to be resolved based on chloroplast genome sequences (Peng et al. 2023), but few genome sequences in Papaveraceae are currently available, especially for Fumarioideae and Hypecoideae. Therefore, more data pertaining to genomic features and phylogenetic relationships between species are required in order to resolve the taxonomy and improve utilization of the taxa. In this study, we determined the complete chloroplast genome of *H. erectum* based on Illumina paired-end sequencing data. In addition, phylogenetic relationships for the genus *Hypecoum* and related genera were reconstructed using published sequences of related species.


Materials and methods

Sample leaves were collected on the way to Ulan-Ude in the vicinity of Kyakhta in the Republic of Buryatia, Russia

(50.674167 N, 106.492778 E) on 28 May 2021 (GenBank BioSample, SAMD00617559). The sample was identified and collected by Eugeny V. Boltenkov, and the voucher specimen was deposited in the Herbarium of the National Museum of Nature and Science, Japan (TNS; <https://tbg.kahaku.go.jp/english/index.php>, Atsushi Ebihara, ebihara@kahaku.go.jp) under accession number TNS1347885 (Figure 1).

Total genomic DNA was extracted from dried leaves using a DNeasy Plant Mini kit (Qiagen, Hilden, Germany). The sequencing library was prepared from extracted DNA using a TruSeq DNA PCR-Free Kit (Illumina, San Diego, CA, USA) and sequenced with 300 bp of pair-end reads on the MiSeq platform (Illumina); the library preparation and sequencing were conducted by Macrogen (Seoul, South Korea). The 3.2 Gbases of raw data were trimmed by clipping adaptor sequences and removing reads of low quality using fastp v. 0.20.0 (Chen et al. 2018). Trimmed reads were deposited in GenBank (accession number, DRR477443). Trimmed reads were assembled using GetOrganelle v.1.7.5 (Jin et al. 2020) with the following settings: -F embplant_pt, and default settings for the other options. The draft contig obtained was examined, and redundant sequences were edited using Bandage v.0.8.1 (Wick et al. 2015). The assembled genome was annotated on the basis of a blast search using GeSeq (Tillich et al. 2017) with default settings. Secondly, to check the entire coverage, the trimmed reads were mapped to the assembled circular genome using BWA (Li and Durbin 2009) with default settings. To facilitate this mapping, the sequence was modified by linearizing it, duplicating 700 bp from one end, and

CONTACT Watanabe Yoichi  ashitaka3776@gmail.com  Graduate School of Horticulture, Chiba University, Matsudo, Chiba 2718510, Japan

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/23802359.2024.2386410>.

© 2024 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.



Figure 1. The *Hypecoum erectum* plant examined (the Republic of Buryatia, Russia; photograph by Eugeny V. Boltenkov on 28 May 2021). (A) Habitat and (B) flowers. *Hypecoum erectum* is an annual plant 10–30 cm in height, pinnate leaves with deeply divided segments, and outer petals of 10–12 × 10–19 mm.

adding these to the opposite end. The maps of the genome and cis/trans-splicing genes were drawn using CPGView (Liu et al. 2023). The annotated chloroplast genome was deposited in GenBank (accession number, LC780190). To investigate the phylogenetic position of *H. erectum*, chloroplast genome sequences of related species were obtained from the GenBank database. A total of 14 species of Papaveraceae based on the study by Peng et al. (2023) and two species of Ranunculaceae and Eupteleaceae, as outgroups, were obtained. Seventy-one genes shared among species that were longer than 150 bp were extracted, and each ortholog gene was aligned using MAFFT v.7.505 (Katoh and Standley 2013) with default settings. These alignments were then combined. A phylogenetic tree was constructed by the maximum likelihood method using RAxML v.8.2.0 (Stamatakis 2014) under the GTR+G + I substitution model with 1000 bootstrap replicates. The substitution model was selected based on the Akaike information criterion, as implemented in jModeltest 2.1.6 (Darriba et al. 2012).

Results

The total length of the complete chloroplast genome of *Hypecoum erectum* was 169,241 bp, with an average coverage of 3126.4× and GC content of 38% (Figure 2, Supplementary Figure 1). A large single copy (LSC: 93,301 bp), a small single copy (SSC: 17,316 bp), and two inverted repeat (IR: 29,312 bp) regions made up the quadripartite structure of the chloroplast genome (Figure 2).

In total, 140 genes were annotated, including 91 protein-coding genes, 41 transfer RNA (tRNA) genes, and eight ribosomal RNA (rRNA) genes. Twenty-one genes and a partial *ycf1* were duplicated in the IRs; of these, *trnI-CAU*, *rpl23*, and *rpl2* were duplicated on both edges of an IR. In addition, two genes of *rpl2* in the IRs on the side of the SSC were identified as pseudogenes. Ten cis-splicing genes and one trans-splicing gene, *rps12*, were identified (Supplementary Figures 2 and 3).

The phylogenetic tree showed that Papaveraceae and each subfamily except Pteridophylloideae were monophyletic with the highest support (100%, Figure 3). The subfamilies Fumarioideae, Hypecoideae, and Pteridophylloideae formed a monophyletic group. *Hypecoum zhukanum* was the closest relative to *H. erectum*.

Discussion and conclusion

The structure of the chloroplast genome of *H. erectum* is similar to that of *H. zhukanum* (Peng et al. 2023). In the phylogenetic analysis, the relationships between the four subgenera of Papaveraceae examined in this study are in accord with those of the previous study of Peng et al. (2023). This phylogenetic analysis also supports the suggestion of monophyly of the two species of *Hypecoum* that were included in this study.

This study provides a reliable chloroplast genome of *H. erectum*, which will be valuable in future phylogenetic, taxonomic, pharmacological, and evolutionary studies of the family. The subfamily Hypecoideae consists of 18 *Hypecoum* species, and the subfamily Pteridophylloideae includes a

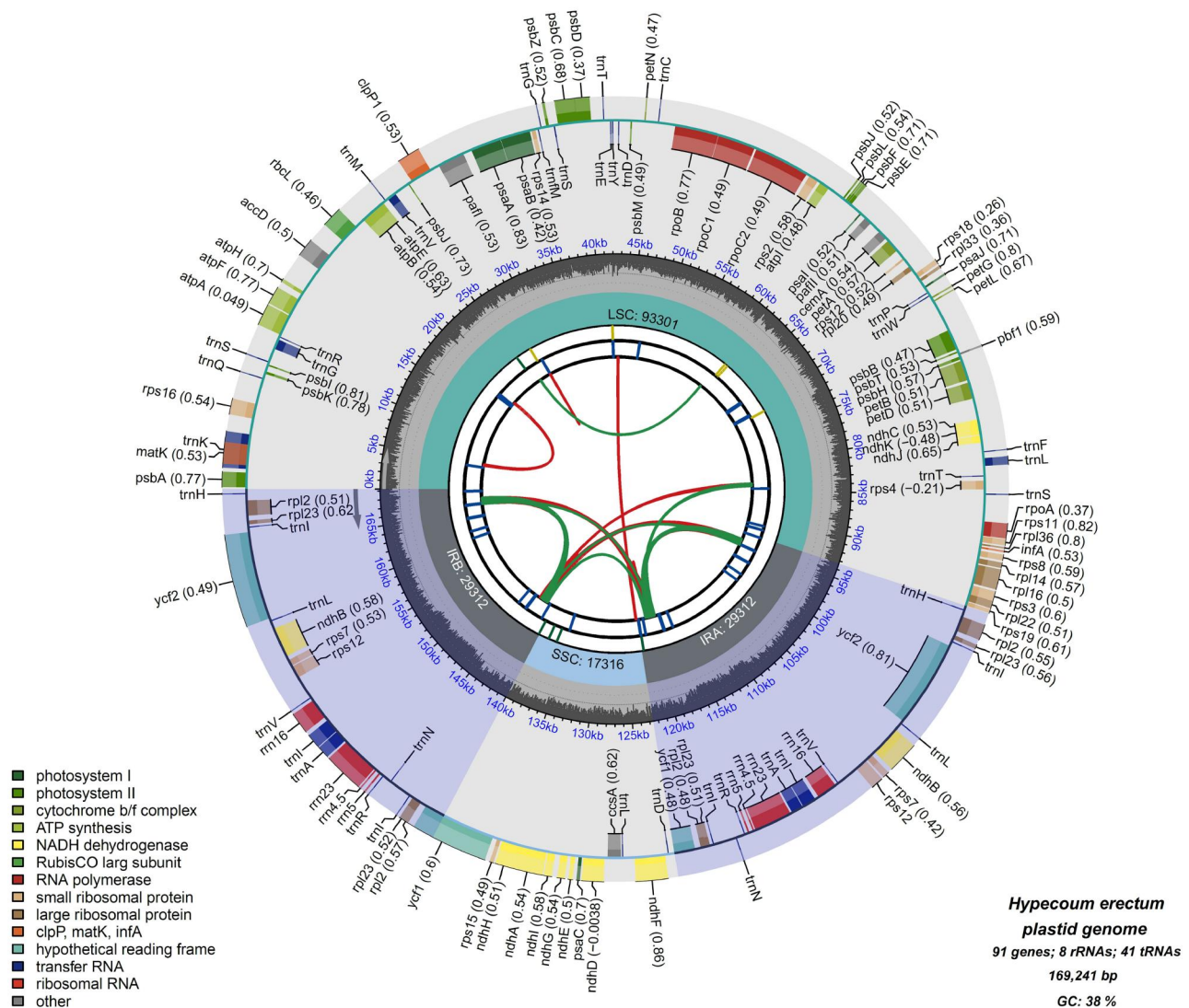


Figure 2. Complete chloroplast genome of *Hypecoum erectum*. The map contains six tracks. From the center outwards, the first track shows the dispersed repeats. The dispersed repeats consist of direct and palindromic repeats, connected with red and green arcs. The second track shows long tandem repeats as short blue bars. The third track shows short tandem repeats or microsatellite sequences as short bars of different colors. The small single copy (SSC), inverted repeat (IRA and IRB), and large single copy (LSC) regions are shown on the fourth track. The GC content along the genome is plotted on the fifth track. The genes are shown on the sixth track. The optional codon usage bias is displayed in parentheses after the gene name. Genes belonging to different functional groups are color-coded. Genes on the inside and outside the map are transcribed in clockwise and counterclockwise directions, respectively.

single species, *Pteridophyllum racemosum* (Zhang and Lidén 2008). The subfamily Fumarioideae contains 20 genera, but there are limited data on their chloroplast genomes, with the exception of the largest genus, *Corydalis* (Yu et al. 2021). Therefore, further studies of chloroplast genomes will help in understanding the evolution of the family (Peng et al. 2023).

Acknowledgments

We are grateful to Dr. Atsushi Ebihara in the Herbarium of the National Museum of Nature and Science, Japan for his help in deposition of the voucher specimen.

Author contributions

E.A.M., K.N., and W.Y. designed the study. E.V.B., E.A.M., and K.N. collected the sample. H.M. conducted molecular experiments. H.M. and W.Y.

generated, analyzed, and visualized the data. W. Y. wrote the manuscript. All the authors have discussed and agreed to the published version of the manuscript.

Ethical approval

There are no ethical issues associated with the material used in the study. This species is not endangered and not on the CITES and IUCN Red list, and the sample was not collected from a protected area. Therefore, no special permission was required in the Russian Federation. The sample was collected by a collaboration between Far East Branch of the Russian Academy of Science and Hokkaido University under Mutually Agreed Terms, MAT.

Disclosure statement

The authors declare no conflicts of interest.

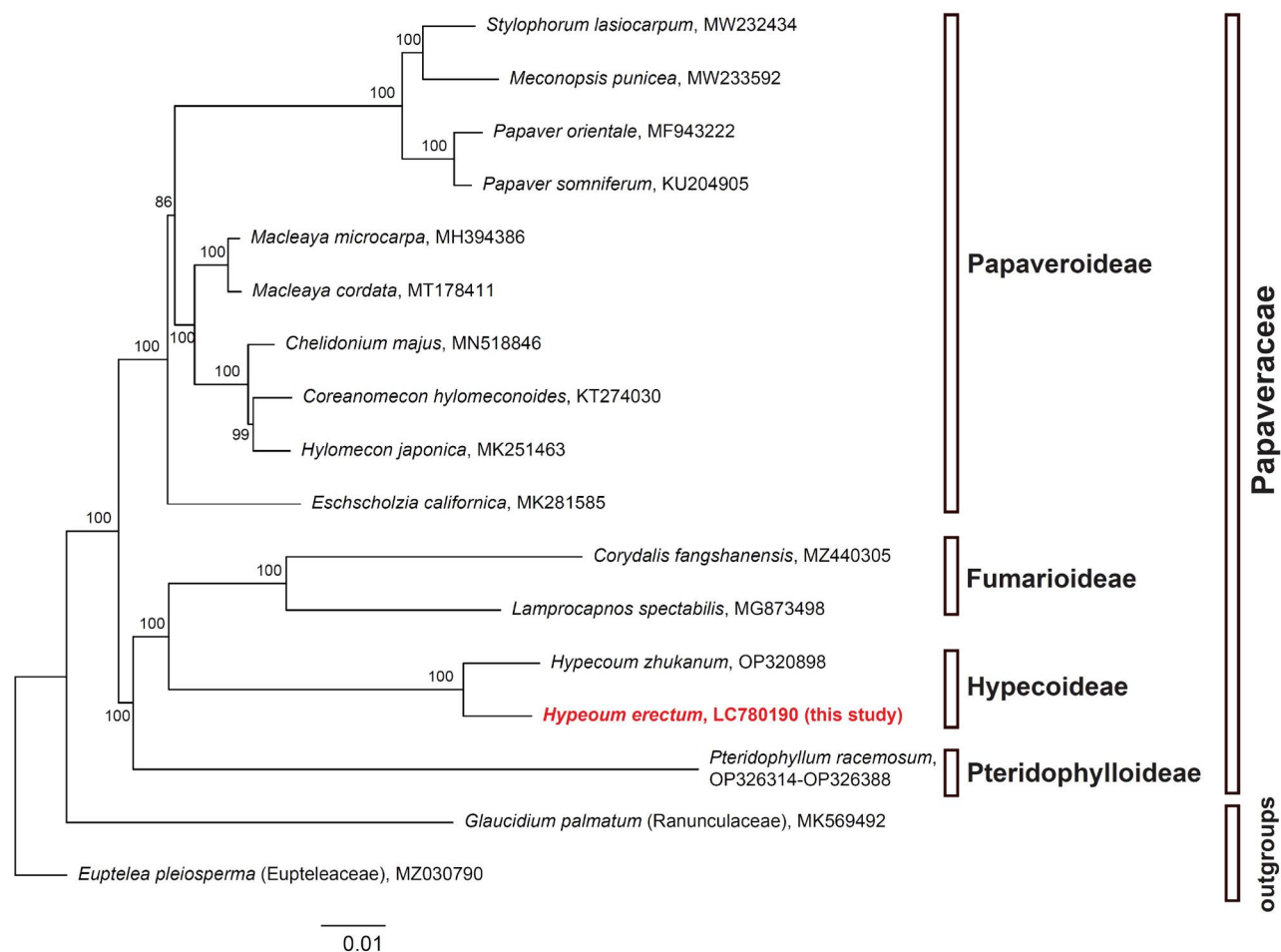


Figure 3. A maximum likelihood (ML) tree of 15 species of Papaveraceae, with two species of Ranunculaceae and Eupteleaceae, as outgroups. Bootstrap probabilities are shown above nodes. The genome of the species determined is highlighted in red-bold. The following sequences were used for phylogenetic analysis: *Stylophorum lasiocarpum* MW232434, *Meconopsis punicea* MW233592 (Liang et al. 2021), *Papaver orientale* MF943222 (Zhou et al. 2018), *Papaver somniferum* KU204905 (Sun et al. 2016), *Macleaya microcarpa* MH394386 (Zeng et al. 2018), *Macleaya cordata* MT178411 (Dong et al. 2020), *Chelidonium majus* MN518846, *Coreanomecon hylomeconoides* KT274030 (Kim and Kim 2016), *Hylomecon japonica* MK251463 (Zhang et al. 2019), *Eschscholzia californica* MK281585, *Corydalis fangshanensis* MZ440305 (Yu et al. 2021), *Lamprocapnos spectabilis* MG873498 (Park et al. 2018), *Hypecoum zhukanum* OP320898 (Peng et al. 2023), *Pteridophyllum racemosum* OP326314-OP326388 (Peng et al. 2023), *Glaucidium palmatum* (Ranunculaceae), MK569492 (Zhai et al. 2019), *Euptelea pleiosperma* MZ030790.

Funding

This research was supported by a Grant-in-Aid for Scientific Research for the Japan Society for the Promotion of Science, KAKENHI [Grant No. 20H03020].

Data availability statement

The data that support the findings of this study are openly available in GenBank (NCBI, <https://www.ncbi.nlm.nih.gov>, reference number LC780190. The associated BioProject, SRA, and BioSample numbers are PRJDB15948, DRR477443, and SAMD00617559, 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.
- Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods*. 9(8): 772–772. doi:10.1038/nmeth.2109.
- Dong S, Gao C, Wang Q, Ge Y, Cheng R. 2020. Characterization of the complete chloroplast genome of *Macleaya cordata* and its phylogenetic position within the subfamily Papaveroideae. *Mitochondrial DNA B Resour*. 5(2):1714–1715. doi:10.1080/23802359.2020.1749172.
- 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.
- Katoh K, Standley DM. 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 HW, Kim KJ. 2016. Complete plastid genome sequences of *Coreanomecon hylomeconoides* Nakai (Papaveraceae), a Korea endemic genus. *Mitochondrial DNA B Resour*. 1(1):601–602. doi:10.1080/23802359.2016.1209089.
- 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.
- Liang R, Caraballo-Ortiz MA, Liu Y, Su X. 2021. Characterization of the complete chloroplast genome of *Meconopsis punicea* (Papaveraceae), an endemic species from the Qinghai-Tibet plateau in China. *Cytol Genet*. 55(2):183–187. doi:10.3103/S0095452721020092.
- Linnaeus C. 1753. *Species Plantarum*: a facsimile of the first edition. London: Ray Society; p. 124.
- 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.
- Park S, An B, Park SJ. 2018. Reconfiguration of the plastid genome in *Lamprocapnos spectabilis*: IR boundary shifting, inversion, and intraspecific variation. *Sci Rep*. 8(1):13568. doi:10.1038/s41598-018-31938-w.

- Peng HW, Xiang KL, Erst AS, Lian L, Ortiz RDC, Jabbour F, Chen ZD, Wang W. 2023. A complete genus-level phylogeny reveals the Cretaceous biogeographic diversification of the poppy family. *Mol Phylogenet Evol.* 181:107712. doi:10.1016/j.ympev.2023.107712.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 30(9):1312–1313. doi:10.1093/bioinformatics/btu033.
- Sun Y, Moore MJ, Zhang S, Soltis PS, Soltis DE, Zhao T, Meng A, Li X, Li J, Wang H. 2016. Phylogenomic and structural analyses of 18 complete plastomes across nearly all families of early-diverging eudicots, including an angiosperm-wide analysis of IR gene content evolution. *Mol Phylogenet Evol.* 96:93–101. doi:10.1016/j.ympev.2015.12.006.
- 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.
- Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage – interactive visualisation of de novo genome assemblies. *Bioinformatics.* 31(20):3350–3352. doi:10.1093/bioinformatics/btv383.
- Yu Z, Zhou T, Li N, Wang D. 2021. The complete chloroplast genome and phylogenetic analysis of *Corydalis fangshanensis* W.T. Wang ex S.Y. He (Papaveraceae). *Mitochondrial DNA B Resour.* 6(11):3171–3173. doi:10.1080/23802359.2021.1987172.
- Yuan HL, Zhao YL, Qin XJ, Liu YP, Yang XW, Luo XD. 2021. Diverse isoquinolines with anti-inflammatory and analgesic bioactivities from *Hypecoum erectum*. *J Ethnopharmacol.* 270:113811. doi:10.1016/j.jep.2021.113811.
- Zeng CX, Hollingsworth PM, Yang J, He ZS, Zhang ZR, Li DZ, Yang JB. 2018. Genome skimming herbarium specimens for DNA barcoding and phylogenomics. *Plant Methods.* 14(1):43. doi:10.1186/s13007-018-0300-0.
- Zhai W, Duan X, Zhang R, Guo C, Li L, Xu G, Shan H, Kong H, Ren Y. 2019. Chloroplast genomic data provide new and robust insights into the phylogeny and evolution of the Ranunculaceae. *Mol Phylogenet Evol.* 135:12–21. doi:10.1016/j.ympev.2019.02.024.
- Zhang Y, Lee J, Liu X, Sun Z. 2019. The first complete chloroplast genome of *Hylomecon japonica* and its phylogenetic position within Papaveraceae. *Mitochondrial DNA B Resour.* 4(2):2349–2350. doi:10.1080/23802359.2019.1573125.
- Zhang M, Lidén M. 2008. *Hypecoum*. Flora of China. In: Wu Z, Raven PH, Hong D, editors. Menispermaceae through Capparaceae. Vol. 7. St Louis (MO): Missouri Botanical Garden Press; p. 288–290.
- Zhou J, Cui Y, Chen X, Li Y, Xu Z, Duan B, Li Y, Song J, Yao H. 2018. Complete chloroplast genomes of *Papaver rhoeas* and *Papaver orientale*: molecular structures, comparative analysis, and phylogenetic analysis. *Molecules.* 23(2):437. doi:10.3390/molecules23020437.