MITOCHONDRIAL DNA PART B: RESOURCES
2024, VOL. 9, NO. 8, 1048–1052
https://doi.org/10.1080/23802359.2024.2391087

Taylor & Francis
Taylor & Francis Group

PLASTOME REPORT

OPEN ACCESS Check for updates

The complete chloroplast genome of *Lindernia crustacea* (L.) F. Muell 1882 (Linderniaceae) and its phylogenetic analysis

Xinhua Wang[a], Liqiang Wang[a], Jiaojiao Kong[b] and Na Kong[a]

[a]College of Pharmacy, Heze University, Heze, Shandong Province, P. R. China; [b]Personnel Office, Heze Medical College, Heze, Shandong Province, P. R. China

ABSTRACT

Lindernia crustacea (L.) F. Muell 1882, a species in the Linderniaceae family, holds traditional medicinal value in China. To investigate its genetic diversity, we assembled, annotated, and characterized the first complete chloroplast genome of *L. crustacea* using Illumina sequencing data and various bioinformatics tools. The genome is 153,647 bp in length, with a GC content of 37.6%. It exhibits a typical quadripartite structure, consisting of a large single-copy region (LSC) of 85,411 bp, a small single-copy region (SSC) of 18,724 bp, and two inverted repeat sequences (IRa and IRb) of 25,816 bp each. The genome was predicted to contain 131 genes, including 87 protein-coding genes, 36 tRNA genes, and eight rRNA genes. Phylogenomic analysis indicated that *L. crustacea* is closely related to *L. stricta*. These findings provide a foundation for further research on the evolution and potential medicinal applications of the Linderniaceae family.

ARTICLE HISTORY
Received 11 January 2024
Accepted 6 August 2024

KEYWORDS
Chloroplast genome; *Lindernia crustacea*; Linderniaceae; phylogenetic analysis

Introduction

The *Lindernia* genus, a key member of the newly established Linderniaceae family, is distributed across warm regions in the Eastern and Western Hemispheres. The genus includes 30 species in its most recent classification, with China hosting up to 26 species. In China, *Lindernia crustacea* (L.) F. Muell 1882 is an annual plant characterized by diffuse branching and heights ranging from 3.5 to 30.0 cm. It commonly inhabits moist environments such as riverbeds, rice fields, and open grassy areas. This plant is extensively utilized in ethnomedicine worldwide and is known for treating diverse conditions such as earaches, injuries, fever, thrush, skin disorders, dysentery, ringworm, and postpartum complications (Smriti et al. 2019).

With its smaller size and slower mutation rate compared to nuclear and mitochondrial genomes, the chloroplast genome has become a crucial tool for resolving phylogenetic relationships across different taxonomic levels (Yan et al. 2023). Morphologically, *Lindernia* was originally classified within the Scrophulariaceae (Lewis 2000). However, with the aid of molecular data, particularly chloroplast genomic data, APG III reclassified it into the new family Linderniaceae (Angiosperm Phylogeny Group 2009). In this study, we aimed to assemble and characterize the complete chloroplast genome of *L. crustacea*, providing a foundation for future investigations into the evolution, conservation, and potential medicinal applications of this species within the Linderniaceae family.

Materials

Fresh leaves were collected from natural *L. crustacea* plants (Figure 1) for sequencing. The collection site of the species was located in Peony District (35°15′9.27"N, 115°29′44.44"E), Heze City, Shandong Province, China. A specimen is stored at Heze University Herbarium under the sample number HZ220832 (contact: Hongqin Li, 463056627@qq.com). Its whole genomic DNA was extracted using a plant genomic DNA kit from Tiangen Biotech, Beijing.

Methods

The extracted whole genomic DNA was fragmented into approximately 300 bp segments to construct a 150 bp paired-end sequencing library. The library was sequenced using the Illumina NovaSeq 6000 platform at Wuhan Benagen Technology Company Limited (Wuhan, China). Raw sequencing reads were trimmed using Trimmomatic software (v0.39) (Bolger et al. 2014) with parameters: LEADING:10, TRAILING:10, SLIDINGWINDOWS:4:20, and MINLEN:50. The cleaned dataset, containing 44,877,348 reads, was deposited in the SRR database (SRR27354904). The cleaned reads were used to assemble the complete chloroplast genome using

CONTACT Na Kong nakong1982@163.com College of Pharmacy, Heze University, Heze 274015, Shandong Province, P. R. China.
Supplemental data for this article can be accessed online at https://doi.org/10.1080/23802359.2024.2391087.

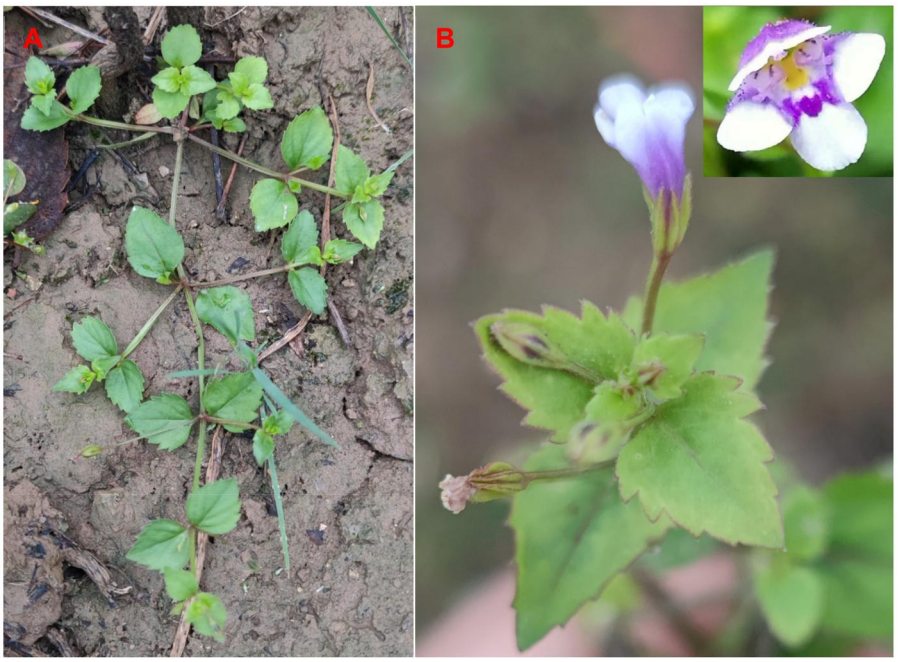

Figure 1. Panorama (A) and detail (B) habitat photos of *Lindernia crustacea*. The author Liqiang Wang took the photo at the position of 35°15′9.27″ N, 115°29′44.44″E. Main identifying traits: Annuals, 10–20 cm tall, much branched. Petiole 1–8 mm; leaf blade triangular-ovate to broadly ovate. Flowers axillary and solitary or in short apical racemes. Pedicel slender, 0.5–2.2 cm, subglabrous. Calyx urnlike, 3–5 mm, shallowly lobed; lobes triangular-ovate, outside sparsely pubescent. Corolla purple, 5–8 mm; tube slightly longer than calyx; lower lip 3-lobed, Middle lobe larger and slightly longer than upper lip; upper lip ovate, sometimes shallowly 2-lobed. Stamens didynamous.

GetOrganelle software (v1.7.1) (Jin et al. 2020). Annotation of the assembled genome was performed with CPGAVAS2 (Shi et al. 2019) and manually calibrated using Apollo software (Pontius 2018). The final genome annotation information was submitted to the GenBank (OQ411035). Finally, the comprehensive circular genome map of the assembled genome was generated using CPGview (Liu et al. 2023).

For phylogenetic analysis of *L. crustacea*, we downloaded whole chloroplast genome sequences of another seven *Lindernia* species from GenBank, with *Adenosma glutinosum* and *Scoparia dulcis* as outgroups. Ten whole chloroplast genomes were aligned using MAFFT software (https://mafft.cbrc.jp/alignment/software/) with the parameter '–auto' (Katoh and Standley 2016). A maximum-likelihood (ML) phylogenetic tree was constructed using IQ-TREE software (v2.2.2.7) (Minh et al. 2020) with the script 'iqtree -s aligned.genomes.fasta -m MFP -bb 1000 -bnni'. The ModelFinder algorithm embedded in IQ-TREE was used to identify the most suitable nucleotide substitution model. Then, an ML phylogenetic tree was constructed based on the selected model TVM + F + I + I+R3. The tree's confidence was assessed using an ultra-fast bootstrap method with 1000 replicates. Further confidence assessments of the tree were performed using the Shimodaira-Hasegawa (SH) test (Shimodaira and Hasegawa 1999) and the approximately unbiased (AU) test (Shimodaira 2002) within the IQ-TREE, using the script: 'iqtree -s input.aligned.phy -m TVM + F + I + I+R3 -z genome.unconstrain.constrain.treels -zb 10000 -zw -au'. Additionally, polymorphisms in the eight *Lindernia* chloroplast genomes were analyzed based on this alignment, and sequence variations were calculated using DnaSP (v6) (Rozas et al. 2017) with a 600 bp sliding window and 200 bp step (Cui et al. 2020).

Results

The chloroplast genome of *L. crustacea* is 153,647 bp in length and exhibits a classic quadripartite structure. It comprises a pair of inverted repeat regions (IRs) of 24,756 bp each, separated by a large single-copy region (LSC) of 85,411 bp and a small single-copy region (SSC) of 18,724 bp (Figure 2). Mapping experiments demonstrated the high reliability of the genome assembly, with an average and minimum depth of 4083.20× and 504×, respectively (Figure S1). The overall GC content of the chloroplast genome is 37.6%, with the IR regions exhibiting the highest GC content of 43.5% and the LSC and SSC regions showing GC contents of 35.4% and 32.0%, respectively (Table S1).

The chloroplast genome of *L. crustacea* was predicted to contain 131 genes, including 87 protein-coding genes (PCGs), eight rRNA genes, and 36 tRNA genes. Seven unique PCGs (*rps*12, *rps*7, *rpl*2, *rpl*23, *ndh*B, *ycf*2, and *ycf*15), seven unique tRNA genes (*trn*A, *trn*E, *trn*L, *trn*M, *trn*N, *trn*R, and *trn*V), and four unique rRNA genes (*rrn*16S, *rrn*23S, *rrn*4.5S, *rrn*5S) are located in the IR regions. Nine PCGs (*rps*16, *atp*F, *rpo*C1, *pet*B, *pet*D, *rpl*16, *rpl*2, *ndh*B, and *ndh*A) each contain one intron, while three PCGs (*rps*12, *ycf*3, and *clp*P) contain two introns. Additionally, five tRNA genes (*trn*K-UUU, *trn*S-CGA, *trn*L-UAA, *trn*E-UUC, and *trn*A-UGC) contain one intron. In total, 17 genes with introns were identified. Among them, *ycf*3, *rps*12, and *clp*P each contained two introns, while the remaining 14 genes

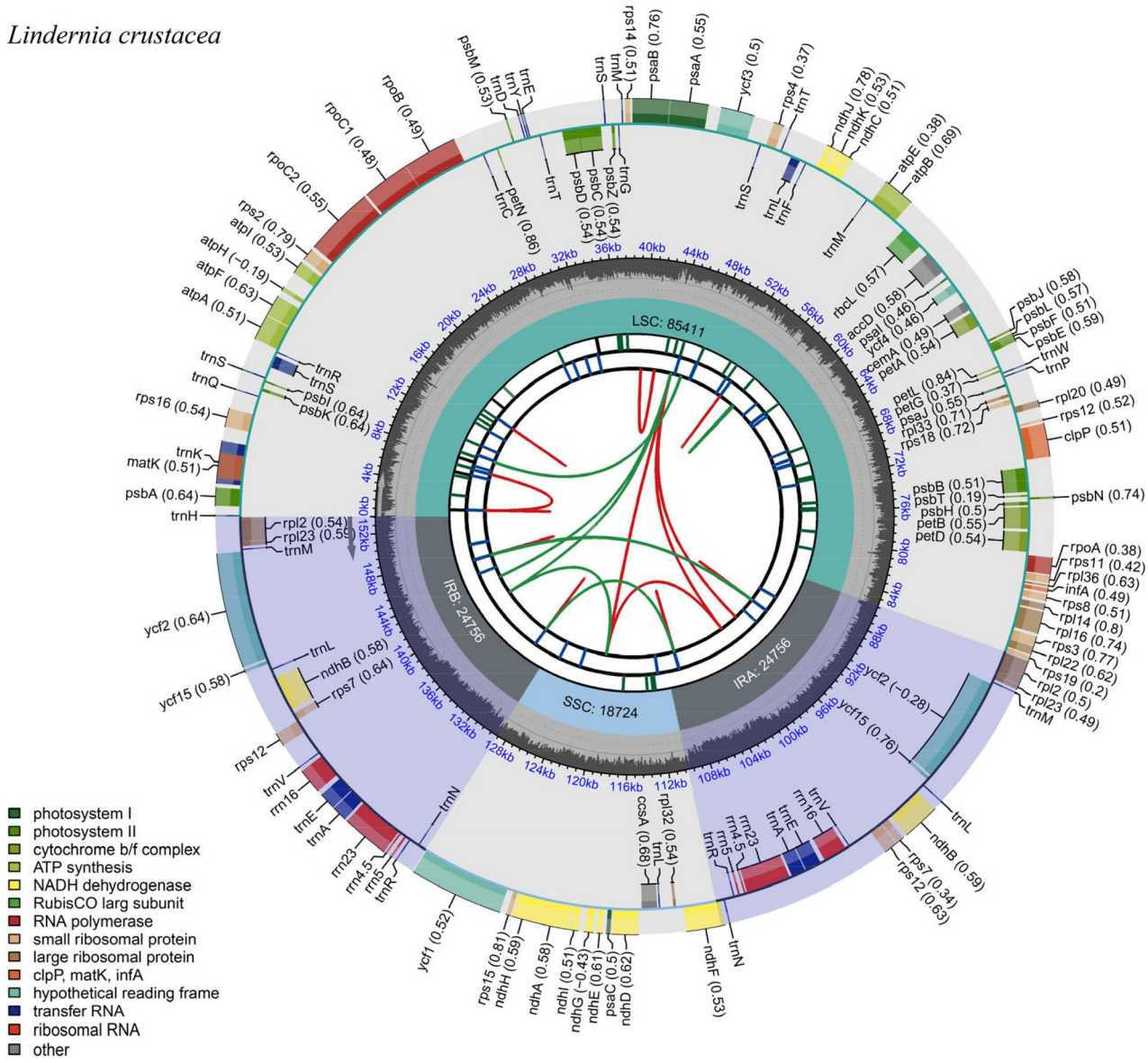

Figure 2. Gene map of the complete chloroplast genome of *Lindernia crustacea*. The species name is shown in the top left corner. The map contains six tracks by default. From the center outward, the first track shows dispersed repeats, including direct and palindromic repeats connected by red and green arcs. The second track displays long tandem repeats as blue bars, while the third displays short tandem repeats or microsatellite sequences as differently colored short bars. These colors correspond to the type and description of each repeat, with black representing complex repeats, green for repeat unit size 1, yellow for size 2, purple for size 3, blue for size 4, orange for size 5, and red for size 6. The fourth track displays the SSC, IRa, IRb, and LSC regions. The fifth track shows the GC content along the genome, while the sixth track sounds the genes. The gene names are followed by optional information about codon usage bias and color-coded based on their functional classification. The inner genes are transcribed clockwise, and the outer genes are transcribed anticlockwise. The available type of the genes is shown in the bottom left corner.

contained a single intron. The structures of the cis-splicing genes and trans-splicing gene (*rps*12) are shown in Figure S2.

Phylogenetic analysis revealed that *L. crustacea* is closely related to *L. stricta* within the *Lindernia* genus. *Lindernia crustacea* and *L. stricta* formed a monophyletic clade with 100% bootstrap support (Figure 3). The topology of the ML tree was also strongly supported by both the Shimodaira-Hasegawa (SH) test and the approximately unbiased (AU) test, with *p*-values of 1.00 and 0.485, respectively. Thus, this phylogenetic analysis provides a reliable account of the evolutionary relationship of *L. crustacea*. Based on the multiple sequence alignments of the eight *Lindernia* species chloroplast genomes, we further identified four regions with the highest variability within the eight *Lindernia* species chloroplast genomes, with pi values ranging from 0.07006 to 0.08024 (Figure S3, Table S2).

Discussion and conclusions

In this study, we reported the first *L. crustacea* chloroplast genome. The genome exhibits a typical quadripartite structure with a size of 153,647 bp and a GC content of 37.6%. The genome contains 131 genes. Phylogenetic analysis indicated that *L. crustacea* is closely related to *L. stricta*. Additionally, four regions within the *Lindernia* species chloroplast genomes were identified as potential targets for designing DNA barcodes.

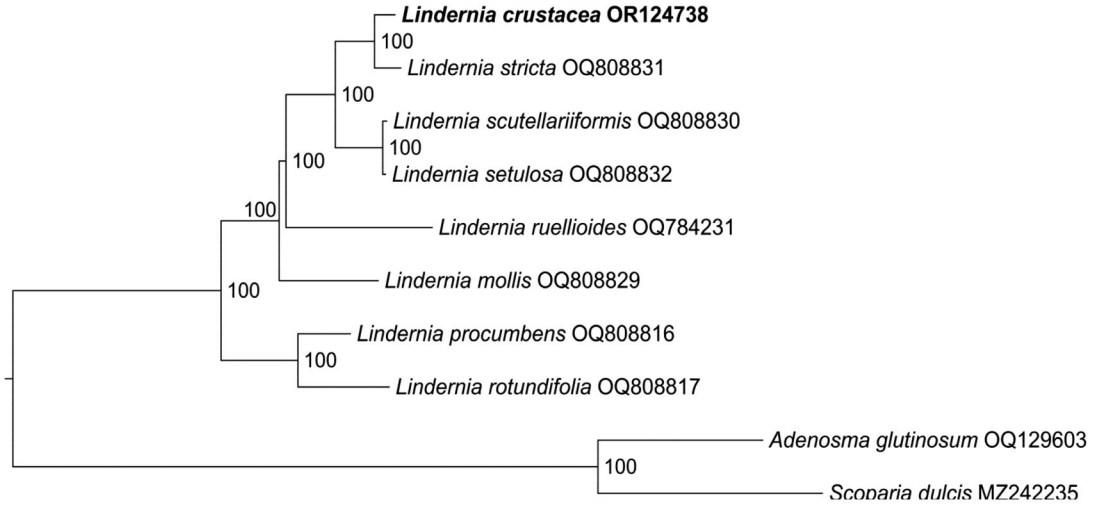

Figure 3. The maximum likelihood phylogeny of *Lindernia crustacea* and its close relatives using whole chloroplast genome sequences. Bootstrap values based on 1000 replicates are shown at each node in the cladogram. We downloaded chloroplast genome sequences of another seven *Lindernia* species from GenBank: *L. stricta* (OQ808831) (Yan et al. 2023), *L. scutellariiformis* (OQ808830) (Yan et al. 2023), *L. setulosa* (OQ808832) (Yan et al. 2023), *L. ruellioides* (OQ784231), *L. mollis* (OQ808829) (Yan et al. 2023), *L. procumbens* (OQ808816) (Yan et al. 2023), *L. rotundifolia* (OQ808817) (Yan et al. 2023). *Adenosma glutinosum* (OQ129603, outgroup) (Xie et al. 2023) and *Scoparia dulcis* (MZ242235, outgroup) (Lin et al. 2021) were used as outgroups. The new *L. crustacea* (OR124738) chloroplast genome sequenced in this study is highlighted in bold.

The *L. crustacea* chloroplast genome is only 238 bp shorter than that of *L. stricta*, with one fewer predicted gene compared to that of *L. stricta* (OQ808831) (Yan et al. 2023). Comparative analyses indicate that Linderniaceae chloroplast genomes are highly conserved in terms of genome structure, GC content, gene order, and gene content (Yan et al. 2023).

Complete chloroplast genomes have shown a powerful ability to resolve backbone phylogeny compared to universal DNA markers (Chen et al. 2022; Peng et al. 2023). In this study, we identified four regions with notable variability, mainly concentrated in the LSC and SSC regions, consistent with findings by Cui et al. (2020). However, *rrn*23S showed high sequence polymorphism in the relatively conserved IR region, differing from that of Cui et al. (2020). The results of this study will facilitate future research into the pharmaceutical applications and evolution of *L. crustacea* and the *Lindernia* genus.

Acknowledgments

Thanking for identifiing the species and collecting samples of the species by Hongqin Li.

Author contributions

The manuscript includes contributions from all authors. Liqiang Wang was primarily responsible for the study design. Jiaojiao Kong extracted DNA for sequencing. Xinhua Wang assembled and annotated the chloroplast genome. Jiaojiao Kong performed the phylogenetic analysis of the species. Na Kong analyzed the structure of the chloroplast genome and drafted the manuscript. All authors approved the publication of the version and agreed to be accountable for all aspects of the work.

Disclosure statement

The authors reported no potential conflict of interest.

Funding

This work was supported by the Doctoral Fund Project of Heze University [XY12BS04] and the Shandong Provincial Natural Science Foundation [ZR2021MC136].

ORCID

Liqiang Wang http://orcid.org/0000-0002-8366-7016

Data availability statement

In this study, the complete chloroplast genome sequence of *L. crustacea* has been submitted to the NCBI database under the accession number OR124738. https://www.ncbi.nlm.nih.gov. The associated BioProject, BioSample, and SRA numbers are PRJNA928567, SAMN39131260, SRR27354904.

Ethical approval

The *L. crustacea* specimen is not an endangered species and does not require any specific permissions or licenses. The collection of plant material and research on this species was conducted in accordance with the guidelines provided by Heze University in this study.

References

Angiosperm Phylogeny Group. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Bot J Linn Soc. 161(2):105–121. doi:10.1111/j.1095-8339.2009.00996.x.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30(15):2114–2120. doi:10.1093/bioinformatics/btu170.

Chen Q, Hu H, Zhang D. 2022. DNA barcoding and phylogenomic analysis of the genus *Fritillaria* in China based on complete chloroplast genomes. Front Plant Sci. 13:764255. doi:10.3389/fpls.2022.764255.

Cui N, Liao BS, Liang CL, Li SF, Zhang H, Xu J, Li XW, Chen SL. 2020. Complete chloroplast genome of *Salvia plebeia* L. (Lamiaceae): organization, specific barcode and phylogenetic analysis. Chin J Nat Med. 18(8):563–572. doi:10.1016/S1875-5364(20)30068-6.

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. 2016. A simple method to control over-alignment in the MAFFT multiple sequence alignment program. Bioinformatics. 32(13):1933–1942. doi:10.1093/bioinformatics/btw108.

Lewis DQ. 2000. A revision of the new world species of *Lindernia* (Scrophulariaceae). Castanea. 65(2):93–122.

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.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 37(5):1530–1534. doi:10.1093/molbev/msaa015.

Peng C, Guo XL, Zhou SD, He XJ. 2023. Backbone phylogeny and adaptive evolution of *Pleurospermum* s.l.: new insights from phylogenomic analyses of complete plastome data. Front Plant Sci. 14:1148303. doi:10.3389/fpls.2023.1148303.

Pontius E. 2018. Apollo. Ann Emerg Med. 72(5):616. doi:10.1016/j.annemergmed.2018.06.016.

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A. 2017. DnaSP 6: DNA sequence polymorphism analysis of large datasets. Mol Biol Evol. 34(12):3299–3302. doi:10.1093/molbev/msx248.

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.

Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. Syst Biol. 51(3):492–508. doi:10.1080/10635150290069913.

Shimodaira H, Hasegawa M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol Biol Evol. 16(8):1114–1116. doi:10.1093/oxfordjournals.molbev.a026201.

Smriti RCD, Abdul BA, Dipankar S. 2019. Scientific evidence of *Lindernia crustacea* (l) f. *muell*, an indigenous plant: a folklore medicine used traditionally. Int Res J Pharm. 10(1):176–183. doi:10.7897/2230-8407.100130.

Yan RR, Geng YF, Jia YH, Xiang CL, Zhou XX, Hu GX. 2023. Comparative analyses of Linderniaceae plastomes, with implications for its phylogeny and evolution. Front Plant Sci. 14:1265641. doi:10.3389/fpls.2023.1265641.