

## The complete chloroplast genome of *Polyspora axillaris* (Theaceae)

Hui-Ting Tang<sup>a</sup>, Chun-Fang Li<sup>b</sup>, Shi-You Zuo<sup>b</sup> and Pei-Feng Liu<sup>b</sup>

<sup>a</sup>School of Tropical Agriculture and Forestry of Hainan University, Haikou, China; <sup>b</sup>Sanya Nanfan Research Institute of Hainan University, Sanya, China

### ABSTRACT

*Polyspora axillaris* (Roxb. ex Ker Gawl.) Sweet 1825, is a shrub or tree that is about 9 meters tall in the Theaceae family, mainly distributed in China and Vietnam, and it is widely used as a green tree species in many regions owing to its rapid growth and good adaptability. It is rich in various beneficial extracts for humans, but there are limited studies on it. In this study, we sequenced and annotated the complete plastome of *P. axillaris*. The chloroplast genome length of *P. axillaris* is 156,770 bp, with a total of 132 genes, including 37 tRNA genes, 8 rRNA genes and 87 protein-coding genes. The complete chloroplast genome of *P. axillaris* contains two Inverted Repeats (IRs) of 26,077 bp, a Large Single-Copy (LSC) region of 86,286 bp and a Small Single-Copy (SSC) region of 18,330 bp. The overall G/C content in the chloroplast is 37.3%. Phylogenetic inference shows that *P. axillaris* formed a sister relationship with *P. hainanensis*, along with 10 Theaceae species. The research result of *P. axillaris* will contribute to the genetic preservation of the species and the phylogenetic study of *Polyspora*.

### ARTICLE HISTORY

Received 14 February 2024  
Accepted 18 July 2024

### KEYWORDS

Genomic structure;  
*Polyspora*; Theaceae;  
plastome

### Introduction

*Polyspora axillaris* (Roxb. ex Ker Gawl.) Sweet 1825, is a shrub or tree that is about 9 meters tall in the Theaceae family, mainly distributed in China and Vietnam (Ming and Bartholomew 2007; Nguyet et al. 2020). It grows rapidly and is adaptable, so it is widely used as a greening tree (Qin 1990; Fan et al. 2021; Fan et al. 2023). *P. axillaris* not only has ecological value, but also has health value. The extracts from its roots, stems, and fruits have antiviral and preventive effects on cardiovascular and cerebrovascular diseases (Li et al. 2017, 2019; Fan et al. 2021). However, although *P. axillaris* has many valuable values, the research on it is still insufficient. Based on this, we reports the complete plastome of *P. axillaris*, with the aim of improving the quality of related collection and phylogenetic research on *Polyspora* Sweet plants.

### Materials


The fresh leaves of *Polyspora axillaris* was sampled from the Jianfengling in Hainan, China (108.79°E, 18.69°N) (Figure 1). The voucher specimen (voucher code: D. J. Chen, X. R. Ke, SK19, HUTB) was deposited at Hainan University, Hainan province, China (contact person Hua-Feng Wang, [hfwang@hainanu.edu.cn](mailto:hfwang@hainanu.edu.cn)).

### Methods

Total genomic DNA extraction was performed using the cetyltrimethylammonium bromide (CTAB) protocol (Doyle and Doyle 1987). Subsequently, we constructed a paired-end libraries with insertion fragment sizes of 200–400 bp using genomic DNA of the selected samples. The sequencing work of these libraries was completed on the BGISEQ-500 platform of BGI in Shenzhen, China. The original read data was trimmed using SOAPfilter\_v2. 2. Approximately 5 Gbp paired-end sequencing data were used for chloroplast genome assembly through GetOrganelle v1. 7. 5. 0 (Jin et al. 2020). To verify the accuracy of assembly, we compared the sequencing data with the assembled chloroplast genome using HISAT 2. 1. 0 (Kim et al. 2019), and statistically analyzed the sequencing depth of each site using SAMtools 1. 19. 2 (Danecek et al. 2021). In terms of gene annotation, we used Geneious Prime v2022. 1. 1 (Kearse et al. 2012) with reference to the chloroplast of *Polyspora axillaris* (NC\_035645) and manually validated the results to detect potential errors. For gene structures that are difficult to annotate, such as cis and trans splicing genes, identification was carried out using CPGView online website (<http://www.1kmpg.cn/cpgview>, Liu et al. 2023). The complete chloroplast genome loop diagram was drawn by OGDRAW version 1. 3. 1 (Greiner et al. 2019).

In order to understand the phylogenetic relationship between *P. axillaris* and related species, we reconstructed a phylogenetic tree using the maximum likelihood (ML)

**CONTACT** Hui-Ting Tang  [20213006570@hainanu.edu.cn](mailto:20213006570@hainanu.edu.cn)  School of Tropical Agriculture and Forestry of Hainan University, Haikou 570288, China

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/23802359.2024.2383672>.

© 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.** Photograph of *Polyspora axillaris* (the photograph was taken by the author in Jianfengling of Hainan, China). The specimen was stored at herbarium of Hainan University (<http://en.hainanu.edu.cn/>, contact person Hua-Feng Wang, hfwang@hainanu.edu.cn) under the voucher number D. J. Chen, X. R. Ke, SK19, without any copyright issues. The leaves in the axils exhibit an inverted oval to oblanceolate shape, characterized by a hairless appearance at 6–14 cm × 2.5–4 cm, with a dark green dorsal surface. Flowers are individual or paired.

method based on the complete chloroplast genome of *P. axillaris*, 5 other *Polyspora* species and 6 *Camellia* species, with *Pyrenaria khasiana* (NC\_035644) as the outgroup (Table S1). All sequences were aligned by MAFFT (Kato and Standley 2013) and trimmed by Gblocks (Talavera and Castresana 2007). Maximum likelihood phylogenies were inferred using IQ-TREE (Guindon et al. 2010; Nguyen et al. 2015) in PhyloSuite (Zhang et al. 2020) for 5000 ultrafast (Minh et al. 2013) bootstraps. In ML analysis, IQ-TREE was used with 1,000 bootstrap replicates automatically without setting specific models. The results were visualized by iTOL (<https://itol.embl.de/>) (Letunic and Bork 2016).

## Results

The minimum and average read mapping depths for assembled genomes were ×41, and ×476, respectively (Figure S1). The chloroplast genome length of *Polyspora axillaris* is 156,770 bp, with a total of 132 genes, including 37 tRNA genes, 8 rRNA genes (4.5S rRNA, 5S rRNA, 16S rRNA and 23S rRNA) and 87 protein-coding genes. And the complete chloroplast of *P. axillaris* contains two Inverted Repeats (IRs) of 26,077 bp, a Large Single-Copy (LSC) region of 86,286 bp and a Small Single-Copy (SSC) region of 18,330 bp (Figure 2). Eight protein-coding genes are duplicated in the IR. Seven tRNA genes are duplicated in the IR. Four rRNA genes are duplicated in the IR. Among them, 11 genes (*rsp16*, *atpF*, *rpoC1*, *ycf3*, *clpP*, *petB*, *petD*, *rpl16*, *rp12*, *ndhB*, and *ndhA*) are cis-splicing genes, and *rps12* is also a trans-splicing gene (Figures S2 and S3). The overall G/C content in the chloroplast is 37.3%, of which the corresponding value for the LSC, SSC, and IR regions are 35.3%, 30.5%, and 43%, respectively. The annotated complete chloroplast genome of *P. axillaris* was submitted to GenBank (accession number OR863388).

The raw reads were deposited in the NCBI Sequence Read Archive (accession number SRR27011443).

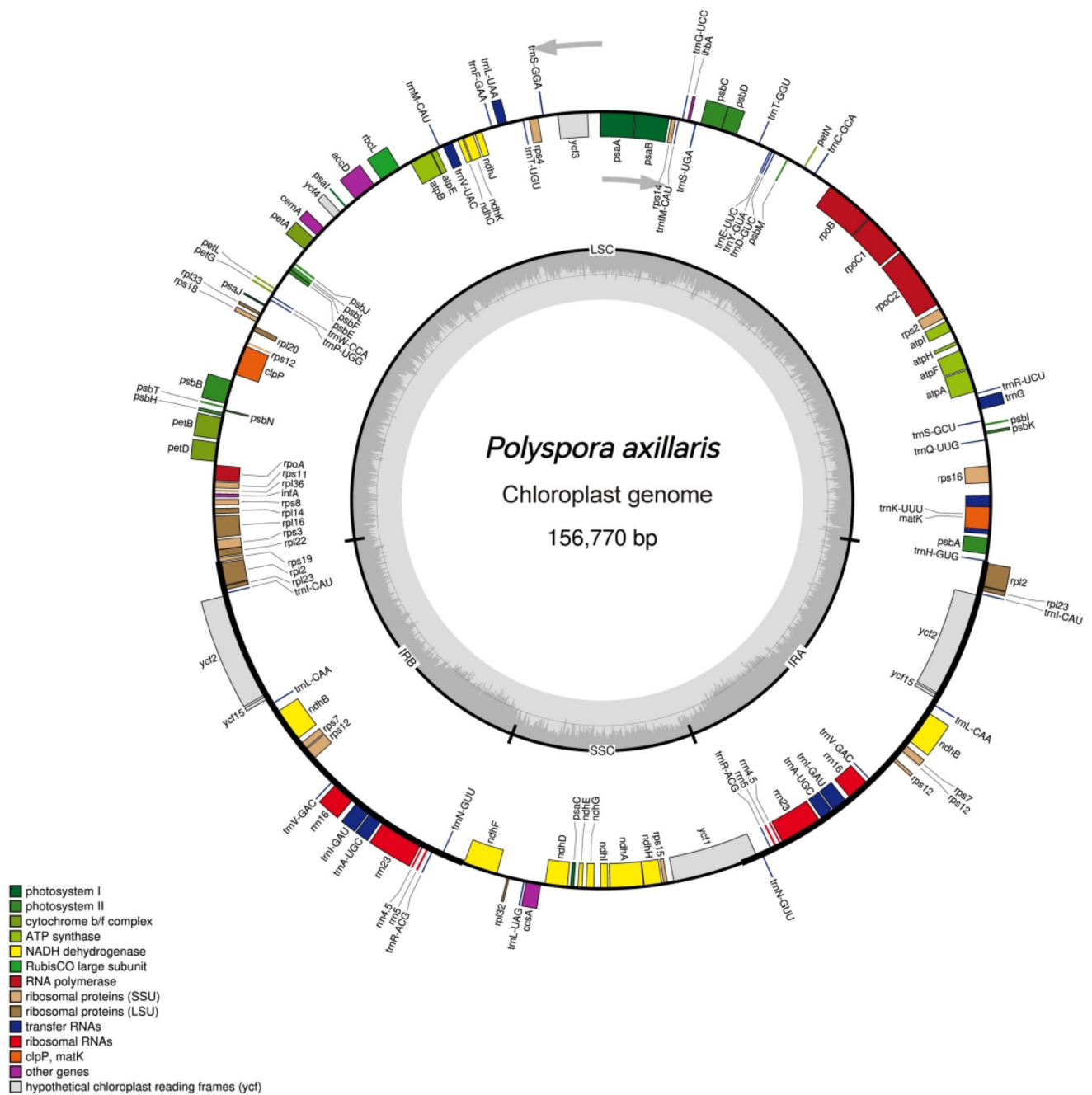
The alignment results of these 13 sequences are displayed in MEGA11 (Tamura et al. 2021). Figure S4 is a partial screenshot of multiple sequence alignment. And the results of phylogenetic research based on ML indicate that *P. axillaris* belongs to the genus *Polyspora* in the Theaceae family. And compared with other species, *P. axillaris* has a closer genetic relationship with *P. hainanensis* (Figure 3). The phylogenetic tree in a Newick text format is as supplementary file. Most nodes in the maximum likelihood (ML) tree of the Theaceae family plastomes are highly supported. In addition, the plastome sequence of *P. axillaris* will be beneficial for the phylogenetic research related to the Theaceae family, deepening the understanding of the complex relationships among Theaceae plants.

## Discussion and conclusions

CP genome analysis has been used to study the genetic diversity of plant species. However, there is limited knowledge about the CP genome of *Polyspora*, especially about *Polyspora axillaris*. Therefore, we sequenced the complete plastome of *P. axillaris*. The genomic structure of *P. axillaris* consists of a pair of IRs, a SSC region, and a LSC region that is similar to the majority of species in Theaceae, especially it has a stronger similarity with the *Polyspora* species (Fan and Ma 2022; Fan et al. 2024). We analyzed the cp genomes of *P. axillaris* and five other *Polyspora* species, which showed a same number of genes, CDS, tRNA, and rRNA. And the cp genomes of six Chinese *Polyspora* species varied in length between 156,452 bp (*P. chrysandra*) and 157,066 bp (*P. speciosa*), which is consistent with Fan and Ma (2022) and Fan et al. (2024) research findings. However, compared to the six species in the *Camellia* genus, we found that the similarity in gene and tRNA numbers between them was not as strong. And basically speaking, the number of genes in the chloroplast genome of *Camellia* species is 1–2 more than that of *Polyspora* species (Table S2). Perhaps during the evolutionary process, these 1–2 genes were the reason for the emergence of the genus *Polyspora* and *Camellia*.

In this study, we also reconstructed a phylogenetic tree using the sequenced chloroplast genomes of *P. axillaris*, as well as the chloroplast genome sequences of 5 other *Polyspora* species and 6 *Camellia* species (these sequences all downloaded from NCBI), confirming the close genetic relationship between the *Polyspora* and *Camellia* genus. *P. pentangensis* is located at the bottom of the genus *Polyspora*, while *P. speciosa*, *P. axillaris* and *P. hainanensis*, *P. chrysandra* and *P. dalgleishiana* formed three branches within the genus. This is similar to the results of Fan and Ma (2022) using chloroplast and ribosomal 18–26S rRNA data to construct a phylogenetic tree of the Theaceae family.

In summary, compared with previous research, we focus on analyzing the whole chloroplast genome sequence of *P. axillaris* and confirming its evolutionary position in the genus *Polyspora*. The chloroplast genome produced in our study provides additional genomic resources for the evolution



**Figure 2.** Gene map of *Polyspora axillaris* chloroplast genome. The circular chloroplast genome map displays 132 genes, including 37 tRNA genes and 8 rRNA genes, as well as 87 protein coding genes. The arrow inside the circle indicates the transcription direction of the gene inside the circle (clockwise). The arrow outside the circle indicates the transcription direction of the gene outside the circle (counterclockwise direction). The light gray line inside the circle represents the at content, and the dark gray area represents the GC content. At the bottom, genes from different functional groups are explained in various colors.

research of the chloroplast genomes in the *Polyspora* genus. We hope our research results can deepen understanding of *P. axillaris* and promote the phylogenetic study of Theaceae.

### Ethical approval

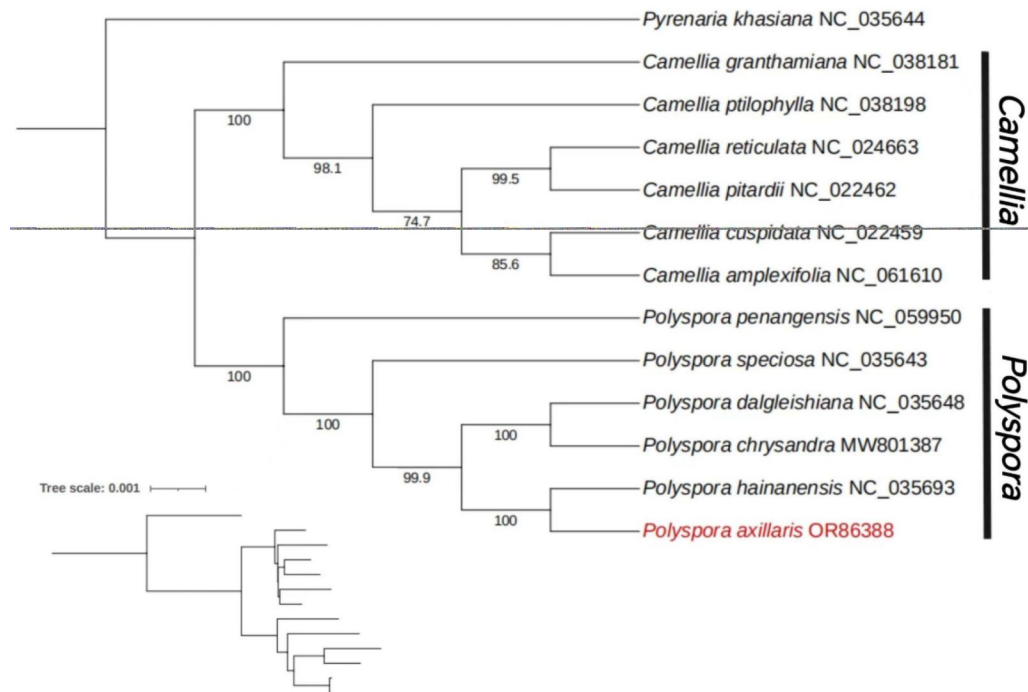
This study was approved by the Institutional Review Committee of Hainan University in Haikou, China. The collection of plant materials is carried out in accordance with the guidelines provided by Hainan University and the regulations of Hainan Province. The field research comply with the legislation of Hainan Province.

### Authors contributions

Hui-Ting Tang designed the project. Hui-Ting Tang, Shi-You Zuo and Pei-Feng Liu analyzed and interpreted the data; Hui-Ting Tang and Chun-Fang Li drafted the manuscript. All authors have made critical revisions to obtain intellectual content and approved the final version of the manuscript.

### Disclosure statement

No potential conflict of interest was reported by the author(s).



**Figure 3.** Maximum likelihood phylogenetic tree based on the complete chloroplast genomes of 12 species of theaceae, with *pyrenaria khasiana* (NC\_035644) as the outgroup. Numbers on the nodes are bootstrap values from 1000 replicates. All sequences were downloaded from NCBI's GenBank. The sequences used to build the tree is as follows: *P. hainanensis* (NC\_035693, Yu et al. 2017), *P. chrysandra* (MW801387, Fan and Ma 2022), *P. dalgleishiana* (NC\_035648, Yu et al. 2017), *P. speciosa* (NC\_035643, Yu et al. 2017), *P. penangensis* (NC\_059950), *camellia pitardii* (NC\_022462, Yang et al. 2013), *camellia reticulata* (NC\_024663, Huang et al. 2014), *camellia amplexifolia* (NC\_061610), *camellia cuspidata* (NC\_022459, Yang et al. 2013), *camellia ptilophylla* (NC\_038198), *camellia granthamiana* (NC\_038181), *pyrenaria khasiana* (NC\_035644, Yu et al. 2017). The chloroplast genome of *polyspora axillaris* is displayed in red and the NCBI GenBank accession numbers are shown in parentheses.

## Funding

No funding was received.

## Data availability statement

The genomic sequence data supporting this research result is publicly available in NCBI GenBank (<https://www.ncbi.nlm.nih.gov/>) with registration number OR863388. The relevant BioProject, Bio-Sample, and SRA numbers are PRJNA1043575, SAMN38414101, and SRR27011443, respectively (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1043575>). A specimen was deposited at Herbarium of Hainan University (<http://en.hainanu.edu.cn/>), contact person Hua-Feng Wang, [hfwang@hainanu.edu.cn](mailto:hfwang@hainanu.edu.cn) under the voucher number D J Chen, X R Ke, SK19.

## References

- Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. 2021. Twelve years of SAMtools and BCFtools. *Gigascience*. 10(2):giab008. doi:10.1093/gigascience/giab008.
- Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19:11–15.
- Fan Z, Gao C, Lin L. 2024. Phylogeographical and population genetics of *Polyspora* sweet in China provides insights into its phylogenetic evolution and subtropical dispersal. *BMC Plant Biol.* 24(1):89. doi:10.1186/s12870-024-04783-5.
- Fan ZF, Han L, Ma CL. 2021. Research advances of *Polyspora* sweet (Theaceae). *Guihaia*. 41(10):1755–1766. doi:10.11931/guihaia.gxzw.202010042.
- Fan ZF, Ma CL. 2022. Comparative chloroplast genome and phylogenetic analyses of Chinese *Polyspora*. *Sci Rep.* 12(1):15984. doi:10.1038/s41598-022-16290-4.

- Fan ZF, Qian SJ, Zhang YH, Ma CL. 2023. Characterization of the complete chloroplast genome of *Polyspora tiantangensis* (Theaceae), an endemic and endangered species in southwestern China. *Mitochondrial DNA B Resour.* 6(3):814–815. doi:10.1080/23802359.2021.1884013.
- Greiner S, Lehwark P, Bock R. 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* 47(W1):W59–W64. doi:10.1093/nar/gkz238.
- Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol.* 59(3):307–321. doi:10.1093/sysbio/syq010.
- Huang H, Shi C, Liu Y, Mao SY, Gao LZ. 2014. Thirteen *Camellia* chloroplast genome sequences determined by high-throughput sequencing: genome structure and phylogenetic relationships. *BMC Evol Biol.* 14(1):151. doi:10.1186/1471-2148-14-151.
- Jin J-J, Yu W-B, Yang J-B, Song Y, dePamphilis CW, Yi T-S, Li D-Z. 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.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* 28(12):1647–1649. doi:10.1093/bioinformatics/bts199.
- Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol.* 37(8):907–915. doi:10.1038/s41587-019-0201-4.
- Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44(W1):W242–W245. doi:10.1093/nar/gkw290.

- Li Y, Cao SY, Lin SJ, Zhang JR, Gan RY, Li HB. 2019. Polyphenolic profile and antioxidant capacity of extracts from *Gordonia axillaris* fruits. *Antioxidants*. 8(6):150. doi:10.3390/antiox8060150.
- Li Y, Li S, Lin SJ, Zhang JJ, Zhao CN, Li HB. 2017. Microwave-assisted extraction of natural antioxidants from the exotic *Gordonia axillaris* fruit: optimization and identification of phenolic compounds. *Molecules*. 22(9):1481. doi:10.3390/molecules22091481.
- 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.
- Ming TL, Bartholomew B. 2007. Theaceae. *Flora China*. 12:366–478.
- Minh BQ, Nguyen MAT, Von Haeseler A. 2013. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol*. 30(5):1188–1195. doi:10.1093/molbev/mst024.
- 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.
- Nguyet HNL, Luong VD, Nguyen VC, Pham TTD, Luu TT, The PV. 2020. An updated checklist of Theaceae and a new species of *Polyspora* from Vietnam. *Taiwania*. 65(2):216–227. doi:10.6165/tai.2020.65.216.
- Qin ZS. 1990. Mountain tree species—*Gordonia axillaris*. *Guangdong For Sci Technol*. 1:20–22.
- Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol*. 56(4):564–577. doi:10.1080/10635150701472164.
- Tamura K, Stecher G, Kumar S. 2021. MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol*. 38(7):3022–3027. doi:10.1093/molbev/msab120.
- Yang JB, Yang SX, Li HT, Yang J, Li DZ. 2013. Comparative chloroplast genomes of *Camellia* species. *PLOS One*. 8(8):e73053. doi:10.1371/journal.pone.0073053.
- Yu X-Q, Gao L-M, Soltis DE, Soltis PS, Yang J-B, Fang L, Yang S-X, Li D-Z. 2017. Insights into the historical assembly of East Asian subtropical evergreen broadleaved forests revealed by the temporal history of the tea family. *New Phytol*. 215(3):1235–1248. doi:10.1111/nph.14683.
- Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT. 2020. PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol Ecol Resour*. 20(1):348–355. doi:10.1111/1755-0998.13096.