


The complete chloroplast genome sequence of *Quercus kerrii* (Fagaceae), and comparative analysis with related species

Ze-Yi Ouyang^{a,b}, Zhi-Hui Li^a, Ran Chen^a, Xiao-Long Jiang^a  and Ji-You Wu^c

^aCollege of Forestry, Central South University of Forestry and Technology, Changsha, China; ^bHunan Botanical Garden, Changsha, China; ^cHunan Academy of Forestry, Changsha, China

ABSTRACT

Quercus kerrii Craib 1911 (section *Cyclobalanopsis*) is a widespread tree species in the tropical seasonal forests of southwest China and Northern Indo-China areas. In this study, we sequenced, assembled and annotated the complete chloroplast genome of *Q. kerrii*. The circular genome was 160,743 bp in length and had a GC content of 36.89%. The *Q. kerrii* chloroplast genome has a typical quadripartite structure, including two inverted repeat regions (length, 25,825 bp; GC content, 42.76%), a large single-copy region (length, 90,196 bp; GC content, 34.74%), and a small single-copy region (length, 18,897 bp; GC content, 30.60%). Genome annotation has indicated that the *Q. kerrii* chloroplast genome contained 131 genes, including 86 protein-coding genes, 37 tRNA, and eight rRNA. The phylogenetic tree showed that *Q. kerrii* had a close relationship with *Q. schottkyana* Rehder & E.H.Wilson 1916.

ARTICLE HISTORY

Received 9 December 2022
Accepted 13 April 2023

KEYWORDS

Quercus section
Cyclobalanopsis; southwest
China and Northern Indo-
China; phylogenetic analysis

Introduction





Oaks (*Quercus* L.) are the most widespread lineage in the Northern Hemisphere. The recent phylogeny well resolved two subgenera of the genus: subgenus *Quercus* and subgenus *Cerris* (Denk et al. 2017; Deng et al. 2018). Three sections, *Cyclobalanopsis*, *Cerris*, and *Ilex*, belong to the subgenus *Cerris*. Section *Cyclobalanopsis* (also known as ring-cupped oaks) consists of ca. 100 species that are endemic to East Asia. Oaks have important economic and ecological values, and their local adaptation and evolutionary history have attracted the attention of many researchers (Cavender-Bares 2019). Benefiting from the development of sequencing technology and the reduction in sequencing costs, more than 40% of the Chinese oak chloroplast genome have been assembled (Fang et al. 2022). These data provide an important resource for understanding the genetic diversity and evolutionary history of oaks. Southwest China and Northern Indo-China are the ancestral distribution area and the species divergence center for the section *Cyclobalanopsis* (Deng et al. 2018); however, the chloroplast genome information of the ring-cupped oaks in this region is rather limited.

Quercus kerrii Craib 1911 (Figure 1) belongs to the section *Cyclobalanopsis* and is predominantly distributed in southwest China, as a keystone tree species of the tropical seasonal broadleaved forests in southwest China and Northern Indo-China. Comparing the specimen collection records with the distribution records of *Q. kerrii* from our field surveys, many of the distribution sites where the species was present



Figure 1. Morphological characteristics of *Quercus kerrii*. *Quercus kerrii* is a tree to 30 m tall. Branchlets densely tawny tomentose, glabrescent, or rarely remaining hairy. Petiole 1–2 cm, tomentose; leaf blade oblong-elliptic, lanceolate, or oblong-oblancoelate, abaxially stellate hairy. Cupule discoid, covering base to 1/2 of nut, outside grayish to pale brown tomentose. Nut oblate apex depressed to flat. Photos were taken by Xiao-Long JIANG in Lincang, Yunnan Province, China.

decades ago did not exist presently. This means that the habitat of *Q. kerrii* has experienced extensive degradation, which was driven by a contemporary climate change and human activities. Our previous species distribution modeling and population genetic study by using 10 nrSSR and the DNA intergenic spacers of two chloroplasts showed the Asian summer monsoon and the Indian summer monsoon boosted the local adaptation of its population (Jiang et al. 2018).

CONTACT Xiao-Long Jiang  xiaolongjiang1@gmail.com  College of Forestry, Central South University of Forestry and Technology, Changsha, Hunan, China; Ji-You Wu  805610788@qq.com  Hunan Academy of Forestry, Changsha, Hunan, China

© 2023 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.

However, these molecular markers only contain very few informative loci to profile the population's genetic structure. To provide more information on variation, in this study, we assembled and annotated the complete chloroplast genome of *Q. kerrii*. Meanwhile, a phylogenetic tree was constructed to determine the phylogenetic position of *Q. kerrii*.

Materials and methods

The samples were collected from a wild population of *Q. kerrii* in August 2014 in Lincang County, Yunnan Province, China (23.65° N, 100.32° E, 1203 m). Fresh leaf samples were immediately cleaned and stored in silica gels. After returning to the laboratory, the dried samples were stored in a refrigerator at -20°C before DNA extraction. Voucher specimens

were deposited in the Herbarium of Shanghai Chenshan Botanical Garden (CSH, <https://cfh.ac.cn/Subsite/Default.aspx?siteid=CSH>, Contact Person: Bin-Jie Ge, gebinjie123@163.com, under the voucher number DM9122). Total genomic DNA was extracted from the silica-dried leaves using the modified CTAB method (Doyle 1987). After the quality determination of the genomic DNA, the qualified DNA was preserved in the Laboratory of Forestry Genetics, Central South University of Forestry and Technology (contact person: Xiaolong Jiang, xiaolongjiang1@gmail.com, under the accession number DM9122).

The genome skimming sequencing was conducted by the Personalbio Company (Shanghai, China) using the Illumina NovaSeq 6000 platform in paired-end 150bp mode. Raw reads were cleaned using fastp v0.21.0 (Chen et al. 2018).

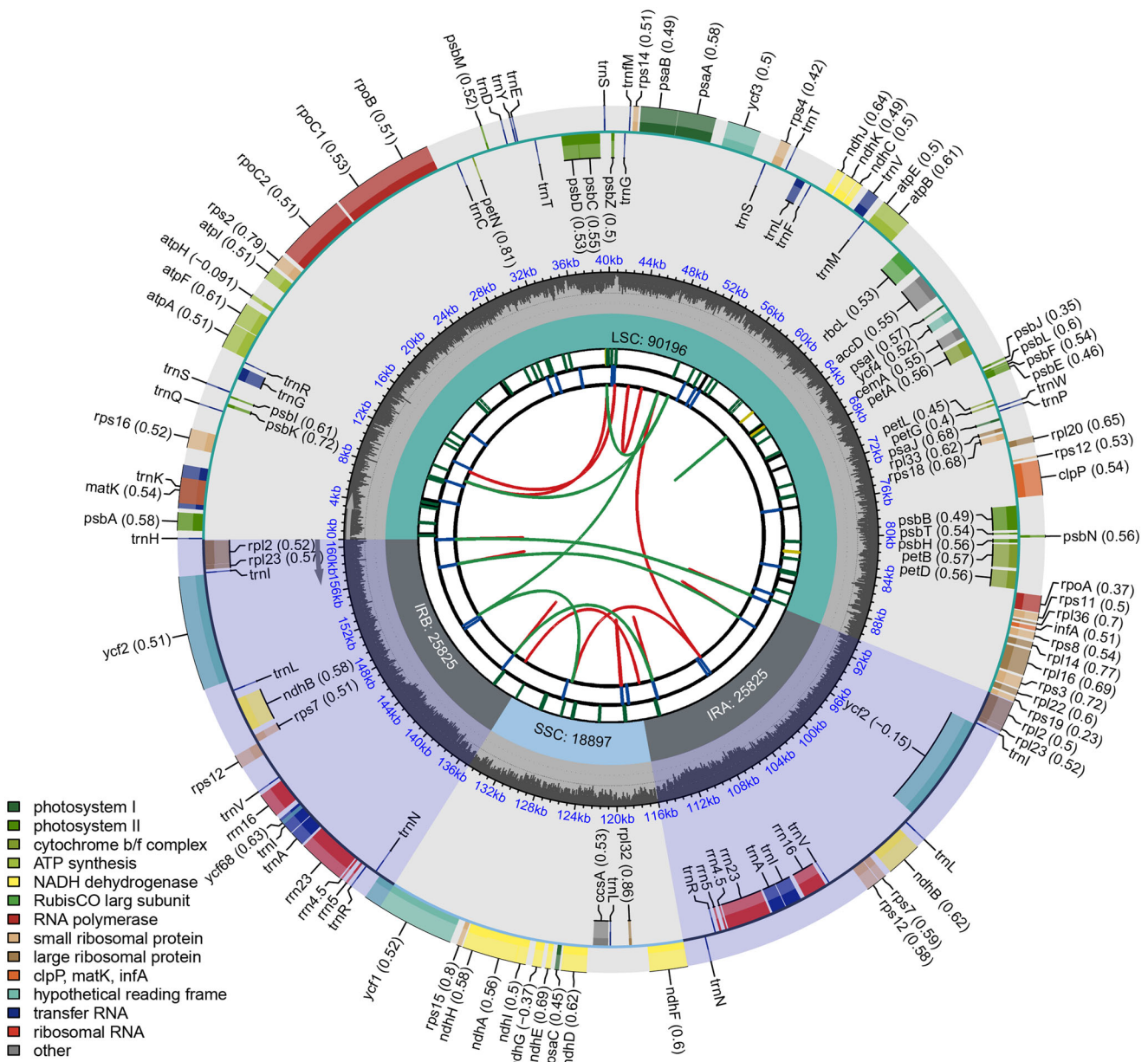


Figure 2. Characterization of the chloroplast genome of *Quercus kerrii*. From the center going outward, the first circle shows the distribution of the repeats connected with red (the forward direction) and green (the reverse direction) arcs. The second circle shows the tandem repeats marked with short bars. The third circle shows the LSC, SSC, IRa, and IRb regions. The fourth circle shows the percent of GC content. The next circle shows the genes having different colors based on the functional groups. The functional classification is shown at the bottom left.

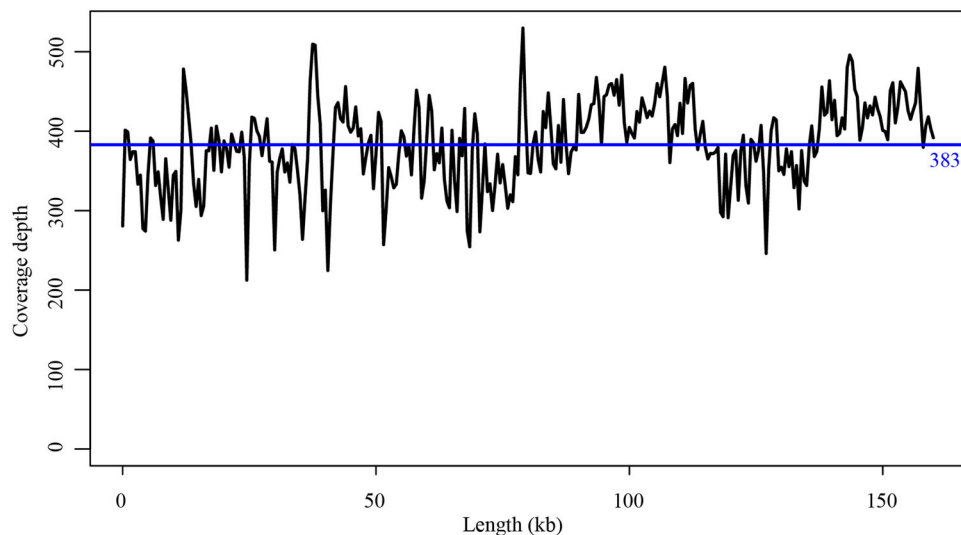


Figure 3. Depth of coverage of the data used to assemble the chloroplast genome of *Quercus kerrii*.

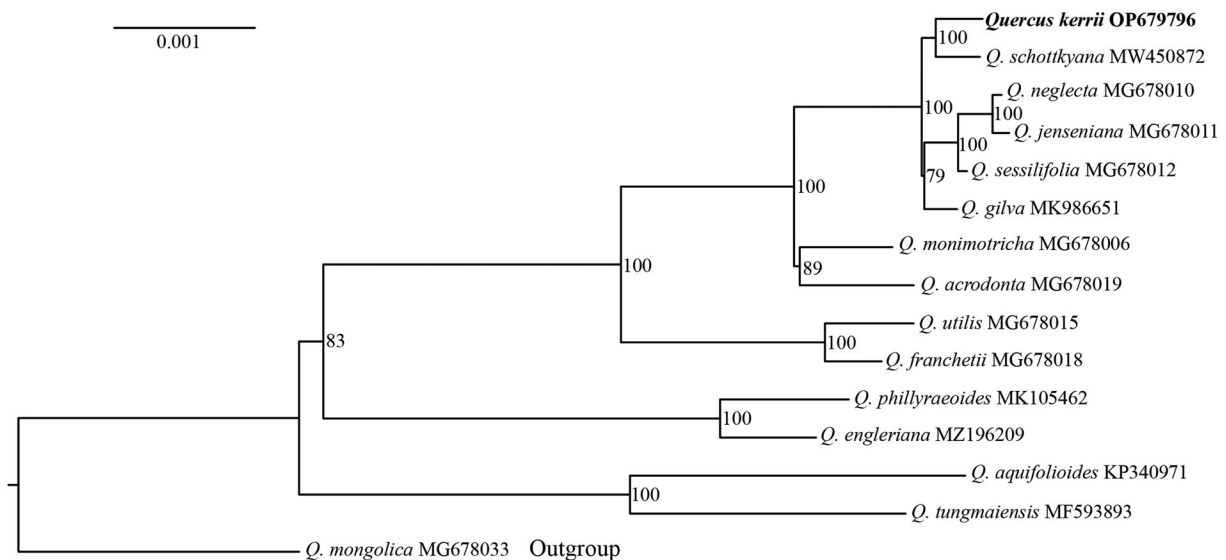


Figure 4. The maximum-likelihood phylogenetic tree of *Quercus kerrii* and other 14 East Asian oaks using complete chloroplast genome sequences. *Quercus mongolica* (subgenus *Quercus* section *Quercus*) was used as outgroup to root the tree. The values at the node indicate the bootstrap support value.

The chloroplast genome was assembled using GetOrganelle v1.7.2beta (Jin et al. 2020) with the default parameters. To save time and computer resources, 5,000,000 reads were randomly selected from the clean reads for *de novo* assembly. The coverage depth of the sequences in the genome was calculated using bbmap (<https://github.com/BioInfoTools/BBMap>). Chloroplast genome annotation was performed using the Plastid Genome Annotator (PGA) (Qu et al. 2019) with the default parameters. The chloroplast genome of another oak of section *Cyclobalanopsis*, *Quercus schottkyana* Rehder & E.H.Wilson 1916, was used as the reference genome. The chloroplast genome of *Q. kerrii* was visualized and SSR was identified using CPGview (Figure 2) (Liu et al. 2023).

To infer the phylogenetic relationship of *Q. kerrii*, a maximum-likelihood (ML) tree was constructed based on the chloroplast genome of *Q. kerrii* in this study and other 14 oaks species downloaded from NCBI (<https://www.ncbi.nlm.nih.gov/>).

The downloaded chloroplast genome sequences and their sources were summarized as below: *Q. monimotricha* Hand.-Mazz. 1929 MG678006, *Q. neglecta* (Schottky) Koidz. 1916 MG678010, *Q. jenseniana* Hand.-Mazz. 1922 MG678011, *Q. sessilifolia* Blume 1851 MG678012, *Q. utilis* Hu & Cheng 1951 MG678015, *Q. franchetii* Skan 1899 MG678018, *Q. acrodonta* Seemen 1897 MG678019, *Q. mongolica* Fisch. ex Turcz. 1850 MG678033 (Yang et al. 2021), *Q. schottkyana* MW450872 (Li, Luo, et al. 2021), *Q. gilva* Blume 1851 MK986651 (Zeng et al. 2019), *Q. phillyraeoides* A.Gray 1859 MK105462 (Pang et al. 2019), *Q. engleriana* Seemen 1897 MZ196209 (Li, Li, et al. 2021), *Q. aquifolioides* Rehder & E.H.Wilson 1916 KP340971 (unpublished), and *Q. tungmaiensis* Y.T.Chang 1966 MF593893 (Yang et al. 2018).

Quercus consists of two subgenera, subgenus *Quercus* and subgenus *Cerris* (Denk et al. 2017). Except *Q. mongolica*, which belongs to the subgenus *Quercus*, other species are

from the subgenus *Cerris*. Therefore, *Q. mongolica* was used as the outgroup to root the tree in this study. The chloroplast genome sequences were aligned using MAFFT v7.4.75 (Katoh and Standley 2013). The phylogenetic tree was reconstructed using IQ-TREE v1.6.12 (Chernomor et al. 2016). The best evolution model was estimated using the IQ-TREE software. To test the credibility of the tree topology, 1000 fast bootstrap replicates were performed.

Results and discussion

A total of 133,363,806 reads were retained after removing the low-quality reads from the raw reads. The chloroplast genome of *Q. kerrii* is a circular molecule of 160,743 bp in length with a 36.89% GC content. Similar to other oaks, the *Q. kerrii* chloroplast genome has a typical quadripartite structure, including two inverted repeat regions (length, 25,825 bp; GC content, 42.76%), a large single-copy region (length, 90,196 bp; GC content, 34.74%) and a small single-copy region (length, 18,897 bp; GC content, 30.60%). The range of coverage depth for reads used to assemble the genome was from 213 to 530 (mean: 383, Figure 3). Totally, 131 genes were annotated using PGA, including 86 protein-coding genes, 37 tRNA (transfer RNA) and eight rRNA (ribosomal RNA). A total of 86 SSRs were identified, among which 81 were single-base repeats and five were two-base repeats. The best model used to reconstruct the phylogenetic tree was K3Pu + F + R2. The phylogenetic tree showed that *Q. kerrii* was more closely related to *Q. schottkyana* with 100% bootstrap support (Figure 4).

In this study, the monophyletic status of the section *Cyclobalanopsis* was well resolved by the phylogenetic tree obtained from the complete chloroplast genome sequences. The section *Ilex* species showed a paraphyletic pattern, and it was consistent with the recent study from Zhou et al. (2022), which however exhibited prominent incongruence to the RAD-seq based phylogeny inferred by Jiang et al. (2019) and Hipp et al. (2020). Likewise, within the section *Cyclobalanopsis*, two main clades corresponding to the Southeast China and Southwest China distribution of the species were detected, which showed significant differences in comparison with the previous phylogeny inferred by RAD-seq (Deng et al. 2018). All these discordance between the plastid and nuclear genome suggested that introgression/hybridization might generally occur in East Asia oak lineages. The complete chloroplast genome sequences and cpSSRs markers developed in this study provided useful tools for the future population genetic and biogeographic studies on the Asian ring-cupped oaks.

Author contributions

Ji-You Wu and Zhi-Hui Li designed the study; Xiao-Long Jiang collected the samples; Ze-Yi Ouyang, Xiao-Long Jiang, and Ran Chen analyzed and interpreted the data; Ze-Yi Ouyang and Xiao-Long Jiang drafted the manuscript. All authors read and approved the final version of the manuscript.

Ethics statement

Quercus kerrii was not listed as a national and provincial key protected wild plant in China nor a threatened species on the IUCN Red List. For research purpose, therefore, no specific permissions or licenses were needed for the sampling of *Q. kerrii* according to the regulations of the People's Republic of China on the protection of wild plants.

Disclosure statement

The authors declare no conflict of interest.

Funding

This work was supported by the Hunan Natural Science Foundation [2022JJ40861], the Education Department of Hunan Province [21A0158], and Science and Technology Research Project of Hunan Province Forestry Department [XLK201706].

ORCID

Xiao-Long Jiang  <http://orcid.org/0000-0003-3861-1109>

Data availability statement

The assembly chloroplast genome sequence of *Quercus kerrii* is available in GenBank of NCBI (<https://www.ncbi.nlm.nih.gov/>) under the accession numbers OP679796.1. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA891795, SRR21966525, and SAMN31336113, respectively.

References

- Cavender-Bares J. 2019. Diversification, adaptation, and community assembly of the American oaks (*Quercus*), a model clade for integrating ecology and evolution. *New Phytol.* 221(2):669–692.
- Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics.* 34(17):i884–i890.
- Chernomor O, von Haeseler A, Minh BQ. 2016. Terrace aware data structure for phylogenomic inference from supermatrices. *Syst Biol.* 65(6): 997–1008.
- Deng M, Jiang XL, Hipp AL, Manos PS, Hahn M. 2018. Phylogeny and biogeography of East Asian evergreen oaks (*Quercus* section *Cyclobalanopsis*; Fagaceae): insights into the Cenozoic history of evergreen broad-leaved forests in subtropical Asia. *Mol Phylogenet Evol.* 119:170–181.
- Denk T, Grimm GW, Manos PS, Deng M, Hipp AL. 2017. An updated infra-generic classification of the oaks: review of previous taxonomic schemes and synthesis of evolutionary patterns. In: Gil-Pelegrín E, Peguero-Pina JJ, Sancho-Knapik D, editors. *Oaks physiological ecology. Exploring the functional diversity of genus Quercus L.* New York: Springer; p. 13–38.
- Doyle JJ. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull.* 19:11–15.
- Fang YM, Zhu ZY, Li Y, Li X. 2022. Advances and prospects of oak biology based on genomics. *J Nanjing Forest Univ.* 46:64–72.
- Hipp AL, Manos PS, Hahn M, Avishai M, Bodénès C, Cavender-Bares J, Crowl AA, Deng M, Denk T, Fitz-Gibbon S, et al. 2020. Genomic landscape of the global oak phylogeny. *New Phytol.* 226(4):1198–1212.
- Jiang XL, An M, Zheng SS, Deng M, Su ZH. 2018. Geographical isolation and environmental heterogeneity contribute to the spatial genetic patterns of *Quercus kerrii* (Fagaceae). *Heredity.* 120(3):219–233.
- Jiang X-L, Hipp AL, Deng M, Su T, Zhou Z-K, Yan M-X. 2019. East Asian origins of European holly oaks (*Quercus* section *Ilex* Loudon) via the Tibet-Himalaya. *J Biogeogr.* 46(10):2203–2214.

- 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.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
- Li X, Li Y, Sylvester SP, Zang M, El-Kassaby YA, Fang Y. 2021. Evolutionary patterns of nucleotide substitution rates in plastid genomes of *Quercus*. *Ecol Evol.* 11(19):13401–13414.
- Li TT, Luo CS, Mou HL, Jiang XL, Xu GB. 2021. The complete chloroplast genome sequence of *Quercus schottkyana*, and comparative analysis with related species. *Mitochondrial DNA B Resour.* 6(9):2607–2609.
- 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.
- Pang X, Liu H, Wu S, Yuan Y, Li H, Dong J, Liu Z, An C, Su Z, Li B. 2019. Species Identification of Oaks (*Quercus* L., Fagaceae) from gene to genome. *Int J Mol Sci.* 20(23):5940.
- Qu XJ, Moore MJ, Li DZ, Yi TS. 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. *Plant Methods.* 15:50.
- Yang Y, Zhang H, Ren T, Zhao G. 2018. Characterization of the complete plastid genome of *Quercus tungmaiensis*. *Conserv Genet Resour.* 10(3): 457–460.
- Yang Y, Zhou T, Qian Z, Zhao G. 2021. Phylogenetic relationships in Chinese oaks (Fagaceae, *Quercus*): evidence from plastid genome using low-coverage whole genome sequencing. *Genomics.* 113(3): 1438–1447.
- Zeng QM, Liu B, Lin RQ, Jiang YT, Liu ZJ, Chen SP. 2019. The complete chloroplast genome sequence of *Quercus gilva* (Fagaceae). *Mitochondrial DNA B Resour.* 4(2):2493–2494.
- Zhou BF, Yuan S, Crowl AA, Liang YY, Shi Y, Chen XY, An QQ, Kang M, Manos PS, Wang BS. 2022. Phylogenomic analyses highlight innovation and introgression in the continental radiations of Fagaceae across the Northern Hemisphere. *Nat Commun.* 13:1320.