Postprint: Plastome Capture History between Quercus subsect. Cerris and sect. Ilex

Authors: Yu Shuliang, Li Zhimin, Ma Xiangguang, Sun Wenguang

Date: 2023-12-24T00:00:00+00:00

Abstract

Quercus subsect. Campylolepides includes three species: Quercus acutissima, Q. variabilis, and Q. chenii, constituting the East Asian branch of Quercus section Cerris. Previous studies have conducted in-depth investigations into the speciation and phylogeography of Quercus subsect. Campylolepides or species within this subsection, and have also discovered that section Cerris experienced ancient genetic introgression with section Ilex, resulting in plastid capture. However, the specific evolutionary history of plastids between Quercus subsect. Campylolepides and section Ilex remains unclear. This study performed shallow whole-genome sequencing on 15 samples from section Ilex and integrated previously published resequencing data totaling 325 individuals from Quercus subsect. Campylolepides and its close relatives, including 276 individuals from 19 populations of the three species within Quercus subsect. Campylolepides. These data were utilized for plastome assembly and analysis. The results demonstrate: (1) Shared haplotypes exist among the three species of Quercus subsect. Campylolepides, yet the plastid haplotypes of the entire subsect. Campylolepides essentially constitute a monophyletic clade nested within a branch composed of section Ilex species from central China to the Liangshan region of Sichuan; (2) Among Quercus subsect. Campylolepides species, a relict haplotype from the Liaodong Peninsula clusters with the section Ilex species Quercus pseudosetulosa as a single clade; (3) Both plastid capture events occurred in the mid-Miocene, and no plastid capture has occurred between Quercus subsect. Campylolepides and section Ilex since then, suggesting that Quercus subsect. Campylolepides and section Ilex species have currently formed nearly complete reproductive isolation.

Full Text

DOI:10.11931/guihaia.gxzw202309006

The Plastid Capture History of Quercus subsect. Campylolepides and Quercus section Ilex

YU Shuliang[1], LI Zhimin[1], MA Xiangguang[2], *SUN Wenguang*[1] (1. College of Life Sciences, Yunnan Normal University, Kunming 650500, China; 2. CAS Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China)

Abstract: Quercus subsect. Campylolepides includes three species: Q. acutissima, Q. variabilis, and Q. chenii, representing the East Asian clade of Quercus section Cerris. Previous studies have investigated speciation and phylogeography within this subsection and have documented ancient gene introgression between section Cerris and section Ilex that led to plastid capture. However, the specific evolutionary history of plastids between Quercus subsect. Campylolepides and Quercus section Ilex remains unclear. This study performed low-coverage genome sequencing on 15 samples from Quercus section Ilex and integrated previously published resequencing data from Quercus subsect. Campylolepides and its relatives, totaling 325 resequencing datasets, including 276 individuals from 19 populations across three species of Quercus subsect. Campylolepides. These data were used for plastid genome assembly and analysis. The results indicate: (1) Shared haplotypes exist among the three species of Quercus subsect. Campylolepides, but the plastid haplotypes of the entire subsection essentially form a monophyletic clade nested within a branch composed of Quercus section Ilex species distributed from central China to Liangshan Prefecture in Sichuan; (2) A relict haplotype of Q. acutissima from the Liaodong Peninsula clusters with Q. pseudosetulosa from Quercus section Ilex; (3) Both plastid capture events occurred in the middle Miocene, after which no plastid capture occurred between Quercus subsect. Campylolepides and Quercus section Ilex, suggesting that nearly complete reproductive isolation has formed between these groups.

Keywords: Quercus subsect. Campylolepides, Quercus section Cerris, Quercus section Ilex, plastid capture, plastid genome introgression

Chinese Library Classification: Q943 **Document Code:** A

Funding: This study was supported by the National Natural Science Foundation of China (31971393, 31960046), the Second Tibetan Plateau Scientific Expedition and Research Program (2019QZKK0502), and the Yunnan Provincial Basic Research Project-Youth Project (202201AU070057).

Author Information: YU Shuliang (1998-), Master's student, research focus: plant systematics and evolution, E-mail: yushuliang163@163.com. *Corresponding authors: MA Xiangguang, Ph.D., Assistant Researcher, research focus: plant systematics and evolution, E-mail: maxiangguang@mail.kib.ac.cn; SUN Wenguang, Ph.D., Lecturer, research focus: plant systematics and evolution, E-mail: sunwenguang@vip.163.com.

Introduction

The phenomenon of "nuclear-plastid incongruence" is frequently encountered in phylogenetic studies, where the topological structures of phylogenetic trees constructed from plastid and nuclear genomes are inconsistent. This is typically caused by incomplete lineage sorting (Meleshko et al., 2021), hybridization, or gene introgression (Stull et al., 2023). Plastid capture is a special phenomenon resulting from interspecific hybridization. When different species hybridize, repeated backcrossing of the hybrid with one parental species can lead to the incorporation of the plastid genome from one species into another, ultimately resulting in the nucleus and cytoplasm being derived from different species (Rieseberg et al., 1991; Kleinkopf et al., 2019). Genetic modeling studies have found that plastid capture readily occurs if cytoplasmic replacement provides advantages in seed production (Tsitrone et al., 2003). Numerous case studies, such as those on *Heuchera* (Liu et al., 2020), *Salix* (Percy et al., 2014), and *Populus* (Liu et al., 2017), demonstrate that plastid capture is widespread in angiosperm groups. In some taxa, relying solely on plastid genes to construct phylogenetic relationships can lead to misinterpretation of species relationships (Rieseberg et al., 1991). In certain groups, plastid genes of closely related species are often associated with geographic location rather than species phylogeny, a phenomenon that is particularly common in *Quercus* species and has been extensively documented (Simeone et al., 2016; Yang et al., 2016; Tekpinar et al., 2021; Yang et al., 2021).

Quercus acutissima, *Q. variabilis*, and *Q. chenii* are native to East Asia. In the domestic classification system of *Quercus*, these three species are grouped together due to their similar morphology (Zhou, 1992). In the recent phylogenetic framework of *Quercus* by Hipp et al. (2020), these three species constitute the East Asian branch of *Quercus* section *Cerris*—defined by Camus (1936) as Subsect. *Campylolepides*. *Quercus acutissima* and *Q. variabilis* are important components of warm-temperate broadleaf forests in East Asia (Li et al., 2014), while *Q. chenii* is restricted to East China (Li et al., 2016). As major components of warm-temperate broadleaf forests in East Asia, *Quercus* subsect. *Campylolepides* species have been extensively studied by scholars worldwide regarding hybridization, introgression, and speciation. Frequent hybridization and introgression occur among species within this subsection; populations of *Q. acutissima* and *Q. variabilis* in similar habitats may share adaptive introgression fragments, with adaptive introgression loci tending to be distributed in regions with suppressed recombination (Fu et al., 2022). Coding regions of *Quercus* subsect. *Campylolepides* species generally exhibit extensive positive and negative selection, with linked selection further driving genetic differentiation (Liang et al., 2022). Li (2019) conducted a comprehensive study on the phylogeography and evolutionary history of *Quercus* subsect. *Campylolepides*, concluding that the genetic structure of *Q. acutissima* shows an east-west differentiation bounded by the second and third topographic steps, while the genetic structure of *Q. chenii* is constrained by the topography and climate of East China.

Moreover, haplotypes are extensively shared among these species, with interspecific haplotype sharing primarily resulting from interspecific introgression and ancestral polymorphism sharing.

In phylogenetic relationships constructed from reduced-representation genome data, *Quercus* subsect. *Campylolepides* species are sister to European *Quercus* section *Cerris* species and together form a monophyletic clade (Hipp et al., 2020). Zhou et al. (2022) conducted a phylogenomic study of the entire Fagaceae family and discovered numerous potential plastid capture events among different sections of *Quercus* and between *Quercus* and related genera. In that study, *Quercus* subsect. *Campylolepides* species from Hubei, Guangxi, Shandong, Hunan, Jiangsu, and Yunnan did not cluster with European *Quercus* section *Cerris* species but instead formed a sister clade with *Quercus* section *Ilex* species from Hubei, specifically *Q. cocciferoides* and *Q. dolicholepis*. This confirms that East Asian *Quercus* section *Cerris* species experienced plastid capture with *Quercus* section *Ilex* species. Simeone et al. (2018) investigated the phylogeographic relationships of 13 *Quercus* section *Cerris* species and suggested that this section likely originated in Northeast Asia during the early Oligocene before dispersing to West Asia, where it came into contact with local *Quercus* section *Ilex* species and engaged in interspecific gene exchange.

In summary, previous research has conducted extensive and in-depth studies on the phylogeography of *Quercus* subsect. *Campylolepides* or individual species within the subsection, as well as on the patterns of genomic differentiation and speciation mechanisms among these species. However, the plastid capture history between East Asian *Quercus* subsect. *Campylolepides* and *Quercus* section *Ilex* remains unclear. Leveraging the substantial genomic data accumulated from previous studies, we integrated *Quercus* subsect. *Campylolepides* into a *Quercus* plastid genome phylogenetic framework that includes its close relatives to investigate plastid lineage sharing and evolutionary history with these related groups. This study focused on performing low-coverage genome sequencing of *Quercus* section *Ilex* samples located within the distribution range of *Quercus* subsect. *Campylolepides* and integrated previously published *Quercus* resequencing data from Zhou et al. (2022), Fu et al. (2022), and Liang et al. (2022) for complete plastid genome assembly and annotation. Phylogenetic analysis of plastid genomes from *Quercus* subsect. *Campylolepides* and its relatives was conducted to address the following questions: (1) How many plastid capture events have occurred between Chinese *Quercus* subsect. *Campylolepides* species and *Quercus* section *Ilex* species, and have these species shared closely related plastid genotypes with other *Quercus* groups, especially *Quercus* section *Ilex*, in recent history? (2) What were the timing and possible processes of ancient gene exchange between Chinese *Quercus* subsect. *Campylolepides* species and *Quercus* section *Ilex* species?

Materials and Methods

1.1 Sample Information

This study newly sequenced 15 samples from *Quercus* section *Ilex* (Table 1), with the sampling principle of selecting specimens located within the distribution range of *Quercus* subsect. *Campylolepides* whenever possible. These were integrated with three previously published resequencing datasets, totaling 325 sequences. The three resequencing datasets included: 2 sequences of *Trigonobalanus*, 2 of *Castanea*, 2 of *Castanopsis*, 2 of *Lithocarpus*, 5 of *Quercus* section *Cerris*, 7 of *Quercus* section *Ilex*, 8 of *Quercus* section *Cyclobalanopsis*, and 1 of *Fagus* from Zhou et al. (2022), totaling 29 sequences. Samples from Fu et al. (2022) and Liang et al. (2022) were also integrated, including 5 sequences of *Quercus* subgenus *Quercus* from Fu et al.'s study, 7 *Q. acutissima* populations (BWL, GSM, LCX, TCX, TMS, ZHX, ZJS), and 9 sympatric populations of *Q. acutissima* and *Q. variabilis* (DWX, JJL, KUM, KYS, LFS, LSK, MPX, PWX, TBS). From Liang et al.'s study, 2 sympatric populations of *Q. variabilis* and *Q. chenii* (HYX, JJX) and 1 sympatric population of *Q. acutissima*, *Q. variabilis*, and *Q. chenii* (LYG) were selected. In total, 19 populations comprising 276 individuals were included (see Table 2).

1.2 Plastid Genome Assembly

The sequenced samples consisted of silica gel-dried leaves. Total DNA was first extracted using the CTAB method, then fragmented using a Covaris ultrasonic disruptor. After successfully constructing 350 bp short-fragment libraries, sequencing was performed on the Illumina high-throughput sequencing platform. The sequencing data volume for each sample was 2 Gb, and the entire sequencing process was completed by Beijing Novogene Bioinformatics Technology Co., Ltd. The newly sequenced samples from this study and the three previously published resequencing datasets constituted the complete sample dataset for this research. Resequencing data were downloaded using SRA-TOOL with Bioproject numbers: PRJNA773751, PRJNA763710, and PRJNA769460. Subsequently, assembly was performed using GetOrganelle v1.6.2 (Jin et al., 2020) with K values of 105 and 121, and R value of 15. The sequence with the highest similarity obtained from NCBI BLAST was the chloroplast genome of *Q. chenii* (NC_{039428}), which was used as the reference sequence for annotating all plastid genomes in this study using PGA (Qu et al., 2019). The annotations were then checked and supplemented in Geneious v9.0.2. A total of 325 complete plastid genome sequences from *Q. acutissima*, *Q. variabilis*, and their relatives were obtained. Protein-coding sequences (CDS) were extracted and concatenated using the Export Annotations tool in Geneious for subsequent analysis.

1.3.1 Haplotype Sharing

The number and distribution of haplotypes among the 276 *Quercus* subsect. *Campylolepides* samples were determined using DnaSP v6.0 (Rozas et al., 2017).

A maximum likelihood (ML) tree of haplotypes was constructed using RAxML v8.0 (Stamakis, 2014), and haplotype geographic distribution maps were generated in ArcMap v10.4.1.

1.3.2 Phylogenetic Relationships and Divergence Time

The 276 obtained plastid sequences from *Quercus* subsect. *Campylolepides* species were aligned using MAFFT v7.490 (Katoh et al., 2013), followed by sliding window analysis in DnaSP v6.0 (Peakall et al., 2012) with a window size of 600 bp and step size of 200 bp. For phylogenetic tree construction, the 276 *Quercus* subsect. *Campylolepides* sequences were reduced by selecting sequences from different species and populations for each haplotype, yielding 75 plastid genome sequences from *Quercus* subsect. *Campylolepides*. These were combined with 49 sequences from close relatives to form the phylogenetic framework. Protein-coding regions (CDS) from the 124 plastid genomes were extracted in Geneious, and 76 shared protein-coding genes were concatenated. The concatenated sequences were aligned using MAFFT v7.490 (Katoh et al., 2013), trimmed using trimAl v1.4 (Capella-Gutiérrez et al., 2009), and finally used to construct phylogenetic trees with RAxML v8.0. The trees were edited and visualized in iTOL (Letunić et al., 2021).

Divergence times for *Quercus* subsect. *Campylolepides* species and their relatives were inferred using the BEAST 2 software package (Bouckaert et al., 2014). The nucleotide substitution model was calculated using jModelTest 2 (Darriba et al., 2012). The Yule Model was employed with a relaxed molecular clock and an evolutionary rate of 1.0. Two fossil calibration points were used to calibrate divergence times: the first fossil record for the earliest Fagaceae macrofossil (Grímsson et al., 2016) was used to calibrate the divergence between the outgroup *Fagus* and the ingroup; the second fossil record for the earliest *Castanopsis* fossil (Wilf et al., 2019) was used to calibrate the divergence between *Castanopsis* and *Castanea*. Three independent MCMC simulations were run, each with 100,000,000 generations and sampling every 1,000 generations. Convergence was assessed in Tracer v1.7.1 by checking whether the ESS values exceeded 200 (Rambaut et al., 2018). ESS values greater than 200 indicated convergence. Finally, TreeAnnotator v2.6.0 was used to construct the maximum clade credibility tree, with the first 30% discarded as burn-in.

Results

2.1 Haplotype Sharing Among Different Species of *Quercus* subsect. *Campylolepides*

From 276 individuals of Chinese *Quercus* subsect. *Campylolepides* species, 77 shared protein-coding genes were obtained, with a concatenated length of 67,863 bp. A total of 43 haplotypes were detected. Haplotype sharing was observed among the three species (Figure 1 [Figure 1: see original paper]). Twelve haplotypes were shared between *Q. acutissima* and *Q. variabilis*, H14 was shared between *Q. acutissima* and *Q. chenii*, H19 was shared between *Q. variabilis* and *Q. chenii*, and H4 was shared among all three species. Twenty-eight private haplotypes were detected, including 19 private to *Q. acutissima*, 5 private to *Q. variabilis*, and 4 private to *Q. chenii*. Among the 15 shared haplotypes in *Quercus* subsect. *Campylolepides*, 7 were distributed across multiple different populations, primarily involving mixing among 11 populations from Northeast, East, Southwest, and South China. In contrast, the 28 species-specific haplotypes tended to be restricted to single populations, with only 4 species-specific haplotypes appearing in different populations.

Figure 1 Haplotype-sharing patterns among three species of Subsect. *Campylolepides*

Figure 2 [Figure 2: see original paper] Geographical distribution of 43 haplotypes from Subsect. *Campylolepides*

The total length of the 276 *Quercus* subsect. *Campylolepides* plastid genomes ranged from 160,806 to 161,311 bp, with a G+C content of 37.8%. Collinearity analysis of the 276 sequences revealed no gene rearrangements. Sliding window analysis of the complete plastid genomes identified highly variable regions concentrated in the *rbcL-accD*, *ccsA-ndhD*, *ycf1*, and *trnK(UUU)-rps16* segments. Alignment of the 276 complete genome sequences identified 846 valid variable sites, of which 309 were located in protein-coding sequences. In the phylogenetic tree based on protein-coding gene sequences, Chinese *Quercus* subsect. *Campylolepides* species generally clustered together in a single clade, except for one *Q. acutissima* sample from the LSK population that grouped with *Quercus* section *Ilex* species *Q. pseudosetulosa*. *Quercus* subsect. *Campylolepides* species first formed a clade with *Q. dolicholepis* and *Q. engleriana* from Hubei Province, then constituted a larger clade with *Q. guyavifolia* and *Q. pannosa* from southwestern Sichuan, as well as *Q. spinosa* and *Q. dolicholepis* from the northeastern margin of Sichuan. This larger clade was sister to a branch composed of European *Quercus* section *Cerris* species, including *Q. cerris*, *Q. suber*, and *Q. castaneifolia* from Europe. Based on the internal clustering patterns within the main *Quercus* subsect. *Campylolepides* clade, it could be broadly divided into four subclades: Clade 1 encompassed 43 individuals from 18 populations, Clade 2 encompassed 19 individuals from 10 populations, Clade 3 comprised 10 individuals from 7 populations, and Clade 4 consisted of 2 *Q. acutissima* individuals from 2 populations.

Figure 3 [Figure 3: see original paper] ML tree based on 124 plastid genomic protein-coding sequences

2.3 Population Divergence Time

In the phylogenetic tree inferred by BEAST (Figure 4 [Figure 4: see original paper]), the phylogenetic relationships between *Quercus* section *Ilex* and *Quercus* section *Cerris* species were inconsistent with the ML tree. In the BEAST phylogeny, European *Quercus* section *Cerris* species clustered with *Quercus* section *Ilex* species from northeastern Sichuan. However, the internal topology of the *Quercus* subsect. *Campylolepides* clade was largely consistent with the ML tree. Among the four main clades of *Quercus* subsect. *Campylolepides*, the most recent common ancestor (MRCA) times were: Clade 1 (8.37 Ma), Clade 2 (8.98 Ma), and Clade 3 (9.76 Ma). The MRCA time for the clade comprising Clades 1, 2, and 3 with Clade 4 was 12.70 Ma. The MRCA time for *Quercus* subsect. *Campylolepides* species across their entire distribution with *Quercus* section *Ilex* species from Hubei and southwestern Sichuan was 18.41 Ma. The MRCA time for European *Quercus* section *Cerris* species with *Quercus* section *Ilex* species from the northeastern margin of Sichuan was 18.61 Ma. The MRCA time for *Q. acutissima* from Lvshunkou (haplotype H43) and *Q. pseudosetulosa* from Guangdong was 14.42 Ma. The MRCA time for the East Asian *Quercus* section *Cerris*-*Quercus* section *Ilex* clade with European *Quercus* section *Cerris* species was 19.98 Ma. Within *Quercus* subsect. *Campylolepides*, the earliest diverging lineage was haplotype H43 from the LSK population in the Liaodong Peninsula, followed by haplotype H42 (Clade 4) shared between LSK and ZHX populations. Clades 1, 2, and 3 subsequently diverged within a relatively short time frame.

F1. Calibrating the divergence time between outgroup *Fagus* and the ingroup; F2. Calibrating the divergence time between *Castanea* and *Castanopsis*.

Figure 4 [Figure 4: see original paper] Divergence time estimation based on plastid genomic protein-coding sequences

Discussion

3.1 Haplotype Sharing Among Three Species of *Quercus* subsect. *Campylolepides*

Analysis of plastid haplotypes from 276 individuals of three Chinese *Quercus* subsect. *Campylolepides* species revealed haplotype sharing among the three species. In terms of quantitative characteristics, 15 shared haplotypes were identified with high frequency, totaling 172 individuals and accounting for 62.32% of all individuals. In contrast, 28 private haplotypes were detected with low frequency, representing 104 individuals and comprising only 37.68% of all individuals.

Li et al. (2022) previously documented extensive haplotype sharing among the three *Quercus* subsect. *Campylolepides* species and hypothesized that these species once shared an ancestral plastid gene pool that may have differentiated into distinct lineages before modern speciation events. The widespread haplotype sharing within *Quercus* subsect. *Campylolepides* was primarily attributed to past and ongoing introgression/hybridization and the retention of ancestral polymorphism. Due to the use of complete plastid genome data, this study achieved higher resolution in haplotype analysis compared to previous research. We found that most interspecific shared haplotypes were shared at the same or adjacent locations, suggesting that interspecific plastid sharing is mainly caused by recent sympatric hybridization leading to plastid capture. This is because lineage sorting is a random process and unlikely to produce such significant sympatric interspecific plastid sharing patterns.

Notably, multiple different types of haplotypes appearing on different evolutionary clades were detected in Northeast China, indicating that this region may represent a long-term stable glacial refuge for *Quercus* subsect. *Campylolepides*. Numerous phylogeographic studies on *Populus davidiana* (Hou et al., 2018), *Quercus mongolica* (Zeng et al., 2015), and *Schisandra chinensis* (Ye et al., 2019) have also demonstrated that Northeast China served as a long-term stable refuge for forest plants during the Last Glacial Maximum. The Liaodong Peninsula and Bohai region may have allowed *Quercus* subsect. *Campylolepides* species to persist stably in Northeast China during range contraction in glacial periods and subsequent postglacial expansion. In contrast, Central and East China represent the regions where haplotype sharing occurs most frequently, with Li et al. (2021) reporting that 70% of hybrid individuals of *Quercus* subsect. *Campylolepides* originate from these areas.

3.2 Plastid Capture Between *Quercus* subsect. *Campylolepides* and East Asian *Quercus* section *Ilex*

Phylogenies constructed from nuclear genomic data indicate that *Quercus* section *Cerris* as a whole forms a monophyletic group (Hipp et al., 2020), with the East Asian and European branches diverging in the early Oligocene (~30 Ma). In contrast, phylogenetic trees and divergence times based on plastid genomes show that *Quercus* subsect. *Campylolepides* first formed an East Asian *Quercus* section *Cerris*-*Quercus* section *Ilex* (Cerris-Ilex) clade with East Asian *Quercus* section *Ilex* species. Within the East Asian Cerris-Ilex clade, *Quercus* subsect. *Campylolepides* did not form a sister branch with the entire East Asian *Quercus* section *Ilex* but only with *Q. engleriana* and *Q. dolicholepis* from Hubei. Other early-diverging *Quercus* section *Ilex* branches were composed of species from eastern Sichuan and Liangshan regions. The East Asian Cerris-Ilex clade and the European *Quercus* section *Cerris* clade began to diverge in the early Miocene (19.98 Ma), which is relatively close to the time inferred from plastid genomes in Zhou et al. (2022). Compared to the divergence time between East Asian and European branches of *Quercus* section *Cerris* inferred from nuclear genomes, the plastid genome estimate is nearly 10 Ma later, indicating that contact and gene exchange with *Quercus* section *Ilex* species occurred later than the geographic divergence between East Asian and European branches of *Quercus* section *Cerris*. The disjunct distribution between East Asian and European branches of *Quercus* section *Cerris* differs from the classic East Asia-Mediterranean disjunction patterns observed in groups such as *Helleborus* (Sun et al., 2001), *Meconopsis* (Zhuang, 1981), and *Paliurus* (Chen et al., 2017), which resulted from Himalayan uplift and Tethys retreat (Sun, 2002). Instead, the disjunct distribution of *Quercus* section *Cerris* was likely caused by the disappearance of temperate deciduous forest vegetation in northern Asia (Denk et al., 2023). We therefore hypothesize that ancestral *Quercus* subsect. *Campylolepides* species likely experienced a bottleneck effect due to the disappearance of northern temperate deciduous forest vegetation, forcing their distribution southward with relatively small population sizes. This brought them into contact with widely distributed subtropical *Quercus* section *Ilex* species, enabling them to capture the plastids of *Quercus* section *Ilex*.

Quercus section *Cerris* and *Quercus* section *Ilex* are closely related and have experienced multiple plastid capture events, with "traces" of hybridization with *Quercus* section *Ilex* species found in both East Asian and European branches of *Quercus* section *Cerris* (Simeone et al., 2016). Within the East Asian branch, this study identified a haplotype (H43) in the LSK population that clustered with *Quercus* section *Ilex* species *Q. pseudosetulosa*, suggesting that more than one plastid capture event may have occurred between *Quercus* section *Cerris* and *Quercus* section *Ilex* species. However, considering that *Q. pseudosetulosa* is a relict species from the Wanshan Islands with no direct geographic contact with *Quercus* subsect. *Campylolepides* species in Northeast China, this plastid capture likely originated from other *Quercus* section *Ilex* species sharing the same plastid type as *Q. pseudosetulosa*, such as *Q. phillyraeoides*. Although *Quercus* section *Ilex* is currently absent from Northeast China, the historical geographic distribution of *Quercus* section *Ilex* in China may have extended further north during geological history.

The extant plastids of *Quercus* subsect. *Campylolepides* primarily represent the outcome of one capture event with *Quercus* section *Ilex*. This capture event occurred between the early Oligocene and middle Miocene—after the divergence time between East Asian and European branches of *Quercus* section *Cerris* inferred from nuclear genes but before the divergence time of plastid haplotypes between *Quercus* subsect. *Campylolepides* and *Quercus* section *Ilex* inferred from plastid genes. The likely geographic location of plastid capture between modern *Quercus* subsect. *Campylolepides* and *Quercus* section *Ilex* was probably in the region from Northeast to Central China. After hybridization and plastid capture with *Quercus* section *Ilex* species, *Quercus* subsect. *Campylolepides* gradually formed reproductive isolation from *Quercus* section *Ilex* and dispersed throughout East Asia. Recent interspecific introgression in *Quercus* subsect. *Campylolepides* has primarily occurred among the three species within the subsection. Subsequently, *Quercus* subsect. *Campylolepides* species diverged into four geographically unstructured clades during the late Miocene. During this period, continuous global cooling (Guo et al., 2008) and intensification of the East Asian monsoon climate (Sun et al., 2005) caused frequent vegetation shifts between northern and southern regions, enabling multiple contacts among different *Quercus* subsect. *Campylolepides* lineages and gradually eroding geographic patterns in their genetic lineages. Quaternary glaciation further fragmented the distribution of this group, with independent refugia existing in Northeast China. Previous studies have also indicated that mountainous areas in Central and East China harbored common refugia for the three *Quercus* subsect. *Campylolepides* species (Li et al., 2021). Local expansion during interglacial periods and migration events among different refugia intensified haplotype sharing among populations.

Our study demonstrates that plastid genome sharing occurs among the three *Quercus* subsect. *Campylolepides* species. *Quercus* section *Cerris* and *Quercus* section *Ilex* have experienced multiple plastid capture events, but the modern plastids of *Quercus* subsect. *Campylolepides* primarily derive from a single capture event. Contact between *Quercus* subsect. *Campylolepides* species and *Quercus* section *Ilex* species, along with subsequent plastid capture, occurred before the formation of modern lineages of *Quercus* subsect. *Campylolepides*, with timing constrained between the early Oligocene and middle Miocene. No further plastid replacement occurred between the two groups after the middle Miocene, suggesting that strong reproductive isolation has developed between them.

References

LI Y, 2019. Genetic structure and evolutionary history of Chinese oak species in Quercus section Cerris[D]. Nanjing: Nanjing Forestry University: 183.

LI Y, ZHANG XW, FANG YM, 2014. Predicting the impact of global warming on the geographical distribution pattern of Quercus variabilis in China[J]. Chinese Journal of Applied Ecology, 25(12): 3381-3389.

LI Y, ZHANG XW, FANG YM, 2016. Responses of the distribution pattern of Quercus chenii to climate change following the Last Glacial Maximum[J]. Chinese Journal of Plant Ecology, 40(11): 1164-1178.

SUN H, 2002. Tethys retreat and Himalayas-Hengduanshan Mountains uplift and their significance on the origin and development of the Sino-Himalayan elements and alpine flora[J]. Acta Botanica Yunnanica, (3): 273-288.

ZHOU ZK, 1992. Origin, phylogeny and dispersal of Quercus from China[J]. Acta Botanica Yunnanica, (03): 227-236.

ZHUANG X, 1981. The systematic evolution and the geographical distribution of Meconopsis Vig[J]. Acta Botanica Yunnanica, (2): 139-146.

BOUCKAERT RR, HELED J, KÜHNERT D, et al., 2014. BEAST 2: A software platform for bayesian evolutionary analysis[J]. PLoS Comput Biol, 10(4): e1003537.

CAMUS A, 1936. Les chênes: Monographie du genre Quercus et monographie du genre Lithocarpus[M]. Paris: Encyclopédie économique de Sylviculture.

CAPELLA-GUTIÉRREZ S, SILLA-MARTíNEZ JM, GABALDÓN T, 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses[J]. Bioinformatics, 25(15): 1972-1973.

CHEN YS, MESEGUER AS, GODEFROID M, et al., 2017. Out-of-India dispersal of *Paliurus* (Rhamnaceae) indicated by combined molecular phylogenetic and fossil evidence[J]. Taxon, 66(1): 78-90.

DARRIBA D, TABOADA GL, DOALLO R, et al., 2012. jModelTest 2: more models, new heuristics and parallel computing[J]. Nat Methods, 9(8): 772.

DENK T, GRIMM GW, HIPP AL, et al., 2023. Niche evolution in a northern temperate tree lineage: biogeographical legacies in cork oaks (*Quercus* section *Cerris*)[J]. Ann Bot, 131(5): 615-633.

FU RR, ZHU Y, LIU Y, et al., 2022. Genome-wide analyses of introgression between two sympatric Asian oak species[J]. Nat Ecol Evol, 6(7): 924-935.

GRÍMSSON F, GRIMM GW, ZETTER R, et al., 2016. Cretaceous and paleogene Fagaceae from North America and Greenland: evidence for a Late Cretaceous split between *Fagus* and the remaining Fagaceae[J]. Acta Palaeobot, 56(2): 247-305.

GUO ZT, SUN BN, ZHANG ZS, et al., 2008. A major reorganization of Asian climate by the early Miocene[J]. Clim Past, 4(3): 153-174.

HIPP AL, MANOS PS, HAHN M, et al., 2020. Genomic landscape of the global oak phylogeny[J]. New Phytol, 226(4): 1198-1212.

HOU Z, WANG ZS, YE ZY, et al., 2018. Phylogeographic analyses of a widely distributed *Populus davidiana*: Further evidence for the existence of glacial refugia of cool-temperate deciduous trees in northern East Asia[J]. Ecol Evol, 8(24): 13014-13026.

JIN JJ, YU WB, YANG JB, et al., 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes[J]. Genome Biol, 21(1): 241.

KATOH K, STANDLEY DM, 2013. MAFFT multiple sequence alignment software Version 7: Improvements in performance and usability[J]. Mol Biol Evol, 30(4): 772-780.

KLEINKOPF JA, ROBERTS WR, WAGNER WL, et al., 2019. Diversification of Hawaiian *Cyrtandra* (Gesneriaceae) under hybridization and influence of incomplete lineage sorting[J]. J Syst Evol, 57(6): 561-578.

LETUNIĆ I, BORK P, 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation[J]. Nucl Acid Res, 49(W1): W293-W296.

LI Y, WANG L, ZHANG XW, et al., 2022. Extensive sharing of chloroplast haplotypes among East Asian Cerris oaks: The imprints of shared ancestral polymorphism and introgression[J]. Ecol Evol, 12(8): e9142.

LI Y, ZHANG XW, WANG L, et al., 2021. Influence of pliocene and pleistocene climates on hybridization patterns between two closely related oak species in China[J]. Ann Bot, 129(2): 183-194.

LIANG YP, SHI Y, YUAN S, et al., 2022. Linked selection shapes the landscape of genomic variation in three oak species[J]. New Phytol, 233(1): 555-568.

LIU L, DU YX, FOLK RA, et al., 2020. Plastome Evolution in Saxifragaceae and multiple plastid capture events involving *Heuchera* and *Tiarella*[J]. Front Plant Sci, 11: 361.

LIU X, WANG Z, SHAO W, et al., 2017. Phylogenetic and taxonomic status analyses of the abaso section from multiple nuclear genes and plastid fragments reveal new insights into the North America origin of *Populus* (Salicaceae)[J]. Front Plant Sci, 7: 2022.

MELESHKO O, MARTIN MD, KORNELIUSSEN TS, et al., 2021. Extensive genome-wide phylogenetic discordance is due to incomplete lineage sorting and not ongoing introgression in a rapidly radiated bryophyte genus[J]. Mol Biol Evol, 38(7): 2750-2766.

PERCY DM, ARGUS GW, CRONK QCB, et al., 2014. Understanding the spectacular failure of DNA barcoding in willows (*Salix*): Does this result from a trans-specific selective sweep?[J]. Mol Ecol, 23(19): 4737-4756.

QU XJ, MOORE MJ, LI DZ, et al., 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes[J]. Plant Methods, 15: 50.

RAMBAUT A, DRUMMOND AJ, XIE D, et al., 2018. Posterior summarization in Bayesian phylogenetics using tracer 1.7[J]. Syst Biol, 67(5): 901-904.

RIESEBERG LH, SOLTIS DE, 1991. Phylogenetic consequences of cytoplasmic gene flow in plants[J]. Am J Bot, 5(1): 65-84.

ROZAS J, FERRER-MATA A, SÁNCHEZ-DELBARRIO JC, et al., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets[J]. Mol Biol Evol, 34(12): 3299-3302.

SIMEONE MC, CARDONI S, PIREDDA R, et al., 2018. Comparative systematics and phylogeography of *Quercus* Section *Cerris* in western Eurasia: inferences from plastid and nuclear DNA variation[J]. PeerJ, 6: e5793.

SIMEONE MC, GRIMM GW, PAPINI A, et al., 2016. Plastome data reveal multiple geographic origins of *Quercus* Group Ilex[J]. PeerJ, 4: e1897.

STAMATAKIS A, 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies[J]. Bioinformatics, 30(9): 1312-1313.

STULL GW, PHAM KK, SOLTIS PS, et al., 2023. Deep reticulation: the long legacy of hybridization in vascular plant evolution[J]. Plant J, 114(4): 743-766.

SUN H, MCLEWIN W, FAY MF, 2001. Molecular phylogeny of *Helleborus* (Ranunculaceae), with an emphasis on the East Asian-Mediterranean disjunction[J]. Taxon, 50(4): 1001-1018.

SUN XJ, WANG PX, 2005. How old is the Asian monsoon system?—Palaeobotanical records from China[J]. Palaeogeogr Palaeoclimatol Palaeoecol, 222(3-4): 181-222.

TEKPINAR AD, AKTAŞ C, KANSU Ç, et al., 2021. Phylogeography and phylogeny of genus *Quercus* L. (Fagaceae) in Turkey implied by variations of *trnT(UGU)-L(UAA)-F (GAA)* chloroplast DNA region[J]. Tree Genet Genom, 17: 1-18.

TSITRONE A, KIRKPATRICK M, LEVIN DA, 2003. A model for chloroplast capture[J]. Evolution, 57(8): 1776-1782.

WILF P, NIXON KC, GANDOLFO MA, et al., 2019. Eocene Fagaceae from Patagonia and Gondwanan legacy in Asian rainforests[J]. Science, 364(6444): eaaw5139.

YANG J, DI XY, MENG X, et al., 2016. Phylogeography and evolution of two closely related oak species (*Quercus*) from north and northeast China[J]. Tree Genet Genom, 12(89): 1-14.

YANG YY, QU XJ, ZHANG R, et al., 2021. Plastid phylogenomic analyses of Fagales reveal signatures of conflict and ancient chloroplast capture[J]. Mol Phylogenet Evol, 163: 107232.

YE JW, ZHANG ZK, WANG HF, et al., 2019. Phylogeography of *Schisandra chinensis* (Magnoliaceae) reveal multiple refugia with ample gene flow in Northeast China[J]. Front Plant Sci, 10: 199.

ZENG YF, WANG WT, LIAO WJ, et al., 2015. Multiple glacial refugia for cool-temperate deciduous trees in northern East Asia: the Mongolian oak as a case study[J]. Mol Ecol, 24(22): 5678-5689.

ZHOU BF, YUAN S, CROWL AA, et al., 2022. Phylogenomic analyses highlight innovation and introgression in the continental radiations of Fagaceae across the Northern Hemisphere[J]. Nat Commun, 13(1): 1320.

Note: Figure translations are in progress. See original paper for figures.