



OPEN ACCESS

EDITED BY

Gang Yao,
South China Agricultural University, China

REVIEWED BY

Qiang Zhang,
Guangxi Zhuang Autonomous Region and
Chinese Academy of Sciences, China
Cen Guo,
Chinese Academy of Sciences (CAS), China

*CORRESPONDENCE

Chunce Guo

✉ chunceguo@jxau.edu.cn

RECEIVED 25 July 2024

ACCEPTED 14 October 2024

PUBLISHED 11 November 2024

CITATION

Wang C, Li Y, Yang G, Zhang W and Guo C
(2024) Comparative analysis of chloroplast
genomes and phylogenetic relationships
in the endemic Chinese bamboo
Gelidocalamus (Bambusoideae).
Front. Plant Sci. 15:1470311.
doi: 10.3389/fpls.2024.1470311

COPYRIGHT

© 2024 Wang, Li, Yang, Zhang and Guo. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Comparative analysis of chloroplast genomes and phylogenetic relationships in the endemic Chinese bamboo *Gelidocalamus* (Bambusoideae)

Chengkun Wang, Yonglong Li, Guangyao Yang,
Wengen Zhang and Chunce Guo*

Jiangxi Provincial Key Laboratory of Improved Variety Breeding and Efficient Utilization of Native Tree
Species, Forestry College, Jiangxi Agricultural University, Nanchang, China

Introduction: *Gelidocalamus* Wen is a small yet taxonomically challenging genus within the Arundinarieae tribe. Recent molecular studies have suggested it may not be monophyletic. However, limited species sampling and insufficient molecular marker information have resulted in poorly resolved phylogenetic relationships within this genus.

Methods: The complete chloroplast genomes covering all 16 species and one variant of *Gelidocalamus* were sequenced, and comparative analyses were conducted. Phylogenetic analyses were performed using different molecular markers, including chloroplast data, the nuclear ribosomal DNA (nrDNA) repeats region, and 29 mitochondrial protein-coding genes. Additionally, the divergence times of *Gelidocalamus* were estimated to reveal their evolutionary history.

Results: The plastomes of *Gelidocalamus* ranged in size from 139,500 bp to 139,801 bp, with a total of 137 identified genes, including 90 protein-coding genes, 39 tRNA genes, and 8 rRNA genes. The size of the nrDNA repeats ranged from 5,802 bp to 5,804 bp. Phylogenetic analysis based on chloroplast data revealed that *Gelidocalamus* is polyphyletic, with different subclades distributed within the IV and V clades. However, phylogenetic analysis based on nrDNA and mitochondrial genes did not effectively resolve the relationships within the genus.

Discussion: Comparative analysis of chloroplast genomes indicated that *Gelidocalamus* shares a high degree of similarity with closely related genera in terms of chloroplast genome collinearity, codon usage bias, and repetitive sequences. Divergence time estimation suggests that it is a relatively young group, with all members appearing successively over the past four million years. The complex phylogenetic patterns may arise from the rapid radiation of Arundinarieae. This study provides a preliminary foundation for further in-depth research on the phylogeny, genomic structural features, and divergence times of this genus.

KEYWORDS

Gelidocalamus Wen, phylogenetic relationships, plastome evolution, molecular markers, divergence times

1 Introduction

Gelidocalamus Wen, a small genus belonging to the Arundinarieae tribe, is endemic bamboo to the humid evergreen broad-leaved forests of eastern China. Established in 1982, the genus comprises approximately 16 species and one variety (Wen, 1982; Li et al., 2006; Zhang et al., 2017; Cai et al., 2021; Wang et al., 2023; Li et al., 2023). *Gelidocalamus* is characterized by several distinct morphological features, such as multiple branches per node, cylindrical and grooveless internodes, persistent sheaths, a single foliage leaf on each ultimate branch (except *G. multifolius* B. M. Yang) (Yang, 1986), and semelauctant inflorescence. One notable characteristic is the emergence of new shoots during the autumn and winter (Liu et al., 2017; Nie et al., 2018). This genus primarily inhabits hilly and low-mountain areas south of the Yangtze River, at elevations below 1,000 m (Li et al., 2016; Nie et al., 2018). *Gelidocalamus* typically grows alongside mountain streams, gullies, and waterways under the canopy of evergreen broad-leaved forest, forming distinctive strip-like patterns. As a stable component of middle-subtropical forests, the *Gelidocalamus* plays a crucial role in water conservation and ecological balance. Additionally, it is valued for its elegant foliage and edible young shoots, offering significant potential for ornamental and culinary uses. However, as a complex genus, some bamboo species are scarce, and our current understanding is primarily based on morphological studies. Little is known about their phylogenetic relationships, origin, and evolution, which restricts further utilization and conservation efforts (Wen, 1982; Liao, 1991; Zhang et al., 2017).

Traditionally, bamboo species classification has relied on vegetative organ characteristics due to the unpredictability of bamboo flowering phenology. However, bamboo species exhibit a high degree of homogeneity in vegetative traits, posing significant challenges to taxonomy and systematics (Peng et al., 2008; Zeng et al., 2010). Molecular phylogenetic studies have sought to address these challenges using various markers such as the internal transcribed spacer (ITS) and single-copy nuclear genes like *Granule-Bound Starch Synthase I* (*GBSSI*) gene. Despite these efforts, the resolution of phylogenetic relationships based on ITS and single-copy genes has been limited (Peng et al., 2008; Zhang et al., 2012; Tong et al., 2020). In recent years, the rapid advancement of Restriction-site Associated DNA sequencing (RAD) sequencing technology has significantly enhanced our understanding of the phylogeny and evolutionary diversification of bamboo species. Ye et al. (2019) employed ddRAD-seq data to reconstruct the phylogeny of bamboo species, focusing on 178 individuals representing alpine bamboo. Similarly, Guo et al. (2021) utilized ddRAD-seq data to establish a phylogenetic framework for temperate bamboo involving 213 individuals from 200 species across 32 genera. These studies provided high-resolution phylogenies for temperate bamboo species and offered insights into their diversification patterns. Additionally, the investigation into the evolution of key traits sheds light on the complex evolutionary history of temperate bamboo (Ye et al., 2019; Guo et al., 2021). In addition, the use of chloroplast gene

fragments or complete genomes has provided superior resolution (Triplett and Clark, 2010; Zeng et al., 2010; Yang et al., 2013; Attigala et al., 2014; Zhou et al., 2022). Since the sequencing of the complete chloroplast genomes of *Nicotiana tabacum* and *Marchantia polymorpha* in 1986 (Ohyama et al., 1986; Sugiura et al., 1986), a wealth of chloroplast genome data have become available. Chloroplast genomes are favored for phylogenetic studies due to their highly conserved gene structures, moderate evolutionary rates, uniparental inheritance, minimal influence from paralogs, and ease of genome extraction (Zhang and Li, 2011; Wang et al., 2012; Li et al., 2020). Using plastid loci or chloroplast genome data, the phylogenetic framework of the Arundinarieae tribe has been progressively constructed and refined (Triplett and Clark, 2010; Zeng et al., 2010; Yang et al., 2013; Attigala et al., 2014; Ma et al., 2014; Zhang et al., 2018; Zhang et al., 2020). This tribe is robustly supported as a monophyletic group, with a sister relationship between the Bambuseae and Olyreae tribes. In contrast, the phylogenetic relationships constructed from chloroplast genes differ from those based on nuclear genes. In the nuclear gene tree, Olyreae diverges first, with Arundinarieae and Bambuseae forming sister groups. In the chloroplast tree, however, Arundinarieae diverges first, and Bambuseae clusters with Olyreae (Guo et al., 2021; Gallaher et al., 2022; Huang et al., 2022).

Although a clear framework has been defined in the Arundinarieae tribe. However, studies on the chloroplasts of *Gelidocalamus* have been limited, and only a few species have been sporadically sampled. The relationships within *Gelidocalamus* and its associations with closely related genera such as *Indocalamus* Nakai, *Ferrocalamus* Hsueh & Keng f., *Shibataea* Makino ex Nakai, *Sinosasa* L. C. Chia ex N. H. Xia, Q. M. Qin & Y. H. Tong, and others within clades IV and V remain unclear (Peng et al., 2008; Triplett and Clark, 2010; Zeng et al., 2010; Zhang et al., 2019; Qin et al., 2020). According to the chloroplast genome-based study, *G. stellatus*, *G. latifolius*, and *G. solidus* are situated within clade V. *G. stellatus* and *G. latifolius* form into a clade with *Indocalamus*, while *G. solidus* clustered separately with *Chimonobambusa ningnanica*. Additionally, *G. tessellatus* forms a clade with the genus *Shibataea* (Guo et al., 2021). Interestingly, the phylogenetic relationships within the Arundinarieae tribe, as inferred from ddRAD data, included a total of 13 *Gelidocalamus* species. Six of these species (*G. annulatus*, *G. tessellatus*, *G. latifolius*, *G. stellatus*, *G. multifolius*, and *G. monophyllus*) clustered into a monophyletic branch. Three species (*G. solidus*, *G. subsolidus*, and *G. albopubescens*) formed a monophyletic lineage that was sister to *Indocalamus* and *Bashannia*, but their phylogenetic relationships conflicted dramatically with those based on chloroplast genome data. The remaining four (*G. dongdingensis*, *G. rutilans*, *G. kunishii*, *G. longiinternodus*) were nested within *Indocalamus*, which significantly conflicted with chloroplast genome results. Additionally, the clades IV and V, where *Gelidocalamus* is situated, were reconfigured to form a new *Indocalamus* + *Gelidocalamus* + *Chimonocalamus* branch (Guo et al., 2021). This article argues that such a significant nucleoplasmic conflict reflects the differing evolutionary histories of the nucleoplasm.

In this study, we collected 27 samples representing 16 species and one variant of *Gelidocalamus*, along with plastomes from closely related genera obtained from the NCBI database and the data matrix of Guo et al. (2021). We re-annotated and examined these plastomes, assembling and annotating newly sequenced plastomes to investigate the phylogenetic relationships within the *Gelidocalamus*. We also assembled nrDNA repeats and mitochondrial coding genes from low-depth data to perform phylogenetic inference and estimate divergence times for *Gelidocalamus*. Our objectives are to (1) explore the phylogenetic relationships among *Gelidocalamus* species while assessing the contributions of molecular markers, (2) understand the structural characteristics and sequence variations of the chloroplast genomes of *Gelidocalamus* and its closely related species, and (3) reveal the evolutionary history of *Gelidocalamus*.

2 Materials and methods

2.1 Sampling, DNA extraction, sequencing, assembly, and annotation

To conduct a comprehensive study on the relationships of *Gelidocalamus*, we performed field surveys at both the type and nontype localities of each species. Mature leaves were collected from individuals at these localities, and all voucher specimens were deposited in the herbarium of Jiangxi Agricultural University, China (JXAU).

Genomic DNA was isolated from dried foliage leaves over silica gel using a modified CTAB method (Murray and Thompson, 1980). Illumina paired-end libraries (2 × 150 bp) were constructed and sequenced by Novogene Bioinformatics Technology Co. Ltd. (Beijing, China), yielding approximately 8 GB of raw data per sample. To enhance assembly accuracy, FastQC 0.11.9 (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc>) and Fastp 0.12.4 (Chen et al., 2018) were used with default parameters to filter out unpaired and low-depth reads.

In the assembly process, the chloroplast genome was assembled using GetOrganelle 1.7.4 (Jin et al., 2020) with k-mers of 45, 65, 85, 105, and 121. Following this, the filtered reads underwent processing through Bandage (Wick et al., 2015) to facilitate the attachment of chloroplast genome scaffolds. The clean reads were mapped to the draft genome using Geneious 9.1.4 (Kearse et al., 2012) to check the concatenation of contigs. The resultant chloroplast genome sequences were annotated using CPGAVAS2 (Shi et al., 2019), followed by a meticulous manual review in Geneious. To visualize the sequenced chloroplast genomes, Chloroplot (Zheng et al., 2020) was employed. Similarly, the assembly of nrDNA repeats was executed using GetOrganelle with k-mers of 35, 85, and 115, supplemented by manual corrections within Geneious. For the mitochondrial genome assembly, Geneminer (Xie et al., 2023) was utilized under default parameters, referencing the complete mitochondrial genomes of *Bambusa oldhamii* (EU365401) and *Ferrocalamus strictus* (JN120789). The resulting output underwent further enhancement through subsequent corrections in Geneious.

2.2 Phylogenetic analyses

To elucidate the phylogenetic relationships among 16 species and one variant within *Gelidocalamus*, as well as their relationships with related genera, we selected and reannotated additional 44 plastomes from 24 genera (Supplementary Table S1) to represent the 12 major clades of the Arundinarieae tribe. Additionally, we chose four plastomes (*Dendrocalamus latiflorus*, *Bambusa multiplex*, *Olyra latifolia*, and *Raddia brasiliensis*) from the Bambuseae and Olyreae tribes as outgroups. All plastomes were sourced from the NCBI database (<https://www.ncbi.nlm.nih.gov/>) and the data matrix of Guo et al. (2021).

Phylogenetic relationships were inferred using both maximum likelihood (ML) and Bayesian inference (BI) methods. For the ML analysis, whole chloroplast genomes and one IR region sequence were aligned using MAFFT 7.450 (Katoh and Standley, 2013) and trimmed with trimAl (Capella-Gutiérrez et al., 2009), respectively. IQ-TREE (Nguyen et al., 2015) was employed with 5,000 bootstrap replications, and the best-fit BIC model GTR + F + I + G4 was selected using ModelFinder (Kalyaanamoorthy et al., 2017). For Bayesian Inference (BI), MrBayes 3.2.6 (Ronquist et al., 2012) was utilized with the same model. Markov Chain Monte Carlo (MCMC) simulations ran for 20,000,000 generations, ensuring an average standard deviation of split frequencies (ASDFs) < 0.01, with sampling every 1,000 generations. The initial 25% burn-in samples were discarded, and the optimized topology was obtained. Nuclear ribosomal DNA (nrDNA) sequences and mitochondrial coding genes were also used in the phylogenetic analysis, using the same methods as described above.

2.3 Estimation of divergence time

To elucidate the origin and evolutionary history of the *Gelidocalamus*, divergence times were estimated from chloroplast genomic data using the BEAST v2.6.3 (Bouckaert et al., 2019). Estimating divergence times within the Arundinarieae tribe or even the Bambusoideae is challenging due to the scarcity of credible fossil evidence and the ambiguity and fragmentary nature of several bamboo fossil records (Guo et al., 2021). As a result, we adopted a conservative approach by utilizing a Bambusoideae fossil, *Bambusoideae* cf. *Chusquea* (Strömberg, 2011), to calibrate the crown Bambusoideae, and supplemented this with three additional secondary calibration points to assist in estimating divergence times (Zhang et al., 2016). Calibration point information is as follows: (1) Bambusoideae cf. *Chusquea* (35–90 Mya); (2) Arundinarieae crown (6.88–20.96 Mya, median age: 12.72 Mya); (3) clade IV crown (2.166.58 Mya, median age: 4.01 Mya); and (4) Clade V crown (1.24–3.82 Mya, median age: 2.38 Mya). Each calibration point was implemented as a uniform distribution between the minimal and maximal age of the constraint. We used a Yule process tree prior, the uncorrelated lognormal relaxed clock model, and the HKY substitution model. The MCMC chain length was set at 250,000,000, with a sampling frequency of 1,000. To assess parameter convergence, we utilized Tracer v1.7.2, stipulating an effective sample size (ESS) exceeding 200. Subsequently, we

summarized the age statistics for internal nodes using TreeAnnotator v2.6.6.

2.4 Plastome structure comparison and sequence divergence analyses

To explore the structural variations in chloroplast genomes within *Gelidocalamus*, we statistically analyzed the total length of the chloroplast genome, the number of genes, and the GC content of each region using Geneious (Kearse et al., 2012). Subsequently, all sequences were aligned and compared using MAFFT (Katoh and Standley, 2013). The alignments were then analyzed for potential rearrangements and inversions through covariance analysis using the mauve plugin within Geneious. Furthermore, events of expansion and contraction within the inverted repeat (IR) regions among the chloroplast genomes were visualized using the IRscope online program (Amiryousefi et al., 2018). In order to further understand the sequence divergence and hypervariable regions detected, the nucleotide diversity (π) values of plastomes were calculated using DnaSP v.5.1 (Librado and Rozas, 2009) with a sliding window analysis.

2.5 Codon usage analyses

To determine the codon usage pattern of *Gelidocalamus*, we extracted protein-coding genes from chloroplast genomes using PhyloSuite (Zhang et al., 2019). Subsequently, we calculated the relative synonymous codon usage (RSCU) values for each protein-coding gene using CodonW (<https://codonw.sourceforge.net/>) and visualized the resulting data using TBtools (Chen et al., 2020). An RSCU value equal to 1.00 indicates no codon usage bias, an RSCU value less than 1.00 suggests less frequent codon usage, while an RSCU value greater than 1.00 indicates the presence of codon usage bias.

2.6 Repeats and SSR analyses

To detect repetitive sequences in chloroplast genome sequences, an online tool MiSa-web (Beier et al., 2017) was employed for the analysis of each genome. The parameter settings included a minimum repeat count of 10 for mononucleotide repeats, a minimum repeat count of five for dinucleotide repeats, a minimum repeat count of four for trinucleotide repeats, and a minimum repeat count of three for tetranucleotide to hexanucleotide repeats. Simultaneously, the identification of scattered repeat sequences within the chloroplast genomes was conducted using REPuter (Kurtz et al., 2001). This analysis covered forward, reverse, complementary, and palindromic repeat sequences, with parameter settings specifying a Hamming Distance of 3, a Minimal Repeat Size of 30, and a Maximum Computed Repeats of 80.

3 Results

3.1 Sampling, sequencing, and data assembly

In this study, fresh leaves of *Gelidocalamus* were collected for molecular analyses. A total of 27 populations were recorded, comprising 15 type localities representing 14 species and one variety (Supplementary Table S1). The morphological diversity of *Gelidocalamus* was documented and photographed (Supplementary Figure S1), and specimens were collected during the field survey. However, surveys for two species, *G. kunishii* and *G. fengkaiensis* (recently published), were not conducted. Illumina sequencing yielded a range of 39,708,004 to 207,484,842 paired-end clean reads across 27 *Gelidocalamus* samples (Supplementary Table S2). The complete chloroplast genome, nrDNA repeat segments (including the 18S, ITS1, 5.8S, ITS2, and 26S regions), and 29 shared mitochondrial genes were obtained from each species (Supplementary Figure S2).

The chloroplast genome size ranges from 139,500 to 139,801 bp, with all samples consistently displaying a GC content of 38.9%. All chloroplast genomes exhibit a typical quadripartite structure (Figure 1), consisting of a long single-copy region (LSC), a short single-copy region (SSC), and two internal inverted repeat regions (IRA and IRB). The length of the LSC region ranges from 83,007 bp (*G. zixingensis*) to 83,364 bp (*G. subsolidus* GDHZ), while the SSC varies from 12,799 bp (*G. subsolidus* GXLB) to 12,882 bp (*G. kunsishii*). The lengths of the inverted repeat regions range from 21,796 bp (*G. dongdingensis*) to 21,842 bp (*G. multifolius* and *G. zixingensis*). Across all genomes, a total of 137 genes were identified, including 90 protein-coding genes, 39 tRNA genes, and eight rRNA genes (Supplementary Tables S3, S4).

For nrDNA repeats, the combined length of 27 tandem repeat sequences ranges from 5,802 to 5,804 bp. The 18S region is constant at 1,813 bp, the ITS1 region varies between 213 bp and 214 bp, the 5.8S region remains constant at 160 bp, the ITS2 region ranges from 233 bp to 234 bp, and the 28S region varies between 3,382 bp and 3,384 bp (Supplementary Figure S2).

A total of 29 mitochondrial protein-coding genes were identified, among which 22 were complete, and seven were incomplete. With the exception of the *rps1* (ribosomal protein S1) and *cytb* (cytochrome b) genes, the recovery rates of the remaining genes were low (Supplementary Figure S2).

3.2 Phylogenetic reconstruction

A total of four alignment matrices for phylogenetic analysis were generated from aligned sequences after trimming nonconserved regions: (1) 75 complete chloroplast genomes, totaling 139,357 bp; (2) 75 chloroplast genomes with the inverted repeat region removed, totaling 117,306 bp; (3) complete nrDNA repeat sequences, totaling 5,805 bp; (4) 24 shared mitochondrial

genes from 27 samples, totaling 19,727 bp (Supplementary Data Sheets 1–4).

Based on the matrices of the complete chloroplast genome and the chloroplast genome with the inverted repeat region removed, the reconstructed phylogenetic relationships using maximum likelihood and Bayesian inference methods exhibit largely consistent topologies (Figure 2; Supplementary Figure S3). The Arundinarieae tribe emerges as a strongly supported monophyletic branch (BSML = 100% and PPBI = 1.00), with its twelve major subclades also receiving robust support. Members of the *Gelidocalamus* are grouped into two distinct monophyletic clades (IV and V), clustering with closely related genera such as *Indocalamus*, *Shibataea*, and *Ferrocalamus*.

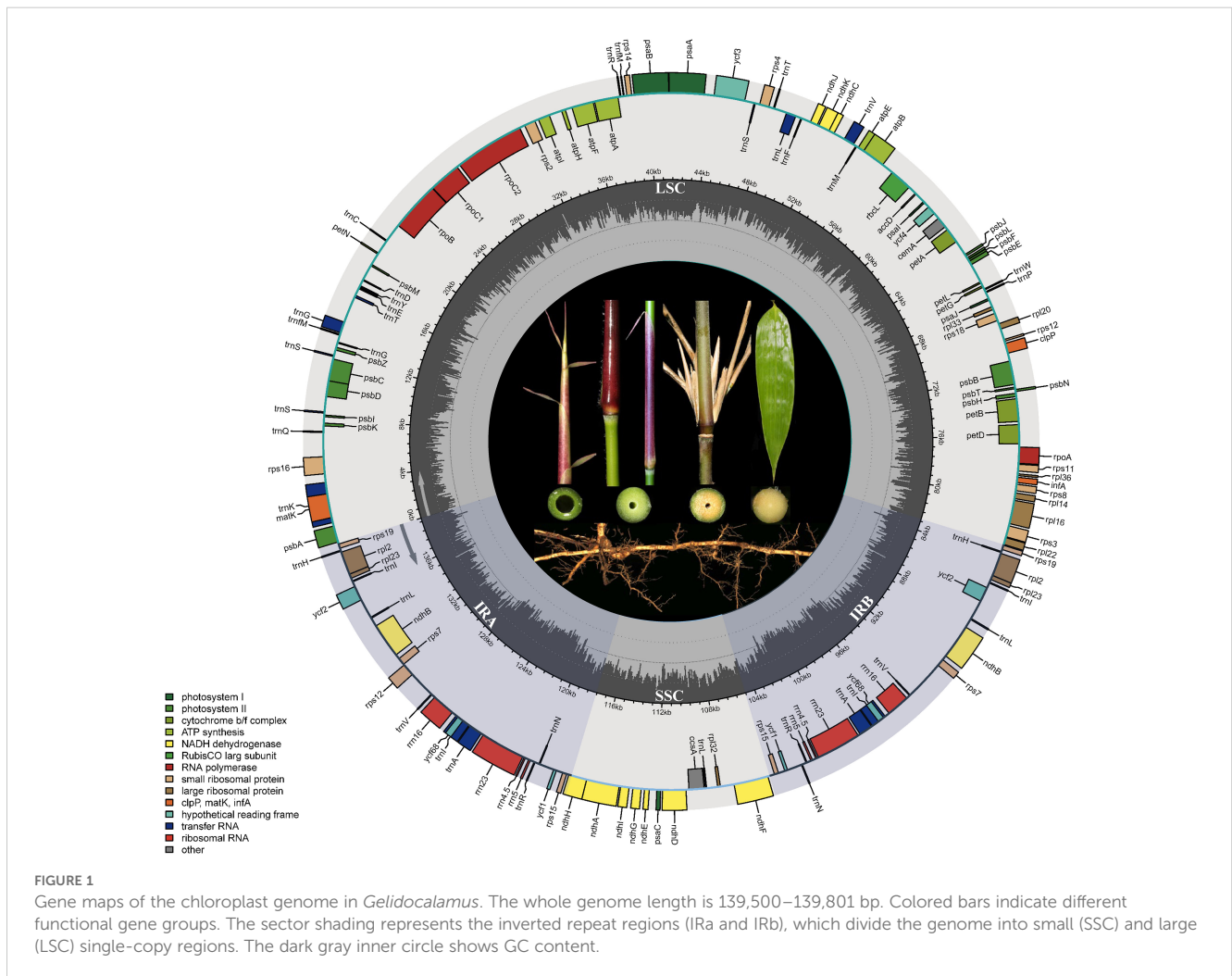
In clade V, there are a total of 10 *Gelidocalamus* samples, comprising eight species and one variety. Among them, *G. stellatus*, *G. stellatus* var. *wugongshanensis* and *G. latifolius* (BSML = 96% and PPBI = 1.00), as well as *G. albozonatus* and *G. annulatus* (BSML = 96% and PPBI = 1.00), each form two distinct monophyletic clades, respectively. The remaining species are distributed throughout the clade and cluster separately with the genera *Indocalamus* and *Phyllostachys*. In clade IV, *G. xunwuensis*, *G. dongdingensis*, and *G. subsolidus* form a monophyletic clade, and cluster

with *Indocalamus barbatus* as a sister group (BSML = 100% and PPBI = 0.66). Meanwhile, *G. multifolius*, *G. tessellatus*, *G. zixingensis*, and *G. fengkaiensis* form a monophyletic group with two *Shibataea* species as a sister group (BSML = 100% and PPBI = 1.00). Additionally, *G. monophyllus* and *Sinosasa longiligulata* cluster together in a clade (BSML = 100% and PPBI = 1.00).

Compared to chloroplast genome data, the phylogenetic relationships based on nrDNA repeats sequences and mitochondrial genes showed weak support and exhibited significant topology differences between the two datasets, both displaying multiple clades (Supplementary Figures S4, S5).

3.3 Divergence time estimation

Estimates of divergence times indicate that the stem age of the Arundinarieae tribe diverged in the Late Oligocene (23.42 Mya, 95% HPD: 16.71–32.21 Mya) and the crown age diverged around 11.20 Mya (95% HPD: 8.22–14.59 Mya) in the Mid–Late Miocene (Figure 3). The crown age for clades IV and V, where *Gelidocalamus* is situated, was 5.74 Mya (95% HPD: 4.50–6.58 Mya) and 2.60 Mya (95% HPD: 1.80–3.56 Mya), respectively. The



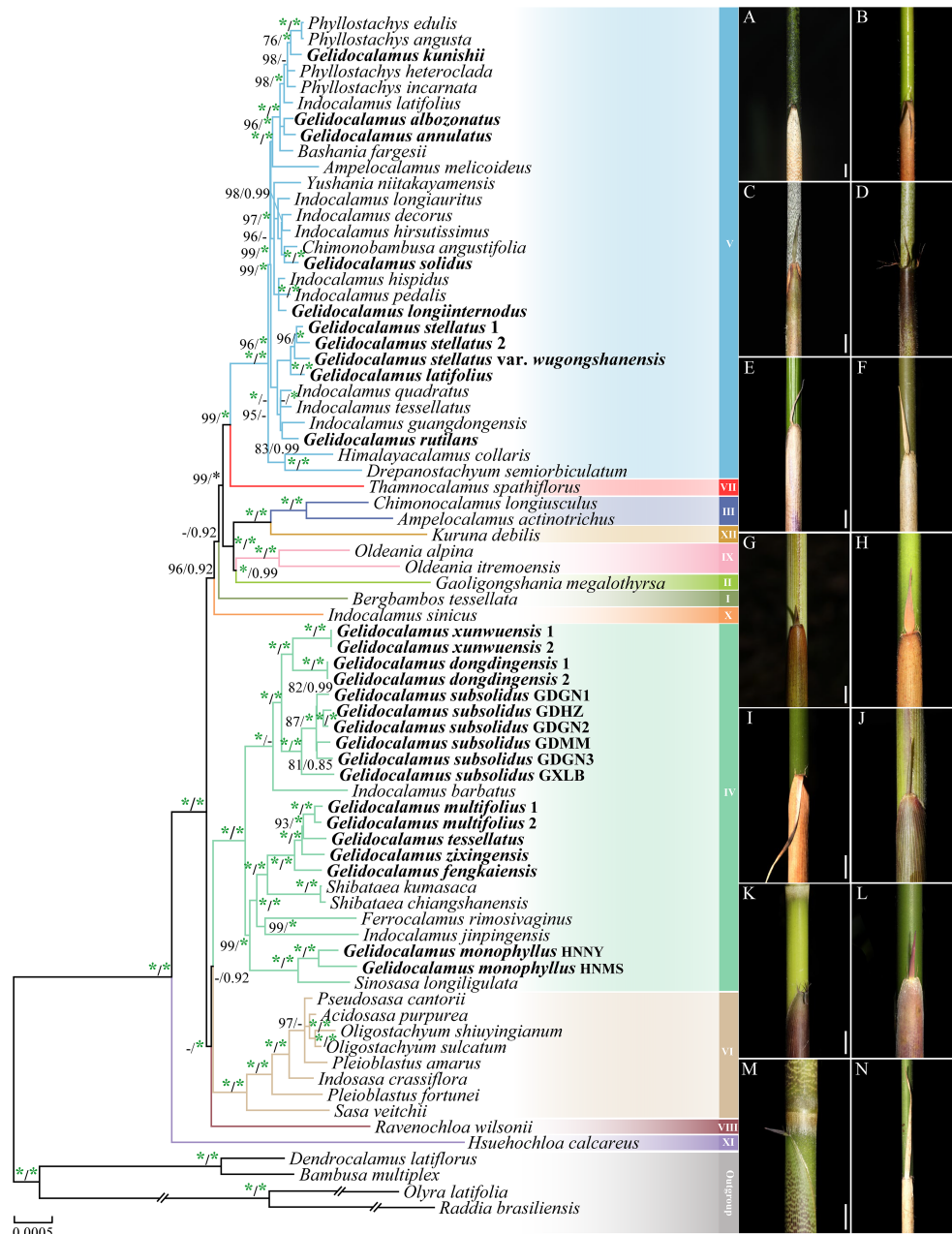


FIGURE 2
Phylogenetic tree of 75 taxa using maximum likelihood (ML) and Bayesian inference (BI) methods based on plastomes. ML tree topology is shown with ML bootstrap values, and BI posterior probabilities are indicated on the nodes. Only bootstrap values (BS) \geq 75% and posterior probabilities (PP) \geq 0.75 are indicated at each node; otherwise, dashes are used. The green asterisk indicates support of 100% BS or 1.00 PP. (A–N) represent the characteristics of the culms and culm sheaths of 14 species of *Gelidocalamus*. (A) *G. albozonatus*, (B) *G. annulatus*, (C) *G. solidus*, (D) *G. longinternodus*, (E) *G. stellatus*, (F) *G. latifolius*, (G) *G. rutilans*, (H) *G. xunwuensis*, (I) *G. dongdingensis*, (J) *G. subsolidus*, (K) *G. multifolius*, (L) *G. zixingensis*, (M) *G. tessellatus*, and (N) *G. monophyllus*.

earliest divergence of *Gelidocalamus* species, including *G. subsolidus* and *G. xunwuensis*, from their common ancestor with *Indocalamus barbatus* occurred at 3.96 Mya (95% HPD: 2.71–5.01 Mya). In contrast, *G. kunsishii* was the latest to diverge at 0.44 Mya (95% HPD: 0.24–0.69 Mya). All other members of *Gelidocalamus* appeared successively within the past three million years (Supplementary Table S5).

3.4 Genome structural variations and sequence divergences

The results of phylogenetic inference indicate that *Gelidocalamus* is a polyphyletic group, with its members intermingled with closely related genera, forming clades IV and V. Therefore, this study provides insights into the structure and evolution of the chloroplast

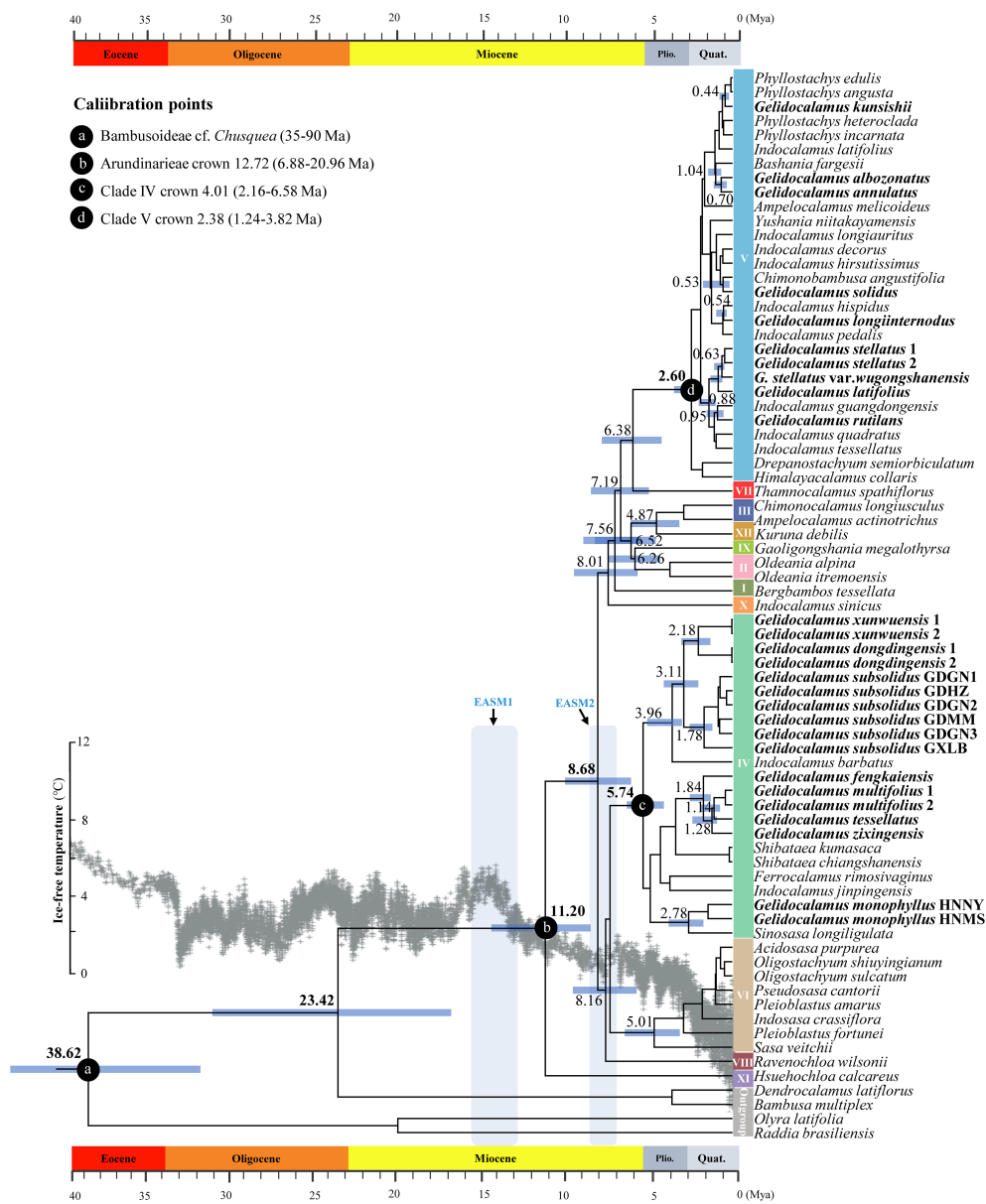


FIGURE 3 Divergence times of major clades in *Gelidocalamus*. The chronogram of the *Gelidocalamus* and relatives is based on BEAST analysis of plastomes DNA. Four calibration points are marked by red circles. The ice-free temperature change curves over the last 40 million years as modified in Zachos et al. (2008). Blue characters and shades of light blue indicate the approximate periods of two intensification of the East Asian monsoon climate. Node bars represent the 95% highest posterior density intervals for node ages, with the mean ages of some important nodes in Arundinarieae indicated.

genomes of *Gelidocalamus* and its close relatives within the two clades.

The collinearity analysis results based on Mauve indicate the absence of any detected rearrangements or inversions within these two clades (Supplementary Figure S6). The boundary expansion and contraction analysis revealed that six genes located at the junctions, namely *rpl22*, *rps19*, *rps15*, *ndhF*, and *ndhH*, exhibited consistent lengths across all samples, except for *psbA*, which showed variations in length in *G. subsolidus* GDGN3. Only the *ndhH* gene spans the boundaries between the SSC/IRA regions. In the boundary comparison analysis, we identified an insertion event specific to clade V. The insertion sequence “GGTTATCCCCG” at the LSC/

IRb junction represents an insertion event before the most recently common ancestor of clade V but was later lost in the ancestor of two *Phyllostachys* species (*P. edulis* and *P. angusta*) (Figure 4A; Supplementary Figure S7). Another insertion sequence “GAGGGGAGATAGAAAA”, located at the SSC/IRA junction, is present in many subclades of the Arundinarieae tribe but has been lost to varying degrees across them. Notably, all species in clade IV entirely lack this sequence (Figure 4A; Supplementary Figure S8). Additionally, in clade IV, a species-specific borderline insertion “GCAGAAAAGCCT” at the LSC/IRA junction was identified in the plastome of *G. subsolidus* GDGN3 (Figure 4B; Supplementary Figure S9).

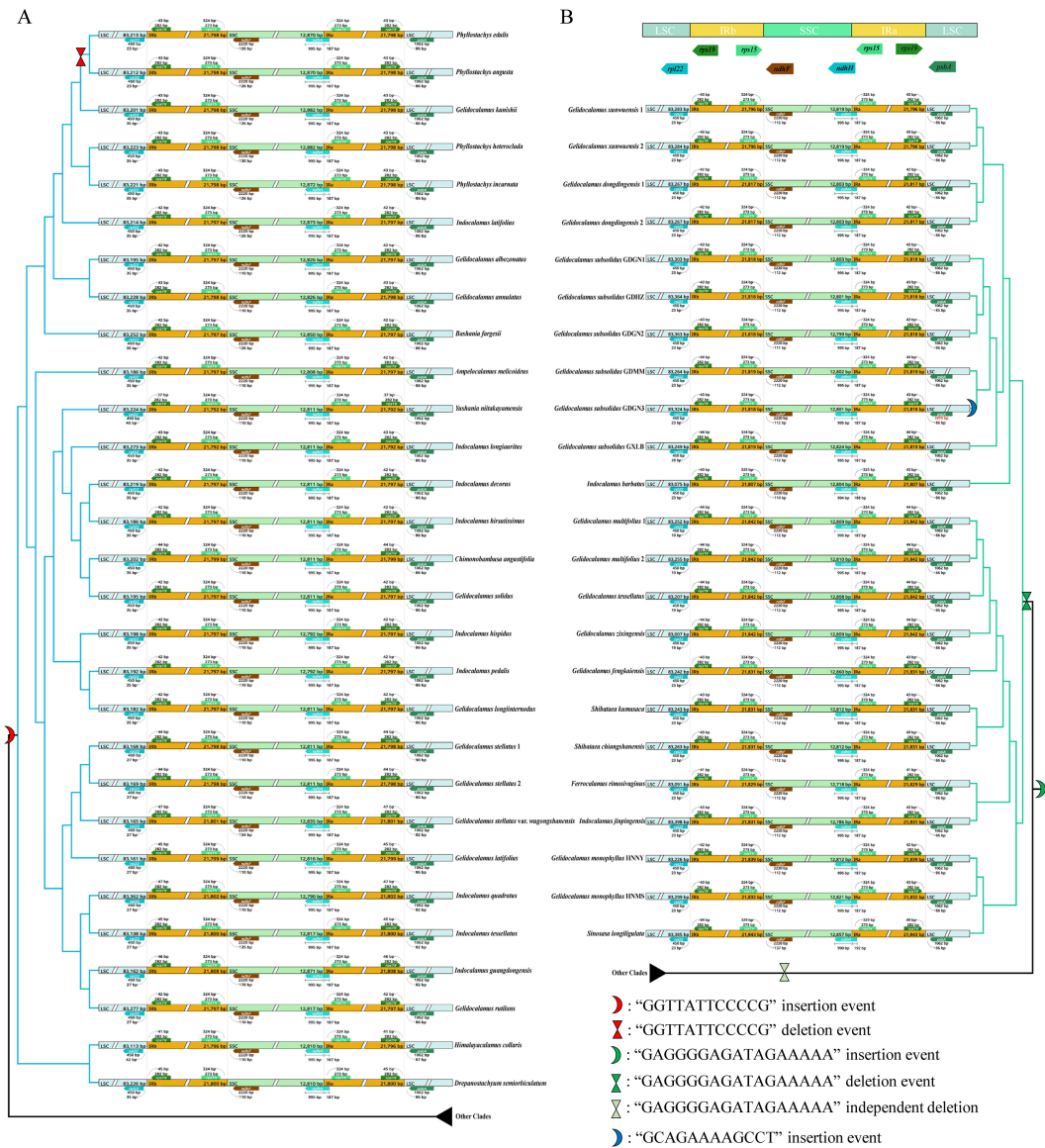


FIGURE 4 Visualization of the expansion and contraction in the inverted repeat region (IR) boundary of chloroplast genomes. The topology is based on the ML tree using complete plastomes. **(A)** Clade V; **(B)** clade IV.

The nucleotide diversity (π) of the two clades ranges from 0 to 0.005 in clade V and 0 to 0.00598 in clade IV, with average π values of 0.0016 and 0.0006, respectively. Notably, the average π value for the SSC region exceeds that of the LSC region, while the inverted repeat region exhibits the lowest nucleotide diversity in both clades (Figure 5). Among the variable sites, clade IV has 27 sites with a π value exceeding 0.004, whereas clade V has only three such sites. As shown in Figure 5, the highly variable regions in the clade IV include intergenic regions such as *trnK-rps16*, *trnG-trnT*, *rbcl-accD*, and *rpl32-trnL*, with *rbcl-accD* region exhibiting the highest π value (0.00598). Coding genes such as *ndhF*, *rpl32*, and *trnT* also show high variability, with *ndhF* having the highest π value (0.00595). In clade V, only three loci exceed a π value of 0.004: *rbcl*, *rbcl-accD*, and *trnG-trnT*, with the *rbcl* gene locus showing the highest π value (0.005).

3.5 Codon usage bias

The codon usage bias of clades IV and V is illustrated in Figure 6. A total of 29 preferred codons (RSCU > 1), 30 low-preference codons (RSCU < 1), and two codons showing no bias (RSCU = 1) were detected in clade V. However, in *Drepanostachyum semiorbiculatum* (clade V) and *G. monophyllus* HNNY (clade IV), 29 low-preference codons and three codons with no codon usage bias were identified. Among the 29 preferred codons, only two codons (UUG/UCC) end with C/G, while the rest end with A/U. The total codon usage among the IV clade chloroplast genomes ranges from 20,750 to 20,893 codons (average 20,866), and from 20,777 to 20,877 codons (average 20,870) among the V clade. Notably, the amino acid leucine, encoded by UUA, UUG, CUU, CUC, CUA, and CUG, has the highest frequency of

usage, ranging from 2,251 to 2,269 codons. In contrast, the amino acid cysteine, encoded by UGU and UGC, has the lowest frequency of usage, ranging from 229 to 233 codons.

3.6 SSRs and long repeats analysis

A total of 1,485 SSRs were identified in clade V and 1,328 in clade IV. The number of SSRs varied across species, ranging from 45 in *Ampelocalamus melicoideus* to 56 in *G. kunsishii* within clade V, and from 48 in *Indocalamus barbatus* to 65 in *G. multifolius* 2 within clade IV. In both clades that include *Gelidocalamus* members, mononucleotide repeats were the most abundant type of SSRs, followed by tetranucleotide repeats. Conversely, hexanucleotide repeats were the least frequent, found only in three species in each clade: *I. decorus*, *G. longinternodus*, and *Drepanostachyum semiorbiculatum* in clade V; and *Ferrocalamus rimosivaginus*, *I. jingpingensis*, and *G. monophyllus* HNMS in clade IV. As shown in Figure 7, among the 868 mononucleotide repeats in clade V, A/T repeats were predominant (97.8%), whereas C/G repeats were much less common (2.2%). Similarly, in clade IV, A/T repeats constituted a high percentage (95.3%) of the 759 mononucleotide repeats. Among the 118 dinucleotide repeats in clade V, TA repeats were the most abundant (49.1%), with AT and TC repeats being approximately equal (around 25%). In clade IV, TA repeats accounted for 66.6%, while AT and TC repeats were about 16.6% each. The number of species with other repeat types did not significantly differ between the two clades. Trinucleotide repeats in both clades were predominantly AAT, TAT, and TCT. Among tetranucleotide repeats, except for less frequent types such

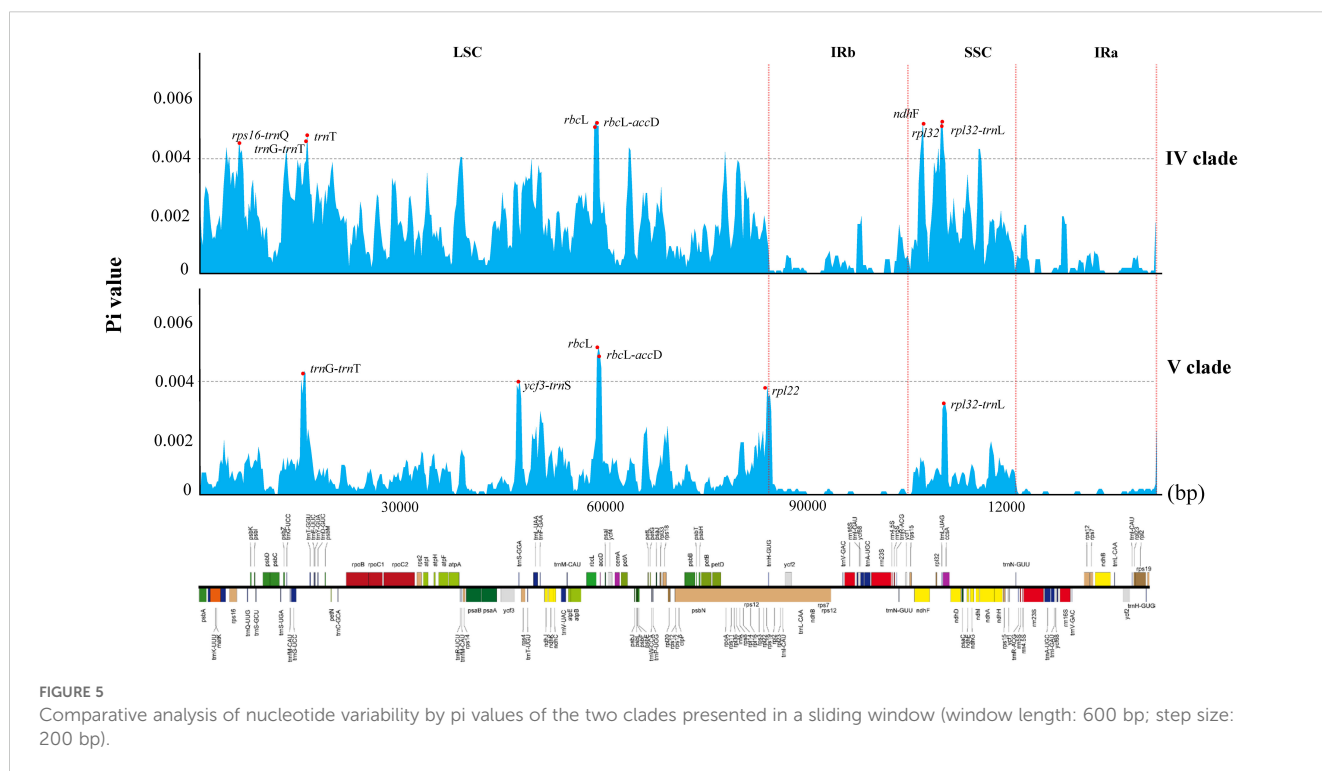
as AAAG, ATTA, TTAT, and TTCT, the quantities of other types were relatively similar. TTTTA was the most prevalent pentanucleotide repeat, and only three occurrences of hexanucleotide repeats were identified in each clade.

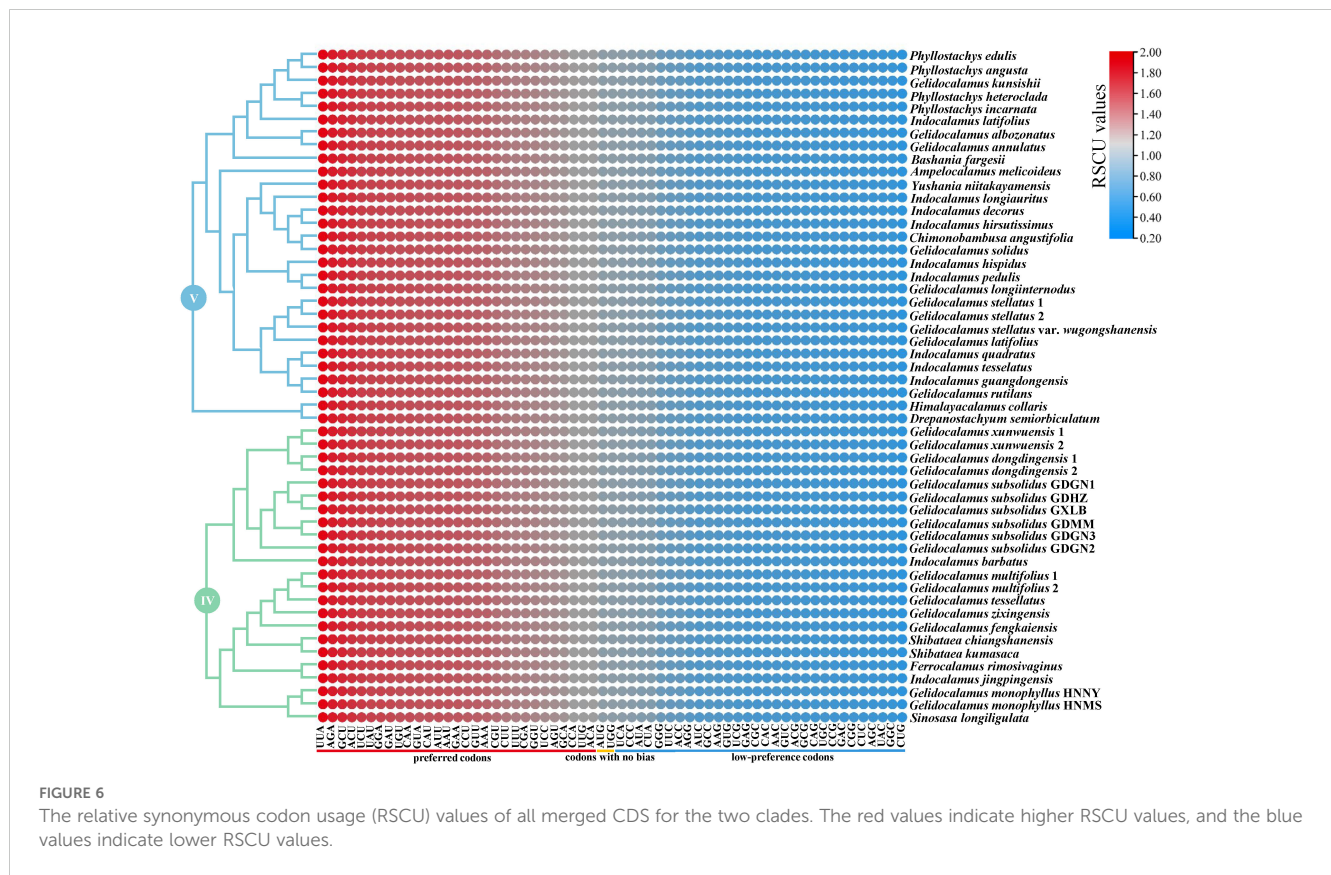
The distribution of SSRs within the chloroplast genomes varied across different regions, with the highest density found in the LSC region and fewer in the SSC and IR regions. The distribution also varied across intron, CDS, and IGS regions, with the IGS region containing the largest number of SSRs. The proportion of distribution across different regions was relatively consistent among species (Supplementary Figures S10–S12). Additionally, we identified 1,907 and 1,647 long repeats in clades V and IV, respectively. Forward repeats were the predominant type of long repeat. Similar to SSRs, the distribution of long repeats is primarily concentrated in the LSC region. However, unlike SSRs, the distribution of long repeats showed that the highest number was found in the CDS region, followed by the IGS region, with the intron region having the fewest (Supplementary Figures S13–S15).

4 Discussion

4.1 Assembly efficiency of different molecular markers from genome skimming data of *Gelidocalamus*

Previous studies have utilized low-depth genomic sequencing data to capture highly copied chloroplast genomes within the total genomic DNA, while the effectiveness of other molecular markers, such as mitochondrial genome and nuclear ribosomal DNA repeats





has been limited. Recent studies indicated that the organelle genome data comprises only a small portion of the total genomic DNA (Straub et al., 2012). By increasing the sequencing depth to around 10X, it is possible to retrieve complete chloroplast, mitochondrial genomes, nrDNA repeats, and a substantial number of single-copy nuclear genes (SCNs) from genome skimming data (Vargas et al., 2019; Liu et al., 2021). To evaluate the feasibility of retrieving complete mitochondrial genome data and nuclear ribosomal DNA repeats from low-depth genome skimming data (2-4 X coverage) in bamboo, we performed assemblies using the sequencing data generated in this study. The results indicated successful retrieval of complete chloroplast genomes and nrDNA repeats, although the reconstruction of the mitochondrial genome remained incomplete. A total of 29 conserved mitochondrial protein-coding genes were identified. Previous studies have shown that a sequencing depth of less than 5X coverage is sufficient to assemble the complete chloroplast genome, some mitochondrial coding genes and the entire nrDNA repeats (Straub et al., 2012; Fonseca and Lohmann, 2020). For the Arundinarieae tribe, it has been reported that approximately 2G of data (1X coverage) is adequate for assembling the complete chloroplast genome and nrDNA repeats, which aligns with our findings. Consequently, in bamboo plants, the complete chloroplast genome and nrDNA repeats can still be assembled with shallower sequencing coverage. Additionally, our study found that approximately two times coverage was adequate to assemble partially complete mitochondrial coding genes and entire nrDNA repeats. However, low-coverage genomic data proved insufficient for successfully assembling single-copy nuclear genes (Guo et al., 2021).

4.2 Phylogenetic relationship of *Gelidocalamus*

The incorporation of new molecular evidence into analyses has progressively refined the framework of phylogenetic relationships within the Arundinarieae tribe and its major clades. Previous studies, however, involved only a limited number of *Gelidocalamus* to represent the genus, leaving the internal phylogenetic relationships within *Gelidocalamus* unclear (Triplett and Clark, 2010; Zeng et al., 2010; Yang et al., 2013; Attigala et al., 2014; Zhang et al., 2020; Guo et al., 2021). Our study utilized extensive sampling, encompassing all 16 species and one variety of *Gelidocalamus*, totaling 27 samples from diverse populations. Representative samples from closely related genera were included to unravel both intrageneric and intergeneric relationships within the *Gelidocalamus*. The results based on chloroplast genomes strongly support the monophyly of the Arundinarieae tribe (BSML = 100% and PPBI = 1.00), which is divided into twelve well-supported clades. *Gelidocalamus* species are distributed across two of these clades (clades IV and clades V) with high support (BSML = 100% and PPBI = 1.00). This research achieved enhanced resolution compared to previous studies with low support values at various nodes. The findings indicate that species such as *G. tessellatus*, *G. xuwuensis*, *G. rutilans*, *G. latifolius*, and *G. stellatus* are distributed across two distinct clades, consistent with previous research (Ma et al., 2017; Guo et al., 2019; Zhang et al., 2019). Species not included in previous studies are also found within these two clades, and the phylogenetic relationships are strongly supported. These results affirm the polyphyletic nature of the

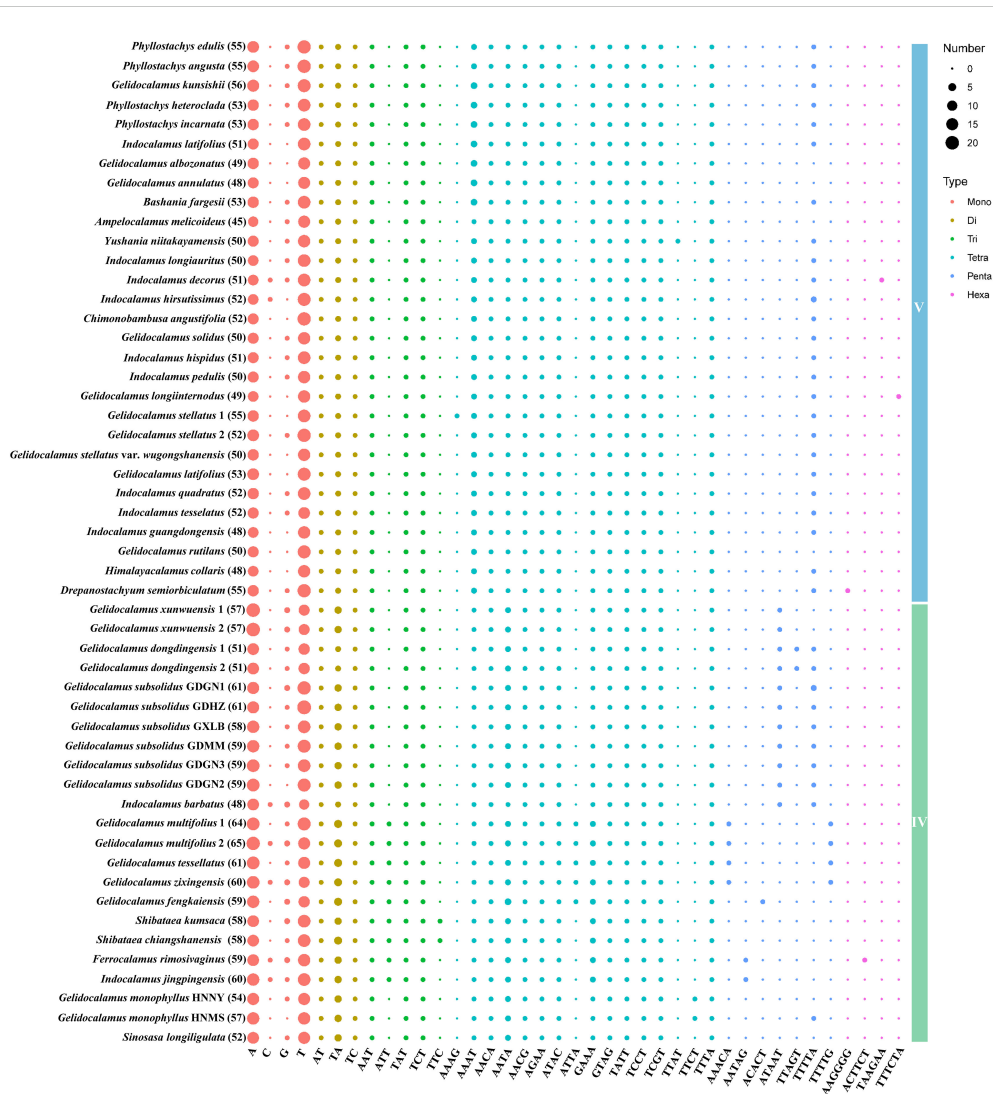


FIGURE 7 Plot of SSR repeat pattern numbers in two clades. The number following each species name indicates the total number of SSRs in each plastome.

Gelidocalamus, challenging traditional morphological classifications. For instance, *G. stellatus*, despite its morphological similarity to *G. tessellatus* and *G. xuwuensis*, belongs to a separate clade. Additionally, species with significant morphological differences, such as *G. longiinternodus*, which exhibit traits like sickle-shaped auricles, single branch on one-year-old plants, and spring sprouting (Wen, 1986), which display characteristics typical of *Indocalamus*. Molecular evidence supports a close relationship between *G. longiinternodus* and the *Indocalamus* genus. The classification boundaries of *Gelidocalamus* remain contentious and require further revision.

The use of different molecular markers to infer phylogenetic relationships often leads to controversies. For instance, Peng et al. (2008) found that *G. stellatus* is closely related to *Ferrocalamus strictus* based on nuclear single-copy genes *GBSSI* and *ITS*, contrasting sharply with results based on complete plastid genomes, indicating significant incongruence between nuclear and plastid. Similar conflicts have been observed in previous research, Guo et al. (2021) investigated the phylogenetic relationships within

the Arundinarieae tribe using ddRAD data, nrDNA sequences, and complete chloroplast genomes. Their results indicated that the phylogenetic relationships inferred from ddRAD data closely resembled those derived from nrDNA sequences, with *Gelidocalamus* consistently grouped within the “*Indocalamus* + *Gelidocalamus* + *Chimonocalamus*” clade. Most *Gelidocalamus* species formed a monophyletic branch, which aligns with the “core species of *Gelidocalamus*” (Li et al., 2024). However, the limited sampling of species for nrDNA sequences and chloroplast genomes restricted comprehensive comparisons across the different molecular datasets. Nevertheless, the phylogenetic relationships based on the chloroplast genome corroborate previous findings, although they exhibit significant conflict with those derived from nuclear genes. Potentially due to factors such as hybridization (Yang et al., 2013), incomplete lineage sorting, and chloroplast capture (Maddison, 1997; Peng et al., 2008; Triplett and Clark, 2010; Zhang et al., 2012; Ma et al., 2017; Guo et al., 2019, Guo et al., 2021). We attempted to establish intrageneric phylogenetic relationships using complete nrDNA repeats and 24 shared mitochondrial protein-coding genes

from each species. However, the results indicated that these highly conserved sequences provided limited informative sites, which impeded the construction of a clear phylogenetic framework. In some instances, the variation within species exceeded that between species (Supplementary Figures S4, S5), likely due to incomplete concerted evolution and pseudogene phenomena affecting ITS sequences in temperate woody bamboos (Gernandt et al., 2001; Wei et al., 2003). The use of different molecular markers introduced variability in phylogenetic relationships, adding complexity to the phylogenetic inference of *Gelidocalamus*. The observed conflict between nuclear and plastid genes suggests that these datasets may have experienced divergent evolutionary histories. Integrating morphological taxonomy with the ddRAD data, we propose that the phylogenetic relationships inferred from ddRAD sequencing offer a more robust framework for understanding the phylogeny of *Gelidocalamus*. This study has elucidated the phylogenetic relationships of *Gelidocalamus* using complete chloroplast genomes and assessed the utility of other molecular markers. However, for complex taxa with reticulate evolutionary histories, such as *Gelidocalamus*, phylogenetic resolution may not be fully achieved through molecular data alone. The morphological challenges associated with their taxonomy are equally complex. Therefore, a comprehensive approach that integrates both morphological features and molecular evidence is crucial for gaining deeper insights into their classification and evolutionary history.

4.3 Evolution history of *Gelidocalamus*

To investigate the origin and evolution of *Gelidocalamus*, we used data from complete chloroplast genomes to estimate divergence times. The findings reveal that the crown-group divergence time of the Arundinarieae tribe, to which *Gelidocalamus* belongs, is approximately in the Late Miocene, around 11.20 Mya (95% HPD: 8.22–14.59 Mya). Our estimated divergence time is slightly younger than those reported in earlier studies, such as 18.81 Mya (Christin et al., 2013) and 18.73 Mya (Guo et al., 2021), but older than 8.96 Mya (Bouchenak-Khelladi et al., 2010) and 10.01 Mya (Hodkinson et al., 2010), and closely aligned with 12.72 Mya (Zhang et al., 2016). The results suggest that *Gelidocalamus* is a relatively young genus, with all its members having differentiated within a short period of time. However, chloroplast genome data indicate that *Gelidocalamus* has a nonmonophyletic origin.

Previous studies have indicated that the Arundinarieae experienced a period of rapid radiation and differentiation during the Mid- or Late Miocene (Zhang et al., 2016; Guo et al., 2021). Our results align with those of Zhang et al. (2016), showing that the 12 major clades began to diverge rapidly in the Late Miocene. This rapid divergence is commonly linked to significant climate changes (Janssens et al., 2009; Hampe and Jump, 2011; Zou et al., 2013), particularly evident during two intensification of the East Asian monsoon (EASM) climate around 15.5–13 Mya and 8–7 Mya (Zhisheng et al., 2001; Zheng et al., 2004; Zhang et al., 2016). The

cooler temperatures and increased summer rainfall during these periods created favorable conditions for temperate woody bamboos. These bamboos swiftly colonized moist slopes, establishing themselves as dominant species. Their effective underground rhizome development further supported their rapid spread and diversification. Previous studies have suggested that incomplete lineage sorting (ILS) may be prevalent in the Arundinarieae tribe due to rapid radiative divergence, which led to species differentiation within a short time frame. This rapid divergence resulted in significant morphological changes that outpaced genetic diversification, leading to substantial morphological differences among species while they remain closely related genetically. Additionally, overlapping distributions and potential hybridization among species may have further complicated the phylogenetic relationships within the genus *Gelidocalamus*. Given the limitations of the dataset used in this study, these issues may require further investigation (Zhang et al., 2012; Yang et al., 2013; Guo et al., 2019, Guo et al., 2021).

4.4 The diversity of plastome characteristics

Chloroplasts, the primary site of photosynthesis, are crucial for plant physiological and developmental processes (Verma and Daniell, 2007; Li et al., 2020). In angiosperms, chloroplast genomes are mostly maternally inherited and highly conserved, with moderate substitution rates. Recently, chloroplast genomes have been widely used for plant phylogenetic studies (Zhang and Li, 2011; Wang et al., 2012; Cui et al., 2023; Peng et al., 2023).

To investigate the evolution of the chloroplast genome in *Gelidocalamus*, we conducted comparative analyses among species within the genus and its close relatives across the two clades in which it is distributed. The gene order and gene content are uniform across all chloroplast genomes. No significant expansion or contraction phenomena were detected at the SSC/IR region boundaries, nor were structural rearrangements or inversion events observed. Insertion events were observed at the boundaries of clades V and IV, with varying degrees of loss and new insertions occurring across different clades. Despite these variations, the overall similarity between the two clades remains high. The expansion and contraction of regions can largely be attributed to these insertions and deletions.

Synonymous codon preference, a crucial feature influencing gene expression and biological evolution, is impacted by variations and natural selection (Duret, 2000; Romero et al., 2000; Duret, 2002; Yuan et al., 2021). In this study, 29 preferred codons were identified among the protein-coding genes of two clades. Only two codons (UUG/UCC) ended with C/G, while the rest ended with A/U, consistent with previous research indicating a high level of consistency in codon preference and usage patterns within the subfamily Bambusoideae (Li et al., 2019). SSRs, a class of short-sequence repeat units widely distributed across genomes, are excellent genetic markers due to their polymorphic nature and high mutation rates (Cavalier-Smith, 2002; Yang et al., 2013; Asaf et al., 2018). In this study, we identified SSRs and long repeats

within the two clades containing *Gelidocalamus*. The results show that *Gelidocalamus* exhibits minimal variation compared to closely related genera within the same clade. However, substantial differences in the number of SSRs and long repeats were observed between congeneric species across different clades.

DNA barcoding employs standardized, sufficiently variable, easily amplifiable, and relatively short DNA sequences for species identification (Hebert et al., 2003; Kress et al., 2005). Various molecular barcodes have been developed for identifying species within the subfamily Bambusoideae (Zhang and Li, 2011). In this study, we identified 17 hypervariable regions ($\pi > 0.004$) in clade IV, including *trnK-rps16*, *trnG-trnT*, *rbcL-accD*, *rpl32-trnL*, *rbcL*, and *rpl32*. Conversely, only three hypervariable regions (*rbcL*, *rbcL-accD*, and *trnG-trnT*) were identified from clade V. Many of these identified highly variable loci have been previously used as molecular markers for rapid identification and phylogenetic analysis (Zhang and Li, 2011), demonstrating their stability for plant identification. However, variability among species within the two clades containing *Gelidocalamus* differs, suggesting that the degree of species differentiation within this genus is not uniform across the chloroplast genomes. This variability may contribute to the separation into two major clades and reflects similar patterns observed in related genera, underscoring the complexity of the phylogenetic relationships among bamboo species.

5 Conclusions

In this study, all species and varieties of *Gelidocalamus* were collected (except for *G. kunishii* and *G. fengkaiensis*, which lacked morphological survey records). A total of 27 chloroplast genomes were assembled from genome skimming data for 16 species and one variety within *Gelidocalamus*. Additionally, complete nrDNA repeats and 29 mitochondrial protein-coding genes shared by each species were assembled. Through extensive sampling, we reconstructed high-resolution intragenomic and intergeneric phylogenetic relationships within *Gelidocalamus* based on complete chloroplast genome data. However, the phylogenetic relationships based on nrDNA repeats and mitochondrial genes exhibited extremely low resolution. Estimations of Divergence times of *Gelidocalamus* suggested it to be a young genus. The cause of intragenomic polyphyly may be linked to a rapid radiation event experienced by the Arundinarieae tribe. Furthermore, a comparative analysis of chloroplast genome structure and variation was conducted on the two clades involving *Gelidocalamus* species. The results revealed a high degree of consistency among all chloroplast genomes in terms of structural collinearity, gene count, gene order, and synonymous codon preference. This study comprehensively elucidated the phylogenetic relationships within *Gelidocalamus* for the first time, estimated its evolutionary times, and compared structural differences in various chloroplast genomes. However, conflicts arising from different molecular markers, such as nuclear genes and plastid genomes, may require further evidence and in-depth exploration.

Data availability statement

The sequence data have been submitted to the GenBank databases under accession numbers OP920757-OP920759, and PP999719-PP999742.

Author contributions

CW: Methodology, Writing – original draft, Writing – review & editing, Data curation, Formal analysis, Investigation, Resources. YL: Conceptualization, Data curation, Investigation, Visualization, Writing – original draft, Writing – review & editing. GY: Conceptualization, Data curation, Investigation, Writing – review & editing, Supervision. WZ: Conceptualization, Investigation, Supervision, Writing – review & editing, Data curation, Methodology. CG: Conceptualization, Methodology, Writing – review & editing, Supervision, Funding acquisition, Project administration, Visualization, Writing – original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Natural Science Foundation of China (No. 31960051) and the Natural Science Foundation for Young Scientists of Jiangxi Province (No. 20192ACB21005).

Acknowledgments

We thank the core facility of the Jiangxi Provincial Key Laboratory of Improved Variety Breeding and Efficient Utilization of Native Tree Species (2024SSY04093) for providing the experimental platform. We are grateful to Chunling Long (Zhejiang Normal University), Ting Kong (Jiangxi Matou Mountain National Nature Reserve), Weijian Li (Nanchang Business College, Jiangxi Agricultural University), Yuguang Liu (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), and Prof. Nianhe Xia (South China Botanical Garden, Chinese Academy of Sciences) for their work in field surveys and sampling.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2024.1470311/full#supplementary-material>

SUPPLEMENTARY DATA SHEET 1

The alignment of 75 complete chloroplast genomes.

SUPPLEMENTARY DATA SHEET 2

The alignment of 75 complete chloroplast genomes with the inverted repeat region removed.

SUPPLEMENTARY DATA SHEET 3

The alignment of nrDNA repeat sequences.

SUPPLEMENTARY DATA SHEET 4

The alignment of 24 shared mitochondrial genes from 27 samples.

References

- Amiryousefi, A., Hyvönen, J., and Pocza, P. (2018). IRscope: an online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* 34, 3030–3031. doi: 10.1093/bioinformatics/bty220
- Asaf, S., Budak, H., Khan, A. L., Khan, M. A., Shahzad, R., Lubna, et al. (2018). Complete chloroplast genome sequence and comparative analysis of loblolly pine (*Pinus taeda* L.) with related species. *PLoS One* 13, e0192966. doi: 10.1371/journal.pone.0192966
- Attigala, L., Triplett, J. K., Kathiriarachchi, H. S., and Clark, L. G. (2014). A new genus and a major temperate bamboo lineage of the Arundinarieae (Poaceae: Bambusoideae) from Sri Lanka based on a multi-locus plastid phylogeny. *Phytotaxa* 174, 187. doi: 10.11646/phytotaxa.174.4.1
- Beier, S., Thiel, T., Münch, T., Scholz, U., Mascher, M., and Valencia, A. (2017). MISA-web: a web server for microsatellite prediction. *Bioinformatics* 33, 2583–2585. doi: 10.1093/bioinformatics/btx198
- Bouchenak-Khelladi, Y., Verboom, G. A., Savolainen, V., and Hodkinson, T. R. (2010). Biogeography of the grasses (Poaceae): a phylogenetic approach to reveal evolutionary history in geographical space and geological time. *Bot. J. Linn. Soc.* 162, 543–557. doi: 10.1111/j.1095-8339.2010.01041.x
- Bouckaert, R., Perte, M., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., et al. (2019). BEAST 2.5: an advanced software platform for bayesian evolutionary analysis. *PLoS Comput. Biol.* 15, e1006650. doi: 10.1371/journal.pcbi.1006650
- Cai, Z. Y., Zhou, X. X., Wong, K. M., and Xia, N. H. (2021). *Gelidocalamus fengkaiensis* (Poaceae: Bambusoideae), a new bamboo species from Guangdong, China, with an analysis of branch development in relation to flowering. *Bot. Stud.* 62, 12. doi: 10.1186/s40529-021-00319-4
- Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348
- Cavalier-Smith, T. (2002). Chloroplast evolution: secondary symbiogenesis and multiple losses. *Curr. Biol.* 12, R62–R64. doi: 10.1016/S0960-9822(01)00675-3
- Chen, C. J., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y. H., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009
- Chen, S. F., Zhou, Y. Q., Chen, Y. R., and Gu, J. (2018). fastp: an ultra-fast all-in-one fastq preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560
- Christin, P. A., Spriggs, E., Osborne, C. P., Stromberg, C. A. E., Edwards, E., and Salamin, N. (2013). Molecular dating, evolutionary rates, and the age of the grasses. *J. Syst. Biol.* 63, 153–165. doi: 10.1093/sysbio/syt072
- Cui, Y. F., Zhou, P., Xiang, K. L., Zhang, Q., Yan, H., Zhang, L. G., et al. (2023). Plastome evolution and phylogenomics of *Trichosporeae* (Gesneriaceae) with its morphological characters appraisal. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1160535
- Duret, L. (2000). TRNA gene number and codon usage in the *C. elegans* genome are co-adapted for optimal translation of highly expressed genes. *Trends Genet.* 16, 287–289. doi: 10.1016/S0168-9525(00)02041-2
- Duret, L. (2002). Evolution of synonymous codon usage in metazoans. *Curr. Opin. Genet. Dev.* 12, 640–649. doi: 10.1016/S0959-437X(02)00353-2
- Fonseca, L. H. M., and Lohmann, L. G. (2020). Exploring the potential of nuclear and mitochondrial sequencing data generated through genome-skimming for plant phylogenetics: A case study from a clade of neotropical lianas. *J. Syst. Evol.* 58, 18–32. doi: 10.1111/jse.12533
- Gallaher, T. J., Peterson, P. M., Soreng, R. J., Zuloaga, F. O., Li, D. Z., Clark, L. G., et al. (2022). Grasses through space and time: an overview of the biogeographical and macroevolutionary history of Poaceae. *J. Syst. Evol.* 60, 522–569. doi: 10.1111/jse.12857
- Gernandt, D. S., Liston, A., and Piñero, D. (2001). Variation in the nrDNA ITS of *Pinus* subsection *Cembroides*: implications for molecular systematic studies of pine species complexes. *Mol. Phylogenet. Evol.* 21, 449–467. doi: 10.1006/mpev.2001.1026
- Guo, C., Guo, Z. H., and Li, D. Z. (2019). Phylogenomic analyses reveal intractable evolutionary history of a temperate bamboo genus (Poaceae: Bambusoideae). *Plant Divers.* 41, 213–219. doi: 10.1016/j.pld.2019.05.003
- Guo, C., Ma, P. F., Yang, G. Q., Ye, X. Y., Guo, Y., Liu, J. X., et al. (2021). Parallel ddRAD and genome skimming analyses reveal a radiative and reticulate evolutionary history of the temperate bamboos. *Syst. Biol.* 70, 756–773. doi: 10.1093/sysbio/syaa076
- Hampe, A., and Jump, A. S. (2011). Climate relicts: past, present, future. *Annu. Rev. Ecol. Syst.* 42, 313–333. doi: 10.1146/annurev-ecolsys-102710-145015
- Hebert, P. D. N., Cywinska, A., Ball, S. L., and deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proc. R. Soc. B* 270, 313–321. doi: 10.1098/rspb.2002.2218
- Hodkinson, T. R., Ni Chonghaile, G., Sungkaew, S., Chase, M. W., Salamin, N., and Stapleton, C. M. A. (2010). Phylogenetic analyses of plastid and nuclear DNA sequences indicate a rapid late Miocene radiation of the temperate bamboo tribe Arundinarieae (Poaceae, Bambusoideae). *Plant Ecol. Divers.* 3, 109–120. doi: 10.1080/17550874.2010.521524
- Huang, W., Zhang, L., Columbus, J. T., Hu, Y., Zhao, Y. Y., Tang, L., et al. (2022). A well-supported nuclear phylogeny of Poaceae and implications for the evolution of C4 photosynthesis. *Mol. Plant* 15, 755–777. doi: 10.1016/j.molp.2022.01.015
- Janssens, S. B., Knox, E. B., Huysmans, S., Merckx, V., and Smets, E. F. (2009). Rapid radiation of *Impatiens* (Balsaminaceae) during Pliocene and Pleistocene: result of a global climate change. *S. F. T. Mol. Phylogenet. Evol.* 52, 806–824. doi: 10.1016/j.jympev.2009.04.013
- Jin, J. J., Yu, W. B., Yang, J. B., Song, Y., dePamphilis, C. W., Yi, T. S., et al. (2020). GetOrganelle: a fast and versatile toolkit for accurate *de novo* assembly of organelle genomes. *Genome Biol.* 21, 241. doi: 10.1186/s13059-020-02154-5
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Jermiin, L., and von Haeseler, A. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *S. Nat. Methods* 14, 587–589. doi: 10.1038/nmeth.4285
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199
- Kress, W. J., Wurdack, K. J., Zimmer, E. A., Weigt, L. A., and Janzen, D. H. (2005). Use of DNA barcodes to identify flowering plants. *Proc. Natl. Acad. Sci.* 102, 8369–8374. doi: 10.1073/pnas.0503123102
- Kurtz, S., Choudhuri, J. V., Ohlebusch, E., Schleiermacher, C., Stoye, J., and Giegerich, R. (2001). REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* 29, 4633–4642. doi: 10.1093/nar/29.22.4633
- Li, D. Z., Wang, Z. P., Zhu, Z. D., Xia, N. H., Jia, L. Z., Guo, Z. H., et al. (2006). Bambuseae (Poaceae). In: *Flora of China*, Eds. Z. Y. Wu, P. H. Raven, and D. Y. Hong. Science Press and Missouri Botanical Garden Press, Beijing and St. Louis.
- Li, J. P., Qin, Z., Guo, C. C., Zhang, W. G., and Yang, G. Y. (2019). Codon bias in the chloroplast genome of *Gelidocalamus tessellatus*. *J. Bamboo Res.* 38, 79–87. doi: 10.19560/j.cnki.issn1000-6567.2019.02.013
- Li, W. J., Zhang, W. G., Tang, M., Ji, C. F., and Yang, G. Y. (2016). The specimen collection situation and species distribution of *Gelidocalamus*. *J. Bamboo Res.* 35, 1–7.
- Li, Y. H., Ren, Y. K., Zhao, X. H., Liu, J., Han, B., Wang, C. B., et al. (2020). Research progress on chloroplast genome of major Gramineous crops. *Biotechnol. Bull.* 36, 112–121. doi: 10.13560/j.cnki.biotech.bull.1985.2020-0285
- Li, Y. L., Guo, R., Zhang, H. J., Yi, S. R., Yang, G. Y., and Zhang, W. G. (2023). *Gelidocalamus albozonatus* (Poaceae, Bambusoideae), a new species from the southeast of Chongqing, China, and analysis of the morphological diversity in the core group of *Gelidocalamus*. *PhytoKeys* 236, 17–27. doi: 10.3897/phytokeys.236.111290
- Liao, G. R. (1991). The stand features and exploitative prospects of *Gelidocalamus stellatus* forests. *J. Bamboo Res.* 10, 19–22.
- Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187

- Liu, Y. G., Li, W. J., Tang, M., Yang, G. Y., and Zhang, W. G. (2017). Taxonomic re-evaluation of some *Gelidocalamus* (Poaceae: Bambusoideae) taxa from Southeast China. *Phytotaxa* 299, 111–117. doi: 10.11646/phytotaxa.299.1.9
- Liu, B. B., Ma, Z. Y., Ren, C., Hodel, R. G. J., Sun, M., Liu, X. Q., et al. (2021). Capturing single-copy nuclear genes, organellar genomes, and nuclear ribosomal DNA from deep genome skimming data for plant phylogenetics: a case study in Vitaceae. *J. Syst. Evol.* 59, 1124–1138. doi: 10.1111/jse.12806
- Ma, P. F., Vorontsova, M. S., Nanjarisoa, O. P., Razanatsoa, J., Guo, Z. H., Haevermans, T., et al. (2017). Negative correlation between rates of molecular evolution and flowering cycles in temperate woody bamboos revealed by plastid phylogenomics. *BMC Plant Biol.* 17, 260. doi: 10.1186/s12870-017-1199-8
- Ma, P. F., Zhang, Y. X., Zeng, C. X., Guo, Z. H., and Li, D. Z. (2014). Chloroplast phylogenomic analyses resolve deep-level relationships of an intractable bamboo tribe Arundinarieae (poaceae). *Syst. Biol.* 63, 933–950. doi: 10.1093/sysbio/syu054
- Maddison, W. P. (1997). Gene trees in species trees. *Syst. Biol.* 46, 523–536. doi: 10.1093/sysbio/46.3.523
- Murray, M. G., and Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8, 4321–4326. doi: 10.1093/nar/8.19.4321
- Nguyen, L. T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300
- Nie, T. J., Li, W. J., Ji, X. N., Liu, Y. G., Li, Z. Y., Yang, G. Y., et al. (2018). Re-evaluation of the taxonomy of *Gelidocalamus stellatus* (Poaceae: Bambusoideae) and its infraspecific taxa from southern China. *Phytotaxa* 356, 215–225. doi: 10.11646/phytotaxa.356.3.3
- Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., et al. (1986). Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* 322, 572–574. doi: 10.1038/322572a0
- Peng, C., Guo, X. L., Zhou, S. D., and He, X. J. (2023). Backbone phylogeny and adaptive evolution of *Pleurospermum* s. l.: new insights from phylogenomic analyses of complete plastome data. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1148303
- Peng, S., Yang, H. Q., and Li, D. Z. (2008). Highly heterogeneous generic delimitation within the temperate bamboo clade (Poaceae: Bambusoideae): evidence from GBSSI and ITS sequences. *Taxon* 57, 799–810. doi: 10.1002/tax.573011
- Qin, Q. M., Tong, Y. H., Zheng, X. R., Ni, J. B., and Xia, N. H. (2020). *Sinosasa* (Poaceae: Bambusoideae), a new genus from China. *Taxon* 70, 27–47. doi: 10.1002/tax.12422
- Romero, H., Zavala, A., and Musto, H. (2000). Codon usage in *Chlamydia trachomatis* is the result of strand-specific mutational biases and a complex pattern of selective forces. *Nucleic Acids Res.* 28, 2084–2090. doi: 10.1093/nar/28.10.2084
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012). MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542. doi: 10.1093/sysbio/syr029
- Shi, L. C., Chen, H. M., Jiang, M., Wang, L. Q., Wu, X., Huang, L. F., et al. (2019). CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* 47, W65–W73. doi: 10.1093/nar/gkz345
- Straub, S. C. K., Parks, M., Weitemier, K., Fishbein, M., Cronn, R. C., and Liston, A. (2012). Navigating the tip of the genomic iceberg: next-generation sequencing for plant systematics. *Am. J. Bot.* 99, 349–364. doi: 10.3732/ajb.1100335
- Strömberg, C. A. E. (2011). Evolution of grasses and grassland ecosystems. *Annu. Rev. Earth Planet.* 39, 517–544. doi: 10.1146/annurev-earth-040809-152402
- Sugiura, M., Shinozaki, K., Zaita, N., Kusuda, M., and Kumano, M. (1986). Clone bank of the tobacco (*Nicotiana tabacum*) chloroplast genome as a set of overlapping restriction endonuclease fragments: mapping of eleven ribosomal protein genes. *Plant Sci.* 44, 211–217. doi: 10.1016/0168-9452(86)90093-2
- Tong, Y. H., Zheng, X. R., Zhang, Y. Y., Qin, Q. M., Ni, J. B., Vu, T. C., et al. (2020). *Schizostachyum dakrongense* (Poaceae, Bambusoideae), a new species from Dakrong nature reserve, Vietnam. *PhytoKeys* 138, 179–186. doi: 10.3897/phytokeys.138.39512
- Triplett, J. K., and Clark, L. G. (2010). Phylogeny of the temperate Bamboos (Poaceae: Bambusoideae: Bambuseae) with an emphasis on Arundinaria and allies. *Syst. Biol.* 35, 102–120. doi: 10.1600/036364410790862678
- Vargas, O. M., Heuert, M., Smith, S. A., and Dick, C. W. (2019). Target sequence capture in the Brazil nut family (Lecythidaceae): Marker selection and in silico capture from genome skimming data. *Mol. Phylogenet. Evol.* 135, 98–104. doi: 10.1016/j.ympev.2019.02.020
- Verma, D., and Daniell, H. (2007). Chloroplast vector systems for biotechnology applications. *Plant Physiol.* 145, 1129–1143. doi: 10.1104/pp.107.106690
- Wang, L., Dong, W. P., and Zhou, S. L. (2012). Structural mutations and reorganizations in chloroplast genomes of flowering plants. *Acta Bot. Boreal.-Occident. Sin.* 32, 1282–1288.
- Wang, C. K., Guo, R., Guo, C. C., Yang, G. Y., and Zhang, W. G. (2023). *Gelidocalamus zixingensis* (Poaceae, Bambusoideae, Arundinarieae), a new species from southern China revealed by morphological and molecular evidence. *PhytoKeys* 218, 29–45. doi: 10.3897/phytokeys.218.96849
- Wei, X. X., Wang, X. Q., and Hong, D. Y. (2003). Marked intragenomic heterogeneity and geographical differentiation of nrDNA ITS in *Larix potaninii* (Pinaceae). *J. Mol. Evol.* 57, 623–635. doi: 10.1007/s00239-003-2512-8
- Wen, T. H. (1982). A new genus and some new species of Bambusoideae from China. *J. Bamboo Res.* 1, 20–45.
- Wen, T. H. (1986). Some issues of taxa of Bambusoideae in China. *J. Bamboo Res.* 02, 10–27.
- Wick, R. R., Schultz, M. B., Zobel, J., and Holt, K. E. (2015). Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics* 31, 3350–3352. doi: 10.1093/bioinformatics/btv383
- Xie, P., Guo, Y. L., Zhou, W. B., Zhang, Z., and Yu, Y. (2023). GeneMiner: a tool for extracting phylogenetic markers from next generation sequencing data. *Authorea*. 24 (3), e13924. doi: 10.22541/au.168172406.69677221/v1
- Yang, B. M. (1986). Three new species of bamboo native of Hunan. *J. Hunan Univ. Nat. Sci.* 3, 89–94.
- Yang, J. B., Tang, M., Li, H. T., Zhang, Z. R., and Li, D. Z. (2013). Complete chloroplast genome of the genus *Cymbidium*: lights into the species identification, phylogenetic implications and population genetic analyses. *BMC Evol. Biol.* 13, 84. doi: 10.1186/1471-2148-13-84
- Yang, H. M., Zhang, Y. X., Yang, J. B., and Li, D. Z. (2013). The monophyly of *Chimonocalamus* and conflicting gene trees in Arundinarieae (Poaceae: Bambusoideae) inferred from four plastid and two nuclear markers. *Mol. Phylogenet. Evol.* 68, 340–356. doi: 10.1016/j.ympev.2013.04.002
- Ye, X. Y., Ma, P. F., Yang, G. Q., Guo, C., Zhang, Y. X., Chen, Y. M., et al. (2019). Rapid diversification of alpine bamboos associated with the uplift of the Hengduan Mountains. *J. Biogeogr.* 46, 2678–2689. doi: 10.1111/jbi.13723
- Yuan, X. L., Li, Y. Q., Zhang, J. F., and Wang, Y. (2021). Analysis of codon usage bias in the chloroplast genome of *Dalbergia odorifera*. *Guihaia* 41, 622–630.
- Zachos, J., Dickens, G., and Zeebe, R. (2008). An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* 451, 279–283. doi: 10.1038/nature06588
- Zeng, C. X., Zhang, Y. X., Triplett, J. K., Yang, J. B., and Li, D. Z. (2010). Large multi-locus plastid phylogeny of the tribe Arundinarieae (Poaceae: Bambusoideae) reveals ten major lineages and low rate of molecular divergence. *Mol. Phylogenet. Evol.* 56, 821–839. doi: 10.1016/j.ympev.2010.03.041
- Zhang, D., Gao, F. L., Jakovlić, I., Zou, H., Zhang, J., Li, W. X., et al. (2019). PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* 20, 348–355. doi: 10.1111/1755-0998.13096
- Zhang, Y. X., Guo, C., and Li, D. Z. (2020). A new subtribal classification of Arundinarieae (Poaceae, Bambusoideae) with the description of a new genus. *Plant Divers.* 42, 127–134. doi: 10.1016/j.pld.2020.03.004
- Zhang, Y. T., Guo, C. C., Yang, G. Y., Yu, F., and Zhang, W. G. (2019). The complete chloroplast genome sequence of *Gelidocalamus xunwuensis* (Bambusoideae: Arundinarieae): a shrubby bamboo endemic to China. *Mitochondrial DNA B Resour.* 4, 3352–3353. doi: 10.1080/23802359.2019.1673679
- Zhang, W. G., Ji, X. N., Liu, Y. G., Li, W. J., and Yang, G. Y. (2017). *Gelidocalamus xunwuensis* (Poaceae, Bambusoideae), a new species from southeastern Jiangxi, China. *PhytoKeys* 85, 59–67. doi: 10.3897/phytokeys.85.13804
- Zhang, Y. J., and Li, D. Z. (2011). Advances in phylogenomics based on complete chloroplast genomes. *Plant Divers.* 33, 365–375.
- Zhang, Y. X., Ma, P. F., and Li, D. Z. (2018). A new genus of temperate woody bamboos (Poaceae, Bambusoideae, Arundinarieae) from a limestone montane area of China. *PhytoKeys* 109, 67–76. doi: 10.3897/phytokeys.109.27566
- Zhang, Y. X., Zeng, C. X., and Li, D. Z. (2012). Complex evolution in Arundinarieae (Poaceae: Bambusoideae): Incongruence between plastid and nuclear GBSSI gene phylogenies. *Mol. Phylogenet. Evol.* 63, 777–797. doi: 10.1016/j.ympev.2012.02.023
- Zhang, X. Z., Zeng, C. X., Ma, P. F., Haevermans, T., Zhang, Y. X., Zhang, L. N., et al. (2016). Multi-locus plastid phylogenetic biogeography supports the Asian hypothesis of the temperate woody bamboos (Poaceae: Bambusoideae). *Mol. Phylogenet. Evol.* 96, 118–129. doi: 10.1016/j.ympev.2015.11.025
- Zheng, S., Poczai, P., Hyvönen, J., Tang, J., and Amiryousefi, A. (2020). Chloroplot: an online program for the versatile plotting of organelle genomes. *Front. Genet.* 11. doi: 10.3389/fgene.2020.576124
- Zheng, H. B., Powell, C. M., Rea, D. K., Wang, J. L., and Wang, P. X. (2004). Late Miocene and mid-Pliocene enhancement of the East Asian monsoon as viewed from the land and sea. *Glob. Planet. Change* 41, 147–155. doi: 10.1016/j.gloplacha.2004.01.003
- Zhisheng, A., Kutzbach, J. E., Prell, W. L., and Porter, S. C. (2001). Evolution of Asian monsoons and phased uplift of the Himalaya–Tibetan plateau since Late Miocene times. *Nature* 411, 62–66. doi: 10.1038/35075035
- Zhou, M. Y., Liu, J. X., Ma, P. F., Yang, J. B., and Li, D. Z. (2022). Plastid phylogenomics shed light on intergeneric relationships and spatiotemporal evolutionary history of *Melocanninae* (Poaceae: Bambusoideae). *J. Syst. Evol.* 60, 640–652. doi: 10.1111/jse.12843
- Zou, X. H., Yang, Z., Doyle, J. J., and Ge, S. (2013). Multilocus estimation of divergence times and ancestral effective population sizes of *Oryza* species and implications for the rapid diversification of the genus. *New Phytol.* 198, 1155–1164. doi: 10.1111/nph.12230