

Chloroplast genome study of *Gueldenstaedtia monophylla* (Fabaceae): unique genome within Chesneyeae

Dashzeveg Oyuntsetseg¹, Nudkhuu Nyamgerel², Shukherdorj Baasanmunkh^{1,2},
Dariganga Munkhtulga², Batlai Oyuntsetseg¹, Gun-Aajav Bayarmaa¹,
Alexey A. Kechaykin³, Courtney H. Babin⁴, Hyeok Jae Choi²

1 Department of Biology, School of Arts and Sciences, National University of Mongolia, Ulaanbaatar, 14201, Mongolia

2 Department of Biology and Microbiology, Changwon National University, Changwon, 51140, Republic of Korea

3 South Siberian Botanical Garden, Altai State University, 61 Lenina Ave., Barnaul, 656049, Russia

4 Department of Biological Sciences, Louisiana State University New Orleans, New Orleans, LA, United States

Corresponding authors: Shukherdorj Baasanmunkh (baasanmunkh.sh@gmail.com), Hyeok Jae Choi (hjchoi1975@changwon.ac.kr)

Academic editor: R. Yakovlev | Received 3 September 2026 | Accepted 2 October 2026 | Published 23 October 2026

<http://zoobank.org/E29868C4-1CDA-48BA-B406-5A25D149A99D>

Citation: Oyuntsetseg D, Nyamgerel N, Baasanmunkh S, Munkhtulga D, Oyuntsetseg B, Bayarmaa G, Kechaykin AA, Babin CH, Choi HJ (2026) Chloroplast genome study of *Gueldenstaedtia monophylla* (Fabaceae): unique genome within Chesneyeae. Acta Biologica Sibirica 12: 1335–1356. <https://doi.org/10.5281/zenodo.23533124>

Abstract

Gueldenstaedtia Fisch. is the smallest genus of the Fabaceae and contains only four species. In higher classification, this genus belonging to tribe Chesneyeae, together with *Chesneya* Lindl. ex Endl. and *Tibetia* (Ali) H.P. Tsui. However, the phylogenetic position and comparative genetic relationships between species remain unresolved due to the scarcity of available genetic data. In this study, we utilized high-throughput sequencing of complete chloroplast genomes of *Gueldenstaedtia monophylla* Fisch. and conducted comparative analyses with related species within the tribe. The complete chloroplast genome of *G. monophylla* is 118,980 bp in length and smallest genome within the tribe. It belongs to the inverted repeat-lacking clade (IRLC) of legumes and encodes 110 genes, including 76 protein-coding genes, 30 tRNAs, and 4 rRNAs. As typical of IRLC legumes, the *rpl22* and *rps16* genes were completely lost and the *clpP* and *rps12* genes each lost one intron. *Gueldenstaedtia* plastid genomes exhibited several large-scale rearrangements that distinguish them from other genera within the tribe.

Furthermore, *G. monophylla* retains some ancestral genomic architecture compared to rest of the genus. Phylogenetic analysis showed that *G. monophylla* exhibited a distinct position within the genus, with an estimated divergence time of 7.21 MYA. In addition, codon usage pattern and selective pressure analysis of protein-coding genes supported the evolutionary distinctiveness of *G. monophylla* within the genus. This reflects a unique evolutionary history and adaptive divergence in core genes, highlighting the importance of *G. monophylla* for understanding the origin and diversification of the genus.

Keywords

IRLC, chloroplast genome rearrangement, molecular evolution, ancient species, original taxonomic position at the subgenus level

Introduction

Gueldenstaedtia Fisch. is the smallest genus of the Fabaceae family and contains only four species worldwide, including *G. guangxiensis* W.L. Sha & X.X. Chen, *G. henryi* Ulbr., *G. monophylla* Fisch., and *G. verna* (Georgi) Boriss. (Zhu 2004; Xie et al. 2016; Baasanmunkh et al. 2022, 2026a). The native range of this genus is distributed from the Sino-Himalayan region to Mongolia, Siberia, and South Korea (Seol et al. 2025). In the past, Zhu (2004) first revised the genus, including the global distribution, illustration, taxonomic notes, and nomenclature. Later, Zhu (2005) studied the pollen and seed morphology of this genus. The morphology of *Gueldenstaedtia* is characterized by perennial herbs, thick and woody roots, and flowers that are pink, purple, or yellow (Zhu 2004). The morphological characteristics of *Gueldenstaedtia* are very similar to the genus *Tibetia* (Ali) H.P. Tsui; however, several morphological characters distinguish it from *Tibetia* by its two upper calyx lobes being free, and stipules free and adnate, as well as pollen grains that are 3-colporate, psilate, and spheroidal (Zhu 2004, 2005). In addition, *Gueldenstaedtia* and *Tibetia* were separated from each other approximately 17.23 million years ago, which is congruent with the Himalayan orogeny and the uplift of the Tibetan Plateau in the mid-Miocene (Xie et al. 2016). Traditionally, *Gueldenstaedtia* was placed in the subtribe Astragalinae of the tribe Galageae (Polhill 1981; Zhu 2004). Later, it has been transferred to the subtribe Chesneyinae of the tribe Caraganeae based on nrDNA and plastid markers (Duan et al. 2016). In recent, Duan et al (2021) upgraded monophyletic subtribe Chesneyinae (*Chesneya* Lindl. ex Endl., *Gueldenstaedtia* and *Tibetia*) to tribe Chesneyeae. In recent phylogenomic framework, monophyly of Chesneyeae is well-supported and placed sister to Caraganeae and Hedysareae based on genome skimming approach (Wang et al. 2025).

Recently, several phylogenetic studies have investigated *Gueldenstaedtia* and *Tibetia* based on partial molecular markers (Zhang et al. 2015; Xie et al. 2016; Matiz-Ceron et al. 2022) and complete chloroplast genomes (Son and Choi 2022; Byun et al. 2025). However, none of them included *G. monophylla* because of lacking genetic

data. A few studies have examined the reproductive biology, genetic variability, and conservation of this species (Selyutina et al. 2017, 2021; Baasanmunkh et al. 2022). *G. monophylla* is a perennial herb with short stem, pink flowers, orbicular or orbicular-reniform leaves, and thin root (Fig. 1). It grows on dry stony and debris slopes, rocks, and alkaline waterside pebbles in Mongolia (western part) and Russia (Altai, Tuva) (Zhu 2004; Baasanmunkh et al. 2022). In Mongolia, the conservation status of this species has been assessed as Near Threatened (Baasanmunkh et al. 2022).

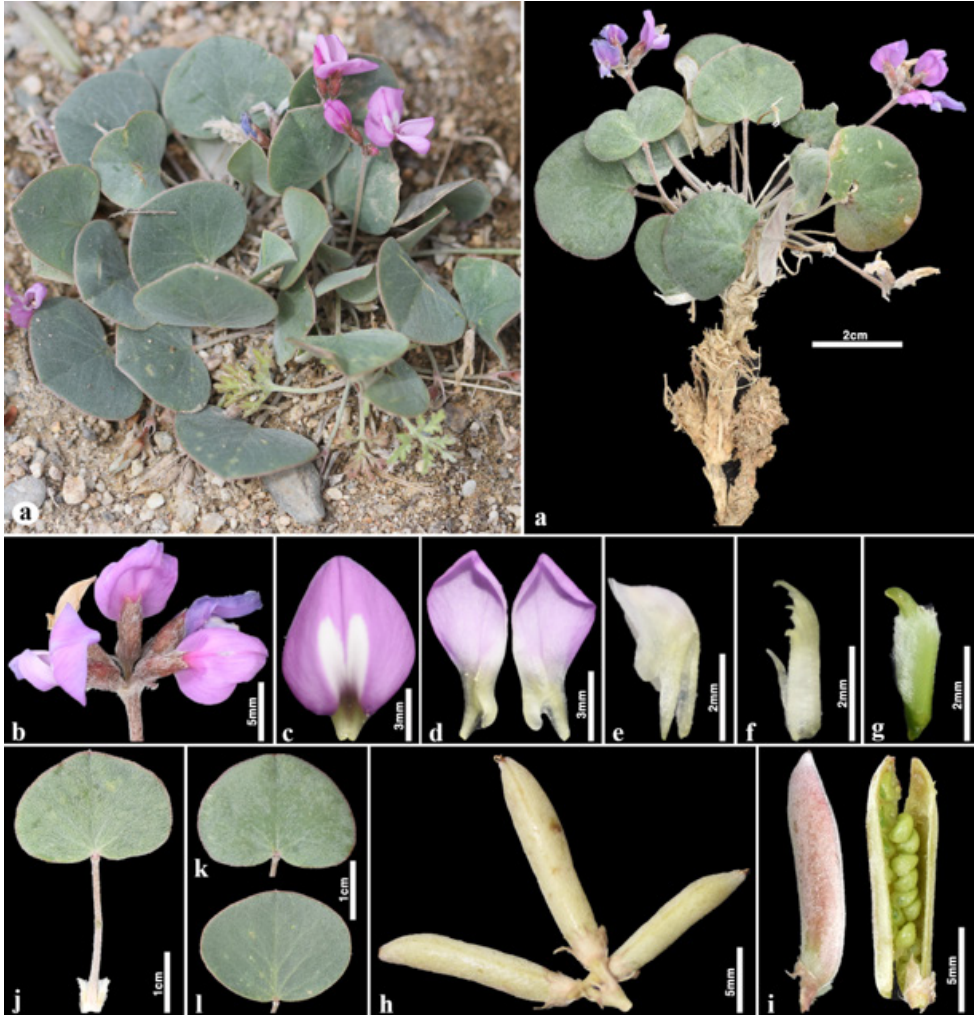


Figure 1. Photo illustration of *Gueldenstaedtia monophylla* in Mongolia. **a.** General habits; **b.** Raceme; **c.** Standard; **d.** Wings; **e.** Keel; **f.** Stamens; **g.** Pistil; **h.** Pods; **i.** Pod valve; **j.** Leaf; **k.** Leaflet, adaxial view; **l.** Leaflet, abaxial view (Photo credits: B. Oyuntsetseg, S. Baasanmunkh, D. Munkhtulga).

The chloroplast genomes (plastomes) of angiosperms have been widely used for various analyses, such as phylogeny, plastome evolution, and DNA molecular markers, because of their generally conserved structure, uniparental inheritance, and moderate substitution rates (Wang et al. 2018; Liao et al. 2021; Shanker et al. 2022; Nyamgerel et al. 2023, 2024; Oyuntsetseg et al. 2024a, 2024b; Qin et al. 2025; Baasanmunkh et al. 2026b). However, some angiosperm lineages exhibit extensive structural variation, including gene loss, intron loss, and genome rearrangements (Dobrogowski et al. 2020). The inverted repeat-lacking clade (IRLC) of legumes represents one of the most structurally dynamic groups, characterized by the loss of one copy of the inverted repeat region, frequent gene and intron losses, and numerous rearrangements such as inversions and translocations, making it an important model for studying chloroplast genome evolution (Sabir et al. 2014; Choi et al. 2019; Moghaddam et al. 2022, 2023; Seol et al. 2025). Within this context, recent plastome analyses of *Gueldenstaedtia verna* (Georgi) Boriss. revealed several notable structural features, including the loss of one IR region, complete loss of *rps16*, loss of an intron of *rpoC1* and *clpP* genes, four inversions (6–50 kb), and one repositioning (Son and Choi 2022). These findings highlight the importance of comparative plastome analyses for improving our understanding of chloroplast genome evolution and phylogenetic relationships within Fabaceae.

This study aims to: (i) characterize the plastome of *G. monophylla* and examine structural genomic features, including gene loss, intron loss, inversion, and repositioning and (ii) reconstruct the phylogenetic relationships among *Gueldenstaedtia* and the closely related genera.

Materials and methods

Plant material, DNA extraction, sequencing, assembly and annotation

The fresh leaves of *G. monophylla* were collected from natural populations in Khovd city, Khovd Province, Mongolia (48.0535° N and 91.6732° E) on 04 Jun 2025. The specimen was deposited in the Herbarium of the National University of Mongolia (contact person: Batlai Oyuntsetseg, email: oyunaa@num.edu.mn). Genomic DNA was extracted from silica gel-dried leaf material using the CTAB method (Doyle and Doyle 1987). A whole-genome sequencing library was constructed from total genomic DNA using the TruSeq DNA Nano Kit protocol and sequenced on a Next-Seq 500 platform (Illumina, USA). Raw paired-end reads were quality filtered by FastQC v.0.11 (Andrews 2010), adapter trimmed by fastp (Chen et al. 2018) with default settings, and reads were assembled by GetOrganelle v1.7.5 (Jin et al. 2020) with the default k-mer settings (e.g., 21, 45, 65, 85, 105, 127) as implemented in the Canonical Ubuntu 2026.04 operating system (<https://ubuntu.com>). To validate the assembly, the clean reads were mapped back to the assembled circular chloroplast genome using BWA-MEM, and the mapping depth was assessed with SAMtools (Li

and Durbin 2009). The assembly and orientation were confirmed using the BLAST tool of NCBI (<https://www.ncbi.nlm.nih.gov/>), and graphical views were checked using Geneious Prime® 2026.0.2 (<https://www.geneious.com>). Chloroplast genome annotations were performed using Geseq Annotation of Organellar Genomes (Tillich et al. 2017) to predict the location of genes, and tRNAs were identified using tRNAscan-SE (Lowe and Chan 2016). A circular chloroplast genomic map was visualized using the Chloroplot online software (<https://irscope.shinyapps.io/Chloroplot/>) (Zheng et al. 2020).

Comparative chloroplast genome analysis

Chloroplast genome rearrangements of all available species within Chesneyeae were identified using collinear blocks in Mauve v.2.4.0 (Darling et al. 2010) with *Caragana* Fabr. from Caraganeae as reference. To compare relative synonymous codon usage (RSCU) of Chesneyeae species, the consensus exons of the protein-coding were separately extracted and screened by MEGA11 (Tamura et al. 2021). RSCU <1 indicated a codon used less frequently than expected, whereas RSCU >1 indicated a codon used more frequently than expected. The codon usage distribution of Chesneyeae species was visualized using the Heatmapper tool with average linkage clustering and Euclidean distance methods (Babicki et al. 2016). Selective pressure analysis was conducted between *Gueldenstaedtia* species and measured by the nonsynonymous (Ka) and synonymous (Ks) substitution in consensus exons using KaKs_Calculator 2.0 (Wang et al. 2010) with calculation mode NY and genetic code 11 (bacterial and plant plastid code). Genes with Ka/Ks ratios greater than one were considered to be under positive selection, whereas those with ratios less than one were considered to be under purifying selection. Statistical significance was assessed using Fisher’s exact test, with $P < 0.05$ considered statistically significant.

Table 1. Pairwise comparison of synonymous and nonsynonymous differences and Ka/Ks ratios among *Gueldenstaedtia* species

Compared species	Genes	Synonymous differences (bp)	Nonsynonymous differences (bp)	Ka/Ks
<i>G. monophylla</i>	<i>ndhI</i>	8.00	5.00	0.18
vs	<i>petB</i>	5.00	0.00	0.00
<i>G. guangxiensis</i>	<i>rps11</i>	1.50	10.50	2.33
	<i>clpP</i>	5.25	57.75	3.37
<i>G. monophylla</i>	<i>ndhI</i>	9.00	5.00	0.16
vs	<i>petB</i>	4.00	0.00	0.00
<i>G. verna_OP</i>	<i>rps11</i>	1.50	10.50	2.33
	<i>clpP</i>	5.25	57.75	3.37
<i>G. monophylla</i>	<i>ndhI</i>	9.00	5.00	0.16
vs	<i>petB</i>	4.00	0.00	0.00
<i>G. verna_PQ</i>	<i>rps11</i>	1.50	10.50	2.33
	<i>clpP</i>	5.25	57.75	3.37

Compared species	Genes	Synonymous differences (bp)	Nonsynonymous differences (bp)	Ka/Ks
<i>G. guangxiensis</i>	<i>ndhI</i>	1.00	0.00	0.00
vs	<i>petB</i>	1.00	0.00	0.00
<i>G. verna_OP</i>	<i>rps11</i>	0.00	0.00	0.00
	<i>clpP</i>	0.00	0.00	0.00
<i>G. guangxiensis</i>	<i>ndhI</i>	1.00	0.00	0.00
vs	<i>petB</i>	1.00	0.00	0.00
<i>G. verna_PQ</i>	<i>rps11</i>	0.00	0.00	0.00
	<i>clpP</i>	0.00	0.00	0.00
<i>G. verna_OP</i>	<i>ndhI</i>	0.00	0.00	0.00
vs	<i>petB</i>	0.00	0.00	0.00
<i>G. verna_PQ</i>	<i>rps11</i>	0.00	0.00	0.00
	<i>clpP</i>	0.00	0.00	0.00

Phylogenetic analysis and divergence time estimation

For the phylogenetic analysis, protein-coding gene sequences of four *Gueldenstaedtia* and 26 species belonging to other genera of the Fabaceae were obtained from the GenBank database based on previous studies (Zhao et al. 2021; Son and Choi 2022; Byun et al. 2025) (Suppl. material 2: Table S1). The protein-coding sequences of the genomes were concatenated and aligned using MAFFT v.7.388 (Katoh and Standley 2013) as implemented in Geneious Prime® 2024.0.5 (<http://www.geneious.com>). The alignment was filtered to remove ambiguously aligned regions using GBlocks v.0.91.1 (Talavera and Castresana 2007). The best-fitting model for nucleotide substitutions was determined using the Akaike Information Criterion in jModelTest v.2.1.10 (Darriba et al. 2012). Maximum likelihood (ML) analysis was performed using RaxML v.8.2.11 (Stamatakis 2014) with 1000 bootstrap replicates under the GTR+G+I model. Bayesian Inference (BI) analysis was performed using MrBayes v.3.2.2 (Ronquist et al. 2012), with two independent runs of four simultaneous chains, executed for 5,000,000 generations using the Markov chain Monte Carlo algorithm. Trees were sampled every 5000 generations, and the first 25% were discarded as burn-in. The trees were summarized using a 50% majority-rule consensus to estimate posterior probabilities (PPs). The GTR+G model was selected for BI analysis. The resulting phylogenetic trees were visualized using FigTree v.1.4.2 (Rambaut 2012).

Divergence time was estimated using BEAST v.2.7.8 (Bouckaert et al. 2019) with 69 protein-coding genes (PCGs) from nine species of Chesneyeae and 15 additional species from closely related genera within Fabaceae. A GTR replacement model was implemented with gamma distributions across four rate categories and estimated using a Yule Prior. For divergence time analysis, three calibration points were established based on a previous study (Zhao et al. 2021): four-tribe crown at 44.25 (40.31–49.23) million years ago (MYA), three-tribe crown at 39.46 (35.07–44.24) MYA, and Astragaleae crown at 20.92 (19.13–22.92) MYA. MCMC chains were run

for 100 million generations, and sampled every 10,000 generations. The effective sample size (ESS) was evaluated using Tracer v.1.7.2 (Rambaut et al. 2018), with all parameters having ESS values greater than 200, indicating sufficient sampling. The first 10% of trees were discarded as burn-in, and maximum clade credibility trees were generated using TreeAnnotator v.2.7.6 in the BEAST package. These were visualized using Figtree (Rambaut 2012), and the results were summarized using the mean height and 95% high posterior densities (HPDs).

Results

Chloroplast genome feature of *Gueldenstaedtia monophylla*

A total of 97,835,743 paired-end reads were generated, of which 97,784,497 high-quality clean reads remained after quality filtering. De novo assembly generated a single contig with 100% coverage, and the minimum and average read mapping depth of the assembled genome were 548× and 13,867.81×, respectively (Suppl. material 1: Fig. S1). The plastome of *G. monophylla* was 118,980 bp in length and exhibited a single inverted repeat (IR) region, representing an IRLC-type plastome structure (Fig. 2). The plastid genome contained 110 genes, which comprised 76 protein-coding genes (PCGs), 30 tRNA genes, and four rRNA genes (Suppl. material 2: Table S2). Eight genes (*atpF*, *pafI*, *ndhA*, *ndhB*, *petB*, *petD*, *rpl16*, *rpl2*, and *rps16*) contained one intron and they were cis-spliced (Suppl. material 1: Fig. S2A). The *rps12* gene was a trans-spliced gene, with the 5'-end located in the LSC and the 3'-end located in an IR region (Suppl. material 1: Fig. S2B).

Comparative genome analysis within Chesneyeae

Collinearity in gene placement among the Chesneyeae revealed large-scale inversions and translocations (Fig. 3). Among the tribe, *Tibetia* showed strong similarity with Caragana, indicating that genome have remained stable during the evolution. Meanwhile, *Chesneya* had ~15 kb genus-specific inversion from the *trnN* to *ndhB* (E). In contrast, *Gueldenstaedtia* revealed several large-scale inversions and translocations: *trnK-psbK* (A, ~50 kb), *rps15-trnL* (C, ~10 kb), and *ycf2* (D) inversions in all speceis; and *accD-rpl16* (B, ~28 kb) and *trnL-rps15* (G, ~6 kb) inversions in *G. verna* and *G. guangxiensis*; D, C, F (*ycf1*), and G blocks translocated in all species.

The RSCU was calculated on the concatenate of 68 core protein-coding (CDS) genes of the Chesneyeae species. These CDSs contained 19,431–19,648 codons in most species, whereas *G. monophylla* contained 18,369 codons due to the partial loss of the *ycf1* gene. Among the amino acids, leucine (10.75%) was the most frequently detected, whereas cysteine (1.18%) was the least ubiquitous. All species exhibited similar patterns in 64 codons, with high RSCU values for AGA (arginine), TTA (leucine), and GCT (alanine) (Fig. 4). Codons with A or T in the third position had

a strong codon bias. Among the three stop codons, TGA was the most commonly detected. Based on hierarchical clustering of RSCU pattern, *Chesneya*, *Gueldenstaedtia*, and *Tibetia* species formed separate clusters. Within the *Gueldenstaedtia*, the two accessions of *G. verna* were clustered together which was closely related to *G. guangxiensis*. While *G. monophylla* displayed a distinct position within the genus. Overall, only minor variation in codon usage bias was observed among species, indicating strong conservation of plastid codon preferences within *Chesneyeae*.

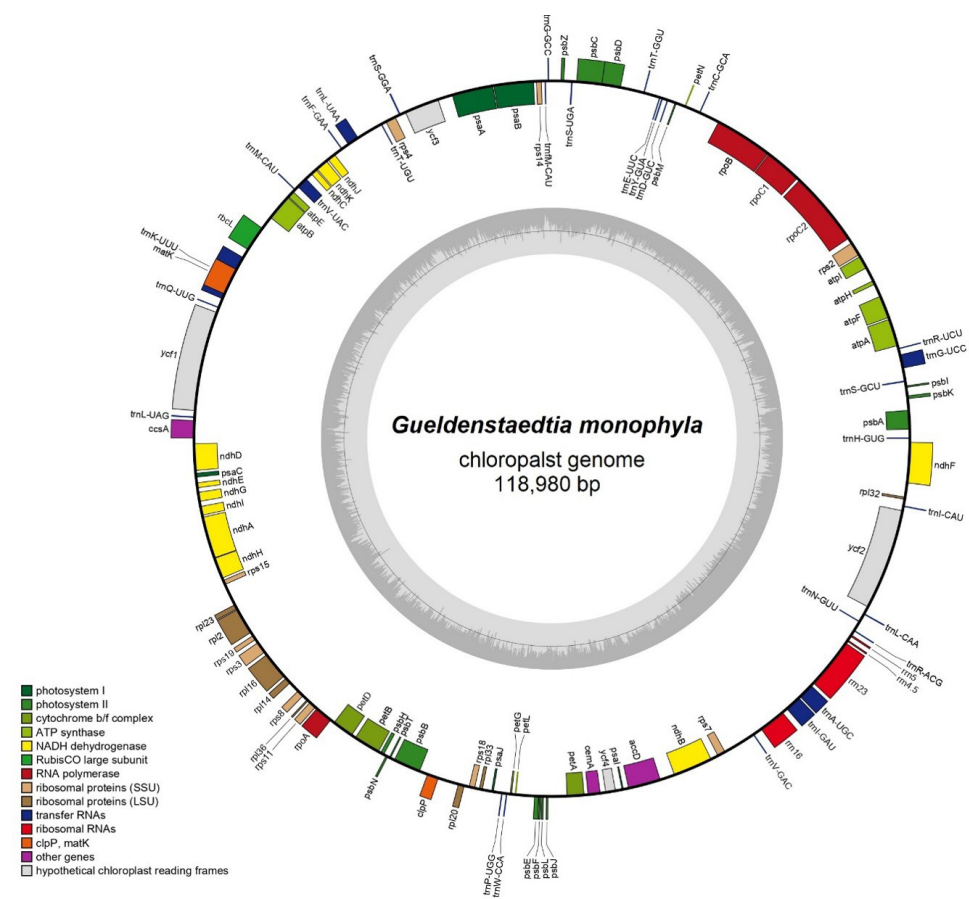


Figure 2. Circular gene map of the complete chloroplast genome of *Gueldenstaedtia monophylla*. Genes drawn inside the circle are transcribed clockwise, and those outside, counter-clockwise. The function of genes is color-coded. The darker gray in the inner circle represents GC content and light gray represents AT content.

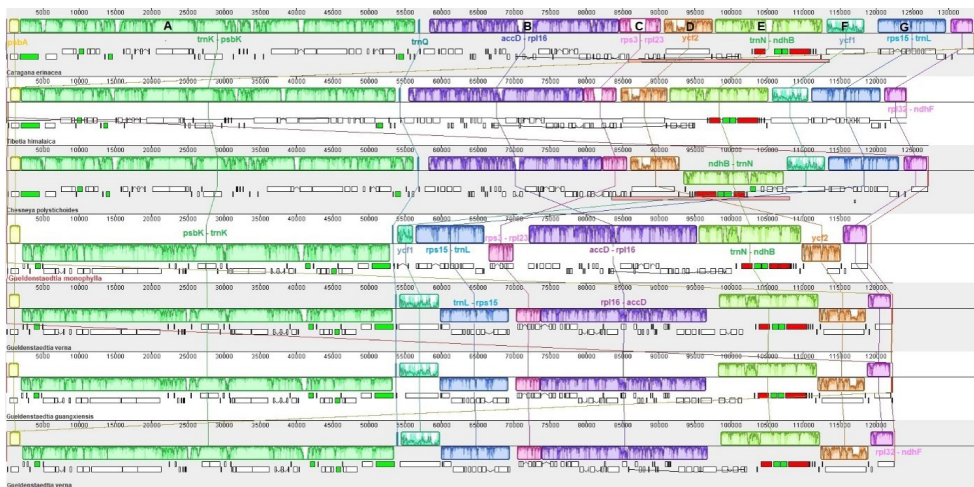


Figure 3. MAUVE alignment of the chloroplast genome of *Gueldenstaedtia* species. The blocks in the top row are in the same orientation, whereas those in the bottom row are in inverse orientation. Boxes in each genome map represent protein-coding genes (white), rRNAs (red), and tRNAs (green).

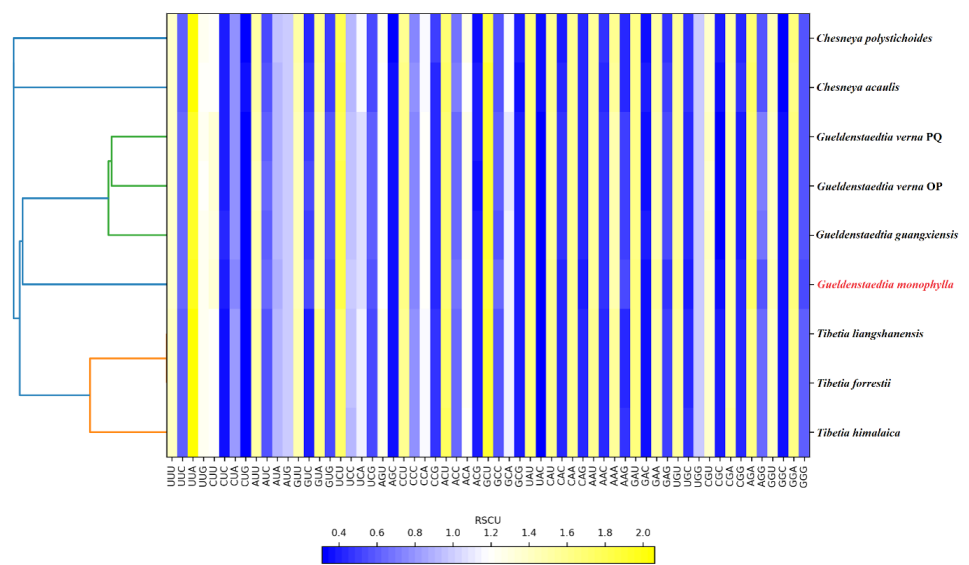


Figure 4. The heatmap of Relative Synonymous Codon Usage (RSCU) values for species of the tribe Chesneyae (*Chesneya*, *Gueldenstaedtia*, and *Tibetia*). The dendrogram on the left clusters species according to similarities in codon usage patterns, while the heatmap colors represent RSCU values from underrepresented (blue) to overrepresented (yellow) codons.

The Ka/Ks analysis revealed substantial variation in selective pressure among protein-coding genes across the three *Gueldenstaedtia* species, with most genes showing purifying selection (Fig. 5). Several genes exhibited a Ka/Ks ratio of 0, including *atpH*, *ndhJ*, *petL*, *petN*, *psaI*, *psbE*, *psbF*, *psbH*, *psbK*, *psbM*, *rpl20*, *rps12*, and *rps18*, reflecting the absence of detected nonsynonymous substitutions and indicating strong conservation at the protein level. However, few genes showed the strongest signal of positive selection ($Ka/Ks > 1$) between *G. monophylla* and others, particularly *ndhI*, *petB*, *rps11*, and *clpP* (Table 1). As expected, synonymous and nonsynonymous differences of positively selected genes were highly pronounced between *G. monophylla* and others.

Phylogenetic analysis and divergence time estimation

To identify the phylogenetic position of *G. monophylla*, 68 core protein-coding genes were extracted from the chloroplast genome of Fabaceae species and aligned for phylogenetic construction. The final dataset comprised 55,339 bp, of which 42,556 bp (76.9%) were constant and 3,043 bp (5.5%) were parsimony-informative. Three *Cicer* species were used as outgroups. The phylogenetic analysis recovered all sampled species of Chesneyeae as a strongly supported monophyletic clade (BI=1, BS=100) (Suppl. material 1: Fig. S3). Within the tribe, *Gueldenstaedtia* and *Tibetia* placed as a sister groups that together were sister to *Chesneya* (BI=1, BS=100). Within *Gueldenstaedtia*, *G. monophylla* formed a well-supported lineage sister to the clade comprising *G. quangxiensis* and *G. verna* (BI=1, BS=100). The close relationship between *G. quangxiensis* and *G. verna* was strongly supported, whereas *G. monophylla* occupied a distinct phylogenetic position as the earliest-diverging extant lineage within the genus.

Divergence time estimation indicated that the major lineages of the IRLC originated during the Eocene and Oligocene (Fig. 6). The split between the Cicerinae lineage (*Cicer*) and the remaining taxa was estimated at approximately 44.05 MYA. Subsequently, Astragaleae diverged from the lineage containing Chesneyeae, Caraganeae, and Hedysareae around 39.70 MYA. The separation of Hedysareae from the Caraganeae-Chesneyeae clade was estimated at 35.07 MYA, suggesting an Oligocene origin for these tribes. The divergence between the tribes Caraganeae and Chesneyeae occurred approximately 33.37 MYA. The genus *Chesneya* diverged from the lineage comprising *Gueldenstaedtia* and *Tibetia* around 27.51 MYA, whereas *Gueldenstaedtia* and *Tibetia* separated at approximately 21.09 MYA. Within *Gueldenstaedtia*, *G. monophylla* diverged from the common ancestor of *G. quangxiensis* and *G. verna* around 7.21 MYA. The divergence between *G. quangxiensis* and *G. verna* was estimated at approximately 0.65 MYA. *G. monophylla* diverged earlier than other species of the genus and may preserve ancestral genomic characteristics.

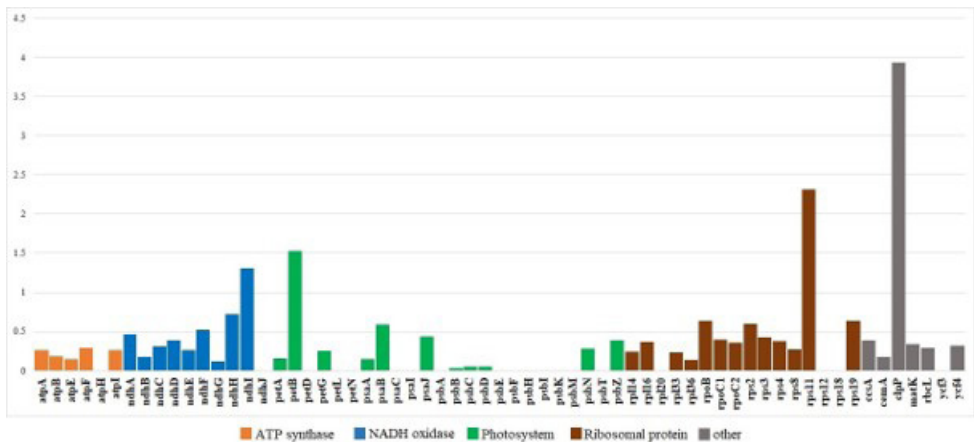


Figure 5. Nonsynonymous/synonymous substitution (K_a/K_s) ratios of protein-coding genes across four *Gueldenstaedtia* species. The genes are grouped by functional groups: ATP synthase, NADH oxidase, Ribosomal protein, and other genes shown in orange, blue, green, brown, and gray color, respectively. Genes with $P > 0.05$ were excluded.

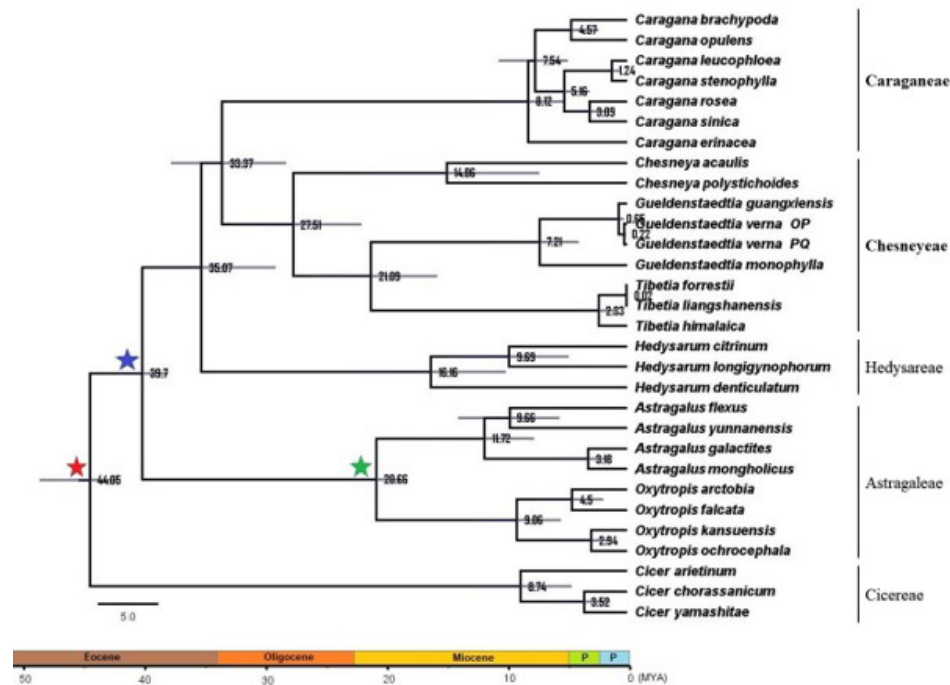


Figure 6. Dated phylogeny of *Gueldenstaedtia* and its related species based on a maximum clade credibility tree obtained from BEAST analysis. Mean divergence time estimates are shown with 95% highest posterior density (HPD; blue bars).

Discussion

Chloroplast genome feature of *Gueldenstaedtia*

The complete chloroplast genome size, structure, and gene content are highly conserved in angiosperms (Dobrogojski et al. 2020). However, rare genome characteristics, such as gene loss, inversion, and IR loss, have been reported in the IRLC legumes of Fabaceae (Moghaddam et al. 2022, 2023; Seol et al. 2025). *Gueldenstaedtia* is one of the most structurally rearranged chloroplast genomes, characterized by multiple inversions, translocations, and partial gene loss (Son and Choi 2022). In this study, the chloroplast genome of *G. monophylla* exhibited distinct structural variation compared with previously reported species within the tribe Chesneyeae.

Gene and intron losses are the most distinctive evolutionary features of the IRLC legumes (Lei et al. 2016; Jo et al. 2022; Cauz-Santos 2025), including *rps16*, *rpl22*, *infA*, *accD*, and *ycf4*, as well as the loss of introns from *clpP*, *rpoC1*, and *atpF* genes. These events have been documented across multiple IRLC legumes, including *Astragalus* L., *Oxytropis* DC., *Vicia* L., *Onobrychis* Mill., and others (Jo et al. 2022; Moghaddam et al. 2022, 2023; Son and Choi 2022; Cui et al. 2024; Liu et al. 2024; Li et al. 2025; Seol et al. 2025; Munkhtulga et al. 2026). Among these changes, the loss of *rps16* and *rpl22* appears to be particularly widespread and may be associated with functional transfer of these genes to the nuclear genome (Lei et al. 2016). Furthermore, the recurrent loss of the intron of *rps12* and *clpP* has been recognized as a common feature of many IRLC and is considered an important evolutionary marker within the group (Jansen et al. 2008). Such as IRLC legumes, *Gueldenstaedtia* have completely lost *rpl22* and *rps16* genes, as well as the introns of *clpP* and *rps12* (Son and Choi 2022; Byun et al. 2025). These gene and intron loss were appeared within the tribe, including *Chesneya* (Nie et al. 2021) and *Tibetia* (Guo and Wariss 2021).

The chloroplast genome size of tribe Chesneyeae showed a decreasing trend, from 128–127 kb in *Chesneya* species to 124–123 kb in *Tibetia* and approximately 122 kb in *Gueldenstaedtia* (*G. guangxiensis* and *G. verna*), whereas *G. monophylla* has the smallest genome size at 118,980 bp. This gradual reduction in chloroplast genome size across Chesneyeae suggests lineage-specific genome contraction during the evolutionary diversification of the tribe. The *ycf1* gene of *G. monophylla* has a relatively smaller size (1,521 bp) compared with all examined species (5,094–5,466 bp). The *ycf1* gene is among the largest and most rapidly evolving plastid genes and frequently serves as a hotspot for structural variation in angiosperms (Dong et al. 2015). The association between rearrangement breakpoints and partial gene loss suggests that large-scale inversion events may have contributed directly to chloroplast genome reorganization in *Gueldenstaedtia*.

Evolutionary distinctiveness of *Gueldenstaedtia monophylla*

Phylogenetic analyses strongly support the position of *Gueldenstaedtia* within tribe Chesneyeae and confirm a sister relationship between *Gueldenstaedtia* and *Tibetia*, consistent with previous molecular studies based on ITS and plastid markers (Zhang et al. 2015; Xie et al. 2016; Zhao et al. 2021; Son and Choi 2022; Byun et al. 2025). However, the inclusion of *G. monophylla* for the first time provides a more complete understanding of evolutionary relationships within the genus. Within *Gueldenstaedtia*, *G. monophylla* occupies the earliest-diverging position and separated from the common ancestor of *G. guangxiensis* and *G. verna* approximately 7.21 MYA (Fig. 6). This divergence is noticeable longer than the recent split between *G. guangxiensis* and *G. verna* (0.65 MYA), suggesting independent evolutionary history for *G. monophylla*.

Gueldenstaedtia chloroplast genome has large-scale inversions and translocation within tribe Chesneyeae (Fig. 3). After the ancestor of *Gueldenstaedtia* diverged from *Chesneya* and *Tibetia*, structural changes occurred within its chloroplasts. These included ~50 kb inversion of *trnK-psbK* (A), ~7 kb translocation of *ycf1-rps15* (F-G), and ~10 kb inversion and translocation of *rps15-trnL* (C) and *ycf2* (D). The ancestor of *Gueldenstaedtia* diverged into two lines. *G. verna* and *G. guangxiensis* belonged to the same line and first evolutionary change of this line was represented by the *accD-rpl16* (B) and *trnL-rps15* (G) inversions, which were not detected in the chloroplast genome of *G. monophylla*. It may suggest an ancestral condition retained within *G. monophylla*, whereas the alternative arrangement likely originated later during the evolution of the genus.

In addition, codon usage patterns within tribe Chesneyeae were relatively consistent in contrast to the structural rearrangements. Similar codon preferences were observed across all taxa, with a strong bias toward A/T-ending codons, due to the AT-rich feature of chloroplast genomes (Dobrogojski et al. 2020). However, significant differences in codon usage were revealed between the species, and RSCU-based clustering closely mirrored phylogenetic positions. In particular, *G. verna* and *G. guangxiensis* formed a closely related group, whereas *G. monophylla* formed a distinct position. These species-specific differences in codon usage and nucleotide variation provide additional support for the phylogenetic relationships among *Gueldenstaedtia* species.

Selective pressure analysis revealed positive selection in *G. monophylla* relative to other species, particularly in *ndhI*, *petB*, *rps11*, and *clpP*, genes involved in essential metabolic processes. Positively selected genes can play a role in adaptive evolution (Yang and Nielsen 2000). When plants grow in harsh environments, specific sites of photosystem-related genes tend to exhibit positive selection signals, potentially pointing to adaptations for improved photosynthetic efficiency and stress tolerance (Gao et al. 2019; Geng et al. 2026). *G. monophylla* is adapted to xeric, rocky-alkaline habitats, in contrast to *G. verna* and *G. guangxiensis*, which grow in meadows and on shrubby hillslopes (Zhu 2004; Stupnikova 2018; Baasanmunkh

et al. 2022). Four positively selected genes in *G. monophylla* have a functional role that plausibly connects to this ecological contrast. *ndhI* encodes a subunit of the NADH dehydrogenase, which drives cyclic electron flow around photosystem I and contributes to photoprotection under high light, drought, and temperature stress (Sirpio et al. 2009). Higher Ka/Ks ratio in *ndhI* is frequently reported in plants from arid or high-altitude habitats (Chen et al. 2022; Sabater 2021). *petB* is a core subunit of the cytochrome b6/f complex, central to the photosynthetic electron transport chain (Stoppel et al. 2011). Similarly, adaptive substitutions here could reflect selection to maintain photosynthetic efficiency under fluctuating light and temperature (Ghotbi and Soorni 2026; Liu et al. 2023; Zhang et al. 2024). Ribosomal protein *rps11* has been linked to adjusting protein synthesis rates under abiotic stress (Robles and Quesada 2022). In parallel, *clpP* encodes a subunit of the ATP-dependent protease, responsible for degrading damaged/misfolded chloroplast proteins (Shikanai et al. 2001). This is one of the more commonly reported stress adaptation plastid genes, since efficient protein turnover helps sustain photosynthesis under oxidative and thermal stress (Erixon and Oxelman 2008). The functional profile of these genes suggests that positive selection in *G. monophylla* may be linked to physiological adaptation to its harsher habitat. Future studies incorporating experimental validation, such as physiological or transcriptomic analyses under stress conditions, will be needed to confirm this link.

The congruence among phylogenetic relationships, chloroplast genome rearrangements, codon usage clustering, and selective pressure analysis provides independent lines of evidence supporting the evolutionary distinctiveness of *G. monophylla*. As the earliest lineage within the genus, *G. monophylla* may retain ancestral genomic characteristics and therefore represents a key taxon for understanding the origin and diversification of *Gueldenstaedtia*. Future studies incorporating the remaining species (*G. henryi*) and nuclear genomic data will help clarify the origin of the observed rearrangements and the evolutionary history of *Gueldenstaedtia*.

Conclusion

This study presents the first complete chloroplast genome of *Gueldenstaedtia monophylla* and places it in a comparative framework with related genera of tribe Chesneyae. The genome retains the typical gene content of IRLC legumes but shows a distinct, smaller architecture than other species within the genus, driven largely by partial *ycf1* loss. The evidence from phylogenetic, structural, codon-usage, and selective pressure analysis converges on the same conclusion: *G. monophylla* is the earliest-diverging lineage within the genus and retains genomic features that are ancestral relative to *G. verna* and *G. guangxiensis*. These combined lines of evidence indicate that *G. monophylla* occupies a distinct evolutionary position within *Gueldenstaedtia* and offers a useful reference point for reconstructing the ancestral plastome architecture of the genus. Future work incorporating *G. henryi* and nu-

clear genomic data will help clarify when and how the structural rearrangements identified here arose, and will further resolve the evolutionary history of *Gueldenstaedtia* within tribe Chesneyeae.

Data availability

The complete chloroplast genomes of *Gueldenstaedtia monophylla* was submitted to GenBank under accession numbers PZ504763 (<https://www.ncbi.nlm.nih.gov/nucleotide/PZ504763.1/>).

Acknowledgments

This work was supported by the National University of Mongolia (Grant No. P2025-5052), the Korea National Arboretum (Grant No. KNA1-2-42, 22-2) and the Korea Basic Science Institute (National Facilities and Equipment Center) grant funded by the Ministry of Education (Grant No. 2023R1A6C101B022).

References

- Andrews S (2010) FastQC: A quality control tool for high throughput sequence data. Babraham Bioinformatics. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Baasanmunkh S, Oyuntsetseg B, Tsegmed Z, Illarionova ID, Nyambayar N, Choi HJ (2022) Taxonomic notes and distribution of *Gueldenstaedtia* (Chesneyinae, Fabaceae) in Mongolia. Korean Journal of Plant Taxonomy 52: 64–70. <https://doi.org/10.11110/kjpt.2022.52.1.64>
- Baasanmunkh S, Kechaykin A, Munkhtulga D, Oyuntsetseg B, Shmakov A, Choi HJ (2026a) Updates on the checklist of native vascular plants of Mongolia. Journal of Asia-Pacific Biodiversity. <https://doi.org/10.1016/j.japb.2026.03.006>
- Baasanmunkh S, Tsegmed Z, Nyamgerel N, Munkhtulga D, Oyuntsetseg B, Park BK, Chen S, Sun H, Choi HJ (2026b) Integrative study of the rare *Incarvillea potaninii* (Bignoniaceae) in Mongolia: conservation, comparative plastome, distribution modelling, phylogeny, and taxonomic insights. Scientific Reports 16: 14814. <https://doi.org/10.1038/s41598-026-45263-0>
- Babicki S, Arndt D, Marcu A, Liang Y, Grant JR, Maciejewski A, Wishart DS (2016) Heatmapper: web-enabled heat mapping for all. Nucleic Acids Research 44: W147–W153. <https://doi.org/10.1093/nar/gkw419>
- Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Kühnert D, De Maio N, Matschiner M, Mendes FK, Müller NF, Ogilvie HA, du Plessis L, Poppinga A, Rambaut A, Rasmussen D, Siveroni I, Suchard MA, Wu Ch, Xie D, Zhang Ch, Stadler T, Drummond AJ (2019) BEAST 2.5: An advanced

- software platform for Bayesian evolutionary analysis. PLoS Computational Biology 15: 1–28. <https://doi.org/10.1371/journal.pcbi.1006650>
- Byun HJ, Kim S-J, Seol B-Y, Choi KS, Choi I-S (2025) Distribution and genetic diversity of *Gueldenstaedtia verna* in Korea. Korean Journal of Plant Taxonomy 55(2): 103–113. <https://doi.org/10.11110/kjpt.2025.55.2.103>
- Cauz-Santos LA (2025) Beyond conservation: the landscape of chloroplast genome rearrangements in angiosperms. New Phytologist 247(6): 2571–2580. <https://doi.org/10.1111/nph.70364>
- Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34: i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
- Chen Z, Yu X, Yang Y, Wei P, Zhang W, Li X, Liu C, Zhao S, Li X, Liu X (2022) Comparative analysis of chloroplast genomes within *Saxifraga* (Saxifragaceae) takes insights into their genomic evolution and adaption to the high-elevation environment. Genes 13(9): 1673. <https://doi.org/10.3390/genes13091673>
- Choi IS, Jansen RK, Ruhlman T (2019) Lost and found: return of the inverted repeat in the legume clade defined by its absence. Genome Biology and Evolution 11(4): 1321–1333. <https://doi.org/10.1093/gbe/evz076>
- Cui X, Liu K, Li E, Zhang Z, Dong W (2024) Chloroplast genomes evolution and phylogenetic relationships of Caragana species. International Journal of Molecular Sciences 25(12): 6786. <https://doi.org/10.3390/ijms25126786>
- Darling AE, Mau B, Perna NT (2010) progressiveMauve: Multiple genome alignment with gene gain, loss and rearrangement. PLoS ONE 5: e11147. <https://doi.org/10.1371/journal.pone.0011147>
- Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. Nature Methods 9: 772–772. <https://doi.org/10.1038/nmeth.2109>
- Dobrogojski J, Adamiec M, Luciński R (2020) The chloroplast genome: a review. Acta Physiologiae Plantarum 42: 98. <https://doi.org/10.1007/s11738-020-03089-x>
- Dong W, Xu C, Li C, Sun J, Zuo Y, Shi S, Cheng T, Guo J, Zhou S (2015) *ycf1*, the most promising plastid DNA barcode of land plants. Scientific Reports 5(1): 8348. <https://doi.org/10.1038/srep08348>
- Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19: 11–15.
- Duan L, Harris AJ, Su C, Zhang Z, Arslan E, Ertuğrul K, Loc PK, Hayashi H, Wen J, Chang Z (2016) A molecular phylogeny of Caraganeae (Leguminosae, Papilionoideae) reveals insights into new generic and infrageneric delimitations. PhytoKeys 70: 111–137. <https://doi.org/10.3897/phytokeys.70.9641>
- Duan L, Li SJ, Su C, Sirichamorn Y, Han LN, Ye W, Loc PK, Wen J, Compton JA, Schrire B, Nie ZL (2021) Phylogenomic framework of the IRLC legumes (Leguminosae subfamily Papilionoideae) and intercontinental biogeography of tribe Wisterieae. Molecular Phylogenetics and Evolution 163: 107235. <https://doi.org/10.1016/j.ympev.2021.107235>

- Erixon P, Oxelman B (2008) Whole-gene positive selection, elevated synonymous substitution rates, duplication, and indel evolution of the chloroplast *clpP1* gene. *PloS One* 3(1): e1386. <https://doi.org/10.1371/journal.pone.0001386>
- Gao C, Deng Y, Wang J (2019) The complete chloroplast genomes of *Echinacanthus* species (Acanthaceae): phylogenetic relationships, adaptive evolution, and screening of molecular markers. *Frontiers in Plant Science* 9: 1989. <https://doi.org/10.3389/fpls.2018.01989>
- Geng X, Xue Y, Wang H, Tu H, Jing R, Li Z, Gai Y (2026) Comprehensive analysis of chloroplast genome evolution in Poaceae: codon usage patterns, selection pressures, and phylogenomic relationships. *BMC Genomics* 27: 217. <https://doi.org/10.1186/s12864-025-12498-6>
- Ghotbi FS, Soorni A (2026) Chloroplast phylogenomics and barcode discovery in medicinal stachys species reveal evolutionary relationships and adaptive signatures. *Ecology and Evolution* 16(5): e73618. <https://doi.org/10.1002/ece3.73618>
- Guo Y, Wariss HM (2021) The complete chloroplast genome of *Tibetia liangshanensis* PC Li (Leguminosae: Papilionoideae), an endemic species of China. *Mitochondrial DNA Part B* 6(7): 1917–1918. <https://doi.org/10.1080/23802359.2021.1935342>
- Jansen RK, Wojciechowski MF, Sanniyasi E, Lee SB, Daniell H (2008) Complete plastid genome sequence of the chickpea (*Cicer arietinum*) and the phylogenetic distribution of *rps12* and *clpP* intron losses among legumes (Leguminosae). *Molecular Phylogenetics and Evolution* 48(3): 1204–1217. <https://doi.org/10.1016/j.ympev.2008.06.013>
- Jin JJ, Yu WB, Yang JB, Song Yu, dePamphilis CW, Yi T-Sh, Li D-Zh (2020) GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biology* 21(1): 241. <https://doi.org/10.1186/s13059-020-02154-5>
- Jo IH, Han S, Shim D, Ryu H, Hyun TK, Lee Y, Kim D, So Y-S, Chung J-W (2022) Complete chloroplast genome of the inverted repeat-lacking species *Vicia bungei* and development of polymorphic simple sequence repeat markers. *Frontiers in Plant Science* 13: 891783. <https://doi.org/10.3389/fpls.2022.891783>
- Katoh K, Standley DM (2013) MAFFT Multiple sequence alignment software v.7: Improvements in performance and usability. *Molecular Biology and Evolution* 30: 772–780. <https://doi.org/10.1093/molbev/mst010>
- Lei W, Ni D, Wang Y, Shao J, Wang X, Yang D, Wang J, Chen H, Liu Ch (2016) Intraspecific and heteroplasmic variations, gene losses and inversions in the chloroplast genome of *Astragalus membranaceus*. *Scientific Report* 6(1): 21669. <https://doi.org/10.1038/srep21669>
- Liao M, Dong Z, Qin R, Wang Y, Yang H, Liu H (2021) Comparative chloroplast genomics of *Sophora* species: evolution and phylogenetic relationships in the early-diverging legume subfamily Papilionoideae (Fabaceae). *Frontiers in Plant Science* 12: 778933. <https://doi.org/10.3389/fpls.2021.778933>
- Liu L, Li H, Li J, Li X, Hu N, Sun J, Zhou W (2024) Chloroplast genomes of *Caragana tibetica* and *Caragana turkestanica*: structures and comparative analysis. *BMC Plant Biology* 24(1): 254. <https://doi.org/10.1186/s12870-024-04979-9>

- Liu TJ, Zhang SY, Wei L, Lin W, Yan HF, Hao G, Ge XJ (2023) Plastome evolution and phylogenomic insights into the evolution of *Lysimachia* (Primulaceae: Myrsinoideae). *BMC Plant Biology* 23: 359. <https://doi.org/10.1186/s12870-023-04363-z>
- Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25(14): 1754–60. <https://doi.org/10.1093/bioinformatics/btp324>
- Li X, Liu L, Ren Q, Zhang T, Hu N, Sun J, Zhou W (2025) Analysis of synonymous codon usage bias in the chloroplast genome of five *Caragana*. *BMC Plant Biology* 25(1): 322. <https://doi.org/10.1186/s12870-025-06351-x>
- Lowe TM, Chan PP (2016) tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Research* 44: W54–W57. <https://doi.org/10.1093/nar/gkw413>
- Matiz-Ceron L, Reyes A, Anzola J (2022) Taxonomical evaluation of plant chloroplastic markers by bayesian classifier. *Frontiers in Plant Science* 12: 782663. <https://doi.org/10.3389/fpls.2021.782663>
- Moghaddam M, Ohta A, Shimizu M, Terauchi R, Kazempour-Osaloo S (2022) The complete chloroplast genome of *Onobrychis gaubae* (Fabaceae-Papilionoideae): comparative analysis with related IR-lacking clade species. *BMC Plant Biology* 22: 75. <https://doi.org/10.1186/s12870-022-03465-4>
- Moghaddam M, Wojciechowski MF, Kazempour-Osaloo S (2023) Characterization and comparative analysis of the complete plastid genomes of four *Astragalus* species. *PloS ONE* 18(5): e0286083. <https://doi.org/10.1371/journal.pone.0286083>
- Munkhtulga D, Baasanmunkh S, Nyamgerel N, Oyuntsetseg B, Tsegmed Z, Jin DP, Kubentayev S, Bayarmaa GA, Choi HJ (2026) *Oxytropis ikhbogdicus* (section Mesogaea, Fabaceae), a new species from Mongolia based on morphological and molecular analyses. *Ecology and Evolution* 16(6): e73803. <https://doi.org/10.1002/ece3.73803>
- Nie LY, Wang AH, Duan L, Chen HF, Wang FG (2021) The complete chloroplast genome of Plateau herb *Chesneya acaulis* (Fabaceae). *Mitochondrial DNA Part B* 6(2): 641–642. <https://doi.org/10.1080/23802359.2021.1878955>
- Nyamgerel N, Baasanmunkh S, Oyuntsetseg B, Bayarmaa GA, Erst A, Park I, Choi HJ (2023) Insight into chloroplast genome structural variation of the Mongolian endemic species *Adonis mongolica* (Ranunculaceae) in the Adonideae tribe. *Scientific Reports* 13: 22014. <https://doi.org/10.1038/s41598-023-49381-x>
- Nyamgerel N, Baasanmunkh S, Oyuntsetseg B, Tsegmed Z, Bayarmaa GA, Lazkov G, Pyak E, Gil HY, Park I, Choi HJ (2024) Comparative plastome analysis and taxonomic classification of snow lotus species (*Saussurea*, Asteraceae) in Central Asia and Southern Siberia. *Functional and Integrative Genomics* 24(2): 42. <https://doi.org/10.1007/s10142-024-01309-y>
- Oyuntsetseg Nyamgerel N, Baasanmunkh S, Oyuntsetseg B, Urgamal M, Yoon JW, Bayarmaa GA, Choi HJ (2024a) The complete chloroplast genome and phylogenetic results support the species position of *Swertia banzragczii* and *Swertia marginata* (Gentianaceae) in Mongolia. *Botanical Studies* 65(1): 11. <https://doi.org/10.1186/s40529-024-00417-z>
- Oyuntsetseg Nyamgerel N, Baasanmunkh S, Oyuntsetseg B, Bayarmaa GA, Choi HJ (2024b) The complete chloroplast genome of *Swertia obtusa* (Gentianaceae) in Mongolia. *Journal of Asia-Pacific Biodiversity* 18(3): 727–732. <https://doi.org/10.1016/j.japb.2024.12.001>

- Polhill RM (1981) Tribe 16. Galegeae (Bronn) Torrey & Gray (1838). Advances in legume systematics part 1: 357–363.
- Qin Q, Dong Y, Chen J, Wang B, Peng Y, Zhang X, Wang X, Zeng J, Zhong G, Zhang S, Du X (2025) Comparative analysis of chloroplast genomes reveals molecular evolution and phylogenetic relationships within the Papilionoideae of Fabaceae. BMC Plant Biology 25: 157. <https://doi.org/10.1186/s12870-025-06138-0>
- Rambaut A (2012) FigTree v1.4.2. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh. <http://tree.bio.ed.ac.uk/software/figtree>
- Rambaut A, Drummond A, Xie D, Baele G, Suchard MA (2018) Posterior summarization in Bayesian Phylogenetics using Tracer 1.7. Systematic Biology 67: 901–904. <https://doi.org/10.1093/sysbio/syy032>
- Robles P, Quesada V (2022) Unveiling the functions of plastid ribosomal proteins in plant development and abiotic stress tolerance. Plant Physiology and Biochemistry 189: 35–45. <https://doi.org/10.1016/j.plaphy.2022.07.029>
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61: 539–542. <https://doi.org/10.1093/sysbio/sys029>
- Sabater B (2021) On the edge of dispensability, the chloroplast *ndh* genes. International Journal of Molecular Sciences 22(22): 12505. <https://doi.org/10.3390/ijms222212505>
- Sabir J, Schwarz E, Ellison N, Zhang J, Baeshen NA, Mutwakil M, Jansen R, Ruhlman T (2014) Evolutionary and biotechnology implications of plastid genome variation in the inverted-repeat-lacking clade of legumes. Plant Biotechnology Journal 12(6): 743–754. <https://doi.org/10.1111/pbi.12179>
- Selyutina I, Konichenko ES, Dorogina OV (2017) Variability and interpopulation differentiation of the rare species *Gueldenstaedtia monophylla* Fisch. (Fabaceae). Vavilov Journal of Genetics and Breeding 21: 354–359. <https://doi.org/10.18699/18699/VJ16.15-o>
- Selyutina I, Konichenko E, Igay N, Darikhand D (2021) The specific features of the reproductive biology of the rare species *Gueldenstaedtia monophylla* Fisch. (Fabaceae). Northern Asia Plant Diversity, Bio Web of Conference 38: 00111. <https://doi.org/10.1051/bio-conf/20213800111>
- Seol BY, Kim SJ, Byun HJ, Lee CH, Kim SW, Park SY, Kim GY, Do HDK, Choi IS (2025) Complete plastid genome of *Crotalaria sessiliflora* (Fabaceae): Insights into plastome structure, *rps16* pseudogenization, and phylogenetic implications. Journal of Asia-Pacific Biodiversity 19(1): 27–36. <https://doi.org/10.1016/j.japb.2025.05.001>
- Shanker AK, Bhanu BD, Alluri A, Rajah N, Chavez R, Maheswari M (2022) Chloroplast evolution and genome manipulation. Clim Change Crop Stress 2022: 411–440. <https://doi.org/10.1016/B978-0-12-816091-6.00001-8>
- Shikanai T, Shimizu K, Ueda K, Nishimura Y, Kuroiwa T, Hashimoto T (2001) The chloroplast *clpP* gene, encoding a proteolytic subunit of ATP-dependent protease, is indispensable for chloroplast development in tobacco. Plant and Cell Physiology 42(3): 264–73. <https://doi.org/10.1093/pcp/pce031>

- Sirpio S, Allahverdiyeva Y, Holmstrom M, Khrouchtchova A, Haldrup A, Battchikova N, Aro EM (2009) Novel nuclear-encoded subunits of the chloroplast NAD (P) H dehydrogenase complex. *Journal of Biological Chemistry* 284(2): 905–912. <https://doi.org/10.1074/jbc.M805404200>
- Son O, Choi KS (2022) Characterization of the chloroplast genome structure of *Gueldenstaedtia verna* (Papilionoideae) and comparative analyses among IRLC species. *Forests* 13: 1942. <https://doi.org/10.3390/f13111942>
- Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>
- Stoppel R, Lezhneva L, Schwenkert S, Torabi S, Felder S, Meierhoff K, Westhoff P, Meurer J (2011) Recruitment of a ribosomal release factor for light- and stress-dependent regulation of *petB* transcript stability in *Arabidopsis* chloroplasts. *The Plant Cell* 23(7): 2680–2695. <https://doi.org/10.1105/tpc.111.085324>
- Stupnikova TV (2018) Growth and development of *Gueldenstaedtia verna* (Fabaceae) in nature and culture in the south of the Amur region. *Rastitel'nye Resursy* 54(2): 246–259. <https://doi.org/full/10.5555/20183309826>
- Talavera G, Castresana J (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology* 56(4): 564–577. <https://doi.org/10.1080/10635150701472164>
- Tamura K, Stecher G, Kumar S (2021) MEGA11: Molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution* 38: 3022–3027. <https://doi.org/10.1093/molbev/msab120>
- Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S (2017) GeSeq – versatile and accurate annotation of organelle genomes. *Nucleic Acids Research* 45: W6–W11. <https://doi.org/10.1093/nar/gkx391>
- Wang D, Zhang Y, Zhang Z, Zhu J, Yu J (2010) KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genomics, proteomics and bioinformatics* 8(1): 77–80. [https://doi.org/10.1016/S1672-0229\(10\)60008-3](https://doi.org/10.1016/S1672-0229(10)60008-3)
- Wang Y-H, Wicke S, Wang H, Jin JJ, Chen SY, Zhang SD, Li DZ, Yi TS (2018) Plastid genome evolution in the early-diverging legume subfamily Cercidoideae (Fabaceae). *Frontiers in Plant Science* 9: 138. <https://doi.org/10.3389/fpls.2018.00138>
- Wang N, Liu PL, Zhang L, Ma R, Zhao L, Wang H, Chang ZY (2025) The phylogenomic framework and infrageneric classification of temperate Asian Caraganeae (Leguminosae, Papilionoideae). *Ecology and Evolution* 15(12): e72638. <https://doi.org/10.1002/ece3.72638>
- Xie YP, Meng Y, Zhang ML (2016) Molecular phylogeny of *Gueldenstaedtia* and *Tibetia* (Fabaceae) and their biogeographic differentiation within Eastern Asia. *PLoS ONE* 11: e0162982. <https://doi.org/10.1371/journal.pone.0162982>
- Yang Z, Nielsen R (2000) Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Molecular Biology and Evolution* 17(1): 32–43. <https://doi.org/10.1093/oxfordjournals.molbev.a026236>

- Zhao Y, Zhang R, Jiang KW, Qi J, Hu Y, Guo J, Zhu R, Zhang T, Egan AN, Yi TS, Huang CH (2021) Nuclear phylotranscriptomics and phylogenomics support numerous polyploidization events and hypotheses for the evolution of rhizobial nitrogen-fixing symbiosis in Fabaceae. *Molecular Plant* 14(5): 748–773. <https://doi.org/10.1016/j.molp.2021.02.006>
- Zhang ML, Wen ZB, Hao XL, Byalt VV, Sukhorukov AP, Sanderson SC (2015) Taxonomy, phylogenetics and biogeography of *Chesneya* (Fabaceae), evidenced from data of three sequences, ITS, trnS-trnG, and rbcL. *Biochemical Systematics and Ecology* 63: 80–89. <https://doi.org/10.1016/j.bse.2015.09.017>
- Zhang T, Li M, Zhu X, Li S, Guo M, Guo C, Shu Y (2024) Comparative chloroplast genomes analysis provided adaptive evolution insights in *medicago ruthenica*. *International Journal of Molecular Science* 25(16): 8689. <https://doi.org/10.3390/ijms25168689>
- Zheng S, Pocai P, Hyvönen J, Tang J, Amiryousefi A (2020) Chloroplot: an online program for the versatile plotting of organelle genomes. *Frontiers in Genetics* 11: 576124. <https://doi.org/10.3389/fgene.2020.576124>
- Zhu XY (2004) A revision of the genus *Gueldenstaedtia* (Fabaceae). *Annales Botanici Fennici* 41: 283–291. <https://www.jstor.org/stable/23726977>
- Zhu XY (2005) Pollen and seed morphology of *Gueldenstaedtia* and *Tibetia* (Leguminosae): with a special reference to their taxonomic significance. *Nordic Journal of Botany* 23: 373–384. <https://doi.org/10.1111/j.1756-1051.2003.tb00408.x>

Supplementary material 1

Figure S1. The overall coverage depth of *Gueldenstaedtia monophylla* chloroplast genome

Figure S2. Schematic map of the (A) cis-splicing and (B) trans-splicing genes in the complete chloroplast genome of *Gueldenstaedtia monophylla*

Figure S3. Phylogenetic tree from Fabaceae species based on the protein coding genes using Maximum parsimony (MP) and Bayesian inference (BI) analysis

Authors: Dashzeveg Oyuntsetseg, Nudkhuu Nyamgerel, Shukherdorj Baasanmunkh, Dariganga Munkhtulga, Batlai Oyuntsetseg, Gun-Aajav Bayarmaa, Alexey A. Kechaykin, Courtney H. Babin, Hyeok Jae Choi

Data type: images

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://journal.asu.ru/biol/article/view/20478/17421>

Supplementary material 2

Table S1. List of species for phylogenetic and comparative chloroplast genomic analyses

Table S2. Gene content of the complete chloroplast genome of *Gueldenstaedtia monophyla*

Authors: Dashzeveg Oyuntsetseg, Nudkhuu Nyamgerel, Shukherdorj Baasanmunkh, Dariganga Munkhtulga, Batlai Oyuntsetseg, Gun-Aajav Bayarmaa, Alexey A. Kechaykin, Courtney H. Babin, Hyeok Jae Choi

Data type: tables

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://journal.asu.ru/biol/article/view/20478/17423>