



OPEN ACCESS

EDITED BY

Muhammad Saeed,
Government College University, Pakistan

REVIEWED BY

Yanxiang Lin,
Fujian University of Traditional Chinese
Medicine, China
Yanping Xie,
Huaibei Normal University, China

*CORRESPONDENCE

Yan Gou

✉ gouyan@scidc.org.cn

Guihua Jiang

✉ jiangguihua@cdutcm.edu.cn

[†]These authors share first authorship

RECEIVED 22 December 2025

REVISED 31 January 2026

ACCEPTED 09 February 2026

PUBLISHED 04 March 2026

CORRECTED 06 May 2026

CITATION

Gao B, Zeng Y, Tang M, Yuan M, Deng K,
Gou Y and Jiang G (2026) Characteristic
analysis of the chloroplast genome of
Cyathula capitata and comparative
analysis with *Cyathula officinalis* and
their hybrid *Cyathula officinalis* ×
Cyathula capitata.
Front. Plant Sci. 17:1772997.
doi: 10.3389/fpls.2026.1772997

COPYRIGHT

© 2026 Gao, Zeng, Tang, Yuan, Deng,
Gou and Jiang. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/)
(CC BY). 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.

Characteristic analysis of the chloroplast genome of *Cyathula capitata* and comparative analysis with *Cyathula officinalis* and their hybrid *Cyathula officinalis* × *Cyathula capitata*

Bixing Gao^{1†}, Yisha Zeng^{2†}, Min Tang², Maohua Yuan³, Ke Deng²,
Yan Gou^{1*} and Guihua Jiang^{2*}

¹Industrial Technology Basic Public Service Platform, Ministry of Industry and Information Technology, Sichuan Institute for Drug Control (Sichuan Medical Device Testing Center, Sichuan Musk Deer Research Institute), Chengdu, China, ²College of Pharmacy, Chengdu University of Traditional Chinese Medicine, Chengdu, China, ³Department of New Energy Materials and Chemistry, Leshan Normal University, Leshan, China

Introduction: *Cyathula capitata* (Wall.) Moq., a widely used medicinal herb in Yi medicine, is often combined with other herbs to treat traumatic injuries, rheumatism, and rheumatoid arthritis. This study aims to characterize the chloroplast genomes, assess variation levels, and elucidate the phylogenetic relationships among *Cyathula capitata*, *Cyathula officinalis*, and their hybrid (*Cyathula officinalis* × *Cyathula capitata*). The findings provide valuable references for species identification, genetic background analysis, and quality control of *Cyathula* medicinal materials.

Methods: We employed second-generation high-throughput sequencing technology to sequence the chloroplast genomes of *Cyathula capitata*, *Cyathula officinalis*, and their hybrid (*Cyathula officinalis* × *Cyathula capitata*). Comparative genomic analyses were conducted to examine their genomic structures, simple sequence repeats (SSRs), codon usage preferences, and inverted repeat (IR) regions. Additionally, a phylogenetic tree based on chloroplast genome sequences was constructed to clarify their evolutionary relationships.

Results: The chloroplast genomes of all three taxa displayed the typical quadripartite structure, comprising a large single-copy region, a small single-copy region, and two IR regions (IRa and IRb). The *Cyathula capitata* genomes ranged from 151,428 to 151,436 bp, showing notable intraspecific diversity likely influenced by geographic factors. However, two *C. capitata* samples shared an identical genome length of 151,518 bp and exhibited consistent genotypes across locations, indicating relative genomic conservation. Repeat sequence analysis identified hexanucleotide SSRs, a unique 16-bp single nucleotide insertion in the *rpl22* gene, and a 30–40 bp forward repeat within the *rps16_1-trnQ-UUG* intergenic region as distinguishing markers for *Cyathula officinalis* and its hybrids. Codon usage analysis revealed no strong bias at the third codon position among the three species, although codons ending with thymine (T) were used more frequently. IR boundary analysis showed variation only among isolated *C. capitata* samples. Comparative genomics highlighted *psbI-trnS-GCU*

and rps16_1-trnQ-UUG as highly variable regions. Phylogenetic analysis indicated that the hybrid (Z4) grouped within *C. capitata*, consistent with the maternal inheritance pattern of chloroplast genomes. Given the potential for variations in maternal parents among different hybrid batches, the positions of these elements in the phylogenetic tree may accordingly vary.

Conclusion: This study elucidated the chloroplast genome features and phylogenetic relationships of *C. capitata*, *C. officinalis*, and their hybrid. The findings offer significant molecular insights that facilitate species identification, genetic analysis, and quality assessment of *Cyathula* medicinal resources. These insights support the sustainable utilization and conservation of these resources.

KEYWORDS

chloroplast genome, *Cyathula capitata*, *Cyathula officinalis*, phylogenetic analysis, sequence characteristic

1 Introduction

Naisibuni (侧金盏), derived from the dried root of *Cyathula capitata* (Wall.) Moq., a species within the genus *Cyathula* of the Amaranthaceae family (Azi, 1994), is widely used in Yi medicine. Locally, it is known as Lebu or Naibu in the Liangshan Yi Autonomous Prefecture and is also referred to as Ma Niuxi or Ku Niuxi. The herb is characterized by a bitter, astringent, and slightly numbing taste, and is classified as neutral in nature (Li and He, 1990). Its primary functions include dispelling wind and dampness, removing blood stasis, and unblocking meridians (Rao, 2006). Commonly used in combination with other herbs, Naisibuni is primarily employed to treat “Sise”, a condition described as obstruction of qi and blood caused by the interaction of internal and external factors, and traumatic injuries. Clinical research has demonstrated that the Yi medicine formula “Wosi,” which features Naisibuni as its principal component, achieves high cure rates, low recurrence rates, and minimal adverse reactions, effectively addressing both symptoms and root causes in the treatment of gouty arthritis (GA) (Zhang et al., 2020; Zhou, 2020; Gu et al., 2024).

The medicinal plants “Naisibuni” (*Cyathula capitata*) and *C. officinalis* from the genus *Cyathula* (Amaranthaceae) have historically been well-documented in medicinal applications. However, due to the morphological similarities between the two source plants, they have long been prone to misidentification and confused use in the market. Conventional identification methods principally depend on morphological and histological characteristics, which are inherently subjective. Therefore, this study conducts a comparative analysis at the chloroplast genome level to provide molecular evidence for accurate identification. As stated in the “Niuxi Jie” section of Zhang Xichun’s Medical Records Combining Chinese and Western Medicine from the Republican Period, the following is documented: “Sichuan-produced ones exhibit two distinct colors, purple and white, with the purple ones being considered superior” (Zhang, 2016). When considered in conjunction with the description of plant morphology in the Flora of China (Flora of China Editorial Committee, 1979), it can be inferred that the *C. officinalis* utilized historically encompassed both *C. capitata* and *C. officinalis*, with Naisibuni being regarded as of

superior quality. However, due to factors such as introduction and cultivation, resource changes, and market circulation, the primary botanical origin of modern commercial *C. officinalis* has gradually shifted to *C. officinalis* as recorded in the Chinese Pharmacopoeia, while Naisibuni (often referred to as “Maniuxi” in the market) is commonly circulated as a regional customary substitute or an adulterant. Additionally, the natural hybrid between these two species—hybrid *Cyathula* (*C. officinalis* × *C. capitata*)—is also present in the market, further complicating accurate identification (Shi et al., 2017). A thorough investigation has revealed that “Niuxi” medicinal materials exhibit a variety of nomenclature and complex origins (Gao et al., 2024). This investigation has identified issues such as “different materials under the same name” and “same material under different names.” It is worth noting that a common misrepresentation of Naisibuni is to refer to him as *C. officinalis*. Consequently, Naisibuni is frequently designated as a significant adulterant of *C. officinalis* in identification studies. However, Naisibuni has been shown to possess distinctive therapeutic efficacy in the treatment of conditions such as rheumatism and arthralgia. The utilization of morphological and histological methods for identification is subject to limitations inherent in the methods themselves. Therefore, this study investigates the chloroplast genomes of Naisibuni and *C. officinalis* to provide supplementary molecular evidence for their differentiation.

Chloroplasts are semiautonomous organelles unique to green plants and certain protists, capable of photosynthesis and containing their own independent genomes (Margulis and Goode, 1981). The chloroplast genome of higher plants typically exhibits a highly conserved quadripartite structure, with sizes ranging from 115 to 165 kilobases (Jansen et al., 2005). This circular genome is divided by a pair of IR sequences (IRs) into a large single-copy (LSC) region and a small single-copy (SSC) region (Qian et al., 2021). Most angiosperm chloroplast genomes contain approximately 120 to 160 distinct genes (Jianguo et al., 2018), encoding proteins involved primarily in photosynthesis and gene expression, alongside transfer RNA (tRNA) and ribosomal RNA (rRNA) genes. Due to its high degree of conservation, abundant genetic information, and predominantly maternal inheritance, the chloroplast genome serves as an ideal resource for studying plant

genetic diversity and conducting phylogenetic analyses (McFadden, 2001; Olmstead and Palmer, 1994). Furthermore, with the rapid advancement of high-throughput sequencing technologies and the moderate nucleotide substitution rate of chloroplast DNA (cpDNA), which is significantly lower than that of nuclear DNA, the sequencing and assembly of chloroplast genomes have become increasingly efficient and accessible (Carbonell-Caballero et al., 2015). Consequently, chloroplast genome sequencing has emerged as a vital molecular marker approach for investigating plant phylogenetics and genetic differentiation.

Currently, *Cyathula officinalis* Kuan is frequently used as a substitute for *Cyathula capitata* in research. Adding to the complexity, hybrids between these two species, referred to as hybrid *Cyathula*, are commonly found in the market, further complicating the understanding of the genetic background and parental origins of *Cyathula* medicinal materials. Previous studies have compared the chloroplast genomes of *Cyathula officinalis* Kuan with those of Naisibuni and three other *Cyathula* species (Guo et al., 2022). However, molecular approaches specifically designed to identify the parental origins of hybrid *Cyathula* remain undeveloped. Moreover, significant intraspecific variation may occur within the same Naisibuni cultivar due to geographical differences (Fu et al., 2013). This study systematically compares the chloroplast genomes of *Cyathula capitata*, *Cyathula officinalis*, and their hybrid (*Cyathula officinalis* × *Cyathula capitata*) from diverse origins. It aims to elucidate the structural features and patterns of inter- and intraspecific variation within the chloroplast genomes, identify the maternal lineage of the hybrid, and infer parental preferences.

2 Materials and methods

2.1 Plant materials

Based on a comprehensive literature review and field survey, fresh plant samples were collected in July 2024, including three batches of *Cyathula capitata*, two batches of *Cyathula officinalis*, and six batches of their hybrid (*Cyathula officinalis* × *Cyathula capitata*). Young leaves were carefully selected, rinsed with clean water, and immediately dried using absorbent paper to remove surface moisture. The samples were then placed into cryogenic tubes and stored on dry ice for subsequent DNA extraction. Detailed records of sampling time, location, latitude, longitude, and altitude were maintained. Species identification of the original plants was confirmed by Professor Jiang Guihua from Chengdu University of Traditional Chinese Medicine and Senior Pharmacist Gao Bixing of the Sichuan Provincial Institute for Drug Control (Supplementary Table 1). Based on preliminary surveys, samples of *C. capitata* (M2, M3) were collected from Liangshan, Sichuan, where *C. officinalis* is absent, and samples of *C. officinalis* (C1) were obtained from Hubei, where *C. capitata* does not occur—representing allopatric populations without coexistence of the other parental species. Conversely, hybrid samples (Z1–Z6) and potential parental samples were collected from regions such as Ya'an and Leshan in Sichuan, where the two species may exhibit sympatric distribution.

2.2 Sequencing and assembly

2.2.1 DNA extraction and sequencing

Genomic DNA was extracted from fresh samples using a whole-genome shotgun approach (Kahraman and Lucas, 2019) to construct libraries with insert sizes of approximately 400 base pairs (bp). Each library underwent paired-end (PE) sequencing utilizing second-generation sequencing (NGS) technology (Zheng et al., 2025) on the Illumina NovaSeq platform (Liu et al., 2022), with one library prepared per sample.

2.2.2 Chloroplast genome assembly and annotation

Quality control of the raw sequencing data was performed using fastp (<https://github.com/OpenGene/fastp>) to filter and generate high-quality clean reads. The chloroplast genomes were then assembled using GetOrganelle v1.7.7.0 (Jin et al., 2020) (<https://github.com/Kinggerm/GetOrganelle>), resulting in complete circular chloroplast sequences.

As *Amaranthus* and *Cyathula* belong to the same family (Amaranthaceae), they share sufficient sequence similarity for read mapping. The plastid genome of *Amaranthus tricolor* possesses a conserved and well-annotated quadripartite structure (Mower and Vickrey, 2017; Viljoen et al., 2018). To ensure the accuracy of chloroplast genome assembly and annotation, the fully annotated sequence of *Amaranthus tricolor* (NC_065013) was retrieved from the National Center for Biotechnology Information (NCBI) database and used as the reference. Annotation of the 11 complete chloroplast genomes listed in Supplementary Table 1 was performed using the online tool CPGAVAS2 (Shi et al., 2019) (<http://47.96.249.172:16019/analyzer/annotate>). The annotated features, including start and stop codons, introns, and exons, were manually verified to ensure accuracy. After confirming the correct annotation of protein-coding genes (CDs), rRNA genes, and tRNA genes, all sequencing data were submitted to the GenBank database at NCBI, with accession numbers provided in Supplementary Table 1. The CPGAVAS2 tool was also used to generate graphical maps of the annotated chloroplast genomes.

2.3 Basic characteristics and comparative analysis of chloroplast genomes

2.3.1 Basic characteristics of chloroplast genomes

Genome-wide length, lengths of the LSC, SSC, and IR regions, as well as GC content, were statistically analyzed using Geneious Prime 2025.0.2 (Kearse et al., 2012).

2.3.2 Analysis of simple sequence repeats and local sequence repeats in the chloroplast genome

SSRs in *Cyathula capitata*, *Cyathula officinalis*, and their hybrid (*Cyathula officinalis* × *Cyathula capitata*) were statistically analyzed

TABLE 1 Chloroplast genome information for phylogenetic inference.

Species name	Accession number
<i>Achyranthes bidentata</i>	MN652923
<i>Achyranthes aspera</i>	MN953051.1
<i>Achyranthes longifolia</i>	MN953049.1
<i>Amaranthus tricolor</i>	KX094399.1
<i>Alternanthera philoxeroides</i>	NC_0427989.1
<i>Celosia cristata</i>	NC045887
<i>Gomphrena globosa</i>	NC069834
<i>Suaeda japonica</i>	NC_02675.1
<i>Haloxylon persicum</i>	NC_027669.1
<i>Salsola collina</i>	NC066995
<i>Dysphania botrys</i>	NC_042166.1
<i>Dysphania pumilio</i>	NC_041159.1
<i>Oxybasis glauca</i>	NC_047226
<i>Chenopodium ficifolium</i>	NC_041200.1
<i>Chenopodium quinoa</i>	NC_034949
<i>Limonium franchetii</i>	OR374019
<i>Limonium bicolor</i>	NC_059915.1

using the online tool MISA (<https://webblast.ipk-gatersleben.de/misa/>) (Beier et al., 2017). The parameters were set as follows: mononucleotide repeats ≥ 10 , dinucleotide repeats ≥ 6 , trinucleotide repeats ≥ 5 , and tetranucleotide, pentanucleotide, and hexanucleotide repeats ≥ 4 .

Long repeat sequences were identified using the online software Reputer (<https://bibiserv.cebitec.uni-bielefeld.de/reputer>) (Kurtz and Schleiermacher, 1999), with a Hamming distance of 3 and a minimum repeat length of 30 bp. Four types of repeats were detected: forward, reverse, complementary, and palindromic repeats.

2.3.3 Codon usage bias analysis

Protein-coding genes were extracted and screened from *Cyathula capitata*, *Cyathula officinalis*, and their hybrid (*Cyathula officinalis* $\times$ *Cyathula capitata*) using the following criteria: (1) sequences shorter than 300 bp were excluded; (2) sequences not beginning with the start codon ATG were removed; and (3) only one copy was retained for genes with multiple repeats. After filtering, the remaining gene sequences were merged, and the relative synonymous codon usage (RSCU) values were calculated using the formula: $\text{RSCU} = \text{number of codons} / \text{total number of synonymous codons} / \text{number of synonymous codon types}$. Finally, codon usage preference analysis was conducted on the chloroplast genomes.

2.3.4 Chloroplast genome comparative analysis

Variation analysis of the IR regions was conducted using IRscope software (<https://irscope.shinyapps.io/irapp/>) (Amiryousefi et al., 2018). Comparative and variation analyses of the chloroplast genomes of *Cyathula capitata*, *Cyathula officinalis*,

and their hybrid (*Cyathula officinalis* $\times$ *Cyathula capitata*) were performed using mVISTA (<https://genome.lbl.gov/vista/mvista/submit.shtml>) (Frazer et al., 2004).

2.3.5 Sequence variation analysis

Single-nucleotide polymorphisms (SNPs) represent sequence variations arising from single nucleotide changes, primarily through nucleotide transitions and transversions (Zou and Ge, 2003). To identify potential DNA barcode regions, nucleotide diversity (P_i) was calculated for both intergenic and coding regions to pinpoint SNP hotspots within *Cyathula capitata*, *Cyathula officinalis*, and their hybrid (*Cyathula officinalis* $\times$ *Cyathula capitata*). To identify highly variable regions, polymorphic sites and nucleotide diversity (P_i) in the six MAFFT-aligned CPGs were assessed using a sliding window analysis in DnaSP 6.10 (Yang et al., 2023). This analysis was complemented by mVISTA-based visual alignments of the chloroplast genomes.

2.3.6 Phylogenetic analysis

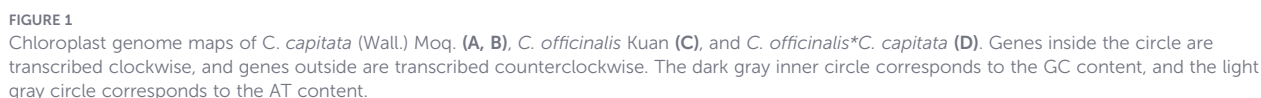
In addition to the three target plants, we downloaded the complete chloroplast genome sequences of 15 other published Amaranthaceae species from NCBI, along with a white-flowered *Dianthus* species as an outgroup (Table 1). Multiple sequence alignment was performed with MAFFT v7.526 to optimize alignment accuracy and retain reliable phylogenetic signals (Zhang et al., 2020). Subsequently, the aligned sequences were entered into IQ-TREE v1.6.8, which was utilized to construct a phylogenetic tree based on the maximum likelihood method.

3 Results

3.1 Basic characteristics of the chloroplast genomes of Naisibuni, *Cyathula officinalis*, and *Cyathula officinalis* $\times$ *Cyathula capitata*

The genome structures of Naisibuni, *Cyathula officinalis*, and the hybrid *Cyathula officinalis* $\times$ *Cyathula capitata* exhibit a typical closed quadripartite circular configuration, consistent with the chloroplast genome architecture of angiosperms. Analysis of the FASTA files revealed that sample C1 shares identical sequences with C2, Z1, Z3, Z5, and Z6; M1 is identical to Z2; while M2, M3, and Z4 each possess unique sequences. Notably, only C1, M1, M2, M3, and Z4 represent distinct sequence types. These results indicate that all collected *Cyathulae Radix* samples share identical sequences, whereas the three Naisibuni batches display sequence heterogeneity. Among the six batches of *Radix Achyranthis Bidentatae* var. *heterophylla*, only three distinct sequences were identified. Given that identical sequences share the same characteristics, subsequent analyses focused solely on these distinct sequence types.

The chloroplast genome of *Cyathula officinalis* spans 151,518 bp, comprising a pair of 25,229 bp IR regions separated by an 83,615 bp LSC region and a 17,445 bp SSC region. The complete chloroplast



sequences for all samples have been submitted to GenBank, with accession numbers listed in [Supplementary Table 1](#). Gene maps illustrating the chloroplast genome structures for the different sequences are presented in [Figure 1](#).

SSRs, also known as microsatellites, are short tandem repeats consisting of 1 to 6 nucleotides. These sequences are characterized by the consecutive repetition of a small nucleotide motif, and the number, type, and distribution of SSR loci vary among species (Chen and Huang, 2004).

Our analysis revealed that *Cyathula capitata* contains four types of SSRs: mononucleotide, dinucleotide, tetranucleotide, and hexanucleotide repeats. In contrast, *Cyathula officinalis* and the hybrid *Cyathula officinalis* $\times$ *Cyathula capitata* exhibit only three types: mononucleotide, dinucleotide, and tetranucleotide repeats. Mononucleotide repeats predominate across all three species, with T/

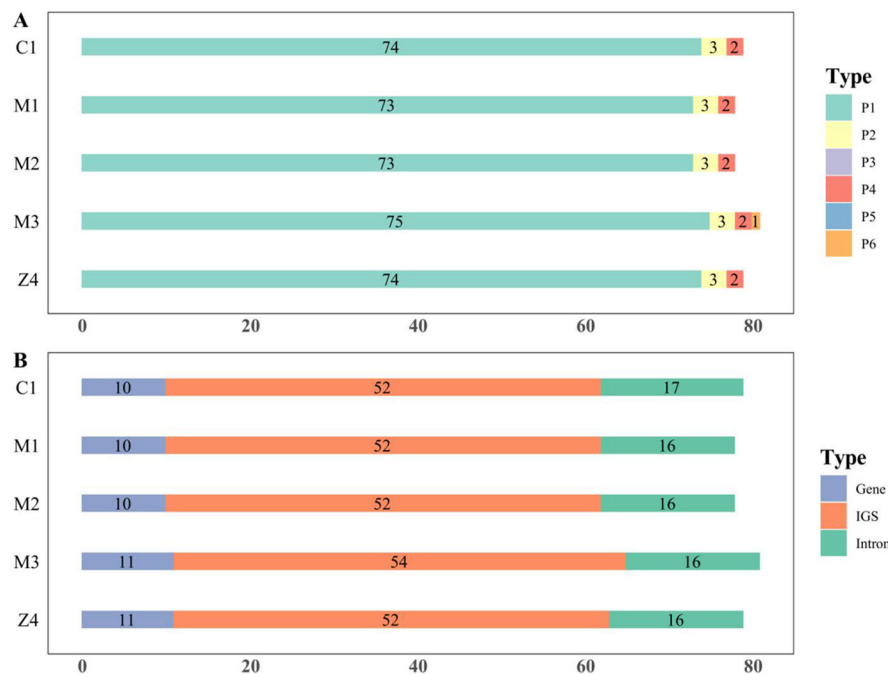


FIGURE 2

SSR distribution in the chloroplast genomes of *Cyathula capitata*, *Cyathula officinalis*, and *Cyathula officinalis* × *Cyathula capitata*. P1-Mononucleotide repeats P2-Dinucleotide repeats P3-Trinucleotide repeats. P4-Tetranucleotide repeats P5-Pentanucleotide repeats P6-Hexanucleotide repeats. (A) Nucleotide repeat types and counts (B) SSR distributions across genomic re.

A repeats being the most frequent. This prevalence aligns with previous research indicating that high AT content is a common feature of higher plant chloroplast genomes. SSRs are primarily distributed in IGS regions, predominantly within the LSC region of the chloroplast genome (Yang et al., 2019). While overall SSR characteristics are similar among the species, differences in the number and distribution of SSRs were observed (Supplementary Table 4, Figure 2). Specifically, mononucleotide SSRs accounted for 93.75–94.81% of SSRs in *Cyathula capitata*, with repeat lengths ranging from 10 to 15. In *Cyathula officinalis*, mononucleotide SSRs made up 94.87%, with lengths ranging from 10 to 16.

The analysis of long repeat sequences (Figure 3) showed that all three species possess only palindromic and forward repeats, with palindromic repeats being more abundant. *Cyathula capitata* contained 29–30 long repeats, including 13 forward repeats; *Cyathula officinalis* had 30 total repeats with 14 forward repeats; and the hybrid exhibited 29–30 repeats with 13–14 forward repeats. These long repeats were primarily located within IGS regions. Notably, *Cyathula officinalis* contained an additional forward repeat within the *rps16_1-trnQ-UUG* intergenic region, ranging from 30 to 40 bp in length. Similarly, the M3 batch of the hybrid harbored an extra palindromic repeat within the *psbN* gene region, measuring 40 to 50 bp. Lastly, the Z1, Z3, Z5, and Z6 batches of the hybrid *Achillea* were identical to *Achillea chinensis*.

3.3 Codon usage bias analysis

The RSCU values for *Cyathula capitata*, *Cyathula officinalis*, and their hybrid (*Cyathula officinalis* × *Cyathula capitata*) (Supplementary Table 5), along with the synonymous codon

relative abundance plots (Figure 4), reveal consistent patterns across all three species. The codons with the highest and lowest relative abundances are both associated with methionine (Met). Specifically, the codon GCT, encoding alanine (Ala), exhibited the highest relative abundance, while CTG, encoding leucine (Leu), showed the lowest. Although slight variations in relative codon usage frequencies were observed among the three chloroplast genomes, their overall codon usage patterns were largely similar. Codons ending with any of the four nucleotides (A, T, G, or C) displayed a usage frequency of approximately 16%. In *Cyathula officinalis* and *Cyathula capitata*, codons with differing relative usage frequencies, beyond GCA (Ala), were predominantly those ending with thymine (T). These included CCT (proline, Pro), GCT (Ala), and ACT (threonine, Thr). In *Cyathula capitata*, the only codon showing a distinctive relative usage frequency among T-ending codons was TCT, which encodes serine (Ser). As illustrated in Figure 4 and Supplementary Table 5, all amino acids, except tryptophan encoded by TGG and Met encoded by ATG, are encoded by multiple codons. For example, arginine, Leu, and Ser each have six codons; Ala, glycine, Pro, Thr, and valine (Val) have four codons; isoleucine has three; and the remaining amino acids are encoded by two codons differing only at the third nucleotide position.

3.4 Comparison of IR region boundary positions in chloroplast genomes

During the evolution of chloroplast genomes, contraction and expansion of boundary regions have frequently occurred (Zhang et al., 2024; Liu et al., 2023; Ju et al., 2017). The chloroplast genomes

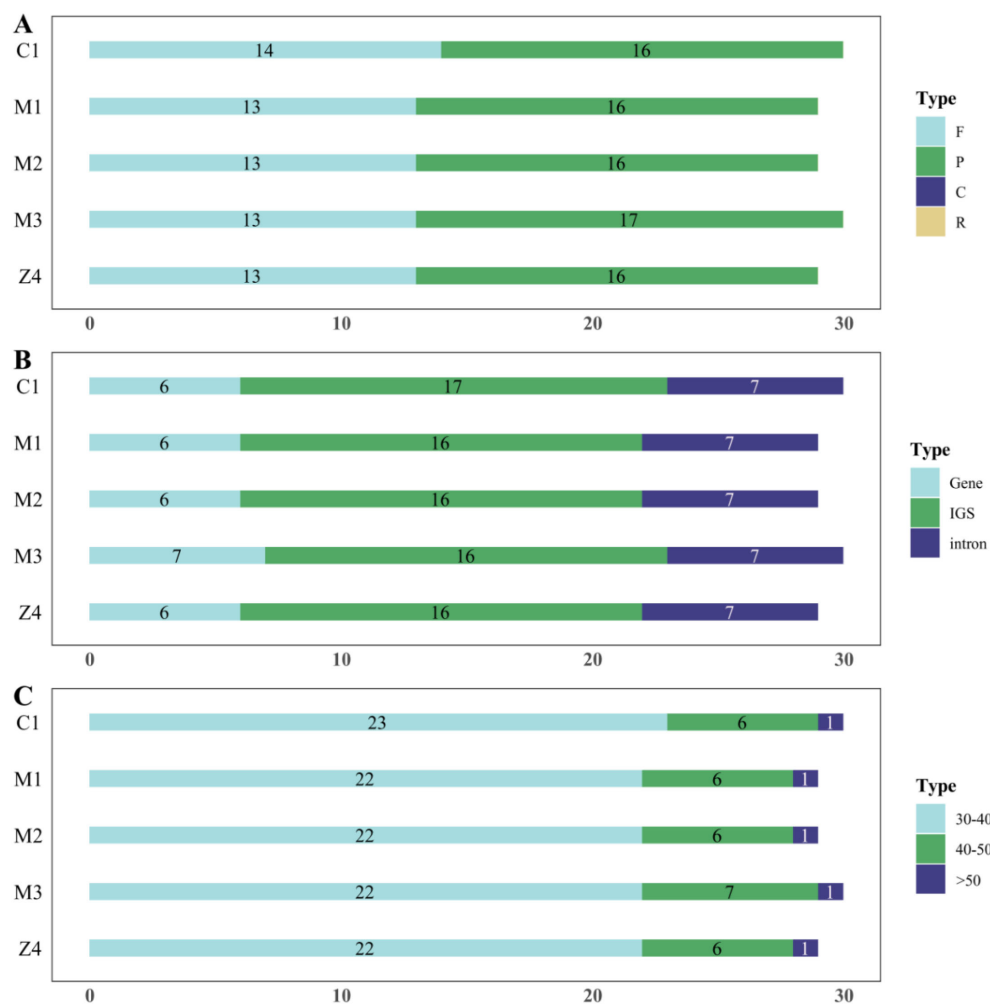
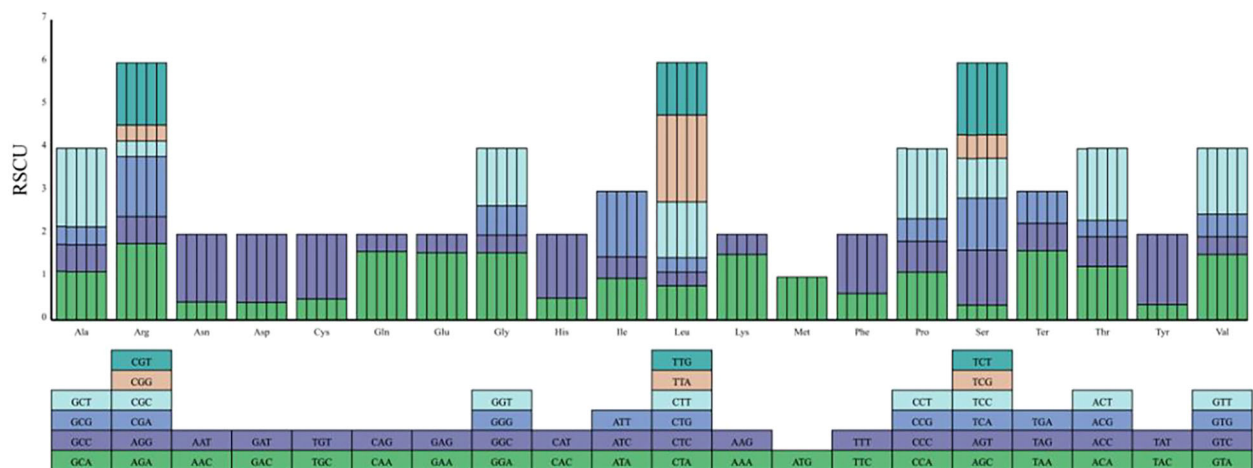


FIGURE 3 Distribution of LSRs in the chloroplast genomes of *Cyathula capitata*, *Cyathula officinalis*, and *Cyathula officinalis* × *Cyathula capitata*. F-Forward repeats P-Palindromic repeats C-Complementary repeats R-Reverse repeats. **(A)** Types of long repeats. **(B)** Genomic regions of LSR distribution. **(C)** Frequency of LSRs by length ranges.



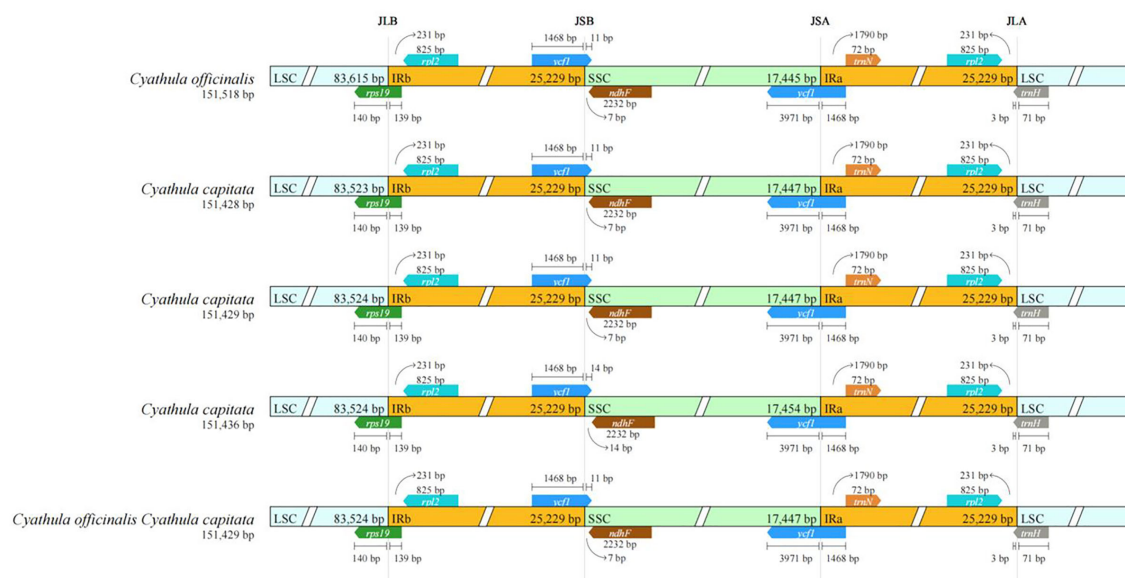


FIGURE 5

Comparison of IR boundary sequences in the chloroplast genomes of *Cyathula capitata*, *Cyathula officinalis*, and *Cyathula officinalis* × *Cyathula capitata*.

of *Cyathula capitata*, *Cyathula officinalis*, and their hybrid (*Cyathula officinalis* × *Cyathula capitata*) all exhibit the typical circular quadripartite structure, featuring four junctions: LSC–IRb, IRb–SSC, SSC–IRa, and IRa–LSC. Using the online tool IRscope, we compared the IR boundary regions across these species. The results (Figure 5) revealed that the LSC/IR and SSC/IR boundaries showed minimal variation among the three species. The SSC/IRa, IRa/LSC, and LSC/IRb junctions were highly conserved, with no observable differences. At the IRb/SSC boundary, the junction was located immediately adjacent to the *ndhF* gene in all species, except in *Cyathula capitata* (batch M3), where a 14 bp gap separated this boundary from the gene. These findings suggest that although the chloroplast genome boundaries among these species are relatively conserved with respect to gene content and sequence length at contraction and expansion sites, some variation still exist.

3.5 Analysis of chloroplast genome single nucleotide polymorphisms and sequence variability in plants

Using the online software mVISTA, we analyzed sequence variation across the chloroplast genomes of the three species (Figure 6). The results confirmed that all three species belong to the Amaranthaceae family, exhibiting high overall sequence similarity while retaining certain distinct differences. Regions showing greater variability included *psbI-trnS-GCU*, *psbM-trnD-GUC*, *trnG-GCC-trnF-CAU*, and *trnL-UAA-trnF-GAA*, with *trnH-GUG* identified as the most divergent gene. These hypervariable regions are promising candidates for developing specific molecular markers for species identification and phylogenetic studies.

To further characterize variation within these highly variable regions, nucleotide diversity (π) values were calculated for the chloroplast genomes of *Cyathula officinalis*, *Cyathula capitata*, and their hybrid (*Cyathula officinalis* × *Cyathula capitata*). Intergenic

regions exhibited π values ranging from 0 to 0.03462 (Figure 7B), while coding regions showed lower variability, with π values ranging from 0 to 0.00143 (Figure 7A), indicating that coding sequences are relatively conserved across the three species. Among intergenic regions, the *psbI-trnS-GCU* interval displayed the highest nucleotide diversity ($\pi = 0.03462$). Located within the LSC region, this site represents a hypervariable zone and holds potential as a specific molecular marker for species identification. Within coding regions, the *rps19* gene exhibited the greatest nucleotide diversity ($\pi = 0.00143$). Since all π values for coding regions were below 0.0100, no coding genes displayed notably high variability.

3.6 Phylogenetic analysis

The species from the *Plumbaginaceae* family constitute a distinct clade as the outgroup (Figure 8), while those from the *Amaranthaceae* family are grouped together and further divided into two major branches. One branch encompasses the genera *Chenopodium*, *Oxybasis*, *Dysphania*, *Salsola*, *Haloxylon*, and *Suaeda*; the other branch comprises the genera *Cyathula*, *Achyranthes*, *Gomphrena*, *Alternanthera*, *Amaranthus*, and *Celosia*. Among the species under consideration, those of *Achyranthes* and *Cyathula* exhibit the closest phylogenetic relationship. Within the genus *Cyathula*, *Cyathula officinalis* and *Cyathula capitata* form a monophyletic clade with 100% bootstrap support, and the two species are separated into distinct subclades. These findings indicate that phylogenetic analysis based on chloroplast genomes can clearly distinguish between them. The chloroplast genome of the hybrid (Z4) is nested within the *C. capitata* clade, which is consistent with the pattern of maternal chloroplast inheritance. That is, the offspring's chloroplast genome is derived from the maternal parent. Therefore, the offspring's chloroplast genome cannot be separated from the maternal lineage in the phylogenetic tree. Given the potential variability in

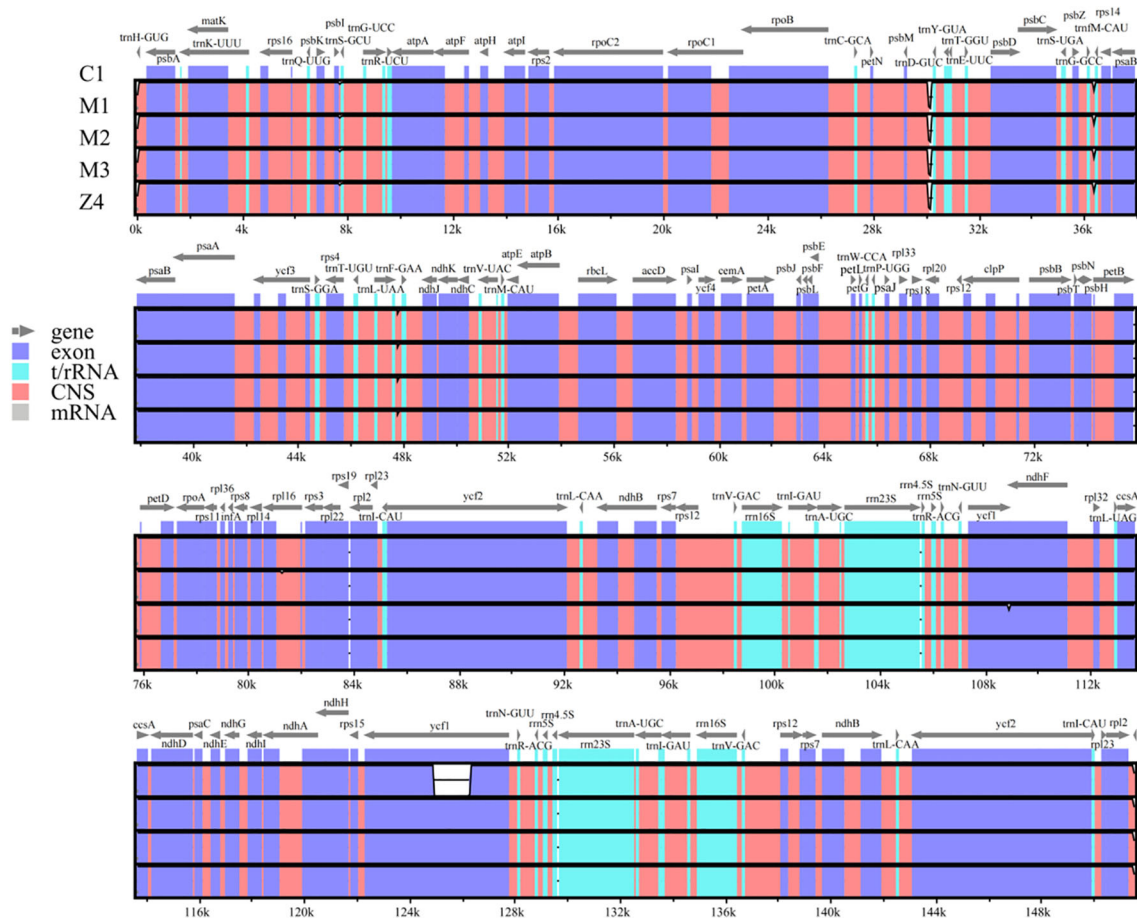


FIGURE 6
Complete chloroplast genome sequences of *Cyathula capitata*, *Cyathula officinalis*, *Cyathula officinalis* × *Cyathula capitata*. The vertical scale indicates the percentage of identity, ranging from 50% to 100%. The horizontal axis represents coordinates within the chloroplast genome. Genome regions are color-coded for protein-coding, rRNA, tRNA, intron, and conserved non-coding sequences.

maternal parents across different batches of the hybrid, the positions of these batches in the tree may also exhibit variation.

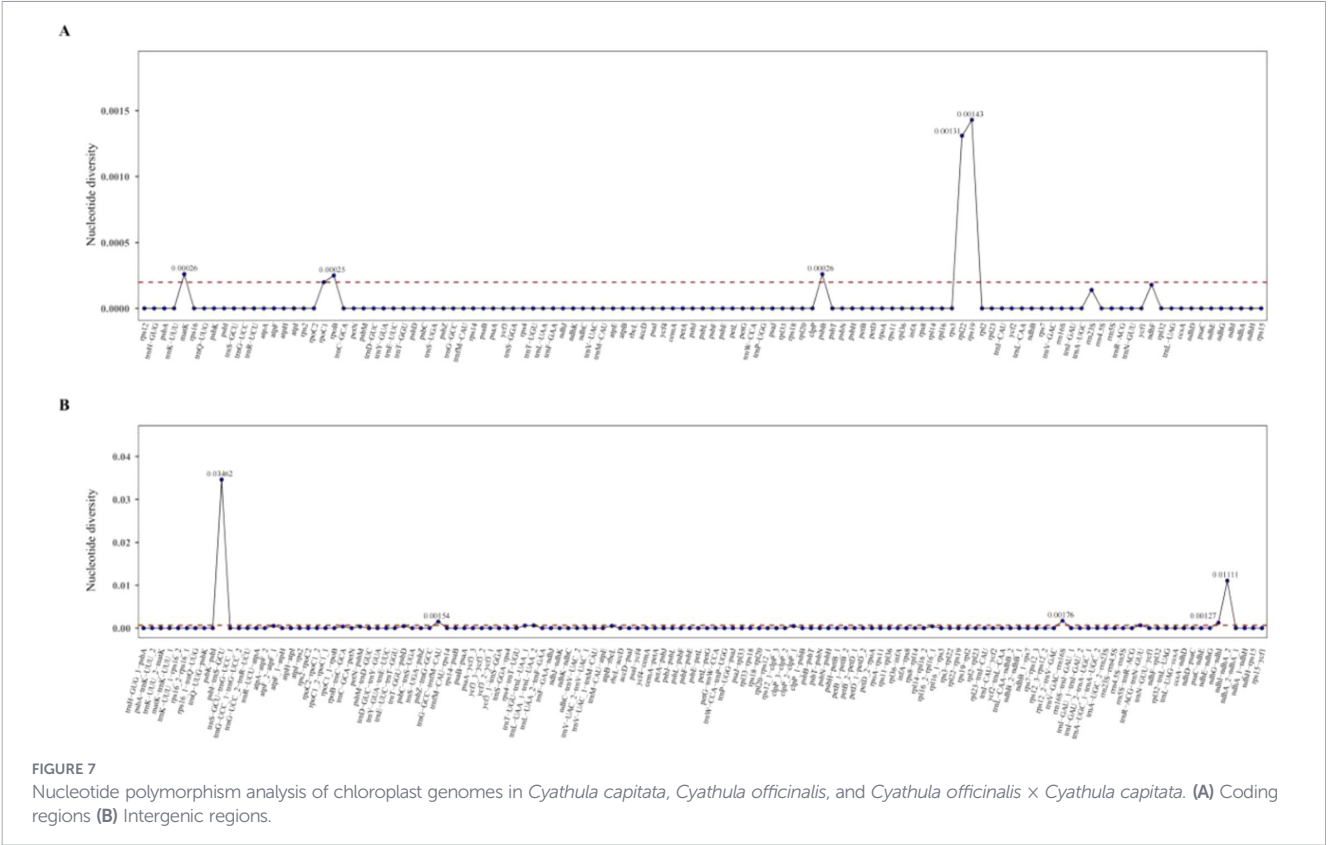
4 Discussion

Chloroplasts originated from free-living prokaryotes, but over time they have retained fewer genes from their ancestor, resulting in a progressive reduction in chloroplast genome size (Kleine et al., 2009). In angiosperms, chloroplasts are relatively conserved in structure, gene content, and gene arrangement (Xiao et al., 2025). Despite this overall conservation in structure and sequence, chloroplast genomes have been shown to vary in size, structure, and evolutionary rates of genes, even among closely related species (Williams et al., 2019). Due to features such as conserved architecture, lack of recombination, uniparental inheritance, and ease of amplification, chloroplast genomes play a significant role in phylogenetic and biogeographic studies (Song et al., 2024). Furthermore, chloroplast sequences are well-suited for high-throughput sequencing due to their high copy number, structural simplicity, highly conserved gene content, and relatively small size. With the ongoing decrease in the cost of second-generation

sequencing technologies, efficient, rapid, and cost-effective high-quality sequencing of complete chloroplast genomes has become widely adopted (Ji et al., 2019). For instance, in genera such as *Scrophularia* and *Artemisia* L., comparative phylogenomic analyses based on chloroplast genomes have been effectively utilized for species delimitation and authentication of adulterants (Guo et al., 2023; Chen et al., 2024).

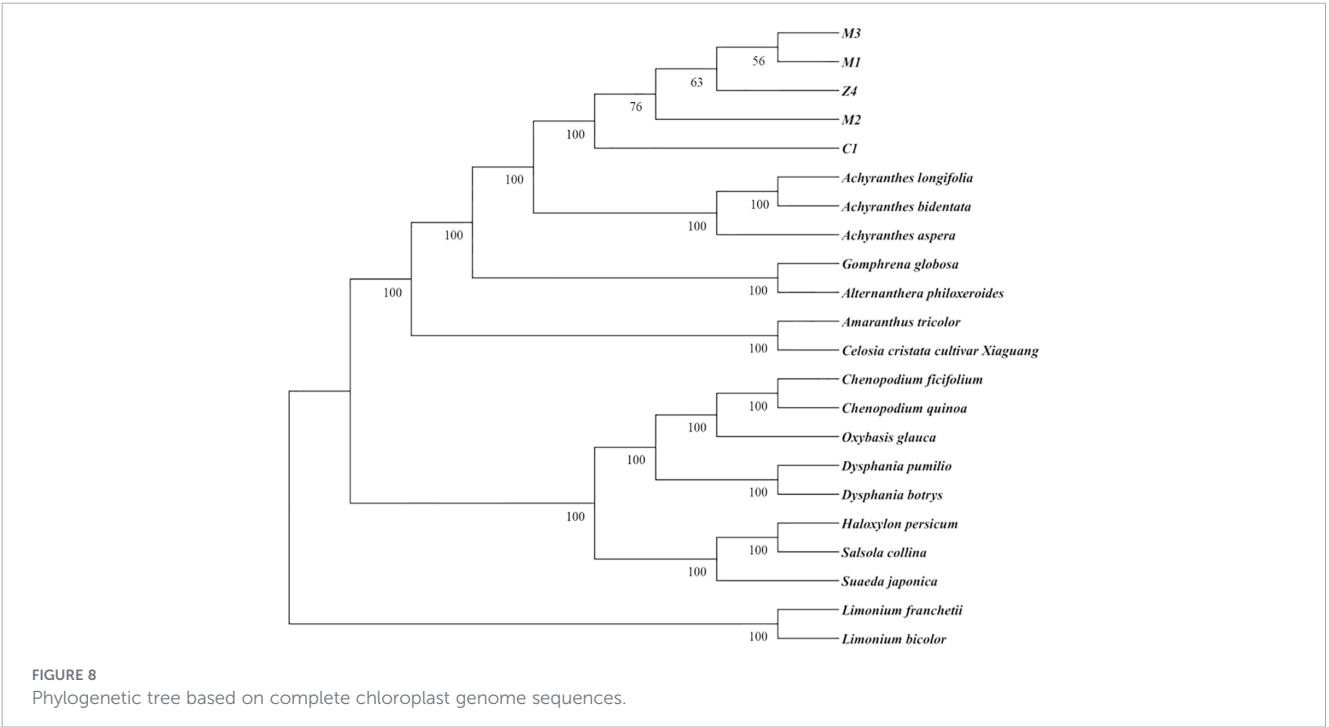
The *Amaranthaceae* family comprises approximately 60 genera and 850 species worldwide, with 13 genera and about 39 species recorded in China. Within this family, the genus *Cyathula* includes roughly 27 species distributed across Asia, Oceania, Africa, and the Americas. However, chloroplast genome data remain limited to only a few species, constraining our understanding of their evolutionary relationships and taxonomic classification.

Despite the evident variability among chloroplasts in terms of size, structure, and the rates at which genes evolve, the preservation of their structural and sequence characteristics is of paramount importance in the context of phylogenetic analysis and species identification (Yin et al., 2017). In this study, complete chloroplast genomes of *Cyathula capitata*, *C. officinalis*, and their hybrid were successfully obtained, thereby elucidating the structural features of their chloroplast genomes. The chloroplast genomes of *C. capitata*, *C. officinalis*, and the hybrid exhibited the characteristic closed



circular quadripartite structure, with total lengths of 151,428–151,436 bp, 151,518 bp, and 151,429–151,518 bp, respectively. These lengths are comparable to those reported for *Corispermum patelliforme* (150,614 bp, 128 genes) (Yang et al., 2025), *Dysphania pumilio* (151,962 bp, 128 genes) (Kim et al., 2019), and other

Amaranthaceae species (ranging from 149,726 to 153,474 bp based on NCBI data), aligning with the length characteristics observed in most angiosperm chloroplast genomes. Their complete chloroplast genomes consistently exhibited the typical quadripartite structure, consisting of the SSC, LSC, and two IR



regions. Overall, the basic chloroplast genomic features of *C. capitata* and *C. officinalis* are consistent with those of most *Amaranthaceae* species, indicating high conservation in genome size and structure. The GC content was found to be higher in the IR regions compared to the LSC and SSC regions. This variation can be attributed to the presence of GC-rich genes within the IRs. Furthermore, while angiosperm chloroplast genomes typically exhibit an average GC content of approximately 35%, the three studied species demonstrated a slightly elevated GC content of 36.46%, which exceeds the typical average for angiosperms (Yang et al., 2025). The presence of introns within plant chloroplast genomes is a salient feature that merits close examination. In this study, it was found that *rps12*, *ycf3*, and *clpP* each contained two introns, while other intron-containing genes contained only one intron. These characteristics are consistent with those observed in the chloroplast genomes of most angiosperms.

Chloroplast SSR markers have been found to facilitate species identification due to their maternal inheritance (Wei et al., 2024). SSR markers, known for their polymorphism, codominance, high reproducibility, and stability, are extensively applied in plant germplasm research (Drouin et al., 2008). We identified 79 SSRs in both *C. officinalis* and the hybrid. For *C. capitata*, batches M1 and M2 contained 78 SSRs each, while batch M3 exhibited 81 SSR loci. SSRs were predominantly located in the LSC region and mainly distributed within IGS regions of the single-copy regions across all three species. SSRs with repeat units of one to five nucleotides were present in all species, but hexanucleotide repeats were exclusive to *C. officinalis* and certain hybrid batches, likely reflecting maternal inheritance from *C. officinalis*. The *C. capitata* samples collected as Naisibuni lacked hexanucleotide repeats. Mononucleotide repeats reached a maximum length of 15 bp in *C. capitata* and some hybrids, while a unique 16 bp mononucleotide repeat was detected within the *rpl22* gene in *C. officinalis* and some hybrid batches. Additionally, an extra forward repeat sequence (30–40 bp) was identified in the *rps16_1-trnQ-UUG* region of *C. officinalis* and certain hybrid batches. These SSR characteristics may prove valuable for clarifying hybrid parental origins and developing precise molecular identification markers.

Codon usage bias analysis facilitates the discovery of novel genes and the prediction of gene function (Luo et al., 2024). In the three *Cyathula* species studied, thirty codons exhibited RSCU values greater than 1, indicating they are frequently used codons. Notably, 96.67% (29 out of 30) of these preferred codons ended with adenine (A) or uracil (U), reflecting a strong bias toward A- or U-ending codons. This bias correlates closely with the high AT content (63.54%) observed in the chloroplast genomes and suggests that an AT preference is a common feature within the genus *Cyathula*. These findings align well with previous codon preference analyses reported for 18 species in the Ampelopsidae tribe, four *Polygonatum* species (*Polygonatum odoratum* (Mill.) Druce, *Polygonatum sibiricum*, *Polygonatum cyrtoneura*, and *Polygonatum kingianum*), and three *Astragalus* species (*Astragalus dahuricus* (Pall.) DC., ..., *Astragalus melilotoides* Pall. var. *tenuis* Ledeb. and *Astragalus laxmannii* Jacq.), where preferred codons similarly ended predominantly with A or U (Hu et al., 2024;

Shi et al., 2024; Yin et al., 2024). The protein-coding sequences of the three *Cyathula* species comprise 64 codons, of which 61 encode the 20 standard amino acids, and the remaining three function as stop codons. Interestingly, the codons encoding Met exhibited both the smallest and largest relative abundance values. RSCU analysis revealed no strong bias at the third codon position among the three species, although codons with differing usage frequencies tended to end mostly with U. The results indicated that the chloroplast genomes of the three species were relatively conserved, with no significant differences observed in the IR boundaries. However, a specific variation was detected in the IRb/SSC boundary of the M3 batch of *Cyathula capitata*. Despite the genomic structure's conservation at the species level, this finding points to the potential presence of low-frequency structural polymorphisms or minor intraspecific diversity within the population, thereby underscoring the biodiversity of this species.

C. officinalis × *C. capitata* is derived from *C. capitata* and *C. officinalis*. Given the maternal inheritance of the chloroplast genome, it is observed that certain hybrid samples exhibit characteristics indicative of the maternal parent. A comparative analysis of chloroplast genomes revealed variations among three batches of *C. capitata* collected from Sichuan, while *C. officinalis* samples from both Hubei and Sichuan provinces exhibited identical genotypes. These results suggest that the chloroplast genome of *C. capitata* exhibits genetic diversity, potentially influenced by geographical factors, while *C. officinalis* appears to exhibit a more conserved genetic profile. Among the six hybrid batches, over half exhibited genotypes identical to *C. officinalis*, suggesting a preferential maternal role of *C. officinalis* in hybridization. Preliminary research has demonstrated that the analysis of chloroplast gene sequencing data can yield ample information for the delineation of species and the reconstruction of evolutionary relationships (Bezeng et al., 2017; Li et al., 2023). Given the predominantly maternal inheritance of chloroplasts, several studies have demonstrated the utility of nuclear single-nucleotide polymorphism (SNP) markers in resolving plant phylogenies (Viljoen et al., 2018). Each chloroplast is subject to uniparental inheritance and is non-recombinant. A comparative analysis further indicated that regions of high divergence and hypervariable sites are primarily located within the LSC region. The *psbI-trnS-GCU* and **rps16_1-trnQ-UUG** intergenic spacers were identified as hypervariable loci. Furthermore, cluster analysis indicated that within the *Amaranthaceae* family, *C. officinalis* and *C. capitata* exhibit the strongest genetic affinity, constituting a monophyletic group with 100% bootstrap support. This finding serves to substantiate the reliability of the constructed phylogenetic tree. This outcome indicates that *C. officinalis* and *C. capitata* can be distinctly differentiated in the phylogeny that has been established from chloroplast genome sequences.

This study reports, for the first time, the complete chloroplast genomes of cultivated *Cyathula officinalis*, *C. capitata*, and *C. officinalis* × *C. capitata* from different production regions, thereby filling a significant research gap at the genomic level for these geographic germplasm resources. In summary, the findings of this study demonstrate a high degree of consistency with previous

research across different plant groups. This consistency is evident in the analysis of chloroplast genome structure, codon usage patterns, and molecular identification techniques. The results suggest that these evolutionary and diagnostic patterns may be broadly applicable. This study provides a methodological foundation for authenticating “*C. officinalis*” in the market by analyzing the chloroplast genome information, phylogenetic relationships, and species identification of *C. capitata*.

5 Conclusions

This study sequenced, assembled, and annotated the chloroplast genomes of three *Cyathula capitata* samples collected from different regions. The chloroplast genomes ranged from 151,428 to 151,436 base pairs, comprising a pair of 25,229-bp IR regions separated by an 83,523–83,524 bp LSC region and a 17,454–17,477 bp SSC region. Comparative analyses with the chloroplast genomes of *Cyathula officinalis* and *Cyathula officinalis* var. *chinensis* provided further insights into their fundamental genomic features, including long and simple repetitive sequences, codon usage bias, and IR contraction and expansion patterns. Finally, a phylogenetic tree was constructed and analyzed, incorporating these three species along with 15 other published *Amaranthaceae* species. This study offers comprehensive information on the chloroplast genomes, phylogenetic relationships, and species identification of *Cyathula officinalis* and *Cyathula capitata*, establishing a reliable method for distinguishing *C. officinalis*, *C. capitata*, and their hybrid (*C. officinalis* × *C. capitata*) in the marketplace.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repository and accession number(s) can be found in the article/Supplementary Material.

Author contributions

BG: Writing – original draft, Writing – review & editing. YZ: Writing – original draft. MT: Writing – original draft. MY: Writing – review & editing. KD: Writing – review & editing. YG: Writing – review & editing. GJ: Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This study was supported by the Scientific Research Project of the National Institutes for Food and Drug Control, China (Project No. 2023KFKT001).

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Correction note

A correction has been made to this article. Details can be found at: [10.3389/fpls.2026.1825796](https://doi.org/10.3389/fpls.2026.1825796).

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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.2026.1772997/full#supplementary-material>

SUPPLEMENTARY TABLE 1

Sampling information and GenBank accession numbers for *Naisibunia*, *Cyathula chinensis*, and *Cyathula heterophylla*.

SUPPLEMENTARY TABLE 2

Basic Characteristics of Chloroplast Genomes.

SUPPLEMENTARY TABLE 3

Chloroplast genome genes of *Cyathula capitata*, *Cyathula officinalis*, and *Cyathula officinalis* × *Cyathula capitata*. () indicates number of gene copies; Gene* denotes one intron; Gene** denotes two introns.

SUPPLEMENTARY TABLE 4

SSR distribution in chloroplast genomes. IGS, CDS, and intron denote the intergenic region, protein-coding region, and intron region of the SSR within the chloroplast genome, respectively; LSC, IRa, SSC, and IRb represent the large single-copy region, inverted repeat region B, small single-copy region, and inverted repeat region A, respectively.

SUPPLEMENTARY TABLE 5

RSCU Values of *Cyathula capitata*, *Cyathula officinalis*, and *Cyathula officinalis* × *Cyathula capitata*.

References

- Amiryousefi, A., Hyvönen, J., and Pocai, P. (2018). Irscope: an online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* 17, 17. doi: 10.1093/bioinformatics/bty220
- Azi, A. (1994). Miscellaneous discussions on ancient Yi medical historical materials. *J. Chin. Med. History* 24, 229–232.
- Beier, S., Thiel, T., Münch, T., Scholz, U., and Mascher, M. (2017). Misa-web: a web server for microsatellite prediction. *Bioinformatics* 16, 2583. doi: 10.1093/bioinformatics/btx198
- Bezeng, B., Davies, T. J., Daru, B. H., Kabongo, R. M., Maurin, O., Yessoufou, K., et al. (2017). Ten years of barcoding at the African Centre for DNA barcoding. *Genome* 60, 629–638. doi: 10.1139/gen-2016-0198
- Carbonell-Caballero, J., Alonso, R., Ibañez, V., Terol, J., Talon, M., and Dopazo, J. (2015). A phylogenetic analysis of 34 chloroplast genomes elucidates the relationships between wild and domestic species within the genus citrus. *Mol. Biol. Evol.* 8, 2015. doi: 10.1093/molbev/msv082
- Chen, C., Luo, D., Wang, Z., Miao, Y., Liu, Q., Zhao, T., et al. (2024). Complete chloroplast genomes of eight *Artemisia* species: Comparative analysis, molecular identification, and phylogenetic analysis. *Plant Biol. (Stuttgart Germany)* 26, 257–269. doi: 10.1111/plb.13608
- Chen, R., and Huang, Y. (2004). Simple repetitive sequences in genomes. *J. Shaanxi Normal University: Natural Sci. Edition* S2, 5.
- Drouin, G., Daoud, H., and Xia, J. (2008). Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. *Mol. Phylogenet. Evol.* 49, 827–831. doi: 10.1016/j.ympev.2008.09.009
- Flora of China Editorial Committee, Chinese Academy of Sciences (1979). *Flora of China* Vol. 25 (Beijing, China: Science Press).
- Frazer, K. A., Pachter, L., Poliakov, A., Rubin, E. M., and Dunchak, I. (2004). Vista: computational tools for comparative genomics. *Nucleic Acids Res.* 32, W273. doi: 10.1093/nar/gkh458
- Fu, H., Yuan, G., Zhong, J., Cao, T., Ni, L., Xie, P., et al. (2013). Environmental and ontogenetic effects on intraspecific trait variation of a macrophyte species across five ecological scales. *PLoS One* 8, e62794. doi: 10.1371/journal.pone.0062794
- Gao, B. X., Qi, J. L., Li, Q., Geng, Z., Zhong, L., Gao, C., et al. (2024). Problems and suggestions on the quality of *Cyathula Radix* based on national drug sampling survey. *Chin. Pharm. J.* 59, 275–280. doi: 10.11669/cpj.2024.03.012
- Gu, Z., Xu, S., Li, H., Tong, Y., Jiang, H., Luo, L., et al. (2024). Current status and reflections on Yi medicine research. *Asia-Pacific Traditional Med.* 20, 64–68. doi: 10.11954/ytctty.202403013
- Guo, L., Wang, X., Wang, R., and Li, P. (2023). Characterization and comparative analysis of chloroplast genomes of medicinal herb *Scrophularia ningpoensis* and its common adulterants (*Scrophulariaceae*). *Int. J. Mol. Sci.* 24, 10034. doi: 10.3390/ijms241210034
- Guo, H., Wang, L., Xu, W., Huo, Z., Yang, P., Zhang, Q., et al. (2022). The complete chloroplast genome sequence of *Cyathula officinalis* and comparative analysis with four related species. *Gene* 839, 146728. doi: 10.1016/j.gene.2022.146728
- Hu, Q., Wu, J., Fan, C., Luo, Y., Liu, J., Deng, Z., et al. (2024). Comparative analysis of codon usage bias in the chloroplast genomes of eighteen Ampelopsidae species (Vitaceae). *BMC genomic Data* 25, 80. doi: 10.1186/s12863-024-01260-8
- Jansen, R. K., Raubeson, L. A., Boore, J. L., Depamphilis, C. W., and Cui, L. (2005). Methods for obtaining and analyzing whole chloroplast genome sequences. *Methods Enzymology* 395, 348. doi: 10.1016/S0076-6879(05)95020-9
- Ji, X., Che, N., Lin, R., Chen, J., and Wu, X. (2019). Efficient ten-gene analysis of NSCLC tissue samples by next-generation sequencing. *Pathology Res. Pract.* 215, 1066–1070. doi: 10.1016/j.prp.2019.02.017
- Jianguo, Z., Yingxian, C., Xinlian, C., Ying, L., Zhichao, X., Baozhong, D., et al. (2018). Complete chloroplast genomes of papaver rhoeas and papaver orientale: molecular structures, comparative analysis, and phylogenetic analysis. *Molecules* 23, 437. doi: 10.3390/molecules23020437
- Jin, J. J., Yu, W. B., Yang, J. B., Song, Y., and Li, D. Z. (2018). GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Cold Spring Harbor Lab.* 21, 241. doi: 10.1101/256479
- Ju, S., Yue-Hong, L., Li-Mei, L., Feng, Y., and Zhao-Bin, X. (2017). Analysis on structure and phylogeny of chloroplast genomes in araliaceae species. *Chin. Traditional Herbal Drugs* 48, 5070–5075. doi: 10.7501/j.issn.0253-2670.2017.24.002
- Kahraman, K., and Lucas, S. J. (2019). Comparison of different annotation tools for characterization of the complete chloroplast genome of *corylus avellana* cv tombul. *BMC Genomics* 20, 874. doi: 10.1186/s12864-019-6253-5
- 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. *Bioinf. (Oxford England)* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199
- Kim, Y., Chung, Y., and Park, J. (2019). The complete chloroplast genome sequence of *dysphania pumilio* (r.br.) mosyakin & clemants (amaranthaceae). *Mitochondrial DNA Part B* 4, 403–404. doi: 10.1080/23802359.2018.1547156
- Kleine, T., Maier, U. G., and Leister, D. (2009). DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. *Annu. Rev. Plant Biol.* 60, 115–138. doi: 10.1146/annurev.arplant.043008.092119
- Kurtz, S., and Schleiermacher, C. (1999). Reputer: fast computation of maximal repeats in complete genomes. *Bioinformatics* 5, 426–427. doi: 10.1093/bioinformatics/15.5.426
- Li, G. D., and He, T. C. (1990). *History of Yi medicine* (Chengdu, China: Sichuan Ethnic Publishing House). doi: 10.15983/j.cnki.jsnu.2004.s2.004
- Li, C. J., Xie, X. T., Liu, H. X., Wang, R. N., and Li, D. Z. (2023). Plastome evolution in the East Asian lobelias (*Lobelioideae*) using phylogenomic and comparative analyses. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1144406
- Liu, X., Sun, Y., Zhu, G., Huang, L., and Yu, B. (2022). Complete chloroplast genomes and comparative analyses of *Hippeastrum 'milady'*, *Hippeastrum albertii* and *Hippeastrum reticulatum* (*Amaryllidaceae*). *PLoS One* 17, e0271335. doi: 10.1371/journal.pone.0271335
- Liu, C., Zeng, X. F., Yang, X. Y., Zhao, Y., Zhang, T., Zhou, S. X., et al. (2023). Comparative analysis and phylogenetic study of chloroplast genomes in *Polygonum multiflorum*. *Chin. Herbal Drugs* 54, 2558–2565. doi: 10.7501/j.issn.0253-2670.2023.08.023
- Luo, Y. J., Wang, R., Zhao, R. F., Lu, X. X., Yin, G. K., Deng, Z. J., et al. (2024). Analysis of synonymous codon usage bias in the chloroplast genome of *Davidia involucreata*. *J. Beijing Forestry Univ.* 46, 8–16. doi: 10.12171/j.1000-1522.20220199
- Margulis, L., and Goode, D. (1981). *Symbiosis in cell evolution: life and its environment on earth* Vol. 12 (Netherlands, Dordrecht: Springer).
- McFadden, G. I. (2001). Chloroplast origin and integration. *Plant Physiol.* 125, 50–53. doi: 10.1104/pp.125.1.50
- Mower, J. P., and Vickrey, T. L. (2017). Structural diversity among plastid genomes of land plants. *Adv. Botanical Res.* 0, 263–292. doi: 10.1016/bs.abr.2017.11.013
- Olmstead, R. G., and Palmer, J. D. (1994). Chloroplast dna systematics: a review of methods and data analysis. *Am. J. Bot.* 81, 1205–1224. doi: 10.1002/j.1537-2197.1994.tb15615.x
- Qian, Y., Gao-Fei, F., Zhi-Qiang, W., Li, L., Jian-Li, Z., and Qing-Jun, L. (2021). Chloroplast genome evolution in four montane zingiberaceae taxa in China. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.774482
- Rao, W. J. (2006). Research on the origins of Yi medicine. *J. Yunnan Univ. Chin. Med.* (S1), 82–86. doi: 10.19288/j.cnki.issn.1000-2723.2006.s1.031
- Shi, L., Chen, H., Jiang, M., Wang, L., Wu, X., Huang, L., et al. (2019). CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* 47, W65–W73. doi: 10.1093/nar/gkz345
- Shi, C. J., Wang, S. S., Liu, X. M., Cheng, Z. Q., Yuan, Q. H., and Song, Y. (2017). Discussion on the current status of *Cyathula officinalis*, a genuine regional medicinal material from Sichuan. *Asia-Pacific Traditional Med.* 13, 55–58. doi: 10.11954/ytctty.201722019
- Shi, N., Yuan, Y., Huang, R., and Wen, G. (2024). Analysis of codon usage patterns in complete plastomes of four medicinal *Polygonatum* species (*Asparagaceae*). *Front. Genet.* 15. doi: 10.3389/fgene.2024.1401013
- Song, C., Zhu, J., and Li, H. (2024). Complete chloroplast genomes of eight *Delphinium* taxa (*Ranunculaceae*) endemic to Xinjiang, China: insights into genome structure, comparative analysis, and phylogenetic relationships. *BMC Plant Biol.* 24, 600. doi: 10.1186/s12870-024-05279-y
- Viljoen, E., Odeny, D. A., Coetzee, M. P. A., Berger, D. K., and Rees, D. J. G. (2018). Application of chloroplast phylogenomics to resolve species relationships within the plant genus *amaranthus*. *J. Mol. Evol.* 86, 216–239. doi: 10.1007/s00239-018-9837-9
- Wei, P., Li, Y., Ke, M., Hou, Y., Aikebaier, A., and Wu, Z. (2024). Complete chloroplast genome of *Krascheninnikovia ewersmanniana*: comparative and phylogenetic analysis. *Genes* 15, 546. doi: 10.3390/genes15050546
- Williams, A. M., Friso, G., van Wijk, K. J., and Sloan, D. B. (2019). Extreme variation in rates of evolution in the plastid Clp protease complex. *Plant journal: Cell Mol. Biol.* 98, 243–259. doi: 10.1111/tpj.14208
- Xiao, X., Chen, J., Ran, Z., Huang, L., and Li, Z. (2025). Comparative analysis of complete chloroplast genomes and phylogenetic relationships of 21 sect. *Camellia* (*Camellia* L.) plants. *Genes* 16, 49. doi: 10.3390/genes16010049
- Yang, Z., Huang, Y., An, W., Zheng, X., Huang, S., Liang, L., et al. (2019). Sequencing and structural analysis of the complete chloroplast genome of the medicinal plant *lyciumchinense* mill. *Plants. (Basel Switzerland)* 8, 87. doi: 10.3390/plants8040087
- Yang, C., Zhang, N., Wu, S., Jiang, C., Xie, L., Yang, F., et al. (2023). A comparative analysis of the chloroplast genomes of three *Lonicera* medicinal plants. *Genes* 14, 548. doi: 10.3390/genes14030548

- Yang, Z., Zhao, J., Li, W., Chen, Y., Zhang, C., and Lin, H. (2025). The complete chloroplast genome of *Corispermum patelliforme* (Amaranthaceae): genome characterization and phylogenetic consideration. *Mitochondrial DNA. Part B Resour.* 10, 976–980. doi: 10.1080/23802359.2025.2561080
- Yin, D., Mu, Y., Li, X., Hua1, M., Li, X., Zhang, P., et al. (2024). Comparison of the complete chloroplast genomes in the *Astragalus*. *Legume Res.* 47, 420–427. doi: 10.18805/LRF-763
- Yin, D., Wang, Y., Zhang, X., Ma, X., He, X., and Zhang, J. (2017). Development of chloroplast genome resources for peanut (*Arachis hypogaea* L.) and other species of *Arachis*. *Sci. Rep.* 7, 11649. doi: 10.1038/s41598-017-12026-x
- Zhang, X. C. (2016). *Records of Chinese medicine with reference to western medicine* (Beijing, China: Traditional Chinese Medicine Ancient Books Publishing House).
- Zhang, J. Z., Chen, X. Y., Wu, Y. J., Li, L. M., and Liu, Y. (2020). Identification of active compounds from yi nationality herbal formula wosi influencing cox-2 and vcam-1 signaling. *Front. Pharmacol.* 11. doi: 10.3389/fphar.2020.568585
- Zhang, D., Gao, F., Jakovli, I., Zou, H., and Wang, G. T. (2020). 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, N., Huang, K., Xie, P., Deng, A., Tang, X., Jiang, M., et al. (2024). Chloroplast genome analysis and evolutionary insights in the versatile medicinal plant *calendula officinalis* l. *Sci. Rep.* 14, 9662. doi: 10.1038/s41598-024-60455-2
- Zheng, L., Duan, M., Zhang, Z., Liu, C., Wang, Z., Wang, G., et al. (2025). Comparative analysis of Chloroplast genomes in 48 samples from 5 species and 2 varieties of *Agropyron Gaertn.* (Poaceae, Triticeae). *BMC Genomics* 26, 912. doi: 10.1186/s12864-025-12026-6
- Zhou, S. C. (2020). An overview of the development history of Yi medicine in Yunnan. *J. Chin. Ethnic Med.* 26, 73–77. doi: 10.16041/j.cnki.cn15-1175.2020.04.045
- Zou, Y., and Ge, S. (2003). New generation molecular markers, SNPs and their applications. *Biodiversity Sci.* 05, 370–382.