Plant Gene 46 (2026) 100581

Contents lists available at ScienceDirect

Plant Gene

journal homepage: www.elsevier.com/locate/plantgene

Structural and evolutionary analysis of *Nepeta nuda* plastid genome reveals lineage-specific structure variation and signature of positive selection

Fatemeh Alsadat Hejazi, Aboozar Soorni *

Department of Biotechnology, College of Agriculture, Isfahan University of Technology, Isfahan 84156-83111, Iran

ARTICLE INFO

Edited by Bingye Xue

Keywords:
Nepeta nuda
Next-generation sequencing
Chloroplast genome
Lamiaceae
Phylogenetic analysis
Molecular markers

ABSTRACT

Nepeta nuda L., a medicinally significant member of the Lamiaceae family, exhibits diverse bioactive compounds and ecological adaptability, yet its chloroplast (cp) genome remained uncharacterized, limiting phylogenetic and evolutionary insights. This study aimed to sequence and annotate the complete cp genome of *N. nuda*, compare its structural and evolutionary features with those of eight congeneric species, and identify potential molecular markers for species discrimination. Using high-throughput Illumina sequencing, we assembled a 152,195 bp quadripartite cp genome encoding 131 functional genes, including key photosynthesis-related and self-replication genes. Comparative analyses revealed distinct codon usage biases (A/U preference), moderate simple sequence repeats (SSR) diversity (18 loci), and hypervariable regions (e.g., *ycf1*, *matK*, *rpl32-trnL-UAG*) suitable for phylogenetic markers. Positive selection signals were detected in seven genes (e.g., *matK*, *ycf1*, *rpoC1*), suggesting adaptive evolution. Inverted repeat (IR) boundary analysis showed conserved synteny but lineage-specific variations, while Mauve alignments highlighted a 20 kb deletion in most *Nepeta* species except *N. stewartiana* and *N. laevigata*. Phylogenomic reconstruction positioned *N. nuda* within a monophyletic *Nepeta* clade, closely related to *N. hemsleyana*, and resolved anomalous placements (e.g., *N. cataria* clustering with *Ocimum*). The *ycf1* marker outperformed *matK* in species-level resolution, though whole-genome analysis provided superior nodal support. These findings enrich genomic resources for *Nepeta*, clarify evolutionary relationships, and identify candidate markers for species identification, with implications for conservation, taxonomy, and medicinal applications leveraging the species' adaptive traits.

1. Introduction

Nepeta nuda L. belongs to the Lamiaceae family, specifically to the subfamily Nepetoideae and tribe Mentheae, which comprises the largest number of genera and species within Nepetoideae and Lamiaceae (Aćimović et al., 2020). This perennial herbaceous plant is characterized by numerous erect stems that can grow 50 to 100 cm tall, with oblong-lanceolate lower leaves and ovate upper leaves. Its flowers are organized in lax or dense spike-like verticillasters that bloom from June to August, depending on the altitude (Aćimović et al., 2022; Aćimović et al., 2020). *Nepeta nuda* exhibits a wide geographical distribution across Europe and Asia, naturally growing in forest clearings, meadows, and ruderal habitats alongside roads at montane and subalpine altitudes up to 2100 m, often dominating other species in the surrounding area (Aćimović et al., 2020; As'ov, 2006; Petrova et al., 2022). In Bulgaria, it is found in all floristic regions up to 1900 m a.s.l. and is known as “naked (or hairless) catmint,” referring to its naked or sparse short hairy stem and leaves (Petrova et al., 2022). Additionally, *N. nuda* is distributed throughout Iran, where populations exhibit various nepetalactone stereoisomers that constitute a large proportion of their essential oils (Narimani et al., 2017).

Nepeta nuda possesses significant medicinal properties that have been utilized in traditional folk medicine. It is traditionally used to remedy gastrointestinal problems, respiratory diseases, cystitis, prostate gland inflammation, dysmenorrhea, and to treat wounds and mastitis. Studies have demonstrated that extracts of *N. nuda* exhibit antiviral, antioxidant, antibacterial, antiparasitic, and anticancer properties (flavonoids and phenolic acids) (Alim et al., 2009; Angelova et al., 2016; Gormez et al., 2013; Kabalay et al., 2018; Sarikurkcu et al., 2018). Phytochemical analyses identify two major classes of bioactive compounds: iridoids (including nepetalactones, 1,5,9-epi-deoxyloganic acid, and nepetariaside) and phenolics (particularly rosmarinic acid and flavonoids), which collectively exhibit antioxidant, antimicrobial, and antiviral properties. Essential oil profiles vary considerably among

* Corresponding author.
E-mail address: soorni@iut.ac.ir (A. Soorni).

https://doi.org/10.1016/j.plgene.2026.100581
Received 30 October 2025; Received in revised form 28 January 2026; Accepted 5 February 2026
Available online 7 February 2026

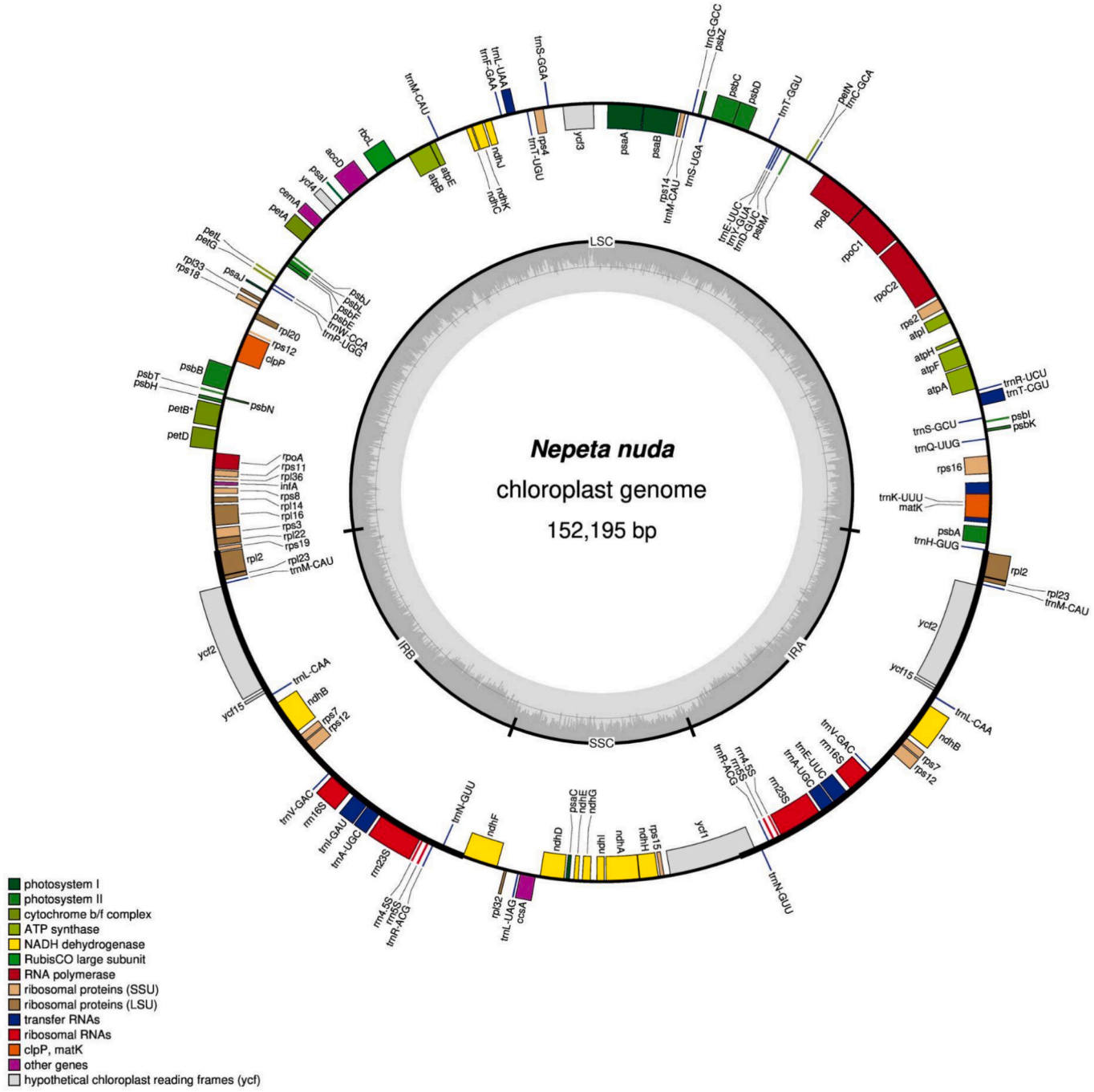

Fig. 1. Circular representation of the *Nepeta nuda* chloroplast genome. The map, constructed using OGDRAW, displays the quadripartite structure with annotated genes color-coded by functional category. Outer rings indicate gene positions in the large single-copy (LSC), small single-copy (SSC), and inverted repeat (IR) regions.

populations, with four distinct chemotypes recognized: 1) 1,8-cineole-dominant, 2) nepetalactone-rich, 3) mixed (germacrene D/1,8-cineole/nepetalactone), and 4) non-specific chemotypes. Notably, environmental factors strongly influence volatile composition, with temperature positively correlating with 1,8-cineole accumulation while negatively affecting nepetalactone production. Organ-specific metabolic profiling reveals higher constitutive levels of defensive compounds in inflorescences compared to leaves, suggesting prioritized resource allocation to reproductive structures (Aćimović et al., 2022; Aničić et al., 2021; Jamzad et al., 2003; Petrović et al., 2024).

As mentioned, the genus *Nepeta* exhibits significant diversity, with numerous species having medicinal and ecological importance. Accurate identification and phylogenetic analysis of these species are crucial for understanding their evolutionary relationships and potential applications. In this regard, cp genome serves as a powerful tool for species identification and phylogenetic studies due to its high conservation, maternal inheritance, and moderate mutation rate (Tonti-Filippini et al., 2017). The cp genome is a circular, double-stranded DNA molecule typically ranging from 120 to 160 kb in size, characterized by a quadripartite structure consisting of a large single-copy (LSC) region, a small single-copy (SSC) region, and two inverted repeat (IR) regions (Akrami et al., 2025; Soorni and Golchini, 2025). It encodes essential genes involved in photosynthesis, transcription, and translation, making it a valuable resource for evolutionary and phylogenetic studies. Due to its conserved nature, the cp genome is widely used in species identification, population genetics, and resolving taxonomic uncertainties.

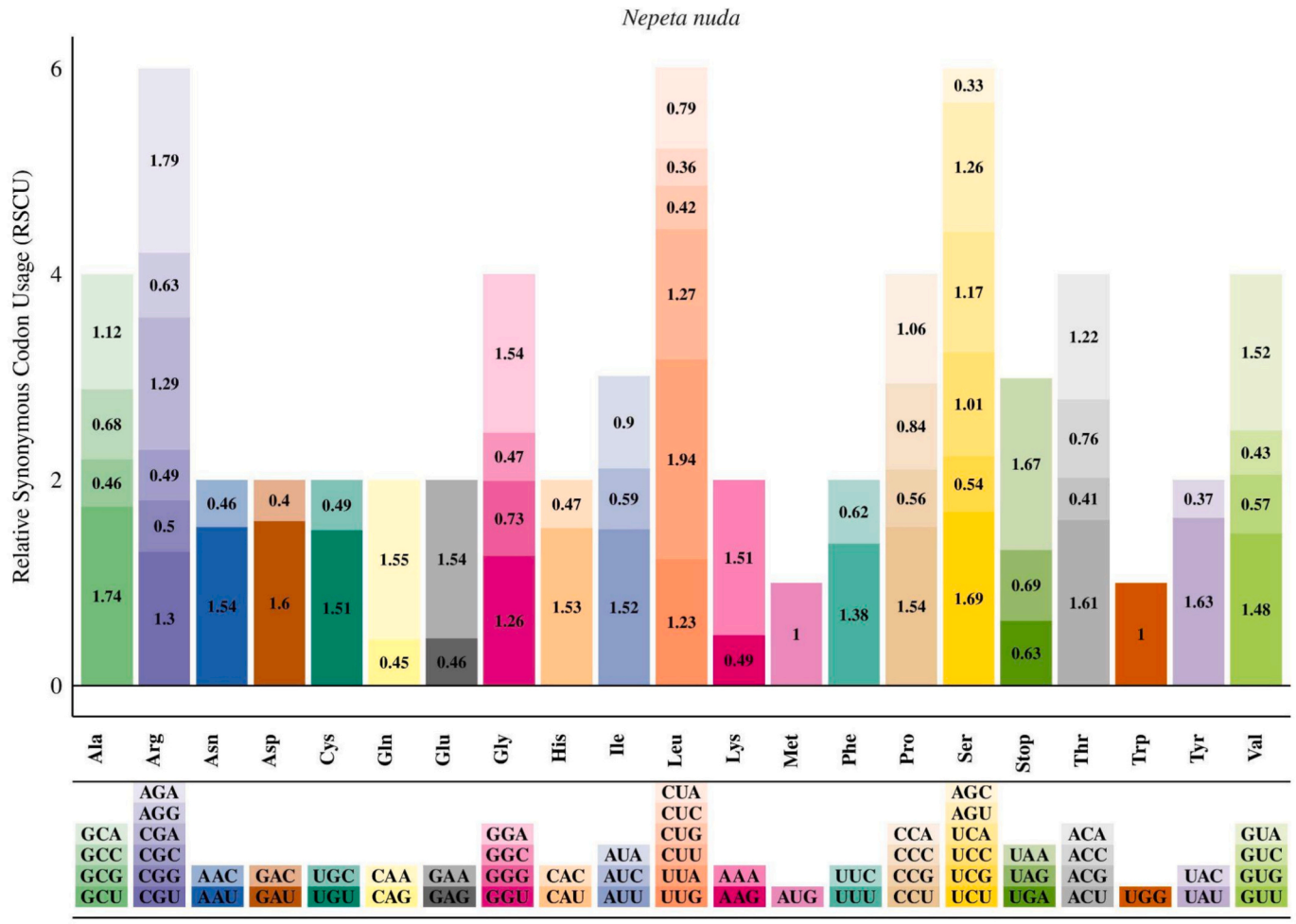

Fig. 2. The bar plot of the Relative Synonymous Codon Usage (RSCU) values for each amino acid of *N. nuda*. Codons are color-coded, and their corresponding amino acids are labeled below the plot.

Additionally, hypervariable regions within the cp genome can serve as molecular markers for distinguishing closely related species (Daniell et al., 2016; Dobrogojski et al., 2020; Zhang et al., 2023).

Several *Nepeta* species have had their cp genomes sequenced, providing insights into their genetic structure and evolutionary relationships. For instance, *Nepeta bracteata* possesses a cp genome of 151,588 bp, with 130 genes, including 87 protein-coding genes and 35 tRNA genes (Chen et al., 2024). Comparative analysis revealed hypervariable regions such as *ndhH-rps15* and *trnH-GUG-psbA*, which could serve as molecular markers for species identification (Chen et al., 2024). Similarly, *Nepeta cataria* has a cp genome of 152,339 bp, encoding 132 genes, and phylogenetic analysis indicated its close relation to *Callicarpa nudiflora* (Zhou et al., 2020). Another study on *N. cataria* reported a slightly larger cp genome (153,526 bp) and identified its close phylogenetic relationship with *Nepeta racemosa* (Luo, 2019).

Additionally, multiple *Nepeta* species from Tibet have been sequenced, including *Nepeta dentata*, *N. hemsleyana*, *N. laevigata*, and *N. thomsonii*. These species exhibit typical quadripartite cp genome structures, with sizes ranging from 151,893 bp (*N. hemsleyana*) to 152,312 bp (*N. dentata*). All four species contain 88 protein-coding genes, 37 tRNA genes, and 8 rRNA genes, consistent with other Lamiaceae members. Phylogenetic analyses suggest close evolutionary relationships among these species, with *N. dentata* and *N. laevigata* forming a distinct clade, while *N. hemsleyana* and *N. thomsonii* exhibit slight divergences in gene arrangement (Niu et al., 2023).

Despite the availability of cp genome data for several *Nepeta* species, *N. nuda* remains unsequenced, creating a gap in understanding its phylogenetic placement and genetic distinctiveness. To address this, we sequenced the complete cp genome of *N. nuda* and conducted comparative analyses with other *Nepeta* species. Our findings contribute to the genomic resources of the genus and provide insights into the evolutionary relationships and potential medicinal applications of *N. nuda.*

2. Material and methods

2.1. DNA isolation, sequencing, assembly, and annotation of cp genome

Fresh leaf specimens from *N. nuda* were collected from *Heris* County, *East Azerbaijan, Iran*. No specific permissions were required for the collection of these plant samples. The plant material was formally identified and is available at the Herbarium of the Ferdowsi University of Mashhad (FUMH) under voucher number of 6302. Genomic DNA was extracted from 100 mg of flash-frozen tissue using the DNeasy Plant Mini Kit (QIAGEN, Germany), followed by stringent quality assessment via 1% agarose gel electrophoresis and NanoDrop 2000c spectrophotometry (A260/A280 $\geq$ 1.8; A260/A230 $\geq$ 2.0). High-purity DNA was processed into Illumina TruSeq libraries (350 bp inserts) and sequenced on a HiSeq 2000 platform (2 $\times$ 150 bp paired-end reads). Raw reads were quality-filtered using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and Trimmomatic (Bolger et al., 2014), then assembled into complete cp genomes via GetOrganelle v1.7.7.1 (Jin et al., 2020a), with assembly validation performed in Bandage (Wick et al., 2015). Annotation was conducted using CPGAVAS2 (Shi et al., 2019), and GeSeq (Tillich et al., 2017), while circular genome

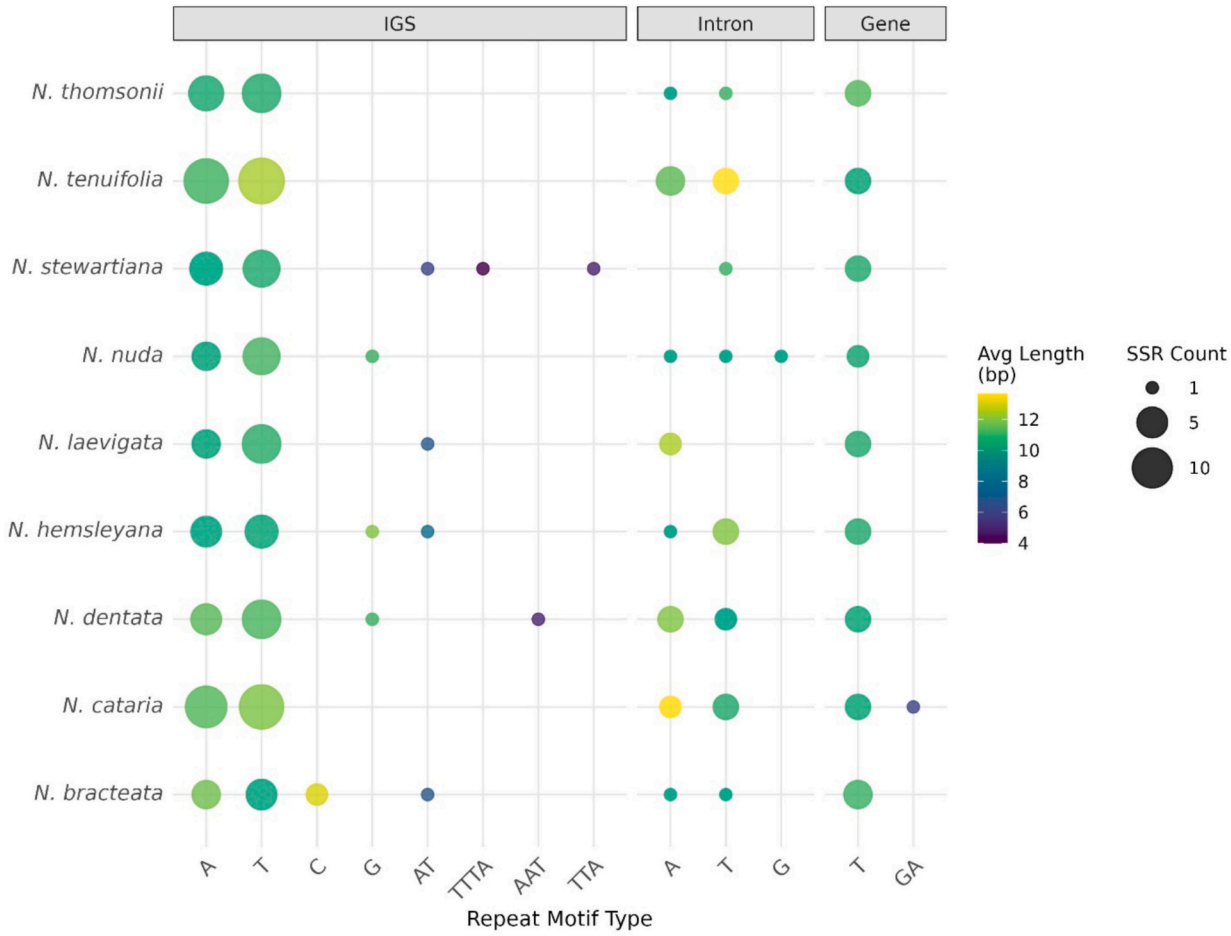

Fig. 3. Comparative analysis of simple sequence repeat (SSR) characteristics across *Nepeta* species. Bubble plot showing the abundance and average length of SSRs categorized by repeat motif type (x-axis) and genomic location (facets). Each bubble represents a specific SSR type within a species, with bubble size proportional to SSR frequency and color intensity (viridis scale) indicating mean repeat length (bp). Color gradients highlight species-specific differences in average SSR length, with warmer tones indicating longer repeats.

visualizations were generated with OGDRAW (Greiner et al., 2019).

2.2. Genome feature characterization

Synonymous codon preference was evaluated through Relative Synonymous Codon Usage (RSCU) analysis, with calculations performed in MEGA6 (Tamura et al., 2013) and subsequent visualization of codon bias patterns achieved through the RSCU-Plot Shiny app (https://pcg-lab.shinyapps.io/RSCU-Plot/)(Akrami et al., 2025; Diani Gohar and Soorni, 2025; Hejazi et al., 2025; Soorni and Golchini, 2025). Microsatellite identification was performed using CPStools (Huang et al., 2024), with the following parameters: mononucleotide repeats required ≥10 units, dinucleotide repeats ≥6 and trinucleotide repeats ≥5 units, while tetra-, penta-, and hexanucleotide motifs required ≥4 repeat units.

2.3. Sequence divergence and selective pressure analysis

To assess sequence polymorphism across coding regions, we computed nucleotide diversity (*Pi*) values for all annotated genes using specialized cp genome analysis software, CPStools (Huang et al., 2024). Besides, selective pressure on consensus protein-coding genes across nine *Nepeta* species genomes was assessed using the EasyCodeML software (Gao et al., 2019). Site models were employed to compare four model pairs (M0 vs. M3, M1a vs. M2a, M7 vs. M8, and M7a vs. M8a), and likelihood ratio tests (LRTs) were conducted with a significance threshold of $p < 0.05$. The ratio of nonsynonymous (Ka) to synonymous (Ks) substitutions (denoted as ω) was calculated for each gene. The resulting ω values, in conjunction with LRT outcomes, were used to infer patterns of selective pressure acting on specific amino acid sites.

The integrated analyses of SSR distribution patterns, nucleotide polymorphism landscapes, and positive selection analysis were conducted through comparative examination of eight additional complete cp genomes representing distinct *Nepeta* species: *N. bracteata* (OQ689450), *N. dentata* (OP186463), *N. laevigata* (OP186465), *N. stewartiana* (MT733874), *N. thomsonii* (OP186466), *N. cataria* (MT663220), *N. hemsleyana* (MW387500), and *N. tenuifolia* (NC_053707), which were retrieved from NCBI's Organelle Genome Resources database to enable systematic evaluation of genomic structural variation and evolutionary dynamics across the genus.

2.4. Comparative structural analysis of Nepeta cp genomes

A comprehensive comparison of nine *Nepeta* cp genomes was conducted to assess structural variations at the genomic level. The expansion and contraction dynamics of inverted repeat (IR) regions, including LSC, IRb, SSC, and IRa, were analyzed and visualized using IRScope (Amiryousefi et al., 2018), with particular attention to the four junction points (JLB, JSB, JSA, and JLA) between IR and single-copy regions. To further evaluate genome architecture, the Mauve alignment algorithm (Darling et al., 2004) was employed to detect structural rearrangements and assess collinearity, revealing conserved synteny while identifying potential lineage-specific inversions or shifts in gene order. This dual-

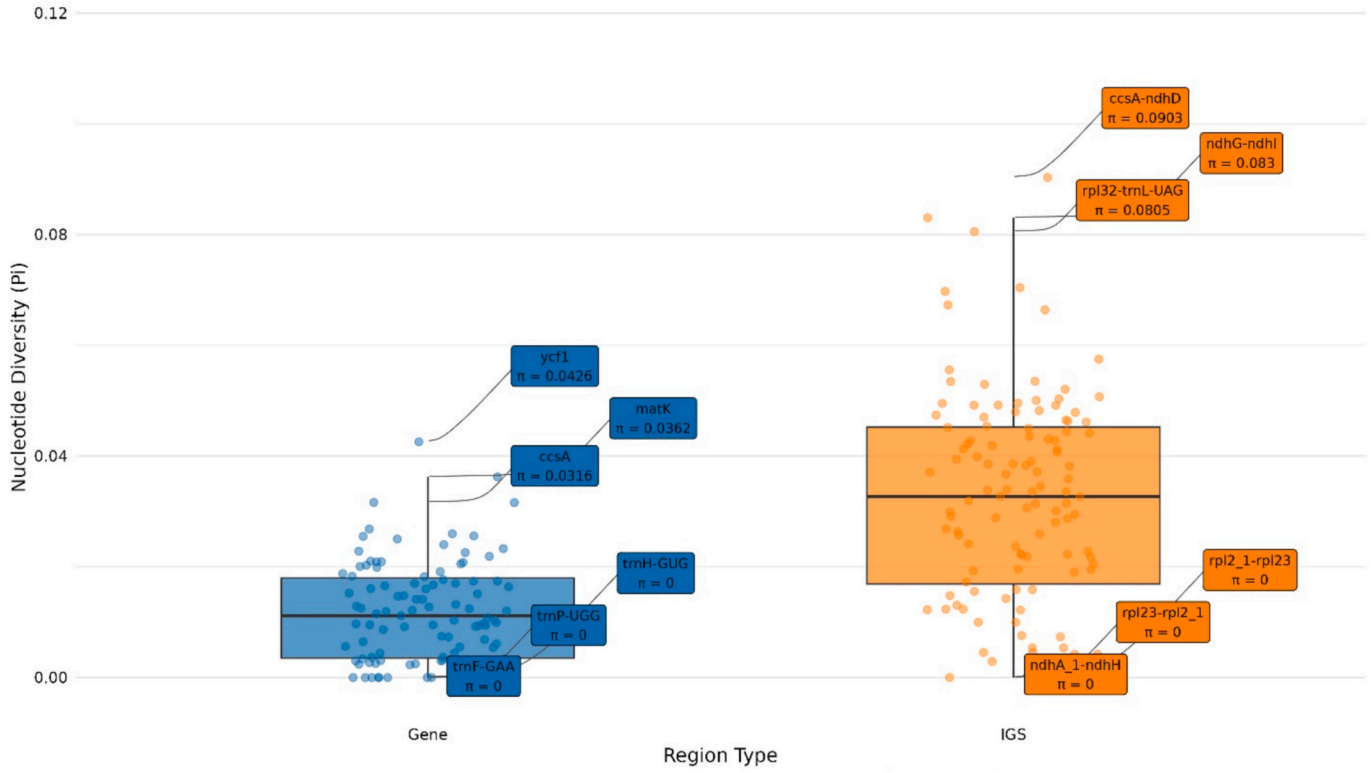

Fig. 4. Comparative analysis of nucleotide diversity (Pi) in *Nepeta* chloroplast genomes. Boxplots depict Pi values for coding genes (blue) and intergenic spacer regions (IGS; orange), with jittered points showing individual values. The top and bottom three most extreme values for each region type are labeled with feature names and exact Pi values. White diamonds indicate mean diversity values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

method approach provided robust insights into the evolutionary dynamics shaping *Nepeta* cp genomes, highlighting both conserved features and structural divergence that may underlie functional adaptation.

2.5. Phylogenetic reconstruction and marker evaluation

To resolve evolutionary relationships within *Nepeta*, we performed phylogenetic analyses following established methodologies (Akrami et al., 2025; Soorni and Golchini, 2025). Whole cp coding sequences from our sequenced taxa were combined with publicly available genomes from representative Lamiaceae genera (*Salvia*, *Mentha*, *Lavandula*, *Thymus*, *Origanum*, *Ocimum*, and *Teucrium*), using *Ajuga bracteosa* as the outgroup. Sequence alignment was conducted with MUSCLE v3.8.1551 (Edgar, 2004) under default parameters, followed by refinement using trimAl v1.4 (Capella-Gutiérrez et al., 2009) to eliminate ambiguously aligned regions (retention threshold: 95% site occupancy; minimum sequence similarity: 0.1%). The curated alignments were concatenated into a supermatrix using SequenceMatrix (Vaidya et al., 2011). Maximum likelihood inference was implemented in IQ-TREE (Nguyen et al., 2015) with the GTR + Γ model, selected via ModelFinder, and nodal support was assessed with 1000 ultrafast bootstrap replicates. The final topology was visualized and annotated in iTOL (Letunic and Bork, 2021). To evaluate the discriminatory power of hypervariable markers, we extracted *matK* and *ycf1* loci from the same taxonomic sampling. These regions were analyzed using identical alignment and phylogenetic protocols to assess their resolution at various taxonomic levels.

3. Results

3.1. Cp genome assembly and annotation

The cp genome of *N. nuda* was fully assembled, revealing a typical quadripartite structure with a total length of 152,195 bp, consisting of a LSC region (83,380 bp), a small SSC region (17,585 bp), and two IR regions (IRa/IRb, each 25,615 bp)(Fig. 1). Genome annotation identified 131 functional genes, including 85 protein-coding genes, 8 ribosomal RNAs (rRNAs), and 36 transfer RNAs (tRNAs), along with two genes containing intronic sequences (*ndhB*, *rpl2* duplicated in IR regions). Photosynthesis-related genes constituted the largest functional category, comprising five subunits of photosystem I (*psaB*, *psaA*, *psaI*, *psaJ*, *psaC*), 15 subunits of photosystem II (*psbA-psbN*), 12 NADH dehydrogenase genes (*ndhA-ndhK*), six cytochrome *b/f* complex components (*petA-petL*), and six ATP synthase genes (*atpA-atpI*). Self-replication machinery included ribosomal proteins (*rpl*, *rps*), RNA polymerase subunits (*rpoA-rpoC2*), and rRNA/tRNA genes. Additionally, several other essential genes were identified, such as the maturase *matK*, protease clpP, and conserved open reading frames (*ycf1-ycf4*). Notably, *ycf1* and *ycf2* were duplicated in the IR regions, while *trnE-UUC* and *trnM-CAU* exhibited multiple copies.

3.2. Genomic feature comparison

The RSCU analysis of *N. nuda* revealed distinct codon preference patterns, with a strong bias toward A/U-ending codons across most amino acids (Fig. 2). The species showed particularly high preference for AUU (Ile, RSCU = 1.52), UUA (Leu, 1.94), and ACU (Thr, 1.61), while strongly avoiding C-ending codons like CUC (Leu, 0.36), GGC (Gly, 0.47), and UGC (Cys, 0.49). Stop codon usage favored UAA (1.67) over UAG (0.69) and UGA (0.63).

Table 1
The results of positive selective pressure analysis in M7 vs. M8 model.

Gene	LnL	ω (M8)	LRT *p*-value	Positively Selected Sites
matK	−3110.78	10.56	2.84E-07	46 P 0.698,81 L 0.948,92 Q 0.548,180 V 0.644,189 L 0.580,203 G 0.501,280 R 0.714,301 R 0.735,308 E 0.618,310 G 0.941,329 S 0.634,340 I 0.728,365 L 0.604,392C 0.733,457 L 0.675,476 Q 0.745,485 L 1.000**,486 R 0.998**
ccsA	−1944.64	10.32	9.49E-05	169 R 0.925,170 R 0.900,183 A 0.999**,190 V 0.786,191 Q 0.633
ndhF	−4596.21	3.17	1.19E-03	15 V 0.651,72 V 0.505,97 S 0.598,106 T 0.931,123 L 0.531,235 A 0.791,317 R 0.706,463 K 0.920,502 S 0.856,514 Q 0.781,520 G 0.699,521 M 0.916,523 R 0.952*,535 A 0.979*,536 L 0.667,571 F 0.548,579 V 0.596,585 L 0.581,652 N 0.855,682 T 0.681,699 V 0.759,710 N 0.583,722 F 0.698
ycf1	−12,623.47	7.75	3.40E-08	276 E 0.508,352 K 0.533,354 S 0.993**,355 F 0.940,356 F 0.526,417 V 0.597,444 T 0.584,492 V 0.618,522 I 0.545,537 P 0.744,576 P 0.635,580 P 0.642,596 G 0.681,599 K 0.984*,600 I 0.746,601 N 0.637,617 N 0.510,620 S 0.811,633 W 0.655,673 T 0.596,674 E 0.569,754 V 0.748,899 S 0.626,918 W 0.651,980 I 0.544,981 G 0.723,1033 R 0.646,1035 W 0.668,1107 S 0.645,1116 I 0.554,1166 S 0.900,1284 L 0.997**,1287 Q 0.668,1288 R 0.586,1324 N 0.992**,1370 T 0.520,1382H 0.730,1398 P 0.628,1481 S 0.905,1533H 0.625,1544 R 0.606,1546 S 0.610,1564 Q 0.740,1566 E 0.566,1712 T 0.509,1716 D 0.962*
rpoC1	−3530.37	42.14	4.30E-06	145 V 0.978*,209 D 0.722,421 Q 0.798,433 K 0.755,473 V 0.743
ndhD	−2644.06	7.43	6.47E-03	1 M 0.939,148 G 0.603,261 Q 0.734,411 G 0.708,412 R 0.688
ndhC	−545.52	6.13	5.68E-03	26 L 0.987*,98 L 0.912

*: means $P < 0.05$, **: means $P < 0.01$.

The analysis of SSRs in the cp genomes of nine *Nepeta* species revealed distinct patterns of SSR distribution, with mononucleotide A/T repeats being the most abundant, followed by dinucleotide and trinucleotide motifs, predominantly located in intergenic spacer regions (IGS), with fewer occurrences in introns and coding sequences (Fig. 3). Notably, *N. nuda* exhibited a moderate SSR frequency (18 loci) characterized by predominant A/T repeats, including a T11 in the *rpoC2* gene - a conserved hotspot across species - and a T10 in the *atpB* gene, mirroring patterns observed in *N. cataria* and *N. dentata*. While the *rpl32-trnL-UAG* locus displayed compound T/G repeats in *N. hemsleyana* and *N. dentata, N. nuda* featured a simpler A11 motif, indicating lower variability at this locus. Unique to *N. nuda* was a G10 repeat in the *petB*_1-*petB*_2 intron, shared only with *N. bracteata's* C15 repeat in *psbZ-trnG-GCC*, and an extended T15 repeat in *rpl32-trnL-UAG*, distinguishing it from other species, which typically had shorter repeats (T10-T12) at this locus.

3.3. Sequence divergence and selective pressure analysis

Analysis of nucleotide diversity (Pi) across nine Nepeta species revealed pronounced heterogeneity in sequence variation among cp genes and IGSs (Fig. 4). Among genes, ycf1 (Pi = 0.0426) and matK (Pi = 0.0362) showed the highest diversity, consistent with their known roles in adaptive evolution, while ribosomal genes (e.g., rps3, Pi = 0.0240) and photosystem components (e.g., psbA, Pi = 0.0075) were more conserved. Notably, tRNA genes (e.g., trnH-GUG, Pi = 0) and some critical housekeeping genes (e.g., rbcL, Pi = 0.0132) exhibited minimal variation, likely due to functional constraints. The IGS regions exhibited significantly higher diversity (mean Pi = 0.038 ± 0.019) compared to coding genes (mean Pi = 0.014 ± 0.009), with the most polymorphic IGS being rpl32-trnL-UAG (Pi = 0.0805), ccsA-ndhD (Pi = 0.0903), and ndhG-ndhI (Pi = 0.0830), suggesting these regions as potential hotspots for evolutionary divergence. The elevated diversity in IGS regions and select genes (ycf1, matK) underscores their utility for phylogenetic and m Figure 1population genetic studies in Nepeta, while conserved loci may serve as robust markers for deeper evolutionary comparisons.

The analysis of 77 cp genes across nine *Nepeta* species revealed seven genes exhibiting strong signatures of positive selection ($P < 0.01$), characterized by elevated ω (dN/dS) ratios and multiple sites under diversifying selection (Table 1). The maturase K gene (*matK*) showed the strongest evidence ($\omega = 10.56$, $P = 2.84 \times 10^{-7}$), with 18 positively selected sites including two with near-certain probability (485 L, 486R), suggesting adaptive evolution in this RNA splicing-related gene. The *ycf1* gene ($\omega = 7.75$, $P = 3.40 \times 10^{-8}$) displayed particularly widespread selection with 34 sites under positive selection, including three with extremely high probability (354S, 1284 L, 1324 N). The *ccsA* gene ($\omega = 10.32$, $P = 9.49 \times 10^{-5}$) showed strong selection at site 183 A, potentially affecting cytochrome *c* biogenesis. Among NADH dehydrogenase genes, *ndhF* ($\omega = 3.17$) and *ndhD* ($\omega = 7.43$) contained multiple selected sites (523R, 535 A in ndhF; 1 M* in ndhD), while *ndhC* showed two highly significant sites (26 L, 98 L). The RNA polymerase subunit *rpoC1* exhibited an exceptionally high ω value (42.14, $P = 4.30 \times 10^{-6}$) with strong selection at 145 V**, indicating potential adaptive changes in transcriptional regulation.

3.4. IR region expansion and contraction in cp genomes

The boundaries of IR regions in cp genomes are known to be dynamic and play a critical role in genome size variation and structural evolution among angiosperms. In the present analysis of nine *Nepeta* species, IR boundary dynamics were examined by comparing the positions of genes at the four junctions: LSC/IRb (JLB), IRb/SSC (JSB), SSC/IRa (JSA), and IRa/LSC (JLA), with particular attention to the behavior of *ycf1*, *ndhF*, *rps19*, *rpl2*, *trnN-GUU*, and *trnH-GUG* (Fig. 5).

Across all *Nepeta* species, several conserved features were observed.

The gene *ndhF* consistently spans the JSB boundary, with a minor portion (32–76 bp) located in IRb and the majority (~2156–2200 bp) in the SSC region. Similarly, all species possess a *ycf1* gene located at the JSA boundary, with the majority of the gene (~4480–4494 bp) in SSC and a smaller portion (~1056–1100 bp) extended into IRa. Furthermore, the *trnN-GUU* gene was uniformly located within IRa in all species, at a conserved distance from JSA (~1384–1430 bp), suggesting stability in the IRa extension. The *trnH-GUG* gene was also consistently located just inside the LSC downstream of JLA, at a narrow range of distances (5–13 bp), except for *N. laevigata*, where it was not detected, potentially due to annotation limits or real absence. Despite this general conservation, species-specific differences highlight the subtle expansion or contraction of IR boundaries. Notably, *N. bracteata*, *N. cataria*, and *N. nuda* lacked a duplicated *ycf1* pseudogene at JSB, whereas all other species exhibited partial duplication of *ycf1* in IRb (~1056–1179 bp), indicating variation in IRb expansion. In species like *N. dentata*, *N. hemsleyana*, *N. stewartiana*, and *N. thomsonii*, this duplicated segment reflects a typical IR expansion pattern resulting in *ycf1* pseudogene formation. Boundary variation is also evident at the JLB junction. In most species, *rps19* spans the LSC/IRb boundary, with ~234–238 bp in LSC and ~ 41–45 bp in IRb. However, *N. tenuifolia* uniquely has *rps19* entirely in LSC, separated by a 67 bp gap from IRb. Similarly, *rpl2* is typically located ~102–107 bp from the JLB, but some structural variations were observed: *N. laevigata* lacks *rpl2* and *rps19* at the junction, instead containing *rpl22* with a 112 bp gap to IRb, and *N. hemsleyana* and

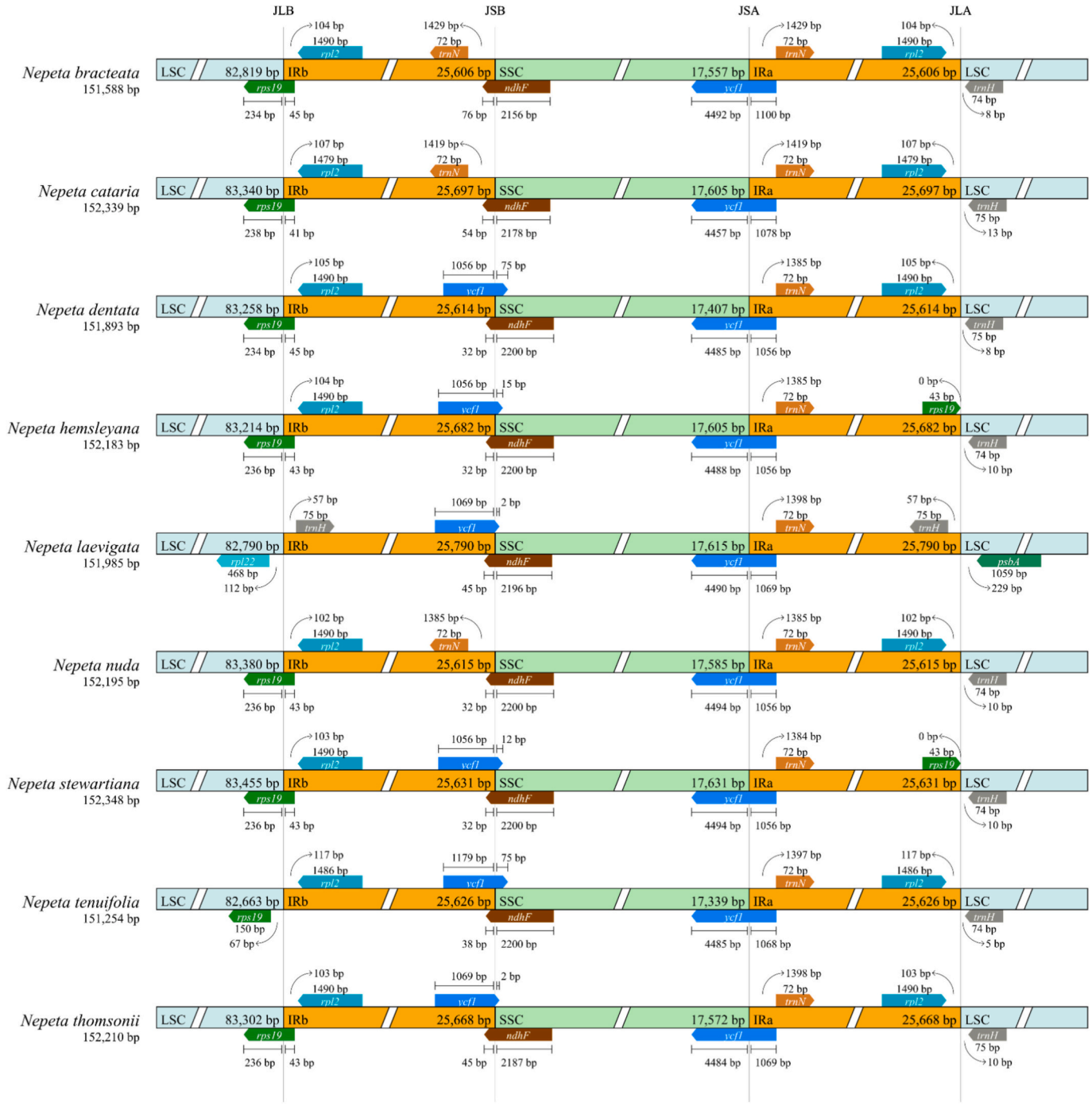

Fig. 5. Comparison of the junctions between the LSC, SSC, and IR regions among *Nepeta* species chloroplast genomes. The gene arrangements and lengths at four junction sites (JLB: LSC/IRb, JSB: IRb/SSC, JSA: SSC/IRa, JLA: IRa/LSC) are shown for each species.

N. stewartiana feature duplicated or missing *rpl2* within IRa, suggesting IR recombination or gene loss.

Among all species, *N. nuda* exhibits a moderate and stable IR configuration, lacking the *ycf1* pseudogene at JSB but maintaining all other standard boundary features. It shares most of its IR structural characteristics with *N. bracteata* and *N. cataria*, indicating a common pattern of IR contraction relative to the other taxa.

3.5. Structural analysis of cp genomes in Nepeta species

Comparative analysis of locally collinear blocks (LCBs) and gene order conservation in the cp genomes of *Nepeta* and related species revealed strong overall synteny (Fig. 6). Using the progressive alignment algorithm in Mauve, we identified a single large conserved collinear block (>150 kb) in all species. However, a notable terminal region (~130–150 kb) within this block is present only in *N. stewartiana* and *N. laevigata*, but absent in the other *Nepeta* species examined. This pattern suggests a lineage-specific deletion of approximately 20 kb in most *Nepeta* species, while *N. stewartiana* and *N. laevigata* retained the ancestral structure. Overall, the cp genomes show high conservation in gene order, punctuated by this terminal variation.

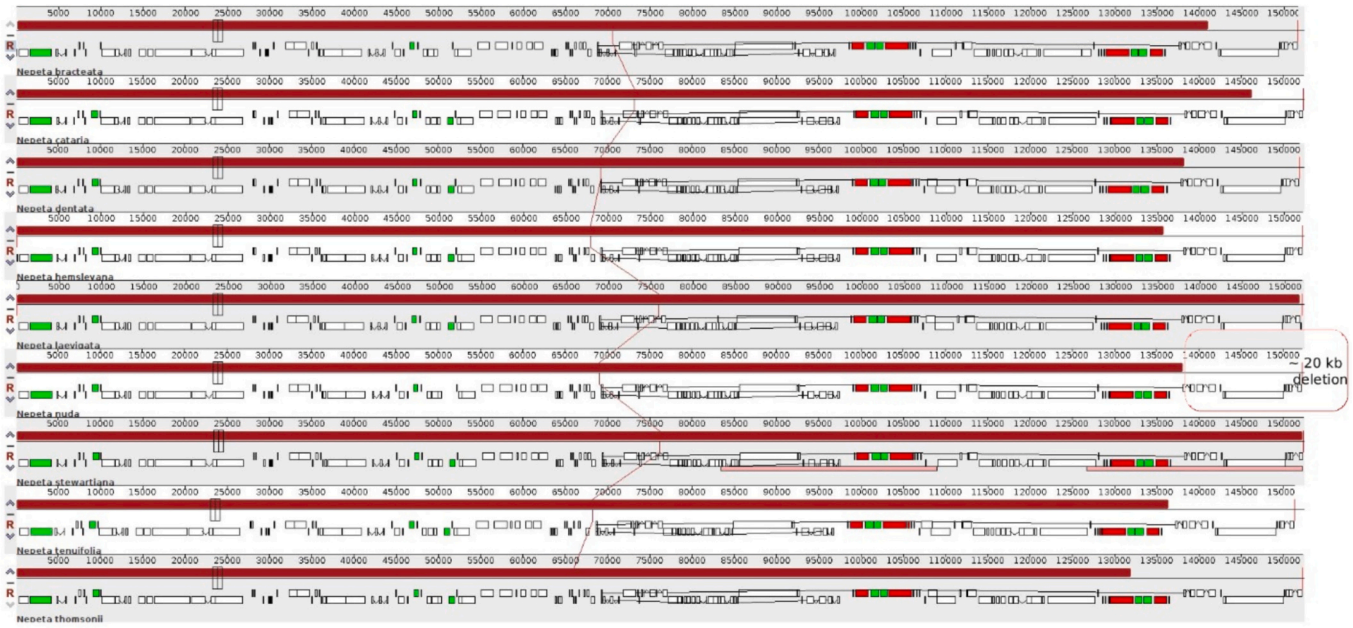

Fig. 6. Comparative alignment of chloroplast genomes from *Nepeta* species and closely related taxa using Mauve. Each horizontal row represents the genome of a species, and colored blocks indicate locally collinear blocks (LCBs) identified by the alignment. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.6. Phylogenetic relationships

The phylogenetic reconstruction based on complete cp genome sequences elucidated the evolutionary relationships among *Nepeta* species and related Lamiaceae taxa (Fig. 7). *N. tenuifolia* emerged as an early-diverging lineage, positioned as sister to all other examined *Nepeta* species. A strongly supported monophyletic clade (100% bootstrap) comprised *N. bracteata*, *N. laevigata*, *N. stewartiana*, *N. thomsonii*, *N. nuda*, *N. hemsleyana*, and *N. dentata*, with *N. bracteata* and *N. laevigata* forming a distinct sister pair. The phylogenetic placement of *N. cataria* outside the core *Nepeta* clade, instead clustering with *Ocimum* species, suggested either convergent evolution or the need for taxonomic reevaluation of this species. Comparative analysis revealed that *Nepeta* species showed closer phylogenetic affinity to *Ocimum*, *Lavandula*, and *Salvia* than to *Mentha* or *Thymus*, providing new insights into the complex evolutionary history of Lamiaceae.

To evaluate the potential of *matK* and *ycf1* as DNA barcodes for species discrimination and phylogenetic reconstruction, we performed separate maximum-likelihood analyses using these highly variable cp markers (Fig. 8). Both genes generally supported the monophyly of major genera, though their resolution for species-level relationships within *Nepeta* showed notable differences compared to the whole cp genome phylogeny.

The *matK*-based tree placed *N. tenuifolia* in a derived position, sister to a clade containing *N. dentata*, *N.* hemsleyana, and a subclade comprising *N. bracteata*/*N. laevigata* and *N. stewartiana*/*N. thomsonii*, albeit with moderate bootstrap support (50–98%). Of particular interest, *N. nuda* occupied an intermediate position (61% support) between these groups, suggesting a potential transitional evolutionary role. In contrast, the *ycf1* phylogeny provided stronger support (100% bootstrap) for *N. nuda* as a distinct lineage closely related to *N. hemsleyana*, better matching the whole-genome topology.

Both markers consistently placed *N. cataria* outside the core *Nepeta* clade, clustering it with *Ocimum* species (57–88% support), reinforcing observations from the whole-genome analysis. The *ycf1* tree showed superior resolution for *Nepeta* relationships overall, particularly for our focal species *N. nuda*, though neither single-gene analysis fully replicated the robust nodal support of the complete cp genome.

These results demonstrate that while *ycf1* performs better than *matK* for resolving *Nepeta* phylogeny, particularly for key species like *N. nuda*, multi-gene approaches remain preferable for comprehensive phylogenetic reconstruction. The persistent anomalous placement of *N. cataria* across all analyses suggests this species may require taxonomic re-evaluation or additional genomic investigation to clarify its evolutionary history.

4. Discussion

The cp genome of *N. nuda* exhibited a typical quadripartite structure, consistent with other *Nepeta* species, including *N. bracteata*, *N. cataria*, *N. hemsleyana*, and *N. dentata* (Bautista et al., 2022; Chen et al., 2024; Niu et al., 2023; Zhou et al., 2020). This structural conservation is a hallmark of land plant chloroplast genomes, as demonstrated by extensive comparative studies in Lamiaceae and other angiosperms (Daniell et al., 2016; Jansen et al., 2007). The minimal length variation among these species underscores the strong evolutionary constraint on plastid genome architecture, a phenomenon attributed to the essential roles of chloroplasts in photosynthesis and other metabolic processes (Ruhlman et al., 2010). However, subtle differences in the size of IR regions and single-copy (SC) regions contribute to these variations, likely due to expansion/contraction events at IR boundaries (Chen et al., 2024; Niu et al., 2023). Such structural dynamics are common in plastid genomes and have been extensively documented in Lamiaceae, where IR boundary shifts are often lineage-specific (Du et al., 2022; Parks et al., 2009; Qian et al., 2013; Soorni and Golchini, 2025; Wolf et al., 2011). For example, the contraction of IR regions into SSC regions has been linked to the loss or pseudogenization of *ycf1* in some lineages, a pattern observed across diverse plant families (Jin et al., 2020b; Wang et al., 2018).

The cp genome of *N. nuda* encodes 131 functional genes, a count consistent with the range observed in congeneric species (*N. bracteata:* 130; *N. cataria*: 132) (Chen et al., 2024; Zhou et al., 2020). These subtle interspecific variations in gene content reflect pervasive evolutionary dynamics shaping plastome architecture across Lamiaceae. Three primary mechanisms likely underlie these discrepancies: IR boundary shifts, lineage-specific gene losses, and annotation heterogeneity. IR

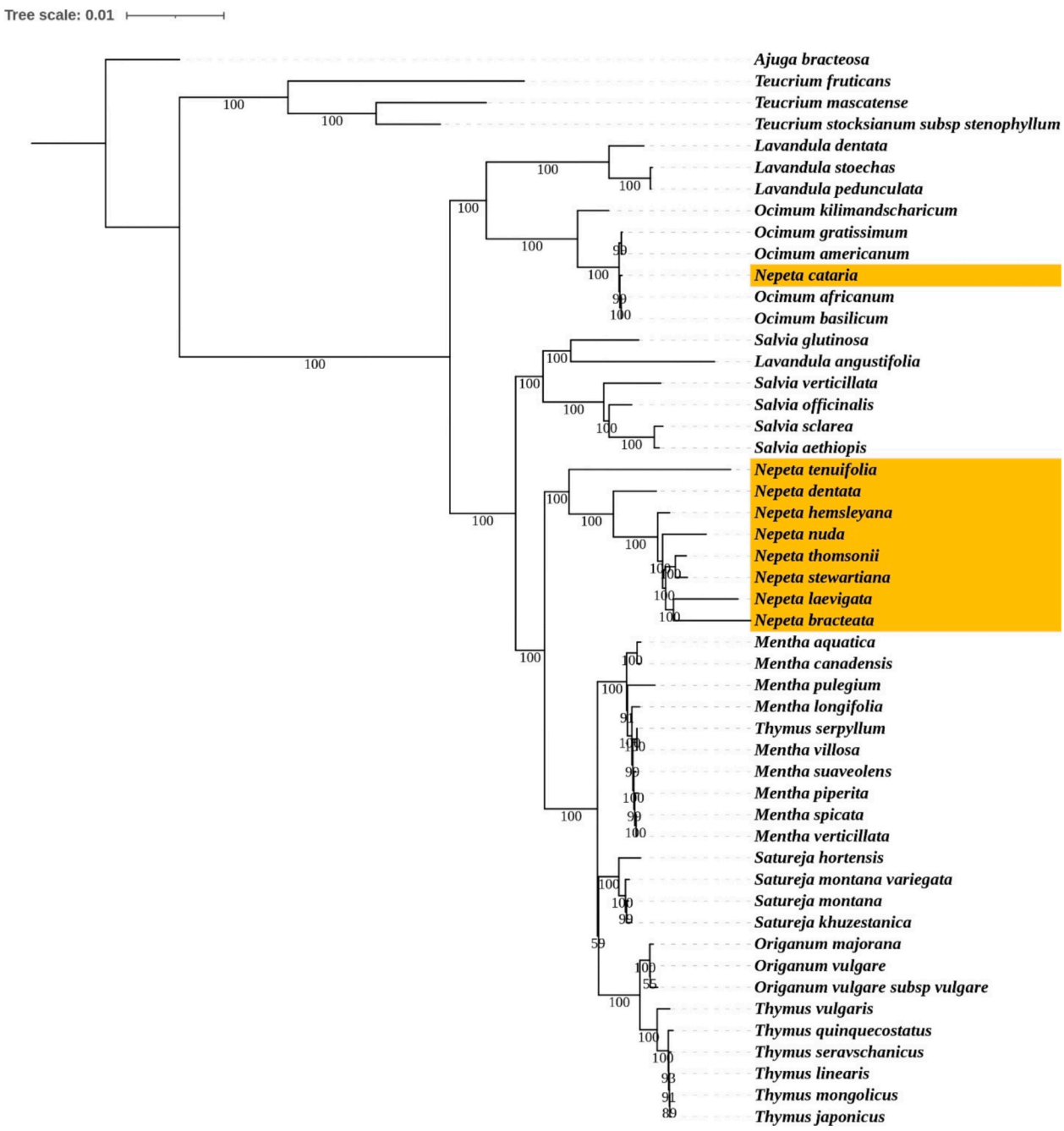

Fig. 7. Phylogenetic tree of 50 cp genomes, including 9 *Nepeta* species and other species such as *Salvia, Thymus, Menthia, Lavandula, Origanum, Ocimum*, and one outgroup (*A. bracteosa*), reconstructed using maximum likelihood in IQ-TREE under the GTR + Gamma model. The highlighted region emphasizes the position and relationships of the sample sequences from this study, illustrating their clustering with closely related species. Branch support was assessed with 100 bootstrap replicates.

expansions or contractions, well-documented in related genera (Daniell et al., 2016; Wang et al., 2017), can alter gene copy numbers through the duplication or truncation of boundary-spanning genes (e.g., *rps19* or *ycf1*). According to our results, *ycf1* and *ycf2* were frequently duplicated in IR regions, while *rps19* and *rpl2* exhibited lineage-specific shifts across species. These findings align with broader studies on plastid genomes, which highlight the role of *ycf1* as a hypervariable marker for species discrimination due to its high substitution rates and structural variability (Dong et al., 2015). For instance, *Salvia* species exhibit IR-mediated variation in *ndhF* and *ycf1* positioning, directly impacting gene tallies (Akrami et al., 2025; Du et al., 2022). Similarly, in *Mentha aquatica*, the partial extension of *rps19* into the IRb region results in its duplication, whereas its complete localization within the LSC in *M. canadensis* yields a single copy (Soorni and Golchini, 2025; Zubair Filimban et al., 2022). Beyond structural rearrangements, gene loss and pseudogenization contribute significantly to observed variation. Non-essential genes (e.g., *ndh* subunits, *accD*) are frequently lost or rendered nonfunctional, as demonstrated in IRLC legumes and *Trifolium* (Cai et al., 2008; Sabir et al., 2014; Sveinsson and Cronk, 2014). Although *Nepeta* species maintain photosynthetic competence, the absence of specific tRNA or ribosomal protein genes (e.g., *rpl23*) in some taxa may reflect nuclear relocation of plastid-targeted proteins, a phenomenon documented in legumes. Additionally, discrepancies in gene annotation, particularly for hypothetical ORFs (*ycf* genes) and short

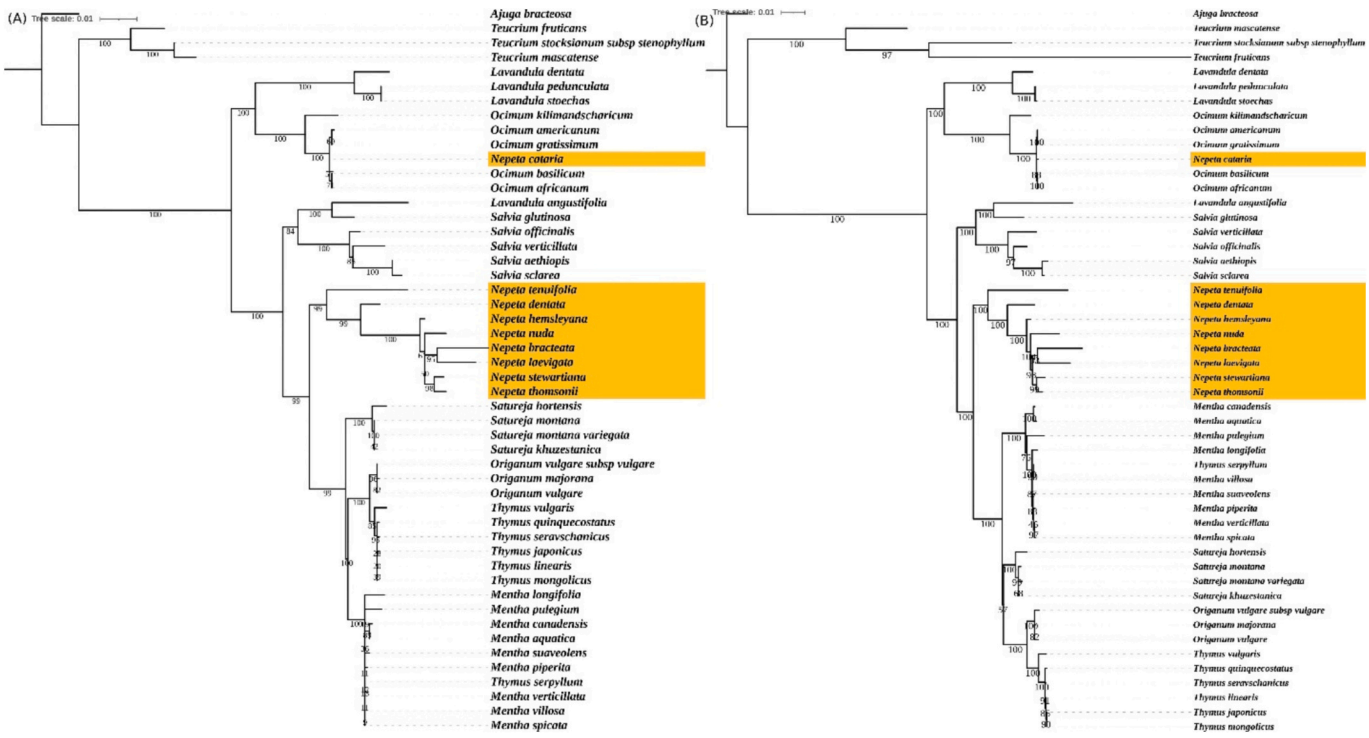

Fig. 8. Phylogenetic tree of 50 cp genomes, including 9 *Nepeta* species, other species such as *Salvia*, *Thymus*, *Menthia*, *Lavandula*, *Origanum*, *Ocimum*, and one outgroup (*A. bracteosa*), reconstructed by *matK* (A) and *ycf1* (B) using maximum likelihood in IQ-TREE under the GTR + Gamma model. The highlighted region emphasizes the position and relationships of the sample sequences from this study, illustrating their clustering with closely related species. Branch support was assessed with 100 bootstrap replicates.

tRNA sequences, can artificially inflate or reduce counts. For example, variable *ycf1* annotation due to fragmentation in draft assemblies has led to inconsistent reporting across studies (Turudić et al., 2023; Xiao-Ming et al., 2017), while tRNA duplications (e.g., *trnE-UUC* triplication in *Satureja khuzestanica*) may be overlooked by automated pipelines. The 131–132 gene range in *Nepeta* aligns with trends in derived Lamiaceae lineages (e.g., *Mentha*: 132; *Satureja*: 128–130) but represents a reduction compared to early-diverging angiosperms (~140 genes), suggesting progressive genome streamlining. This pattern mirrors observations in other Nepetoideae genera, potentially linked to ecological specialization (e.g., drought adaptation in *N. cataria*)(Chen et al., 2024). However, the retention of IRs distinguishes *Nepeta* from radically reduced plastomes (e.g., IRLC legumes), implying that non-IR mechanisms, such as small deletions or relaxed selection on noncoding regions, drive its modest gene loss.

The presence of positive selection in *matK*, *ycf1*, *rpoC1*, *ndhF*, *ndhD*, *ndhC*, and *ccsA* genes indicates that adaptive evolution has acted on functionally important regions of cp genome in *N. nuda*. Several of these genes (*ndhC*, *ndhD*, and *ndhF*) encode subunits of the NADH dehydrogenase–like complex, which is involved in cyclic electron flow and photoprotection, suggesting that selection may be associated with adaptation to variable environmental conditions. The *ccsA* gene plays a key role in cytochrome *c* biogenesis and efficient electron transport, while *rpoC1*, a core subunit of the plastid RNA polymerase, may be subject to selection due to its role in regulating cp gene transcription. In addition, *matK*, a maturase involved in intron splicing, and ycf1, a highly variable and essential chloroplast gene, have frequently been identified as targets of positive selection in angiosperms (Carbonell-Caballero et al., 2015; Wang et al., 2024; Yang et al., 2022). Collectively, these results suggest that selection on genes related to photosynthesis, transcriptional regulation, and RNA processing has contributed to the evolutionary diversification of *N. nuda* within the Lamiaceae family (Akrami et al., 2025; Diani Gohar and Soorni, 2025; Hejazi et al., 2025).

The nucleotide diversity analysis revealed hypervariable regions (*ycf1*, *matK*, *rpl32-trnL-UAG*) in *N. nuda*. The superior discriminatory power of *ycf1* and *matK* has been previously demonstrated in multiple plant groups, reinforcing their utility as core barcodes. The extensive variability in *ycf1* likely stems from the gene's large size and relatively relaxed selective constraints compared to core photosynthetic genes. While the full-length *ycf1* sequence (approximately 5–6 kb) presents practical challenges for standard barcoding applications, targeted sequencing of its most variable regions (particularly the 3′ end) has proven effective for resolving closely related species (Dong et al., 2015). The *matK* gene also demonstrated significant variability, reinforcing its status as one of the core plant DNA barcodes. As a maturase involved in group II intron splicing, *matK* appears to tolerate a higher degree of sequence variation than genes involved in core photosynthetic functions. This pattern of constrained functionality coupled with elevated substitution rates makes *matK* particularly valuable for phylogenetic reconstruction at intermediate taxonomic levels. Our findings in *N. nuda* mirror those reported for other Lamiaceae genera, including *Mentha* (Zubair Filimban et al., 2022) and *Ocimum* (Kirankumar et al., 2023), suggesting conserved evolutionary dynamics across the family. The consistent performance of *matK* across diverse plant lineages (Hollingsworth et al., 2011) further supports its utility as a reliable marker for *Nepeta* taxonomy. The identification of the *rpl32-trnL-UAG* region as hypervariable in *N. nuda* corresponded to findings in *Ardisia* (Yuan et al., 2024), *Scutellaria* (Shan et al., 2021), *Diospyros* (Li et al., 2018), and *Fritillaria* (Chen et al., 2022), where *rpl32* in the SSC region was identified as one of the most hypervariable DNA barcodes.

Phylogenetic analysis placed *N. nuda* in a clade with *N. hemsleyana* and *N. dentata*. Surprisingly, *N. cataria* clustered with *Ocimum* rather than core *Nepeta*, suggesting convergent evolution or historical misclassification (Luo, 2019). Whole-genome phylogenies outperformed single-gene trees (*ycf1* and *matK*), emphasizing the need for multi-locus approaches.

5. Conclusion

This study presents the first complete cp genome assembly of *N. nuda,* providing critical insights into its genomic architecture, evolutionary dynamics, and phylogenetic relationships within the genus *Nepeta*. The cp genome exhibits a typical quadripartite structure, with notable features such as codon usage bias favoring A/U-ending codons, moderate SSR diversity, and highly variable regions (*ycf1*, *matK*, *rpl32-trnL-UAG*) that serve as potential molecular markers for species identification. Positive selection analysis revealed seven genes under strong diversifying selection (*matK*, *ycf1*, *rpoC1*, *ndhF*, *ndhD*, *ndhC, ccsA*), suggesting adaptive evolution, possibly linked to environmental adaptation or metabolic specialization. Comparative analysis of IR boundaries and structural variations highlighted conserved synteny across *Nepeta* species, with a lineage-specific 20 kb deletion distinguishing most taxa from *N. stewartiana* and *N. laevigata*. Phylogenetic reconstruction confirmed the monophyly of *Nepeta* and resolved *N. nuda* as closely related to *N. hemsleyana*, while the anomalous placement of *N. cataria* with *Ocimum* species warrants further taxonomic investigation. The superior resolution of *ycf1* compared to *matK* underscores its utility as a barcode marker, though whole-genome analysis remains optimal for robust phylogenetics. These findings expand genomic resources for *Nepeta*, enhance understanding of its evolutionary history, and provide valuable markers for future studies in species identification, conservation, and medicinal plant research.

CRediT authorship contribution statement

Fatemeh Alsadat Hejazi: Writing – review & editing, Methodology, Investigation, Formal analysis. **Aboozar Soorni:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Consent for publication

Not applicable.

Ethics approval and consent to participate

The authors declared that experimental research works on the plants described in this paper comply with institutional, national and international guidelines. This article does not involve any endangered or protected species.

Funding

No funding was received for conducting this study.

Declaration of competing interest

The authors declare no conflicts of interest.

Acknowledgments

Not applicable.

Data availability

The assembled and annotated genome is accessible in NCBI database under the research accession PV930134.

References

Aćimović, M., Stanković-Jeremić, J., Cvetković, M., 2020. Phyto-pharmacological aspects of *Nepeta nuda* L.: a systematic review. Lek. sirovine 75–83.

Aćimović, M., Lončar, B., Pezo, M., Stanković Jeremić, J., Cvetković, M., Rat, M., Pezo, L., 2022. Volatile compounds of Nepeta nuda L. from Rtanj Mountain (Serbia). Horticulturae 8, 85.

Akrami, A.M., Meratian Esfahani, S., Soorni, A., 2025. Decoding the chloroplast genomes of five Iranian Salvia species: insights into genomic structure, phylogenetic relationships, and molecular marker development. BMC Genomics 26, 545.

Alim, A., Goze, I., Cetin, A., Atas, A.D., Cetinus, S.A., Vural, N., 2009. Chemical composition and in vitro antimicrobial and antioxidant activities of the essential oil of Nepeta nuda L. subsp. Albiflora (Boiss.) gams. Afr. J. Microbiol. Res. 3, 463–467.

Amiryousefi, A., Hyvönen, J., Poczai, P., 2018. IRscope: an online program to visualize the junction sites of chloroplast genomes. Bioinformatics 34, 3030–3031.

Angelova, P., Hinkov, A., Tsvetkov, V., Shishkova, K., Todorov, D., Shishkov, S., 2016. Inhibition of human herpes virus type 2 replication by water extract from Nepeta nuda L. Acta Microbiol. Bulg. 32, 148–149.

Aničić, N., Gašić, U., Lu, F., Ćirić, A., Ivanov, M., Jevtić, B., Dimitrijević, M., Anđelković, B., Skorić, M., Nestorović Živković, J., 2021. Antimicrobial and immunomodulating activities of two endemic Nepeta species and their major iridoids isolated from natural sources. Pharmaceuticals 14, 414.

As'ov, B., 2006. Conspectus of the Bulgarian Vascular Flora: Distribution Maps and Floristic Elements. Boris ASSYOV.

Bautista, M.A.C., Wen, X., Ma, C., Tu, Y., Chen, T., 2022. Complete chloroplast genome sequence of the Tibetan catnip Nepeta hemsleyana Oliver ex Prain (Lamiaceae). Mitochondrial DNA Part B 7, 552–553.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Cai, Z., Guisinger, M., Kim, H.-G., Ruck, E., Blazier, J.C., McMurtry, V., Kuehl, J.V., Boore, J., Jansen, R.K., 2008. Extensive reorganization of the plastid genome of Trifolium subterraneum (Fabaceae) is associated with numerous repeated sequences and novel DNA insertions. J. Mol. Evol. 67, 696–704.

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348.

Carbonell-Caballero, J., Alonso, R., Ibañez, V., Terol, J., Talon, M., 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. 32, 2015–2035.

Chen, H., Zhang, X., Zhang, G., Zhang, Z., Ma, G., Sun, Z., Liu, C., Huang, L., 2024. The complete chloroplast genome sequence of Nepeta bracteata and comparison with congeneric species. Gene 893, 147919.

Chen, Q., Hu, H., Zhang, D., 2022. DNA barcoding and phylogenomic analysis of the genus Fritillaria in China based on complete chloroplast genomes. Front. Plant Sci. 13, 764255.

Daniell, H., Lin, C.-S., Yu, M., Chang, W.-J., 2016. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. Genome Biol. 17, 1–29.

Darling, A.C.E., Mau, B., Blattner, F.R., Perna, N.T., 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 14, 1394–1403.

Diani Gohar, S., Soorni, A., 2025. Plastome diversity and phylogenomic analysis of Satureja (Lamiaceae): uncovering evolutionary patterns and diagnostic markers. Planta 262, 149.

Dobrogojski, J., Adamiec, M., Luciński, R., 2020. The chloroplast genome: a review. Acta Physiol. Plant. 42, 98.

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. Sci. Rep. 5, 8348.

Du, Q., Yang, H., Zeng, J., Chen, Z., Zhou, J., Sun, S., Wang, B., Liu, C., 2022. Comparative genomics and phylogenetic analysis of the chloroplast genomes in three medicinal Salvia species for bioexploration. Int. J. Mol. Sci. 23, 12080.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797. https://doi.org/10.1093/nar/gkh340.

Gao, F., Chen, C., Arab, D.A., Du, Z., He, Y., Ho, S.Y.W., 2019. EasyCodeML: a visual tool for analysis of selection using CodeML. Ecol. Evol. 9, 3891–3898.

Gormez, A., Bozari, S., Yanmis, D., Gulluce, M., Agar, G., Sahin, F., 2013. Antibacterial activity and chemical composition of essential oil obtained from Nepeta nuda against phytopathogenic bacteria. J. Essent. Oil Res. 25, 149–153.

Greiner, S., Lehwark, P., Bock, R., 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3. 1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. 47, W59–W64.

Hejazi, F.A., Mohammadi, P., Soorni, A., 2025. Comparative chloroplast genomics of Teucrium species reveals genome evolution, phylogenetic relationships, and candidate molecular markers. Sci. Rep. 15, 44318.

Hollingsworth, P.M., Graham, S.W., Little, D.P., 2011. Choosing and using a plant DNA barcode. PLoS One 6, e19254.

Huang, L., Yu, H., Wang, Z., Xu, W., 2024. CPStools: a package for analyzing chloroplast genome sequences. iMetaOmics 1, e25.

Jamzad, Z., Grayer, R.J., Kite, G.C., Simmonds, M.S.J., Ingrouille, M., Jalili, A., 2003. Leaf surface flavonoids in Iranian species of Nepeta (Lamiaceae) and some related genera. Biochem. Syst. Ecol. 31, 587–600.

Jansen, R.K., Cai, Z., Raubeson, L.A., Daniell, H., Depamphilis, C.W., Leebens-Mack, J., Müller, K.F., Guisinger-Bellian, M., Haberle, R.C., Hansen, A.K., 2007. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. Proc. Natl. Acad. Sci. 104, 19369–19374.

Jin, D.-M., Wicke, S., Gan, L., Yang, J.-B., Jin, J.-J., Yi, T.-S., 2020a. The loss of the inverted repeat in the putranjivoid clade of Malpighiales. Front. Plant Sci. 11, 942.

Jin, J.-J., Yu, W.-B., Yang, J.-B., Song, Y., DePamphilis, C.W., Yi, T.-S., Li, D.-Z., 2020b. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 21, 1–31.

Kabalay, B., Mutlu, D., Arslan, S., Semiz, G., Kocabıyık, K., 2018. Chemical composition and cytotoxicity of Nepeta nuda subsp. lydiae PH Davis essential oil towards colon carcinoma. In: Proceedings of the 4th International Symposium EuroAsian Diversity, Kiev, Ukraine, pp. 3–6.

Kirankumar, S.I., Balaji, R., Tanuja, Parani, M., 2023. The complete chloroplast genome of *Ocimum basilicum* L. var. *basilicum* (Lamiaceae) and its phylogenetic analysis. Mitochondrial DNA Part B 8, 1169–1173.

Letunic, I., Bork, P., 2021. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 49, W293–W296. https://doi.org/10.1093/nar/gkab301.

Li, W., Liu, Y., Yang, Y., Xie, X., Lu, Y., Yang, Z., Jin, X., Dong, W., Suo, Z., 2018. Interspecific chloroplast genome sequence diversity and genomic resources in Diospyros. BMC Plant Biol. 18, 210.

Luo, J., 2019. Characterization of the complete chloroplast genome of Nepeta cataria L. of the herbal medicine from China and phylogenetic relationships. Mitochondrial DNA Part B 4, 1452–1454.

Narimani, R., Moghaddam, M., Ghasemi Pirbalouti, A., Mojarab, S., 2017. Essential oil composition of seven populations belonging to two Nepeta species from northwestern Iran. Int. J. Food Prop. 20, 2272–2279.

Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274. https://doi.org/10.1093/molbev/msu300.

Niu, Y., Qin, Q., Dong, Y., Wang, X., Zhang, S., Mu, Z., 2023. Chloroplast genome structure and phylogenetic analysis of 13 Lamiaceae plants in Tibet. Front. Biosci. 28, 110.

Parks, M., Cronn, R., Liston, A., 2009. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. BMC Biol. 7, 84.

Petrova, D., Gašić, U., Yocheva, L., Hinkov, A., Yordanova, Z., Chaneva, G., Mantovska, D., Paunov, M., Ivanova, L., Rogova, M., 2022. Catmint (Nepeta nuda L.) phylogenetics and metabolic responses in variable growth conditions. Front. Plant Sci. 13, 866777.

Petrović, L., Skorić, M., Šiler, B., Banjanac, T., Gašić, U., Matekalo, D., Lukić, T., Nestorović Živković, J., Dmitrović, S., Aničić, N., 2024. Patterns of genetic variation of Nepeta nuda L. from the Central Balkans: understanding drivers of chemical diversity. Plants 13, 1483.

Qian, J., Song, J., Gao, H., Zhu, Y., Xu, J., Pang, X., Yao, H., Sun, C., Li, X., Li, C., 2013. The complete chloroplast genome sequence of the medicinal plant Salvia miltiorrhiza. PLoS One 8, e57607.

Ruhlman, T., Verma, D., Samson, N., Daniell, H., 2010. The role of heterologous chloroplast sequence elements in transgene integration and expression. Plant Physiol. 152, 2088–2104.

Sabir, J., Schwarz, E., Ellison, N., Zhang, J., Baeshen, N.A., Mutwakil, M., Jansen, R., Ruhlman, T., 2014. Evolutionary and biotechnology implications of plastid genome variation in the inverted-repeat-lacking clade of legumes. Plant Biotechnol. J. 12, 743–754.

Sarikurkcu, C., Ceylan, O., Targan, S., Zeljković, S.Ć., 2018. Chemical composition and biological activities of the essential oils of two endemic Nepeta species. Ind. Crop. Prod. 125, 5–8.

Shan, Y., Pei, X., Yong, S., Li, J., Qin, Q., Zeng, S., Yu, J., 2021. Analysis of the complete chloroplast genomes of Scutellaria tsinyunensis and Scutellaria tuberifera (Lamiaceae). Mitochondrial DNA Part B 6, 2672–2680.

Shi, L., Chen, H., Jiang, M., Wang, L., Wu, X., Huang, L., Liu, C., 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. Nucleic Acids Res. 47, W65–W73.

Soorni, A., Golchini, M.M., 2025. Complete chloroplast genome of Mentha aquatica reveals hypervariable regions and resolves phylogenetic position within the genus Mentha. Mol. Biol. Rep. 52, 677.

Sveinsson, S., Cronk, Q., 2014. Evolutionary origin of highly repetitive plastid genomes within the clover genus (Trifolium). BMC Evol. Biol. 14, 228.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.

Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E.S., Fischer, A., Bock, R., Greiner, S., 2017. GeSeq–versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 45, W6–W11.

Tonti-Filippini, J., Nevill, P.G., Dixon, K., Small, I., 2017. What can we do with 1000 plastid genomes? Plant J. 90, 808–818.

Turudić, A., Liber, Z., Grdiša, M., Jakše, J., Varga, F., Šatović, Z., 2023. Variation in chloroplast genome size: biological phenomena and technological artifacts. Plants 12, 254.

Vaidya, G., Lohman, D.J., Meier, R., 2011. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. Cladistics 27, 171–180. https://doi.org/10.1111/j.1096-0031.2010.00329.x.

Wang, K., Li, L., Hua, Y., Zhao, M., Li, S., Sun, H., Lv, Y., Wang, Y., 2017. The complete chloroplast genome of *Mentha spicata*, an endangered species native to South Europe. Mitochondr. DNA B 2, 907–909. https://doi.org/10.1080/23802359.2017.1413311.

Wang, Y.-H., Wicke, S., Wang, H., Jin, J.-J., Chen, S.-Y., Zhang, S.-D., Li, D.-Z., Yi, T.-S., 2018. Plastid genome evolution in the early-diverging legume subfamily Cercidoideae (Fabaceae). Front. Plant Sci. 9, 138.

Wang, Yinghui, Xu, C., Guo, X., Wang, Yan, Chen, Y., Shen, J., He, C., Yu, Y., Wang, Q., 2024. Phylogenomics analysis of Scutellaria (Lamiaceae) of the world. BMC Biol. 22, 185.

Wick, R.R., Schultz, M.B., Zobel, J., Holt, K.E., 2015. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics 31, 3350–3352.

Wolf, P.G., Der, J.P., Duffy, A.M., Davidson, J.B., Grusz, A.L., Pryer, K.M., 2011. The evolution of chloroplast genes and genomes in ferns. Plant Mol. Biol. 76, 251–261.

Xiao-Ming, Z., Junrui, W., Li, F., Sha, L., Hongbo, P., Lan, Q., Jing, L., Yan, S., Weihua, Q., Lifang, Z., 2017. Inferring the evolutionary mechanism of the chloroplast genome size by comparing whole-chloroplast genome sequences in seed plants. Sci. Rep. 7, 1555.

Yang, L., Abduraimov, O., Tojibaev, K., Shomurodov, K., Zhang, Y.-M., Li, W.-J., 2022. Analysis of complete chloroplast genome sequences and insight into the phylogenetic relationships of Ferula L. BMC Genomics 23, 643.

Yuan, L., Ni, Y., Chen, H., Li, J., Lu, Q., Wang, L., Zhang, X., Yue, J., Yang, H., Liu, C., 2024. Comparative chloroplast genomes study of five officinal Ardisia species: unraveling interspecific diversity and evolutionary insights in Ardisia. Gene 912, 148349.

Zhang, Y., Tian, L., Lu, C., 2023. Chloroplast gene expression: recent advances and perspectives. Plant Commun. 4.

Zhou, H., Du, L., Fang, Z., 2020. The complete chloroplast genome sequence of Nepeta cataria L.(Labiatae). Mitochondrial DNA Part B 5, 3280–3281.

Zubair Filimban, F., Yarádua, S.S., Bello, A., Choudhry, H., 2022. Complete chloroplast genomes of two mint species (Lamiaceae) from Al-Madinah, Saudi Arabia: phylogenetic and genomic comparative analyses. Mitochondrial DNA Part B 7, 1797–1799.