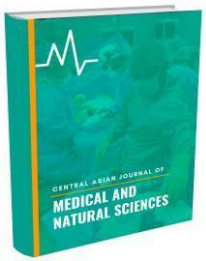

CENTRAL ASIAN JOURNAL OF MEDICAL AND NATURAL SCIENCES
https://cajmns.centralasianstudies.org/index.php/CAJMNS
Volume: 06 Issue: 04 | October 2025 ISSN: 2660-4159

Article

Systematic Analysis of *Rosa* L. Genus and Some of its Species Based on Morphological Traits and DNA Markers

G. Ismonova[*1]

1. Andijan State University, Andijan, Uzbekistan
* Correspondence: bio.gulmira@gmail.com

Abstract: This study presents a systematic analysis of the *Rosa* L. genus and selected species based on morphological characteristics and DNA markers. Species such as *Rosa* kokanica, *Rosa* canina, *Rosa* ecae, *Rosa* fedtschenkoana.were examined in detail through the assessment of diagnostic morphological traits, including leaf shape, flower structure, fruit morphology, and prickles. In addition to morphological evaluation, molecular phylogenetic analysis was conducted using DNA markers, particularly the Internal Transcribed Spacer (ITS) region and chloroplast DNA (cpDNA) sequences. The results revealed genetic relationships and phylogenetic affiliations among the studied species, contributing to a clearer understanding of their systematic placement within the genus. The integration of morphological and molecular data provided deeper insights into the taxonomy of *Rosa* species, clarifying certain ambiguities in traditional classifications. These findings are valuable for the advancement of modern botanical systematics, conservation strategies, and the effective management of genetic resources within the *Rosa* genus.

Keywords: *Rosa* L, Morphological Traits, DNA Markers, ITS Region, Chloroplast DNA (cpDNA), Taxonomy, Botanical Classification

1. Introduction

The genus *Rosa* L., commonly known as roses, comprises approximately 150 to 200 species and is widely distributed across the temperate and subtropical regions of the Northern Hemisphere. More than half of these species are found in Asia, making it the primary center of diversity for this genus (Fougère-Danezan et al., 2015; Bruneau et al., 2007). Roses have long held significant economic and cultural importance due to their use as ornamental plants, sources of essential oils, and applications in traditional medicine. This has increased scientific interest in their genetic diversity and evolutionary history (Qi et al., 2018; Zhu et al., 2015).

Nevertheless, phylogenetic relationships within the *Rosa* genus remain incompletely resolved. Contributing factors include frequent interspecific hybridization, widespread polyploidy, low sequence divergence in plastid and nuclear genomes, high morphological variability, and complex taxonomy shaped by historical hybridization events (Bruneau et al., 2007; Debray et al., 2022; Jeon & Kim, 2019). As a result, traditional classifications—such as those proposed by Rehder (1940), which divided the genus *Rosa* into four subgenera and several sections based on morphology and chromosome counts—often do not align with phylogenetic trees inferred from molecular data (Fougère-Danezan et al., 2015; Zhang et al., 2022)

Citation: Ismonova, G. Systematic Analysis of *Rosa* L. Genus and Some of its Species Based on Morphological Traits and DNA Markers. Central Asian Journal of Medical and Natural Science 2025, 6(4), 1521-1529.

Received: 30th May 2025
Revised: 18th Jun 2025
Accepted: 4th Jul 2025
Published: 21st Jul 2025

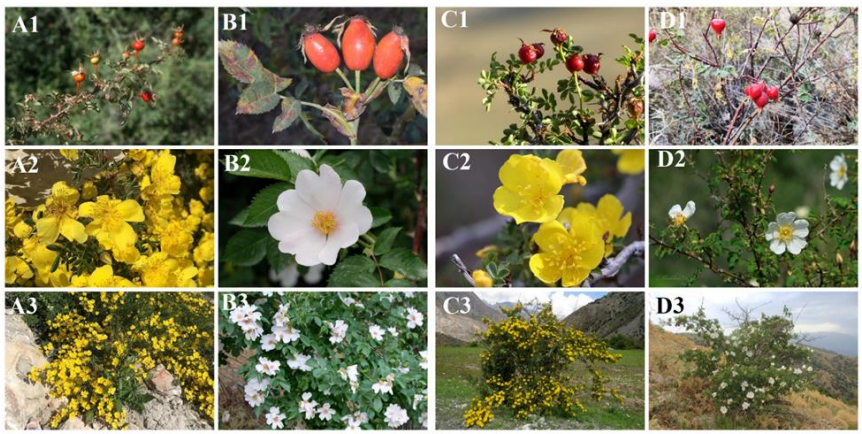

Figure 1. Photographs of four species belonging to the genus *Rosa* (1 – fruits, 2 – flowers, 3 – whole plant appearance). A – *Rosa* kokanica, B – *Rosa* canina, C – *Rosa* ecae, D – *Rosa* fedtschenkoana

2. Materials and Methods

Analysis of plastid genomes (i.e., plastomes) is emerging as an important tool for resolving phylogenetic questions at various taxonomic levels. This is due to the relatively conserved structure of plastid genomes, their maternal inheritance, and moderate levels of sequence variability (Jeon & Kim, 2019; Zhang et al., 2022). In recent years, comparative studies of plastid genomes in *Rosa* species have contributed to the identification of mutation hotspots and to understanding gene divergence and interspecific relationships. However, comprehensive analyses covering wild species are still insufficient (Zhang et al., 2022; Gao et al., 2024). Plastid-based phylogenetic studies often reveal deep phylogenetic splits corresponding to major subgenera and sections. However, detecting fine-scale differences and understanding hybrid evolution requires integration with nuclear genome data (Debray et al., 2022; Jeon et al., 2025). Hybridization and polyploidization have played significant roles in the evolutionary history of the *Rosa* genus, complicating phylogenetic reconstruction and species delimitation (Bruneau et al., 2007; Zhu et al., 2015; Debray et al., 2022). The discrepancies between plastid- and nuclear-based phylogenetic trees highlight historical and ongoing hybridization events within the genus *Rosa* (Jeon et al., 2025; Vozárová et al., 2021).

These evolutionary processes have also contributed to the complex distribution patterns of species. Asia is considered the primary center of genetic diversity, from which species may have later spread to Europe and North America (Fougère-Danezan et al., 2015; Debray et al., 2022). In the context of Uzbekistan and Central Asia, *Rosa* ecae, *Rosa* fedtschenkoana, *Rosa* kokanica, and *Rosa* canina are important representatives of the regional wild rose flora. However, these species have not been sufficiently studied using modern plastomic analyses (Zhang et al., 2023; Gao et al., 2024). Including their plastid genomes in research would allow for the clarification of their phylogenetic placement, genetic differences, and biogeographic relationships.

Such studies contribute to refining the phylogenetic framework of the *Rosa* genus, resolving taxonomic ambiguities, and strengthening selection and conservation strategies based on molecular data (Jeon & Kim, 2019; Debray et al., 2022).

3. Results and Discussion

Complete chloroplast genome analysis of *Rosa* species (*Rosa*ceae) was conducted using newly collected leaf samples of *Rosa* ecae, *Rosa* fedtschenkoana, *Rosa* kokanica, and *Rosa* canina, gathered in June 2024 from wild habitats in Uzbekistan. Species identification was performed by N. Beshko at the National Herbarium of Uzbekistan, and representative herbarium specimens for each species were deposited at the TASH Herbarium. The collected plant materials were immediately dried in silica gel and stored at room

temperature until DNA extraction. Genomic DNA was extracted from leaf tissues using a modified CTAB (cetyltrimethylammonium bromide) method (Doyle & Doyle, 1987). High-throughput sequencing was performed on the Illumina NovaSeq 6000 platform, generating 150 bp paired-end reads. Poor-quality reads and adapter sequences were removed using Trimmomatic v0.39 (Bolger et al., 2014).

The cleaned reads were assembled using GetOrganelle v1.7.5, resulting in complete chloroplast genomes (Jin et al., 2020). The assembled chloroplast genomes were annotated using Geneious Prime v2025.1.2 (Tillich et al., 2017), with *Dianthus ruyschiana* (PQ963003) serving as the reference genome. Annotations were subsequently manually reviewed and curated for accuracy. Final annotated chloroplast genome maps were visualized using OGDRAW (Greiner et al., 2019).

The boundaries of the large single-copy (LSC), small single-copy (SSC), and inverted repeat (IR) regions in the chloroplast genomes of the four species were identified and compared. The shifts in junctions at JLB (LSC/IRb), JSB (SSC/IRb), JSA (SSC/IRa), and JLA (IRa/LSC) were visualized using IRscope (Amiryousefi et al., 2018) and manually aligned in Geneious Prime v2025.1.2 (Figure 3).

To detect nucleotide variability across the chloroplast genomes, complete genome sequences were aligned using MAFFT v7.471 (Katoh & Standley, 2013). Based on the aligned sequences, nucleotide diversity (π) was calculated using DnaSP v6.12.03 (Rozas et al., 2017). A sliding window analysis was performed with a window length of 800 bp and a step size of 200 bp to identify highly variable regions. These highly variable regions may serve as useful molecular markers for phylogenetic and DNA barcoding studies.

Phylogenetic relationships were reconstructed based on the complete chloroplast genomes of *Rosa* species and closely related genera (Potentilla, Rubus) (see Supplementary Figure). Sequences were aligned using MAFFT v7.520 (Katoh & Standley, 2013), and phylogenetic trees were constructed using RAxML v8.2.12 under the GTRGAMMA substitution model with 1,000 bootstrap replicates (Stamatakis, 2014). Rubus peltatus, Rubus tsangii, Potentilla parvifolia, and Potentilla lineata were used as outgroup taxa.

The sequenced plastid genomes of 71 *Rosa* species exhibited the typical quadripartite structure, comprising large single-copy (LSC) and small single-copy (SSC) regions, separated by two inverted repeat (IR) regions. The physical map of the chloroplast genome of *Rosa* kokanica, a representative species of the genus, is shown in Figure 2. The total genome size and the lengths of each region varied slightly among species. The chloroplast genomes ranged from 156,333 bp (R. laevigata) to 157,396 bp (R. minutifolia). Among the four newly sequenced species, the sizes of the LSC and SSC regions ranged from 85,640 bp (R. canina) to 86,300 bp (R. fedtschenkoana), and from 18,772 bp (R. fedtschenkoana) to 18,841 bp (R. kokanica), respectively. The IR regions ranged from 26,048 bp (R. fedtschenkoana) to 26,068 bp (R. canina). The GC content across the 71 *Rosa* chloroplast genomes was nearly uniform, ranging between 37.2% and 37.3% (Table 1).

Table 1. General characteristics of the complete chloroplast genomes of four *Rosa* species.

Takson	Umumiy uzunligi (bp)	LSC uzunligi (bp)	SSC uzunligi (bp)	LSC ning GC foizi (%)	SSC ning GC foizi (%)	IR ning GC foizi (%)	Jami genlar soni	Jami protein kodlovchi genlar soni	tRNA	rRNA
R. kokanica	157,231	86,260	18,841	35.2	31	42.7	132	87	37	8
R. fedtschenkoana	157,168	86,300	18,772	35.2	31	42.7	132	87	37	8

R. ecae	157,215	86,253	18,830	35.2	31	42.7	132	87	37	8
R. canina	156,613	85,640	18,837	35.2	31	42.7	132	87	37	8

Table 2. Chloroplast genome genes of four *Rosa* species.

Genlarning toifasi	Genlarning nomi
Fotosintezga oid genlar (PS)	
ATP sintaza bo'linmalari (ATP)	*atpA, atpB, atpE, atpF, atpH, atpI*
Sitoxrom b/f kompleksi (PET)	*petA, petB*[b], *petD*[a], *petG, petL, petN*
Fotosistema I bo'linmalari (PSA)	*psaA, psaB, psaC, psaI, psaJ*
Fotosistema II bo'linmalari (PSB)	*psbA, psbB, psbC, psbD, psbE, psbF, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ, ycf3*[a]
NADH-degidrogenaza bo'linmalari (NDH)	*ndhA*[a], *ndhB*[ab], *ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK*
Rubisko bo'linmasi (RBC)	*rbcL*
O'z-o'zini replikatsiya qiluvchi genlar (SR)	
Katta ribosomal bo'linma genlari (RPL)	*rpl14, rpl16*[a], *rpl2*[ab], *rpl20, rpl22, rpl23*[b], *rpl32, rpl33, rpl36*
DNKga bog'liq RNK polimaraza (RPO)	*rpoA, rpoB, rpoC1*[a], *rpoC2*
Kichik ribosomal bo'linma genlari (RPS)	*rps11, rps12*[ab], *rps14, rps15, rps16*[a], *rps18, rps19, rps2, rps3, rps4, rps7*[b], *rps8*
Boshqa genlar (OG)	
Asetil-KoA-karboksilaza bo'linmasi (ACC)	*accD*
c-turdagi sitoxrom sintezlovchi gen (CCS)	*ccsA*
Qobiq membrana oqsili (CEM)	*cemA*
Proteaza (CLP)	*clpP*[a]
Maturaza (MAT)	*matK*
Barqaror ochiq o'qiluvchi genlar (ORFs)	*ycf1, ycf2*[b], *ycf4, orf42b, orf188*
RNK genlari	
Ribosomal RNK genlari	*rrn4.5*[b], *rrn5*[b], *rrn16*[b], *rrn23*[b]
Transport RNK genlari	*trnA-UGC*[ab], *trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA, trnfM-CAU, trnG-GCC, trnG-UCC, trnH-GUG, trnI-CAU*[b], *trnI-GAU*[ab], *trnK-UUU*[a], *trnL-CAA*[b], *trnL-UAA*[a], *trnL-UAG, trnM-CAU, trnN-GUU*[b], *trnP-UGG, trnQ-UUG, trnR-ACG*[b],

trnR-UCU, trnS-GCU, trnS-GGA, trnS-UGA, trnT-GGU[b], trnT-UGU, trnV-GAC[b], trnV-UAC[a], trnW-CCA, trnY-GUA

A total of 114 unique genes were identified in the plastid genomes of *Rosa* species, comprising 80 protein-coding genes (PCGs), 30 tRNA genes, and 4 rRNA genes (Table 2). Our data confirm the stable gene set previously reported within the *Rosa* genus (Table 2). In *Rosa* kokanica, *Rosa* fedtschenkoana, *Rosa* ecae, and *Rosa* canina, the gene order and the junction regions between the inverted repeats (IR) and single-copy regions (LSC and SSC) are highly conserved, with only minor differences observed in the lengths of the boundaries. This indicates the evolutionary stability of these regions. In particular, the genes rpl22, rps19, rpl2, and psbA are consistently located at the junctions between IR and single-copy regions, with only slight variations in their precise positions or sequence lengths (Figure 3).

In R. kokanica, a total of 22 intron-containing genes were identified. Among them, 19 genes contain a single intron, including eight protein-coding genes (rpl2, rpl16, rps16, rpoC1, ndhA, ndhB, petB, petD) and six tRNA genes (trnK-UUU, trnG-UCC, trnL-UAA, trnV-UAC, trnI-GAU, trnA-UGC). Three genes (ycf3, rps12, and clpP1) contain two introns.

The trnK-UUU gene contains the largest intron, which is up to 2,497 bp in length and includes the matK gene. The smallest intron is found in the trnL-UAA gene, with a length of 545 bp (Table 3).

Table 3. Intron and exon lengths of split genes in the R. kokanica chloroplast genome.

No.	Gen	Boshlash	Tugash	Ekzon I	Intron I	Ekzon II	Intron II	Ekzon III
1.	trnK-UUU	1279	1,711	37	2497	35		
2.	rps16	6,370	5,230	38	874	229		
3.	trnG-GCC	9,081	9,844	21	693	50		
4.	rpoC1	23,502	20,696	436	761	1610		
5.	ycf3	45,779	43,758	126	735	226	780	155
6.	trnL-UAA	49,281	49,912	37	545	50		
7.	*trnV-UAC*	54,077	53,406	39	596	37		
8.	*rps12*	71,541	143,605	114	71,157	232	536	26
9.	*clpP1*	73,893	71,837	69	827	291	642	228
10.	*petB*	76,837	78,273	6	774	657		
11.	*petD*	78,467	79,668	6	722	474		
12.	*rpl16*	84,556	83,180	9	966	402		
13.	*rpl2*	87,830	86,325	393	678	435		
14.	*ndhB*	99,053	96,845	777	676	756		
15.	*rps12*	100,680	71,541	232	536	26	28,233	114
16.	*trnI-GAU*	104,491	105,516	42	949	35		

17.	*trnA-UGC*	105,581	106,467	38	814	35
18.	*ndhA*	124,595	122,250	551	1,254	541
19.	*trnA-UGC*	137,911	137,025	38	814	35
20.	*trnI-GAU*	139,001	137,976	42	949	35
21.	*ndhB*	144,439	146,647	777	676	756
22.	*rpl2*	155,662	157,167	393	678	435

A sliding window analysis was conducted to assess nucleotide variability (π) across the chloroplast genomes of 71 *Rosa* species, revealing regions with both high and low sequence divergence. The first half of the genome exhibited elevated nucleotide diversity, while the latter half remained highly conserved, with π values approaching zero. This decrease corresponds to regions located within the IR or SSC regions, which are genetically more stable due to copy correction mechanisms. The highly variable regions are primarily situated in the first half of the genome and may include both protein-coding and non-coding regions, indicating their potential utility as phylogenetic markers (Figure 4).

Phylogenetic analysis based on complete chloroplast genomes resolved the *Rosa* genus into four main clades (A, B, C, and D), clearly delineating evolutionary relationships among species (Figure 5).

Clade A contains the early-diverging species *Rosa* berberifolia, which is distinctly separated from all other *Rosa* species.

Clade B includes species such as R. xanthina, R. ecae, R. omeiensis, and R. mairei, exhibiting strong phylogenetic support with high bootstrap values.

Clade C comprises species like R. graciliflora, R. sinobiflora, and R. glomerata, which form closely related subgroups indicating close genetic affinities.

Clade D is the largest and most diverse group, containing numerous species including R. persica, R. sericea, R. fedtschenkoana, R. chinensis, R. multiflora, and R. maximowiczii. This clade also contains several well-supported subclades, reflecting detailed evolutionary divergences within the genus.

Overall, the analysis confirms the monophyly of the *Rosa* genus and provides a robust phylogenetic framework to deepen our understanding of interspecific relationships within this important plant group (Figure 5).

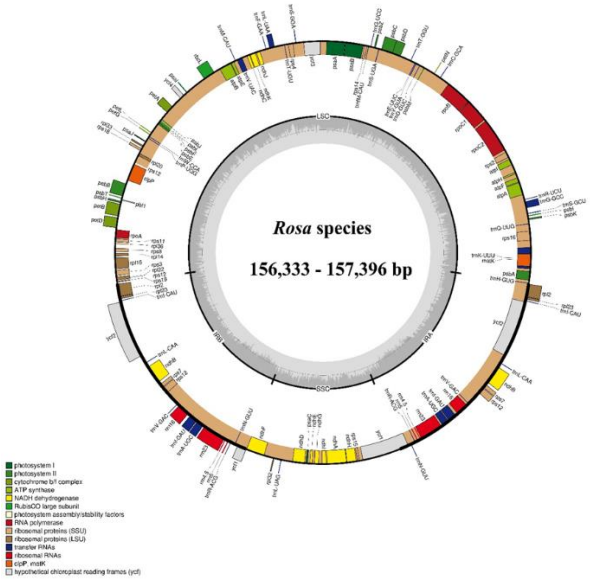

Figure 2. Gene map of the *Rosa* plastome.

Genes located inside the circle are transcribed clockwise, while those outside are transcribed counterclockwise. Genes are grouped into different functional categories and are color-coded accordingly.

The inner circle shows the GC content in dark gray, and the AT content in lighter gray.

LSC – Large Single Copy,

SSC – Small Single Copy,

IRA/B – Inverted Repeat A/B regions.

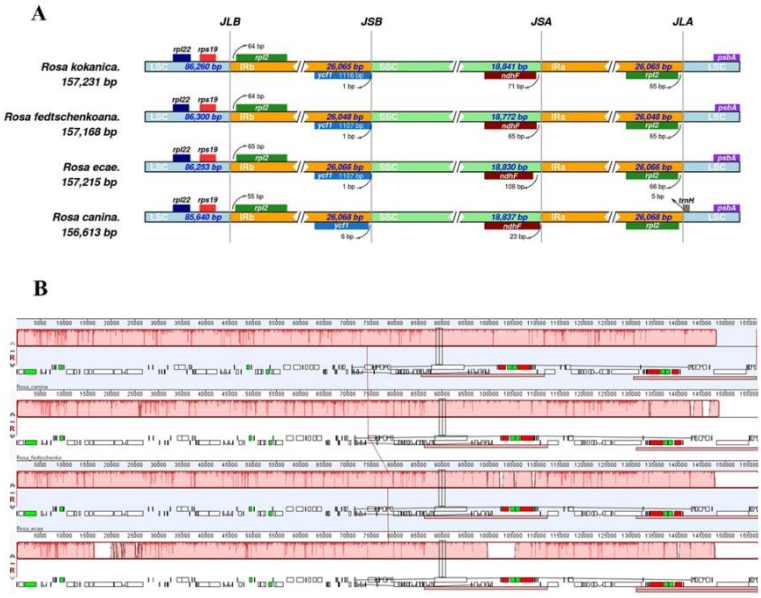

Figure 3. Comparative analysis of chloroplast genomes from four different *Rosa* species.

(a) Comparison of the boundaries of the Large Single-Copy (LSC), Inverted Repeat (IR), and Small Single-Copy (SSC) regions among the four *Rosa* species. Colored boxes represent genes located at or near the boundary regions.

(b) Similarity comparison of *Rosa* chloroplast genomes visualized as a plot. Locally colinear blocks (LCBs) are indicated with different colors, representing homologous regions. Histograms within each block display the degree of sequence similarity.

The results were visualized using IRscope and Mauve software.

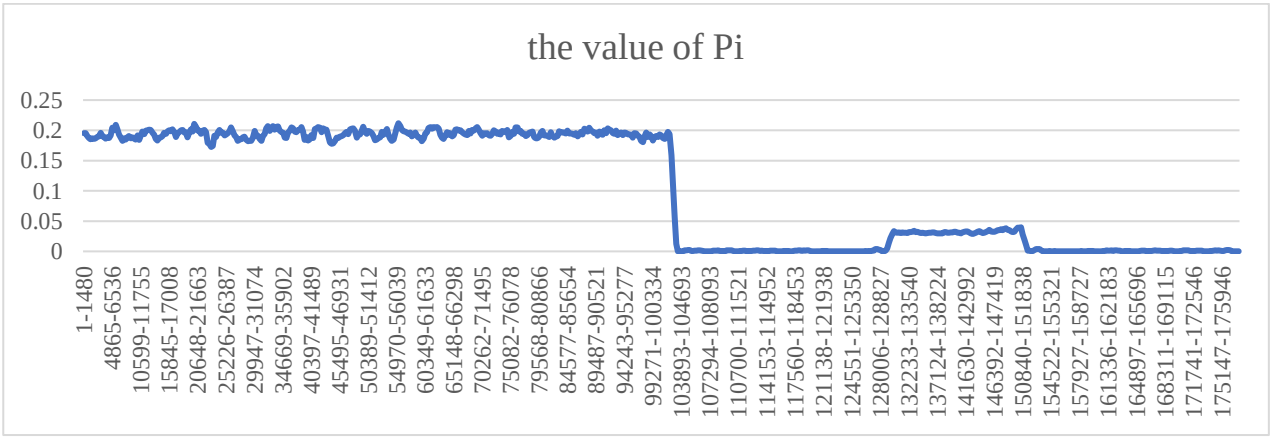

Figure 4. Nucleotide diversity (Pi value) in genomes belonging to the *Rosa* genus visualized using a sliding window approach (window size: 800 bp; step size: 200 bp).

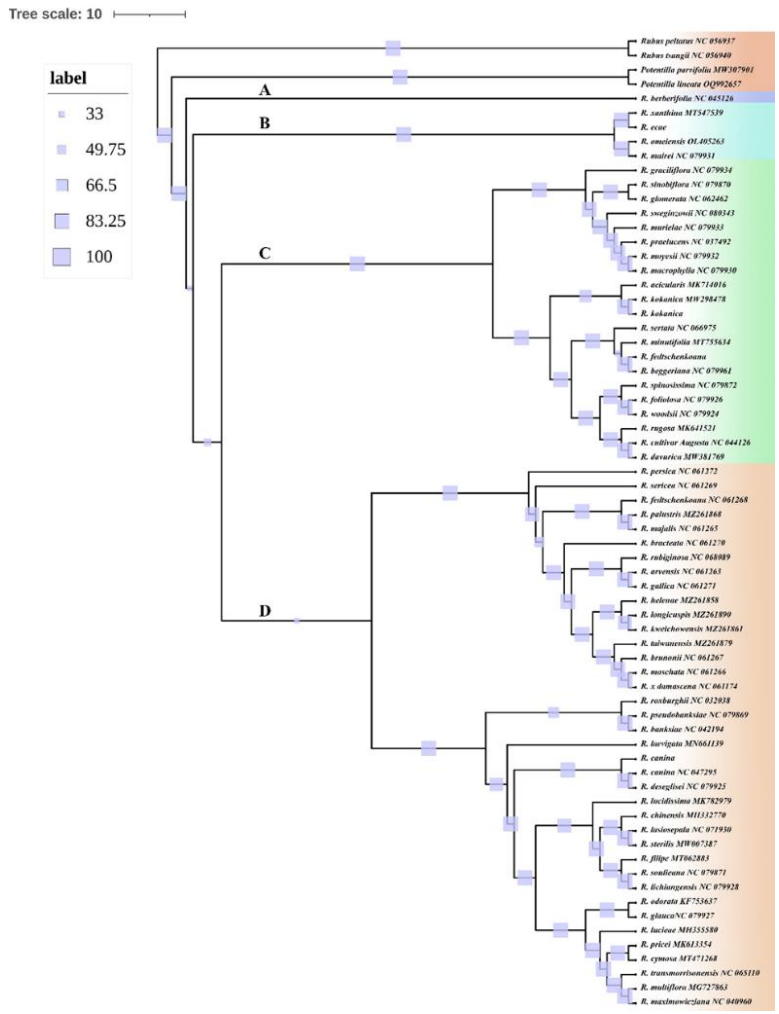

Figure 5. Phylogenetic tree of species belonging to the genus *Rosa*, constructed using the Maximum Likelihood (ML) method based on complete chloroplast genome sequences.

4. Conclusion

This study provides a comprehensive analysis of chloroplast genomes from four wild Rosa species in Uzbekistan—*Rosa ecae*, *R. fedtschenkoana*, *R. kokanica*, and *R. canina*—through high-throughput sequencing and comparative plastome analyses. The results confirm that their chloroplast genomes exhibit high structural conservation, gene content stability, and similar GC content, reflecting evolutionary stability across the genus. Sliding window and phylogenetic analyses identified highly variable regions suitable as potential molecular markers and resolved the Rosa genus into four major clades, affirming its monophyly. This research enhances understanding of the genetic diversity, evolutionary history, and biogeographical relationships of wild Rosa species, contributing valuable molecular data for taxonomy, conservation, and future phylogenomic studies.

REFERENCES

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30(15), 2114-2120.

Bruneau, A., Starr, J. R., & Joly, S. (2007). Phylogenetic relationships in the genus *Rosa*: new evidence from chloroplast DNA sequences and an appraisal of current knowledge. Systematic Botany, 32(2), 366-378.

Debray, K., Le Paslier, M. C., Bérard, A., Thouroude, T., Michel, G., Marie-Magdelaine, J., ... & Malécot, V. (2022). Unveiling the patterns of reticulated evolutionary processes with phylogenomics: hybridization and polyploidy in the genus *Rosa*. Systematic Biology, 71(3), 547-569.

Fougère-Danezan, M., Joly, S., Bruneau, A., Gao, X. F., & Zhang, L. B. (2015). Phylogeny and biogeography of wild roses with specific attention to polyploids. Annals of Botany, 115(2), 275-291.

Gao, C., Li, T., Zhao, X., Wu, C., Zhang, Q., Zhao, X., ... & Li, Z. (2023). Comparative analysis of the chloroplast genomes of *Rosa* species and RNA editing analysis. BMC Plant Biology, 23(1), 318.

Gao, W., Zhou, X., Yu, Q., Lin, G., Fu, C., Kang, T., & Zeng, H. (2024). Assembly and comparative analyses of the Chloroplast genomes of the threatened plant *Rosa* Anemoniflora. Forests, 15(6), 940.

Greiner, S., Lehwark, P., & Bock, R. (2019). OrganellarGenomeDRAW (OGDRAW) version 1.3. 1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic acids research, 47(W1), W59-W64.

Jeon, J. H., Maki, M., Chiang, Y. C., & Kim, S. C. (2024). Inferring complex evolutionary history of the closely related East Asian wild roses in *Rosa* sect. Synstylae (*Rosa*ceae) based on genomic evidence from conserved orthologues. Annals of Botany, mcae170.

Jeon, J.H., & Kim, S.C. (2019). Comparative Analysis of the Complete Chloroplast Genome Sequences of Three Closely Related East-Asian Wild Roses (*Rosa* sect. Synstylae; *Rosa*ceae). Genes, 10(1), 23.

Jian, H., Zhao, L., Zhang, H., Ma, C., Wang, Q., Yan, H., ... & Zhang, T. (2022). Phylogeography and Population Genetics of *Rosa* chinensis var. s pontanea and R. lucidissima Complex, the Important Ancestor of Modern Roses. Frontiers in Plant Science, 13, 851396.

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

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular biology and evolution, 30(4), 772-780.

Qi, W., Chen, X., Fang, P., Shi, S., Li, J., Liu, X., ... & Zhang, Z. (2018). Genomic and transcriptomic sequencing of *Rosa* hybrida provides mic*Rosa*tellite markers for breeding, flower trait improvement and taxonomy studies. BMC Plant Biology, 18, 1-11.

Vozárová, R., Herklotz, V., & Kovařík, A. (2021). Ancient Origin of Two 5S rDNA Families Dominating in the Genus *Rosa* and Their Behavior in the Canina-Type Meiosis. Frontiers in Plant Science, 12, 643548.

Yan, H., Liu, Y., Wu, Z., Yi, Y., & Huang, X. (2021). Phylogenetic relationships and characterization of the complete chloroplast genome of *Rosa* sterilis. Mitochondrial DNA Part B, 6(4), 1544-1546.

Zhang, C., Li, S. Q., Xie, H. H., Liu, J. Q., & Gao, X. F. (2022). Comparative plastid genome analyses of *Rosa*: Insights into the phylogeny and gene divergence. Tree Genetics & Genomes, 18(3), 20.