Taylor & Francis
Taylor & Francis Group

MITOGENOME ANNOUNCEMENT

🔓 OPEN ACCESS Check for updates

The complete chloroplast genome sequences of the *Rosa kokanica* (Regel) Regel ex Juz. (Rosaceae)

Tae-Young Choi[a] 🔟, Aleksey Kim[b], Dong-Kap Kim[b], Su-Young Jung[b] and Soo-Rang Lee[a] 🔟

[a]Department of Biology Education, Chosun University, Gwangju, Republic of Korea; [b]Forest Biodiversity Division, Korea National Arboretum, Pocheon, Republic of Korea

ABSTRACT

Rosa kokanica is a deciduous shrub distributed in Central Asia. We determined the genomic characteristics of the complete chloroplast genome in *R. kokanica* with a de novo assembly strategy. The chloroplast genome was 156,802bp in length harboring 89 protein coding genes, 37 tRNA genes and eight rRNA genes. It exhibits typical quadripartite structure comprising a large single-copy (LSC) (85,899bp), a small single-copy (SSC) (18,773bp) and a pair of inverted repeats (IRs) (26,065bp). Phylogenetic analysis of 16 chloroplast genomes from Rosaceae revealed that the genus *Rosa* is a monophyletic group and the *R. kokanica* is clustered together with the congener, *R. acicular*.
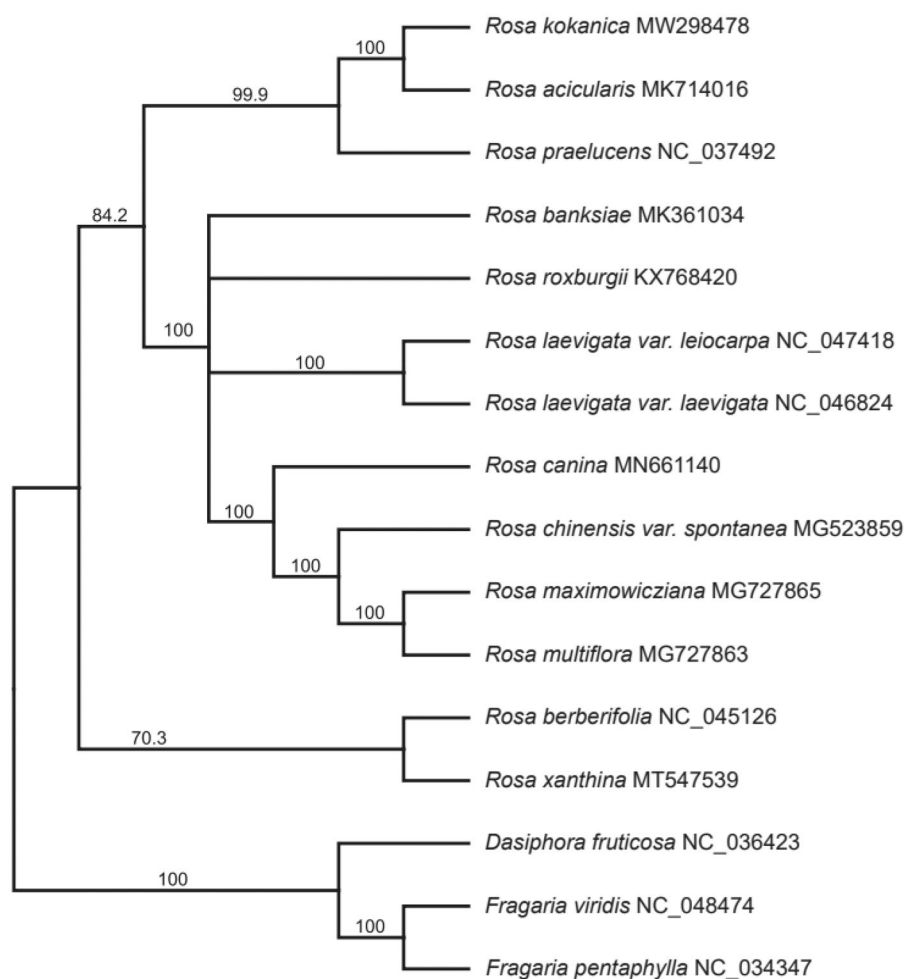
Rosa kokanica (Regel) Regel ex Juz. (Rosaceae) is a deciduous shrub distributed in high mountain slopes of Central Asia including Afghanistan, China (Xinjiang), Iran, Kazakhstan, Kyrgyzstan, Pakistan, Tadzhikistan and Uzbekistan (Gu and Robertson 2003). Genus *Rosa* is well known for its economic importance as ornamental and medicinal plants (Raymond et al. 2018). *R. kokanica,* which is one of ca. 30 *Rosa* taxa distributed in Central Asia, has been widely used for medicines, and has potential ornamental values (Tolekova et al. 2020). However, the genomic information applicable for Central Asian *Rosa* is still scarce. In the present study, we investigated the genomic architecture in the whole chloroplast genome of *R. kokanica* using whole genome shotgun sequencing.

We collected young leaves of *R. kokanica* from Namangang province, Uzbekistan (N 41°01′43.4″, E 70°36′27.5″). The voucher specimen was prepared and deposited at the Herbarium of Korea National Arboretum (KH) with the accession number KHB1547488. The total genomic DNA was extracted followed by manufacturer's protocol (Qiagen, Hilden, Germany). After library preparation, the prepared libraries were sequenced on Illumina MiSeq platform (Illumina, San Diego, CA). Seven million high-quality 300 bp paired-end reads were obtained. We assembled 2.19GB reads with *de novo* strategy using Geneious Prime (ver. 4.2.1) according to the manufacturer's instruction. The genes were predicted with GeSeq (Tillich et al. 2017), and manually curated based on Blast search result. The simple sequence repeats were investigated with MISA (Beier et al. 2017).

The complete chloroplast genome of *R. kokanica* has been submitted to Genbank (accession no. MW298478). It is 156,802bp in length with the typical quadripartite structure comprising a large single copy (LSC) (85,899bp), a small single-copy (SSC) (18,773bp) and a pair of inverted repeats (IRs) (26,065bp). The chloroplast genome contained 131 genes including 89 protein coding genes, 37 tRNA genes and eight rRNA genes. 51 simple sequence repeats were identified in the cp genome, most of which was mono-nucleotide repeat.

To investigate its phylogenetic relationship, the concatenated CDs sequences from entire chloroplast genome of 16 *Rosa* and outgroup taxa were aligned in MAFT (Katoh et al. 2019). All sequences except *R. kokanica* were downloaded from NCBI Genbank. We assigned *Dasiphora* and *Fragaria* as an outgroup following phylogenetic relationships based on a previous study (Potter et al. 2007). We inferred the phylogeny using Maximum likelihood (ML) algorithm implemented in RAxML v. 4.0 with GTR GAMMA model. For the clade support, 1000 bootstrap replicates were used. The 13 species of *Rosa* formed a monophyletic group (BP = 100) with strong support on ML tree (Figure 1). In ML tree, section *Pimpinellifoliae* did not form a monophyletic group, which is consistent with previous phylogenetic studies (Fougère-Danezan et al. 2015). The ML tree also indicated that *R. kokanica* is most closely related with *R. acicularis*.

CONTACT Soo-Rang Lee ✉ ra1130@chosun.ac.kr, ra1130@hotmail.com 🖥 Department of Biology Education, Chosun University, Gwangju, 61452, Republic of Korea

Figure 1. Maximum Likelihood (ML) tree based on chloroplast genome sequences of 16 species of Rosaceae, numbers on the nodes indicated the bootstrap support value (>50%).

Disclosure statement

Funding

ORCID

Tae-Young Choi (iD) http://orcid.org/0000-0002-6070-6676
Soo-Rang Lee (iD) http://orcid.org/0000-0003-0277-4926

Data availability statement

The genome sequence data that support the findings of this study are openly available in GenBank of NCBI at (https://www.ncbi.nlm.nih.gov/) under the accession no. MW298478. The associated BioProject and Bio-Sample numbers are PRJNA706063 and SAMN18115976, respectively.

References

Beier S, Thiel T, Münch T, Scholz U, Mascher M. 2017. MISA-web: a web server for microsatellite prediction. Bioinformatics. 33(16):2583–2585.

Fougère-Danezan M, Joly S, Bruneau A, Gao XF, Zhang LB. 2015. Phylogeny and biogeography of wild roses with specific attention to polyploids. Ann Bot. 115(2):275–291.

Gu C, Robertson KR. 2003. *Rosa* L. In: Wu ZY, Raven PH, Hong DY, editors. Flora of China. Vol 9. St. Louis and Beijing: Missouri Botanical Garden and Science Press; p. 339–381.

Katoh K, Rozewicki J, Yamada KD. 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform. 20(4):1160–1166.

Potter D, Eriksson T, Evans RC, Oh S, Smedmark JEE, Morgan DR, Kerr M, Robertson KR, Arsenault M, Dickinson TA, et al. 2007. Phylogeny and classification of Rosaceae. Plant Syst Evol. 266(1–2):5–43.

Raymond O, Gouzy J, Just J, Badouin H, Verdenaud M, Lemainque A, Vergne P, Moja S, Choisne N, Pont C, et al. 2018. The *Rosa* genome provides new insights into the domestication of modern roses. Nat Genet. 50(6):772–777.

Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. GeSeq - versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 45(W1):W6–W11.

Tolekova S, Sharmanov T, Sinyavskiy Y, Berzhanova R, Mammadov R, Aksoy OK, Yusifli R. 2020. Antioxidant, pharmacological, medical properties and chemical content of *Rosa* L. extracts. Int. J Second Metab. 7:200–212.